SITE DIRECTED MUTATIONS OF ONE OF THE PLANT PATHOGEN EFFECTOR GENE FOR FUNCTIONAL ANALYSIS

A THESIS SUBMITTED TO
THE GRADUATE SCHOOL OF NATURAL AND APPLIED SCIENCES
OF
MIDDLE EAST TECHNICAL UNIVERSITY

BY

ZEMRAN MUSTAFA

IN PARTIAL FULFILLMENT OF THE REQUIREMENTS
FOR
THE DEGREE OF MASTER OF SCIENCE
IN
BIOCHEMISTRY

FEBRUARY 2015

Approval of the thesis:

SITE DIRECTED MUTATIONS OF ONE OF THE PLANT PATHOGEN EFFECTOR GENE FOR FUNCTIONAL ANALYSIS

Submitted by **ZEMRAN MUSTAFA** in partial fulfillment of the requirements for the degree of **Master of Science in Biochemistry Department, Middle East Technical University** by,

Prof. Dr. Gülbin Dural _____
Dean, **graduate school of Natural and Applied Sciences**

Prof. Dr. Orhan Adalı _____
Head of Department, **Biochemistry**

Prof. Dr. Mahinur S. Akkaya _____
Supervisor, **Chemistry Dept., METU**

Examining Committee Members:

Assoc. Prof. Dr. Çağdaş Devrim Son _____
Biology Dept., METU

Prof. Dr. Mahinur S. Akkaya _____
Chemistry Dept., METU

Assoc. Prof. Dr. Erdoğan Eşref Hakkı _____
Faculty of Agriculture, Selçuk University

Assoc. Prof. Dr. Tülin Yanık _____
Biology Dept., METU

Dr. Banu Avcıoğlu Dündar _____
BIYANS R & D and Training Manager

Date: 20.02.2015

I hereby declare that all information in this document has been obtained and presented in accordance with academic rules and ethical conduct. I also declare that, as required by these rules and conduct, I have fully cited and referenced all materials and results that are not original to this work.

Name, Last Name: Zemran MUSTAFA

Signature :

ABSTRACT

SITE DIRECTED MUTATIONS OF ONE OF THE PLANT PATHOGEN EFFECTOR GENE FOR FUNCTIONAL ANALYSIS

Mustafa, Zemran

M.S., Department of Biochemistry

Supervisor: Prof. Dr. Mahinur S. Akkaya

February 2015, 88 pages

Wheat (*Triticum aestivum* L.) is one of the most important crops planted in the world and used as major food resource. Yet, there are major losses of productivity due to different biotic and abiotic destructive agents. Among the biotic factors especially the yellow rust is one of the most frequent.

Yellow rust or also called "stripe rust", due to stripe like appearances on the leaves, disease is caused by *Puccinia striiformis* f. sp. *tritici,* which is an obligate plant fungal pathogen. To cope with high yield losses caused by these pathogens, investigation of host-pathogen interaction at the molecular level can empower the scientists in fight against the pathogen.

In this thesis research, one of the *Puccinia striiformis* f. sp. *tritici* candidate effector genes (PstHa2a5 / Yin *et al*., 2009) was studied for functional analysis by conducting site directed mutagenesis (SDM) on the gene. The amino acid sequence of this effector

indicates the presence of a conserved FXC motif. The function of this motif is currently unknown but it is suspected that it is needed for intracellular targeting of effector proteins. Mutated versions of the gene were cloned into pK7WGF2 and pK7FWG2 gateway destination vectors to be transformed first into *Agrobacterium tumefaciens* by electroporation and then delivered into *Nicotiana benthamiana* leaves *via* agrobacterium-mediated gene transformation. The distorted motif of PstHa2a5 by SDM was tested for its role in intracellular localization of the protein. By expressing the mutant protein in *N. benthamiana* leave cells and tracking green fluorescent protein (GFP) expressed as a fusion to the modified gene, under microscope functional analysis were conducted.

Keywords: Wheat, *Triticum aestivum* L., yellow rust, *Puccinia striiformis* f. sp. *tritici*, site directed mutagenesis, PstHa2a5, *Nicotiana benthamiana*, *Agrobacterium*, gene transformation, Gateway cloning, GFP, subcellular localization.

ÖZ

İŞLEVSEL ANALİZ İÇİN BİTKİ PATAOJEN EFEKTÖR GENİNDE YÖNLENDİRİLMİŞ NOKTA MUTASYONU

Mustafa, Zemran

Yüksek Lisans, Biyokimya Bölümü

Tez Yöneticisi: Prof. Dr. Mahinur S. Akkaya

Şubat 2015, 88 sayfa

Buğday (*Triticum aestivum* L.) dünyada ekilen en önemli tahıllardan biridir ve ana besin kaynaklarından biri olarak kullanılır. Ancak değişik biyotik ve abiyotik etkenler nedeniyle verimlilikte önemli kayıplar meydana gelmektedir. Biyotik etkenler arasında özellikle sarı pas en sık görülenlerdendir.

Puccinia striiformis f. sp. *tritici*, sarı pas hastalığının nedensel etkeni olarak, bir obligat mantar patojenidir. Bu patojenlerin neden olduğu yüksek verim kayıplarıyla başa çıkmak için, moleküler düzeyde konak-patojen etkileşim araştırmaları, bilim insanlarının patojene karşı savaşta konumlarını kuvvetlendirecektir.

Bu tez araştırmasında, *Puccinia striiformis* f. sp. *tritici*'nin aday efektör genlerinden birinde (PstHa2a5 / Yin *et al*., 2009) işlevsel analiz için yönlendirilmiş nokta mutasyonu

(SDM) yapılmıştır. Bu efektörün amino asit dizisi, korunmuş bir FXC motifinin varlığını göstermektedir. Bu motifin işlevi henüz bilinmemektedir ancak proteinlerin hücre içi hedefleme için gerekli olduğundan şüphelenilmektedir.

Mutasyona uğratılmış gen pK7WGF2 ve pK7FWG2 gateway hedef vektörlerine klonlanıp ilk olarak *Agrobacterium tumefaciens*'a elektroporasyon ile transfer edilip daha sonra *Nicotiana benthamiana* yapraklarına agrobacterium yolu ile gen transferi yapılmıştır. PstHa2a5'in SDM ile değişikliğe uğratılmış motifi hücre içi lokalizasyonundaki rolünü tespiti için *N. benthamiana* yaprak hücrelerinde ifade edilip, mikroskop altında modifiye edilmiş gene birleşik olarak ifade edilen yeşil floresan proteinini (GFP) takip ederek test edilmiştir.

Anathar Kelimeler: Buğday, *Triticum aestivum* L., sarı pas, *Puccinia striiformis* f. sp. *tritici*, yönlendirilmiş nokta mutasyonu, PstHa2a5, *Nicotiana benthamiana*, *Agrobacterium*, gen transferi (transformasyon), Gateway klonlama, GFP, hücre içi lokalizasyon.

To my family,

To myself…

ACKNOWLEDGEMENTS

First of all, I want to express my deepest appreciation and gratitude to my supervisor Prof. Dr. Mahinur S. AKKAYA. Her tireless work, willingness for cooperation and valuable advices, encouraged me to push myself to the limits. Her advices both during the research process as well as in planning my career and have been priceless. I feel very fortunate to have such a helpful supervisor.

I want also to thank all my present and former lab mates: Bayantes DAGVADORJ, Ahmet Çağlar ÖZKETEN, Ayşe ANDAÇ, Kübra NARCI, Adnan YARAMIŞ, Sait ERDOĞAN, Bahtiyar YILMAZ, Serkan TUNA, Dilay KIZIŞAR, Evrim ELÇİN and Taylan BEYAZTAŞ for their friendship and collaboration during my master study, thesis research and other projects.

I would like to thank Bayantes DAGVADORJ also for PstHa2a5 gene constructs.

I am also grateful to my friends: Murat ERDEM, Ahmet Çağlar ÖZKETEN, Ömer Ender AYKANLI, Özgür YALÇIN for their friendship and support.

I would like to show gratefulness to my beloved parents Zana and Fikret MUSTAFA, my darling wife Deniz DÜNDAR MUSTAFA and my precious siblings for their endless support and motivation.

I would like to thank to TÜBİTAK (113Z350), TÜBİTAK (110T445) and METU-BAP for research supports.

TABLE OF CONTENTS

LIST OF TABLES

TABLES

LIST OF FIGURES

FIGURES

LIST OF ABBREVIATIONS

Avr:	Avirulence
bp:	Base pair
cDNA:	Complementary DNA
DNA:	Deoxyribonucleic acid
dNTP:	Deooxy-nucleotidetriphosphate
dpi:	Days post inoculation
ETI:	Effector-triggered immunity
FAO:	Food and Agriculture Organization of the United Nations
Gen:	Gentamicin
GFP:	Green fluorescent protein
HR:	Hypersensitive response
Kan:	Kanamycin
kb:	Kilobase
LB:	Lysogeny broth
LRR:	Leucine rich repeat
M:	Molar

m:	Mili (10^{-3})
µ:	Micro (10^{-6})
g:	Gram
mL:	Mililiter
n:	Nano (10^{-9})
PAMP:	Pathogen-associated molecular patterns
PCD:	Programmed cell death
PCR:	Polymerase chain reaction
pmol:	Picomole
PRR:	Pattern recognition receptor
Pst:	*Puccinia striiformis* f. sp. *tritici*
PTI:	PAMP-triggered immunity
R:	Resistance
Rif:	Rifampicin
RNA:	Ribonucleic acid
ROS:	Reactive oxygen species
SAR:	Systemic acquired resistance
SDM:	Site directed mutagenesis

Spec:	Spectinomycin
T-DNA:	Transfer DNA
Taq:	*Thermus aquaticus*
Ti:	Tumor-inducing
T3SS:	Type-III secretion system
T4SS:	Type-IV secretion system
U:	Unit
Vir:	Virulence

CHAPTER 1

INTRODUCTION

1.1 Wheat and its significance

Wheat is one of the earliest cultivated crops in the world, (around 8000 BC during 'Neolithic Revolution'), and today is still used as main food source for human and livestock in many countries. Genetic studies reveal that wheat has been originated in Near East from wild einkorn wheat, specifically at the outskirt of Karacadağ Mountain, Diyarbakır of today's modern Turkey (Heun, 1997). During domestication of wheat two traits of great importance were selected: 1) Loss of shattering of the spike at maturity and 2) The change from hulled forms to free-threshing naked forms (Shewry, 2009). The range of growing temperature for wheat is between 3 °C to 32 °C with an optimum growing temperature at about 25 °C (Briggle, 1980).

Wheat is grown yearly on around 215 million hectares around the world, an area equivalent to that of Greenland. About one of third of arable lands in Turkey are devoted to wheat production (http://wheatatlas.org/country/TUR) (Table 1.1).

Wheat flour contains about 70 – 75% starch, 12 – 14% water, 8 – 16% proteins, 2 – 3% fibers, 2% lipids and 1% minerals (Scheuer *et al*., 2011). Because of its high starch content which can easily be converted into sugars, it is an attractive option to be used as biofuel in future but because of current world wide hunger problem it is not yet favorable to use it as the source of energy production.

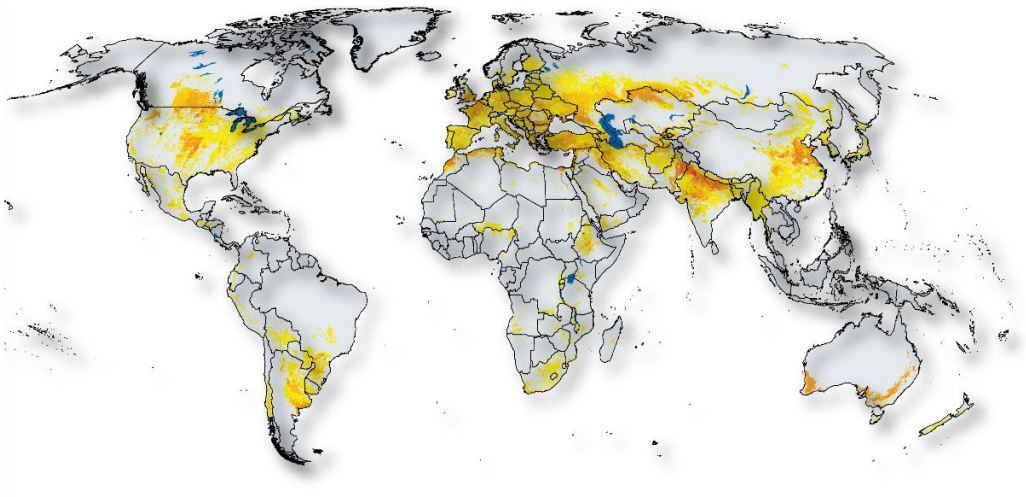

Figure 1.1 Global Wheat Cultivation (You *et al*., 2014). Darker colors indicate areas where more wheat is grown.

Table 1.1 Wheat data for Turkey based on FAO (Food and Agriculture Organization of the United Nations).

Wheat Data for Turkey		
Indicator (Source)	**Value**	**Year**
Total Land (FAO)	769,630 km^2	2012
Arable Land (FAO)	20,577,000 ha	2012
Wheat area (FAO)	7,772,600 ha	2013
Wheat production (FAO)	22,050,000 t	2013
Wheat yield (FAO)	2.84 t/ha	2013
Percentage of total daily calories from wheat (FAO)	36%	2011

1.2 Wheat Yellow (Stripe) Rust

Yellow (stripe) rust is one of the most common and destructive disease of wheat world-wide which can cause 10% to 90% yield loss depending on susceptibility of the wheat line, stage of infection, disease development rate, and the period of infection. The most severe epidemics of this disease occur in temperate-cool and wet climates and they result in severe socio-economic consequences. Yellow rust disease is caused by an obligatory biotrophic fungal pathogen *Puccinia striiformis* f. sp. *tritici* (*Pst*) (Chen, 2005). This type of rust forms yellow to orange powdery structures called urediniospore which contains many single uredia. The disease causes lesions on leaves that are likely to be spread in stripe like structure on entire length of leaves (Markell *et al*., 2000).

Figure 1.2 Stripe rust lesion on wheat.

Germination occurs at humid (more than 50%) and cold conditions, but not in free water. The wheat leaf rust fungus is adapted to a range of different climates, and the disease can be found in diverse wheat growing areas throughout the World. The range of temperature at which yellow rust can grow vary from 0°C to 26 °C with optimum temperatures between 9 °C to 12 °C (Schröder *et al*., 1964). Spreading of the disease is completed by urediniospores which travel with the wind. Viability of urediniospores decrease with exposure to ultraviolet radiation from the sun, however, in suitable conditions they can spread even overseas (Maddison and Manners, 1972).

Even though fungicides can be used to control spreading of the disease, they increase the cost of wheat production and cause environmental pollution. Thus, breeders constantly try to obtain new resistant wheat cultivars against the disease, but new aggressive strains of yellow rust appear constantly causing disease even in resistant cultivars (Chen, 2005; Rapilly, 1979). At this point, understanding of the resistance mechanism of plants and the pathogen mechanism of virulence at the molecular level may allow us to generate better and faster solutions against the disease.

1.3 Plant immunity system

Plants live constantly under thread of bacterial, viral and fungal pathogens. To survive these threads, they have developed a sophisticated defense mechanisms (Morel & Dangl, 1997). For a pathogen to invade plant tissue firstly they have to overcome preinvasive immunity of plant which consists of pre-existing physical and chemical barriers such as the leaf cuticle and tissue architecture like cell wall and remain undetected by suppressing plant immunity system. Prior to pathogen invasion, plant cells may undergo systemic-acquired resistance (SAR) which is characterised by the induction of a number of defence-related proteins, including many PR proteins. SAR is induced by salicylic acid, jasmonic acid and ethylene (Cohn, Sessa, & Martin, 2001). Beside SAR plant immune system

consist of pathogen-associated molecular patterns triggered immunity (PTI) and effector triggered immunity (ETI) (Göhre & Robatzek, 2008).

1.3.1 PAMP triggered immunity (PTI)

Pathogen-associated molecular patterns (PAMP) triggered immunity (PTI) is a primary plant defense mechanism. PAMPs or sometimes called MAMPs (microbe-associated molecular patterns) contain conserved regions and are recognized by plant Pattern Recognition receptors (PRRs). After being recognized by these receptors, PAMPs trigger defense responses. PRRs reside in plasma membrane. PTI also recognizes damage associated molecular patterns (DAMPs) that means it recognizes plants' damaged cellular parts (Figure 1.4). They stimulate signaling cascades involving Ca^{2+} fluxes and mitogen activated protein kinases (MAPKs). These lead to defense reactions such as production of reactive oxygen species (ROS), medium alkalinization, deposition of callose in the cell wall, expression of pathogenesis related (PR) proteins and defensins. Some pathogens inject a range of effectors to suppress defense responses at the level of perception by PRRs, signaling, or defense action, making host susceptible to colonization (Göhre & Robatzek, 2008).

1.3.2 Effector triggered immunity (ETI)

Effectors are products of the respective avirulance (*Avr*) genes of pathogens secreted to suppress PTI of plants and are vital for their virulence. Effectors also intermediate penetration of pathogens into host tissues, access to nutrients, and multiplication. Resistant plant varieties have developed R genes which identify effector molecules or their actions.

R-genes are characterized having the nucleotide binding domain (NB) and the leucine rich repeat (LRR) domain(s) and are often called NB-LRR proteins (Figure 1.4)

After recognition, R genes initiate hypersensitive response (HR) which lead to programed cell death (PCD) which comprise ETI effector triggered immunity. The cell death response likely benefits the plant as whole by depriving pathogens from access to further nutrient sources and limiting pathogen multiplication (Cohn, Sessa, & Martin, 2001). R-gene products (receptors) are localized in plasma membrane and in intracellular matrix and are race specific. This type of action is called gene-for-gene resistance as specific R proteins account for recognition of specific effector molecules. It is an evolutionary race of pathogen virulence and plant defense systems (Göhre & Robatzek, 2008). At least two steps are necessary to induce the HR: recognition of the pathogen and transduction of the perceived signals to the programed cell death pathway (Morel & Dangl, 1997).

