Inland Norway University of Applied Sciences

Faculty of Applied Ecology, Agricultural Sciences and Biotechnology

Bibek Singh Karki

Master Thesis

Designing and assessment of different CRISPR/Cas9 RNP complexes to establish basis for *eIF4E* genome editing in potato protoplasts.

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

This thesis is the end of my journey in obtaining my master's degree. Without the support and encouragement from numerous people including colleagues, well-wishers and institutions, completion of this study would have been impossible. At the end of this research, I feel a great pleasure to express my gratitude to all those people who have contributed to making this study possible. Firstly, I would like to thank *Inland Norway University of Applied Sciences*, Campus Hamar, for providing the opportunity to work on this MSc project.

I would like to sincerely acknowledge to *Frøydis Deinboll Myromslien*, Head of Department (Department of Biotechnology, INN, Hamar) and Department of Biotechnology INN for the permission to work on this project and for the valuable discussion and support. Correspondingly, I would like to acknowledge Prof. *Dennis Eriksson* for his support, inspiration and constant supervision, and guidance throughout the study of this project.

Furthermore, I am very grateful to *Wenche Johansen* and *Rob Wilson*, professors, INN, Hamar for their constant supervision, irreplaceable support and inspiration, and valuable guidance throughout the study of the whole project. Notably, I would like to express my humble gratitude towards *Diana Katherine Castillo Avila* for co-supervision, close collaboration, and unconditional support throughout the project. Additionally, I would like to express gratitude towards *Carl Jonas Jorge Spetz*, *NIBIO*, as well as *Graminor* for the support during this project.

Moreover, I am also thankful to all the lab engineers, *Wenche Kristiansen*, *Hanne Greaker*, *Anne Bergljot Falck-Ytter*, and *Teklu Tewoldebrhan Zeremichael* for their support and guidance in laboratory procedures and to those who supported us in the laboratory settings. Finally, I would like to thank all the staffs working at INN, Hamar for their direct or indirect support in the fulfilment of this project.

Abbreviations

BLAST – Basic local alignment search tool

CDS – Coding sequence

cDNA – Complementary DNA

CRISPR – Clustered regularly interspaced short palindromic repeats

Cas – CRISPR associated protein

cv. – Cultivar

DNA – Deoxy ribonucleic acid

DSB – Double stranded DNA break

DTT – Dithiothreitol

E. coli – *Escherichia coli*

ER – Extreme dominance resistance

gDNA – genomic DNA

GM/GMO – Genetically modified organism

HDR – Homology directed DNA repair

HR – Hypersensitive dominant resistance

IVT – In- vitro transcription / CA – Cleavage assay

LB - kanamycin – Luria-Bertani medium containing Kanamycin antibiotic

mRNA – Messenger RNA

NCBI – National Center for Biotechnology Information (database collection)

nCBP – New cap-binding protein

NHEJ – Non homologous end joining

PAM – Protospacer adjacent motifs

PCR – Polymerase chain reaction

PEG – Polyethylene glycol

PLRV – Potato Leafroll Virus

PTNRD – Potato tuber necrotic ringspot disease

PVY – Potato Virus Y

RNA – Ribonucleic acid

RNP complex – Ribonucleoprotein (Cas9 + sgRNA) complex

sgRNA – single guide RNA

SpCas9 – *Streptococcus pyogenes* derived Cas9 protein

ssRNA – Single stranded RNA

S. t. eIF4E – *Solanum tuberosum* specific Eukaryotic translation initiation factor 4E

TALEN – Transcription activator-like effector nucleases

VPg – Virus genome–linked protein

ZFNs – Zinc-finger nucleases

Table of Contents

Abstract

Potato crop is considered as a key to food security; however, potato virus Y infection have been causing seed potato and primary yield loss of up to 80%. Moreover, none of the commercially available potato cultivars has been reported to present natural resistance against potato virus Y. Recently, the CRISPR/Cas9 system has been widely used for precise genome editing and efficient plant mutation research and breeding. This system depends on single guide RNA to guide the Cas9 protein mediated genome editing. Thus, designing and assessing of target gene or allele specific single guide RNA is a crucial step in precise genome editing. Furthermore, commercial potato cultivars are highly heterozygous and encompasses several single nucleotide polymorphisms. Hence, this study aimed to identify allelic variants of *eIF4E* gene in Desirée, Kuras, Celandine, and Innovator potato cultivars, and to design and analyse suitable candidate sgRNAs to perform CRISPR/Cas9 ribonucleoprotein (RNP) based genome editing to obtain potato virus Y-resistant potatoes. Firstly, using PCR amplification combined with Sanger sequencing, different allelic variants specific to four cultivars were identified. So far, 3 allelic coding sequences were identified from Desirée cultivar, and 2, 4, and 5 variants were detected from Celandine, Kuras and Innovator, respectively. Using the coding sequences seven different sgRNAs were designed, in-vitro transcribed and assessed by performing in-vitro cleavage assay. The results indicated that the SpCas9 derived by all seven sgRNAs, cleaved the three allelic CDS complementary DNA specific to Desirée cultivar. Finally, two different protoplast isolation tests were performed, resulting in isolation of moderate number (1×10^4 protoplasts / µl, and 2×10^6 protoplasts /µl) of protoplasts from Desirée and Kuras cultivars respectively. Thus, this study concluded that a basic platform was established for further RNP complex based in-vivo *eIF4E* genome editing, on potato protoplasts to produce PVY resistant potatoes.

Key words: Plant mutation breeding, CRISPR/Cas9 system, eukaryotic translation initiation factor (eIF4E), Potato virus Y, Protoplasts, RNP complex, in-vitro transcription, and cleavage assay.

1. Introduction

1.1 Background of study

The world population is projected to be approximately 9.7 billion by the year 2050, thus increasing food demand by 70% (Thatcher et al., 2017). Food security has thus become one of the major topics of concern. However, (Devaux et al., 2021; Haverkort et al., 2009) strongly states that potato can play vital role in the food security for the increased population due to its high nutritional content, economic importance, producibility in all climatic regions, and high production volume.

Currently, potato (*Solanum tuberosum* L.) is cultivated throughout the world as the foremost non-cereal crop and fourth most cultivated crop (Lucioli et al., 2022). However, potatoes are disease prone crops. Since last few decades potato virus Y (PVY) infection has become one of the leading cause of seed potatoes and primary yield loss, constituting up to 80% of yield loss (Kreuze et al., 2020). The PVY virus is an RNA virus, and the 5' end of the viral RNA is covalently linked with virus encoded VPg protein. Viral VPg hijacks the potato eIF4E protein thus binding to the cap of eIF4E and introducing PVY infection (Grzela et al., 2006). Although few wild type potatoes (Phureja) has shown natural resistance against PVY (Torrance et al., 2020), none of the commercially available potato cultivars presents natural resistance against PVY (Lucioli et al., 2022).

Plant breeders have been adopting conventional breeding to introduce PVY resistance in potatoes. However, the most commercial cultivars of potato are autotetraploid and it makes potato breeding complicated. Hence, due to the varying ploidy levels between commercial and wild PVY resistant potato, conventional breeding in potato is limited (Muthoni et al., 2015). This led to the necessity of implementing molecular methods for developing PVY resistant potatoes. CRISPR/Cas9 is one of such methods of gene editing, that has revolutionized plant genome modification (Zhang et al., 2021). Several studies have shown that editing *eIF4E* gene family using CRISPR/Cas9 brings resistance against potyviruses such as PVY (Lucioli et al., 2022). Hence this study aimed to develop a CRISPR/Cas9 RNP based gene editing platform for introducing PVY resistant potatoes.

1.2 Potato as a ‘Key to food security’

With the global population projecting to reach around 9.7 billion within 2 and a half decades, food security for the growing population has become one of the main topics of discussion. Nevertheless, potato crop can help maintain the food security for the increasing population (Devaux et al., 2021; Haverkort et al., 2009). Additionally, in most of the developing countries potato persists to be ‘local for local’ crop due to its production volume and storability (Haverkort et al., 2009).

Potato is cultivated in almost 20 million hectares of land (Devaux et al., 2021), across all climatic regions around the globe. Approximately 368 million tons of potatoes are produced each year globally from 4,000 different varieties (Berdugo-Cely et al., 2021; Zaheer & Akhtar, 2016). Scrutinizing the context of global potato production, Europe is the second largest potato producer (Goffart et al., 2022) and covers 60% of the global potato market, thus generating net worth up to 12 billion euros per year (Devaux et al., 2020).

Potato tubers are affluent source of nutrients (Zaheer & Akhtar, 2016), minerals, and medicinal compounds, while consumed in lower concentrations (Burgos et al., 2020). Hence, due to its high nutritional value, and health beneficiaries for the consumers and a good source of income for the producers as a cash crop, it continues to be a farmers beloved crop. Potato is believed to help maintain the food security for the over growing population (Kreuze et al., 2020). Based on previous market prices during deranged food demand and supply chains, potato crop can be anticipated to be resistant against global price volatility (Devaux et al., 2021). Furthermore, potatoes have relatively shorter maturity period, producibility in almost all climatic regions, high nutritional contents, and is a good income resource (Devaux et al., 2020; Haverkort et al., 2009). Thus, potato crops can help withstand all changes regarding climatic effects and market effects to ensure the sustenance of people across the globe.

However, the increased demand of food supply requires mass production of potatoes. Additionally, the increased food demand will eventually lead to requirement of enhanced cultivars; with high nutritional contents, pest resistance and increased yield (Berdugo-Cely et al., 2021; Govindaraj et al., 2015). Nonetheless, potato is a disease prone crop and during recent decades the production of seed potatoes and overall yield of potatoes have been severely affected by several disease outbreaks (Kreuze et al., 2020).

1.3 Potato Virus Y (PVY) causes 60-80% of yield loss.

Approximately, 50 virus species and one viroid have been reported to infect potato worldwide (Kreuze et al., 2020). However, currently two viruses; Potato virus Y (PVY) and Potato Leafroll Virus (PLRV) are reported to be the most damaging and leading cause for high yield loss (i.e., exceeding 80%). Whereas, PVY has outstripped PLRV recently and has become the most important virus causing loss of seed potatoes and around 60- 80% primary tuber yield loss (Kreuze et al., 2020).

PVY belongs to the Potyvirus genus in potyviridae family, affecting solanaceous crops specifically potato and is widely spread over the globe (Kreuze et al., 2020). Based on biological properties, seven different strains of PVY (PVY^{O}, PVY^{C}, PVY^{Z}, PVY^{E}, PVY^{N}, PVY^{N}–Wi, and PVY^{NTN}) has been reported so far (Lacomme & Jacquot, 2017; Lucioli et al., 2022). Moreover, studies have shown that these strains have been presenting different levels of infection in potato. PVY^{O} and PVY^{C} strains has been reported to bring about mild infection symptoms on leaves such as mosaic lesions as represented in Figure 1.1 a, and halting, crumpling as presented in Figure 1.1 b. Whereas no visible leaf symptoms were presented by PVY^{N} and PVY^{N}–Wi strains. PVY^{NTN} strain, on the other hand presented severe potato tuber necrotic ringspot disease (PTNRD) as shown in Figure 1.1 c, aside from severe leaf symptoms (Grech-Baran et al., 2018).

Furthermore, PVY has been developing more newer and alarming recombinants. Additionally, studies have reported that the newer recombinants of PVY have been presenting more severe and more progressive symptoms on potatoes. Two new recombinants, PVY^{NTN}-HN1 or, PVY^{NTN}-HN2 infected ‘Yukon Gold’ potatoes presented severe leaf deformation and clear necrotic ringspots in tuber potatoes (Hu et al., 2009, 2011).

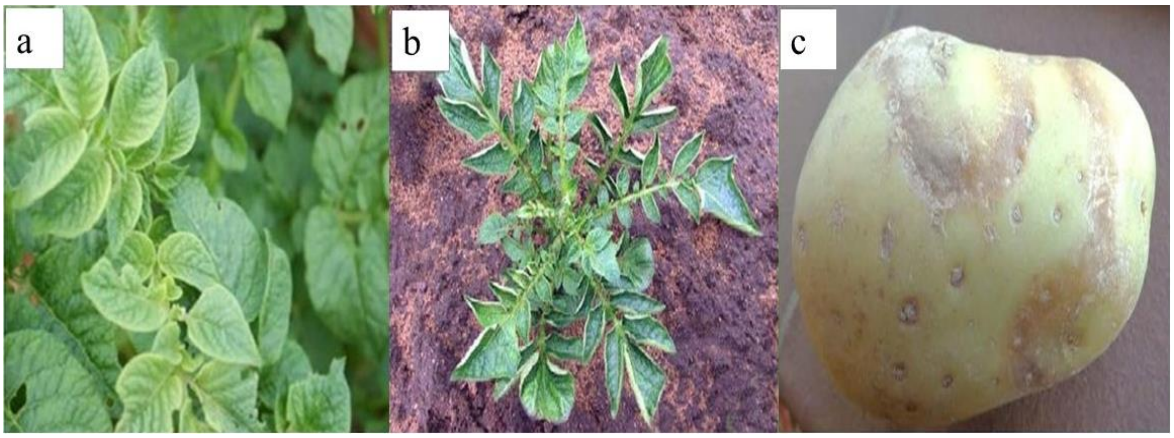


Figure 1.1. Various signs and symptoms of PVY infection in potato plant and tuber potato. (a) mosaic PVY symptoms on leaves of potato crop. (b) severe halting, and crumpling symptoms in potato leaves. (c) severe potato tuber necrotic ringspot disease (PTNRD) symptom induced by PVY^{NTN} infection in tuber potato [Adapted from (Kruger & Waals, 2020)].

Similarly, a study was conducted by Nie et al., (2012), analysing the response of five different PVY isolates (PVY^{O}-FL, PVY^{O}-RB, PVY^{NTN}-Sl, PVY^{N}-Jg, and $PVY^{N:O}$) from four strains against 14 different potato cultivars. The results of the study showed that severity of symptoms was dependent on cultivars and the isolates. However, a conclusion was made that the cultivars presenting necrotic responses were mostly infected with PVY^{O}-FL, PVY^{O}-RB, and PVY^{NTN}-Sl isolates and led to greater yield loss. Whereas the cultivars infected with PVY^{N}-Jg and $PVY^{N:O}$ isolates presented mosaic symptoms during early developmental stage of plants. Nonetheless, up to 46.1% yield loss was reported in cultivars with primary infection presenting only mosaic symptoms (Nie et al., 2012).

During analysis of the causative agents of PVY infection, sixty different aphid species have been noted to be responsible in non-persistent transmission of PVY as demonstrated in Figure 1.2 (Lacomme & Jacquot, 2017; Lucioli et al., 2022). Additionally it rarely occurs by plant to plant contact (Fuentes et al., 2019). Furthermore, the use of PVY infected seed potatoes i.e., tubers used to cultivate new plants carries PVY, can also facilitate the transmission of the disease as presented in Figure 1.2 (Fuentes et al., 2019; Sahi et al., 2016).

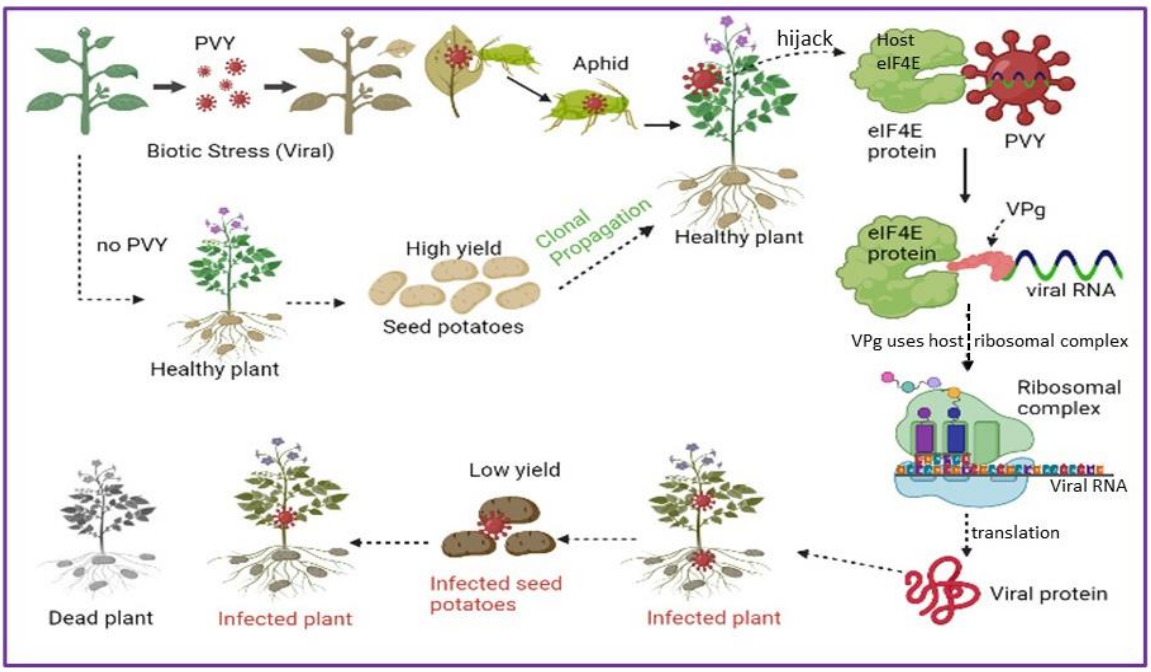


Figure 1.2. Cartoon representing detailed overview of the PVY transmission and induction of infection in potatoes. [Concept adapted from (Fuentes et al., 2019; Gingras et al., 1999; Lacomme & Jacquot, 2017; Lucioli et al., 2022; Robaglia & Caranta, 2006) and created in Biorender.com]

Further investigation for underlying cause of PVY infection showed that the PVY viral genome comprises of 9.7 kb monopartite positive-sense single-stranded RNA. This viral RNA encodes 11 various functional proteins. Virus genome–linked protein (VPg- virulence factor) is one of these proteins, covalently bound to 5' terminal of viral genome, and is the responsible one for infecting potatoes (Lacomme & Jacquot, 2017). The viral VPg (virus genome-linked protein) plays a decisive role by "hi-jacking" the eukaryotic translation initiation factor (eIF4E) of the host (Figure 1.2) i.e. competing with the host mRNA and taking over the eIF4E (Gingras et al., 1999). Thereby, it uses the host ribosomal complex for its own translation (Robaglia & Caranta, 2006) thus infecting the potato cells.

It is illustrious that, PVY has been causing tuber infection and huge yield loss, bringing about severe effects on the potato market. Therefore, it is essential to look after the possible preventive measures. The commonly applied preventive measure for such transmissions is the use of insecticides mainly to control aphids (Lucioli et al., 2022). However, insecticides also have been proven non-effective against PVY along with other viruses. The main reason for this is that there is higher possibility of aphids transmitting virus to the plant prior to the onset of action by insecticides. As will be discussed later, conventional breeding has also been commonly employed over the past century by plant breeders, however conventional breeding

in potato is limited (Muthoni et al., 2015). Additionally, it was difficult to keep up with the evolving pathogens (Lucioli et al., 2022).

1.4 *eIF4E* mediates host resistance against PVY in potato

The chemical (insecticidal) treatments proved to be ineffective against viruses and other intracellular pathogens and hazardous to the environment (Grech-Baran et al., 2018). So due to all these circumstances host resistance deployment started gaining attention (Torrance et al., 2020). Hence, the scientists started looking for the resistance genes against PVY in potatoes. Potato presents two types of host resistance, extreme dominant resistance (ER) mediated by Ry (Ryadg, Rysto, and Rycbc) genes (Karasev & Gray, 2013; Lucioli et al., 2022) and hypersensitive dominant resistance (HR), mediated by N (Ny, Nc, and Nz) genes (Lucioli et al., 2022; Valkonen, 2015).

In addition, plant genes referred to as susceptibility 'S' genes, that encodes proviral factors (usually proteins that are essential for virus in gross infection cycle) can also be considered for introducing virus resistance. Mutation in these S genes often leads to inability for viruses to utilize them, thus inhibiting viral infection and acting as a virus resistance gene. However, this type of mutation in S genes rarely occurs in the nature (Lucioli et al., 2022). Furthermore, all such genes reported till date have been noted to encode; eukaryotic translation initiation factors (eIFs); *eIF4E*, *eIF4G* and respective isoforms (Wang & Krishnaswamy, 2012).

Moreover, eIF4E are 7-methylguanosine triphosphate cap- binders and usually exits in three different forms; eIF4E, eIF(iso)4E, and nCBP (new cap-binding protein), in most of the angiosperms (Browning & Bailey-Serres, 2015; Lucioli et al., 2022; Patrick & Browning, 2012). Nonetheless, potato is heterozygous, tetraploid crop (Nadakuduti et al., 2018). Translation initiation factor in potato i.e., potato *eIF4E* gene are prevalent in four different forms, namely, *eIF4E1* and its paralog *eIF4E2*, *eIF(iso)4E*, and *nCBP* (Lucioli et al., 2022). The *eIF4E* mediated resistance against potyvirus was first reported in mutated *Arabidopsis thaliana* against tobacco etch virus (TEV), due to deficiency in *eIF(iso)4E* (isoform of *eIF4E*) (Hashimoto et al., 2016; Lellis et al., 2002). Later on studies reported that this type of resistance mediated by *eIF4E*s against PVY and few other potyviruses are naturally existent in several crops such as pepper (Ruffel et al., 2002), tomato (Ruffel et al., 2005), and lettuce (Hashimoto et al., 2016).

Additionally, two similar studies Torrance et al., (2009, 2020) reported that, potato group Phureja clones derived resistance against several strains of PVY such as PVY^{O}, PVY^{C}, PVY^{N} and PVY^{NTN}. Additionally, these studies reported that these resistances were being mediated mostly by Ry genes and extreme dominant resistance (ER) type. However, none of the commercially available potato cultivars has been reported till date to be presenting natural resistance against PVY (Lucioli et al., 2022). Thus, it continues to pose a greater threat in the field of potato cultivation and hence brings about a strong demand for genome editing to deploy host resistance.

1.5 Plant mutation breeding as a technique for introducing genetic mutation

Since centuries, geneticists and plant breeders had been relying on conventional breeding for the purpose of obtaining the desired traits in the crops (Haverkort et al., 2009). Mostly, wild species of plants are cross bred with other varieties for the core purpose of crop improvement. However, as mentioned previously, conventional breeding is tedious (Haverkort et al., 2009). Additionally, the difference in the ploidy levels between the commercial cultivar potatoes and wild species (source of resistance genes) limits the conventional breeding in potato (Muthoni et al., 2015). This created strong demand for development and use of molecular methods for breeding.

During recent decades, three types of mutagenesis, random, and site-directed mutagenesis (Udage, 2021) have been commonly used for plant mutation breeding. Randomly induced mutagenesis in the plant genome, mainly uses, physical (fast neutrons, γ-radiations) and chemical such as, ethyl methane sulphonate (EMS) and nitroso methyl urea (NMU) (Alonso & Ecker, 2006). Physical means induces chromosomal expunction and rearrangement in the plant cells. Whereas chemical means provokes a wide range of gene alterations, such as insertions, deletions, and substitution of single base-pair. However, the radiation and chemical use can have severe effects on the plant growth and often makes it difficult to trace back the mutation (Alonso & Ecker, 2006; Shikazono et al., 2005). Hence, these traditional plant mutation breeding in autotetraploid potato are time-consuming as they require series of crossings (Schaart et al., 2021).

Recently, the demand has been to consolidate new attributes as, drought resistance, cold resistance, disease resistance, but retaining the yield (Baret, 2017). Thus, scientists have

been developing genome editing technologies using nucleases such as, meganucleases, transcription activator-like effector nucleases (TALENs), zinc-finger nuclease (ZFNs), and clustered regularly interspaced short palindromic repeats (CRISPR)/associated Cas proteins, which depends on DNA-protein binding (Barman et al., 2020; Gaj et al., 2013). These nucleases mostly induce DNA double-strand breaks (DSBs) at specific target site within the genome, thus activating cellular DNA repair machinery. This often results in insertion-deletion (indels) mutation (Charbonnel et al., 2011; Gorbunova & Levy, 1997; Lloyd et al., 2012; Schaart et al., 2021). However, TALENs and ZFNs based editing requires target specific protein designing for each experiments (Barman et al., 2020; Barrangou & Doudna, 2016) hence can be time consuming, and impractical. All these issues were finally addressed by a recently uncovered genome editing technology: CRISPR-Cas system.

1.6 Application of CRISPR-Cas system in Plant mutation breeding

Plant mutation breeding has been the flourishing technique during recent decades (Udage, 2021). However, the scientific revolutionary era of plant mutation breeding inaugurated mainly after, (Ishino et al., 1987) first discovered CRISPR sequences in *iap* gene in *Escherichia coli* (*E. coli*) genome. The CRISPR/Cas system, is a method of genome editing which induces site-specific mutagenesis either by insertion, deletion, or substitution of the nucleotides (Knott & Doudna, 2018). CRISPR/Cas system has been playing an unprecedented role in plant genome editing. It ensures precise genome editing via Cas protein directed to specific position within the gene by using a guide RNA (Andersson et al., 2018). Hence, CRISPR-Cas system is a RNA based DNA editing system (Jinek et al., 2012). Unlike TALENs AND ZFNs based method, it does not require specific designing of the nucleases for each experiment (Barman et al., 2020).

The CRISPR repeats (Jansen et al., 2002; Mojica et al., 2005) arrays or loci, are the major defining factors of CRISPR/Cas system. These arrays are naturally found in archaeal and bacterial genomes and are fragments of viral genome incorporated during previous infections (Hille et al., 2018). They consists of 17-84 bases long viral DNA fragments (spacer) apportioned by short palindromic repeats of 23-50 bases, clustered into intergenic regions (Gostimskaya, 2022; Popkov et al., 2016). Previous studies showed that CRISPR system in proximity with CRISPR-associated genes containing helicase and nuclease motifs, forms a

CRISPR-Cas organization (Bolotin et al., 2005; Mojica et al., 2005; Pourcel et al., 2005). This organization works as an acquired immune system in prokaryotes to protect against both the foreign DNA (Gostimskaya, 2022; Marraffini & Sontheimer, 2008) and RNA (Gostimskaya, 2022; Shmakov et al., 2015, 2017).

The CRISPR locus is transcribed to produce CRISPR-associated RNA (crRNA), which is complementary to specific viral genomic fragment (Deltcheva et al., 2011). This crRNA in combination with tracrRNA (trans-activating CRISPR RNA) and Cas protein, recognizes the PAM and protospacer sequences (foreign DNA complementary to crRNA) upon later infections. Once, the PAM sequence is recognized and complementarity between crRNA and foreign DNA is confirmed, Cas nuclease or, cascades of Cas protein cleaves or introduces mutation on that foreign DNA (Gostimskaya, 2022; Mojica et al., 2009). This system helps to mutate or degrade the viral DNA thus preventing the prokaryotic cells from infection.

Currently, CRISPR/Cas systems are divided into two classes, based on involvement of Cas proteins. The systems involving a complex of several Cas proteins (protein cascades) are classified into class 1, while class 2 includes systems with single multidomain proteins (e.g. Cas9, Cas12, Cas13) (Gostimskaya, 2022; Koonin & Makarova, 2019; Shmakov et al., 2015, 2017). The recent studies have been using class 2 system of CRISPR/Cas as a genome editing tool. Among class 2 CRISPR proteins, Cas9 recognizes PAM sequence in single strand of DNA and directs blunt ended DSB, whereas Cas12 protein recognizes PAM on both strands and induces staggered cleavage on both strands (Fonfara et al., 2016; Hille et al., 2018; Yamano et al., 2016; Yang et al., 2016; Zetsche et al., 2015). Cas13 protein on the other hand cleaves single stranded RNA (i.e., ssRNA) instead of DNA and they are activated by target ssRNA. Unlike other Cas proteins, Cas13 and Cas12a doesn't require tracrRNA (Hille et al., 2018). Among all these Cas proteins of class 2 CRISPR/Cas system, Cas9 nuclease (so called 'genetic scissors') in association with CRISPR loci induces directional cleavage of foreign DNA or, RNA (Bolotin et al., 2005). Thus, most of the studies have been using CRISPR/Cas9 system as a tool for genome editing.

CRISPR-Cas9 based genome engineering uses Cas9 i.e., DNA endonuclease which has the capacity to cleave both the strands of DNA to introduce DSBs (Barman et al., 2020). The Cas9 protein is guided to the specific site in a genome by a target specific single guide RNA (Hahn et al., 2020). The SpCas9 endonuclease along with sgRNA binds and targets the DNA adjacent to the PAM sequence. The complementarity between 20 nucleotides (spacer sequence) of sgRNA and target DNA, activates the two domains of SpCas9 i.e., HNH domain and RuvC domain which then cleaves both the strands of DNA at position 3 base pairs

upstream to PAM hence inducing blunt ended double strand break (DSB) (Barman et al., 2020). The cells have break repair mechanisms, involving homology-directed repair (HDR) or, non-homologous end joining (NHEJ) repair. NHEJ is an error prone pathway that induces mainly indels, thus incorporating frame-shift mutation. Whereas, HDR depends mainly on addition of donor template and has been on focus in gene editing techniques (Barman et al., 2020).

However, the species edited by HDR with addition of a foreign DNA / fragment into the genome are considered as genetically modified GMO/ GM species. The cultivation and commercial use of such species falls under GMO legislations and are subject to risk assessment and authorisation before cultivation in EU and several countries around the globe (Eriksson et al., 2018). Nonetheless, various forms of Cas9 protein have been developed recently for performing DNA-free genome editing. Cas9 nickase (nCas9) introduces DNA nicks i.e., single strand cleaves. Correspondingly, modified forms of nCas9 have been used in plant genome editing, for example, single to multiple base substitution within the targeted gene via base-editors and editing of short fragments of genomic DNA by prime-editors (Anzalone et al., 2020; Schaart et al., 2021). The most recent technology prime editing has been further expanding the CRISPR-Cas9 based genome editing in plants (Kantor et al., 2020).

Moreover, CRISPR/Cas system is dependent on RNA guided working mechanism, so it is essential to design the sgRNAs to specifically target the gene of interest. sgRNA is a chimeric molecule synthesized by connecting crRNA and tracrRNA together (Gostimskaya, 2022; Jinek et al., 2012). Furthermore, in CRISPR/Cas system based genome editing, the target specificity is totally dependent upon the 20 nucleotide spacer sequence of the crRNA (Jansing et al., 2019). Hence, designing of sgRNA spacer sequence is essential for the selection of precise location of the targets (protospacers) within the gene thus determining the site of edition (Gerashchenkov et al., 2020). The sgRNA can be designed by using online sgRNA design tools. There are several design tools that helps design specific sgRNAs corresponding to the organism specific genomic sequences and perform in-silico analysis (Heigwer & Boutros, 2021).

As, it is evident that CRISPR/Cas9 system of genome editing mainly depends on two main components, Cas9 and sgRNA, so the co-expression of Cas9 protein and its guide RNA is essential. Furthermore, for plant genome editing it is essential to deliver the multiple expression units into the plant cells (Hahn et al., 2020). For the delivery of CRISPR/Cas components, basically two different delivery methods are available, DNA-based method and

DNA-free method. DNA-based delivery method involves use of Agrobacterium- mediated gene transformation by inserting Cas protein genes and sgRNA encoding gene into Ti plasmid under various promoters (Zhang et al., 2021). Thus, DNA based method involves various viral vector based or electroporation based stable (Knott & Doudna, 2018) transformation of the genes encoding CRISPR/Cas components into plant genome (Zhang et al., 2021). As a foreign DNA must be incorporated into the plant genome, so the plants modified by this method are often subjected to GMO legislations before culivation.

DNA-free delivery method on the other hand involves direct delivery of Cas protein and sgRNA as complex i.e., Ribonucleoprotein (RNP) complex into the plant cell. The commonly employed method or mode is PEG mediated transfection of protoplast (Knott & Doudna, 2018). For the process of isolating protoplasts, the plant cell walls are removed using enzymatic means (Reed & Bargmann, 2021). The removal of cell walls facilitates easy transfection of CRISPR/Cas components into the plant cell usually by employing polyethylene glycol (PEG) mediated transfection and usually RNP complex can be directly transformed into protoplasts (Reed & Bargmann, 2021; Svitashev et al., 2016). Additionally, transformation of protoplasts with RNP complex is more efficacious and promising approach of genome alteration to provoke transgene-fee edits. Furthermore, RNP complexes are able to act faster as they do not require transcription and/or, translation (Andersson et al., 2018).

1.7 Impact of the study

The overall objective of this study is to establish a basic platform for PEG mediated CRISPR RNP-based genome editing in potato protoplast to develop PVY resistant potato plantlets, as illustrated in Figure 1.3. Hence, the study aims to characterize different allelic variants in *eIF4E* gene and to design and assess candidate sgRNAs suitable for CRISPR-Cas9 RNP based-genome editing and assess protoplasts isolation as one of the critical steps during gene editing. The study will be following the steps as illustrated in Figure 1.3 and Figure 3.1, however, protoplast transfection will not be performed in this study due to the time limits for the project. Nevertheless, the findings from this study are expected to help design and proceed further with CRISPR/ Cas RNP based in-vivo potato *eIF4E* gene editing, thus contributing production of PVY resistant potatoes.

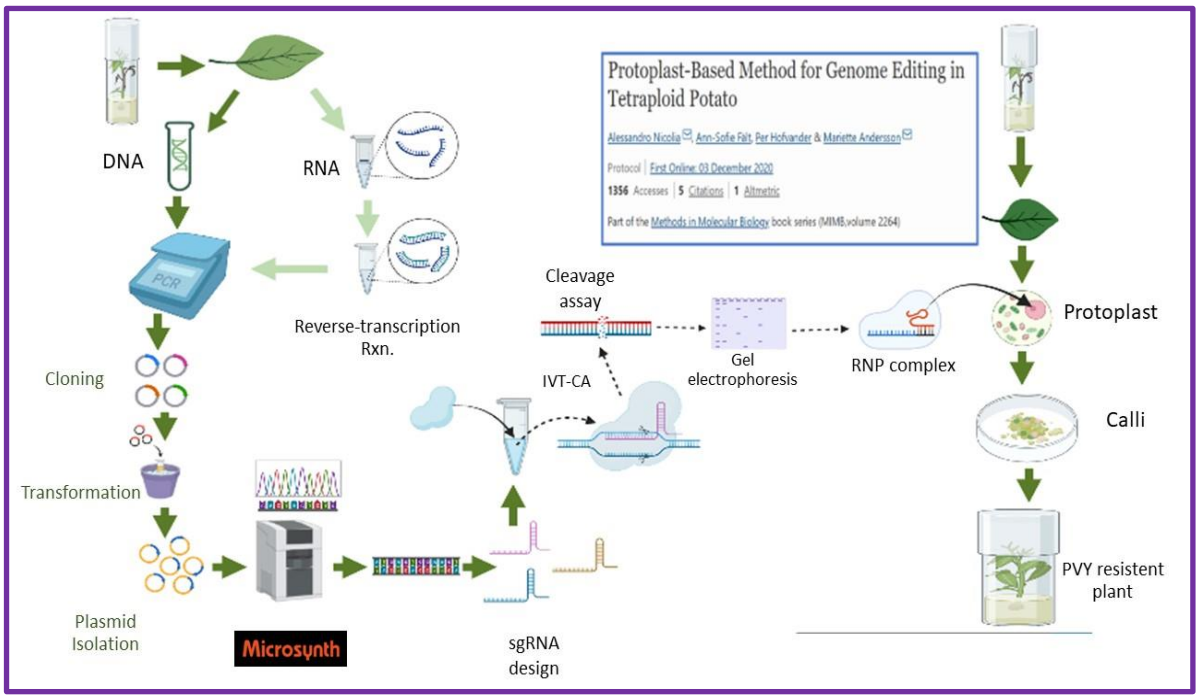


Figure 1.3. Schematic representation of an overall experimental design to obtain PVY resistant potato plant by using RNP complex based genome editing of potato protoplast. However, the isolated protoplasts were not transformed with RNP complexes in this study and further studies were not conducted on isolated protoplasts.