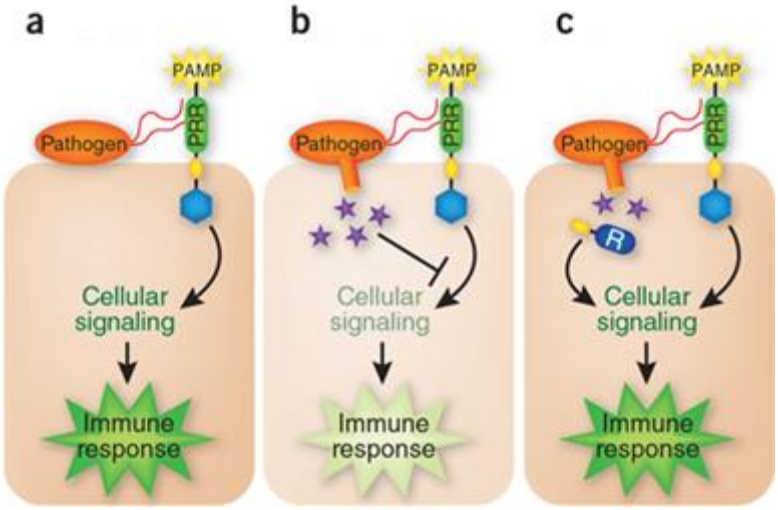

Figure 1.3 Simplified representation of the plant immune system (Pieterse *et al*., 2009). **a)** PAMP-triggered immunity; **b)** Effector-triggered susceptibility; **c)** Effector-triggered immunity.

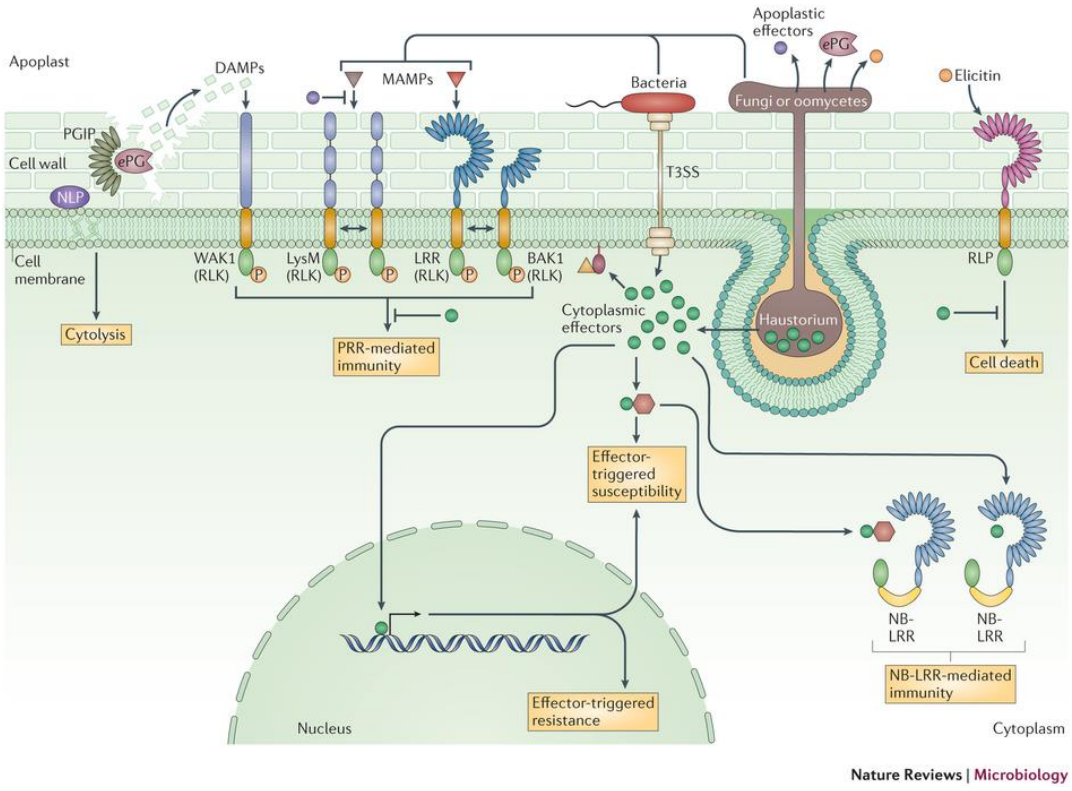

Figure 1.4 Schematic representation of plant PTI and ETI immune systems (Wirthmueller *et al*., 2013).

1.4 Site directed mutagenesis (SDM)

There are three ways available for site directed mutagenesis to be complete: cassette mutagenesis, primer extension and procedures based on polymerase chain reaction (PCR). Cassette mutagenesis is based on synthetic DNA fragment containing the desired mutant sequence which is then ligated to desired fragment. This method is limited to sites where suitable restriction sites on both flanking region of DNA fragment are available.

Primer extension method is easy to be done. It requires only one primer with mismatch, insertion or deletion with regard to the single stranded plasmid (Figure 1.5). The disadvantage of this method is the requirement to obtain single stranded plasmid and the need to separate the final product which contains both mutant and wild type plasmids.

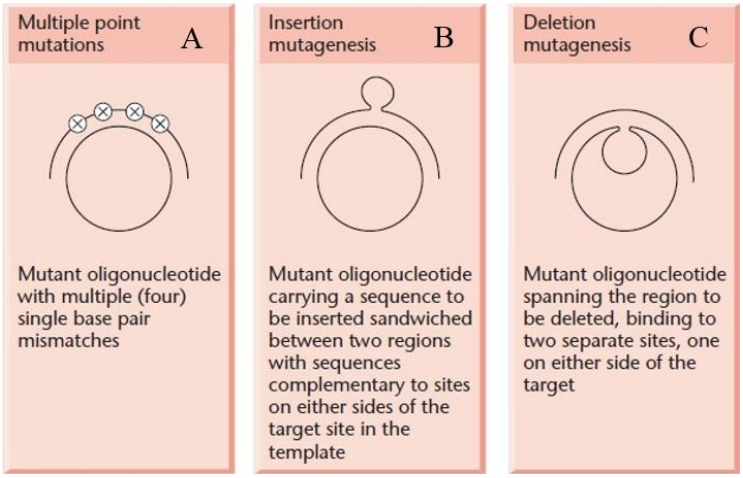

Figure 1.5 Site directed mutagenesis based on primer extension method. **A)** Multiple point mutation; **B)** Insertion mutation; and **C)** Deletion mutation procedures (Primorse *et al*., 2001).

PCR based SDM is the most convenient method for site directed mutagenesis. It contains three different PCRs in two steps process. Four different primers, two of which spans the site of mutation, are required for a single mutation of the DNA fragment (Figure 1.6).

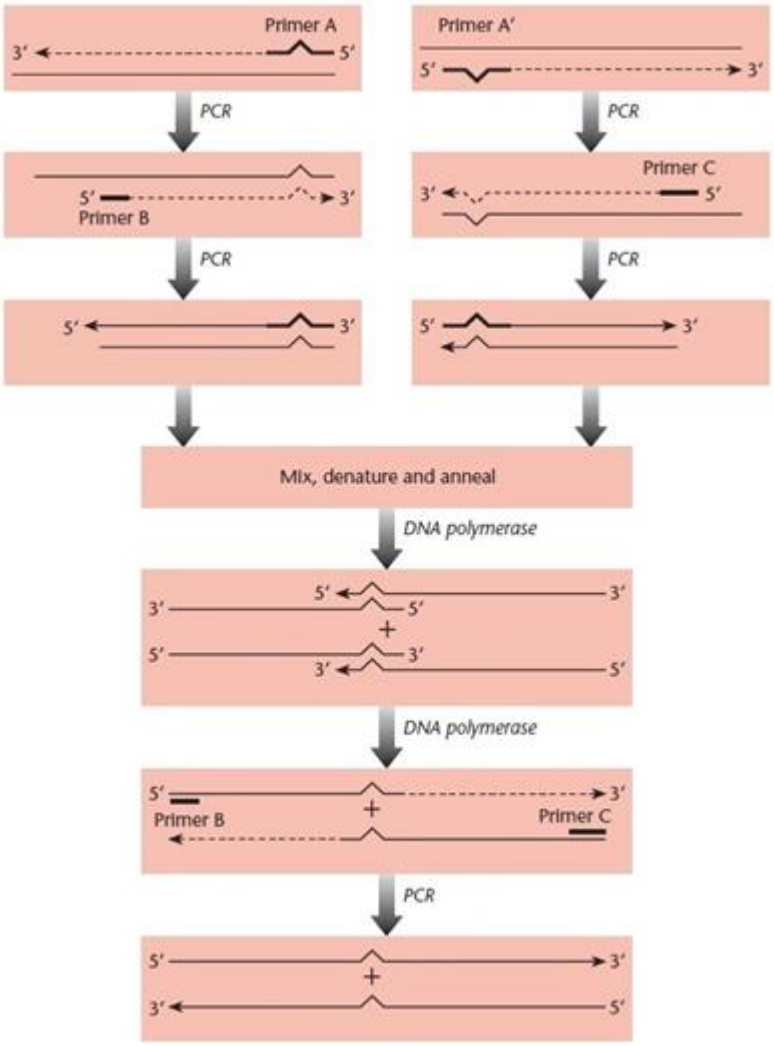

Figure 1.6 PCR based method for site directed mutagenesis. In the first step two different PCRs are performed. Four different primers are needed; the B and C primers defining edges of DNA fragment, A and A' primers both spanning site of mutation generation. In first step two PCRs with two different products coinciding in the place of mutation are generated. The second step is a joining fragment PCR where products from step one are joined to form a whole sequence with desired mutation in desired place.

1.5 Agrobacterium-mediated transformation

Agrobacterium tumefaciens is a gram negative bacterial pathogen with a great range of plant hosts (mainly dicotyledonous), where it causes crown gall disease. It causes tumor formation. *Agrobacterium* contains a plasmid named Ti-plasmid which is vital for its virulence. It delivers only T-strand (T-DNA) located inside Ti-plasmid, with guidance of virulence (vir) genes also found in Ti-plasmid, to the plant cell. 30 kb of virulence (vir) region is organized in six operons that are essential for the T-DNA transfer (Opabode, 2006). The T strand is inserted by Type Four Secretion System (T4SS) inside plant cell where it is inserted into the plant cell chromosome and become transcriptionally active. T strand products induce tumor growth on the plant trunk with transformed plant cells producing nutrition to be exploited by *Agrobacterium*. It is a magnificent model and the only known example in nature where pathogen transfer DNA fragment for its virulence, to organisms of other kingdom (Pitzschke & Hirt, 2010). After its discovery, scientists modified the Ti-plasmid to use it as a successful way for transformation of plant cell by replacing T-DNA region with desired DNA fragment. It can also be used for transformation of fungi and even some mammalian cells. It allowed construction of the first vector from bacterial kingdom for plant transformation. Usually Agrobacterium-mediated expression of a transgene achieves the highest level 2 – 3 days following argoinfiltration, after which the expression level decreases rapidly (Wydro *et al*., 2006; Pitzschke *et al*., 2010). It is an easy and cheap method which is used for many researches in plants. However, it has some restrictions as its efficiency is low, and the part of plant genome to which T strand attaches cannot be determined definitely.

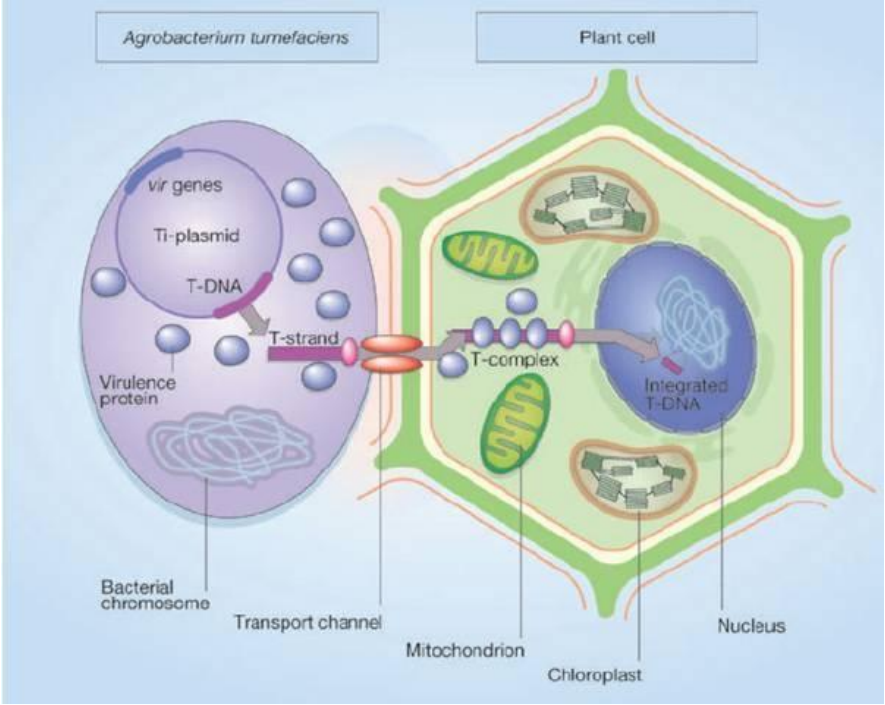

Figure 1.7 Agrobacterium-mediated transformation (Gelvin, 2005).

1.6 Gateway cloning

For genomic functional studies cloning of the gene of interest into desired plasmids for different uses is crucial. Traditionally, gene cloning has relied on restriction enzyme digestion and ligation. On the other hand, Gateway cloning, developed by Invitrogen Co. has advantages as it utilizes λ bacteriophage site-specific recombination strategy without necessity of using restriction enzymes (Xu & Li, 2008). In general, the Gateway cloning technology involves a two-step process. First step is to clone the gene of interest into donor vector (pENTR/D-TOPO) through so-called BP reaction that consists of mix containing the phage integrase, topoisomerase, and the integration host factor. For a DNA fragment to be inserted into donor vector it is flanked at 5' end by CACC sequence which

is recognized and joined to pENTR vector. DNA fragment inserted into pENTR is flanked by two *att*L regions. The pENTR vector has *ccdB* gene in site of insertion, which is disturbed as the gene is inserted. If not disturbed, *ccdB* is lethal gene for *Escheria coli*, to which the plasmid is transformed. This way non-recombinant plasmid carrying *E. coli* bacteria are eliminated. Moreover, the donor vector contain antibiotic resistance gene for positive selection of transformed *E. coli*. The purpose of using donor vector is that the gene of interest can easily be transformed form this vector to wide variety of vectors named as destination vectors (pDEST), which are more specialized vectors that incorporate particular functional properties to the cloned DNA fragment (Hartley *et al*., 2000). Subcloning of our DNA fragment from donor vector to destination vector is performed with so-called LR reaction which consist of a mix containing integrase, integration host factor, and the phage excisionase. pDEST has two attR sites, flanked to the *ccdB* gene. In LR reaction *att*L1 and *att*L2 of the pENTR vector react with *att*R1 and *att*R2 on the pDEST respectively, to recombine the DNA of interest with *ccdB* gene in correct direction (Karimi *et al*., 2007).

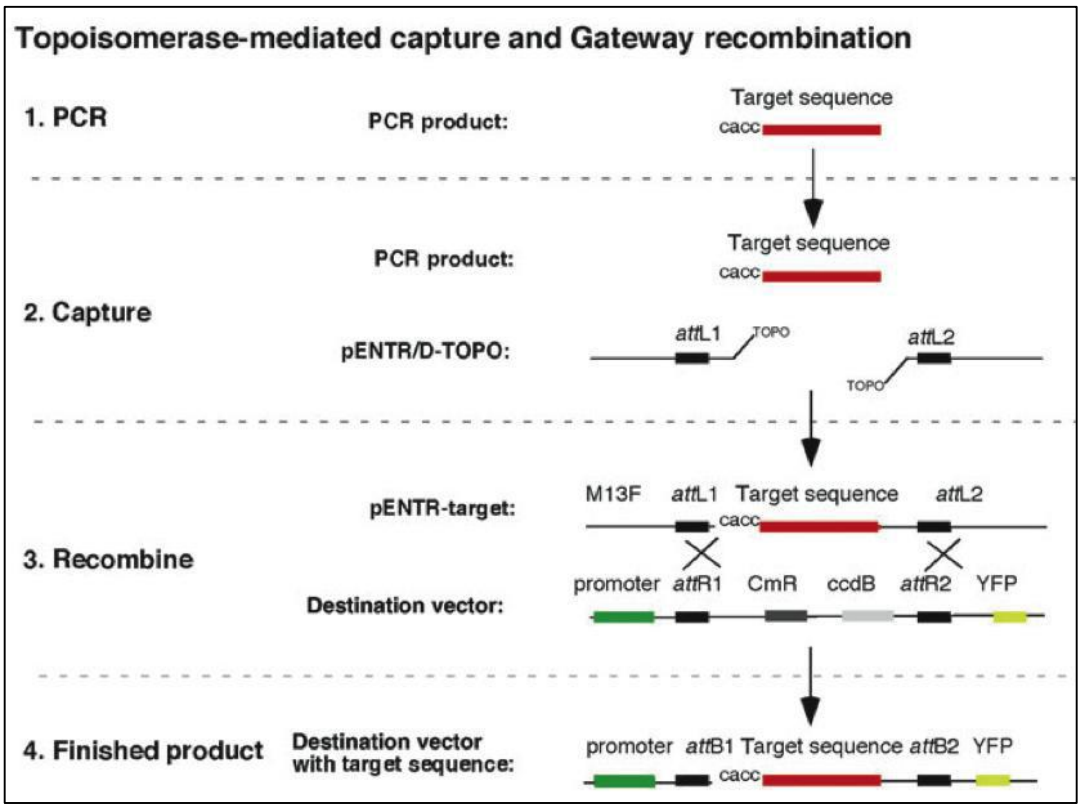

Figure 1.8 Overview of the process for cloning a DNA fragment into a desired destination vector by gateway cloning technology: 1) After DNA containing fragment to be amplified is isolated, PCR is done by joining a CACC flanking sequence *via* properly designed forward primer. 2) The amplified target sequence is captured into pENTR/D-TOPO by recognition of CACC sequence in 5'end of the targeted sequence. The vector has *att*L regions on both sides of insertion. 3) Target sequence is subcloned into destination vector which has *att*R regions on both sites of *ccdB* gene. *att*R regions recombine with *att*L regions of pENTR/D-TOPO vector. 4) After two-step procedure the target sequence is inserted in specific destination vector (Early *et al*., 2006).

1.6.1 pK7WGF2 and pK7FWG2 vector

pK7WGF2 and pK7FWG2 vectors are binary T-DNA vectors used as destination vectors in Gateway cloning technology and are readily transformed into *Agrobacterium tumefaciens* for agrobacterium-mediated gene transfer. Target sequence translocate between *att*L sites of pENTR/D-TOPO vector to *att*R sites of pK7WGF2 or pK7FWG2 vectors. *ccdB* gene is located beween *att*R sites of pK7WGF2 and pK7FWG2 vectors. The *ccdB* gene is needed for negative selection of non-recombinant plasmid carrying *E. coli*. pK7WGF2 and pK7FWG2 vectors have spectinomycin (Sm) and streptomycin (Sp) resistances in bacterium, and kanamycin (Kan) resistance in plant (Karimi *et al*., 2002). The map of both vectors are represented in Figure 1.9.

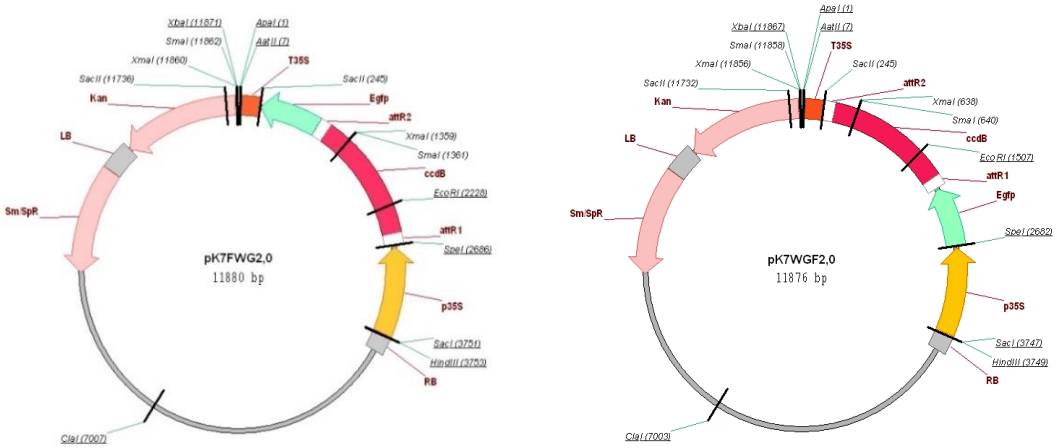

Figure 1.9 Vector map of pK7WGF2 and pK7FWG2 vector map (Karimi *et al*., 2002). Note that Egfp (Enhanced green fluorescent protein) is located at 5'end of DNA insert in pK7WGF2 **(a)**; and in 3'end of DNA insert in pK7FWG2 **(b)**.

1.7 PstHa2a5 gene

Genomic studies of *Puccinia striiformis* f. sp. *tritici* reveal genome composition and its genes. One of these genes is PstHa2a5 gene. It was found that PstHa2a5 gene expression increases slightly in germinated uredinospores and up to 120 times in *Pst*-infected infected leaves, comparing to uredinospores that have not yet invaded any plant cell (Yin *et al*., 2009). This give us a clue that PstHa2a5 gene must be an important gene for the pathogen virulence. Its sequence resembles an effector molecule as it has the FXC motif, one of YXC, FXC or WXC (X can be any amino acid) motifs found in effector molecules. (Godfrey *et al*., 2010). These motifs are found in the first 27 amino acids. The PstHa2a5 gene contains also N-terminal signal peptide and is considered a putative secreted protein. The presence of FKC motif 24 amino acids away from beginning and highly probable cleavage site supports even more the hypothesis of being an effector molecule (Yin *et al*., 2009).

PstHa2a5 gene is a newly found gene whose function and interactions are still unknown and it is in focus of researchers.

Table 1.2 Predicted secreted proteins from cDNA library of *Pst* haustoria (Yin *et al*., 2009b). Prediction completed *via* signal P 3.0 algorithm and iPSORT.

Unisequence	GenBank accession	Size (aa)	No. of Cysteine residues	Homology in databases	E value
PSTha2a5	GH737102	117	6	predicted protein of *Puccinia graminis*	5.23E-23
PSTha9F18	GH737274	259	14	hypothetical protein of *Puccinia graminis*	8.96E-41
PSTha12a4	GH737444	289	10	predicted protein of *Puccinia graminis*	4.00E-23
PSTha12j12	GH737467	133	6	no homology	-
PSTha15N21	GH737567	98	7	no homology	-
PSTha21O8	GH737139	92	1	putative sulfate transporter	1.42E-38
PSTha5a23	GH737046	108	4	no homology	-
PSTha6i16	GH737950	73	5	predicted protein of *Puccinia graminis*	3.00E-10
PSTha8F13	GH738007	116	6	no homology	-
PSTha2c7	GH737231	204	13	predicted protein of *Puccinia graminis*	1.15E-15
PSTha16B3	GH737129	87	2	no homology	-
PSTha10F24	GH737323	56	3	no homology	-
PSTha16D6	GH737598	66	3	hypothetical protein of *Aspergillus niger*	2.00E-16
PSTha9C13	GH738022	65	1	predicted protein of *Puccinia graminis*	6.00E-10
PSTha12h2	GH737173	70	5	no homology	-

1.8 Aim of the study

The objective of this thesis study was to investigate the role of a conserved motif present in PstHa2a5, whether it is involved in the subcellular localization in the plant cell upon secretion from the haustorium. As in many other plant pathogens, in some rust pathogens too, (W/Y/F)XC motif which is assumed to play a role in localization of the effector. Wheat Yellow Rust disease agent, *Puccinia striiformis* f. sp. *tritici* (*Pst*) predicted effector gene, PstHa2a5 (Yin *et al*., 2009) also has "FKC" conserved motif, which is positioned adjacent to the signal peptide putative cleavage site. The uniqueness of FXC motif of this effector is its position to the Signal Peptide sequence, and it happens to start at the possible cleavage site of the SP. However, the (W/Y/F)XC conserved motif usually present further downstream of the SP of those candidate effectors possessing it. The obvious distinction of Cysteine in the motif is that it may form disulphide bond with another cysteine residue present in this effector. Since PstHa2a5 has 8 cysteine residues in its sequence, any of the remaining ones might be involved in disulphide bond formation contributing in the stability of the protein, which is considered one of the critical features in assigning the effector candidates, especially for those effectors targeting the plant apoplastic region. To test the role of FXC motif in subcellular location, the substitution mutations were decided on the cysteine residue of FXC motif. For substitution, Serine was selected due to the similarities in the electron density the chemistry of –OH and –SH groups: although hydroxyl group of Serine cannot form a covalent bond, it can involve in Hydrogen bonding. Another selected substitution was to Valine; since it is an inert amino acid with closest electron density to Cysteine.

CHAPTER 2

MATERIALS AND METHODS

2.1 *Nicotiana benthamiana* growth and maintenance

Nicotiana benthamiana seeds were planted in a small pot semi-covered with stretch film to sustain the humidity. The pot was placed in plant growth room which was set for 16 hour light and 8 hour dark and 24 ^{o}C temperature. After 10 days the grown sprouts were transferred into bigger pots (650 cm^3 volume) as one sprout per pot. *N. benthamiana* leaves of 4 – 6 week old were used for agro-infiltration. Every six months two subcultured *N. benthamiana* plants were kept for more than 6 weeks to use as seed source for later plantings. Seeds were matured and gathered around 8 – 10 weeks after planting.

2.2 PstHa2a5 gene construct

PstHa2a5 gene (GenBank accession number: GH737102.1) sequence was retrieved from Expressed Sequence Tag (EST) records of National Center of Biotechnology Information (NCBI). 'SignalP 4.0 Server' tool was used for signal peptide prediction of the PstHa2a5 gene product (http://www.cbs.dtu.dk/services/SignalP/).

The sequence was sent for synthesis contained 5' FLAG-tag for immunoprecipitation, restriction enzyme sites *Pac*I at 5' and *Not*I at 3' for directional cloning (Figure 2.1). The gene construct was synthesized by GeneScript and sent cloned in pBluescript II SK vector

using *EcoR*I site (Figure 2.2). This construct was used for amplification of PstHa2a5 gene for cloning in Gateway pENTR vector.

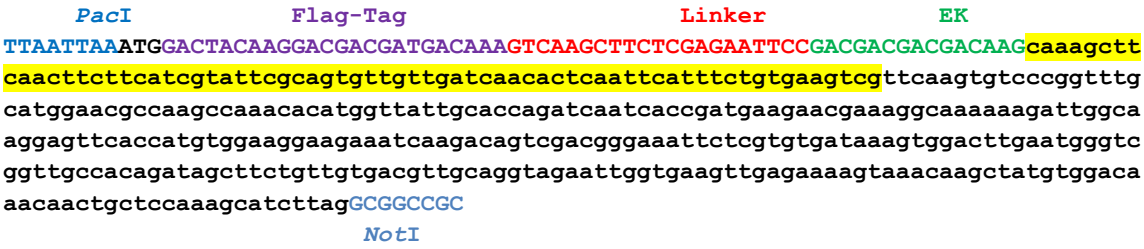

Figure 2.1 PstHa2a5 gene construct cloned in pBlueScript II SK. *Pac*I/*Not*I restriction sites (blue), flag-tag (purple), linker sequence (red), enterokinase sequence (green). PstHa2a5 gene's own signal peptide is highlighted in yellow.

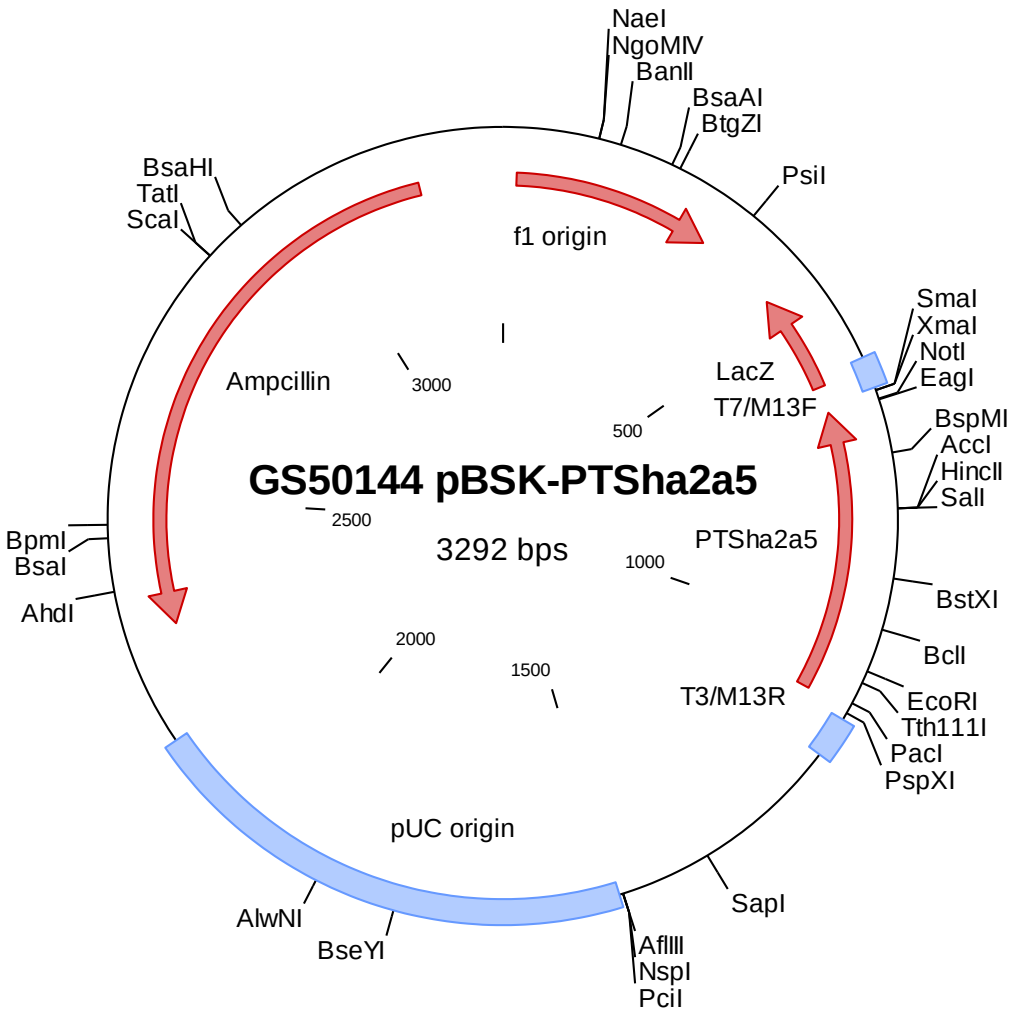

Figure 2.2 pBSK-PstHa2a5 plasmid map. Synthetic PstHa2a5 gene was cloned into *Eco*RV restriction enzyme digested pBluescript II SK plasmid derivative.

2.2.1 PstHa2a5 gene and primers

Amplification of PstHa2a5 gene with 5'CACC flanking region for Gateway cloning was carried out with PstHa2a5-CACC-fwd and PstHa2a5-rev-with/without-stop primers. As

DNA template, PstHa2a5 gene construct cloned in pBlueScript II SK was used. PstHa2a5-C27S-fwd and PstHa2a5-C27S-rev primers were used for generating point mutation of 27th amino acid of PstHa2a5 gene from "Cysteine (C)" to "Serine (S)". PstHa2a5-C27V-fwd and PstHa2a5-C27V-rev primers were used for generating point mutation of 27th amino acid of PstHa2a5 gene from "Cysteine (C)" to "Valine (V)". For generation of the fragments with point mutations, PstHa2a5-rev-with-stop primer was used for gene construct where GFP was flanked on 5' end of the gene whereas PstHa2a5-rev-without-stop primer was used for gene construct where GFP was flanked on 3' of the gene. Flanked GFP region had its own stop codon.

Table 2.1 PstHa2a5 amplification and point mutation primers.

Primer names	**Sequences (5' to 3')**
PstHa2a5-CACC-fwd	CCACATGCAAAGCTTCAACTTCTTC
PstHa2a5-C27S-fwd	CTGTGAAGTCGTTCAAGTCTCCCG
PstHa2a5-C27S-rev	TCCATGCAAACCGGGAGACTTGAA
PstHa2a5-C27V-fwd	TCTGTGAAGTCGTTCAAGGTTCCCG
PstHa2a5-C27V-rev	TCCATGCAAACCGGGAACCTTGAA
PstHa2a5-rev-with-stop	CTAAGATGCTTTGGAGCAGTTGTTTGT
PstHa2a5-rev-without-stop	AGATGCTTTGGAGCAGTTGTTTGT

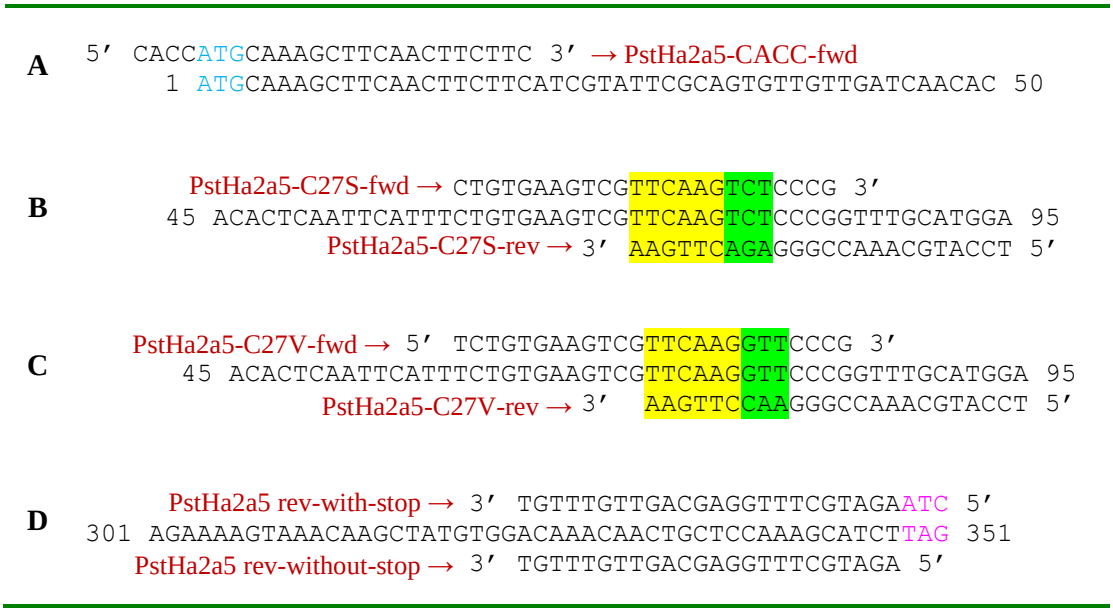

Figure 2.3 Locations of primers bound to PstHa2a5 gene: **A)** PstHa2a5-CACC-fwd (above). ATG in cyan marks beginning of gene with methionine; **B)** PstHa2a5-C27S-fwd (above), PstHa2a5-C27S-rev (below); **C)** PstHa2a5-C27V-fwd (above), PstHa2a5-C27V-rev (below). In (**B**) and (**C**) yellow highlighted sequence marks "F" and "K" of the "FKC" motif, while green highlighted sequence marks mutated sequence of "FKC" to "S" in (**B**) and to "V" in (**C**); **D)** PstHa2a5-rev-with-stop (above), PstHa2a5-rev-without-stop (below). Sequences in pink marks the stop codon.

2.3 Polymerase chain reaction (PCR)

The gene of interest was amplified using PCR. PCR was performed using Q5 High Fidelity *Taq* DNA Polymerase (NEB, Cat# M0491S, Lot# 0041209) with the appropriate primer pairs following the manufacturer's protocol as represented in Table 2.2.

Table 2.2 Components and their amounts used in PCR.

PCR Components in 200 µL tube	Amount	Final concentration
5X Q5 Reaction Buffer	5 µL	1X
5X Q5 High GC Enhancer	5 µL	1X
10 mM dNTPs	0.5 µL	200 µM
10 µM Forward Primer (10pmol/ µL)	1.25 µL	0.5 µM
10 µM Reverse Primer (10pmol/ µL)	1.25 µL	0.5 µM
Q5 High Fidelity DNA Polymerase (2000U/mL)	0.5 µL	1 U
ddH_2O	variable	
Template DNA	variable; 20 – 150 ng	0.8 – 6 ng/µL
	Total volume: 25 µL	

Table 2.3 Conditions in thermocycler for PCR using Q5 high fidelity DNA polymerase.