2. Aims and Objectives

The main aim of this study was to identify allelic variants of the *eIF4E* gene in the potato cultivars Desirée, Kuras, Celandine, and Innovator potato cultivars, and to design and assess the candidate sgRNAs to perform CRISPR-Cas9 genome editing to obtain PVY-resistant potatoes.

2.1 Specific Objectives:

1. To determine the sequence of the *eIF4E* gene and its corresponding coding sequence (CDS) for potatoes belonging to Desirée, Kuras, Celandine, and Innovator cultivars.
2. To design candidate sgRNAs specific to *eIF4E* gene in Desirée, Kuras, Celandine, and Innovator cultivars.
3. To analyze the DNA-target cleavage efficiency of synthesized sgRNAs by in-vitro cleavage assay.
4. To assess (Nicolia et al., 2021) protocol to isolate protoplasts from Desirée and Kuras suitable for DNA transfection.

Designing experimental workflow has a great impact on the effective and consistent operation of a research project. Thus, an experimental design is crucial in securing the aims and objective. Hence, in order to achieve the goals of this study an overall design was prepared (Figure 1.3) and an experimental workflow was prepared (Figure 3.1). In addition to experimental workflow Figure 3.1, protoplast isolation was also performed, aiming for the assessment of (Nicolia et al., 2021) protocol.

3. Material and Methods

3.1 Study design and area

The study design of this project work was experimental, and the research project was conducted in the laboratory facility in biohus of Inland Norway University of Applied Sciences, Hamar, Norway. To achieve the aims and objectives of this study (2, p-22), this study was conducted following the experimental design or workflow as illustrated in Figure 3.1.

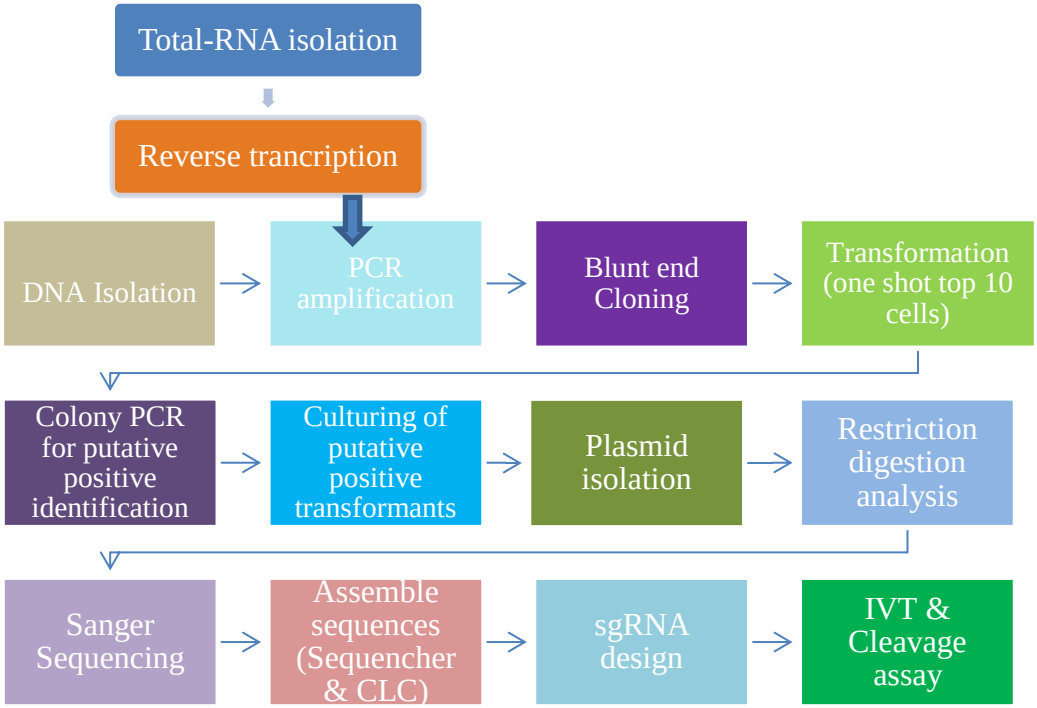


Figure 3.1. Experimental workflow for obtaining sequence information of potato *eIF4E* gene and mRNA CDS specific to four different cultivars and to design and assess various sgRNAs.

3.2 Plant material and growth parameters

Plantlets of *S. tuberosum* (Potato), Desirée and Birkeland cultivar, grown in-vitro were provided by Graminor, Norway. Whereas in-vitro grown plantlets of Celandine, Kuras, and Innovator cultivar potato were provided by Overhalla Klonavelssenter, Norway. The nodes were propagated in-vitro, every 3-4 weeks using 50-100 ml of Medium (4.4 g Murashige and Skoog medium, 30 g sucrose, 8 g of agar in 1 L, 5.8 pH, 0.1% Gibberellic acid 3) and left to

grow in Growth chamber with 16 hours at 250 µE light intensity, 8 hours at dark, 24℃/ 20℃ temperature and 70% humidity.

3.3 Retrieval of potato *eIF4E* CDS and gene sequence information from databases

3.3.1 Database mining for potato *eIF4E* gene and CDS sequences

In pursuance of retrieving the potato specific *eIF4E* gene sequence information, two different approaches were implemented: keyword search in database, and accession number retrieval from published literatures. Firstly using "potato eIF4E" and "Solanum tuberosum eIF4E" as keywords, a thorough search was conducted in NCBI (https://www.ncbi.nlm.nih.gov/) Gene database and Nucleotide database (see Supplementary Figure 8.1 and Figure 8.2).

Alongside, articles containing accession numbers for previously reported and published potato *eIF4E* gene and its mRNA CDS, were searched in NCBI PubMed (https://pubmed.ncbi.nlm.nih.gov/). However, most of the related articles (Cavatorta et al., 2011; Duan et al., 2012; Gutierrez Sanchez et al., 2020) were referring mainly to accession number NM_001288431.1 for the *eIF4E* mRNA CDS sequence, and few allelic mRNA CDS sequences (accession number, FN666435.1, FN666436.1, JN831440.1, JN831441.1, and JN831442.1). Furthermore, the sequence information of all the *eIF4E* mRNA allelic CDS sequences were retrieved from NCBI Nucleotide database in FASTA format.

To retrieve the potato *eIF4E* specific genomic sequence for designing gene specific primers, the precise chromosomal location of the gene within the potato genome was identified. For this purpose, a BLAST search was performed in Phytozome database (https://phytozome-next.jgi.doe.gov/blast-search) using the FASTA sequence for accession number NM_001288431.1 on the target: Solanum tuberosum v6.1. Once the genomic location of *eIF4E* was assured genomic sequence information for potato *eIF4E* (Soltu.DM.03G000970) flanking 294 bases upstream and 189 bases downstream (92), was retrieved from Phytozome; Potato genome v6.1 (Stuberosum_686_v6.1.softmasked.fa.gz).

3.3.2 Identification of exon and intron borders of potato *eIF4E* gene

To identify the intron exon borders of the potato *eIF4E* gene, an RNA-seq exon coverage was analysed in (https://www.ncbi.nlm.nih.gov/genome/gdv/browser/gene/?id=102580433) NCBI Genome Data Viewer. To further confirm the exon-intron location, the *eIF4E* genomic sequence retrieved from Phytozome was aligned with the potato *eIF4E* mRNA CDS sequences retrieved from NCBI (accession no.: NM_001288431.1, FN666435.1, FN666436.1, JN831440.1, JN831441.1, and JN831442.1) in CLC main workbench 7.9.3 (using default parameters).

3.4 Determining the genomic *eIF4E* sequence of the Desirée cultivar and the mRNA *eIF4E* sequences of four different cultivars

3.4.1 Primers designing for PCR amplification of potato *eIF4E* gene and mRNA CDS

Hence, in order to determine the sequence information by Sanger sequencing it was essential to design different overlapping primer pairs as mentioned in Table 4.2 (with minimum of ~200 bp overlapping) to cover full length of the eIF4E gene as shown in Figure 4.3. The primer pairs were designed in Primer-3web software version 4.1.0 (https://primer3.ut.ee/) employing design parameters and salt correction correlation according to (SantaLucia, 1998). The primer length were set to be 20 nucleotides long, Tm (DNA melting temperature) was set to be between 50°C - 62°C and GC content falling between range of 45-60% (Álvarez-Fernández, 2013). The amplicon size range was set to be 500-1500bp as the fragments were to be sequenced using Sanger sequencing. Furthermore, to analyse the unique genomic representation, all the 14 individual primers were analysed by running a BLAST search in Phytozome database (https://phytozome-next.jgi.doe.gov/blast-search) against *Solanum tuberosum v6.1* genome. Finally, the designed primers were synthesized and delivered by Invitrogen. In total seven different primer pairs were designed (Table 4.2).

3.4.2 Genomic DNA isolation from potato leaves

Approximately 50-100 mg leaves from each Desirée and Birkeland cultivar plantlets were pulverised into fine powder using pre-cooled mortar and pestle with liquid nitrogen, then lysed by adding 400 µl AP1 buffer from DNeasy Plant Mini Kit (Qiagen, Germany). Further employing the protocol provided by the company, genomic DNA (gDNA) for both cultivars were isolated and finally eluted in 50 µl – 70 µl of Buffer AE. The quality of the extracted gDNA samples was then assured by 0.8% TAE agarose gel whereas, quantity and purity were evaluated by using NanoDrop Spectrophotometer (Thermo Scientific, USA).

3.4.3 Total RNA extraction from potato leaves

Fresh leaves from plantlets belonging to four different cultivars (Desirée, Kuras, Celandine, and Innovator) were cleaved and pulverised into fine powder using pre-cooled mortar and pestle with liquid nitrogen. Then maximum of 100 mg of tissue powder was lysed by adding 450 µl RLT buffer containing 10 µl β-mercaptoethanol (β-ME) per ml of Buffer RLT, from RNeasy® Plant Mini Kit (Qiagen, Germany) and employing the protocol provided by the company, total RNA was eluted in 30 µl – 50 µl of RNAse-free water. The isolated RNA samples were then analysed in 1.5% TAE agarose gel and purity was assured using NanoDrop Spectrophotometer (Thermo Scientific, USA).

3.4.4 Reverse transcription of total RNA to cDNA

Approximately 2 µg of isolated total RNA was taken and the gDNA content in the sample was digested by using 1 µl ezDNase enzyme from SuperScript™ IV First-Strand Synthesis System Cat. num.: 18091050 (Invitrogen, Lithuania), and incubating at 37°C for 2 min following the company's manual. The ezDNase enzyme was inactivated by adding 10 mM DTT to the sample and incubating for 5 minutes at 55°C. The gDNA free RNA was reverse transcribed to obtain complementary DNA (cDNA), by using 50 µl Oligo d(T) primer and employing SuperScript™ IV First-Strand Synthesis System Cat. num.: 18091050 (Invitrogen, Lithuania) including SuperScript™ IV Reverse Transcriptase reaction, by following the protocol provided by the manufacturer.

3.4.5 PCR amplification of the genomic DNA and cDNA

Total 2 µl (34.2 ng/µl and 40.2 ng/µl for Desirée and Birkeland cultivar) of isolated gDNA was used as template in total 50 µl PCR reactions to amplify the gDNA fragments by employing Phusion Hot Start II High Fidelity- DNA polymerase (Thermoscientific, Lithuania) following the manufacturer's manual (Pub. No. MAN0012397). Seven different reactions were set for gDNA fragments amplification employing seven different primer pairs (Table 4.2). (For Master Mix preparation see Supplementary Table 8.1) The program setup for PCR was set as mentioned in Table 3.1. Later for amplification of the template gDNA using primer pair seven i.e., St_eIF4E-7F, St_eIF4E-7R (master mix prepared as per Supplementary Table 8.2), the annealing temperatures were set to be 60℃, 64℃, and 65℃, i.e., gradient PCR (see Supplementary Table 8.3) was performed to optimise the PCR product generation. After, the completion of reactions, 11 µl of each PCR reactions were separated for analysis, and the remaining reactions were stored at -20℃.

Table 3.1. PCR program used for the PCR amplification of the genomic DNA fragments.

(PCR Program)	Stage 01		
		Step 01: 98°C x 10 sec	
	Stage 02		
		Step 01: 98°C x 20sec	
		Step 02: 60°C x 20 sec	35 cycles
		Step 03: 72°C x 20 sec	
	Stage 03		
		Step 01: 72°C x 7 mins	
		Hold at 4°C	

Correspondingly for PCR amplification of cDNA, 2 µl for Desirée, Kuras, Celandine and Innovator cultivar) of reverse transcribed cDNA samples were used as templates in total 50 µl PCR reactions. The cDNA fragments were amplified using first forward primer and seventh reverse primer (i.e., St_eIF4E-1F and St_eIF4E-7R), by employing Phusion Hot Start II High Fidelity- DNA polymerase (Thermoscientific, Lithuania) following the manufacturer's manual (Pub. No. MAN0012397).

Later the purity of PCR products (amplified gDNA, and cDNA fragments) were analysed by NanoDrop Spectrophotometer (Thermo Scientific, USA), and by 1.5 % (1X TAE buffer) agarose gel electrophoresis at 90V for 30- 40 mins. For this purpose, 10 µl of each PCR reactions were transferred to separate tubes, 2 µl of 6X loading buffer (New England

Biolab) was added to each tube. The PCR products were visualized using SYBR safe (Invitrogen, USA) staining and analysed on reference with 1kb ladder (New England Biolab). Afterwards, the gels were placed in G: box (SYNGENE, USA) and photographs were taken.

3.4.6 Cloning PCR amplified *eIF4E* gene fragments and cDNA into pCR® Blunt vector and transformation in Top 10 cells

After PCR amplification, 0.5 - 2µl of blunt PCR products i.e., the seven different PCR amplified DNA fragments belonging to Desirée cultivar and, the PCR amplified cDNA amplicons belonging to Desirée, Kuras, Celandine, and Innovator cultivar were cloned into pCR® Blunt vector (see Supplementary Table 8.4), based on the protocol of Zero Blunt® PCR Cloning kit (Invitrogen, USA). Afterwards, the vials of OneShot Top 10 chemically competent cells were thawed on ice, divided into two tubes: 25 µl each. Then, 25 µl of OneShot Top 10 chemically competent cells (Invitrogen, USA) were transformed by using 2 µl of ligation reaction, as per instructed in the company's protocol and the cells were cultured overnight on LB plates with 50 µg/ml Kanamycin.

Afterwards, 14 well-isolated colonies for all the Desirée specific gDNA fragments transformation reactions (seven reactions), 19 colonies for Desirée and Innovator specific cDNA transformation reactions, while for Celandine, and Kuras specific cDNA transformation reactions, 15 colonies were picked, and the putative positive colonies were identified. For this, colony PCR was performed by using Hot firepol® DNA polymerase, 5 U/µl (Solis BioDyne, Estonia) and insert fragment specific primer pairs (using the same primers as used for amplification). (For Master Mix preparation and PCR conditions, see Supplementary Table 8.5 and Table 8.6). After completion of PCR, the reactions were analysed using 1% TAE agarose gel to analyse the putative positive colonies.

Ten positive colonies per transformation reactions were picked into sterile plastic culture tubes containing 3 ml of LB broth containing 50 µg/ml Kanamycin and incubated at 37°C with agitation (225 rpm) for 16-18 hours. Afterwards, the plasmid DNA were isolated by using the PureYield Plasmid Miniprep System (Promega, USA), following the manufacturer's instructions, and eluted out using 30 µl of PCR water. After elution, the plasmid DNA concentration and purity were measured using NanoDrop Spectrophotometer (Thermo Scientific, USA). To further validate the presence of the intended insert fragments, ~ 400 ng of plasmid DNA was restriction digested by either, Sma I enzyme (New England Biolab) in presence of rCutSmart™ buffer (New England Biolab) and 1 hour incubation at RT

or, by 1× EcoRI enzyme in presence of EcoRI buffer (New England Biolab) incubating at 37°C for one hour (see reaction details on Supplementary Table 8.7 and Table 8.8). The reactions were then analysed by 1% TAE agarose gel electrophoresis.

3.4.7 Sanger Sequencing

All isolated plasmid samples (i.e., 10 plasmid samples per amplicons) were first diluted to 80 ng/µl and 12 µl of each plasmid samples sent to Microsynth Seqlab GmbH, Germany for Sanger sequencing. After the sequence information were received, trimming, and assembling of the sequences was performed in Sequencher™ 5.4.6- Build 46289 software then alignment of different contigs and consensus was performed in CLC Main Workbench 7.9.3 (see Supplementary Figure 8.5, Figure 8.6, Figure 8.7 and Figure 8.8).

3.5 Designing of single guide RNAs and analysing their efficacy

3.5.1 Designing several sgRNAs for gene editing

In order to design single guide RNAs (sgRNAs), two different tools were implemented, CRISPOR (http://crispor.tefor.net/) and CRISPRdirect (https://crispr.dbcls.jp/) (Gerashchenkov et al., 2020). Using Desirée cultivar specific three different allelic *eIF4E* cDNA sequences, seven different sgRNAs (20 nucleotide spacer sequences) were designed (i.e., five spacer sequences from CRISPOR tool and two spacer sequences from CRISPRdirect tool) and retrieved (Table 4.6).

3.5.2 High copy plasmid isolation for IVT cleavage assay

Alongside the sgRNA designing, the three different plasmids i.e., PCR z.b._De-eIF4E-1, PCR z.b._De-eIF4E-2, and PCR z.b._De-eIF4E-5, (PCR z.b. represents PCR zero blunt vector) containing Desirée cultivar specific three different allelic cDNA sequences as insert (Supplementary Figure 8.3), to be assessed during the in-vitro cleavage assay were isolated to obtain higher concentration of plasmids. Thus, to obtain high copy of plasmids, the starter 100 ml pre-culture was centrifuged at 6000 x g, 4°C for 15 minutes. QIAGEN® Plasmid Maxi Kit (Qiagen, Germany) was employed following the company's protocol. However, changes were made regarding the centrifugation speeds, i.e., instead of centrifugation at 20,000 x g, 4°C,

just 15,000 x g, 4℃. Also, the filter columns to isolate plasmid DNA were used from QIAGEN® Plasmid Plus Maxi Kit (25) (Qiagen, Germany).

3.5.3 sgRNA synthesis and purification

The designed spacer sequences for sgRNAs (from section 3.5.1, and spacer sequences in result Table 4.6) were used to design seven different IVT- forward primers (TaKaRa, Germany) as listed in Table 3.2 to PCR amplify DNA templates containing sgRNA encoding sequences as per instructed in manufacturer's manual. The forward primers were synthesized and provided by Invitrogen. The seven different forward primers were then availed to PCR amplify sgRNA templates and later in-vitro transcribed to synthesize higher amount of sgRNA using Guide-it™ sgRNA *In Vitro* Transcription kit (TaKaRa, Germany) following the manufacturer's instructions. The resulting sgRNAs were purified using Guide-it™ IVT RNA Clean-Up Kit (TaKaRa, Germany) following company's manual. After purification of all the sgRNAs, purity and quantity were evaluated by applying 1 µl of solution in NanoDrop Spectrophotometer (Thermo Scientific, USA).

Table 3.2. IVT- forward primers designed to PCR amplify the DNA templates containing sgRNA encoding sequences for the in-vitro synthesis of seven different sgRNAs.

Oligo sequence (5' to 3') Mandatory Field	**Oligo name**
CCTCTAATACGACTCACTATAggAATGATACGGCGTCGTATTTGTTTAAGAGCTATGC	StDe1_IVT_T7_eIF4E
CCTCTAATACGACTCACTATAggAGTCGTTAGTGTCCGGTCTAGTTTAAGAGCTATGC	StDe2_IVT_T7_eIF4E
CCTCTAATACGACTCACTATAggCAGCAGCGTATATAGCCAGCGTTTAAGAGCTATGC	StDe3_IVT_T7_eIF4E
CCTCTAATACGACTCACTATAgGTCGTTAGTGTCCGGTCTAAGTTTAAGAGCTATGC	StDe4_IVT_T7_eIF4E
CCTCTAATACGACTCACTATAgGTGGAGCAGTCGTTAGTGTCGTTTAAGAGCTATGC	StDe5_IVT_T7_eIF4E
CCTCTAATACGACTCACTATAgGTTTCTAGATCATAGCGATTGTTTAAGAGCTATGC	StDe6_IVT_T7_eIF4E
CCTCTAATACGACTCACTATAggTGGAGCAGTCGTTAGTGTCCGTTTAAGAGCTATGC	StDe7_IVT_T7_eIF4E

3.5.4 Analysing the efficacy of sgRNAs by in-vitro cleavage assay

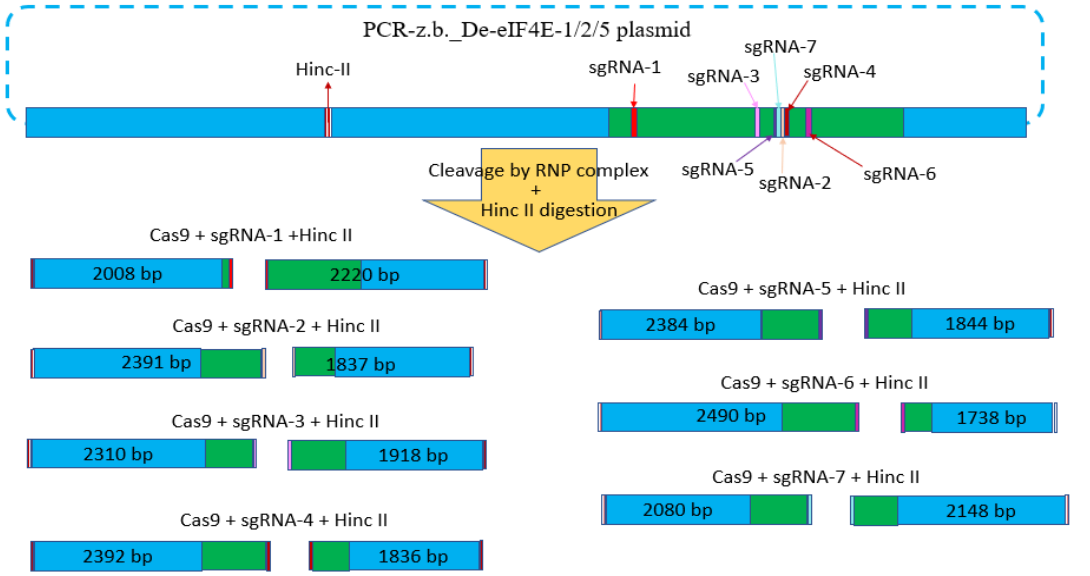


Figure 3.2. Cartoon illustrating the working mechanism of the in-vitro cleavage assay for this study. On top, the positioning of the protospacer sequences complementary to different sgRNA spacer sequences are represented within the 4228 bp long circular plasmids (PCR z.b._De-eIF4E-1/2/5 as represented in Supplementary Figure 8.3). On lower half, the resulting DNA fragments after cleavage by various CRISPR-Cas9- RNP complexes (SpCas9 + in-vitro transcribed sgRNAs 1-7) and further restriction digestion by Hinc II enzyme.

Table 3.3. Preparation of reaction mixture for performing in-vitro cleavage assay to analyse the efficiency of different RNP complexes (SpCas9 + sgRNAs 1-7).

Components	**Stock conc.**	**Final conc.**	**Volume/20 µl rxn.**
SpCas9	1 µM	~30 nM	0-1
sgRNA (1-7)	1 µM	~30 nM	0-1
NEB r 3.1 Buffer	10X	1X	2
H_2O	-	-	q.s. to 16 µl
Template DNA	50 ng/ µl	200 ng/ 30 µl	4
		Total	**20 µl**
Hinc II			2

To analyse the efficiency of the in-vitro transcribed sgRNAs, in-vitro cleavage assay was performed. For this purpose, 1:1 molar ration i.e., 1µM sgRNA and 1µM SpCas9 protein

(M0386T from New England Biolab), was added in 2 µl of NEB r 3.1 buffer (New England Biolab), and volume was made up to 16 µl with nuclease free water (Table 3.3). The reaction mixture was incubated for 10 min at 25℃ then, approximately 200 ng of target plasmid (i.e., 4 µl of 50 ng/ µl plasmid stocks) was added into the reaction to make total volume of 20 µl and was incubated for 1 hour at 37℃. The Cas9 protein was heat in-activated at 65℃ for 5 minutes as per instructed by manufacturer. Afterwards, the reaction was cooled down, briefly centrifuged to collect and 2 µl of Hinc II enzyme was added to digest the plasmid thus producing two distinctly different sized fragments (Figure 3.2) for ease of analysis. Finally, the 20 µl of the reactions were analysed by 1.5% TAE agarose gel electrophoresis at 90 volt for 60 minutes. The gel was analysed using G: BOX (SYNGENE, USA).

3.5.5 Protoplast isolation, Transfection, and Regeneration

Two different tests of protoplast isolation from Desirée and Kuras leaves was performed following the (Nicolia et al., 2021) protocol with few modifications, to analyse the efficiency of the protocol. The major modification was medium C (Enzyme Solution) was prepared without addition of vitamin 3 i.e., vitamin D3 solution (due to unavailability) also, instead of PVP-10, PVP-40 was used. Hence, the sliced leave tissues from medium B were washed with 10 ml of plasmolysis solution, then 20 ml of plasmolysis solution was added to the dish, covered with aluminium foil, and incubated at room temperature for 30 min. Afterwards, the plasmolysis solution was removed and 25 ml of enzymatic digestion solution (i.e., medium C-without vitamin 3) was added then sealed, and wrapped in aluminium foil, and incubated overnight (about 14 hours) at 25℃ without shaking (see Figure 3.3).

Afterwards, the petri dish was incubated for 30 min at RT with very gentle shaking (25 RPM), then the green solution containing protoplasts was gently sieved through two sterile filters of 100 and 70 µm were mounted together, then remaining protoplasts were washed from the filters using 10 ml of wash solution. The sieved protoplast suspension was transferred to sterile 15 ml centrifuge tubes (8 ml per tube), and the tubes were topped up to 15 ml with additional wash solution. The suspension was then centrifuged at 50×g for 5 min. Supernatant was discarded and protoplasts were gently resuspended in 2 ml of wash solution. 6 ml of resuspended protoplasts was layered on the top of 6 ml sucrose solution with a sterile Pasteur pipette, with slight disruption of the interface. The tubes were subsequently centrifuged at 50×g for 15 min, a thick dark band of protoplasts appeared at the interface of the two solutions. The band of protoplast was carefully extracted out on 2 ml transformation buffer 1 (Figure 3.3).

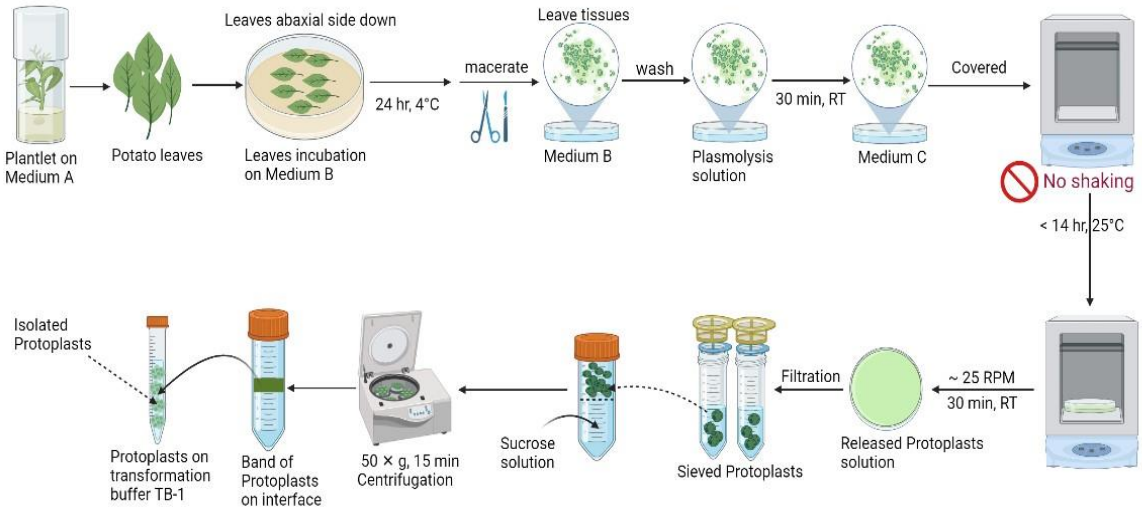


Figure 3.3. Schematic representation of the working mechanism of protoplast isolation from potato leaves following Nicolia protocol. [concept adopted from (Nicolia et al., 2021) and created in Biorender.com]

Finally, a pilot test of protoplast transfection was performed according to (Nicolia et al., 2021), using three plasmids expressing Cas9 and RFP (Red fluorescent protein) provided from Dr. Carl Spetz from NIBIO to evaluate the transfectability of isolated protoplast (Figure 3.3) and later regeneration. However, Medium E was prepared and used without addition of vitamin 3. The alginate gels were prepared and incubated in 10 mL of Medium E (without vitamin 3) at 25℃. However, the results regarding the transfection are not presented as results in this study.

4. Results

4.1 Retrieval of potato *eIF4E* CDS and gene sequence information from databases

4.1.1 Fourteen nucleotide hits of potato *eIF4E* gene & CDS belongs to 3 different cultivars

The potato *eIF4E* gene sequence information was retrieved by employing two different approaches: keyword search in NCBI database, and accession number retrieval from published literatures. The search criteria used were "(("Solanum tuberosum"[Organism] OR potato[All Fields]) AND eIF4E[All Fields]) AND alive[prop]" and (("Solanum tuberosum"[Organism] OR Solanum tuberosum[All Fields]) AND eIF4E[All Fields]) AND alive[prop]. This keyword search on NCBI Gene and Nucleotide database rendered in total 14 different nucleotide hits belonging to potato *eIF4E*, and the nonredundant sequences are listed in Table 4.1. Sequence information retrieved from NCBI were noted to be corresponding to three different cultivars: Russet Burbank, Zhukovskiy ranniy, and Solyntus. Additionally, literature search resulted with accession numbers specific to potato *eIF4E* allelic mRNA CDS sequences (Table 4.1).

Table 4.1. Various accession numbers belonging to nonredundant sequences specific to potato *eIF4E* gene & CDS. These accession numbers were retrieved by keyword search against the NCBI Gene and Nucleotide database or, from published literatures.

Method used	**Accession no.**	**Sequence**	**Alleles**	**Reference**
Keyword search	NW_006239139.1	Genomic		(Cavatorta et al., 2011; Duan et al., 2012; Gutierrez Sanchez et al., 2020)
	NW_006239211.1	Genomic		
	NM_001288431.1	mRNA CDS		
Literature search	FN666435.1	CDS	allele a	(alaoikela et al., 2011)
	FN666436.1	CDS	allele b	
	JN831440.1	CDS	*eIF4E*-1	(Cavatorta et al., 2011)
	JN831441.1	CDS	*eIF4E*-2	
	JN831442.1	CDS	*eIF4E*-3	

Using the FASTA sequence for potato *eIF4E* mRNA cds (accession number NM_001288431.1), a BLAST search was performed in Phytozome on the target: Solanum tuberosum v6.1-potato genome, to substantiate the precise chromosomal location of *eIF4E* in the potato genome. Once the gene position was identified at chromosome number- 3, the *eIF4E* gene sequence (i.e., S. tuberosum v6.1|Soltu.DM.03G000970) was retrieved in FASTA format with 294 bases flanking upstream and 189 bases flanking downstream (Supplementary documents A, p-92). Finally, all the mRNA CDS sequences from Table 4.1 were aligned in CLC and this alignment indicated presence of 22 SNPs between the sequences of 696 bp lengths (alignment in Supplementary Figure 8.15).

4.1.2 Potato *eIF4E* gene encompasses 5 exons & 4 introns.

To identify the intron-exon borders of the potato *eIF4E* gene, an RNA-seq exon coverage analysis was performed in NCBI (Figure 4.1). As clearly suggested, potato *eIF4E* gene encompasses five exons and four introns. The exon-intron location was further confirmed by aligning all the CDS sequences (Table 4.1) to the genomic sequence (i.e., sequence retrieved from Phytozome) of potato *eIF4E* (alignment result in Figure 4.2).

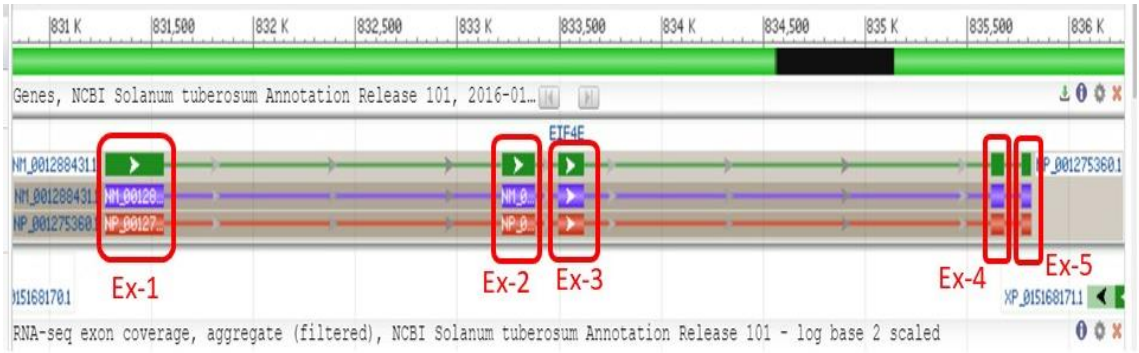


Figure 4.1. RNA-seq exon coverage analysis in NCBI to identify the exon- intron distribution along the potato *eIF4E* gene. Ex-(1-5) represents the exonic regions and the regions between two exons represents the intronic regions.

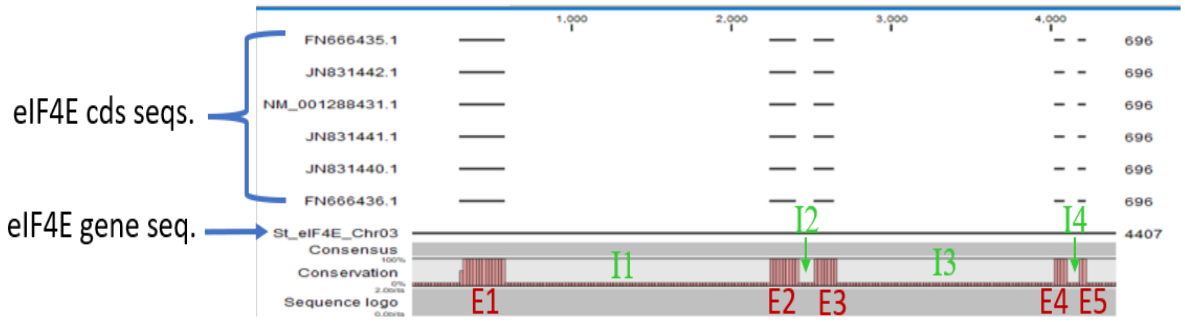


Figure 4.2. Alignment of mRNA cds sequences with the potato *eIF4E* gene sequence in CLC main workbench 7.9.3, to further validate the precise distribution of exonic and intronic regions within the potato *eIF4E* gene. Letter E represents Exons and letter I represents Introns.