Steps	Temperature ºC	Duration	Repetition (cycles)
Initial Denaturation	98	1 min	1
Denaturation	98	10 sec	35
Annealing	55	30 sec	
Extension	72	20 sec	
Final Extension	72	2 min	1

2.3.1 Colony PCR

Colony PCR was performed to verify the presence of gene in the colonies of *Escherichia coli* or *Agrobacterium* GV3101. Since the product of colony PCR was used only for verification not for further usage *Taq* DNA Polymerase with Standard *Taq* Buffer (New England Biolabs, Cat# M0320S, Lot# 0141203) with the appropriate primer pairs was preferred. In colony PCR instead of DNA template, single-colony cells were placed directly in the PCR mixture. Colony PCR was performed following the manufacturer's protocol as represented at Table 2.4.

Table 2.4 Components and their amounts used in colony PCR.

PCR Components in 200 µL tube	Amount	Final concentration
10X Standard *Taq* Reaction Buffer	2.5 µL	1X
$MgCl_2$ (25mM)	1.5 µL	1.5 mM
dNTPs (10mM)	0.5 µL	200 µM
10 µM Forward Primer (10pmol/ µL)	0.5 µL	0.2 µM
10 µM Reverse Primer (10pmol/ µL)	0.5 µL	0.2 µM
Taq Polymerase (5000 U/mL)	0.2 µL	1 U
ddH_2O	variable	
Colony from liquid media	1 µL at OD_{600} = 1.0	
	Total volume: 25 µL	

Table 2.5 Conditions in thermocycler for colony PCR using *Taq* DNA polymerase.

Steps	Temperature °C	Duration	Repetition (cycles)
Initial Denaturation	95	3 min	1
Denaturation	95	45 sec	35
Annealing	55	45 sec	
Extension	68	45 sec	
Final Extension	68	5 min	1

2.3.2 Fragment joining PCR

Fragment joining PCR was used for joining two DNA fragments which overlaps head to tail to create a bigger fragment consisting both of them. One fragment's 3'end region was identic to other fragment's 5' end region (more than 12 nucleotides should overlap). For this purpose Q5 High Fidelity *Taq* DNA Polymerase was used. Forward primer of the 3' end region overlapping fragment and reverse primer of 5'end overlapping region fragment was used. Protocol described in Table 2.2 was used with one extra DNA template.

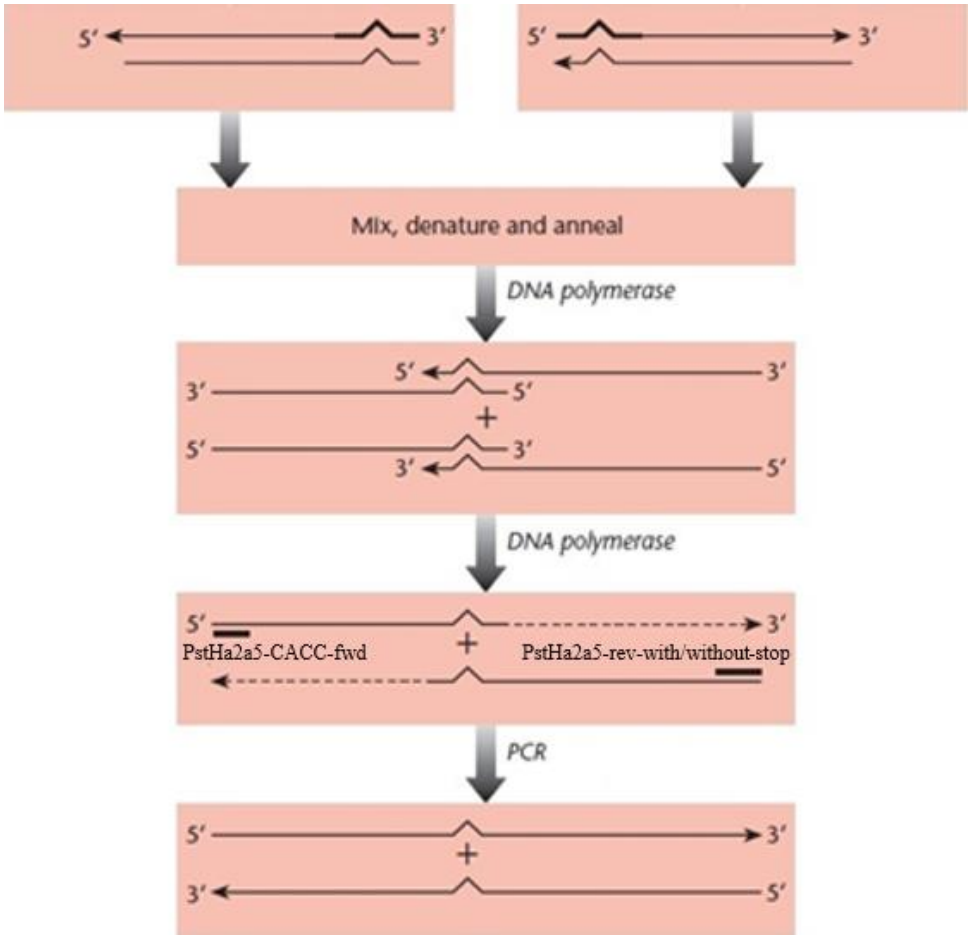

Figure 2.4 Fragment joining PCR procedure for generation of complete fragment with desired point mutation.

2.4 PCR product purification

For purification of PCR products QIAquick PCR purification kit was used. First, PB buffer was poured into PCR tube in 5 : 1 ratio (PCR mixture was transferred to another tube before adding PB buffer if total volume exceeded 200 µL, which is the volume of PCR tube). The mixture was mixed by vortexing. Quick mini spin was carried out to

collect the liquid to the bottom of tube. The mixture was transferred to QIAquick spin column which was placed inside sterile 2 mL collection tube. The tube including column was centrifugated at 13,500 rpm for 1 min. The flow-thorough was discarded. At this stage DNA was stuck in column membrane. The column was washed by adding 0.75 mL of PE Buffer and was kept at room temperature for 5 – 7 minutes for salts to dissolve salts in the buffer. Then, it was centrifuged for 1 minute at 13,500 rpm. Flow-through was discarded and a 1 minute centrifugation at 13,500 rpm was carried out again to remove any remaining of washing buffer. At final stage column was placed into clean 1.5 mL collection tube, and 30 – 50 µL Buffer EB (elution buffer; 10 mM Tris·Cl, pH 8.5) was added. It was kept for 5 minutes at room temperature and then centrifuged at 13,500 rpm for 2 minutes. The collected purified PCR product was measured in NanoDrop (ND-1000 spectrophotometer), recorded, and kept in -20 ^{o}C for later usage.

2.5 Gateway cloning

2.5.1 pENTR/D-TOPO cloning

For gateway cloning procedure first step was to ligate the gene of interest into donor vector (pENTR^{TM}/D-TOPO) by topoisomerase-mediated cloning. PstHa2a5 gene flanked with CACC sequence at 5' end was amplified using PstHa2a5-CACC-fwd and PstHa2a5-rev-with-stop primer pair for insertion into pK7WGF2 destination vector and PstHa2a5-CACC-fwd and PstHa2a5-rev-without-stop primer pairs for insertion into pK7FWG2 destination vector. PCR conditions were the same as mentioned in Section 2.3. Invitrogen pENTR^{TM} Directional TOPO^{R} cloning kit and its protocol with slight changes were used for the cloning reaction. First of all, 0.5 µL of purified PCR product (30 – 150 ng/µL) was mixed with 0.5 µL salt solution, 0.25 µL pENTR/D-TOPO solution and 1.75 µL ddH_2O in 200 µL PCR tube. The PCR tubes were put in thermocycler to be kept at 23 ^{o}C for 30 minutes. After that, tubes were kept at 4 ^{o}C overnight to increase ligation efficiency. The

next day mixture was ready to be transformed in competent *E. coli* TOP10 cells. pENTR/D-TOPO vector contains kanamycin resistance gene, thus the transformed cells were selected in LB agar containing kanamycin (50 μg/mL) after growing for 16 – 18 hours at 37 ^{o}C.

2.5.2 LR clonase reaction

As second step in gateway cloning procedure subcloning of the gene of interest to destination vector (pK7WGF2 or pK7FWG2) *via* LR reaction was conducted. Invitrogen Gateway^{R} LR Clonase^{TM} II Enzyme Mix kit and its protocol with slight changes was used for the reaction. Firstly, transformed *E. coli* TOP10 cells containing PstHa2a5 gene cloned into pENTR/D-TOPO vector were grown overnight in LB containing kanamycin (50μg/mL). Also *E. coli* TOP10 cells containing destination vectors (pK7WGF2 and pK7FWG2) were grown overnight in LB containing spectinomycin (100 μg/mL). Plasmid isolation was performed from overnight grown *E. coli* TOP10 cells according to the procedure described in Section 2.6.

To perform LR reaction 0.5 μL (around 200 ng/ μL) of pENTR/D-TOPO vector containing PstHa2a5 gene was mixed with 0.5 μL (around 200 ng/ μL) of destination vector in 200 μL PCR tube. 2.5 μL TE buffer, 0.5 LR clonase and 1 μL ddH_2O was added into PCR tubes and kept at 25 ^{o}C for 2 hours at thermocycler. After 2 hours 0.5 μL proteinase K was added to stop the LR reaction. At this point the mixture was ready to be transformed into competent *E. coli* TOP10 cells. pK7WGF2 and pK7FWG2 destination vectors contain spectinomycin resistance gene, thus the transformed cells were selected in LB agar containing spectinomycin (100 μg/mL). After 16 – 18 hours single colony selection with tip of the loop was performed from LB agar plates and inoculated into LB containing spectinomycin (100 μg/mL) and kept at 37 ^{o}C for 16 – 18 hours. Plasmid isolation from grown *E. coli* TOP10 cells containing gene of interest in destination vector was performed.

The isolated plasmids were used to transform electrocompetent *Agrobacterium* GV3101 (pMP90) cells.

2.6 Preparation of *E. coli* TOP10 competent cells

Lysogeny broth agar (LB; 5 g tryptone, 2.5 g yeast extract, 5 g NaCl, 7.5 g agar, 1.6 mL NaOH (0.5 M) dissolved in 0.5 L ddH_2O and autoclaved) was prepared and poured into plates. Streak plate from *E. coli* TOP10 master stock was prepared and incubated at 37 ^oC overnight. Single colony was selected with the tip of a loop, inoculated in 4 mL LB medium and was let grow overnight in shaker at 37 ^oC at 250 rpm. 1 mL of LB containing overnight grown *E. coli* TOP10 cells was poured into 100 mL sterile LB medium and let grow at 37 ^oC for 1.5 – 2 hours until absorbance (A_{600nm}) reached to 0.6 (measured *via* spectrophotometer, Shimadzu, UV-1601). Then the solution was divided into two 50 mL falcons and centrifuged at 4 ^oC, at 5000 rpm for 5 minutes. The supernatant was discarded while cells were stack at the bottom as pellet. 10 mL of cold (4 ^oC) and filtered (MN sterilizer, 0.22 μm cellulose acetate filter) 100 mM $CaCl_2$ was added to both of 50 mL falcons. The cells were suspended by gently shaking the mixture until no pellet remains at the bottom and kept for 10 minutes in ice. Then, the mixture was centrifuged again at 4 ^oC, at 5000 rpm for 5 minutes. The supernatant was discarded while cells were stack at the bottom as pellet. For the second time, 10 mL of cold and filtered 100 mM $CaCl_2$ was added to both of 50 mL falcons and centrifuged at 4 ^oC, at 5000 rpm for 5 minutes. Supernatant was discarded again. Cells were re-suspended in 2 mL of cold and filtered 100 mM $CaCl_2$, and stored at 4 ^oC. These cells could be used for up to one week for transformation.

2.7 Plasmid isolation

For plasmid isolation Qiagen Plasmid Miniprep kit (Cat. No: 27106, Lot. No: 124104532) was used. Isolation was conducted according to the manufacturer's protocol. Overnight grown cells were collected by centrifugation at 8,000 rpm for 2 minutes in 2 mL tubes. Supernatant was discarded. Cells stuck as pellet were suspended in 250 µL Buffer P1 (RNase A added) by vortexing the solution. Then, 250 µL Buffer P2 was added and gently inverted 4 – 6 times until the solution turns completely blue. Following, 350 µL Buffer N3 was added and gently inverted 4 – 6 times until the blue color turns completely white. The tubes were centrifuged for 10 minutes at 13,500 rpm in a microcentrifuge (CLP Model 3410 Microcentrifuge). A white pellet and transparent supernatant was formed. Supernatant was transferred into QIAprep spin column. At this step the supernatant was pipetted carefully not to take any part of white pellet. QIAquick spin column which was placed inside sterile 2 mL collection tube. The tube including column was centrifuged at 13,500 rpm for 1 min. The flow-thorough was discarded. At this stage plasmid DNA was stuck in column membrane. The column was washed by adding 0.75 mL of PE Buffer and was kept at room temperature for 5 – 7 minutes for salts to dissolve salts in the buffer. Then, it was centrifuged for 1 minute at 13,500 rpm. Flow-through was discarded and a 1 minute centrifugation at 13,500 rpm was carried out again to remove any remaining of washing buffer. At final stage column was placed into clean 1.5 mL collection tube, and 30 – 50 µL Buffer EB was added. It was kept for 5 minutes at room temperature and then centrifuged at 13,500 rpm for 2 minutes. The collected purified plasmid DNA was measured in NanoDrop, recorded, and kept in -20 ^{o}C for later usage.

2.7.1 Transformation of plasmid into TOP10 competent cells

The recombined plasmid to be transformed was obtained *via* isolation or ligation procedures. 3 µL of the recombined plasmid was mixed with 100 µL of *E. coli* TOP10

competent cell mixture in 1.5 mL tube and kept for 10 minutes in ice. Heat shock was applied by placing the tubes into 42 ^{o}C water bath for 45 seconds. After heat shock tubes were placed immediately into ice again and kept for 5 minutes. The 500 µL LB medium was added into tube and let grow at 37 ^{o}C at 200 rpm for 2 hours. The mixture was spread at different amounts in LB agar plates containing the apropriate antibiotic to which the transformed plasmid contained the resistance gene. pENTR/D-TOPO vector contains kanamycin resistance gene, whereas both pK7WGF2 and pK7FWG2 vectors contain spectinomycin resistance gene. Thus, only the transformed cells could survive. The spread LB agar plates were kept at 37 ^{o}C in growth chamber overnight. The transformed cells formed single colonies were picked up with the tip of loop and inoculated in 4mL LB containing the same antibiotic.

2.8 Agrobacterium-mediated gene transformation

2.8.1 *Agrobacterium* GV3101 (pMP90) electro-competent cell preparation

LB-tetracycline (Tet) (2.5 µg/mL) agar was prepared and poured into plates. Streak plate from *Agrobacterium* GV3101 (pMP90) cells (obtained from Dr. Csaba Koncz, Max-Planck Institute Plant Breeding Research Department, Cologne, Germany) was performed and incubated at 28 ^{o}C for 2 days. Single colony was selected with the tip of loop and inoculated in 5 mL LB-Tet (5 µg/mL) and let grow overnight at 28 ^{o}C at 250 rpm. 1 mL of LB-Tet containing overnight grown *Agrobacterium* cells was poured into 100 mL LB-Tet (5 µg/mL) media and let grow at 28 ^{o}C overnight at 250 rpm for the second time. Following day, cells were let grow until absorbance (A_{600nm}) was between 0.5 – 0.7. The solution was divided into two 50 mL falcons and kept in ice for 30 minutes. Then, falcons centrifuged at 4 ^{o}C, at 3,500 rpm for 15 minutes. The supernatant was discarded while cells were stack at the bottom as pellet. 50 mL of cold (4 ^{o}C) 10% glycerol was added to both of 50 mL falcons. The cells were suspended by gently shaking the mixture until no

pellet remained. Then the mixture was centrifuged again at 4 ºC, at 3500 rpm for 15 minutes. The supernatant was discarded while cells were stack at the bottom as pellet. 50 mL of cold 10% glycerol was added to both of 50 mL falcons again and the centrifugation procedure were done once more. After centrifugation supernatant was discarded, while cells stack at the bottom as pellet were re-suspended in 200 µL GYT (10% glycerol, 0.125% yeast extract, 0.25% tryptone) medium. The solution was divided into four 1.5 mL tubes each containing 50 µL of GYT suspended electro-competent *Agrobacterium* GV3101 (pMP90) cells. Tubes were stored at -80 ºC for later usage.

2.8.2 Electroporation of *Agrobacterium* GV3101 (pMP90) electro-competent cells

The recombined plasmid to be transformed was obtained *via* isolation or ligation procedures. 2 µL of isolated or ligated plasmid (50 – 300 ng/µL) was mixed with 50 µL of *Agrobacterium* GV3101 (pMP90) electro-competent cell mixture in 1.5 mL tubes and kept for 10 minutes in ice. The mixture was poured into an ice-cold pulser cuvette with gap of 1 mm. The electric shock of 2.2 kV was applied to filled pulser cuvettes (Celljject duo, Therma corporation). 1 mL of SOC (Super Optimal broth with Catabolite repression) medium was added to the mixture and transferred to sterile 1.5 mL tubes. The electroporated *Agrobacterium* cells in medium were let to recover at 28 ºC at 150 rpm for 1.5 hours. The mixture was spread 150 µL per LB agar plate containing the appropriate antibiotics. Thus, only the transformed cells could survive. The spread LB agar plates were kept at 28 ºC in growth chamber for 2 – 3 days. Single colonies were picked up with the tip of loop and verified for presence of our gene of interest in the plasmid *via* colony PCR. Verified colonies were inoculated in 4mL LB containing the same antibiotic for 2 – 3 days at 28 ºC at 150 rpm.

2.8.3 Agro-infiltration

For agro-infiltration first media listed in Table 2.6, 2.7 and 2.8 were prepared.

Table 2.6 Media and their preparation required for agro-infiltration.

Media	Concentration and content
MES pH 5.7	0.5 M MES infiltration buffer
Acetosyringone	0.1 M in DMSO (Dimethyl sulfoxide)
$MgCl_2$	1 M
LB media	5 g yeast extract, 10 g tryptone, 5 g NaCl in 1L ddH_2O (autoclaved)
Gentamicin (Gen)	25 µg/mL
Spectinomycin (Spec)	100 µg/mL
Rifampicin (Rif)	10 µg/mL

Table 2.7 Preparation of L-MESA

Media	Amount	Concentration and content
MES 0.5 M, pH 5.7	1 mL	10 mM MES pH 5.7
Acetosyringone 0.1 M	10 µL	20 µM
LB media	49 mL	
Total	50 mL	

Table 2.8 Preparation of Agroinduction media.