The exon-intron border analysis showed that the actual length of the reported potato *eIF4E* gene was 3924 bp where the lengths of the five different exons were as, 286 bp, 167 bp, 128 bp, 65 bp, and 50 bp from E1- E5 respectively.

4.2 Determining the *eIF4E* genomic sequence of the Desirée cultivar and the *eIF4E* cDNA sequences of four different cultivars

4.2.1 Designing seven different overlapping primer pairs to PCR amplify the potato *eIF4E* gene and CDS

The gene sequence from Phytozome database accession number Soltu.DM.03G000970 with 294 bases upstream and 189 bases downstream flanking sequence was used to design overlapping primer pairs for PCR amplification of the coding and genomic sequences of potato *eIF4E*, using online software, Primer 3. Seven different overlapping primer pairs were designed (Table 4.2), covering almost the entire *eIF4E* gene, (Figure 4.3). As can be seen in Figure 4.3, the expected gDNA amplicon sizes would range from 600 bp to 1167 bp. Finally, one primer set, set: St_eIF4E_1F and St_eIF4E_7R was used to amplify the entire CDS of the *eIF4E* CDS.

Table 4.2. Seven different overlapping primer pairs (1-7) for PCR amplification of potato *eIF4E* gene fragments. Tm represents the melting temperature of the primers in ℃, GC% represents percentage of bases guanine (G) and cytosine (C) in the primer sequence. The expected amplicon sizes are provided in base-pairs (bp). Each individual primer was blasted in Phytozome to analyse their unique representation in the whole genome of potato (Phytozome remark). Length represents number of nucleotides in the primer sequence whereas amplicon size is represented in bp.

Primer Pairs	**Oligo name**	**Length (N)**	**Tm**	**GC%**	**Sequence (5' to 3')**	**Amplicon size (bp)**	**Phytozome remark**
1	St_eIF4E_1F	20	59.5	55	CAGCTGAGAAGTTGAAGGCC	600	Unique
	St_eIF4E_1R	20	57.5	50	TTGCTCTGCTGGCTCAAAAG		
2	St_eIF4E_2F	20	59.27	55	GGCGTCGTATTTGGGGAAAG	990	Unique
	St_eIF4E_2R	20	59.46	60	GTCCTCCTACCACCAAGAGC		
3	St_eIF4E_3F	20	59.04	50	TTTTGAGCCAGCAGAGCAAG	1151	Unique
	St_eIF4E_3R	20	59.1	50	AGGTACACGCATCCATGACA		
4	St_eIF4E_4F	20	57.5	50	TGACCGGTGAAGTTCTAGGA	857	Unique
	St_eIF4E_4R	20	59.5	55	TTTCCACGTCCCTCCATTGG		
5	St_eIF4E_5F	20	57.5	50	TGTCTGTCATGGATGCGTGT	1167	Unique
	St_eIF4E_ 5R	20	59.5	55	CGTGTTGGGAGGAGCTGAAA		
6	St_eIF4E_6F	20	59.04	50	GTGGACCAAGAATGCTGCAA	1058	Unique
	St_eIF4E_6R	20	59.01	50	TAGCCATGGTTCTCGGGTTT		
7	St_eIF4E_7F	20	60.18	55	TCCTCCCAACACGTCTCAGA	1082	Unique
	St_eIF4E_7R	20	59.96	55	AGTGCCTACCAACTTTCCGG		

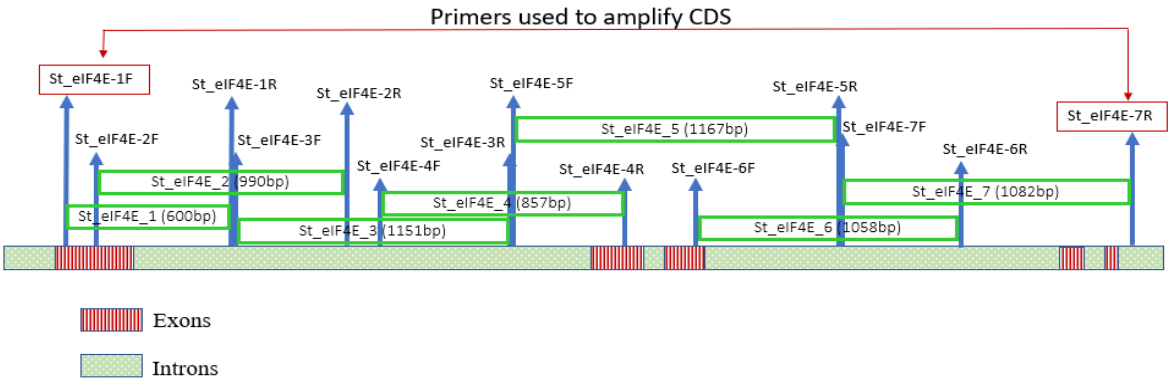


Figure 4.3. Distribution of seven different overlapping primer pairs along the potato *eIF4E* genomic locus. Light green boxes above the gene represents the seven different amplicons as output from PCR upon amplification using the seven different primer pairs (Table 4.2). The numbers in the boxes represents the sizes (in bp) of the expected PCR products.

4.2.2 PCR amplification of the gene and cDNA fragments

Using the seven different primer pairs, the gDNA prepared from two different cultivars (Desirée and Birkeland) was used as templates for PCR amplification. Additionally, using the St_eIF4E_1F and St_eIF4E_7R primers, cDNA prepared from Desirée and Innovator cultivar were used as templates for PCR amplification. All PCR reactions were analysed by agarose gel electrophoresis (Figure 4.4 and Figure 4.5). Correspondingly, reverse transcribed cDNA fragments prepared from Kuras and Celandine cultivars were also PCR amplified (data not shown).

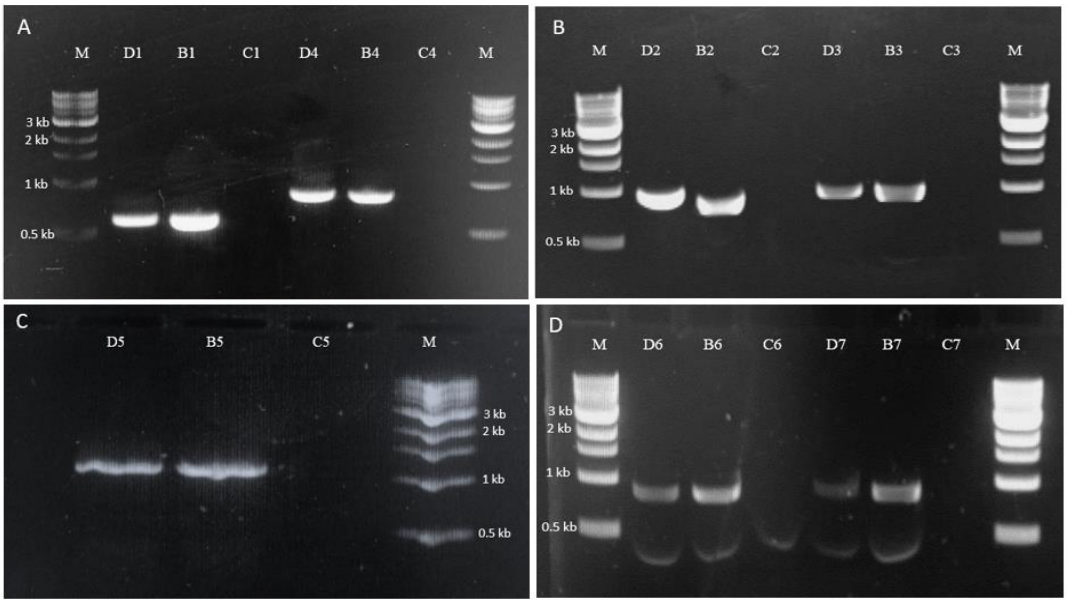


Figure 4.4. Agarose gel electrophoresis of PCR amplified *eIF4E* genomic fragments of Desirée and Birkeland cultivar. M = molecular weight marker (1 kb DNA ladders). D(1-7), seven different amplicon fragments of *eIF4E* gene from Desirée cultivar. B(1-7), seven different amplicon fragments of *eIF4E* gene from Birkeland cultivar. C(1-7), no template controls, containing the primer pairs (1-7) specific PCR reactions. The numbers represent the 7 different amplicons from primer set provided in Table 4.2.

All the PCR reactions, seemed to contain the amplified DNA fragments of expected sizes in the agarose gel (Figure 4.4 and Table 4.2), suggesting the presence of *eIF4E* amplicons of the expected sizes. No DNA amplification was detected in the control reactions (C1-7), suggesting that the PCR amplifications were template specific and that there was no contamination between the samples. The DNA fragment amplified from the Birkeland cultivar using primer set 2(i.e., B2) was of slightly smaller size than the corresponding fragments amplified using gDNA extracted from the Desirée cultivar (Figure 4.4 B, lane 2). This clearly suggested the presence of an indel in the gene sequence. Furthermore, PCR amplification using primer set 7 for the Desirée cultivar (Figure 4.4 D, lane 5) resulted in very low amounts of amplified product, compared to all the other reactions. Hence, to optimise the PCR product generation of the template Desirée gDNA using primer pair seven (i.e., St_eIF4E-7F, St_eIF4E-7R), gradient PCR was performed. The annealing temperatures were set to be 60℃, 64℃, and 65℃ (Supplementary Table 8.2 and Table 8.3). This resulted in generation of higher amount of PCR products (Supplementary Figure 8.4).

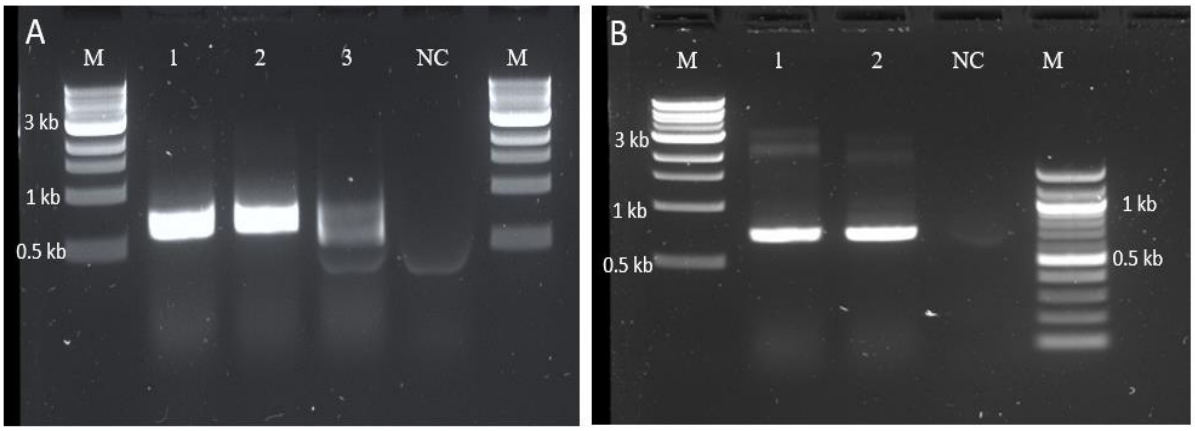


Figure 4.5. PCR amplified fragments of potato *eIF4E* mRNA CDS cDNA. M = molecular weight marker (1 kb DNA ladders); NC = negative control (no template control). Image (A); samples 1 and 2: PCR amplified cDNA fragments of *eIF4E* CDS from Desirée cultivar potato, sample 3: PCR amplified cDNA fragments from Innovator cultivar, Image (B); samples 1 and 2: PCR amplified cDNA fragments from Innovator cultivar.

For cDNA amplification, all the PCR reactions except for the control reactions (NC) showed amplicons of the expected sizes, i.e., approximately 700 bp (Figure 4.5), suggesting specific template amplification. However, PCR amplification using cDNA from the Innovator cultivar resulted in very low amounts of PCR product (Figure 4.5 A, lane 3). This PCR was repeated where the initial denaturation time at 98 °C was adjusted to 20 sec, then both denaturation time and annealing time were adjusted to 30 sec instead of 20 sec each. This resulted in the amplification of high amount of the PCR product of the expected size (data not shown).

4.2.3 Cloning of PCR amplicons to prepare for sequencing

The PCR amplified gDNA (D1-D7) and cDNA fragments (Figure 4.4 and Figure 4.5) were cloned and used to transform *E. coli*. To identify putative positive clones, colony PCR was performed using the same primers as used for amplification and the resulting PCR reactions were analysed on 1% agarose gel (Figure 4.6 and Figure 4.7). In total 14 well-separated colonies for each of the gDNA transformation reactions (i.e., 7 reactions, D1-D7) were tested.

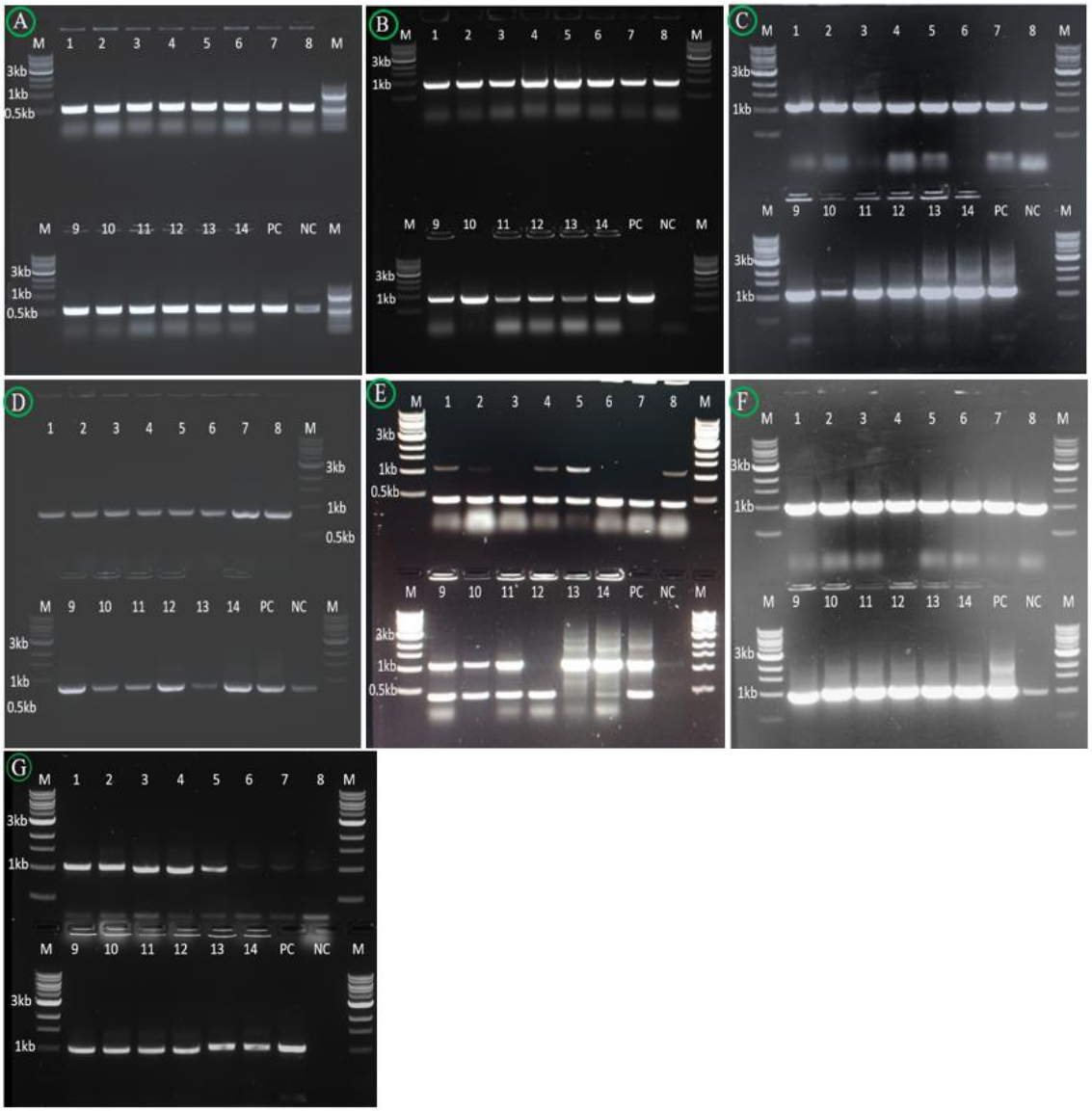


Figure 4.6. Agarose gel electrophoresis of putative positive clones identified by colony PCR. Images A-G represents colony PCR reactions to analyse the presence cloned PCR fragments (genomic fragments) amplified with primer set 1-7, respectively. M: 1 kb or, 100bp DNA ladders; 1-14, colony PCR reactions for 14 different colonies (i.e., reactions showing DNA fragments are putative positive); PC, positive control (i.e., specific amplicon in diluted concentration); NC, no template control reaction.

Table 4.3. Percentage analysis of putative positive clones identified by colony PCR using seven different primer sets (Table 4.2). A total of 14 *E. coli* colonies were analysed to identify the colonies harbouring plasmid with PCR amplified, Desirée *eIF4E* genomic amplicons (Figure 4.4) as insert.

Images	**Transformation rxn**	**Primer sets used**	**Colonies**	**Positive clones (%)**
A	D1	St_eIF4E_1F/ 1R	14	100
B	D2	St_eIF4E_2F/ 2R	14	100
C	D3	St_eIF4E_3F/ 3R	14	100
D	D4	St_eIF4E_4F/ 4R	14	100
E	D5	St_eIF4E_5F/ 5R	14	71
F	D6	St_eIF4E_6F/ 6R	14	100
G	D7	St_eIF4E_7F/ 7R	14	100

The PCR positive clones were analysed for all the seven different transformation reactions. Of this, 6 out of 7 reactions gave 100% positive clones, however for one transformation reaction i.e., D5 (Figure 4.6 E) the percentage of positive clones was 71% (Figure 4.6, and Table 4.3). Additionally, samples 6, 7, and 8 for transformation reaction -7 i.e., D7 (Figure 4.6 G, panel 1, lane 7-9) showed relatively low amount of DNA fragments. On the other hand, the negative controls (NC) for reactions 1, 4, and 6 (Figure 4.6 A, D, and F, panel 2, lane 9, respectively) also seemed to contain low amount of DNA fragments of the same size range as that of positive control fragments (reaction specific amplicons), suggesting the presence of clones in the reaction (i.e., contamination in the negative control reactions). Furthermore, in addition to the DNA fragment of the expected size, another lower-sized DNA fragment was also amplified in most of the reactions for transformation reaction 5 i.e., reaction D5 (Figure 4.6 E), the nature of which was not clear, but probably a result of unspecific amplification.

Correspondingly, for Desirée and Innovator specific cDNA transformation reactions, 19 well-separated colonies were analysed, while for Celandine, and Kuras specific cDNA transformation reactions, 15 colonies were analysed.

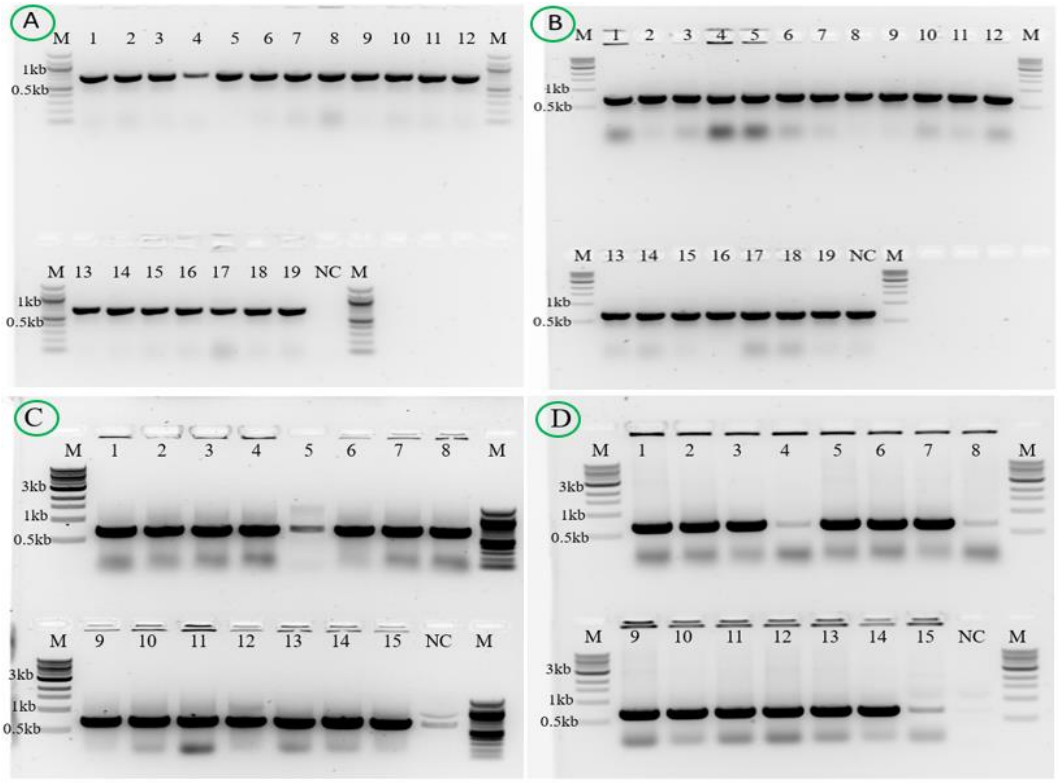


Figure 4.7. Putative positive clones identified by colony PCR. Images A- D represents colony PCR reactions to analyse presence of *eIF4E* cDNA from Desirée, Innovator, Celandine and Kuras cultivar as insert in the *E. coli* harbouring plasmid. M: 1 kb or, 100 bp (for image A and righthand side corner of image C) DNA ladders; 1-14, colony PCR reactions for 14 different colonies (i.e., reactions showing DNA fragments are putative positive); PC, positive control (i.e., specific amplicon in diluted concentration); NC, no template control.

Table 4.4. Percentage analysis of putative positive clones identified by colony PCR using primer sets (St_eIF4E_1F and St_eIF4E_7R, i.e., St_eIF4E_1F/7R). A total of 19 *E. coli* colonies were analysed for Desirée-cDNA and Innovator-cDNA transformation reactions and 15 colonies for Celandine-cDNA and Kuras-cDNA transformation reactions were analysed to identify the colonies harbouring plasmid with PCR amplified, cultivar specific cDNA amplicons as insert.

Image code	Transformation rxn	Primer set used	Colonies	Positive clones (%)
A	Des_eIF4E- cDNA	St_eIF4E_1F/ 7R	19	100
B	In_eIF4E-cDNA	St_eIF4E_2F/ 7R	19	100
C	Ce_eIF4E-cDNA	St_eIF4E_3F/ 7R	15	100
D	Ku_eIF4E-cDNA	St_eIF4E_4F/ 7R	15	100

The PCR positive clones were analysed for all four different transformation reactions, specific to four different cultivars, Desirée, Innovator, Celandine, and Kuras. All the colony PCR reactions for all four transformation reactions, showed presence of DNA fragments of expected size, i.e., around 700 bp. It was clearly indicative that all four transformation reactions gave 100% positive clones (Figure 4.7, and Table 4.4). However, the negative control reactions (NC) for Innovator (Figure 4.7 B) also showed high amount of DNA fragments in the same size range as expected products, whereas that for Celandine showed low amount of DNA fragments in same size range, both indicating contamination in the negative control reactions (presence of cDNA clones). Furthermore, PCR reactions 5 for Celandine and PCR reactions 4, 8, and 15 for Kuras (Figure 4.7 images: C, panel 1, lane 6 and D, panel 1, lane 5, 9 and panel 2, lane 8 respectively) showed relatively low amount of DNA fragments.

4.2.4 Identification of *eIF4E* genomic sequence for Desirée cv. and mRNA sequences for four different cultivars

Ten positive clones per amplicon were cultured overnight and the plasmids were isolated. The presence of Desirée gDNA (not performed for cDNA) fragment insert was further assured by performing restriction digestion analysis (results presented in Supplementary Figure 8.12 - Figure 8.14). A total of ten plasmid samples per transformation reaction were later sequenced by Sanger sequencing. Raw sequence data from Sanger were cleaned and trimmed. All the sequences corresponding to seven amplicons for gDNA were assembled to generate *eIF4E*

consensus sequence. Correspondingly, all the cDNA sequences corresponding to each cultivars were assembled and analysed separately to generate consensus sequences (see Supplementary Figure 8.5 - Figure 8.8). Finally, the total length of sequences retrieved were noted (Table 4.5). Clustering of the CDS sequences resulted in 3 representatives for Desirée cultivar, 2 representatives for Celandine, 4 for Kuras whereas 5 representatives were noted for Innovator, thus indicating multiple allelic variants. Additionally, each individual consensus sequences were aligned with the corresponding reference sequences to analyse the actual gene length or length of mRNA CDS and the number of polymorphisms within the consensus sequences were noted for each cultivar.

Table 4.5. Determination of number of polymorphisms in cultivar specific sequence analysis. The 10 sequences retrieved from Sanger sequencing were aligned and the total length covered by the sequencing along with polymorphisms within these sequences were noted. Further, the actual length of mRNA encoding gene sequence and protein coding (CDS) were also noted.

Cultivar	**Sequence**	**Covered length (bp)**	**Gene/ CDS length (bp)**	**Polymorphism**	**Indels**	**Base substitution**
Desirée	Genomic	3956	3936	114	5	109
Desirée	CDS	716	696	14	-	14
Kuras	CDS	716	696	16	-	16
Celandine	CDS	716	696	5	-	5
Innovator	CDS	716	696	15	-	15

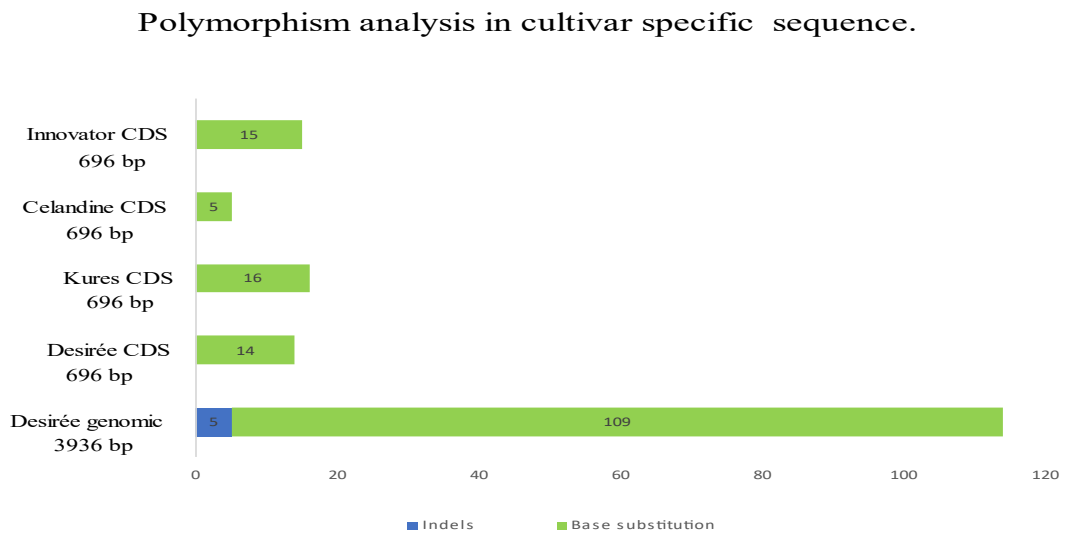


Figure 4.8. Graphical representation of the polymorphism (indels and base substitution) in cultivar specific sequence analysis. The sequences determined from Sanger sequencing were assembled and inclusive consensus was generated. The polymorphisms were identified, and the counts were noted in Table 4.5 and a graph was prepared using the data. The blue box represents number of indels, while the light green boxes represent base substitution (SNPs).

The assembling and aligning of all the *eIF4E* genomic sequences generated a consensus of 3956 bp length. Whereas a total of 716 bp long *eIF4E* cDNA consensus sequences were generated for Desirée, Innovator, Celandine, and Kuras cultivars. The genomic consensus was aligned with pre-reported *eIF4E* genomic sequence (Phytozome; S. tuberosum v6.1|Soltu.DM.03G000970) and the actual length of the coding genomic sequence was noted to be 3936 bp (sequence details on Supplementary documents C 7, p-95). Similarly, the cDNA consensus sequences were aligned with pre-reported *eIF4E* CDS sequence (NM_001288431.1), and the length of mRNA CDS (protein coding sequence) were noted to be 696 bp.

Furthermore, it was noted that the genomic sequence contained 114 different polymorphisms out of which, five were indels located to intronic sequences of the gene, while 109 were base substitutions (or, SNPs). Out of 109 SNPs, 14 SNPs were in the exonic regions whereas the other 95 SNPs were distributed in the intronic regions of the gene (Table 4.5 and Figure 4.8). Additionally, a total of 14, 16, 5, and 15 SNPs were noted in mRNA CDS sequences resulting in 3, 2, 4 and 5 different allelic variants for Desirée, Celandine, Kuras, and Innovator cultivars respectively (see Appendix, Supplementary documents- C, p- 96-101).

4.3 Designing, in-vitro transcription and assessing efficacy of sgRNAs

4.3.1 Designing and in-vitro transcribing seven different sgRNAs

Two different online sgRNA design tools, CRISPOR & CRISPRdirect were used to design sgRNAs using the different CDS sequences identified during this study (Supplementary documents C- 95). Seven different sgRNA spacer sequences with off-target activities were obtained and postulated as potential candidates (Table 4.6).

Table 4.6. sgRNA spacer sequences designed for gene editing of potato *eIF4E* gene in Desirée cultivar. The sizes of each spacer sequences were set to be 20 nucleotide long.

S.No.	sgRNA spacer sequences	Size	PAM	Off-targets	Tool used
1	AATGATACGGCGTCGTATTT	20	GGG	4	CRISPOR
2	AGTCGTTAGTGTCCGGTCTA	20	AGG	1	CRISPOR
3	CAGCAGCGTATATAGCCAGC	20	TGG	4	CRISPOR
4	GTCGTTAGTGTCCGGTCTAA	20	GGG	4	CRISPOR
5	GTGGAGCAGTCGTTAGTGTC	20	CGG	3	CRISPOR
6	GTTTCTAGATCATAGCGATT	20	CGG	3	CRISPRdirect
7	TGGAGCAGTCGTTAGTGTCC	20	GGG	-	CRISPRdirect

Using these seven different sgRNA spacer sequences seven IVT- forward primers were designed to PCR amplify DNA templates containing sgRNA encoding sequences. The forward primers were synthesized and provided by Invitrogen, which were then used to PCR amplify sgRNA templates. These templates were in-vitro transcribed and purified to synthesize higher amount of sgRNA. The purity and quantity (Table 4.7) were evaluated by NanoDrop.

Table 4.7. Quantity and purity analysis of in-vitro transcribed and purified sgRNAs using Nanodrop spectrophotometer. Concentration of sgRNAs were analysed in ng/ µl, and total sgRNA yield was calculated in µg, whereas the purity of sgRNAs were analysed by observing 260/230 values.

sgRNAs	**Conc. (ng/µl)**	**Yield (µg)**	**260/280**	**260/230**
1	429	8.58	2.28	2.60
2	603.8	12.07	2.24	1.94
3	806	16.12	2.16	2.25
4	375.8	7.5	2.29	2.26
5	728	14.56	2.25	2.53
6	126	2.52	2.22	2.25
7	457	9.14	2.28	2.19

It was evident from Table 4.7 that sgRNA number 3, 5, and 2 were synthesized in higher amounts followed closely by sgRNAs 1, 7, and 4 with decent amount of sgRNAs being synthesized, whereas sgRNA-6 was produced in very low amount. However, the amount synthesized would be enough to perform further reactions and the purity value was almost within the expected range (2 - 2.2) so, the process was not repeated. All the sgRNAs seemed to be almost within the expected purity range, except for sgRNAs 1, 5, and 2 were little off the limit, indicating possible contamination.

To have an overview of how these sgRNAs would target the *eIF4E* gene in-vivo (inside the plant cell), each spacer sequences were searched and located on the genomic sequence, and a basic representation of this overview is provided in Figure 4.9. However, during this analysis, sgRNA spacer 3 was noted to be targeting in the exon-exon junction region between exon two and three. Out of 20, 14 Nucleotides were targeting 14 bases from exon-2 whereas six remaining bases were targeting six initial bases from exon-three as illustrated in Figure 4.9, indicating it cannot be used for the *eIF4E* gene editing.

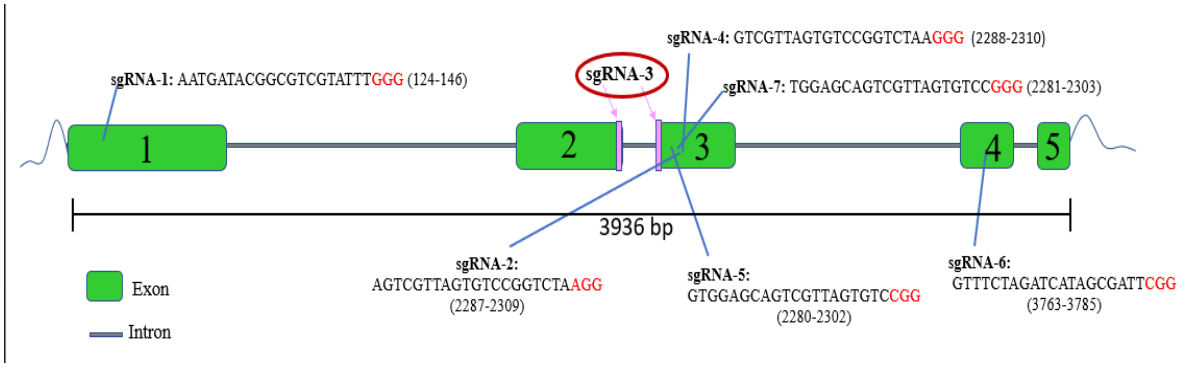


Figure 4.9. Schematic representation of possible resemblance of the target positions for seven different sgRNAs (Table 4.4) in the *eIF4E* genomic locus of the Desirée plant. The location of target sites is presented by engineered sgRNAs along with their sequences and the PAM motifs are indicated in red. Green boxes represent exon 1-5.

4.3.2 Assessing efficacy of sgRNAs by in-vitro cleavage assay

To determine the specificity and efficacy of the in-vitro transcribed and purified sgRNAs, in-vitro cleavage assay was performed. For this assay three different plasmids containing three different Desirée cultivar specific allelic cDNA sequences (St_Des-eIF4E-CDS_1, St_Des-eIF4E-CDS_2, St_Des-eIF4E-CDS_5) were used. The circular plasmids of 4228 bp sizes (Supplementary Figure 8.3) were in-vitro cleaved by RNP complexes (i.e., Cas9 protein + in-vitro transcribed sgRNAs) and then digested by Hinc II enzyme to generate two different sized DNA fragments (Figure 3.2) and the reactions were analysed by agarose gel electrophoresis (Figure 4.10).