Media	Amount	Concentration and content
MES 0.5 M, pH 5.7	1 mL	10 mM MES pH 5.7
$MgCl_2$ 1 M	0.5 mL	10 mM
Acetosyringone 0.1 M	75 µL	150 µM
ddH_2O	48.4 mL	
Total	50 mL	

Acetosyringone was added in L-MESA and Agroinduction media just before usage for acetosyringone not to lose its activity. Streak plate from transformed *Agrobacterium* cells were done to obtain single cell colony. Streak plates were incubated at 28 ^{o}C for 2 – 3 days. Single colony was selected with the tip of a loop and inoculated in 3 mL L-MESA containing Spectinomycin (Spec, 50 μg/mL), Gentamicin (Gen, 25 μg/mL), and Rifampicin (Rif, 10 μg/mL) inside 15 mL falcons and let grow at 28 ^{o}C, at 250 rpm in shaker for 24 – 30 hours. 500 μL of *Agrobacterium* grown culture was pipetted and inoculated into 5 mL L-MESA containing Kan (50 μg/mL), Gen (25 μg/mL), and Rif (10 μg/mL) and let grow at 28 ^{o}C, at 250 rpm in shaker for about 6 – 8 hours until absorbance (A_{600nm}) reached to 0.8 – 1.0. Then, the solution was centrifuged at 3,500 rpm for 10 minutes in microcentrifuge. Supernatant was discarded while *Agrobacterium* cells were stick at the bottom as pellet. The pellet was suspended in agro-induction media until the solution absorbance (A_{600nm}) reached to 1.0. The solution was kept overnight at room temperature in dark, meanwhile growth room temperature of *N. benthamiana* plants was decreased from 24 ^{o}C to 22 ^{o}C overnight, to increase agro-infiltration efficiency. Back sides of the leaves were infiltrated using 0.5 mL syringe without needle, at places away from main veins.

A firm hold of the front side of the leaf and gentle counter pressure was applied to the place of infiltration with the thumb of one hand, while *Agrobacterium* mixture was injected with the syringe. As the *Agrobacterium* mixture enters the intercellular space of the leaf, the infiltrated area turned visibly darker green. The injection continued until the darker green circle stopped expanding (Figure 2.5).

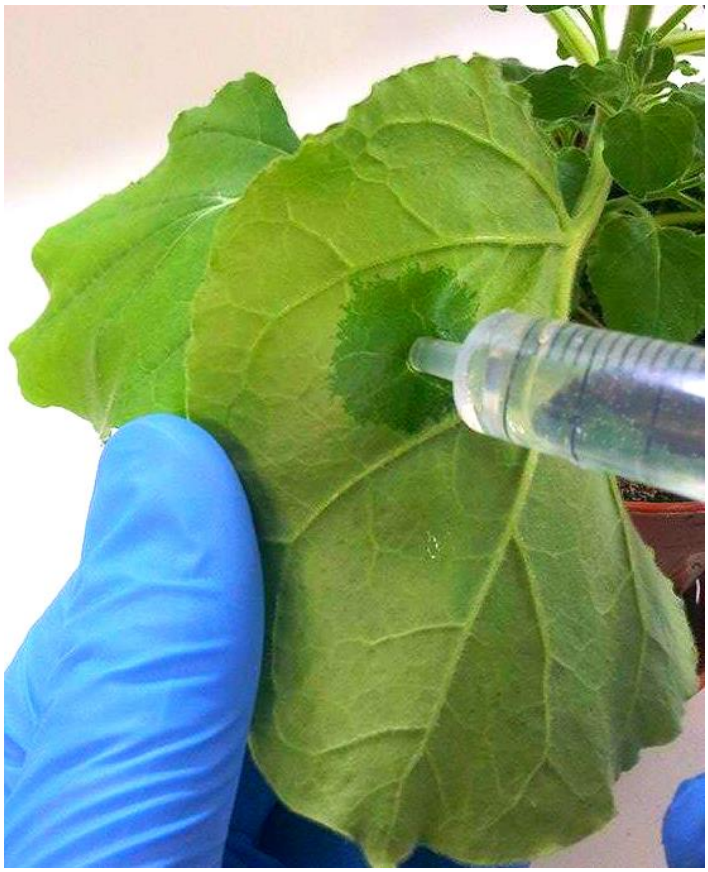

Figure 2.5 Agro-infiltration of *N. benthamiana* leaf using syringe. The darker region around needle tip shows the mixture entered the intercellular space of leaf.

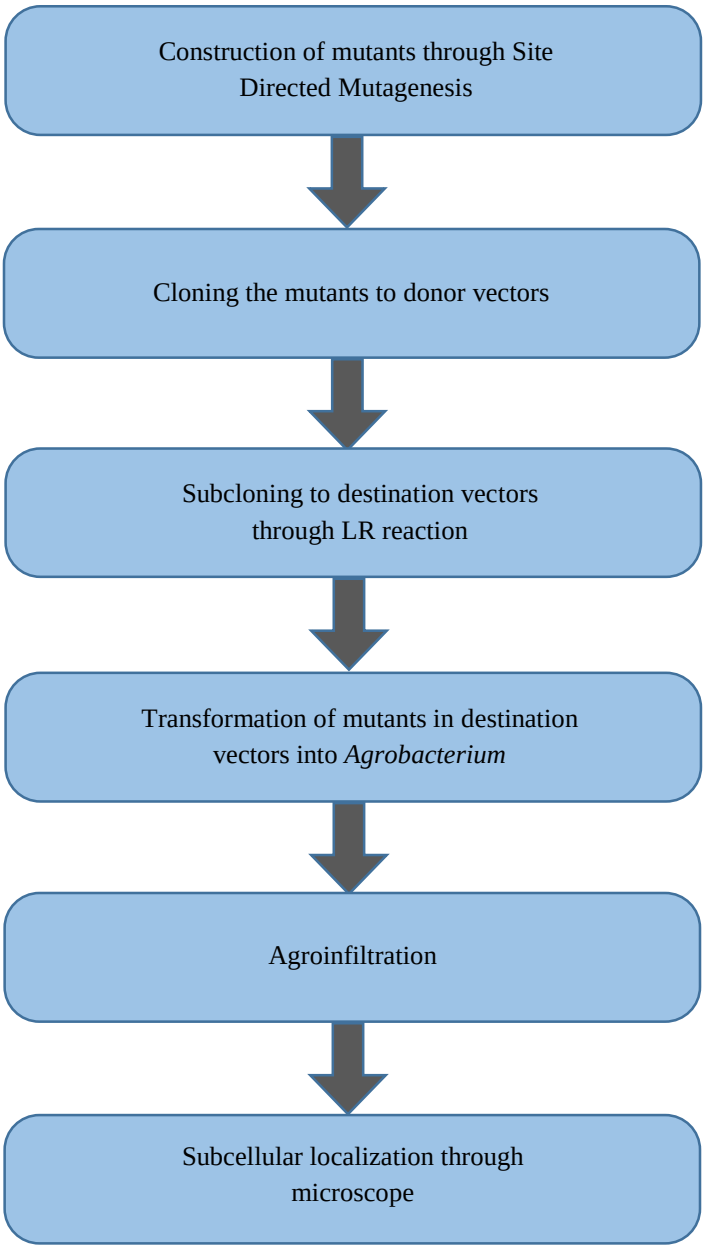

Figure 2.6 Flowchart of the steps of the experimental design.

CHAPTER 3

RESULTS AND DISCUSSION

3.1 Prediction of PstHa2a5 as an effector candidate

PstHa2a5 gene was predicted as an effector gene from data retrieved from haustorial cDNA library and analyzed by Yin and his colleagues (Yin *et al*., 2009a). PstHa2a5 cDNA sequence was obtained from NCBI database. PstHa2a5 gene sequence was converted into possible open reading frames (ORF) using ORF finder and the full length protein sequence (116 amino acids long) was obtained. The signal peptide of the gene product was predicted by 'SignalP 4.0 Server' online bioinformatics tool as shown in Figure 3.1 (http://www.cbs.dtu.dk/services/SignalP/). The probability of containing signal peptide is 0.937, having maximum cleavage site probability of 0.629 between positions 24 and 25. Note that at the site of maximum cleavage probability lies just before amino acids FKC which match exactly the conserved effector motif FXC.

Signal peptides generally are cut off form cleavage site after the protein has reached its destination. Presence of signal peptide and cleavage site in PstHa2a5 gene supports its function as effector protein which are common features of them.

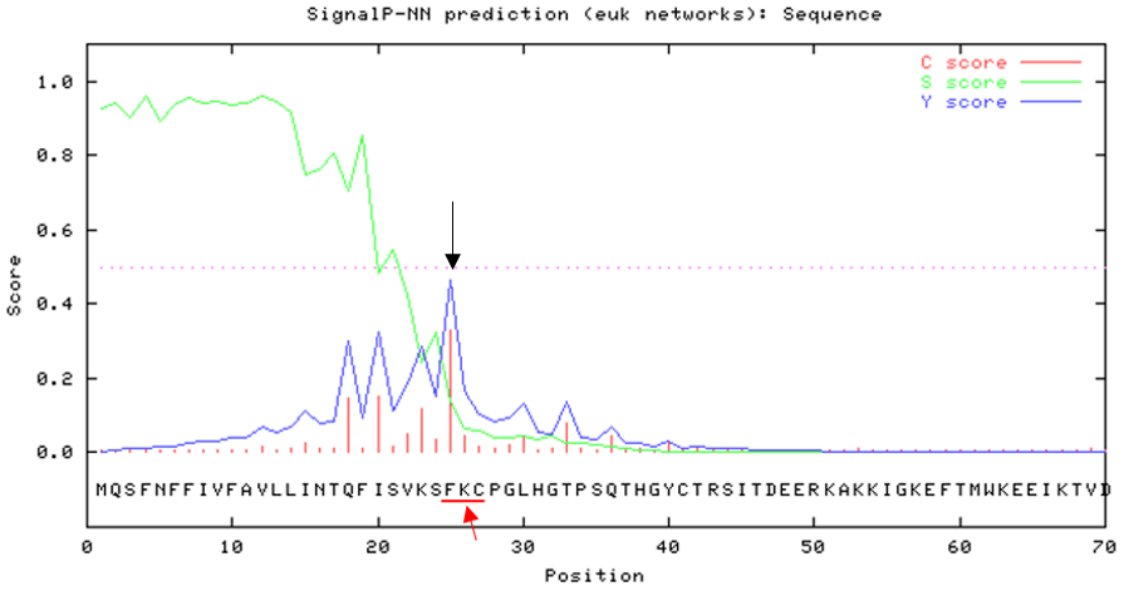

Figure 3.1 SignalP prediction results for PstHa2a5 gene encoded protein sequence. C score (red lines) represents possible signal peptide cleavage sites, S score (green line) represents signal peptide possibility, and Y score (blue line) represents a combination (geometric average) of the C-score and the slope of the S-score. With black arrow is pointed out the point of maximum cleavage probability. Red arrow at bottom marks FKC motif which lies immediately after probable cleavage site.

3.2 Decision for mutations

PstHa2a5 gene consist of FKC motif, which is one of the conserved effector motifs found as YXC, FXC and WXC when candidate effector proteins were investigated. PstHa2a5 is also one of these effectors which are rich in Cysteine residues. It is considered that Cysteines contribute to the stability and appropriate folding of the effectors by forming disulfide bonds. It is known for a long time that effectors rich in Cysteine residues are likely to be apoplastic effectors likely to maintain the structure outside of the cell (Stergiopoulos *et al.*, 2009). The presence of (Y/F/W)XC conserved motif predicted to

have a role in localization (Godfrey *et al*., 2010). Therefore, we decided to substitute this Cysteine with Serine and Valine. Serine was selected so that in the R-group the only difference was hydroxyl group instead of sulfhydryl group, so that although potential disulfide bond formation was eliminated with Serine, yet, the similarity of the amino acid was maintained. We also wanted to substitute Cysteine with Valine, although chemical structure is different the R-groups (–SH and $–CH_3$) have similar sizes. Also, Valine having an inert R-group as in the case of Serine, "– S – S –" formation was eliminated with these mutations. Thus, we intended to check a potential role of "–SH" for the subcellular localization of the gene product.

We decided to change the third amino acid, which is Cysteine to Serine and Valine. The reason of choosing these amino acids for modification was that as both Serine and Valine have similar size R-groups they will not disturb overall protein 3-D conformation but have different chemical properties (Figure 3.3 and 3.4).

1	atgcaaagcttcaacttcttcatcgtattcgcagtgttgttgatc 45
1	M Q S F N F F I V F A V L L I 15
46	aacactcaattcatttctgtgaagtcgttcaagtgtcccggtttg 90
16	N T Q F I S V K S F K C P G L 30
91	catggaacgccaagccaaacacatggttattgcaccagatcaatc 135
31	H G T P S Q T H G Y C T R S I 45
136	accgatgaagaacgaaaggcaaaaaagattggcaaggagttcacc 180
46	T D E E R K A K K I G K E F T 60
181	atgtggaaggaagaaatcaagacagtcgacgggaaattctcgtgt 225
61	M W K E E I K T V D G K F S C 75
226	gataaagtggacttgaatgggtcggttgccacagatagcttctgt 270
76	D K V D L N G S V A T D S F C 90
271	tgtgacgttgcaggtagaattggtgaagttgagaaaagtaaacaa 315
91	C D V A G R I G E V E K S K Q 105
316	gctatgtggacaaacaactgctccaaagcatcttag 351
106	A M W T N N C S K A S * 116

ttcaagtgt : DNA sequence

F K C : Amino acid sequence

Figure 3.2 Original nucleotide and amino acid sequence of PstHa2a5 gene. FKC is the suspected cleavage site. C was targeted to change by side directed mutagenesis.

1	atgcaaagcttcaacttcttcatcgtattcgcagtgttgttgatc 45
1	M Q S F N F F I V F A V L L I 15
46	aacactcaattcatttctgtgaagtcgttcaagtctcccggtttg 90
16	N T Q F I S V K S F K S P G L 30
91	catggaacgccaagccaaacacatggttattgcaccagatcaatc 135
31	H G T P S Q T H G Y C T R S I 45
136	accgatgaagaacgaaaggcaaaaaagattggcaaggagttcacc 180
46	T D E E R K A K K I G K E F T 60
181	atgtggaaggaagaaatcaagacagtcgacgggaaattctcgtgt 225
61	M W K E E I K T V D G K F S C 75
226	gataaagtggacttgaatgggtcggttgccacagatagcttctgt 270
76	D K V D L N G S V A T D S F C 90
271	tgtgacgttgcaggtagaattggtgaagttgagaaaagtaaacaa 315
91	C D V A G R I G E V E K S K Q 105
316	gctatgtggacaaacaactgctccaaagcatcttag 351
106	A M W T N N C S K A S * 116

ttcaagtct tgt was changed to tct : DNA sequence

F K S FKC was changed to FKS : Amino acid sequence

Cysteine
Cys, C

Serine
Ser, S

Figure 3.3 Nucleotide and amino acid sequence of site directed mutated PstHa2a5-C27S gene. FKC motif have been changed to FKS.

1	atgcaaagcttcaacttcttcatcgtattcgcagtgttgttgatc 45
1	M Q S F N F F I V F A V L L I 15
46	aacactcaattcatttctgtgaagtcgttcaaggttcccggtttg 90
16	N T Q F I S V K S F K V P G L 30
91	catggaacgccaagccaaacacatggttattgcaccagatcaatc 135
31	H G T P S Q T H G Y C T R S I 45
136	accgatgaagaacgaaaggcaaaaaagattggcaaggagttcacc 180
46	T D E E R K A K K I G K E F T 60
181	atgtggaaggaagaaatcaagacagtcgacgggaaattctcgtgt 225
61	M W K E E I K T V D G K F S C 75
226	gataaagtggacttgaatgggtcggttgccacagatagcttctgt 270
76	D K V D L N G S V A T D S F C 90
271	tgtgacgttgcaggtagaattggtgaagttgagaaaagtaaacaa 315
91	C D V A G R I G E V E K S K Q 105
316	gctatgtggacaaacaactgctccaaagcatcttag 351
106	A M W T N N C S K A S * 116

ttcaaggtt tgt was changed to gtt : DNA sequence

F K V FKC was changed to FKV : Amino acid sequence

Figure 3.4 Nucleotide and amino acid sequence of site directed mutated PstHa2a5-C27V gene. FKC motif have been changed to FKV.

3.3 Generation of point mutation mutations by site directed mutagenesis

PstHa2a5 gene with 5' end CACC flanking region was amplified with PCR from PstHa2a5-pBlueScript II SK plasmid construct, for construction of PstHa2a5-C27S and PstHa2a5-C27V mutants.

3.3.1 Construction of PstHa2a5-C27S

3.3.1.1 Construction of PstHa2a5-C27S with stop codon

For construction of PstHa2a5-C27S gene with stop codon two consecutive PCRs were performed. PstHa2a5-CACC-fwd and PstHa2a5-C27S-rev pair of primers were used for generating the short fragment from 5' to mutation region, PstHa2a5-C27S-fwd and PstHa2a5-rev-with-stop pair of primers were used for generating the long fragment from mutation region to 3' end containing stop codon (TAG). In second step joining fragment PCR of two fragments were performed to obtain full length of PstHa2a5 gene containing C27S mutation and the stop codon at the end (Figure 3.5).

3.3.1.2 Construction of PstHa2a5-C27S without stop codon

For construction of PstHa2a5-C27S gene without stop codon firstly 2 PCRs were performed. PstHa2a5-CACC-fwd and PstHa2a5-C27S-rev pair of primers were used for generating the short fragment from 5' to mutation region. PstHa2a5-C27S-fwd and PstHa2a5-rev-without-stop pair of primers were used for generating the long fragment from mutation region to 3' end. In second step joining fragment PCR of two fragments were performed to obtain full length of PstHa2a5 gene containing C27S mutation (Figure 3.5).