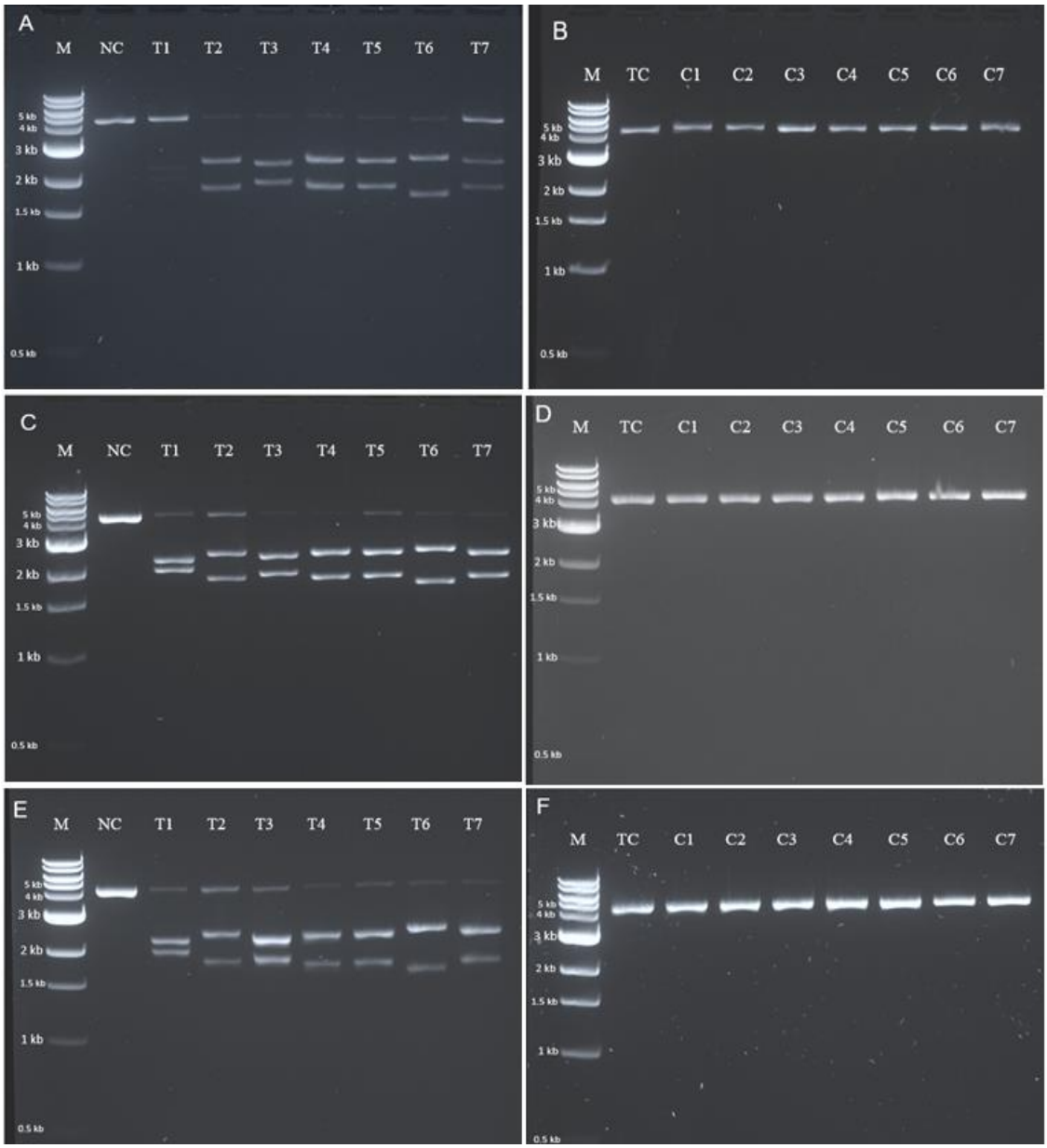


Figure 4.10. Schematic representation of agarose gels harbouring the different DNA fragments from in vitro cleavage assay. Images A & B; cleavage reactions for PCR z.b._De-eIF4E-1 plasmids, images C & D; reactions for PCR z.b._De-eIF4E-2 plasmids, whereas E &F; reactions for PCR z.b._De-eIF4E-5 plasmids. M refers to 1 kb DNA ladder as DNA marker; NC refers to negative control (i.e., circular plasmid – Cas9 – sgRNA + Hinc II), T(1-7) refers to treatment reactions (i.e., circular plasmid + Cas9 + sgRNA 1-7 + Hinc II), whereas TC refers to treatment control (i.e., plasmid + Cas9 – sgRNA + Hinc II) and, C(1-7) refers to sgRNA specific control reactions (i.e., plasmid – Cas9 + sgRNA 1-7 + Hinc II).

All the treatment reactions (T1-T7) i.e., cleavage reactions, presented two DNA fragments within the expected size (Figure 4.10 A, C, and E), suggesting that the circular plasmid DNA used in the reactions were cleaved at two different positions. Furthermore, all the control reactions, i.e., negative controls (NC- no Cas9 and sgRNA control), treatment controls (TC- added Cas9 but no sgRNA), and C(1-7: no Cas9 but added sgRNA) reactions

showed a single DNA fragment at around 4200 bp size, clearly indicating that all three of the circular plasmids were cleaved and linearized approximately to the expected size i.e., 4228 bp (Figure 4.10). Hence, it was confirmed that the cleavage in the treatment reactions were Cas9 specific i.e., sgRNAs together with the recombinant Cas9 protein (RNP complexes) were able to cleave their target regions in cDNA specific to Desirée *eIF4E*. However, relatively low amount of DNA fragments was obtained from treatment reactions, T1 and T7 for PCR z.b._De-eIF4E-1 plasmid (Figure 4.10 A, lane 3 and 9), indicating inefficient cleavage by the Cas9 protein guided by sgRNA 1 and 7.

4.4 Protoplasts were isolated from the potato leaves

Protoplasts were isolated from mesophyll from Desirée and Kures cultivars according to (Nicolia et al., 2021) protocol with few modifications. Two trials were conducted. The first trial employed protoplast isolation from; 24 leaves from 6 weeks old potato plantlets collected and immediately processed from Desirée cultivar, and 24 leaves (mostly tiny leaves) from 6 weeks old Kuras plantlets collected one hour before processing (Figure 4.11). Thus, 1.0 X 10^4 protoplasts/ ml were isolated from Desirée cultivar. However, no viable protoplasts were obtained from Kuras cultivar.

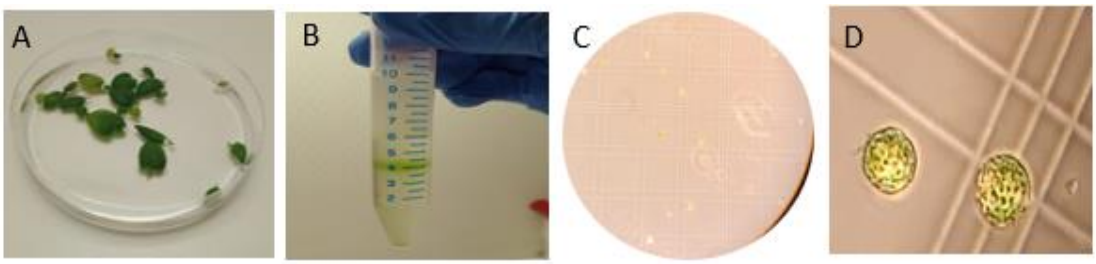


Figure 4.11. Pilot test for isolation and purification of potato protoplasts Desirée and Kuras cultivar by employing (Nicolia et al., 2021) protocol. (A) 24 leaves collected from 6-weeks-old Desirée plantlets. (B &C) Freshly isolated, washed, and purified protoplasts from Desirée leaves. Whereas no protoplasts were obtained from Kuras leaves (hence, figure not shown).

Table 4.8. Analysis of protoplasts isolated from two different cultivars, during two different tests. The number of leaves used were noted and protoplast count was also noted.

Test	Cultivars	Leaves no.	Plant age	Protoplast count/ml
First	Desirée	~ 20	6 weeks	1.0 X 10^4
	Kuras	~ 20	6 weeks	0
Second	Desirée	~ 40	~8 weeks	1.0 X 10^4
	Kuras	~ 40	~8 weeks	2.0 X 10^5

The second trial for protoplast isolation was conducted using 40 moderate sized leaves from almost 8 weeks old Desirée and Kuras plantlets. The layering of Desirée protoplast sample on top of sucrose solution was performed using Pasteur's pipette in a faster pace, whereas that for Kuras cultivar protoplast was performed slowly using cut tip pipette. This resulted in isolation of 1.0 X 10^4 protoplasts/ ml from Desirée cultivar, whereas 2.0 X 10^5 protoplasts/ ml from Kuras cultivar (Figure 4.12 and Table 4.8).

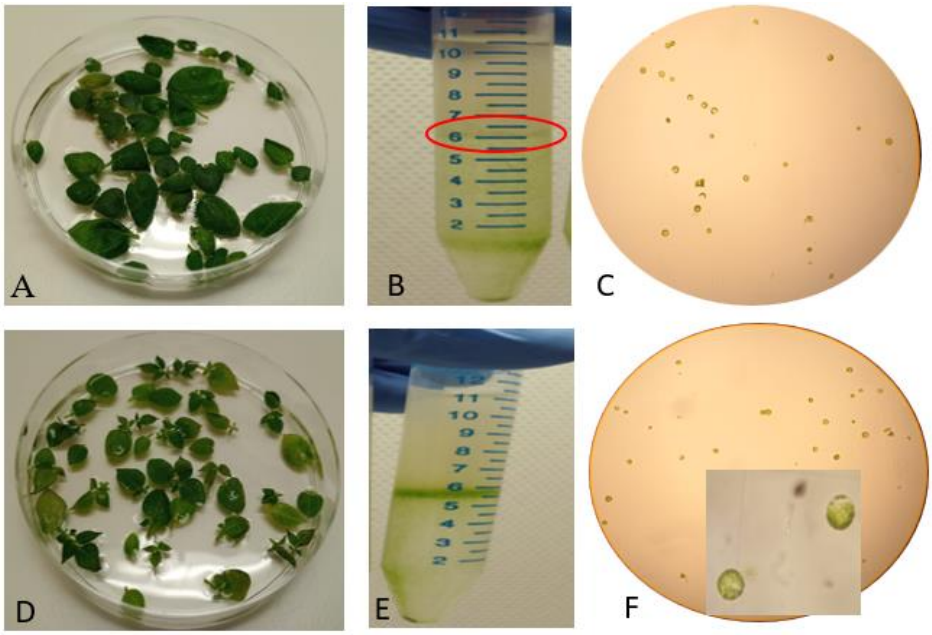


Figure 4.12. Isolation and purification of potato protoplasts by employing (Nicolia et al., 2021) protocol. (A and D): Leaves from 8-weeks-old Desirée and Kuras plantlets respectively. (B and E): Phase separation step for Desirée and Kuras specific protoplasts respectively with the viable protoplast

ring in the interface. (C and F): Freshly isolated and washed protoplasts from Desirée and Kuras cultivar respectively.

Finally, a transfectability analysis of the viable protoplasts was performed by transformation of the isolated viable protoplasts (Figure 4.12) with the plasmid expressing Cas9 protein and RFP reporter. PEG mediated transfection of protoplast was employed in this study. However, the result of this analysis is not presented in this study.

5. Discussion

Since few decade, plant breeders have been developing and widely utilizing various gene technologies in plant research and breeding (Eriksson et al., 2018). The recent breeding method for plants is plant mutation breeding (Udage, 2021), using techniques for targeted genome editing. One of the most efficient technique is CRISPR/Cas system (Eriksson et al., 2018) of gene editing which helps create targeted mutagenesis in the genome, thus altering the gene expression. Conventional CRISPR/Cas9 system relies on a target specific sgRNA which directs the SpCas9 protein to the specific target site to introduce DSB in the DNA(Barman et al., 2020; Hahn et al., 2020). This DSB is then repaired by cell mostly using NHEJ method which introduces mutagenesis in the genome (Barman et al., 2020).

This experiment-based study has thus focused on establishing various candidate sgRNAs for precise and efficacious potato *eIF4E* genome editing in potato protoplast. The designed and in-vitro transcribed sgRNAs together with SpCas9 protein, are intended to introduce DSB in the DNA. A total of seven different combinations of SpCas9 and sgRNAs were assessed by in-vitro cleavage assay. These assays showed that all the RNP complexes assessed in-vitro presented target specific cleavage of the DNA inserts. This confirmed that these RNP complexes can perform a double stranded DNA break (DSB), thus presenting a promising usability of these complexes for in-vivo genome editing in plants.

The reasons for choosing Desirée, Birkeland, Kures, Innovator, and Celandine cultivars for study is that Desirée is the most used potato variety in molecular and physiological studies. Whereas the other cultivars are commercially important cultivars around Scandinavian countries and are susceptible to PVY infection.

5.1 Retrieval of potato *eIF4E* gene and CDS sequence information from databases

Plant genome editing in polyploid crops requires editing of specific genomic sequence as it can affect the time required for editing as it aids in omitting necessity of crossing and screening (Jansing et al., 2019). Hence, retrieval of *eIF4E* gene sequence specific to target cultivars used in this study was crucial. Various databases are established to ease the storage and retrieval of the genomic sequence information. NCBI is the largest online resource for biological information and data from 34 specific databases, including nucleic acid sequence database and the database for published life science journals (Sayers et al., 2021). Thus, during the earlier

phase of this study, thorough searches were conducted in NCBI database, to retrieve the pre-reported *eIF4E* gene and CDS sequence information belonging to the five different cultivars (i.e., cv. Desirée, Birkeland, Kuras, Celandine, and Innovator) specific *eIF4E* gene and its mRNA CDS.

The NCBI database searches showed just 14 nucleotide hits belonging to potato *eIF4E* gene specific to three different cultivars: Russet Burbank, Zhukovskiy ranniy, and Solyntus. Furthermore, it was also noted that potato *eIF4E* gene family consists of four different forms: *eIF4E1*, *eIF4E2*, *eIF(iso)4E*, and *nCBP*. This was further confirmed by Lucioli et al.,(2022), where they have even provided information about chromosomal locations of these genes.

Furthermore, accession number search for such pre-reported sequences was also performed by reviewing various studies from NCBI PubMed database. This search resulted in retrieval of accession numbers belonging to various *eIF4E* allelic mRNA CDS sequences. However, the target cultivars specific gene sequence information could not be located. The reason for this could be that the cultivars of interest are mostly commercial cultivars which are usually not subjected for studies, instead using of model plant is preferred.

Hence, to determine the target cultivar specific *eIF4E* sequences, the potato specific *eIF4E* sequence information was retrieved for designing the primers for amplification of the gene or mRNA. This would also facilitate further in the analysis of polymorphisms, and in target cultivar specific gene editing, which will be discussed later. To identify and retrieve the potato *eIF4E* sequence, FASTA sequence for accession number NM_001288431.1 belonging to potato *eIF4E* mRNA complete CDS was retrieved and used. The reason behind choosing this sequence was that it was emphasized in NCBI as the *eIF4E* mRNA reference sequence. Additionally, it was the only sequence information with (NM_) prefix as per evident from the NCBI searches and accession number searches. NM_ prefix is only given to the curated mRNA sequences. Furthermore, Phytozome is a comprehensive database for analysing and comparing of data corresponding to plant genome and gene family (Goodstein et al., 2012). So, the retrieved sequence was blasted against the target: Solanum tuberosum v6.1-potato in Phytozome database. The blast search helped identify exact location of the gene in the chromosome number 3 and the potato *eIF4E* gene sequence was retrieved with flanking sequences upstream and downstream. This is essential to design overlapping primer sets to ensure the coverage of almost entire genome sequence.

5.2 Primers designing and PCR amplification

PCR is a sensitive enzymatic reaction that helps in amplification of even low amount of target gene fragment, facilitating the nucleic acid analysis process. Nonetheless, its efficacy is fully dependent on the efficacy of the primers used in the reaction (Kumar & Chordia, 2015). Hence, designing of target specific and efficient primers can have a positive effect on PCR. So far, keeping the exon-intron boundaries into consideration, seven different primer sets (Table 4.2) were designed to cover almost the entire *eIF4E* gene sequence during the PCR amplification of fragments, by employing online Primer3 software. Primer3 software was used as it is an open online source for designing primers for application in molecular field. Additionally, Primer3 software uses thermodynamic models to predict the melting temperature, reduces primer dimers formation and improves primer specificity (Untergasser et al., 2012). Each designed primer was further in-silico validated by blasting against the target: S. tuberosum v6.1 in Phytozome database to analyse their unique representation in potato genome. The reason behind this representation analysis is to assure that the primes designed are specific to the target gene.

The PCR amplification of the gDNA and cDNA were performed using Phusion Hot Start II HF DNA polymerase and specific primer pairs. The reason for choosing this polymerase for PCR amplification was that Phusion polymerases has 3´ to 5´ exonuclease (proofreading) activity together with 5´ to 3´ polymerase activity. It means that these enzymes have proofreading activity which helps to ensure that the dNTPs incorporated during each PCR amplification step are template specific. This is essential for the reactions or, amplicons to be sequenced to ensure the sequence information extracted is unbiased. Additionally, the goal was to perform blunt end cloning, a type of PCR cloning method. So, it was essential to get blunt ended PCR amplicons, which can be formed by Phusion DNA polymerase.

The PCR amplification reactions were analysed on agarose gels. The result from this analysis showed that all the PCR reactions contained the amplicons of expected size, and the amplification was template specific. However, very low amount of DNA amplicons was obtained from PCR amplification of template gDNA from Desirée cultivar, using primer set 7. The reason for this could be due to the inadequate time for dissociation and extension, suboptimal annealing temperature. Hence, 4 µl (~140 ng) of template gDNA was added to reaction, the gradient PCR program was performed (according to Supplementary Table 8.2 and Table 8.3). Finally, high amount of expected sized DNA amplicons were obtained. The reason

of choosing gradient PCR was that it helps to optimize the PCR amplification mainly by optimizing the primer annealing condition.

5.3 Importance of blunt-end cloning and colony PCR

The blunt-ended gDNA and cDNA amplicons were then cloned into PCR® zero-blunt vector and the *E. coli* were transformed. The reason behind using this vector was that the Phusion DNA polymerases generates blunt ended PCR products or amplicons. Correspondingly, the vector used is a linearized vector with blunt ends. Hence, this could omit the time-consuming process of digesting of both the vector and insert fragments before ligation. In addition, PCR zero blunt vector also contains kanamycin selective marker, facilitating selective growth of the bacteria that harbours the ligated plasmids.

The bacterial colonies containing the target DNA amplicons (putative positive clones) were identified by performing colony PCR. The reason for this was, colony PCR helps to identify the presence of the targeted DNA amplicon in bacteria grown in presence of selective media after transformation (Bergkessel & Guthrie, 2013). This is essential as the transformed bacteria could be harbouring the empty vector and still forming colony. As the target was to sequence the DNA amplicons, it was mandatory to analyse the samples before isolating the plasmids and sequencing them. The results from this analysis suggested that almost all the transformation reactions were putative positive. However, few negative controls seemed to contain DNA amplicons. Looking back at the process of culturing the transformed *E. coli*, the transformation reaction was spread over the entire LB- kanamycin surface in the plate. So, while picking up negative control from area without visible colonies, there is still possibility that few of the putative positive cells could have been picked. Finally, ten positive clones per transformation reaction were picked, plasmids were isolated, and Sanger sequenced.

5.4 Sanger sequencing of genomic *eIF4E* and its mRNA sequence for four different potato cultivars

As the *eIF4E* sequences specific to the target cultivars were not noted in the databases, the initial step was to determine the cultivar specific *eIF4E* sequence. Since the target cultivar potatoes are tetraploid, they contain higher number of SNPs and indels within and between the cultivars. Hence thorough analysis and assurance of the specific allelic sequences is essential to design the target specific sgRNAs for CRISPR/Cas9 based gene editing (Carlsen

et al., 2022). Correspondingly, Lucioli et al., (2022) also mentions that commercial potato cultivars have heterozygous genome, so it is crucial to analyse the polymorphism between the alleles and cultivars to avoid unintended effects of gene editing.

Thus, the *eIF4E* gene or mRNA CDS sequences were determined for four different cultivars of potato (Desirée, Kures, Celandine, and Innovator) by Sanger sequencing method. The reason for using Sanger sequencing method was that it is most reliable method, relatively cost effective and efficient for DNA fragments sequencing. It also has capacity to analyse long reads in comparison to next-generation sequencing techniques. The sequences extracted were assembled, and the polymorphisms were analysed. Moreover, 114 polymorphisms were observed within Desirée genomic sequence, five of them being indels which were located within the intron regions of the gene. Correspondingly, the CDS sequences belonging to different cultivars also presented multiple polymorphisms all being SNPs. So far, four different allelic variants were detected from Kuras, three from Desirée cultivar, whereas two from Celandine and five from Innovator (sequence information on Appendix, C p- 96- 101).

All these cultivars of potato are tetraploid crops and have four different alleles of a specific gene. The reason behind recovery of just three and two allelic sequences could be that only these alleles were being highly expressed during RNA isolation and they could have competed out the other alleles during RT-PCR amplification. Furthermore, the reason for detection of five allelic variants from Innovator could be due to epigenetic changes (mutation due to environmental effect) or, some somatic variation in the selected tissue culture. Supporting this point, Xu et al., (2011) mentioned that the heterozygous diploid potato clone used in their study reported that the possible effect of the somatic variation and inbreeding depression could introduce mutation thus introducing gene presence/absence variants. Moreover, these sequence information could be very useful for the further studies of these cultivars, like analysing type of mutation and the possible resulting protein study and altering functions. This information could help later for allele or cultivar specific gene targeting for precise and efficient gene editing.

5.5 Designing and in-vitro transcription of sgRNAs

CRISPR/Cas genome editing is powerful tool, which requires three main components for gene editing, Cas protein, sgRNA and PAM sequence. The sgRNA designing is essential to select the precise location of the targets (protospacers) within the gene thus determining the site of edition (Gerashchenkov et al., 2020). For designing the sgRNAs, all 27 online tools as per

mentioned in Table 2 in Gerashchenkov et al., (2020) were analysed, out of 27 tools listed, only nineteen tools were accessible. Furthermore, just seven tools out of them were noted to contain sequence data for potato genome. However, three tools namely, CRISPOR, CRISPRdirect, and Breaking Cas, were found to be most convenient tools. However, Breaking Cas did not provide specific off-target information. So, just CRISPOR and CRISPRdirect tools were used to design five and two sgRNAs respectively.

The allelic CDS sequences from Desirée, Celandine, and Kuras cultivar were used to design the candidate sgRNAs. Seven different sgRNAs were then in-vitro transcribed and purified. The result analysis from (Table 4.7) showed that out of seven, six sgRNAs were synthesized in high amount, whereas sgRNA number 6 was in relatively low amount. The reason for this could be due to the suboptimal amplification conditions, or even loss during purification. Moreover, the reason behind using the *eIF4E* mRNA CDS sequence for designing sgRNAs was that the cDNA was the complete sequence. Additionally, the aim was to design the sgRNA in the exotic region of the cDNA. This could help further target the exonic sequences in the gene in-vivo. However, the analysis of the target positions for these sgRNAs in the *eIF4E* genomic locus of the Desirée plant (Figure 4.9) showed that sgRNA number 3 was designed to target in the exon-exon junction. This means that out of 20 nucleotides, 14 targets at the end of exon-2 whereas remaining 6 nucleotides targets starting 6 bases of the exon-3 of Desirée *eIF4E* gene. Hence, sgRNA number 3 cannot be fully trusted for in vivo gene editing, as the in-vivo target is gDNA and not cDNA.

5.6 In-vitro analysis of efficacy of the sgRNAs

The purified sgRNAs were then analysed for their efficacy, using in-vitro cleavage assay. Karmakar et al., (2021) mentioned in their study that, efficiency of the Cas9 protein to cleave the target DNA strongly depends on engineered sgRNAs guiding it. Even though the online tools for designing sgRNA predicts numerous candidate sgRNAs for a specific target site, only few sgRNAs shows efficient cleavage (Karmakar et al., 2021). Thus, in-vitro cleavage assay helps to analyse the specificity and efficacy of the designed sgRNAs in-vitro before using it for in-vivo gene edition in the plant cells. Additionally, Sagarbarria & Caraan, (2023) mentioned that designing of sgRNA for plant genome editing is quite challenging as many of the in-silico sgRNA design tools use guidelines based on animal experiments. So, it is mandatory to analyse the designed sgRNAs in-vitro before in-vivo gene editing.

In this study, using seven different SpCas9 – sgRNA (RNP) complexes, three different plasmids containing allelic cDNA specific to Desirée were cleaved in-vitro. The assay results clearly indicated that all the RNP complexes were able to cleave their target regions in all allelic cDNA specific to Desirée *eIF4E*. However, RNP complexes 1 and 7 (Cas9 protein guided by sgRNA 1 and 7) were noted to be less efficient while cleaving De-eIF4E-1 allele of Desirée *eIF4E* (Figure 4.10 A). The reason for the low efficiency of sgRNA 1 is not clear as the spacer sequence and PAM sequence are noted to be specific. Whereas the possible reason behind low cleavage efficiency by sgRNA 7 against De-eIF4E-1 allele could be, due to the change in the third nucleotide of PAM sequence, i.e., the target PAM sequence for sgRNA 7 is 'GGG' whereas De-eIF4E-1 allele contains 'GGT' instead. The SpCas9-sgRNA complex identifies and binds to the 'NGG' PAM sequence, thus unwinding the two strands of DNA to check complementarity for performing DSB. Hence, this change in the PAM sequence could have affected the binding of the SpCas9 with the target, finally hindering efficient cleavage of the target DNA.

Additionally, the annotation analysis on the three different insert sequences in the plasmids used, indicated that sgRNA 1, 3, and 5 are specific to all three inserts. Whereas the sgRNA 2 and 4 are specific to De-eIF4E-1 and have single base change within the seed sequence in both De-eIF4E-2 and De-eIF4E-5 alleles. Correspondingly, sgRNA 6 and 7 are specific to alleles De-eIF4E-2 and De-eIF4E-5, and sgRNA 6 has two bases change in De-eIF4E-1.

The result from this study clearly suggests that the sgRNAs designed and synthesized during the study are efficient and have potential to cleave the Desirée specific *eIF4E* gene in-vitro. Thus, it can be indicative that the RNP complexes analysed could possibly display DSBs in the *eIF4E* gene in-vivo. However, due to the timeline of the study, further in-vivo analysis or genome editing was not conducted. Nonetheless, the finding from this study can contribute to a numerous further studies.

Moreover, the findings from this study are totally based on the in-vitro assay. Although it indicated that these sgRNAs or, RNPs have potential to cleave the Desirée specific *eIF4E* gene in-vitro, it cannot be ensured that they will show the same efficacy level in-vivo. A study by Sagarbarria & Caraan, (2023) performed both in-vitro and in vivo assay of four different sgRNAs designed to target the eggplant polyphenol oxidase genes. Their result indicated that even though few of the sgRNAs efficiently cleaved the gene in vitro, they could not produce edits in vivo. The possible reasons for such difference could be that the genes usually exist in their secondary structure inside the plant cell. This suggests that even though the sgRNAs with

high in-vitro cleavage efficacy may not be reliable in in vivo gene editing. So, a quick in-vivo assay could probably help to analyse if these sgRNAs could possibly derive an in-vivo cleavage by guiding a Cas9 protein to the specific target.

5.7 Protoplast were isolated from the potato leaves.

Protoplast has recently been serving as a vital system for performing various genomic and cell-based studies in plant molecular biology and plant genome-editing. It facilitates the delivery of genome -editing tools directly inside the plant cells, thus helping efficient genome editing (Patil et al., 2022). Furthermore, protoplasts-based transfection is often considered to be advantageous in minimizing the risk of stable integration of the recombinant DNA (Nicolia et al., 2021).

In this study, two different protoplast isolation tests were performed from the leave tissues from Desirée and Kuras plantlets. The first study involved use of 24 leaves from 6-week-old plantlets for the protoplast isolation. Whereas the second study involved use of 40 leaves from around 8-week-old plantlets. The first protoplast isolation test resulted in 1×10^4 protoplasts / µl (Table 4.8) from Desirée cultivar, whereas no protoplasts were obtained from Kuras cultivar. The strong reason for not getting protoplast from Kuras could be that the Kuras leaves were picked one hour before processing. Furthermore, these leaves were not placed on any media or water, so this could have led to death of the cells thus affecting isolation of any viable protoplasts.

Additionally, the Nicolia et al., (2021) protocol was not followed strictly which might have contributed to the lower yield of protoplast. The changes were: preparation of Medium C and Medium E without addition of Vitamin 3 (vitamin D3) solution, and substitution of PVP10 with PVP40 due to the unavailability of vitamin D3 and PVP10 respectively. This could have affected the yield of the protoplast; however, further investigations are required to prove this. The second test resulted in production of 1×10^4 protoplasts / µl from Desirée leaves as in prior test, whereas 2×10^5 protoplasts / µl were obtained from Kuras leaves. During the addition of resuspended Desirée protoplasts solution on top of sucrose solution, Nicolia et al., (2021) protocol recommends slow layering on top by using a sterile Pasteur pipette or a micropipette with a cut tip. However, the resuspended protoplasts were added in faster pace so, sudden disruption on the interface was noted. This could have been the reason for the low yield from Desirée leaves as plenty of protoplasts were noted to have sedimented on the bottom

after centrifugation instead of forming a band in the interface, indicating they are not viable (probable cause could be due to cell disruption during the layering).

The overall finding was that the second test produced relatively higher number of protoplasts. This could be due to isolation was performed using almost double count of leaves than in first test and more mature leaves were used. Overall, though vitamin D3 was not added into the medium C, the counts of viable protoplast for Kuras were high and moderate number of viable protoplasts were obtained from Desirée. This clearly indicated that vitamin D3 is not quite essential during protoplast isolation, however further optimization and study is crucial to prove this finding.

Isolation of large number of viable protoplasts plays a vital role in efficient genome editing (Coy et al., 2022). Thus, isolation of high counts of viable protoplasts is crucial. The main reason behind this is that out of millions of protoplasts only few takes up the genome editing components. Furthermore, keeping the transformed protoplasts viable until development into calli has also been a challenge for scientists and breeders.

Finally, the protoplasts isolated during the second test were transfected with the plasmids expressing Cas9, RFP. This study aimed in evaluation of the transfectability of isolated protoplast and later regeneration. However, due to time limit of the project, further evaluations were not conducted regarding this area.

6. Conclusion

Firstly, the *eIF4E* genomic and coding sequence information for four different potato cultivars were determined. The genomic sequence of Desirée specific *eIF4E* gene was noted to contain 109 SNPs and five indels corresponding to different alleles. Correspondingly, *eIF4E* mRNA CDS sequences of Desirée, Kures, Celandine and Innovator showed 14, 16, 5 and 15 SNPs respectively. This clearly indicated the presence of multiple polymorphisms within and between these cultivars of potatoes. The findings suggested that it is essential to design cultivar or allele specific sgRNAs for genome editing in various cultivars.

Secondly, three different plasmids containing Desirée *eIF4E* cDNA fragments were in-vitro cleaved by 7 different CRISPR-RNPs to analyze the efficacy of the designed sgRNAs. The findings from the cleavage assay indicated that most of the CRISPR-RNPs designed and assessed were efficiently in-vitro cleaving the *eIF4E* cDNA. Hence, a similar kind of gene cleavage inside the potato protoplast is expected to finally introduce DSB within the target region.

Finally, for the RNP delivery, a PEG mediated transformation of protoplasts-based method was chosen. So, the protoplasts were isolated from the potato leaves following (Nicolia et al., 2021) protocol. The findings from the result suggest that vitamin D3 is not so vital for protoplast isolation. Additionally, the age of the plantlets used also has some effect on the counts of protoplasts isolated. However, sufficient studies are needed to confirm these findings and the isolated protoplasts in this study were not transformed with CRISPR RNPs.

Thus, the main conclusion of this study was that a basic platform was established for further RNP complex based in-vivo *eIF4E* genome editing, on the Desirée cultivar potato. The sequence information generated during the study could facilitate in the allele specific study and to have further insight into the genomic information of the commercial cultivars used in study and genome editing studies. Also, further study can be conducted where these sgRNAs can be assessed together with multiple other Cas9 proteins for obtaining desired editing like base editing or simple gene knockout. Another area to investigate would be using of several delivery methods to deliver these sgRNAs and Cas9 protein into the cells and analyse if it affects the mutation efficiency. Correspondingly, sgRNAs could be designed and cultivar specific edition can be performed using the sequence information for the other commercial cultivars.

7. Reference

Ala-Poikela, M., Goytia, E., Haikonen, T., Rajamäki, M.-L., & Valkonen, J. P. T. (2011). Helper Component Proteinase of the Genus Potyvirus Is an Interaction Partner of Translation Initiation Factors eIF(iso)4E and eIF4E and Contains a 4E Binding Motif. *Journal of Virology*, *85*(13), 6784–6794. https://doi.org/10.1128/JVI.00485-11

Alonso, J. M., & Ecker, J. R. (2006). Moving forward in reverse: Genetic technologies to enable genome-wide phenomic screens in Arabidopsis. *Nature Reviews Genetics*, *7*(7), Article 7. https://doi.org/10.1038/nrg1893

Álvarez-Fernández, R. (2013). Chapter One - Explanatory Chapter: PCR Primer Design. In J. Lorsch (Ed.), *Methods in Enzymology* (Vol. 529, pp. 1–21). Academic Press. https://doi.org/10.1016/B978-0-12-418687-3.00001-X

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

Anzalone, A. V., Koblan, L. W., & Liu, D. R. (2020). Genome editing with CRISPR–Cas nucleases, base editors, transposases and prime editors. *Nature Biotechnology*, *38*(7), Article 7. https://doi.org/10.1038/s41587-020-0561-9

Baret, P. V. (2017). Acceptance of Innovation and Pathways to Transition Towards More Sustainable Food Systems. *Potato Research*, *60*(3), 383–388. https://doi.org/10.1007/s11540-018-9384-1

Barman, A., Deb, B., & Chakraborty, S. (2020). A glance at genome editing with CRISPR–Cas9 technology. *Current Genetics*, *66*(3), 447–462. https://doi.org/10.1007/s00294-019-01040-3

Barrangou, R., & Doudna, J. A. (2016). Applications of CRISPR technologies in research and beyond. *Nature Biotechnology*, *34*(9), Article 9. https://doi.org/10.1038/nbt.3659

Berdugo-Cely, J. A., Martínez-Moncayo, C., & Lagos-Burbano, T. C. (2021). Genetic analysis of a potato (Solanum tuberosum L.) breeding collection for southern Colombia using Single Nucleotide Polymorphism (SNP) markers. *PLoS ONE*, *16*(3), e0248787. https://doi.org/10.1371/journal.pone.0248787

Bergkessel, M., & Guthrie, C. (2013). Colony PCR. *Methods in Enzymology*, *529*, 299–309. https://doi.org/10.1016/B978-0-12-418687-3.00025-2

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

Browning, K. S., & Bailey-Serres, J. (2015). Mechanism of Cytoplasmic mRNA Translation. *The Arabidopsis Book*, *2015*(13). https://doi.org/10.1199/tab.0176

Burgos, G., Zum Felde, T., Andre, C., & Kubow, S. (2020). The Potato and Its Contribution to the Human Diet and Health. In H. Campos & O. Ortiz (Eds.), *The Potato Crop: Its Agricultural, Nutritional and Social Contribution to Humankind* (pp. 37–74). Springer International Publishing. https://doi.org/10.1007/978-3-030-28683-5_2

Carlsen, F. M., Johansen, I. E., Yang, Z., Liu, Y., Westberg, I. N., Kieu, N. P., Jørgensen, B., Lenman, M., Andreasson, E., Nielsen, K. L., Blennow, A., & Petersen, B. L. (2022). Strategies for Efficient Gene Editing in Protoplasts of Solanum tuberosum Theme: Determining gRNA Efficiency Design by Utilizing Protoplast (Research). *Frontiers in Genome Editing*, *3*. https://www.frontiersin.org/articles/10.3389/fgeed.2021.795644

Cavatorta, J., Perez, K. W., Gray, S. M., Van Eck, J., Yeam, I., & Jahn, M. (2011). Engineering virus resistance using a modified potato gene. *Plant Biotechnology Journal*, *9*(9), 1014–1021. https://doi.org/10.1111/j.1467-7652.2011.00622.x

Charbonnel, C., Allain, E., Gallego, M. E., & White, C. I. (2011). Kinetic analysis of DNA double-strand break repair pathways in Arabidopsis. *DNA Repair*, *10*(6), 611–619. https://doi.org/10.1016/j.dnarep.2011.04.002

Coy, M. R., Abbitt, S. E., & Frank, M. J. (2022). Protoplast Isolation and Transfection in Maize. *Methods in Molecular Biology (Clifton, N.J.)*, *2464*, 91–104. https://doi.org/10.1007/978-1-0716-2164-6_7

Deltcheva, E., Chylinski, K., Sharma, C. M., Gonzales, K., Chao, Y., Pirzada, Z. A., Eckert, M. R., Vogel, J., & Charpentier, E. (2011). CRISPR RNA maturation by trans-encoded small RNA and host factor RNase III. *Nature*, *471*(7340), 602–607. https://doi.org/10.1038/nature09886

Devaux, A., Goffart, J.-P., Kromann, P., Andrade-Piedra, J., Polar, V., & Hareau, G. (2021). The Potato of the Future: Opportunities and Challenges in Sustainable Agri-food Systems. *Potato Research*, *64*(4), 681–720. https://doi.org/10.1007/s11540-021-09501-4

Devaux, A., Goffart, J.-P., Petsakos, A., Kromann, P., Gatto, M., Okello, J., Suarez, V., & Hareau, G. (2020). Global Food Security, Contributions from Sustainable Potato Agri-Food Systems. In H. Campos & O. Ortiz (Eds.), *The Potato Crop: Its Agricultural, Nutritional and Social Contribution to Humankind* (pp. 3–35). Springer International Publishing. https://doi.org/10.1007/978-3-030-28683-5_1

Duan, H., Richael, C., & Rommens, C. M. (2012). Overexpression of the wild potato eIF4E-1 variant Eva1 elicits Potato virus Y resistance in plants silenced for native eIF4E-1. *Transgenic Research*, *21*(5), 929–938. https://doi.org/10.1007/s11248-011-9576-9

Eriksson, D., Brinch-Pedersen, H., Chawade, A., Holme, I. B., Hvoslef-Eide, T. A. K., Ritala, A., Teeri, T. H., & Thorstensen, T. (2018). Scandinavian perspectives on plant gene technology: Applications, policies and progress. *Physiologia Plantarum*, *162*(2), 219–238. https://doi.org/10.1111/ppl.12661

Fonfara, I., Richter, H., Bratovič, M., Le Rhun, A., & Charpentier, E. (2016). The CRISPR-associated DNA-cleaving enzyme Cpf1 also processes precursor CRISPR RNA. *Nature*, *532*(7600), 517–521. https://doi.org/10.1038/nature17945

Fuentes, S., Jones, R. A. C., Matsuoka, H., Ohshima, K., Kreuze, J., & Gibbs, A. J. (2019). Potato virus Y; the Andean connection. *Virus Evolution*, *5*(2), vez037. https://doi.org/10.1093/ve/vez037

Gaj, T., Gersbach, C. A., & Barbas, C. F. (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

Gerashchenkov, G. A., Rozhnova, N. A., Kuluev, B. R., Kiryanova, O. Y., Gumerova, G. R., Knyazev, A. V., Vershinina, Z. R., Mikhailova, E. V., Chemeris, D. A., Matniyazov, R. T., Baimiev, A. K., Gubaidullin, I. M., Baimiev, A. K., & Chemeris, A. V. (2020). [Design of Guide RNA for CRISPR/Cas Plant Genome Editing]. *Molekuliarnaia Biologiia*, *54*(1), 29–50. https://doi.org/10.31857/S0026898420010061

Gingras, A. C., Raught, B., & Sonenberg, N. (1999). eIF4 initiation factors: Effectors of mRNA recruitment to ribosomes and regulators of translation. *Annual Review of Biochemistry*, *68*, 913–963. https://doi.org/10.1146/annurev.biochem.68.1.913

Goffart, J.-P., Haverkort, A., Storey, M., Haase, N., Martin, M., Lebrun, P., Ryckmans, D., Florins, D., & Demeulemeester, K. (2022). Potato Production in Northwestern Europe (Germany, France, the Netherlands, United Kingdom, Belgium):

Characteristics, Issues, Challenges and Opportunities. *Potato Research*, *65*(3), 503–547. https://doi.org/10.1007/s11540-021-09535-8

Goodstein, D. M., Shu, S., Howson, R., Neupane, R., Hayes, R. D., Fazo, J., Mitros, T., Dirks, W., Hellsten, U., Putnam, N., & Rokhsar, D. S. (2012). Phytozome: A comparative platform for green plant genomics. *Nucleic Acids Research*, *40*(D1), D1178–D1186. https://doi.org/10.1093/nar/gkr944

Gorbunova, V., & Levy, A. A. (1997). Non-homologous DNA end joining in plant cells is associated with deletions and filler DNA insertions. *Nucleic Acids Research*, *25*(22), 4650–4657. https://doi.org/10.1093/nar/25.22.4650

Gostimskaya, I. (2022). CRISPR–Cas9: A History of Its Discovery and Ethical Considerations of Its Use in Genome Editing. *Biochemistry. Biokhimiia*, *87*(8), 777–788. https://doi.org/10.1134/S0006297922080090

Govindaraj, M., Vetriventhan, M., & Srinivasan, M. (2015). Importance of genetic diversity assessment in crop plants and its recent advances: An overview of its analytical perspectives. *Genetics Research International*, *2015*, 431487. https://doi.org/10.1155/2015/431487

Grech-Baran, M., Witek, K., Szajko, K., Witek, A. I., Morgiewicz, K., Wasilewicz-Flis, I., Jakuczun, H., Marczewski, W., Jones, J. D., & Hennig, J. (2018). *Extreme resistance to Potato Virus Y in potato carrying the Rysto gene is mediated by a TIR-NLR immune receptor* (p. 445031). bioRxiv. https://doi.org/10.1101/445031

Grzela, R., Strokovska, L., Andrieu, J.-P., Dublet, B., Zagorski, W., & Chroboczek, J. (2006). Potyvirus terminal protein VPg, effector of host eukaryotic initiation factor eIF4E. *Biochimie*, *88*(7), 887–896. https://doi.org/10.1016/j.biochi.2006.02.012

Gutierrez Sanchez, P. A., Babujee, L., Jaramillo Mesa, H., Arcibal, E., Gannon, M., Halterman, D., Jahn, M., Jiang, J., & Rakotondrafara, A. M. (2020). Overexpression

of a modified eIF4E regulates potato virus Y resistance at the transcriptional level in potato. *BMC Genomics*, *21*(1), 18. https://doi.org/10.1186/s12864-019-6423-5

Hahn, F., Korolev, A., Sanjurjo Loures, L., & Nekrasov, V. (2020). A modular cloning toolkit for genome editing in plants. *BMC Plant Biology*, *20*(1), 179. https://doi.org/10.1186/s12870-020-02388-2

Hashimoto, M., Neriya, Y., Yamaji, Y., & Namba, S. (2016). Recessive Resistance to Plant Viruses: Potential Resistance Genes Beyond Translation Initiation Factors. *Frontiers in Microbiology*, *7*. https://doi.org/10.3389/fmicb.2016.01695

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

Heigwer, F., & Boutros, M. (2021). Cloud-Based Design of Short Guide RNA (sgRNA) Libraries for CRISPR Experiments. *Methods in Molecular Biology (Clifton, N.J.)*, *2162*, 3–22. https://doi.org/10.1007/978-1-0716-0687-2_1

Hille, F., Richter, H., Wong, S. P., Bratovič, M., Ressel, S., & Charpentier, E. (2018). The Biology of CRISPR-Cas: Backward and Forward. *Cell*, *172*(6), 1239–1259. https://doi.org/10.1016/j.cell.2017.11.032

Hu, X., He, C., Xiao, Y., Xiong, X., & Nie, X. (2009). Molecular characterization and detection of recombinant isolates of potato virus Y from China. *Archives of Virology*, *154*(8), 1303–1312. https://doi.org/10.1007/s00705-009-0448-z

Hu, X., Nie, X., He, C., & Xiong, X. (2011). Differential pathogenicity of two different recombinant PVYNTN isolates in Physalis floridana is likely determined by the coat protein gene. *Virology Journal*, *8*, 207. https://doi.org/10.1186/1743-422X-8-207

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.

Jansen, Ruud., Embden, Jan. D. A. van, Gaastra, Wim., & Schouls, Leo. M. (2002). Identification of genes that are associated with DNA repeats in prokaryotes. *Molecular Microbiology*, *43*(6), 1565–1575. https://doi.org/10.1046/j.1365-2958.2002.02839.x

Jansing, J., Schiermeyer, A., Schillberg, S., Fischer, R., & Bortesi, L. (2019). Genome Editing in Agriculture: Technical and Practical Considerations. *International Journal of Molecular Sciences*, *20*(12), 2888. https://doi.org/10.3390/ijms20122888

Jinek, M., Chylinski, K., Fonfara, I., Hauer, M., Doudna, J. A., & Charpentier, E. (2012). A programmable dual-RNA-guided DNA endonuclease in adaptive bacterial immunity. *Science (New York, N.Y.)*, *337*(6096), 816–821. https://doi.org/10.1126/science.1225829

Kantor, A., McClements, M. E., & MacLaren, R. E. (2020). CRISPR-Cas9 DNA Base-Editing and Prime-Editing. *International Journal of Molecular Sciences*, *21*(17), Article 17. https://doi.org/10.3390/ijms21176240

Karasev, A. V., & Gray, S. M. (2013). Continuous and emerging challenges of Potato virus Y in potato. *Annual Review of Phytopathology*, *51*, 571–586. https://doi.org/10.1146/annurev-phyto-082712-102332

Karmakar, S., Behera, D., Baig, M. J., & Molla, K. A. (2021). In Vitro Cas9 Cleavage Assay to Check Guide RNA Efficiency. In M. T. Islam & K. A. Molla (Eds.), *CRISPR-Cas Methods: Volume 2* (pp. 23–39). Springer US. https://doi.org/10.1007/978-1-0716-1657-4_3

Knott, G. J., & Doudna, J. A. (2018). CRISPR-Cas guides the future of genetic engineering. *Science*, *361*(6405), 866–869. https://doi.org/10.1126/science.aat5011

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

Kreuze, J. F., Souza-Dias, J. A. C., Jeevalatha, A., Figueira, A. R., Valkonen, J. P. T., & Jones, R. A. C. (2020). Viral Diseases in Potato. In H. Campos & O. Ortiz (Eds.), *The Potato Crop: Its Agricultural, Nutritional and Social Contribution to Humankind* (pp. 389–430). Springer International Publishing. https://doi.org/10.1007/978-3-030-28683-5_11

Kruger, K., & Waals, J. E. (2020). Potato virus Y and Potato leafroll virus management under climate change in sub-Saharan Africa. *South African Journal of Science*. https://doi.org/10.17159/sajs.2020/8579

Kumar, A., & Chordia, N. (2015). In silico PCR primer designing and validation. *Methods in Molecular Biology (Clifton, N.J.)*, *1275*, 143–151. https://doi.org/10.1007/978-1-4939-2365-6_10

Lacomme, C., & Jacquot, E. (2017). General Characteristics of Potato virus Y (PVY) and Its Impact on Potato Production: An Overview. In C. Lacomme, L. Glais, D. U. Bellstedt, B. Dupuis, A. V. Karasev, & E. Jacquot (Eds.), *Potato virus Y: biodiversity, pathogenicity, epidemiology and management* (pp. 1–19). Springer International Publishing. https://doi.org/10.1007/978-3-319-58860-5_1

Lellis, A. D., Kasschau, K. D., Whitham, S. A., & Carrington, J. C. (2002). Loss-of-susceptibility mutants of Arabidopsis thaliana reveal an essential role for eIF(iso)4E during potyvirus infection. *Current Biology: CB*, *12*(12), 1046–1051. https://doi.org/10.1016/s0960-9822(02)00898-9

Lloyd, A. H., Wang, D., & Timmis, J. N. (2012). Single Molecule PCR Reveals Similar Patterns of Non-Homologous DSB Repair in Tobacco and Arabidopsis. *PLOS ONE*, *7*(2), e32255. https://doi.org/10.1371/journal.pone.0032255

Lucioli, A., Tavazza, R., Baima, S., Fatyol, K., Burgyan, J., & Tavazza, M. (2022). CRISPR-Cas9 Targeting of the eIF4E1 Gene Extends the Potato Virus Y Resistance Spectrum of the Solanum tuberosum L. cv. Desirée. *Frontiers in Microbiology*, *13*. https://www.frontiersin.org/articles/10.3389/fmicb.2022.873930

Marraffini, L. A., & Sontheimer, E. J. (2008). CRISPR interference limits horizontal gene transfer in staphylococci by targeting DNA. *Science (New York, N.Y.)*, *322*(5909), 1843–1845. https://doi.org/10.1126/science.1165771

Mojica, F. J. M., Díez-Villaseñor, C., García-Martínez, J., & Almendros, C. (2009). Short motif sequences determine the targets of the prokaryotic CRISPR defence system. *Microbiology*, *155*(3), 733–740. https://doi.org/10.1099/mic.0.023960-0

Mojica, F. J. M., Díez-Villaseñor, C., García-Martínez, J., & Soria, E. (2005). Intervening Sequences of Regularly Spaced Prokaryotic Repeats Derive from Foreign Genetic Elements. *Journal of Molecular Evolution*, *60*(2), 174–182. https://doi.org/10.1007/s00239-004-0046-3

Muthoni, J., Kabira, J., Shimelis, H., & Melis, R. (2015). Tetrasomic inheritance in cultivated potato and implications in conventional breeding. *Australian Journal of Crop Science*. https://search.informit.org/doi/abs/10.3316/informit.075206802413826

Nadakuduti, S. S., Buell, C. R., Voytas, D. F., Starker, C. G., & Douches, D. S. (2018). Genome Editing for Crop Improvement – Applications in Clonally Propagated Polyploids With a Focus on Potato (Solanum tuberosum L.). *Frontiers in Plant Science*, *9*. https://www.frontiersin.org/articles/10.3389/fpls.2018.01607

Nicolia, A., Fält, A.-S., Hofvander, P., & Andersson, M. (2021). Protoplast-Based Method for Genome Editing in Tetraploid Potato. In P. Tripodi (Ed.), *Crop Breeding: Genetic Improvement Methods* (pp. 177–186). Springer US. https://doi.org/10.1007/978-1-0716-1201-9_12

Nie, B., Singh, M., Murphy, A., Sullivan, A., Xie, C., & Nie, X. (2012). Response of Potato Cultivars to Five Isolates Belonging to Four Strains of Potato virus Y. *Plant Disease*, *96*, 1422–1429. https://doi.org/10.1094/PDIS-01-12-0018-RE

Patil, G. B., Stupar, R. M., & Zhang, F. (2022). Protoplast Isolation, Transfection, and Gene Editing for Soybean (Glycine max). *Methods in Molecular Biology (Clifton, N.J.)*, *2464*, 173–186. https://doi.org/10.1007/978-1-0716-2164-6_13

Patrick, R. M., & Browning, K. S. (2012). The eIF4F and eIFiso4F Complexes of Plants: An Evolutionary Perspective. *Comparative and Functional Genomics*, *2012*, 287814. https://doi.org/10.1155/2012/287814

Popkov, V. A., Zorova, L. D., Korvigo, I. O., Silachev, D. N., Jankauskas, S. S., Babenko, V. A., Pevzner, I. B., Danilina, T. I., Zorov, S. D., Plotnikov, E. Y., & Zorov, D. B. (2016). Do Mitochondria Have an Immune System? *Biochemistry. Biokhimiia*, *81*(10), 1229–1236. https://doi.org/10.1134/S0006297916100217

Pourcel, C., Salvignol, G., & Vergnaud, G. (2005). CRISPR elements in Yersinia pestis acquire new repeats by preferential uptake of bacteriophage DNA, and provide additional tools for evolutionary studies. *Microbiology*, *151*(3), 653–663. https://doi.org/10.1099/mic.0.27437-0

Reed, K. M., & Bargmann, B. O. R. (2021). Protoplast Regeneration and Its Use in New Plant Breeding Technologies. *Frontiers in Genome Editing*, *3*. https://www.frontiersin.org/articles/10.3389/fgeed.2021.734951

Robaglia, C., & Caranta, C. (2006). Translation initiation factors: A weak link in plant RNA virus infection. *Trends in Plant Science*, *11*(1), 40–45. https://doi.org/10.1016/j.tplants.2005.11.004

Ruffel, S., Dussault, M.-H., Palloix, A., Moury, B., Bendahmane, A., Robaglia, C., & Caranta, C. (2002). A natural recessive resistance gene against potato virus Y in pepper corresponds to the eukaryotic initiation factor 4E (eIF4E). *The Plant Journal: For Cell and Molecular Biology*, *32*(6), 1067–1075. https://doi.org/10.1046/j.1365-313x.2002.01499.x

Ruffel, S., Gallois, J. L., Lesage, M. L., & Caranta, C. (2005). The recessive potyvirus resistance gene pot-1 is the tomato orthologue of the pepper pvr2-eIF4E gene. *Molecular Genetics and Genomics: MGG*, *274*(4), 346–353. https://doi.org/10.1007/s00438-005-0003-x

Sagarbarria, M. G. S., & Caraan, J. A. M. (2023). *Usefulness of Current sgRNA Design Guidelines and in vitro Cleavage Assays for Plant CRISPR/Cas Genome Editing: A Case in Eggplant (Solanum melongena L.)* (p. 2023.03.19.532877). bioRxiv. https://doi.org/10.1101/2023.03.19.532877

Sahi, G., Hedley, P. E., Morris, J., Loake, G. J., & MacFarlane, S. A. (2016). Molecular and Biochemical Examination of Spraing Disease in Potato Tuber in Response to Tobacco rattle virus Infection. *Molecular Plant-Microbe Interactions®*, *29*(10), 822–828. https://doi.org/10.1094/MPMI-08-16-0169-R

SantaLucia, J. (1998). A unified view of polymer, dumbbell, and oligonucleotide DNA nearest-neighbor thermodynamics. *Proceedings of the National Academy of Sciences*, *95*(4), 1460–1465. https://doi.org/10.1073/pnas.95.4.1460

Sayers, E. W., Beck, J., Bolton, E. E., Bourexis, D., Brister, J. R., Canese, K., Comeau, D. C., Funk, K., Kim, S., Klimke, W., Marchler-Bauer, A., Landrum, M., Lathrop, S.,

Lu, Z., Madden, T. L., O'Leary, N., Phan, L., Rangwala, S. H., Schneider, V. A., … Sherry, S. T. (2021). Database resources of the National Center for Biotechnology Information. *Nucleic Acids Research*, *49*(D1), D10–D17. https://doi.org/10.1093/nar/gkaa892

Schaart, J. G., van de Wiel, C. C. M., & Smulders, M. J. M. (2021). Genome editing of polyploid crops: Prospects, achievements and bottlenecks. *Transgenic Research*, *30*(4), 337–351. https://doi.org/10.1007/s11248-021-00251-0

Shikazono, N., Suzuki, C., Kitamura, S., Watanabe, H., Tano, S., & Tanaka, A. (2005). Analysis of mutations induced by carbon ions in Arabidopsis thaliana*. *Journal of Experimental Botany*, *56*(412), 587–596. https://doi.org/10.1093/jxb/eri047

Shmakov, S., Abudayyeh, O. O., Makarova, K. S., Wolf, Y. I., Gootenberg, J. S., Semenova, E., Minakhin, L., Joung, J., Konermann, S., Severinov, K., Zhang, F., & Koonin, E. V. (2015). Discovery and Functional Characterization of Diverse Class 2 CRISPR-Cas Systems. *Molecular Cell*, *60*(3), 385–397. https://doi.org/10.1016/j.molcel.2015.10.008

Shmakov, S., Smargon, A., Scott, D., Cox, D., Pyzocha, N., Yan, W., Abudayyeh, O. O., Gootenberg, J. S., Makarova, K. S., Wolf, Y. I., Severinov, K., Zhang, F., & Koonin, E. V. (2017). Diversity and evolution of class 2 CRISPR–Cas systems. *Nature Reviews Microbiology*, *15*(3), Article 3. https://doi.org/10.1038/nrmicro.2016.184

Svitashev, S., Schwartz, C., Lenderts, B., Young, J. K., & Mark Cigan, A. (2016). Genome editing in maize directed by CRISPR–Cas9 ribonucleoprotein complexes. *Nature Communications*, *7*(1), Article 1. https://doi.org/10.1038/ncomms13274

Thatcher, A., Waterson, P., Todd, A., & Moray, N. (2017). State of Science: Ergonomics and global issues. *Ergonomics*, *61*, 1–33. https://doi.org/10.1080/00140139.2017.1398845

Torrance, L., Cowan, G. H., McLean, K., MacFarlane, S., Al-Abedy, A. N., Armstrong, M., Lim, T.-Y., Hein, I., & Bryan, G. J. (2020). Natural resistance to Potato virus Y in Solanum tuberosum Group Phureja. *Theoretical and Applied Genetics*, *133*(3), 967–980. https://doi.org/10.1007/s00122-019-03521-y

Torrance, L., Liu, H., Cowan, G., Bradshaw, J., & MacFarlane, S. (2009). Extreme resistance to potyviruses in Solanum tuberosum group phureja. *Aspects of Applied Biology, No.94*, 1–4.

Udage, A. (2021). Introduction to Plant Mutation Breeding: Different Approaches and Mutagenic Agents. *Journal of Agricultural Sciences – Sri Lanka*, *16*, 466. https://doi.org/10.4038/jas.v16i03.9472

Untergasser, A., Cutcutache, I., Koressaar, T., Ye, J., Faircloth, B. C., Remm, M., & Rozen, S. G. (2012). Primer3—New capabilities and interfaces. *Nucleic Acids Research*, *40*(15), e115. https://doi.org/10.1093/nar/gks596

Valkonen, J. P. T. (2015). Elucidation of virus-host interactions to enhance resistance breeding for control of virus diseases in potato. *Breeding Science*, *65*(1), 69–76. https://doi.org/10.1270/jsbbs.65.69

Wang, A., & Krishnaswamy, S. (2012). Eukaryotic translation initiation factor 4E-mediated recessive resistance to plant viruses and its utility in crop improvement. *Molecular Plant Pathology*, *13*(7), 795–803. https://doi.org/10.1111/j.1364-3703.2012.00791.x

Xu, X., Pan, S., Cheng, S., Zhang, B., Mu, D., Ni, P., Zhang, G., Yang, S., Li, R., Wang, J., Orjeda, G., Guzman, F., Torres, M., Lozano, R., Ponce, O., Martinez, D., De la Cruz, G., Chakrabarti, S. K., Patil, V. U., … Wageningen University & Research Centre. (2011). Genome sequence and analysis of the tuber crop potato. *Nature*, *475*(7355), Article 7355. https://doi.org/10.1038/nature10158

Yamano, T., Nishimasu, H., Zetsche, B., Hirano, H., Slaymaker, I. M., Li, Y., Fedorova, I., Nakane, T., Makarova, K. S., Koonin, E. V., Ishitani, R., Zhang, F., & Nureki, O. (2016). Crystal Structure of Cpf1 in Complex with Guide RNA and Target DNA. *Cell*, *165*(4), 949–962. https://doi.org/10.1016/j.cell.2016.04.003

Yang, H., Gao, P., Rajashankar, K. R., & Patel, D. J. (2016). PAM-Dependent Target DNA Recognition and Cleavage by C2c1 CRISPR-Cas Endonuclease. *Cell*, *167*(7), 1814-1828.e12. https://doi.org/10.1016/j.cell.2016.11.053

Zaheer, K., & Akhtar, M. H. (2016). Potato Production, Usage, and Nutrition—A Review. *Critical Reviews in Food Science and Nutrition*, *56*(5), 711–721. https://doi.org/10.1080/10408398.2012.724479

Zetsche, B., Gootenberg, J. S., Abudayyeh, O. O., Slaymaker, I. M., Makarova, K. S., Essletzbichler, P., Volz, S. E., Joung, J., van der Oost, J., Regev, A., Koonin, E. V., & Zhang, F. (2015). Cpf1 is a single RNA-guided endonuclease of a class 2 CRISPR-Cas system. *Cell*, *163*(3), 759–771. https://doi.org/10.1016/j.cell.2015.09.038

Zhang, D., Zhang, Z., Unver, T., & Zhang, B. (2021). 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

8. Appendix

I. List of Supplementary tables:

Table 8.1. Preparation of Master mix (MM) for PCR amplification of the gDNA or reverse transcribed cDNA.

Initial conc.	Component/ Tubes	Final conc.	Volm. (µl) (1rxn)	Volm. (µl) (3.5 rxn)
-	PCR- H_2O	-	31.5	110.25
5×	Phusion HF buffer	1×	10	35
10 mM	dNTPs	0.2 mM	1	3.5
10 µM	Forward Primer	0.5 µM	2.5	8.75
10 µM	Reverse Primer	0.5 µM	2.5	8.75
2 U/mL	Phusion HS II HF-DNA pol.	1 U/ mL	0.5	1.75
	Total volume		**48 µl**	**168 µl**

- Experimental setups: Three setup reactions (x 7 reactions using seven different primer sets) for; template Desiree DNA (2 µl), template Birkeland DNA (2 µl) and one reaction for negative control reaction (no template DNA; PCR- water, 2 µl).
- Four reactions for cDNA as later experiments were performed using cDNA Desirée, Kuras, Celandine, and Innovator cultivar.

Furthermore, as the quantity of *eIF4E* gene fragments amplified by primer pair-7 (St_eIF4E-7F, St_eIF4E-7R) was not quite sufficient as just a very low amount of amplified DNA was acquired as can be clear from (Figure 4.4), it had huge impact in cloning of the fragments into the PCR zero blunt vector. Hence, the PCR amplification for this fragment was repeated by using 4 µl (i.e., 34.2 ng/µl) of template DNA as mentioned in Table 8.2 and using gradient PCR program as per the Table 8.3.

Table 8.2. Preparation of Master mix (MM) for PCR amplification of the template gDNA from Desirée, using primer set-7 (St_eIF4E-7F, St_eIF4E-7R).

Initial conc.	Component/ Tubes	Final conc.	Volm. (µl) (1rxn)	Volm. (µl) (3.5 rxn)
-	PCR- H_2O	-	29.5	110.25
5×	Phusion HF buffer	1×	10	35
10 mM	dNTPs	0.2 mM	1	3.5
10 µM	Forward Primer- 7	0.5 µM	2.5	8.75
10 µM	Reverse Primer-7	0.5 µM	2.5	8.75
2 U/mL	Phusion HS II HF-DNA pol.	1 U/ mL	0.5	1.75
	Template DNA		4	
	Total volume		**50 µl**	**168µl**

Table 8.3. Gradient PCR program used for the PCR amplification of the genomic DNA fragment using the seventh forward and reverse primers as per (Table 8.2).

(Gradient PCR)	Stage 01	Step 01: 98°C x 30 sec	
	Stage 02	Step 01: 98°C x 20sec Step 02: 60°C/64°C/65°C x 20 sec Step 03: 72°C x 30 sec	35 cycles
	Stage 03	Step 01: 72°C x 7 mins Hold at 4°C	

This PCR amplification showed significant effect on the yield of the PCR amplicons. High amount of PCR products of expected size was observed (Supplementary Figure 8.4). This further facilitated effective cloning and transformation reactions.

Table 8.4. Preparation of ligation reaction of PCR amplified fragments into PCR zero blunt vector based on the protocol of Zero Blunt® PCR Cloning kit.

Ligation reaction		
Components	**Initial conc.**	**Volume**
PCR product		0.5-1 µl
T4 DNA ligase Buffer	5 x	2 µl
pCR zero blunt vector		1 µl
PCR grade water		to 9 µl
Express Link T4 DNA ligase	5 U/µl	1 µl
	Total	**10 µl**

Table 8.5. Preparation of Master mix (MM) for performing colony PCR to identify putative positive clones containing the plasmid with intended insert (DNA of interest). Insert specific primer set were used for each transformation reactions.

Initial conc.	**Component/ Tubes**	**Final conc.**	**Volm. (µl) (1rxn)**	**Volm. (µl) (16 X)**
-	PCR- H_2O	-	15.1	241.6
10×	Buffer B1	1×	2	32
25mM	$MgCl_2$	1.5 mM	1.2	19.2
10 mM	dNTPs	0.2 µM	0.4	6.4
10 µM	Forward Primer	0.2 µM	0.4	6.4
10 µM	Reverse Primer	0.2 µM	0.4	6.4
2 U/mL	Hot firepol- DNA pol.	1 U/ mL	0.5	8
	Total volume		**20 µl**	**320 µl**

Table 8.6. PCR program used for performing the colony PCR amplification using insert specific primer sets for PCR amplification.

(PCR program)	Stage 01	Step 01: 98°C x 10 min	30 cycles
	Stage 02	Step 01: 98°C x 10sec Step 02: 58°C x 30 sec Step 03: 72°C x 40 sec	
	Stage 03	Step 01: 72°C x 5 mins Hold at 10°C	

Table 8.7. Preparation of reaction mixture for performing Restriction digestion assay using SmaI enzyme.

Components	**Volume (µl)**
Plasmid DNA	2 (~400 ng)
rCutSmart Buffer	1
SmaI enzyme	1
H_2O	q.s. to 10
Total	**10 µl**

Table 8.8. Preparation of reaction mixture for performing Restriction digestion assay using EcoRI enzyme.

Components	**Volume (µl)**
Plasmid DNA	2 (~400 ng)
EcoRI Buffer	1
EcoRI enzyme	1
H_2O	q.s. to 10
Total	**10 µl**

II. List of Supplementary figures:

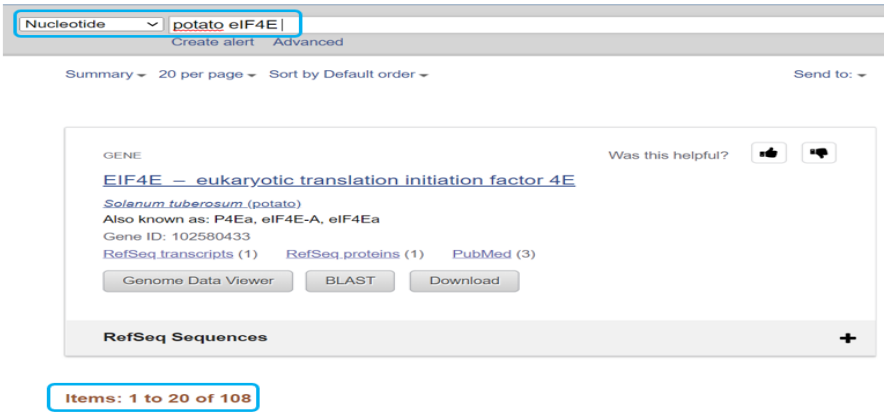


Figure 8.1. NCBI Nucleotide database search using "potato eIF4E" as keyword. This search was performed to retrieve S. tuberosum (potato) specific *eIF4E* gene or mRNA sequence information. It resulted in total 108 nucleotide hits out of which just 14 were noted to be specific to potato.

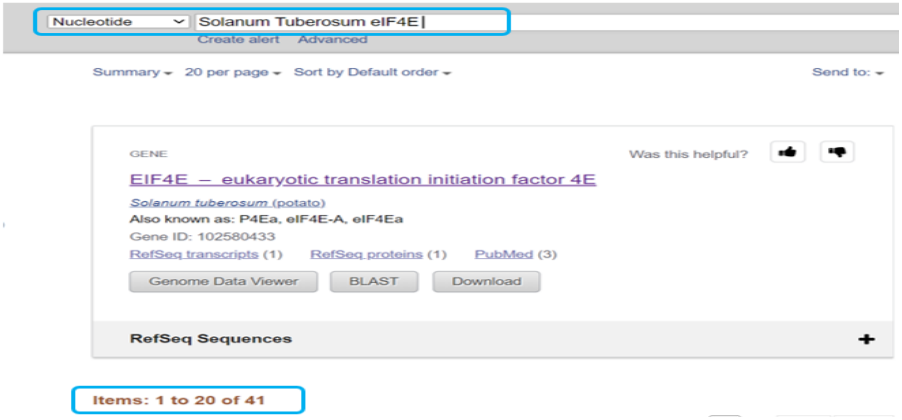


Figure 8.2. NCBI Nucleotide database search using "Solanum tuberosum eIF4E" as keyword. This search was performed to retrieve S. tuberosum specific *eIF4E* gene sequence information. It resulted in total 41 nucleotide hits.

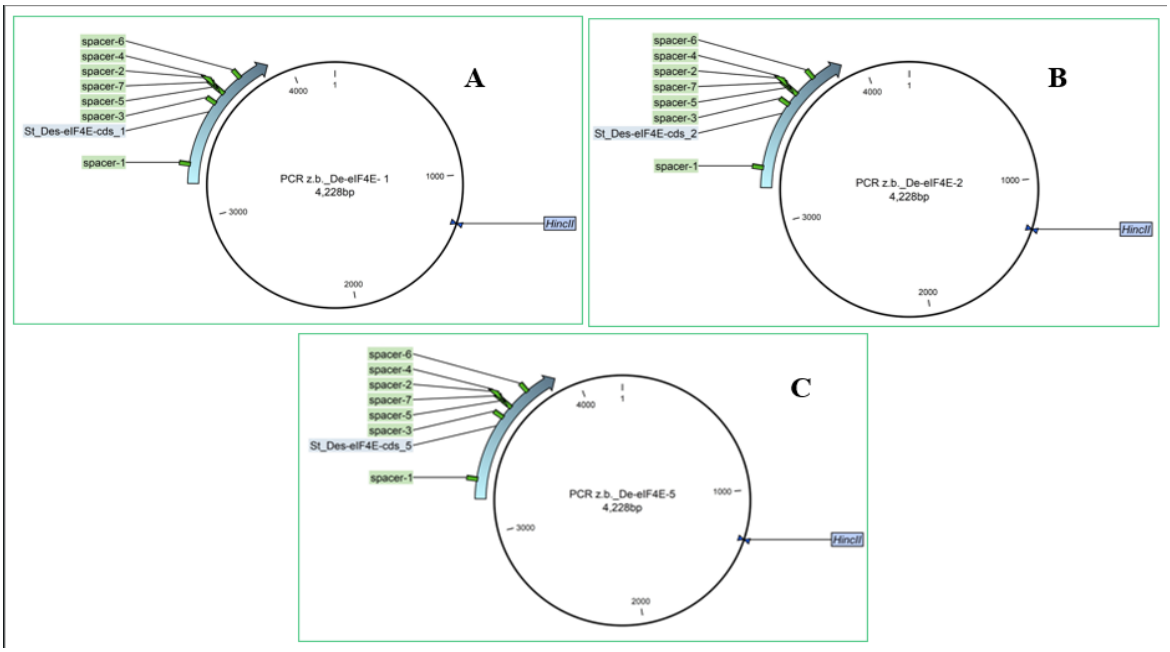


Figure 8.3. Schematic representation of in-silico maps of three different plasmids used in in-vitro cleavage assay. (A) PCR z.b._De-eIF4E-1, (B) PCR z.b._De-eIF4E-2, and (C) PCR z.b._De-eIF4E-5 all-inclusive of three different PCR amplified and Sanger sequence verified, allelic cDNA sequences (St_Des-eIF4E-cds_1, St_Des-eIF4E-cds_2, St_Des-eIF4E-cds_5) inserted into PCR® zero-blunt vector. [designed in CLC Main workbench 7.9.3] (sequence details in Supplementary documents-D, p-101)

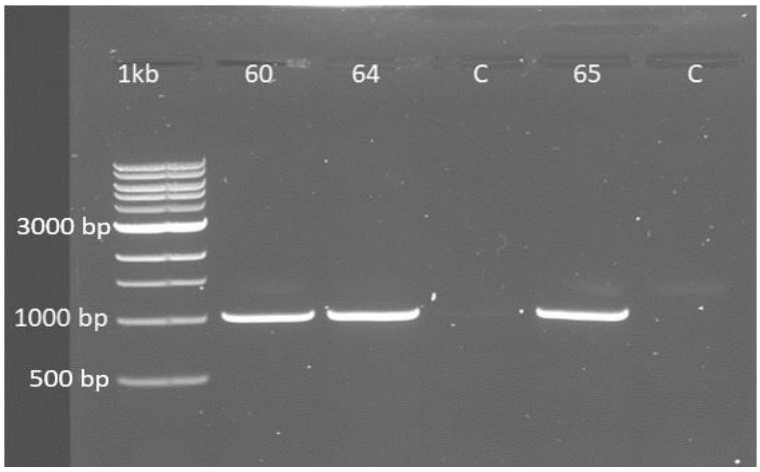


Figure 8.4. Agarose gel electrophoresis of PCR amplified *eIF4E* genomic fragments of Desirée cultivar using primer pair seven (i.e., St_eIF4E-7F, St_eIF4E-7R). To optimise the PCR product generation, gradient PCR was performed (Table 8.3) with the annealing temperatures set to be 60℃, 64℃, and 65℃. 1kb: 1kb DNA marker, C: no template control.