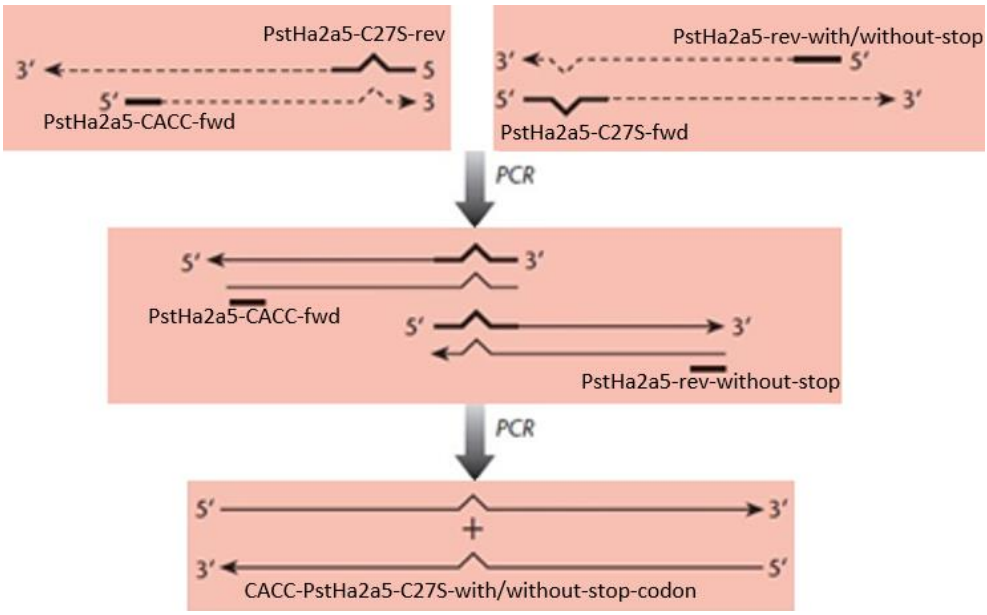

Figure 3.5 Scheme of the procedure for construction of PstHa2a5-C27S mutant gene with or without stop codon.

3.3.2 Construction of PstHa2a5-C27V

3.3.2.1 Construction of PstHa2a5-C27V with stop codon

For construction of PstHa2a5-C27V gene with stop codon firstly 2 PCRs were performed. PstHa2a5-CACC-fwd and PstHa2a5-C27V-rev pair of primers were used for generating the short fragment from 5' end to mutation region. PstHa2a5-C27V-fwd and PstHa2a5-rev-with-stop pair of primers were used for generating the long fragment from mutation region to 3' end containing stop codon (TAG). In second step joining fragment PCR of two fragments were performed to obtain full length of PstHa2a5 gene containing C27V mutation and the stop codon at the end (Figure 3.6).

3.3.2.2 Construction of PstHa2a5-C27V without stop codon

For construction of PstHa2a5-C27V gene without stop codon firstly 2 PCRs were performed. PstHa2a5-CACC-fwd and PstHa2a5-C27V-rev pair of primers were used for generating the short fragment from 5' to mutation region. PstHa2a5-C27V-fwd and PstHa2a5-rev-without-stop pair of primers were used for generating the long fragment from mutation region to 3' end. In second step joining fragment PCR of two fragments were performed to obtain full length of PstHa2a5 gene containing C27V mutation (Figure 3.6).

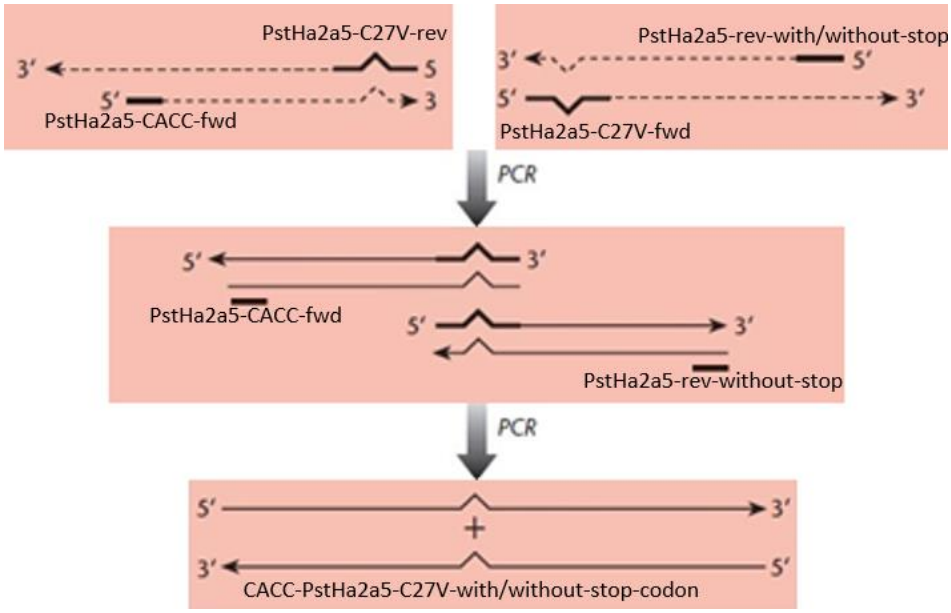

Figure 3.6 Scheme of the procedure for construction of PstHa2a5-C27V mutant gene with or without stop codon.

3.3.3 Confirmation of fragment lengths and mutations

For confirmation of amplified fragment lengths, PCR products were run in 1% agarose gel in 1X TAE buffer and visualized under UV light. As DNA ladder, 100 bp DNA Ladder (0.5 μg) (Cat# N3231S, Lot# 0831006, NEB) was used and the sample fragment length was measured comparing the known ladder fragments. 1μL of DNA ladder and 2μL of samples were mixed with 1μL of 6X loading dye (Thermo Scientific, #R0611) and the volume was completed to 6 μL with ddH_2O prior to loading on the gel.

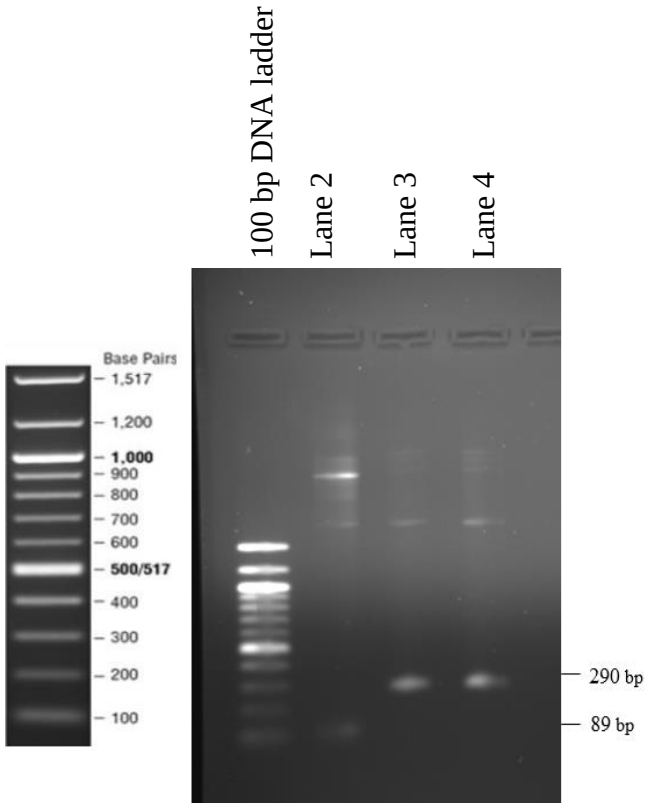

Figure 3.7 Agarose gel electrophoresis image of PCR for generating short and long fragments of PstHa2a5-C27S mutants. Lane 1: 100 bp DNA ladder, Lane 2: short

fragment (5’end to mutation region), Lane 3: long fragment with stop codon (mutation region to 3’ end), Lane 4: long fragment without stop codon (mutation region to 3’ end)

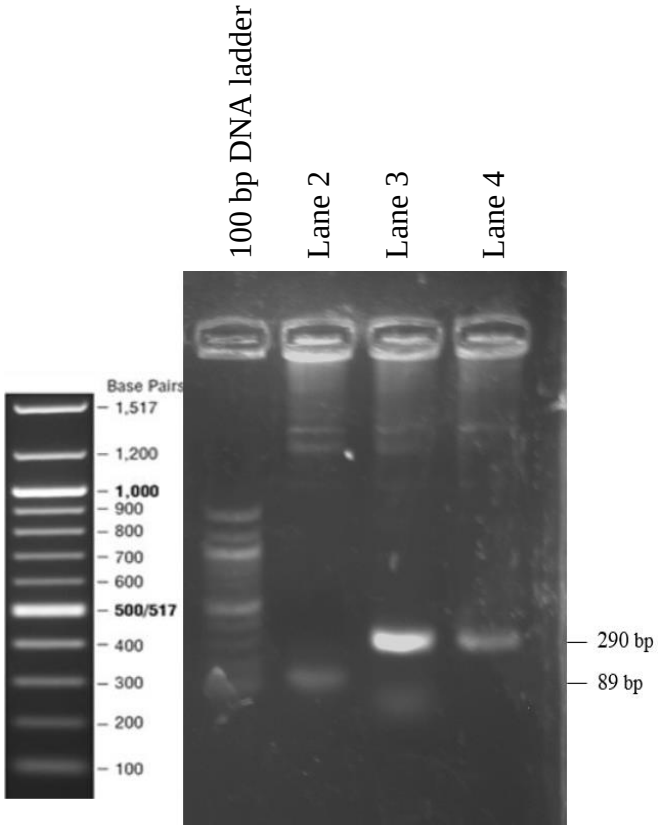

Figure 3.8 Agarose gel electrophoresis image of PCR for generating short and long fragments of PstHa2a5-C27V mutants. Lane 1: 100 bp DNA ladder, Lane 2: short fragment (5’end to mutation region), Lane 3: long fragment with stop codon (mutation region to 3’ end), Lane 4: long fragment without stop codon (mutation region to 3’ end).

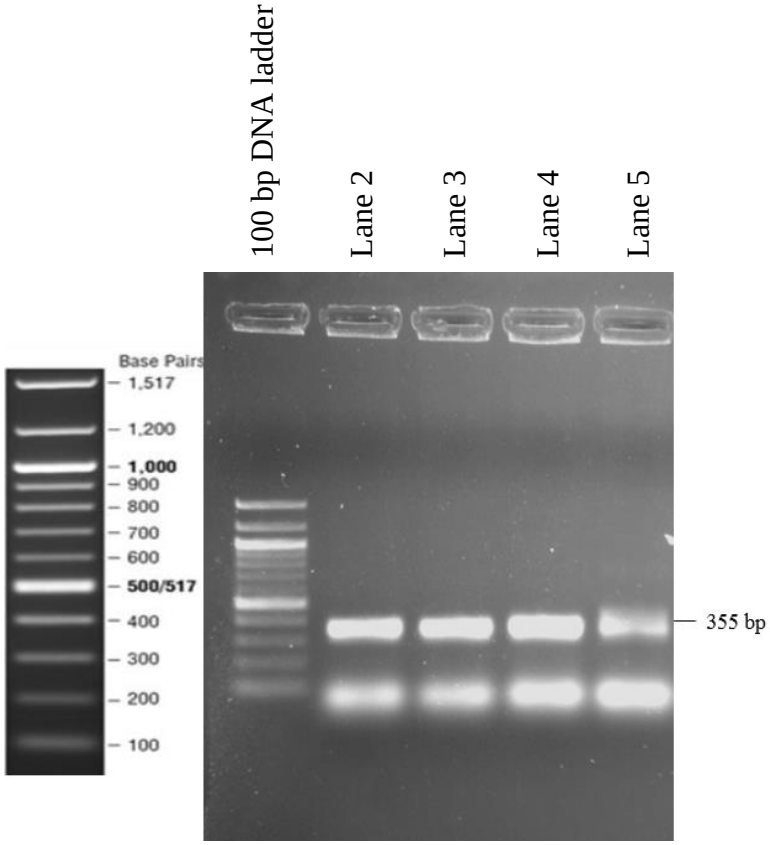

Figure 3.9 Agarose gel electrophoresis image of joining fragment PCR. Lane 1: 100 bp DNA ladder, Lane 2: PstHa2a5-C27S with stop codon, Lane 3: PstHa2a5-C27S without stop codon, Lane 4: PstHa2a5-C27V with stop codon, Lane 5: PstHa2a5-C27S without stop codon. The bright bands observed below 100 bp were due excessive primer usage.

To confirm the mutation of sequences, joining PCR products made sequenced by Sanger sequence method (Figure 3.10).

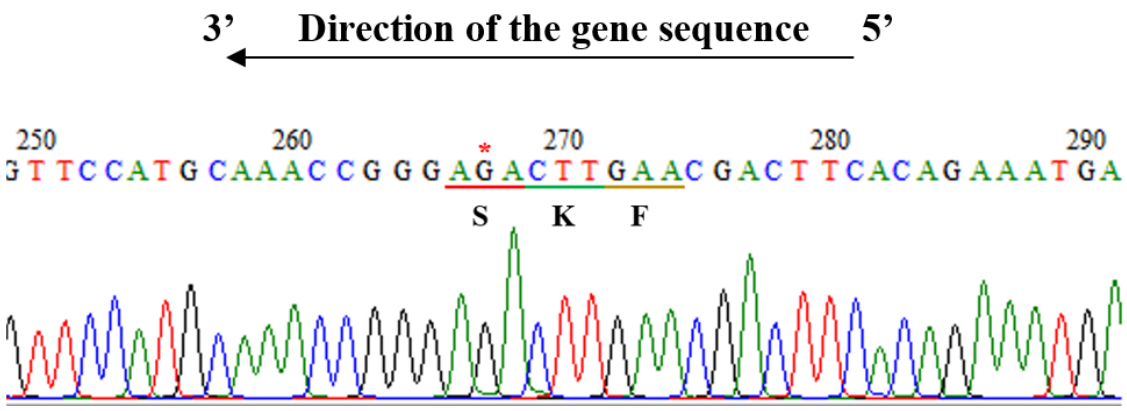

Figure 3.10 Chromatogram showing the C27S mutation site of PstHa2a5 gene. The DNA strand being read was complementary to the coding strand of the gene. The nucleotide indicated with asterisk show the place of mutation. The sequence has been changed from Cysteine coding ACA to Serine coding AGA triplet.

3.4 Ligation of mutants to donor vector and transformation to *E. coli*

Amplified four mutant versions of PstHa2a5 gene with 5’ end CACC flanking region were ligated into pENTR™/D-TOPO donor vector as described in Section 2.5.1. Four constructs were produced: 1) PstHa2a5-C27S-with-stop-codon-pENTR™/D-TOPO, 2) PstHa2a5-C27S-without-stop-codon-pENTR™/D-TOPO, 3) PstHa2a5-C27V-with-stop-codon-pENTR™/D-TOPO, and 4) PstHa2a5-C27V-without-stop-codon-pENTR™/D-TOPO. These constructs were transformed to competent *E. coli* TOP10 cells for

amplification and storage. Kanamycin (50 μg/mL) was used in LB and agar LB for elimination of non-transformed cells.

3.5 Plasmid isolation of donor vectors from *E. coli* TOP10 cells

Four different transformed *E. coli* TOP10 cells were grown in Kan-LB overnight for amplification. Next day, plasmid isolation was performed as described in Section 2.7. After the isolation of plasmid DNA, the content was measured in NanoDrop (Figure 3.11).

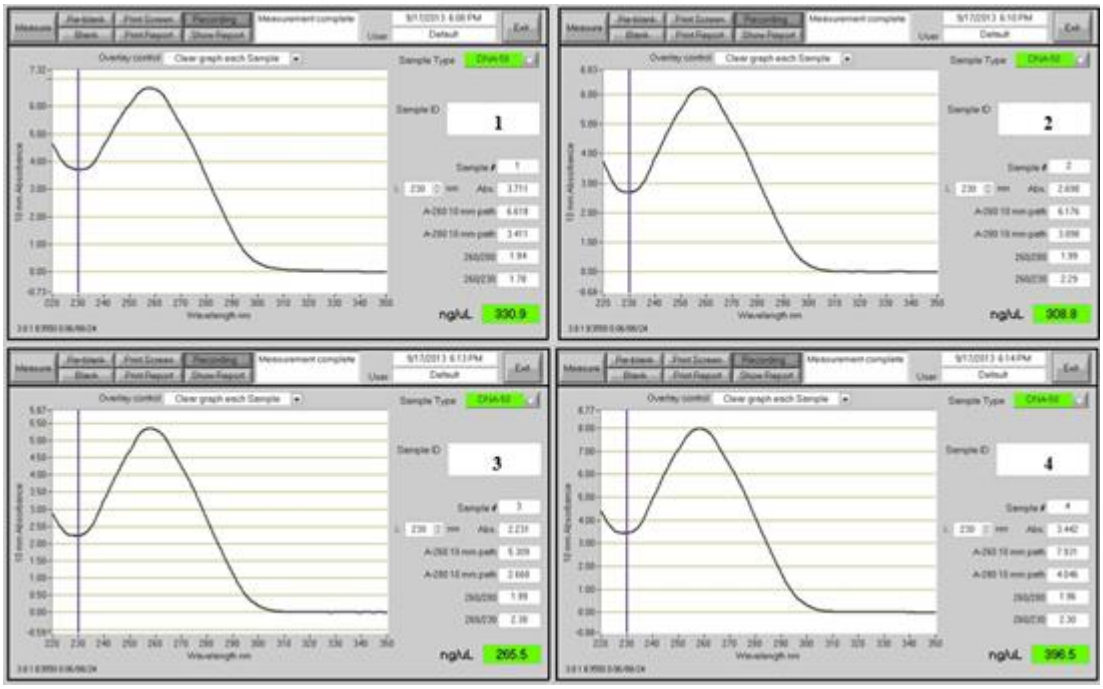

Figure 3.11 NanoDrop results of plasmid isolation of donor vectors. 1) PstHa2a5-C27S-with-stop-codon-pENTR™/D-TOPO, 2) PstHa2a5-C27S-without-stop-codon -pENTR™/D-TOPO, 3) PstHa2a5-C27V-with-stop-codon-pENTR™/D-TOPO, 4) PstHa2a5-C27V-without-stop-codon-pENTR™/D-TOPO.

3.6 Subcloning to destination vector and transformation to *E. coli*

Destination vectors pK7WGF2 and pK7FWG2 were obtained from overnight grown *E. coli* DB3.1 strain cells containing corresponding vectors. Vectors contained *ccdB* gene before LR reaction which is lethal for *E. coli* TOP10 but not lethal for *E. coli* DB3.1 strain. Next day, plasmid isolation from these cells were performed (Figure 3.12).