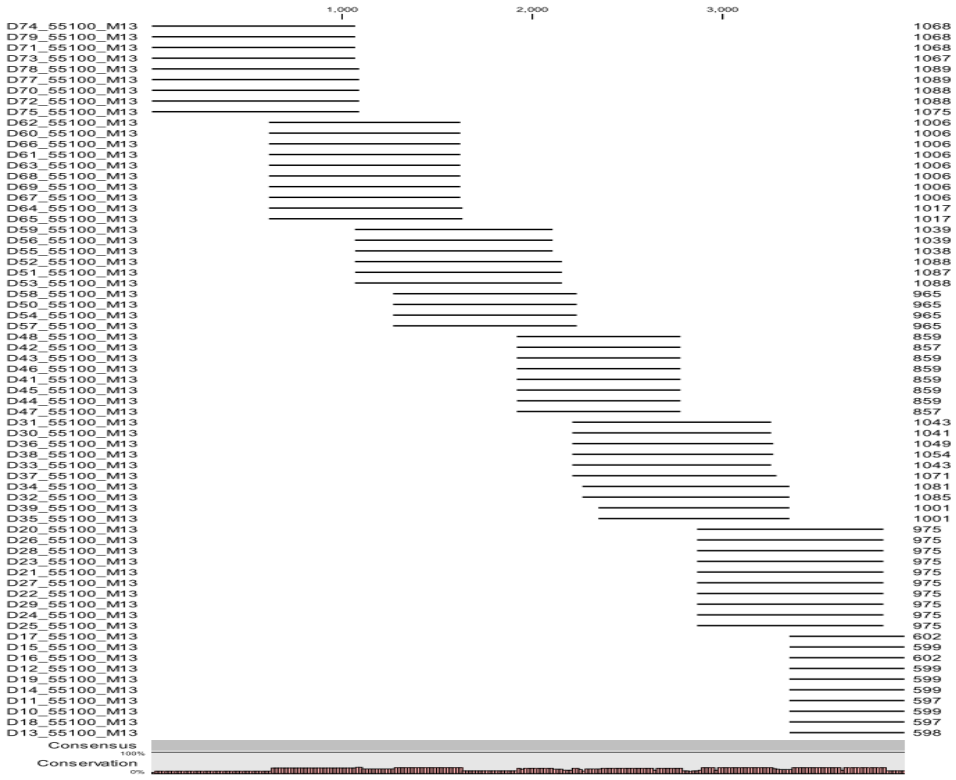


Figure 8.5. Assembling of the sequences corresponding to different gDNA amplicons from Desirée cultivar to generate consensus sequence in CLC Main Workbench 7.9.3.

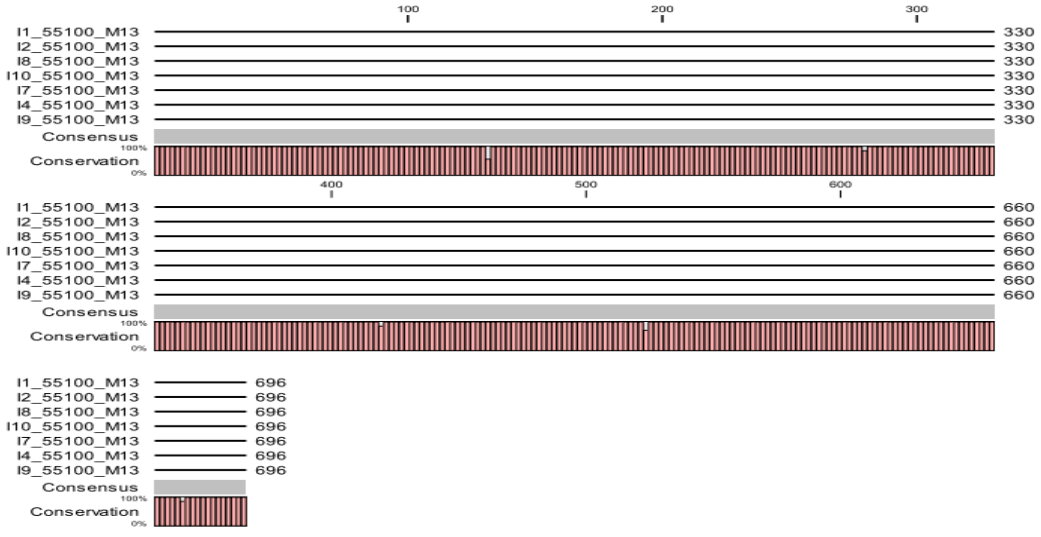


Figure 8.6. Assembling of the sequences corresponding to cDNA amplicons from Innovator cultivar to generate consensus sequence and analyse polymorphism in CLC Main Workbench 7.9.3.

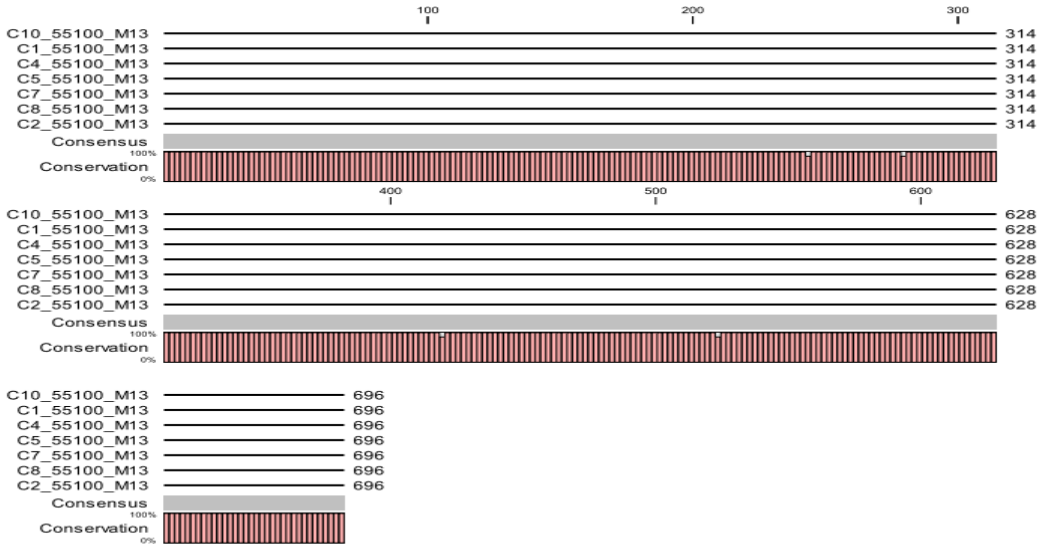


Figure 8.7. Celandine cultivar specific cDNA sequences assembly, to generate consensus sequence and analyse polymorphism.

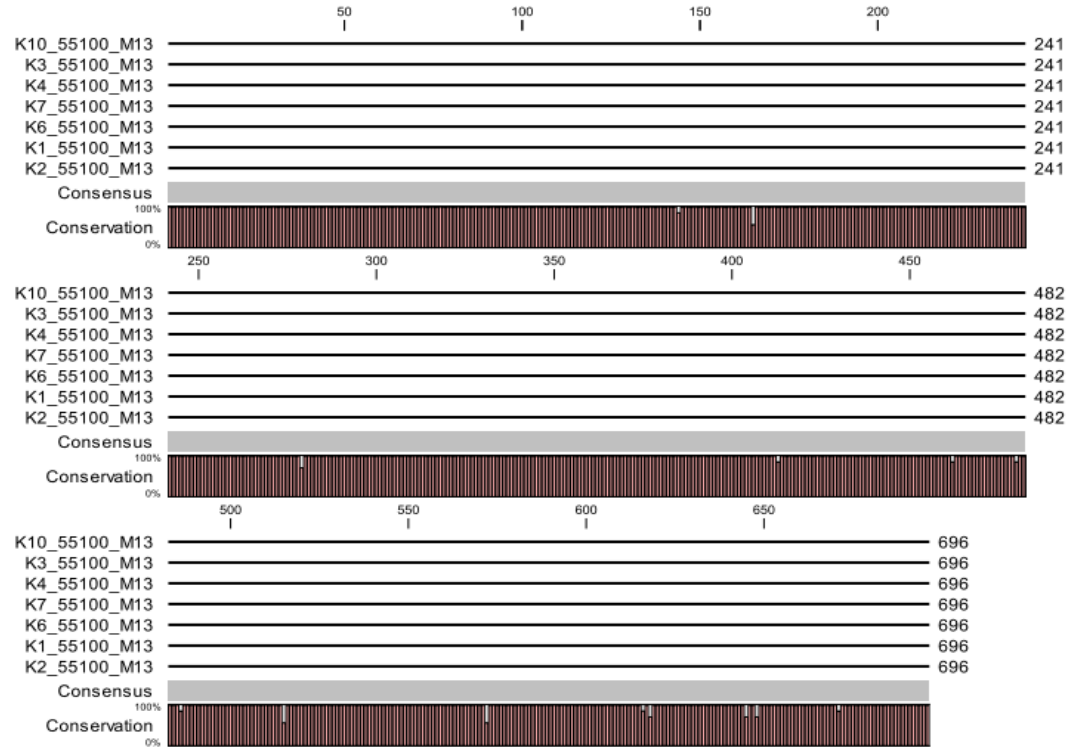
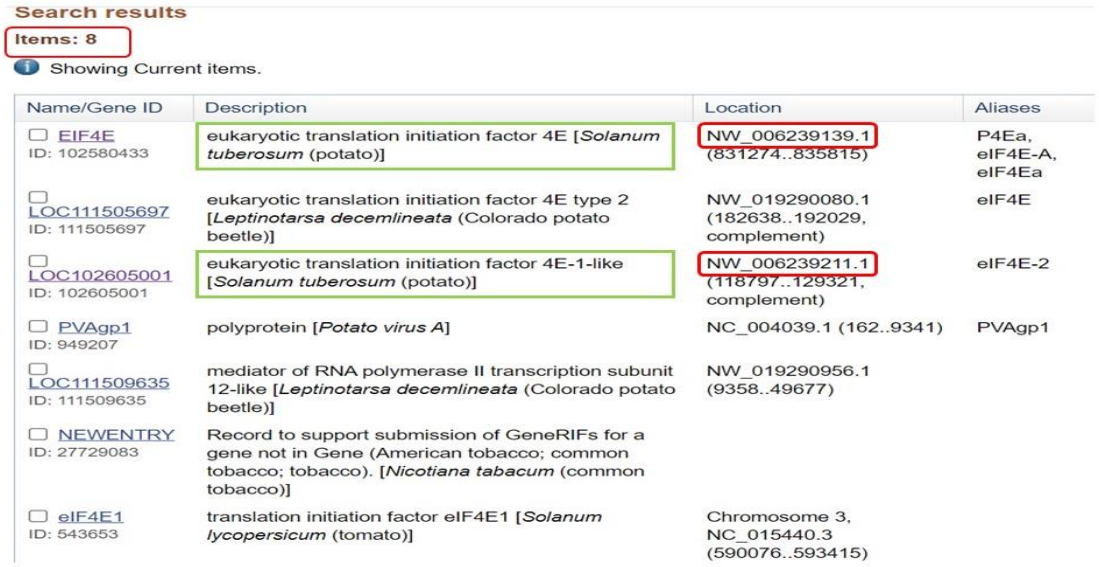


Figure 8.8. Consensus generation by assembling Kuras cultivar specific cDNA sequences.



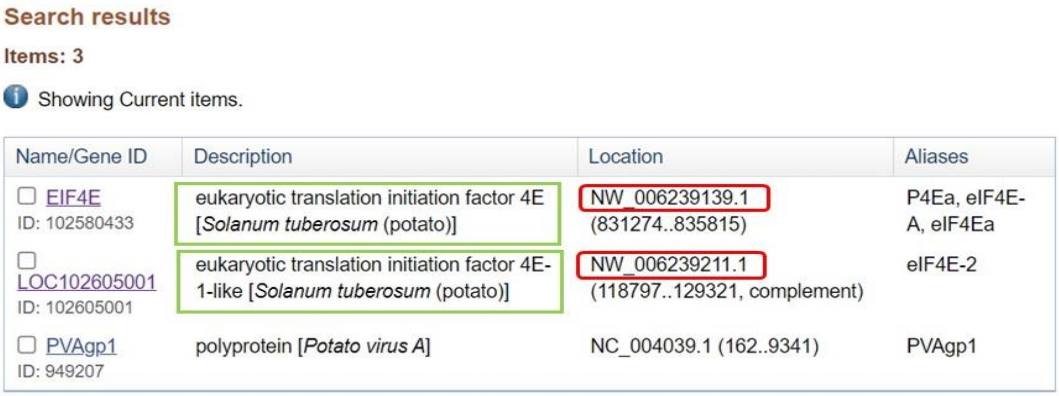
Search results

Items: 8

Showing Current items.

Name/Gene ID	Description	Location	Aliases
EIF4E ID: 102580433	eukaryotic translation initiation factor 4E [*Solanum tuberosum* (potato)]	NW_006239139.1 (831274..835815)	P4Ea, eIF4E-A, eIF4Ea
LOC111505697 ID: 111505697	eukaryotic translation initiation factor 4E type 2 [*Leptinotarsa decemlineata* (Colorado potato beetle)]	NW_019290080.1 (182638..192029, complement)	eIF4E
LOC102605001 ID: 102605001	eukaryotic translation initiation factor 4E-1-like [*Solanum tuberosum* (potato)]	NW_006239211.1 (118797..129321, complement)	eIF4E-2
PVAgp1 ID: 949207	polyprotein [*Potato virus A*]	NC_004039.1 (162..9341)	PVAgp1
LOC111509635 ID: 111509635	mediator of RNA polymerase II transcription subunit 12-like [*Leptinotarsa decemlineata* (Colorado potato beetle)]	NW_019290956.1 (9358..49677)	
NEWENTRY ID: 27729083	Record to support submission of GeneRIFs for a gene not in Gene (American tobacco; common tobacco; tobacco). [*Nicotiana tabacum* (common tobacco)]		
eIF4E1 ID: 543653	translation initiation factor eIF4E1 [*Solanum lycopersicum* (tomato)]	Chromosome 3, NC_015440.3 (590076..593415)	

Figure 8.9. The result showing hits obtained from NCBI Gene database upon search using "potato eIF4E" as keyword. This search was performed to retrieve potato specific *eIF4E* gene sequence information, which are clearly marked in light green boxes (two hits), with their respective gene accession numbers highlighted inside red boxes to their right.



Search results

Items: 3

Showing Current items.

Name/Gene ID	Description	Location	Aliases
EIF4E ID: 102580433	eukaryotic translation initiation factor 4E [*Solanum tuberosum* (potato)]	NW_006239139.1 (831274..835815)	P4Ea, eIF4E-A, eIF4Ea
LOC102605001 ID: 102605001	eukaryotic translation initiation factor 4E-1-like [*Solanum tuberosum* (potato)]	NW_006239211.1 (118797..129321, complement)	eIF4E-2
PVAgp1 ID: 949207	polyprotein [*Potato virus A*]	NC_004039.1 (162..9341)	PVAgp1

Figure 8.10. The search result representing hits obtained from NCBI Gene database search using "Solanum tuberosum eIF4E" as keyword. A total of three different hits were captured, where first two hits with accession numbers NW_006239139.1 and NW_006239211.1 were specific to the *S. tuberosum eIF4E* gene sequence.

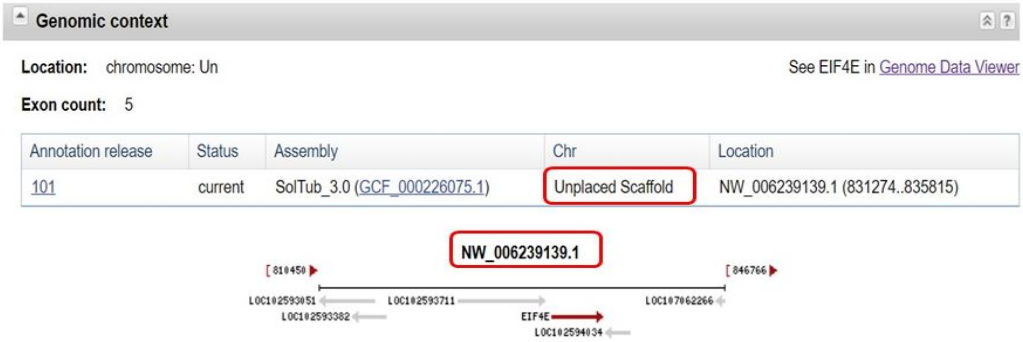


Figure 8.11. NCBI Gene database search showing information for accession number NW_006239139.1. The genomic context information obtained from the NCBI Gene database for the *eIF4E* gene sequence (accession no: NW_006239139.1) clearly mentions that the sequence information is an unplaced scaffold as in-boxed in the figure.

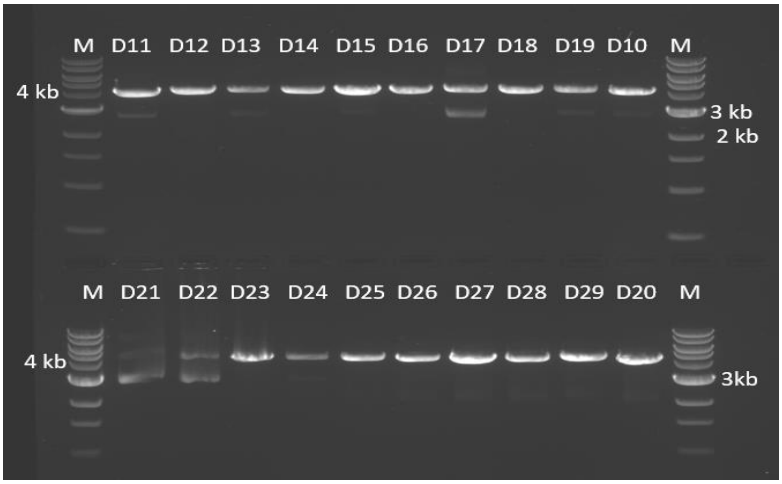


Figure 8.12. Agarose gel electrophoresis of restriction digested plasmids by using Sma I enzyme. M: 1 kb DNA marker. D represents Desirée and 1/ 2 represents the plasmid number according to insert fragment amplified by primer set 1 / 2.

Restriction digestion assays (supplementary Table 8.7 and Table 8.8) were performed to further confirm the presence of Desirée genomic fragments as insert in the plasmids before sending it for sequencing. The result showed that all the plasmid samples for D1 showed fragment of expected size, suggesting presence of intended insert in the plasmid. However, samples D21 and D22 (Figure 8.12, panel 2, lane 2 and 3) were doubtful of containing inserts amplified by primer set-2. So, the plasmids from another two positive clones were isolated and sent for sequencing together with rest of the samples.

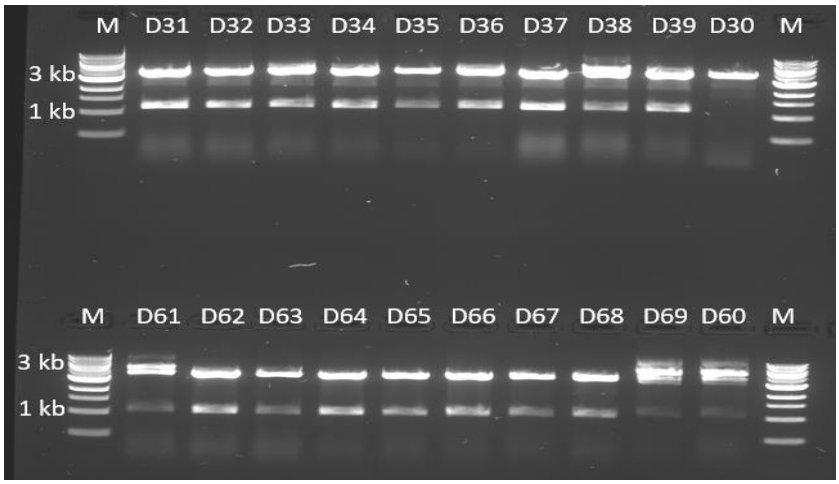


Figure 8.13. Agarose gel electrophoresis of restriction digested plasmids by using EcoRI enzyme. M: 1 kb DNA marker. D represents Desirée and 3/ 6 represents the plasmid number according to insert fragment amplified by primer set 3 / 6.

The result from this analysis showed that almost all samples showed two different fragments of expected size, thus suggesting presence of the insert of expected size. However, sample D30 (Figure 8.13, panel 1, lane 11) showed a single DNA fragment suggesting absence of insert. Thus, another colony was picked, plasmids were isolated and sent for sequencing.

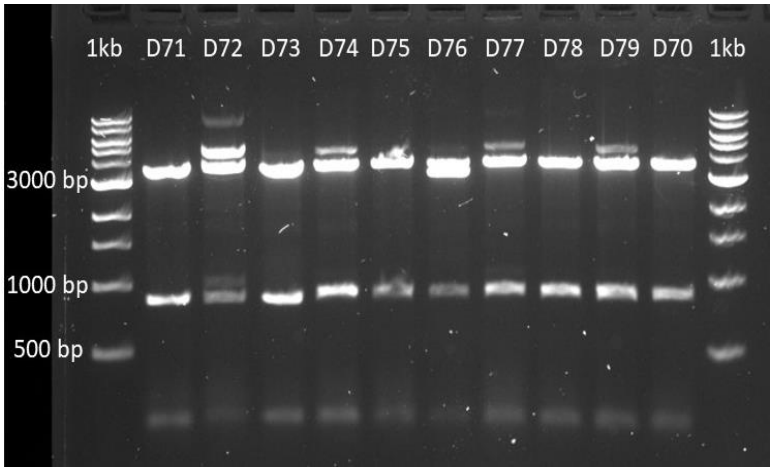


Figure 8.14. Agarose gel electrophoresis of restriction digested plasmids by using EcoRI enzyme. Agarose gel electrophoresis of restriction digested plasmids by using EcoRI enzyme. 1 kb: 1 kb DNA marker. D7: Desirée sample containing gDNA fragment 7.

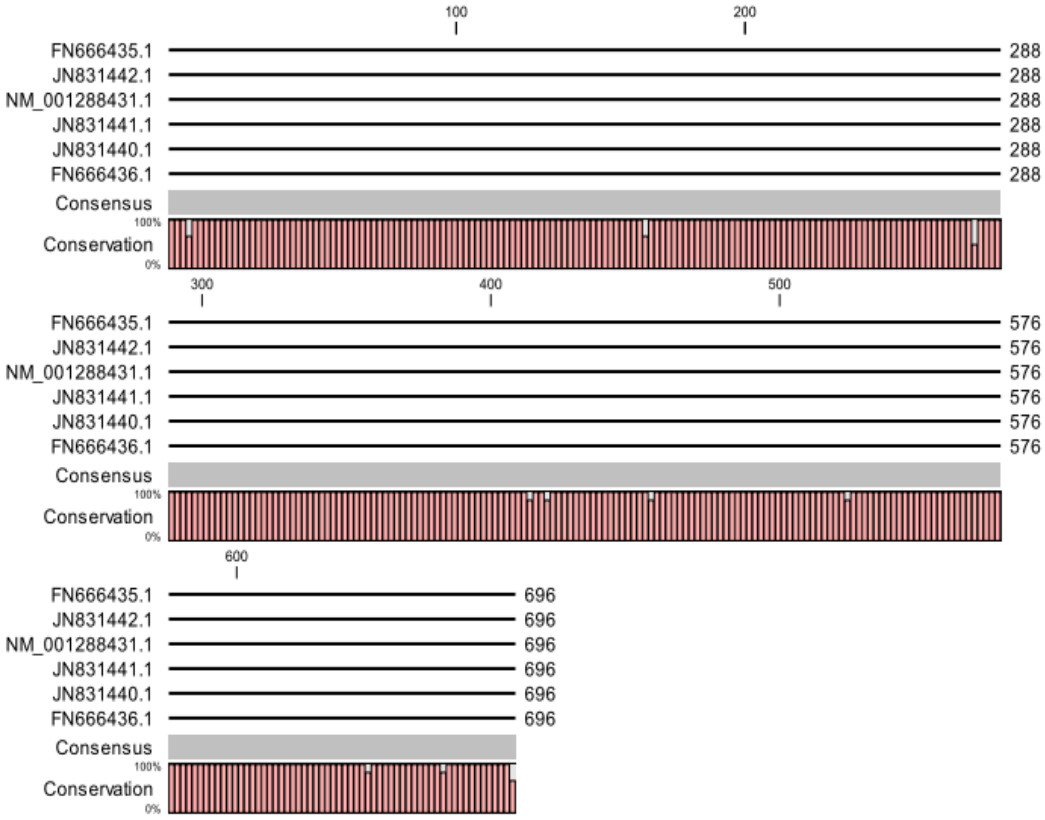


Figure 8.15. Alignment of all the mRNA CDS sequences from Table 4.1. This alignment was done to analyse the total SNPs available within these sequences and to analyse sequence similarities.

III. List of supplementary documents

A. S. tuberosum v6.1|Soltu.DM.03G000970 FASTA sequence retrieved from Phytozome database with 294 bases flanking upstream and 189 bases flanking downstream.

>S.tuberosumv6.1|Soltu.DM.03G000970|chr03:879683..883607forwardupstream=294
|downstream=189 (5 exons highlighted in light green)
ATTTGTAATTATGTTATGACAAACAAACAAATTAATATAACTGTTCATGTGACATTGTTGGTATATC
GTTAGTTTAATTGTTAATTTTTGAAAAATAAAACTTACATATTTTGAACATGCATGATTATTTGATTT
TCGAAATACGAACAGTGCTGCCTAAGTTAGGTATTTACTACATTAGTACTTCCACAATTTAAGTATT
ACACTATGGTCCAAAAGGTCCTTATAAATACCATTTTTCCTAGTAAAGTAAAAAAAAAAAAAAAAGT
ACTCCACAGTCCACAGAGCAGCAAAAATGGCAGCAGCTGAAATGGAGAGAACGACGTCGTTTGATG
CAGCTGAGAAGTTGAAGGCCGCCGATGGAGGAGGAGGGGAGGTAGACGATGAACTTGAAGAAGGT
GAAATTGTTGAAGAATCAAATGATACGGCGTCGTATTTGGGGAAAGAAATCACAGTGAAGCATCCA
TTGGAGCATTCATGGACTTTTTGGTTTGATAGCCCTATTGCTAAATCTCGACAAACTGCTTGGGGAA
GCTCACTTCGAAATGTCTACACTTTCTCCACTGTTGAAGATTTTTGGGGGTAATTTTTTTTTTCTTTCA
AATTGGTGATAGTGTAGTGTAAGAGGAAACAGGGGGAGTGGATCGCAAGGTGTGGAATCGAATCCT
CGTGAAAGTTTAGATAGTCAATTAATTGAGCTACTGAGATTCCCCGGATTTTTTAAAAAAAAGAATT
GGGGATAGAGGAAGGGGAAATGGGGGAGATAATTATAAGGTTAAACCGTCGTTAACAAGGTGAAA
GTTTAGATAGTTAATCAATTGAGCTACTAAGATGGTACGATTTTCGTAAGCTCAAATTGGAAAATAA
AAAAATTAGATTATTACTAAGGAGCTGAGAAATTCAGAAATGAGTTAGCTTTTGAGCCAGCAGAGC
AAGTTAAGTTGAGATTTCATTGATTTTAGCTTGTTTTTGTTGTTGGTTGTGTTCTTTGCTTGTGAATTA
TGGGGTTTTATTATAATGTTCCAAAAAATTTGTTGGTGTTGTAGGAATTTTTTGATTTAGGGTTTTGA
ATGTGTTCTGATTATCTGTCTTCATATAGGTGATATTGAGGTAAAAAAAGGTGTCTTCATTGGTTTTT
CTCTCATAATTGTTAGATTTTCTTCTTGTACATTTTGCATTTTTAGTTAACTTGCCAAGGTGTTTAGTT
AATGCACGACGTGCAGTCTTACTATGGAATGCTGAATTGGTTTGAGCTTTAGTTTATGCAAGCTGTA
TTAGTAATTGCACATTGCTTTGTCCTTTTTCAGTTAGCACTTTCTAAAAAAATGGAATTATCTCTCGG
TTACTTGCCAATTTTAAGTTTAGTTCATGGATCAGATTCAGTGAATAGATATATGCTCTTGGTGGTAG
GAGGACAAGGATAGATAAAGTAATCTCGGTGTAAAGGAGACGAAGTGCAGAATGAATGGAAGTTG
AGACATTTTGGAGGATGGGGTAAAGAATGACCGGTGAAGTTCTAGGAAAGAAAAAGTTAACCTTGT
TGAAAAGTCATATTTCTCTTCCTTGCACAATCTTGTGTAGGAATGTCTTTAATTGGGTTATGTCGTCG
GATAGTTTATATTGCTTCAAATTTAAGAACTTCTTTCATATTTGTCAAATGCTCATATTTACAACTAA
GGATATTTTGAGCTAATAAGGTAGTTGATAATGTGTTGGTCTTAGTTTTCTCTCCAAAGCAAGATAA
TATAGTTACTATTTTCAAGTACAACTACTTTTTCTTTAATTTACATATAAGAGCAGGTCCTGTACTTC
ATAGCTGTGATACCTCTATGTGCATTTGTCTTAATGTACCAAGCTGACTATTCACTTTTAAAAATAAA
GAGAGATGTGAAATTATTTCTTAATTAGGCTAAAATATGCTCATCTTCCCTGTAGATTGATGAGACT
AAACTTGGCTGCAGGCCAGTATTTTCGATCTCACGCATAGACCTTTTTATTAAGTCCTATGATTTGAG
TAACTAGGTGGTAATGGTGTGAATTTGATGTCTGTCATGGATGCGTGTACCTTGTTTGGTGAGCTCTT
TATTGGTCTAATTACTCAAGGCACATAAGTTATCGTACAACTTGGACTATGACATGCCTGTTTGATA
TTCCACATCATGGATTAGGTCTTTTAAATGCTATTATCCTTTTGGCTCATGATGAAATCTTGAACCAT
GTCGCTTATTCTGCAAACAGTGCTTACAATAATATCCATCACCCAAGCAAGTTGGTTATGGGAGCTG
ACTTTCATTGTTTTAAGCATAAAATTGAGCCAAAGTGGGAAGATCCTGTATGTGCCAATGGAGGGAC
GTGGAAAATGAGTTTTTTGAAGGGTAAATCTGATACCAGCTGGCTATATACGGTATGCCGAAGATAT
TTCCATCCAGCTCTTAATGATAGGTCACTCTAGTAATGTTATTTTCCCCTTTGATATAATTTCACCTCT
TGTTTTCTTATATGGGATTACTGTAGCTGCTGGCTATGATTGGACATCAATTTGATCATGGAGATGA
AATTTGTGGAGCAGTCGTTAGTGTCCGGTCTAAGGGAGAAAAAATAGCTTTGTGGACCAAGAATGC
TGCAAATGAAACAGCTCAGGTAATTTGCTTTTTATTTTTGGTTTCAACGGCCACGTTTGTCATTTTCG
TGGATCAAGTGGACAGACATTTTTCGTTGTGTACATACTGTAGTGCTGATGTTTATTCAAGATATGCT
ATGAAATACATTGTACCATTGAGGATATTGGAAGTTAAGAAGGCAAGTCATCAGATATATAGCATG
GTTTCACTATTTTTATCTACATTCTAGGTAAGAGGATATCTCAAACATTGATACACTGTGTATTTGTC

TAACTTTATAGCTTTTGGTAGTGAATTACTAACATAAGAAAACAATACAGCTGGAAGCTAAAAAGA
AATATTCTTGTGGATGAAGGGAAGTGTTAATTTGTCTGCAATAATATGTTAATTTTTTTGATTGCTAT
ATAATATGCTAACTTAATTGTTGGGGTGGTGTTAATATGTATGGAGCACTTTATGATAAACTCATCT
GTATGTTTAATAGAAAGGCAAAGCGTTGAAGTAGGGAGGTGTGCCATACATCAAAGAAGGCTTATT
AACCGACATTAGTAATAGACATAGGTTATTGTTTTCAGCTCCTCCCAACACGTCTCAGACGTTTTGT
AGAACTGGCATTGGTATCTTTAGTACCCAATCATATAATTGCATAATCTAGACTATGATTATTGCTTG
GTTGGCAACTCAGGTATGTGAAGCTTAAATAAAGTGAACTTACATATGAAAAGGATACTATTTTGAT
GTTATTTTCTGGGGAACTGTGGATATTGTTGTCAAATTGCTTGCTACATTCCTGCTTGATATGTCATT
ACTTAGTATTTATTGTTGGGAGATTCTGGTATCATTCCAAGATCCATAGACGATTTGAGAAGTCTTA
AAATTCTTTTATTCCAAGGTTAATTCATTTGGTGTAGTCATCCGGTATCCAATAACCCCCCCCCCCTC
CCCCCCCCGACTAATCCAGATTCATGCCAGTCCATTAAAATAGGAATTGCTCGGTTATGAGAAATTT
CTCCATTAGGGCTCAAACCCGAGAACCATGGCTAAGGGCGGAGGGATCCCCTTGGTGGTCTCAAGG
GTTAAGTATTATTATCATTTCATGTTCATATGATTGTAACATTTAGATATCCTTGCGAAGTGTAATAT
TGTTTTGTCTCGTCATTTGTTTTATGCTTCATATAAATTCCCTTTTCTTTCGCCCTGTCTCTTCGTTTC
TTTTTATCCAGTGTACTTCATTCTGCTAGAATTATATAGTCGTCGGTGTTCCTTTCTGTGCTTCATTGA
TGATATTGAATCCTTTCAGCTGTTTACCGGTTGATTACTTTTATGTTAGTGCTTAGCTTTTTGTTTTTA
TATCTTTGCAACTGTACTTACCATCTTCCTGAAACTTCTCTTTACAGGTTAGCATTGGTAAGCAATGG
AAGCAGTTTCTAGATCATAGCGATTCGGTTGGCTTCATATTTCATGTATGAAATCTTGGTTATCGTAC
GCCTTGAATTCAGTTTCTCTTCAATTAGCAAGACTCACAAGGAATCATCTTCTTTTGCAGGACGATG
CAAAGAGGCTCGACAGAAATGCCAAGAATCGTTACACAGTATAGTTCTTGATGCAATGTGGGATTG
CAAGAAACACAATTCGTACCGGAAAGTTGGTAGGCACTGATTTAGTTTCTCATTTGATAAGCTTCTG
GTTTGACTAACTCGTGTATTGATGTTTGCACTTTCTAATCGCGGAAAACTGTTTGGTTTGAATTCATG
CCTCTACAATTCGCGTGTTCTGCGTTACTTCC

B. Potato *eIF4E* mRNA CDS sequences retrieved from NCBI database.

1. FASTA sequence for potato *eIF4E* mRNA cds (accession number NM_001288431.1).

ATGGCAACAGCTGAAATGGAGAGAACGACGTCGTTTGATGCAGCTGAGAAGTTGAAGGCCGCCGAT
GCAGGAGGAGGAGAGGTAGACGATGAACTTGAAGAAGGTGAAATTGTTGAAGAATCAAATGATAC
GGCGTCGTATTTGGGGAAAGAAATCACAGTGAAGCATCCATTGGAGCATTCATGGACTTTTTGGTTT
GATAGCCCTATTGCTAAATCTCGACAAACTGCTTGGGGAAGCTCACTTCGAAATGTCTACACTTTCT
CCACTGTTGAAGAGTTTTGGGGTGCTTACAATAATATCCATCACCCAAGCAAGTTGGTTATGGGAGC
AGACTTTCATTGTTTTAAGCATAAAATTGAGCCAAAGTGGGAAGATCCTGTATGTGCCAATGGAGG
GACGTGGAAAATGAGTTTTTTGAAGGGTAAATCTGATACCAGCTGGCTATATACGCTGCTGGCTATG
ATTGGACATCAATTTGATCATGGAGATGAAATTTGTGGAGCAGTCGTTAGTGTCCGGTCTAAGGGAG
AAAAAATAGCTTTGTGGACCAAGAATGCTGCAAATGAAACAGCTCAGGTTAGCATTGGTAAGCAAT
GGAAGCAGTTTCTAGATCATAGCGATTCGGTTGGCTTCATATTTCATGACGATGCAAAGAGGCTCGA
CAGAAGTGCCAAGAATCGTTACACCGTATAG

2. FASTA sequence, accession number: FN666435.1.

ATGGCAGCAGCTGAAATGGAGAGAACGACGTCGTTTGATGCAGCTGAGAAGTTGAAGGCCGCCGAT
GCAGGAGGAGGAGAGGTAGACGATGAACTTGAAGAAGGTGAAATTGTTGAAGAATCAAATGATAC
GGCGTCGTATTTGGGGAAAGAAATCACAGTGAAGCATCCATTGGAGCATTCATGGACTTTTTGGTTT
GATAGCCCTATTGCTAAATCTCGACAAACTGCTTGGGGAAGCTCACTTCGAAATGTCTACACTTTCT
CCACTGTTGAAGAGTTTTGGGGTGCTTACAATAATATCCATCACCCAAGCAAGTTGGTTATGGGAGC
AGACTTTCATTGTTTTAAGCATAAAATTGAGCCAAAGTGGGAAGATCCTGTATGTGCCAATGGAGG
GACGTGGAAAATGAGTTTTTTGAAGGGTAAATCTGATACCAGCTGGCTATATACGCTGCTGGCTATG
ATTGGACATCAATTTGATCATGGAGATGAAATTTGTGGAGCAGTCGTTAGTGTCCGGTCTAAGGGAG

AAAAAATAGCTTTGTGGACCAAGAATGCTGCAAATGAAACAGCTCAGGTTAGCATTGGTAAGCAAT
GGAAGCAGTTTCTAGATCATAGCGATTCGGTTGGCTTCATATTTCATGACGATGCAAAGAGGCTCGA
CAGAAGTGCCAAGAATCGTTACACCGTATGA

3. FASTA sequence, accession number: FN666436.1.

ATGGCAGCAGCTGAAATGGAGAGAACGACGTCGTTTGATGCAGCTGACAAGTTGAAGGCCGCCGAT
GCAGGAGGAGGAGAGGTAGACGATGAACTTGAAGAAGGTGAAATTGTTGAAGAATCAAATGATAC
GGCGTCGTATTTAGGGAAAGAAATTACAGTGAAGCATCCATTGGAGCATTCATGGACTTTTTGGTTT
GATAGCCCTATTGCTAAATCTCGACAAACTGCTTGGGGAAGCTCACTTCGAAATGTCTACACTTTCT
CCACTGTTGAAGATTTTTGGGGTGCTTACAATAATATCCATCACCCAAGCAAGTTGGTTATGGGAGC
AGACTTTCATTGTTTTAAGCATAAAATTGAGCCAAAGTGGGAAGATCCTGTATGTGCCAATGGAGG
GACGTGGAAAATGAGTTTTTCGAAGGGTAAATCTGATACTAGCTGGCTATATACGCCGCTGGCAAT
GATTGGACATCAATTCGATCACGGAGATGAAATTTGTGGAGCAGTCGTTAGTGTCCGGGCTAAGGG
AGAAAAAATAGCTTTGTGGACCAAGAATGCTGCAAATGAAACAGCTCAGGTTAGCATCGGTAAGCA
ATGGAAGCAGTTTCTAGATTACAGCGATTCGGTTGGCTTCATATTTCATGACGATGCAAAGAGGCTC
GACAGAAATGCCAAGAATCGTTACACCGTATGA

4. FASTA sequence, accession number: JN831440.1.

ATGGCAACAGCTGAAATGGAGAGAACGACGTCGTTTGATGCAGCTGAGAAGTTGAAGGCCGCCGAT
GCAGGAGGAGGAGAGGTAGACGATGAACTTGAAGAAGGTGAAATTGTTGAAGAATCAAATGATAC
GGCGTCGTATTTAGGGAAAGAAATCACAGTGAAACATCCATTGGAGCATTCATGGACTTTTTGGTTT
GATAGCCCTATTGCTAAATCTCGACAAACTGCTTGGGGAAGCTCACTTCGAAATGTCTACACTTTCT
CCACTGTTGAAGATTTTTGGGGTGCTTACAATAATATCCATCACCCAAGCAAGTTGGTTATGGGAGC
AGACTTTCATTGTTTTAAGCATAAAATTGAGCCAAAGTGGGAAGATCCTGTATGTGCCAATGGAGG
GACGTGGAAAATGAATTTTTTGAAGGGTAAATCTGATACCAGCTGGCTATATACGCTGCTGGCAATG
ATTGGACATCAATTCGATCACGGAGATGAAATTTGTGGAGCAGTCGTTAGTGTCCGGTCTAAGGGA
GAAAAAATAGCTTTGTGGACCAAGAATGCTGCAAATGAAACAGCTCAGGTTAGCATTGGTAAGCAA
TGGAAGCAGTTTCTAGATTACAGCGATTCGGTTGGCTTCATATTTCACGATGATGCAAAGAGGCTCG
ACAGAAGTGCCAAGAATCGTTACACCGTATAG

5. FASTA sequence, accession number: JN831441.1.

ATGGCAACAGCTGAAATGGAGAGAACGACGTCGTTTGATGCAGCTGAGAAGTTGAAGGCCGCCGAT
GCAGGAGGAGGAGAGGTAGACGATGAACTTGAAGAAGGTGAAATTGTTGAAGAATCAAATGATAC
GGCGTCGTATTTAGGGAAAGAAATCACAGTGAAACATCCATTGGAGCATTCATGGACTTTTTGGTTT
GATAGCCCTATTGCTAAATCTCGACAAACTGCTTGGGGAAGCTCACTTCGAAATGTCTACACTTTCT
CCACTGTTGAAGATTTTTGGGGTGCTTACAATAATATCCATCACCCAAGCAAGTTGGTTATGGGAGC
AGACTTTCATTGTTTTAAGCATAAAATTGAGCCAAAGTGGGAAGATCCTGTATGTGCCAATGGAGG
GACGTGGAAAATGAGTTTTTTGAAGGGTAAATCTGATACCAGCTGGCTATATACGCTGCTGGCTATG
ATTGGACATCAATTTGATCATGGAGATGAAATTTGTGGAGCAGTCGTTAGTGTCCGGTCTAAGGGAG
AAAAAATAGCTTTGTGGACCAAGAATGCTGCAAATGAAACAGCTCAGGTTAGCATTGGTAAGCAAT
GGAAGCAGTTTCTAGATCATAGCGATTCGGTTGGCTTCATATTTCATGACGATGCAAAGAGGCTCGA
CAGAAGTGCCAAGAATCGTTACACCGTATAG

6. FASTA sequence, accession number: JN831442.1.

ATGGCAACAGCTGAAATGGAGAGAACGACGTCGTTTGATGCAGCTGAGAAGTTGAAGGCCGCCGAT
GCAGGAGGAGGAGAGGTAGACGATGAACTTGAAGAAGGTGAAATTGTTGAAGAATCAAATGATAC
GGCGTCGTATTTGGGGAAAGAAATCACAGTGAAGCATCCATTGGAGCATTCATGGACTTTTTGGTTT
GATAGCCCTATTGCTAAATCTCGACAAACTGCTTGGGGAAGCTCACTTCGAAATGTCTACACTTTCT
CCACTGTTGAAGAGTTTTGGGGTGCTTACAATAATATCCATCACCCAAGCAAGTTGGTTATGGGAGC
AGACTTTCATTGTTTTAAGCATAAAATTGAGCCAAAGTGGGAAGATCCTGTATGTGCCAATGGAGG
GACGTGGAAAATGAGTTTTTTGAAGGGTAAATCTGATACCAGCTGGCTATATACGCTGCTGGCTATG
ATTGGACATCAATTTGATCATGGAGATGAAATTTGTGGAGCAGTCGTTAGTGTCCGGTCTAAGGGAG
AAAAAATAGCTTTGTGGACCAAGAATGCTGCAAATGAAACAGCTCAGGTTAGCATTGGTAAGCAAT
GGAAGCAGTTTCTAGATCATAGCGATTCGGTTGGCTTCATATTTCATGACGATGCAAAGAGGCTCGA
CAGAAGTGCCAAGAATCGTTACACCGTATAG

C. Sequences determined by Sanger Sequencing

7. Desirée_eIF4E- gDNA_consensus sequence-inclusively (mRNA coding genomic sequence trimmed out) with added 40 first bases according to the alignment with the reference genomic *eIF4E* sequence from Phytozome (S. tuberosum v6.1|Soltu.DM.03G000970).

ATGGCAGCAGCTGAAATGGAGAGAACGACGTCGTTTGATGCAGCTGAGAAGTTGAAGGCCGCCGAT
GSAGGAGGAGGRGAGGTAGACGATGAACTTGAAGAAGGTGAAATTGTTGAAGAATCAAATGATAY
GGCGTCGTATTTAGGGAAAGAAATCACAGTGAAGCATCCATTGGAGCATTCATGGACTTTTTGGTTT
GATAGCCCTATTGCTAAATCTCGACAAACTGCTTGGGGAAGCTCMCTTCGAAATGTCTACACTTTCT
CCACTGTTGAAGATTTTTGGGGGTAATTTTTTTTTTCTTTCAAATTGGTGMTAGTGTAGTGTAAGAGG
AAACAGGGGGAGTGGATCGCAAGGTGTGGAAWCGAATCCTCSTGAAAGTTTAGATAGTCAATTAAT
TGAGCTACTGAGATTCCCCGGATTTTTWWWAAAAAWAAATTGGGGATAGAGGAARGGGAAATGG
GGGAGAGAATTATAAGGTTAAACCGTCGTTAACAARGTGARAGTTTAGGTAGTTAATCAATTGAGC
TWCTRAGATTSTTTGGGTACGATTTTCGTAAGCTCAAATTGGAAAATAAAAAATTAGATTATTACTA
AGGAGCTGAGAAATTCAGAAATGAGTTAGCTTTTGAGCCAGCAGAGCAAGTTAAGTTGAGATTTCA
TTGATTTTAGCTTGTTTTTGTTGTTGGTTGTGTTCTTTGCTTGTGGAWTATGGGGTTTTATTATAATGT
TCCAAAAAAAWTTGTTGGTGTTGTAGGAATTTTTTGATTTAGGGTTTTGAATGTGTTCTGATTATCTGT
CTTCATATAGGTGATATTGAGGTAAAAAAAGGTATCTTCATTGGTTTTTCTCTCATAATTGTTAGATT
TTCTTCTTGTACATTTTGCATTTTTRGTTAACTTGCYAAGRTGTTTAGTTAATGCACRAYGYGCAGTC
TTACTATGGAATGCTGAATYGGTTTGAGCTTTAGTTTATGCARGCTGTATTAGTAATTGCACATTGCT
TTGTCYTTTTTCAGTTAGCWCTTKCTAAAAAAMYKGAATTTTGATCTCYCAGTTAYTTGCCAATTTT
AAGTTTAKTTCATGGATYGGATTCARTGAATAGATATATGCTCTTTGGTGKTWGGAGGACAAGGAT
AGRTAAAKTAATCTYGGTGTAAAGGAGASGAAGTGCARAAWATGAATGGAAGTTGAGACATTTTG
GAGGATGGGGTRAAGAATGACCGGTGAAGTTCTAGGAAAGAAAAAGTTAACCTTGTTGAAAAGTCA
TATTTCTCTTCCTTGCACAATCTTGTGTAGGAATGTCTTTAATTGGGTTATATCGTCGGATAGTTTAT
AYTRCTTCAAATTTAAGAACCTYTTWCATRTTTGTSAAATGCTCATATTTACAACTAAGGATATTTTG

AGCTAATAAGGTAGTTGATAATGCKTTGGTCTTAGTTTTCTCTCCWAAGCAAGATAATATAGTTACT
ATTTTCAAGTACAACTACTTCTGTTTCTTTAATTTACATATAAGAGCAGGYCCTGTACTTCATAGCTG
TGATAYCTCTATGTGCATTTGTCTTAATKTACYRAGCTGACTATTCACTTTTAMAARTAAAGACAGA
TGTGWWWYTATTTCTTAATTAGGCTWAAATATGCTCATCTYCCCTGTAGATTGATGAGACTAAACT
TRGCTGCAGGCCAGTATTTTCRATCTCACGCATAGACCTTTTTATTAAGTCCTATGATTTGAGYMACT
AGGTGGTAATGGTGTGAATTTGATGTCTGTCATGGATGYRTGTACCTTGTTTGGTGARYTCTTTATTG
GTCTAATTACTCAAGGCACTTAAGTTATYGTACAACTTGKACTATGACATGCCTGTTTGATATTCCA
CAYCATGGATTAGGTCTTTTAAATGCTATTATCCTTTTGGCTCATGATGAAATCTTGAACCATGTCRC
TTATTCTGCAARCAGTGCTTACAATAATATCCATCACCCAAGCAAGTTGGTTATGGGAGCAGACTTT
CATTGTTTTAAGCATARAATTGAGCCAAAGTGGGAAGATCCTGTATGTRCCAATGGAGGGACGTGG
AAAATGAGTTTTTCGAAGGGTAAATCTGATACCAGCTGGCTATATACGGTATGCCRAAGATATTTCC
ATCCAGCTCTTAATGATAGGTCACTCTAGTAATGTTATTTTCCCCTTTGATATAATTTCAMCTCTTGT
TTTCTTATATGGGATTACTGTAGCTGCTGGCWATGATTGGACATCAATTYGATCATGGAGATGAAAT
TTGTGGAGCAGTCGTTAGTGTCCGGGCTAAGGGAGAAAAAATAGCTTTGTGGACCAAGAATGCTGC
AAATGAAACAGCTCAGGTAATTTKCTTTTTATTTTTGGTTTCAACGGCYRCGTTTGTCATTTTCGTGG
ATCAAGTGGACAGACATTTTTCGTTGTGTACATACTGTAGTGCTGATGTTTATTCAAGATATGCTATG
AAATACATTGTACCATTGAGGATATTGGAAGTTAAGAAGGSAAGTCATCAGATATATAGCATGGTTT
CACTATTTTTATYTACATTCTAGGTAAGAGGATATCTYAAACRTTGATACACTGTGTATTTGTCTAAC
TTTATAGCTTTTGRTAGTGAATWACYWWCATRAGAAAACAATACAGCTGGAAGCTAAAAARAAAT
ATTCTTGTGGATGAAGGGAAGTGTTAATTTGTCTGCAATAATATGTTAATTTTTTTKATTGCTATATA
ATATGCTAACTTAATYRTTGGGGTGGTGTTAATATGTATGGAGCACTTTATGATAAACTCATCTGTA
TGTTTAATAGAAAGGCAAAGCGTTGAAGTAGGGAGGTGTGCCATACATCAAAGAAGGCTTATTAAC
CGACRTTAGTAATAGACATAGGTTATTGWWTTCAKSTCCTCCCAACACGTCTCAGACGTTTTGTAGA
ACTGGCATTGGTATCTTTAGTACCCAATCATATAATTGCATAATCTAGACTATGATTATTGCTTGGTT
GGCAACTCAGGTATGTGAAGCTTAAATAAAGTGAACTTACATATGAAAAGGATACTATTTTGATGTT
ATTTTCTGGGGAACAGTGGATATTGTTGTCAAATTGCTTGCTACATTCCTGCTTGATATGTCATGTCA
TTACTTATTATTTATTGTTGGGAGATTCTGGTATCATTCCAAGATCCATAGACGATTTGAGAAGTCTT
AAAATTCTTTTATTCCAAGGTTAATTCATTTGGTGTAGTCATCCGGTATCCAATAACCCCCCCCCCCC
CCGACTAATCCAGATTCATGCCAGTCCATTAAAATAGGAATTGCTCGGTTATGAGAAATTTCTCCAT
TAGGGCTCAAACCCGAGAACCATGGCTAAGGGCGGAGGGATCCCCTTGGTGGTCTCAAGGGTTAAG
TATTATTATCATTTCATGTTCATATGATTGTAACATTTAGATATCCTTGCGAAGTGTAATATTGTTTT
GTCTCGTCATTTGTTTTATGCTTCATATAAATTCCCTTTTCTTTCGCCCCTGTCTCTTCGTTTCTTTTTA
TCCAGTGTACTTCATTCTGCTAGAATTATATAGTGGTCGGTGTTCCTTTCTGAGCTTCATTGATGATA
TTGAATCCTTTCAGCTGTTTACCGGTTGATTACTTTTATGTTAGTGCTTAGCTTTTTGTTTTTATATCT
TTGCAACTGTACTTACCATCTTCCTGAAACTTCTCTTTACAGGTTAGCATTGGTAAGCAATGGAAGC
AGTTTCTAGATCATAGCGATTCGGTTGGCTTCATATTTCATGTATGAAATCTTGGTTATCGTACGCCT
TGAATTCAGTTTCTCTTCAATTAGCACGACTCACAAGGAATCATCTTCTTTTGCAGGACGATGCAAA
GAGGCTCGACAGAAATGCCAAGAATCGTTACACAGTATAG

8. Desirée _eIF4E- mRNA-CDS-1 consensus sequence- (St_Des-eIF4E-CDS-1, trimmed CDS, after alignment with reference mRNA sequence NM_001288431.1).

ATGGCAGCAGCTGAAATGGAGAGAACGACGTCGTTTGATGCAGCTGAGAAGTTGAAGGCCGCCGAT
GCAGGAGGAGGAGAGGTAGACGATGAACTTGAAGAAGGTGAAATTGTTGAAGAATCAAATGATAC
GGCGTCGTATTTAGGGAAAGAAATCACAGTGAAACATCCATTGGAGCATTCATGGACTTTTTGGTTT

GATAGCCCTATTGCTAAATCTCGACAAACTGCTTGGGGAAGCTCACTTCGAAATGTCTACACTTTCT
CCACTGTTGAAGATTTTTGGGGTGCTTACAATAATATCCATCACCCAAGCAAGTTGGTTATGGGAGC
AGACTTTCATTGTTTTAAGCATAAAATTGAGCCAAAGTGGGAAGATCCTGTATGTGCCAATGGAGG
GACGTGGAAAATGAATTTTTTGAAGGGTAAATCTGATACCAGCTGGCTATATACGCTGCTGGCAATG
ATTGGACATCAATTCGATCACGGAGATGAAATTTGTGGAGCAGTCGTTAGTGTCCGGTCTAAGGGA
GAAAAAATAGCTTTGTGGACCAAGAATGCTGCAAATGAAACAGCTCAGGTTAGCATTGGTAAGCAA
TGGAAGCAGTTTCTAGATTACAGCGATTCGGTTGGCTTCATATTTCACGATGATGCAAAGAGGCTCG
ACAGAAGTGCCAAGAATCGTTACACAGTATAG

9. Desirée _mRNA-CDS-2 consensus sequence (St_Des-eIF4E-CDS-2, trimmed CDS, after alignment with reference mRNA sequence)

ATGGCAGCAGCTGAAATGGAGAGAACGACGTCGTTTGATGCAGCTGAGAAGTTGAAGGCCGCCGAT
GCAGGAGGAGGAGAGGTAGACGATGAACTTGAAGAAGGTGAAATTGTTGAAGAATCAAATGATAC
GGCGTCGTATTTAGGGAAAGAAATCACAGTGAAACATCCATTGGAGCATTCATGGACTTTTTGGTTT
GATAGCCCTATTGCTAAATCTCGACAAACTGCTTGGGGAAGCTCACTTCGAAATGTCTACACTTTCT
CCACTGTTGAAGATTTTTGGGGTGCTTACAATAATATCCATCACCCAAGCAAGTTGGTTATGGGAGC
AGACTTTCATTGTTTTAAGCATAAAATTGAGCCAAAGTGGGAAGATCCTGTATGTGCCAATGGAGG
GACGTGGAAAATGAGTTTTTCGAAGGGTAAATCTGATACCAGCTGGCTATATACGCTGCTGGCTATG
ATTGGACATCAATTTGATCATGGAGATGAAATTTGTGGAGCAGTCGTTAGTGTCCGGGCTAAGGGA
GAAAAAATAGCTTTGTGGACCAAGAATGCTGCAAATGAAACAGCTCAGGTTAGCATTGGTAAGCAA
TGGAAGCAGTTTCTAGATCATAGCGATTCGGTTGGCTTCATATTTCATGACGATGCAAAGAGGCTCG
ACAGAAATGCCAAGAATCGTTACACAGTATAG

10. Desirée _mRNA-CDS-5 consensus sequence (St_Des-eIF4E-CDS-5, trimmed CDS, after alignment with reference mRNA sequence)

ATGGCAGCAGCTGAAATGGAGAGAACGACGTCGTTTGATGCAGCTGAGAAGTTGAAGGCCGCCGAT
GCAGGAGGAGGAGAGGTAGACGATGAACTTGAAGAAGGTGAAATTGTTGAAGAATCAAATGATAC
GGCGTCGTATTTAGGGAAAGAAATCACAGTGAAGCATCCATTGGAGCATTCATGGACTTTTTGGTTT
GATAGCCCTATTGCTAAATCTCGACAAACTGCTTGGGGAAGCTCCCTTCGAAATGTCTACACTTTCT
CCACTGTTGAAGATTTTTGGGGTGCTTACAATAATATCCATCACCCAAGCAAGTTGGTTATGGGAGC
AGACTTTCATTGTTTTAAGCATAAAATTGAGCCAAAGTGGGAAGATCCTGTATGTACCAATGGAGG
GACGTGGAAAATGAGTTTTTCGAAGGGTAAATCTGATACCAGCTGGCTATATACGCTGCTGGCTATG
ATTGGACATCAATTTGATCATGGAGATGAAATTTGTGGAGCAGTCGTTAGTGTCCGGGCTAAGGGA
GAAAAAATAGCTTTGTGGACCAAGAATGCTGCAAATGAAACAGCTCAGGTTAGCATTGGTAAGCAA
TGGAAGCAGTTTCTAGATCATAGCGATTCGGTTGGCTTCATATTTCATGACGATGCAAAGAGGCTCG
ACAGAAATGCCAAGAATCGTTACACAGTATAG

11. Celandine_mRNA-CDS-1 consensus sequence (trimmed CDS, after alignment with reference mRNA sequence)

ATGGCAGCAGCTGAAATGGAGAGAACGACGTCGTTTGATGCAGCTGAGAAGTTGAAGGCCGCCGAT
GCAGGAGGAGGAGAGGTAGACGATGAACTTGAAGAAGGTGAAATTGTTGAAGAATCAAATGATAC
GGCGTCGTATTTGGGGAAAGAAATCACAGTGAAGCATCCATTGGAGCATTCATGGACTTTTTGGTTT
GATAGCCCTATTGCTAAATCTCGACAAACTGCTTGGGGAAGCTCACTTCGAAATGTCTACACTTTCT
CCACTGTTGAAGAGTTTTGGGGTGCTTACAATAATATCCATCACCCAAGCAAGTTGGTTATGGGAGC
AGACTTTCATTGTTTTAAGCATAAAATTGAGCCAAAGTGGGAAGATCCTGTATGTGCCAATGGAGG
GACGTGGAAAATGAGTTTTTTGAAGGGTAAATCTGATACCAGCTGGCTATATACGCTGCTGGCTATG
ATTGGACATCAATTTGATCATGGAGATGAAATTTGTGGAGCAGTCGTTAGTGTCCGGTCTAAGGGAG
AAAAAATAGCTTTGTGGACCAAGAATGCTGCAAATGAAACAGCTCAGGTTAGCATTGGTAAGCAAT
GGAAGCAGTTTCTAGATCATAGCGATTCGGTTGGCTTCATATTTCATGACGATGCAAAGAGGCTCGA
CAGAAGTGCCAAGAATCGTTACACAGTATAG

12. Celandine_mRNA-CDS-2 consensus sequence (trimmed CDS, after alignment with reference mRNA sequence)

ATGGCAGCAGCTGAAATGGAGAGAACGACGTCGTTTGATGCAGCTGAGAAGTTGAAGGCCGCCGAT
GCAGGAGGAGGAGAGGTAGACGATGAACTTGAAGAAGGTGAAATTGTTGAAGAATCAAATGATAC
GGCGTCGTATTTGGGGAAAGAAATCACAGTGAAGCATCCATTGGAGCATTCATGGACTTTTTGGTTT
GATAGCCCTATTGCTAAATCTCGACAAACTGCTTGGGGAAGCTCCCTTCGAAATGTCTACACTTTCT
CCACTGTTGAAGATTTTTGGGGTGCTTACAATAATATCCATCACCCAAGCAAGTTGGTTATGGGAGC
AGACTTTCATTGTTTTAAGCATAAAATTGAGCCAAAGTGGGAAGATCCTGTATGTACCAATGGAGG
GACGTGGAAAATGAGTTTTTCGAAGGGTAAATCTGATACCAGCTGGCTATATACGCTGCTGGCTATG
ATTGGACATCAATTTGATCATGGAGATGAAATTTGTGGAGCAGTCGTTAGTGTCCGGGCTAAGGGA
GAAAAAATAGCTTTGTGGACCAAGAATGCTGCAAATGAAACAGCTCAGGTTAGCATTGGTAAGCAA
TGGAAGCAGTTTCTAGATCATAGCGATTCGGTTGGCTTCATATTTCATGACGATGCAAAGAGGCTCG
ACAGAAGTGCCAAGAATCGTTACACAGTATAG

13. Kuras_ mRNA-CDS-k1 consensus sequence (trimmed CDS, after alignment with reference mRNA sequence)

ATGGCAGCAGCTGAAATGGAGAGAACGACGTCGTTTGATGCAGCTGAGAAGTTGAAGGCCGCCGAT
GCAGGAGGAGGAGAGGTAGACGATGAACTTGAAGAAGGTGAAATTGTTGAAGAATCAAATGATAC
GGCGTCGTATTTGGGGAAAGAAATCACAGTGAAGCATCCATTGGAGCATTCATGGACTTTTTGGTTT
GATAGCCCTATTGCTAAATCTCGACAAACTGCTTGGGGAAGCTCACTTCGAAATGTCTACACTTTCT
CCACTGTTGAAGAGTTTTGGGGTGCTTACAATAATATCCATCACCCAAGCAAGTTGGTTATGGGAGC
AGACTTTCATTGTTTTAAGCATAAAATTGAGCCAAAGTGGGAAGATCCTGTATGTGCCAATGGAGG
GACGTGGAAAATGAATTTTTTGAAGGGTAAATCTGATACCAGCTGGCTATATACGCTGCTGGCAATG
ATTGGACATCAATTCGATCACGGAGATGAAATTTGTGGAGCAGTCGTTAATGTCCGGTCTAAGGGA
GAAAAAATAGCTTTGTGGACCAAGAATGCTGCAAATGAAATAGCTCAGGTTAGCATTGGTAAGCAA
TGGAAGCAGTTTCTAGATTATAGCGATTCGGTTGGCTTCATATTTCATGACGATGCAAAGAGGCTCG
ACAGAAATGCCAAGAATCGTTACACAGTATAG

14. Kuras_ mRNA-CDS-k2 consensus sequence (trimmed CDS, after alignment with reference mRNA sequence)

ATGGCAGCAGCTGAAATGGAGAGAACGACGTCGTTTGATGCAGCTGAGAAGTTGAAGGCCGCCGAT
GCAGGAGGAGGAGAGGTAGACGATGAACTTGAAGAAGGTGAAATTGTTGAAGAATCAAATGATAC
GGCGTCGTATTTAGGGAAAGAAATCACAGTGAAACATCCATTGGAGCATTCATGGACTTTTTGGTTT
GATAGCCCTATTGCTAAATCTCGACAAACTGCTTGGGGAAGCTCACTTCGAAATGTCTACACTTTCT
CCACTGTTGAAGATTTTTGGGGTGCTTACAATAATATCCATCACCCAAGCAAGTTGGTTATGGGAGC
AGACTTTCATTGTTTTAAGCATAAAATTGAGCCAAAGTGGGAAGATCCTGTATGTGCCAATGGAGG
GACGTGGAAAATGAATTTTTTGAAGGGTAAATCTGATACCAGCTGGCTATATACGCTGCTGGCAATG
ATTGGACATCAATTCGATCACGGAGATGAAATTTGTGGAGCAGTCGTTAGTGTCCGGTCTAAGGGA
GAAAAAATAGCTTTGTGGACCAAGAATGCTGCAAATGAAACAGCTCAGGTTAGCATTGGTAAGCAA
TGGAAGCAGTTTCTAGATTACAGCGATTCGGTTGGCTTCATATTTCACGATGATGCAAAGAGGCTCG
ACAGAAGTGCCAAGAATCGTTACACAGTATAG

15. Kuras_ mRNA-CDS- k3 consensus sequence (trimmed CDS, after alignment with reference mRNA sequence)

ATGGCAGCAGCTGAAATGGAGAGAACGACGTCGTTTGATGCAGCTGAGAAGTTGAAGGCCGCCGAT
GCAGGAGGAGGAGAGGTAGACGATGAACTTGAAGAAGGTGAAATTGTTGAAGAATCAAATGATAC
GGCGTCGTATTTAGGGAAAGAAATCACAGTGAAACATCCATTGGAGCATTCATGGACTTTTTGGTTT
GATAGCCCTATTGCTAAATCTCGACAAACTGCTTGGGGAAGCTCACTTCGAAATGTCTACACTTTCT
CCACTGTTGAAGATTTTTGGGGTGCTTACAATAATATCCATCACCCAAGCAAGTTGGTTATGGGAGC
AGACTTTCATTGTTTTAAGCATAAAATTGAGCCAAAGTGGGAAGATCCTGTATGTGCCAATGGAGG
GACGTGGAAAATGAATTTTTTGAAGGGTAAATCTGATACCAGCTGGCTATATACGCTGCTGGCAATG
ATTGGACATCAATTCGATCACGGAGATGAAATTTGTGGAGCAGTCGTTAGTGTCCGGTCTAAGGGA
GAAAAAATAGCTTTGTGGACCAAGAATGCTGCAAATGAAACAGCTCAGGTTAGCATTGGTAAGCAA
TGGAAGCAGTTTCTAGATTACAGCGATTCGGTTGGCTTCATATTTCACGATGATGCAAAGAGGCTCG
ACAGAAGTGCCAAGAATCGTTACACAGTATAG

16. Kuras_ mRNA-CDS- k6 consensus sequence (trimmed CDS, after alignment with reference mRNA sequence)

ATGGCAGCAGCTGAAATGGAGAGAACGACGTCGTTTGATGCAGCTGAGAAGTTGAAGGCCGCCGAT
GCAGGAGGAGGAGAGGTAGACGATGAACTTGAAGAAGGTGAAATTGTTGAAGAATCAAATGATAC
GGCGTCGTATTTAGGGAAAGAAATCACAGTGAAGCATCCATTGGAGCATTCATGGACTTTTTGGTTT
GATAGCCCTATTGCTAAATCTCGACAAACTGCTTGGGGAAGCTCACTTCGAAATGTCTACACTTTCT
CCACTGTTGAAGAGTTTTGGGGTGCTTACAATAATATCCATCACCCAAGCAAGTTGGTTATGGGAGC
AGACTTTCATTGTTTTAAGCATAAAATTGAGCCAAAGTGGGAAGATCCTGTATGTGCCAATGGAGG
GACGTGGAAAATGAGTTTTTTGAAGGGTAAATCTGATACCAGCTGGCTATATACGCTGCTGGCTATG
ATTGGACATCAATTTGATCATGGAGATGAAATTTGTGGAGCAGTCGTTAGTGTCCGGTCTAAGGGAG
AAAAAATAGCTTTGTGGACCAAGAATGCTGCAAATGAAACAGCTCAGGTTAGCATTGGTAAGCAAT
GGAAGCAGTTTCTAGATCATAGCGATTCGGTTGGCTTCATATTTCATGACGATGCAAAGAGGCTCGA
CAGAAGTGCCAAGAATCGTTACACAGTATAG

17. Innovator_ mRNA-CDS- I1 consensus sequence (trimmed CDS, after alignment with reference mRNA sequence)

ATGGCAGCAGCTGAAATGGAGAGAACGACGTCGTTTGATGCAGCTGAGAAGTTGAAGGCCGCCGAT
GCAGGAGGAGGAGAGGTAGACGATGAACTTGAAGAAGGTGAAATTGTTGAAGAATCAAATGATAC
GGCGTCGTATTTAGGGAAAGAAATCACAGTGAAGCATCCATTGGAGCATTCATGGACTTTTTGGTTT
GATAGCCCTATTGCTAAATCTCGACAAACTGCTTGGGGAAGCTCACTTCGAAATGTCTACACTTTCT
CCACTGTTGAAGATTTTTGGGGTGCTTACAATAATATCCATCACCCAAGCAAGTTGGTTATGGGAGC
AGACTTTCATTGTTTTAAGCATAAAATTGAGCCAAAGTGGGAAGATCCTGTATGTGCCAATGGAGG
GACGTGGAAAATGAGTTTTTCGAAGGGTAAATCTGATACCAGCTGGCTATATACGCTGCTGGCAAT
GATTGGACATCAATTCGATCACGGAGATGAAATTTGTGGAGCAGTCGTTAGTGTCCGGGCTAAGGG
AGAAAAAATAGCTTTGTGGACCAAGAATGCTGCAAATGAAACAGCTCAGGTTAGCATCGGTAAGCA
ATGGAAGCAGTTTCTAGATTACAGCGATTCGGTTGGCTTCATATTTCATGACGATGCAAAGAGGCTC
GACAGAAATGCCAAGAATCGTTACACAGTATAG