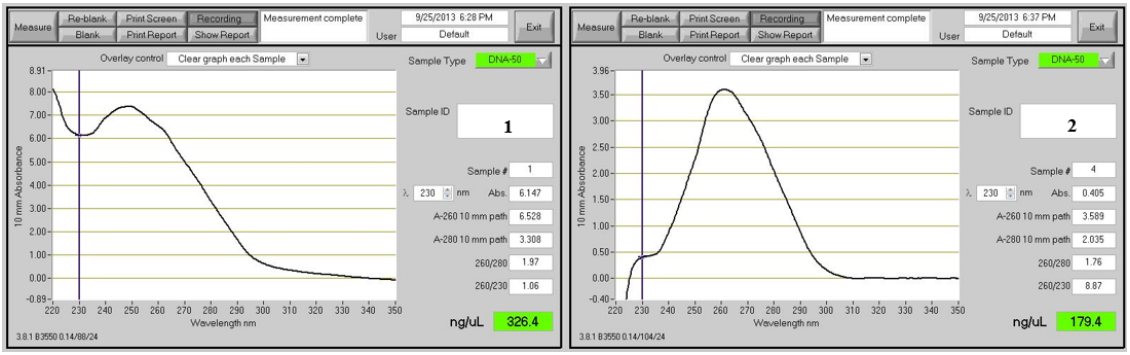

Figure 3.12 NanoDrop results of plasmid isolation of destination vectors. 1) pK7WGF2, and 2) pK7FWG2.

Isolated donor vectors were subcloned to destination vectors as described in Section 2.5.2. Four constructs were produced: 1) PstHa2a5-C27S-with-stop-codon-pK7WGF2, 2) PstHa2a5-C27S-without-stop-codon-pK7FWG2, 3) PstHa2a5-C27V-with-stop-codon-pK7WGF2, and 4) PstHa2a5-C27V-without-stop-codon-pK7FWG2. These constructs were transformed to competent *E. coli* TOP10 cells for amplification and storage. Spectinomycin (100 µg/mL) was used in LB and agar LB for elimination of non-transformed cells.

3.7 Plasmid isolation of destination vectors from *E. coli* TOP10 cells

Four different transformed *E. coli* TOP10 cells were grown in Spec-LB overnight for amplification. Next day, plasmid isolation was performed as described in Section 2.7. After the isolation of plasmid DNA, the content was measured in NanoDrop (Figure 3.13).

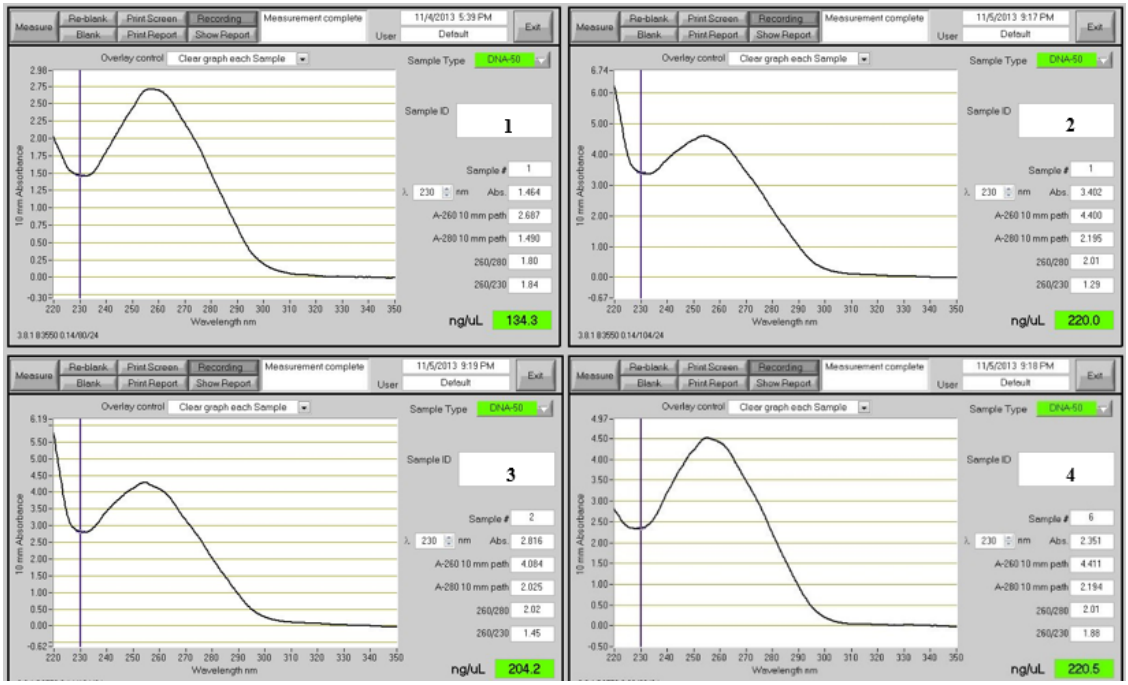

Figure 3.13 NanoDrop results of plasmid isolation of destination vectors. 1) PstHa2a5-C27S-with-stop-codon-pK7WGF2, 2) PstHa2a5-C27S-without-stop-codon- pK7FWG2, 3) PstHa2a5-C27V-with-stop-codon-pK7WGF2, 4) PstHa2a5-C27V-without-stop-codon-pK7FWG2.

3.8 Transformation to *Agrobacterium* GV3101 (pMP90) cells

The four different constructs of destination vectors were transformed into electrocompetent *Agrobacterium* GV3101 (pMP90) cells as described in Section 2.8.2. Electroporaterd cells were grown in LB containing Kanamycin (Kan, 50 μg/mL), Gentamicin (Gen, 25 μg/mL), and Rifampicin (Rif, 10 μg/mL) to eliminate non-transformant *Agrobacterium* cells. Grown cells were used for storage and preparation for the agro-infiltration.

3.9 Agro-infiltration

Four different transformed *Agrobacterium* cells with corresponding destination vectors got prepared for agro-infiltration procedure as described in Section 2.8.3. After preparation, 6 week old *N. benthamiana* leaves were used for agro-infiltration (Figure 3.14). One plant was used for only one type of destination vector deliverance. 3 – 4 leaves of the plant were infiltrated.

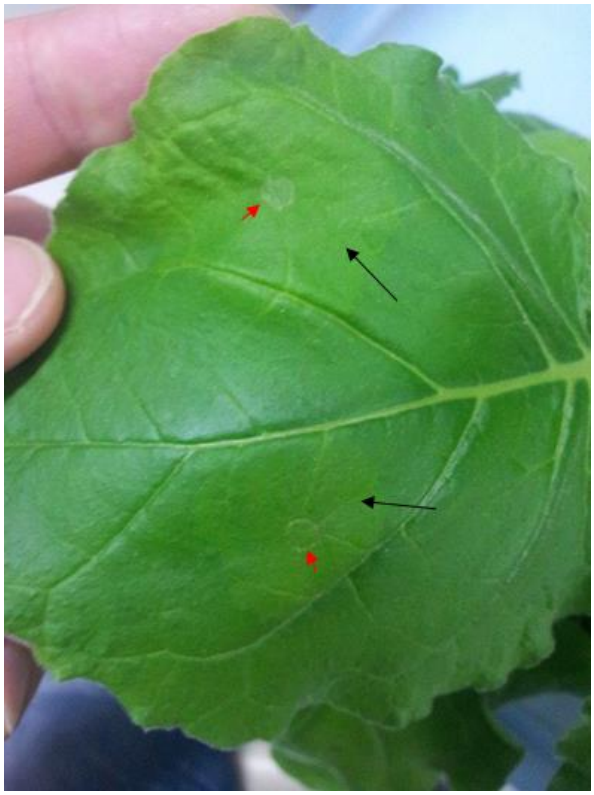

Figure 3.14 Infiltrated *N. benthamiana* leaf 2 days post inoculation (dpi). The red arrows show the spots of infiltration. Small circular scar occurs due to pressure of the tip of syringe. The back arrows show the tissue that was infiltrated. This region has lighter green color and a circular shape around the point of infiltration.

3.10 Subcellular localization

3.10.1 Free GFP

When GFP was expressed alone in *N. benthamina* leaves, it appeared in the cytoplasm and in the nucleus as it is obvious in the following florescent microscope pictures at different focus settings (Figure 3.15).

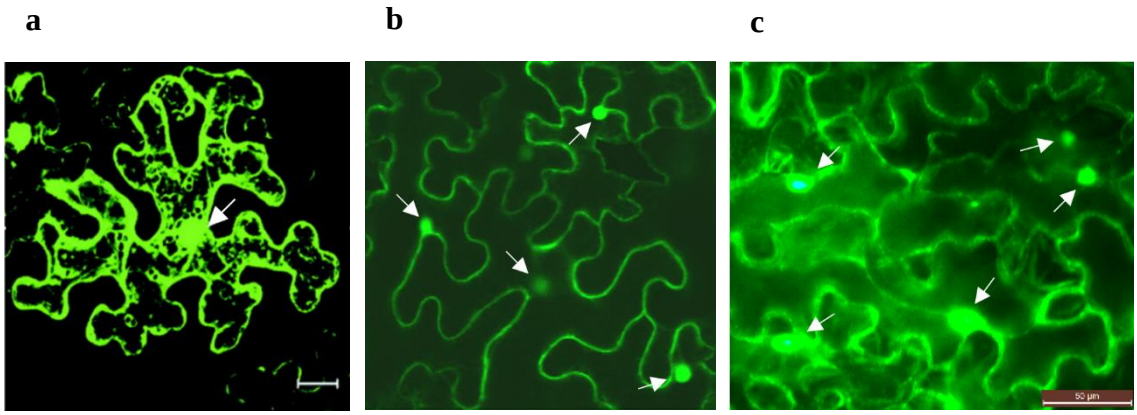

Figure 3.15 Pictures with varying focus settings of subcellular localization of free GFP expressed in *N. benthamiana* leaf epidermal cells. The protein was dispersed in the cytoplasm and nuclei (white arrows). GFP did not penetrate vacuoles, which remained unstained. Due to the vacuole size, the cytoplasm often appears as a thick band in the cell periphery. **a)** (Lavy *et al*., 2002); **b)** Retrieved from: https://plantscientist. wordpress.com /2013/05/04/organism-of-the-week-aequorea-victoria-jellyfish; **c)** Taken from Bayantes Dagvadorj.

3.10.2 Wild type PstHa2a5 gene with signal peptide

The construct (made by Bayantes Dagvadorj) was made with the wild type PstHa2a5 by fusing it to the N-terminus of the GFP, so that the GFP was placed to the C-terminus of the signal peptide (SP)-containing gene (Figure 3.16). For the determination of the subcellular localization, the agro-infiltrated leaves were observed under the light microscope 2 days post infiltration (dpi). As it would be seen in the figure, the signal peptide containing PstHa2a5 was localized in cytoplasm and some has been secreted outside the cell, to the intercellular space as it might be expected, since the plant cell is most likely possessing the signal peptide, thus enabling the gene to be secreted to the intercellular space.

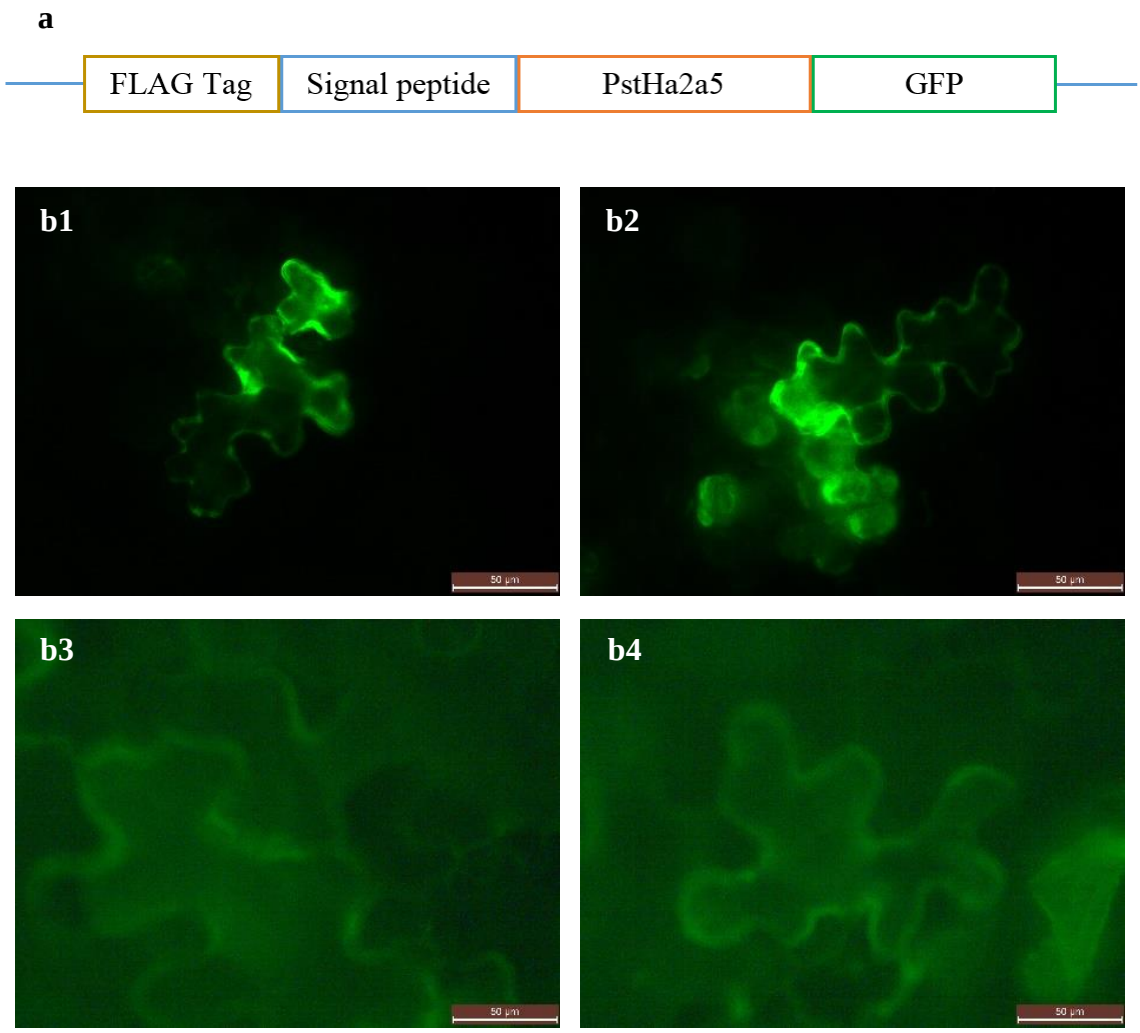

Figure 3.16 Pictures with varying focus settings of subcellular localization of the wild type PstHa2a5 with signal peptide (pK7FWG2 / FLAG-Tag+SP+PstHa2a5+GFP / *N. benthamiana*). The representative pictures out of 25. **a)** Structure of the construct; **b)** Pictures of subcellular localization assay: agro-infiltrated leaves were observed after 2 days post inoculation (dpi) under microscope (Leica DM4000B microscope / DFC 280 camera) under 40X magnification.

Due to the positioning of the cells in these two examples (**b1**, **b2**) at this particular focus setting, the upper part of the cell in **b1** and the lower part of the cell in **b2** appear closer to the ocular of the microscope, thereby they these regions look more intense GFP luminescence. Even though wild type effector with signal peptide does not show accumulation in the nucleus (**b1, b2**). In pictures **b3** and **b4**, the cells appear at the same plane, horizontally with a uniform GFP luminescence. Therefore, these pictures (**b3**, **b4**) better indicate the absence of nucleus targeting, when the effector expressed with N terminal signal peptide suggesting it was secreted outside of the cell. Cloudy appearance of all the pictures in **b1**-**b4** also supports release of the effector out of the cell when it has the signal peptide on its N-terminus.

3.10.3 Wild type PstHa2a5 gene without signal peptide

When the signal peptide was removed (Figure 3.17), secretion of the gene to the intercellular matrix was not observed, as expected. The importance of this observation is that this effector does not appear to be an apoplastic effector, since we were only observing cytoplasmic and nuclear localizations, which is very similar to the expression of the GFP gene alone in *N. benthamiana*.

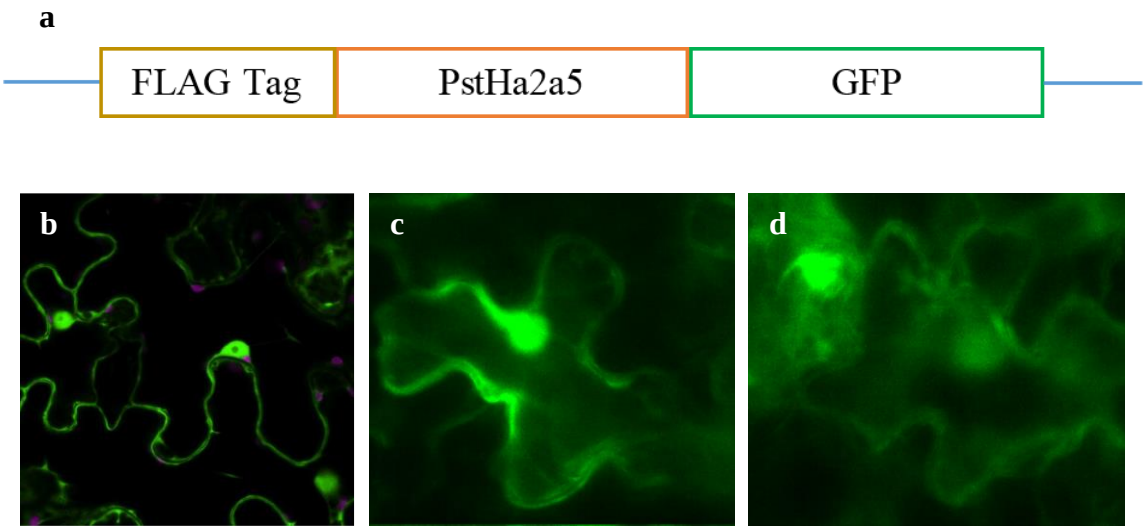

Figure 3.17 Pictures with varying focus settings of subcellular localization of the wild type PstHa2a5 without signal peptide (pK7FWG2 / FLAG-Tag+PstHa2a5+GFP / *N. benthamiana*). The representative pictures out of 25. **a)** Structure of the construct; **b)** Confocal microscopy performed on Leica DM6000B/TCS SP5 confocal microscope (Leica Microsystems CMS GmbH, Germany); **c** and **d)** Pictures of subcellular localization assay. Agro-infiltrated leaves were observed after 2 days post inoculation (dpi) under microscope (Leica DM4000B microscope / DFC 280 camera). Pictures were digitally magnified from 40X magnification.