18. Innovator_ mRNA-CDS- I2 consensus sequence (trimmed CDS, after alignment with reference mRNA sequence)

ATGGCAGCAGCTGAAATGGAGAGAACGACGTCGTTTGATGCAGCTGAGAAGTTGAAGGCCGCCGAT
GCAGGAGGAGGAGAGGTAGACGATGAACTTGAAGAAGGTGAAATTGTTGAAGAATCAAATGATAC
GGCGTCGTATTTAGGGAAAGAAATCACAGTGAAGCATCCATTGGAGCATTCATGGACTTTTTGGTTT
GATAGCCCTATTGCTAAATCTCGACAAACTGCTTGGGGAAGCTCACTTCGAAATGTCTACACTTTCT
CCACTGTTGAAGATTTTTGGGGTGCTTACAATAATATCCATCACCCAAGCAAGTTGGTTATGGGAGC
AGACTTTCATTGTTTTAAGCATAAAATTGAGCCAAAGTGGGAAGATCCTGTATGTGCCAATGGAGG
GACGTGGAAAATGAGTTTTTCGAAGGGTAAATCTGATACCAGCTGGCTATATACGCTGCTGGCAAT
GATTGGACATCAATTCGATCACGGAGATGAAATTTGTGGAGCAGTCGTTAGTGTCCGGGCTAAGGG
AGAAAAAATAGCTTTGTGGACCAAGAATGCTGCAAATGAAACAGCTCAGGTTAGCATTGGTAAGCA
ATGGAAGCAGTTTCTAGATTATAGCGATTCGGTTGGCTTCATATTTCATGACGATGCAAAGAGGCTC
GACAGAAATGCCAAGAATCGTTACACAGTATAG

19. Innovator_ mRNA-CDS- I4 consensus sequence (trimmed CDS, after alignment with reference mRNA sequence)

ATGGCAGCAGCTGAAATGGAGAGAACGACGTCGTTTGATGCAGCTGAGAAGTTGAAGGCCGCCGAT
GGAGGAGGAGGGGAGGTAGACGATGAACTTGAAGAAGGTGAAATTGTTGAAGAATCAAATGATAT
GGCGTCGTATTTAGGGAAAGAAATCACAGTGAAGCATCCATTGGAGCATTCATGGACTTTTTGGTTT
GATAGCCCTATTGCTAAATCTCGACAAACTGCTTGGGGAAGCTCACTTCGAAATGTCTACACTTTCT
CCACTGTTGAAGATTTTTGGGGTGCTTACAATAATATCCATCACCCAAGCAAGTTGGTTATGGGAGC
AGACTTTCATTGTTTTAAGCATAGAATTGAGCCAAAGTGGGAAGATCCTGTATGTGCCAATGGAGG
GACGTGGAAAATGAGTTTTTCGAAGGGTAAATCTGATACCAGCTGGCTATATACGCTGCTGGCTATG
ATTGGACATCAATTTGATCATGGAGATGAAATTTGTGGAGCAGTCGTTAGTGTCCGGTCTAAGGGAG
AAAAAATAGCTTTGTGGACCAAGAATGCTGCAAATGAAACAGCTCAGGTTAGCATTGGTAAGCAAT
GGAAGCAGTTTCTAGATTATAGCGATTCGGTTGGCTTCATATTTCATGACGATGCAAAGAGGCTCGA
CAGAAATGCCAAGAATCGTTACACAGTATAG

20. Innovator_ mRNA-CDS- I7 consensus sequence (trimmed CDS, after alignment with reference mRNA sequence)

ATGGCAGCAGCTGAAATGGAGAGAACGACGTCGTTTGATGCAGCTGAGAAGTTGAAGGCCGCCGAT
GGAGGAGGAGGGGAGGTAGACGATGAACTTGAAGAAGGTGAAATTGTTGAAGAATCAAATGATAT
GGCGTCGTATTTAGGGAAAGAAATCACAGTGAAGCATCCATTGGAGCATTCATGGACTTTTTGGTTT
GATAGCCCTATTGCTAAATCTCGACAAACTGCTTGGGGAAGCTCACTTCGAAATGTCTACACTTTCT
CCACTGTTGAAGATTTTTGGGGTGCTTACAATAATATCCATCACCCAAGCAAGTTGGTTATGGGAGC
AGACTTTCATTGTTTTAAGCATAGAATTGAGCCAAAGTGGGAAGATCCTGTATGTGCCAATGGAGG
GACGTGGAAAATGAGTTTTTCGAAGGGTAAATCTGATACCAGCTGGCTATATACGCTGCTGGCAAT
GATTGGACATCAATTCGATCATGGAGATGAAATTTGTGGAGCAGTCGTTAGTGTCCGGGCTAAGGG
AGAAAAAATAGCTTTGTGGACCAAGAATGCTGCAAATGAAACAGCTCAGGTTAGCATTGGTAAGCA
ATGGAAGCAGTTTCTAGATTATAGCGATTCGGTTGGCTTCATATTTCATGACGATGCAAAGAGGCTC
GACAGAAATGCCAAGAATCGTTACACAGTATAG

21. Innovator_ mRNA-CDS- I9 consensus sequence (trimmed CDS, after alignment with reference mRNA sequence)

ATGGCAGCAGCTGAAATGGAGAGAACGACGTCGTTTGATGCAGCTGAGAAGTTGAAGGCCGCCGAT
GCAGGAGGAGGAGAGGTAGACGATGAACTTGAAGAAGGTGAAATTGTTGAAGAATCAAATGATAC
GGCGTCGTATTTGGGGAAAGAAATCACAGTGAAGCATCCATTGGAGCATTCATGGACTTTTTGGTTT
GATAGCCCTATTGCTAAATCTCGACAAACTGCTTGGGGAAGCTCACTTCGAAATGTCTACACTTTCT
CCACTGTTGAAGAGTTTTGGGGTGCTTACAATAATATCCATCACCCAAGCAAGTTGGTTATGGGAGC
AGACTTTCATTGTTTTAAGCATAAAATTGAGCCAAAGTGGGAAGATCCTGTATGTGCCAATGGAGG
GACGTGGAAAATGAGTTTTTTGAAGGGTAAATCTGATACCAGCTGGCTATATACGCTGCTGGCTATG
ATTGGACATCAATTTGATCATGGAGATGAAATTTGTGGAGCAGTCGTTAGTGTCCGGTCTAAGGGAG
AAAAAATAGCTTTGTGGACCAAGAATGCTGCAAATGAAACAGCTCAGGTTAGCATTGGTAAGCAAT
GGAAGCAGTTTCTAGATCATAGCGATTCGGTTGGCTTCATATTTCATGACGATGCAAAGAGGCTCGA
CAGAAGTGCCAAGAATCGTTACACAGTATAG

D. Sequences for in-silico maps (Figure 8.3) of plasmids used for in-vitro cleavage assay.

1. Plasmid PCR z.b._De-eIF4E-1 (nucleotides in green color is insert sequence)

CTTCCGCTTCCTCGCTCACTGACTCGCTGCGCTCGGTCGTTCGGCTGCGGCGAGCGGTATCAGCTCA
CTCAAAGGCGGTAATACGGTTATCCACAGAATCAGGGGATAACGCAGGAAAGAACATGTGAGCAA
AAGGCCAGCAAAAGGCCAGGAACCGTAAAAAGGCCGCGTTGCTGGCGTTTTTCCATAGGCTCCGCC
CCCCTGACGAGCATCACAAAAATCGACGCTCAAGTCAGAGGTGGCGAAACCCGACAGGACTATAAA
GATACCAGGCGTTTCCCCCTGGAAGCTCCCTCGTGCGCTCTCCTGTTCCGACCCTGCCGCTTACCGG
ATACCTGTCCGCCTTTCTCCCTTCGGGAAGCGTGGCGCTTTCTCATAGCTCACGCTGTAGGTATCTCA
GTTCGGTGTAGGTCGTTCGCTCCAAGCTGGGCTGTGTGCACGAACCCCCCGTTCAGCCCGACCGCTG
CGCCTTATCCGGTAACTATCGTCTTGAGTCCAACCCGGTAAGACACGACTTATCGCCACTGGCAGCA
GCCACTGGTAACAGGATTAGCAGAGCGAGGTATGTAGGCGGTGCTACAGAGTTCTTGAAGTGGTGG
CCTAACTACGGCTACACTAGAAGAACAGTATTTGGTATCTGCGCTCTGCTGAAGCCAGTTACCTTCG

GAAAAAGAGTTGGTAGCTCTTGATCCGGCAAACAAACCACCGCTGGTAGCGGTGGTTTTTTTGTTTG
CAAGCAGCAGATTACGCGCAGAAAAAAAGGATCTCAAGAAGATCCTTTGATCTTTTCTACGGGGTC
TGACGCTCAGTGGAACGAAAACTCACGTTAAGGGATTTTGGTCATGAGATTATCAAAAAGGATCTT
CACCTAGATCCTTTTAAATTAAAAATGAAGTTTTAGCACGTGTCAGTCCTGCTCCTCGGCCACGAAG
TGCACGCAGTTGCCGGCCGGGTCGCGCAGGGCGAACTCCCGCCCCCACGGCTGCTCGCCGATCTCG
GTCATGGCCGGCCCGGAGGCGTCCCGGAAGTTCGTGGACACGACCTCCGACCACTCGGCGTACAGC
TCGTCCAGGCCGCGCACCCACACCCAGGCCAGGGTGTTGTCCGGCACCACCTGGTCCTGGACCGCG
CTGATGAACAGGGTCACGTCGTCCCGGACCACACCGGCGAAGTCGTCCTCCACGAAGTCCCGGGAG
AACCCGAGCCGGTCGGTCCAGAACTCGACCGCTCCGGCGACGTCGCGCGCGGTGAGCACCGGAACG
GCACTGGTCAACTTGGCCATGGTGGCCCTCCTCACGTGCTATTATTGAAGCATTTATCAGGGTTATT
GTCTCATGAGCGGATACATATTTGAATGTATTTAGAAAAATAAACAAATAGGGGTTCCGCGCACATT
TCCCCGAAAAGTGCCACCTGATGCGGTGTGAAATACCGCACAGATGCGTAAGGAGAAAATACCGCA
TCAGGAAATTGTAAGCGTTAATAATTCAGAAGAACTCGTCAAGAAGGCGATAGAAGGCGATGCGCT
GCGAATCGGGAGCGGCGATACCGTAAAGCACGAGGAAGCGGTCAGCCCATTCGCCGCCAAGCTCTT
CAGCAATATCACGGGTAGCCAACGCTATGTCCTGATAGCGGTCCGCCACACCCAGCCGGCCACAGT
CGATGAATCCAGAAAAGCGGCCATTTTCCACCATGATATTCGGCAAGCAGGCATCGCCATGGGTCA
CGACGAGATCCTCGCCGTCGGGCATGCTCGCCTTGAGCCTGGCGAACAGTTCGGCTGGCGCGAGCC
CCTGATGCTCTTCGTCCAGATCATCCTGATCGACAAGACCGGCTTCCATCCGAGTACGTGCTCGCTC
GATGCGATGTTTCGCTTGGTGGTCGAATGGGCAGGTAGCCGGATCAAGCGTATGCAGCCGCCGCAT
TGCATCAGCCATGATGGATACTTTCTCGGCAGGAGCAAGGTGAGATGACAGGAGATCCTGCCCCGG
CACTTCGCCCAATAGCAGCCAGTCCCTTCCCGCTTCAGTGACAACGTCGAGCACAGCTGCGCAAGG
AACGCCCGTCGTGGCCAGCCACGATAGCCGCGCTGCCTCGTCTTGCAGTTCATTCAGGGCACCGGAC
AGGTCGGTCTTGACAAAAAGAACCGGGCGCCCCTGCGCTGACAGCCGGAACACGGCGGCATCAGA
GCAGCCGATTGTCTGTTGTGCCCAGTCATAGCCGAATAGCCTCTCCACCCAAGCGGCCGGAGAACCT
GCGTGCAATCCATCTTGTTCAATCATGCGAAACGATCCTCATCCTGTCTCTTGATCAGAGCTTGATCC
CCTGCGCCATCAGATCCTTGGCGGCAAGAAAGCCATCCAGTTTACTTTGCAGGGCTTCCCAACCTTA
CCAGAGGGCGCCCCAGCTGGCAATTCCGGTTCGCTTGCTGTCCATAAAACCGCCCAGTCTAGCTATC
GCCATGTAAGCCCACTGCAAGCTACCTGCTTTCTCTTTGCGCTTGCGTTTTCCCTTGTCCAGATAGCC
CAGTAGCTGACATTCATCCGGGGTCAGCACCGTTTCTGCGGACTGGCTTTCTACGTGAAAAGGATCT
AGGTGAAGATCCTTTTTGATAATCTCATGCCTGACATTTATATTCCCCAGAACATCAGGTTAATGGC
GTTTTTGATGTCATTTTCGCGGTGGCTGAGATCAGCCACTTCTTCCCCGATAACGGAGACCGGCACA
CTGGCCATATCGGTGGTCATCATGCGCCAGCTTTCATCCCCGATATGCACCACCGGGTAAAGTTCAC
GGGAGACTTTATCTGACAGCAGACGTGCACTGGCCAGGGGGATCACCATCCGTCGCCCCGGCGTGT
CAATAATATCACTCTGTACATCCACAAACAGACGATAACGGCTCTCTCTTTTATAGGTGTAAACCTT
AAACTGCCGTACGTATAGGCTGCGCAACTGTTGGGAAGGGCGATCGGTGCGGGCCTCTTCGCTATTA
CGCCAGCTGGCGAAAGGGGGATGTGCTGCAAGGCGATTAAGTTGGGTAACGCCAGGGTTTTCCCAG
TCACGACGTTGTAAAACGACGGCCAGTGAATTGTAATACGACTCACTATAGGGCGAATTGGGCCCT
CTAGATGCATGCTCGAGCGGCCGCCAGTGTGATGGATATCTGCAGAATTCAGGCAGCTGAGAAGTT
GAAGGCCGCCGATGCAGGAGGAGGAGAGGTAGACGATGAACTTGAAGAAGGTGAAATTGTTGAAG
AATCAAATGATACGGCGTCGTATTTAGGGAAAGAAATCACAGTGAAACATCCATTGGAGCATTCAT
GGACTTTTTGGTTTGATAGCCCTATTGCTAAATCTCGACAAACTGCTTGGGGAAGCTCACTTCGAAA
TGTCTACACTTTCTCCACTGTTGAAGATTTTTGGGGTGCTTACAATAATATCCATCACCCAAGCAAGT
TGGTTATGGGAGCAGACTTTCATTGTTTTAAGCATAAAATTGAGCCAAAGTGGGAAGATCCTGTATG
TGCCAATGGAGGGACGTGGAAAATGAATTTTTTGAAGGGTAAATCTGATACCAGCTGGCTATATAC
GCTGCTGGCAATGATTGGACATCAATTCGATCACGGAGATGAAATTTGTGGAGCAGTCGTTAGTGTC

CGGTCTAAGGGAGAAAAAATAGCTTTGTGGACCAAGAATGCTGCAAATGAAACAGCTCAGGTTAGC
ATTGGTAAGCAATGGAAGCAGTTTCTAGATTACAGCGATTCGGTTGGCTTCATATTTCACGATGATG
CAAAGAGGCTCGACAGAAGTGCCAAGAATCGTTACACAGTATAGTTCTTGATGCAATGTGGGAATG
CAAGAAACACAATTCGTACCGGAAAGTTGGTAGGCACTCCTGAATTCCAGCACACTGGCGGCCGTT
ACTAGTGGATCCGAGCTCGGTACCAAGCTTGATGCATAGCTTGAGTATTCTAACGCGTCACCTAAAT
AGCTTGGCGTAATCATGGTCATAGCTGTTTCCTGTGTGAAATTGTTATCCGCTCACAATTCCACACA
ACATACGAGCCGGAAGCATAAAGTGTAAAGCCTGGGGTGCCTAATGAGTGAGCTAACTCACATTAA
TTGCGTTGCGCTCACTGCCCGCTTTCCAGTCGGGAAACCTGTCGTGCCAGCTGCATTAATGAATCGG
CCAACGCGCGGGGAGAGGCGGTTTGCGTATTGGGCGCT

2. Plasmid PCR z.b._De-eIF4E-2 (nucleotides in green color is insert sequence)

CTTCCGCTTCCTCGCTCACTGACTCGCTGCGCTCGGTCGTTCGGCTGCGGCGAGCGGTATCAGCTCA
CTCAAAGGCGGTAATACGGTTATCCACAGAATCAGGGGATAACGCAGGAAAGAACATGTGAGCAA
AAGGCCAGCAAAAGGCCAGGAACCGTAAAAAGGCCGCGTTGCTGGCGTTTTTCCATAGGCTCCGCC
CCCCTGACGAGCATCACAAAAATCGACGCTCAAGTCAGAGGTGGCGAAACCCGACAGGACTATAAA
GATACCAGGCGTTTCCCCCTGGAAGCTCCCTCGTGCGCTCTCCTGTTCCGACCCTGCCGCTTACCGG
ATACCTGTCCGCCTTTCTCCCTTCGGGAAGCGTGGCGCTTTCTCATAGCTCACGCTGTAGGTATCTCA
GTTCGGTGTAGGTCGTTCGCTCCAAGCTGGGCTGTGTGCACGAACCCCCCGTTCAGCCCGACCGCTG
CGCCTTATCCGGTAACTATCGTCTTGAGTCCAACCCGGTAAGACACGACTTATCGCCACTGGCAGCA
GCCACTGGTAACAGGATTAGCAGAGCGAGGTATGTAGGCGGTGCTACAGAGTTCTTGAAGTGGTGG
CCTAACTACGGCTACACTAGAAGAACAGTATTTGGTATCTGCGCTCTGCTGAAGCCAGTTACCTTCG
GAAAAAGAGTTGGTAGCTCTTGATCCGGCAAACAAACCACCGCTGGTAGCGGTGGTTTTTTTGTTTG
CAAGCAGCAGATTACGCGCAGAAAAAAAGGATCTCAAGAAGATCCTTTGATCTTTTCTACGGGGTC
TGACGCTCAGTGGAACGAAAACTCACGTTAAGGGATTTTGGTCATGAGATTATCAAAAAGGATCTT
CACCTAGATCCTTTTAAATTAAAAATGAAGTTTTAGCACGTGTCAGTCCTGCTCCTCGGCCACGAAG
TGCACGCAGTTGCCGGCCGGGTCGCGCAGGGCGAACTCCCGCCCCCACGGCTGCTCGCCGATCTCG
GTCATGGCCGGCCCGGAGGCGTCCCGGAAGTTCGTGGACACGACCTCCGACCACTCGGCGTACAGC
TCGTCCAGGCCGCGCACCCACACCCAGGCCAGGGTGTTGTCCGGCACCACCTGGTCCTGGACCGCG
CTGATGAACAGGGTCACGTCGTCCCGGACCACACCGGCGAAGTCGTCCTCCACGAAGTCCCGGGAG
AACCCGAGCCGGTCGGTCCAGAACTCGACCGCTCCGGCGACGTCGCGCGCGGTGAGCACCGGAACG
GCACTGGTCAACTTGGCCATGGTGGCCCTCCTCACGTGCTATTATTGAAGCATTTATCAGGGTTATT
GTCTCATGAGCGGATACATATTTGAATGTATTTAGAAAAATAAACAAATAGGGGTTCCGCGCACATT
TCCCCGAAAAGTGCCACCTGATGCGGTGTGAAATACCGCACAGATGCGTAAGGAGAAAATACCGCA
TCAGGAAATTGTAAGCGTTAATAATTCAGAAGAACTCGTCAAGAAGGCGATAGAAGGCGATGCGCT
GCGAATCGGGAGCGGCGATACCGTAAAGCACGAGGAAGCGGTCAGCCCATTCGCCGCCAAGCTCTT
CAGCAATATCACGGGTAGCCAACGCTATGTCCTGATAGCGGTCCGCCACACCCAGCCGGCCACAGT
CGATGAATCCAGAAAAGCGGCCATTTTCCACCATGATATTCGGCAAGCAGGCATCGCCATGGGTCA
CGACGAGATCCTCGCCGTCGGGCATGCTCGCCTTGAGCCTGGCGAACAGTTCGGCTGGCGCGAGCC
CCTGATGCTCTTCGTCCAGATCATCCTGATCGACAAGACCGGCTTCCATCCGAGTACGTGCTCGCTC
GATGCGATGTTTCGCTTGGTGGTCGAATGGGCAGGTAGCCGGATCAAGCGTATGCAGCCGCCGCAT
TGCATCAGCCATGATGGATACTTTCTCGGCAGGAGCAAGGTGAGATGACAGGAGATCCTGCCCCGG
CACTTCGCCCAATAGCAGCCAGTCCCTTCCCGCTTCAGTGACAACGTCGAGCACAGCTGCGCAAGG
AACGCCCGTCGTGGCCAGCCACGATAGCCGCGCTGCCTCGTCTTGCAGTTCATTCAGGGCACCGGAC
AGGTCGGTCTTGACAAAAAGAACCGGGCGCCCCTGCGCTGACAGCCGGAACACGGCGGCATCAGA
GCAGCCGATTGTCTGTTGTGCCCAGTCATAGCCGAATAGCCTCTCCACCCAAGCGGCCGGAGAACCT

GCGTGCAATCCATCTTGTTCAATCATGCGAAACGATCCTCATCCTGTCTCTTGATCAGAGCTTGATCC
CCTGCGCCATCAGATCCTTGGCGGCAAGAAAGCCATCCAGTTTACTTTGCAGGGCTTCCCAACCTTA
CCAGAGGGCGCCCCAGCTGGCAATTCCGGTTCGCTTGCTGTCCATAAAACCGCCCAGTCTAGCTATC
GCCATGTAAGCCCACTGCAAGCTACCTGCTTTCTCTTTGCGCTTGCGTTTTCCCTTGTCCAGATAGCC
CAGTAGCTGACATTCATCCGGGGTCAGCACCGTTTCTGCGGACTGGCTTTCTACGTGAAAAGGATCT
AGGTGAAGATCCTTTTTGATAATCTCATGCCTGACATTTATATTCCCCAGAACATCAGGTTAATGGC
GTTTTTGATGTCATTTTCGCGGTGGCTGAGATCAGCCACTTCTTCCCCGATAACGGAGACCGGCACA
CTGGCCATATCGGTGGTCATCATGCGCCAGCTTTCATCCCCGATATGCACCACCGGGTAAAGTTCAC
GGGAGACTTTATCTGACAGCAGACGTGCACTGGCCAGGGGGATCACCATCCGTCGCCCCGGCGTGT
CAATAATATCACTCTGTACATCCACAAACAGACGATAACGGCTCTCTCTTTTATAGGTGTAAACCTT
AAACTGCCGTACGTATAGGCTGCGCAACTGTTGGGAAGGGCGATCGGTGCGGGCCTCTTCGCTATTA
CGCCAGCTGGCGAAAGGGGGATGTGCTGCAAGGCGATTAAGTTGGGTAACGCCAGGGTTTTCCCAG
TCACGACGTTGTAAAACGACGGCCAGTGAATTGTAATACGACTCACTATAGGGCGAATTGGGCCCT
CTAGATGCATGCTCGAGCGGCCGCCAGTGTGATGGATATCTGCAGAATTCAGGCAGCTGAGAAGTT
GAAGGCCGCCGATGCAGGAGGAGGAGAGGTAGACGATGAACTTGAAGAAGGTGAAATTGTTGAAG
AATCAAATGATACGGCGTCGTATTTAGGGAAAGAAATCACAGTGAAACATCCATTGGAGCATTCAT
GGACTTTTTGGTTTGATAGCCCTATTGCTAAATCTCGACAAACTGCTTGGGGAAGCTCACTTCGAAA
TGTCTACACTTTCTCCACTGTTGAAGATTTTTGGGGTGCTTACAATAATATCCATCACCCAAGCAAGT
TGGTTATGGGAGCAGACTTTCATTGTTTTAAGCATAAAATTGAGCCAAAGTGGGAAGATCCTGTATG
TGCCAATGGAGGGACGTGGAAAATGAGTTTTTCGAAGGGTAAATCTGATACCAGCTGGCTATATAC
GCTGCTGGCTATGATTGGACATCAATTTGATCATGGAGATGAAATTTGTGGAGCAGTCGTTAGTGTC
CGGGCTAAGGGAGAAAAAATAGCTTTGTGGACCAAGAATGCTGCAAATGAAACAGCTCAGGTTAGC
ATTGGTAAGCAATGGAAGCAGTTTCTAGATCATAGCGATTCGGTTGGCTTCATATTTCATGACGATG
CAAAGAGGCTCGACAGAAATGCCAAGAATCGTTACACAGTATAGTTCTTGATGCAATGTGAGAATG
CAAAAAACACAATTCGTACCGGAAAGTTGGTAGGCACTCCTGAATTCCAGCACACTGGCGGCCGTT
ACTAGTGGATCCGAGCTCGGTACCAAGCTTGATGCATAGCTTGAGTATTCTAACGCGTCACCTAAAT
AGCTTGGCGTAATCATGGTCATAGCTGTTTCCTGTGTGAAATTGTTATCCGCTCACAATTCCACACA
ACATACGAGCCGGAAGCATAAAGTGTAAAGCCTGGGGTGCCTAATGAGTGAGCTAACTCACATTAA
TTGCGTTGCGCTCACTGCCCGCTTTCCAGTCGGGAAACCTGTCGTGCCAGCTGCATTAATGAATCGG
CCAACGCGCGGGGAGAGGCGGTTTGCGTATTGGGCGCT

3. Plasmid sequence: PCR z.b._De-eIF4E-5 (nucleotides in green color is the insert sequence)

CTTCCGCTTCCTCGCTCACTGACTCGCTGCGCTCGGTCGTTCGGCTGCGGCGAGCGGTATCAGCTCA
CTCAAAGGCGGTAATACGGTTATCCACAGAATCAGGGGATAACGCAGGAAAGAACATGTGAGCAA
AAGGCCAGCAAAAGGCCAGGAACCGTAAAAAGGCCGCGTTGCTGGCGTTTTTCCATAGGCTCCGCC
CCCCTGACGAGCATCACAAAAATCGACGCTCAAGTCAGAGGTGGCGAAACCCGACAGGACTATAAA
GATACCAGGCGTTTCCCCCTGGAAGCTCCCTCGTGCGCTCTCCTGTTCCGACCCTGCCGCTTACCGG
ATACCTGTCCGCCTTTCTCCCTTCGGGAAGCGTGGCGCTTTCTCATAGCTCACGCTGTAGGTATCTCA
GTTCGGTGTAGGTCGTTCGCTCCAAGCTGGGCTGTGTGCACGAACCCCCCGTTCAGCCCGACCGCTG
CGCCTTATCCGGTAACTATCGTCTTGAGTCCAACCCGGTAAGACACGACTTATCGCCACTGGCAGCA
GCCACTGGTAACAGGATTAGCAGAGCGAGGTATGTAGGCGGTGCTACAGAGTTCTTGAAGTGGTGG
CCTAACTACGGCTACACTAGAAGAACAGTATTTGGTATCTGCGCTCTGCTGAAGCCAGTTACCTTCG
GAAAAAGAGTTGGTAGCTCTTGATCCGGCAAACAAACCACCGCTGGTAGCGGTGGTTTTTTTGTTTG

CAAGCAGCAGATTACGCGCAGAAAAAAAGGATCTCAAGAAGATCCTTTGATCTTTTCTACGGGGTC
TGACGCTCAGTGGAACGAAAACTCACGTTAAGGGATTTTGGTCATGAGATTATCAAAAAGGATCTT
CACCTAGATCCTTTTAAATTAAAAATGAAGTTTTAGCACGTGTCAGTCCTGCTCCTCGGCCACGAAG
TGCACGCAGTTGCCGGCCGGGTCGCGCAGGGCGAACTCCCGCCCCCACGGCTGCTCGCCGATCTCG
GTCATGGCCGGCCCGGAGGCGTCCCGGAAGTTCGTGGACACGACCTCCGACCACTCGGCGTACAGC
TCGTCCAGGCCGCGCACCCACACCCAGGCCAGGGTGTTGTCCGGCACCACCTGGTCCTGGACCGCG
CTGATGAACAGGGTCACGTCGTCCCGGACCACACCGGCGAAGTCGTCCTCCACGAAGTCCCGGGAG
AACCCGAGCCGGTCGGTCCAGAACTCGACCGCTCCGGCGACGTCGCGCGCGGTGAGCACCGGAACG
GCACTGGTCAACTTGGCCATGGTGGCCCTCCTCACGTGCTATTATTGAAGCATTTATCAGGGTTATT
GTCTCATGAGCGGATACATATTTGAATGTATTTAGAAAAATAAACAAATAGGGGTTCCGCGCACATT
TCCCCGAAAAGTGCCACCTGATGCGGTGTGAAATACCGCACAGATGCGTAAGGAGAAAATACCGCA
TCAGGAAATTGTAAGCGTTAATAATTCAGAAGAACTCGTCAAGAAGGCGATAGAAGGCGATGCGCT
GCGAATCGGGAGCGGCGATACCGTAAAGCACGAGGAAGCGGTCAGCCCATTCGCCGCCAAGCTCTT
CAGCAATATCACGGGTAGCCAACGCTATGTCCTGATAGCGGTCCGCCACACCCAGCCGGCCACAGT
CGATGAATCCAGAAAAGCGGCCATTTTCCACCATGATATTCGGCAAGCAGGCATCGCCATGGGTCA
CGACGAGATCCTCGCCGTCGGGCATGCTCGCCTTGAGCCTGGCGAACAGTTCGGCTGGCGCGAGCC
CCTGATGCTCTTCGTCCAGATCATCCTGATCGACAAGACCGGCTTCCATCCGAGTACGTGCTCGCTC
GATGCGATGTTTCGCTTGGTGGTCGAATGGGCAGGTAGCCGGATCAAGCGTATGCAGCCGCCGCAT
TGCATCAGCCATGATGGATACTTTCTCGGCAGGAGCAAGGTGAGATGACAGGAGATCCTGCCCCGG
CACTTCGCCCAATAGCAGCCAGTCCCTTCCCGCTTCAGTGACAACGTCGAGCACAGCTGCGCAAGG
AACGCCCGTCGTGGCCAGCCACGATAGCCGCGCTGCCTCGTCTTGCAGTTCATTCAGGGCACCGGAC
AGGTCGGTCTTGACAAAAAGAACCGGGCGCCCCTGCGCTGACAGCCGGAACACGGCGGCATCAGA
GCAGCCGATTGTCTGTTGTGCCCAGTCATAGCCGAATAGCCTCTCCACCCAAGCGGCCGGAGAACCT
GCGTGCAATCCATCTTGTTCAATCATGCGAAACGATCCTCATCCTGTCTCTTGATCAGAGCTTGATCC
CCTGCGCCATCAGATCCTTGGCGGCAAGAAAGCCATCCAGTTTACTTTGCAGGGCTTCCCAACCTTA
CCAGAGGGCGCCCCAGCTGGCAATTCCGGTTCGCTTGCTGTCCATAAAACCGCCCAGTCTAGCTATC
GCCATGTAAGCCCACTGCAAGCTACCTGCTTTCTCTTTGCGCTTGCGTTTTCCCTTGTCCAGATAGCC
CAGTAGCTGACATTCATCCGGGGTCAGCACCGTTTCTGCGGACTGGCTTTCTACGTGAAAAGGATCT
AGGTGAAGATCCTTTTTGATAATCTCATGCCTGACATTTATATTCCCCAGAACATCAGGTTAATGGC
GTTTTTGATGTCATTTTCGCGGTGGCTGAGATCAGCCACTTCTTCCCCGATAACGGAGACCGGCACA
CTGGCCATATCGGTGGTCATCATGCGCCAGCTTTCATCCCCGATATGCACCACCGGGTAAAGTTCAC
GGGAGACTTTATCTGACAGCAGACGTGCACTGGCCAGGGGGATCACCATCCGTCGCCCCGGCGTGT
CAATAATATCACTCTGTACATCCACAAACAGACGATAACGGCTCTCTCTTTTATAGGTGTAAACCTT
AAACTGCCGTACGTATAGGCTGCGCAACTGTTGGGAAGGGCGATCGGTGCGGGCCTCTTCGCTATTA
CGCCAGCTGGCGAAAGGGGGATGTGCTGCAAGGCGATTAAGTTGGGTAACGCCAGGGTTTTCCCAG
TCACGACGTTGTAAAACGACGGCCAGTGAATTGTAATACGACTCACTATAGGGCGAATTGGGCCCT
CTAGATGCATGCTCGAGCGGCCGCCAGTGTGATGGATATCTGCAGAATTCAGGCAGCTGAGAAGTT
GAAGGCCGCCGATGCAGGAGGAGGAGAGGTAGACGATGAACTTGAAGAAGGTGAAATTGTTGAAG
AATCAAATGATACGGCGTCGTATTTAGGGAAAGAAATCACAGTGAAGCATCCATTGGAGCATTCAT
GGACTTTTTGGTTTGATAGCCCTATTGCTAAATCTCGACAAACTGCTTGGGGAAGCTCCCTTCGAAA
TGTCTACACTTTCTCCACTGTTGAAGATTTTTGGGGTGCTTACAATAATATCCATCACCCAAGCAAGT
TGGTTATGGGAGCAGACTTTCATTGTTTTAAGCATAAAATTGAGCCAAAGTGGGAAGATCCTGTATG
TACCAATGGAGGGACGTGGAAAATGAGTTTTTCGAAGGGTAAATCTGATACCAGCTGGCTATATAC
GCTGCTGGCTATGATTGGACATCAATTTGATCATGGAGATGAAATTTGTGGAGCAGTCGTTAGTGTC
CGGGCTAAGGGAGAAAAAATAGCTTTGTGGACCAAGAATGCTGCAAATGAAACAGCTCAGGTTAGC

ATTGGTAAGCAATGGAAGCAGTTTCTAGATCATAGCGATTCGGTTGGCTTCATATTTCATGACGATG
CAAAGAGGCTCGACAGAAATGCCAAGAATCGTTACACAGTATAGTTCTTGATGCAATGTGAGAATG
CAAAAAACACAATTCGTACCGGAAAGTTGGTAGGCACTCCTGAATTCCAGCACACTGGCGGCCGTT
ACTAGTGGATCCGAGCTCGGTACCAAGCTTGATGCATAGCTTGAGTATTCTAACGCGTCACCTAAAT
AGCTTGGCGTAATCATGGTCATAGCTGTTTCCTGTGTGAAATTGTTATCCGCTCACAATTCCACACA
ACATACGAGCCGGAAGCATAAAGTGTAAAGCCTGGGGTGCCTAATGAGTGAGCTAACTCACATTAA
TTGCGTTGCGCTCACTGCCCGCTTTCCAGTCGGGAAACCTGTCGTGCCAGCTGCATTAATGAATCGG
CCAACGCGCGGGGAGAGGCGGTTTGCGTATTGGGCGCT