As a result, wild PstHa2a5 gene with signal peptide was expressed in cytoplasm and nucleus, whereas GFP luminescence was also observed in the exterior of the cells (Figure 3.16). In the case of the wild type PstHa2a5 gene without the signal peptide, protein remained inside the cell and the cellular nucleus (Figure 3.17). It implies that signal peptide cause protein to be secreted outside of the cell. Similarly, when pathogen secretes

its effectors out of the haustorium, the signal peptide is likely to be cleaved out and the effector is localized inside the plant cell. Here, signal peptide causes the protein to move across the membrane. It is understood that PstHa2a5 gene besides crossing the haustorial membrane, it can also cross the plant membrane and reach to the intracellular matrix.

3.10.4 PstHa2a5-C27S mutants

The subcellular localization of the signal peptide containing PstHa2a5-C27S mutant in *N. benthamiana* was detected with the construct in which the gene was fused to the GFP either to N-terminus or the C-terminus of the gene. Independently to which end of protein GFP was fused, the results were similar; protein was expressed in the cells and secreted outside in both of the cases (Figure 3.18 and 3.19). Comparing to the wild type PstHa2a5 with signal peptide (Figure 3.16), it appears that C27S mutation did not cause distinctive difference in the localization of the effector. This outcome indicates that mutating the Cysteine into Serine, both having similar chemical properties, was not significant in preventing the secretion of the gene to the intercellular matrix. In other words, the mutation did not change the ability of the gene to be secreted. Thus, in the pathogen infected susceptible plant cell, this observation translates into which for the effector to be secreted out of the haustorium, the Cysteine residue within the FXC motif does not have a very significant role. In turn, it may be suggested that this Cysteine is not involved in forming a disulphide bond, if it were, the changing it into Serine, which cannot form a disulphide bond, should have had affected the ability of secretion. The –OH in Serine having the similar chemical properties with –SH in Cysteine appears to be maintaining the effective secretion, if anything, it looks like it enhances the secretion. When compared to the wild type effector with the signal peptide (Figure 3.16), secretion of the effector protein appeared to be reduced (Figure 3.16 versus Figure 3.18).

Below, microscopic pictures of the site specifically mutated effector with N- signal peptide are presented.

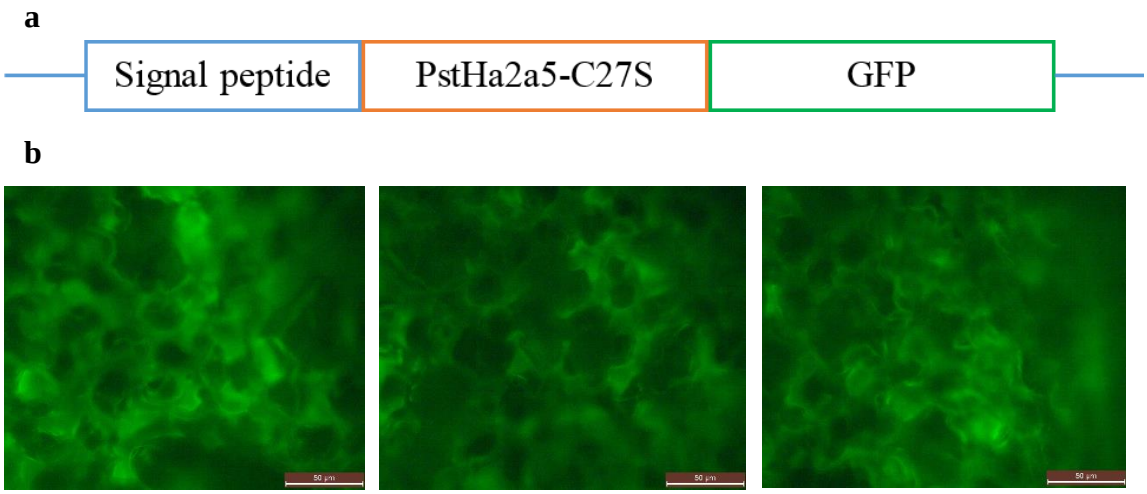

Figure 3.18 Pictures with varying focus settings of subcellular localization of PstHa2a5-C27S (pK7FWG2 / SP+PstHa2a5 - C27S / *N. benthamiana*). The representative pictures out of 25. **a)** Structure of the construct; **b)** Pictures of subcellular localization assay: agro-infiltrated leaves were observed after 2 days post inoculation (dpi) under microscope (Leica DM4000B microscope / DFC 280 camera) under 40X magnification.

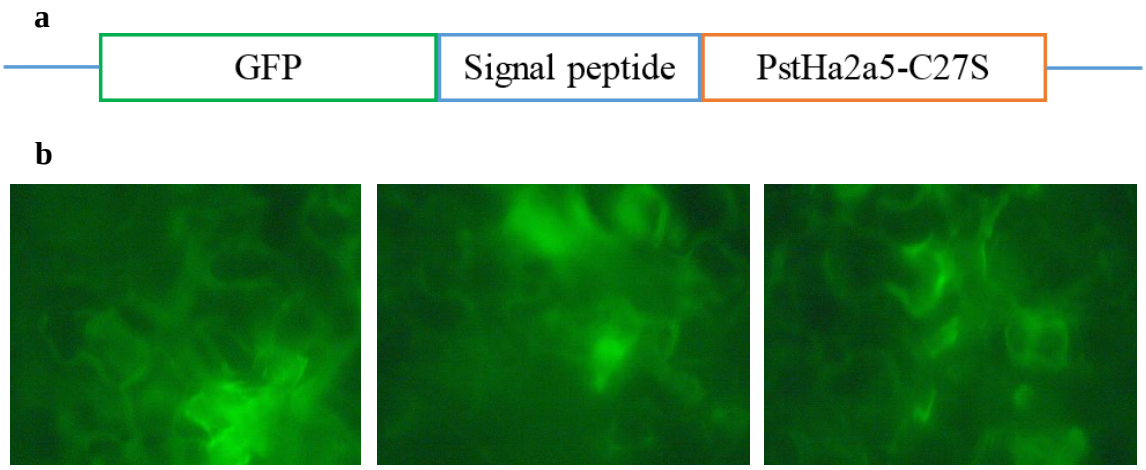

Figure 3.19 Pictures with varying focus settings of subcellular localization of PstHa2a5-C27S (pK7WGF2 / SP+PstHa2a5 - C27S / *N. benthamiana*). The representative pictures out of 25. **a)** Structure of the construct; **b)** Pictures of subcellular localization assay: agro-infiltrated leaves were observed after 2 days post inoculation (dpi) under microscope (Leica DM4000B microscope / DFC 280 camera) under 40X magnification.

Regardless the of position of GFP fusion in both constructs (Figures 3.18 and 3.19) all of the C27S mutants look very similar to the wild type effector with signal peptide with no appearance in nucleus, and successful secretion to the outside of the cell, which is suggested by the cloudy appearance GFP luminescence. Based on these observations the Cysteine to Serine mutation in FKC indicates no effect. The reasons for this may be because of the followings;

i) If wild in the wild type effector, the cysteine in FKC is not involved in disulphide bonding, –SH group function can be well mimicked with the hydroxyl of the Serine due to similar electron densities.

ii) If the native effector is involved in disulphide bonding with another cysteine in its structure, the hydroxyl of the Serine again due to similar electron density and the ability to make hydrogen bonding, may still facilitate maintaining the native structure of the effector regardless of the absence of a covalent bonding.

The above considerations cannot be made independent of the observations obtained with the other mutant namely; C27V in FKC. In the microscopic pictures below the observations of C27V mutant are presented.

3.10.5 PstHa2a5-C27V mutants

The subcellular localization of PstHa2a5-C27V mutant in *N. benthamiana* was conducted with the construct having the mutant with GFP to N-terminus or C-terminus of the gene. Independently to which end of protein GFP was joined, the results were similar. Unlike C27S, substitution to Valine prevents the protein to be maintained in the nucleus, yet it still appears in the cytoplasm and the outside of the cell. The result is suggesting that although substitution of Cysteine with Valine is not interfering with its secretion outside of the cell, it is making the protein easier to leak out of the nucleus, most likely due to its effect on folding. Protein may becoming more relaxed with this mutation, which was not observed with C27S, in which Serine is still having an electronegative group with capability of hydrogen bonding which may help maintaining a structure similar to that of the wild type structure, keeping it compact and as difficult as the wild type to stay partly in the nucleus (Figure 3.19).

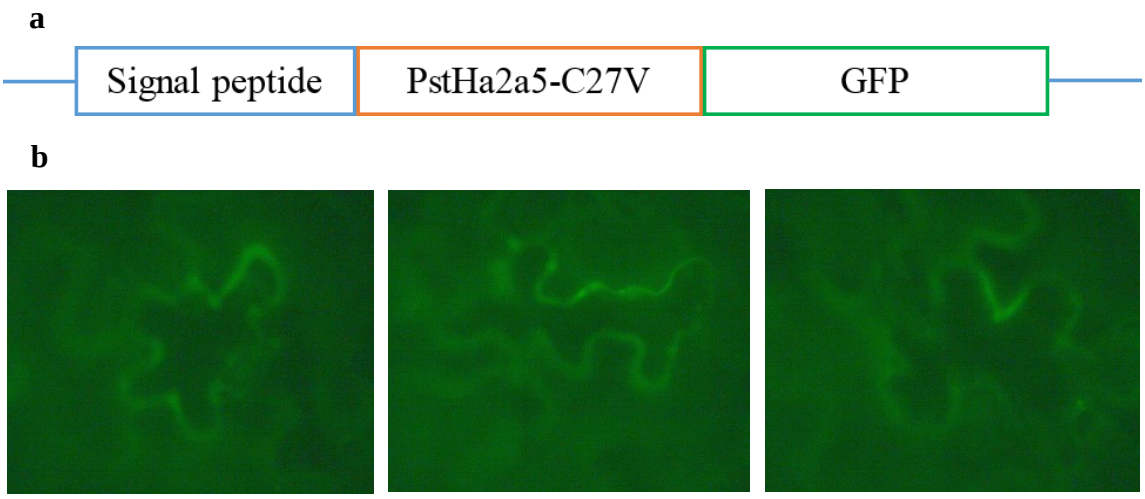

Figure 3.20 Pictures with varying focus settings of subcellular localization of PstHa2a5-C27V (pK7FWG2 / SP+PstHa2a5 - C27S / *N. benthamiana*). The representative pictures out of 25. **a)** Structure of the construct; **b)** Pictures of subcellular localization assay: agro-infiltrated leaves were observed after 2 days post inoculation (dpi) under microscope (Leica DM4000B microscope / DFC 280 camera) under 40X magnification.

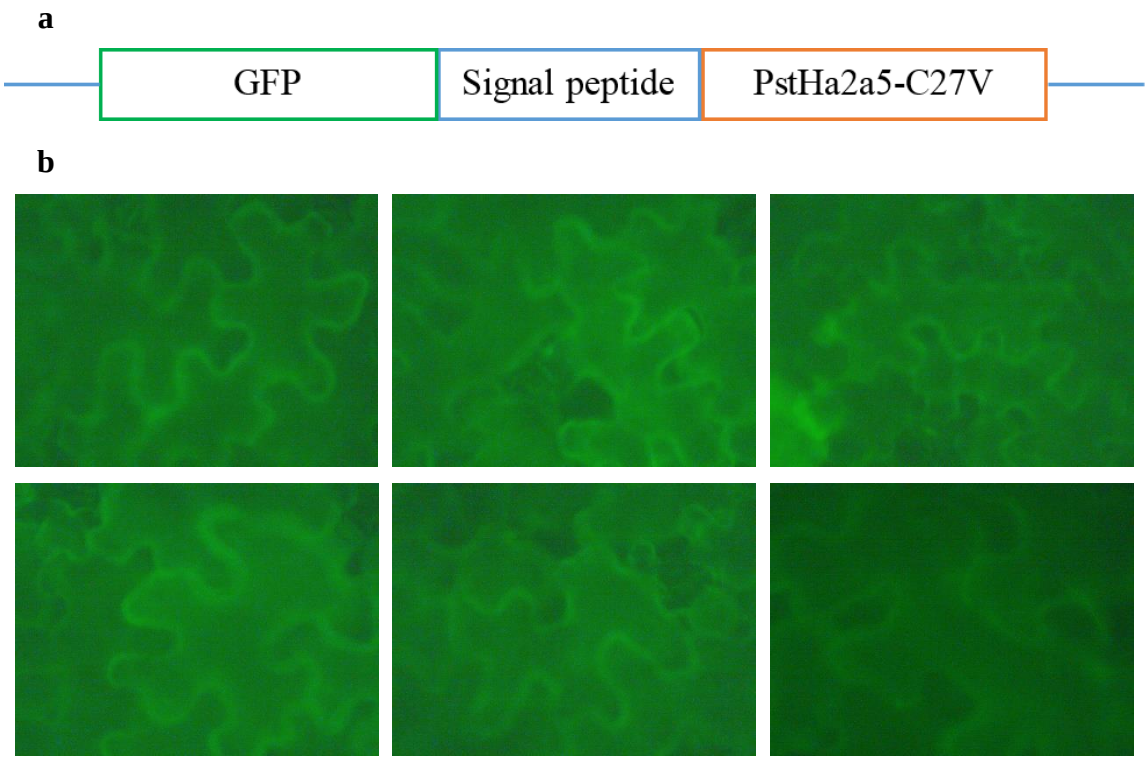

Figure 3.21 Pictures with varying focus settings of subcellular localization of PstHa2a5-C27V (pK7WGF2 / SP+PstHa2a5 - C27S / *N. benthamiana*). The representative pictures out of 25. **a)** Structure of the construct; **b)** Pictures of subcellular localization assay: agro-infiltrated leaves were observed after 2 days post inoculation (dpi) under microscope (Leica DM4000B microscope / DFC 280 camera) under 40X magnification.

Table 3.1 Observed subcellular localization of different constructs

	Nucleus	Cytoplasm	Membrane	Apoplasm	Outside
Free GFP	Yes	Yes	No	No	No
Wild SP + PstHa2a5	No	Yes (partially)*	Maybe	?	Yes*
Wild PstHa2a5 (No SP)	Maybe	Yes (patially)	No	No	No*
PstHa2a5 C27S (N' GFP)	No	Yes (partially)	?	?	Yes**
PstHa2a5 C27S (C' GFP)	No	Yes (partially)	?	?	Yes**
PstHa2a5 C27V (N' GFP)	No	Yes (partially)	?	?	Yes**
PstHa2a5 C27V (C' GFP)	No	Yes (partially)	?	?	Yes**

(*) As expected; (**) Not expected; (?) Inconclusive

CHAPTER 4

CONCLUSION

In this thesis, a predicted *Puccinia striiformis* f. sp. *tritici* effector gene, PstHa2a5, was studied for its subcellular localization. The gene was chosen as its expression level is increased highly during invasion of the plant cell, following pathogen infection. The presence of signal peptide and many cysteine residues are also supporting characters for this protein to be a candidate effector. Moreover, PstHa2a5 gene contains the conserved FKC motif which suits to the (Y/F/W)XC motif found in rust pathogen candidate effectors.

First of all, the decision of generation of mutants was an important part of the experimental design. FKC motif was decided to be changed to FKS and FKV separately, by site directed mutagenesis. Site directed mutagenesis was performed in two steps by means of PCR. First step was to generate two fragments, which overlap at site of mutation while second step was to join the fragments to generate a whole gene sequence containing the mutation.

After generation of two mutants of PstHa2a5 gene with flanking CACC sequence at 5' end, the fragments were cloned first into donor vector (pENTR/D-TOPO), then to desired destination vector (pK7WGF2 and pK7FWG2), possessing GFP gene. Thus, any differences in intracellular localization while GFP is joined at N-terminus or C-terminus was investigated. Two mutant forms of the PstHa2a5 were compared with two wild type of the gene constructs, one consisting the signal peptide and one not. Two days post infiltrated leaves were examined under light microscope with appropriate filters (Leica

DM4000B microscope / DFC 280 camera) at 40X magnification to track GFP fluorescence, for investigating the subcellular localizations.

The subcellular localization of wild type PstHa2a5 gene was studied with constructs, which were designed and cloned by Bayantes Dagvadorj. Two types of wild type of PstHa2a5 gene, with and without signal peptide, to which GFP was fused at C-terminus of the gene was used. As expected the wild gene with signal peptide was secreted to the intercellular space, while the wild gene without signal peptide remained in the cell, localized in the cytoplasm and the nucleus.

In case of PstHa2a5-C27S mutant, where GFP was joined to either the N-terminus or the C-terminus, the protein was expressed inside cell end secreted outside of the cell, similar to wild type PstHa2a5 with signal peptide. Here it is deduced that C27S mutation did not cause distinctive difference in the localization of the gene comparing to wild type PstHa2a5 with signal peptide. It could be concluded that Cysteine 27 is not essential for the signal peptide function or Serine acted same or similarly as Cysteine, by leaving the function of signal peptide intact. If Serine can be used as replacement for Cysteine, it may suggest that Cysteine-27 does not form disulfide bond, as Serine cannot form disulfide bond. Maybe similar size of R-groups and similar polar characters of Cysteine and Serine enabled normal function of the signal peptide.

In case of PstHa2a5-C27V mutants, where GFP was joined to either the N-terminus or the C-terminus, the protein was expressed inside the cell, and was observed in cytoplasm and outside the cell. However, with this mutation the absence of localization to the plant cell nucleus was very clearly observed, suggesting that although Valine did not cause distinctive changes in the localization of the gene comparing to wild type PstHa2a5 with the signal peptide, it makes it easier for the protein to escape from the nucleus or not enter at all. The observed results may be due to conformational changes of protein caused by substitution of Cysteine-27 to Valine. The protein may have changed to more relaxed

conformation because of the mutation. If Cysteine-27 do not form disulphide bond, as implied by experiments with PstHa2a5-C27S mutants, then the polarity of the amino acid may be an important property of the protein, since Valine has similar size but non-polar R-group, comparing to polar amino acid, Cysteine, the change in polarity of 27^{th} Cysteine may be the explanation of observed difference in subcellular localization of the gene.

In order to verify the explanations of above observations, X-ray structure analyses are needed. However, prior to X-Ray analysis, some of the inconclusive microscopic analysis of the mutants due to lack of confocal microscopic images, thus for the future studies, in addition to observing subcellular beavers of the mutants by confocal microscopy but image analysis should be conducted in the presence of the membrane and nucleus markers.

The hypothesis of this thesis study was to assess the function of Cysteine residue on the conserved (W/F/Y)XC motif in the subcellular location of this effector by SDM. Based on the results, there is a need to re-examine the function of this motif in subcellular localization to predict the function of the effecter as a whole, by completely removing it, which can be achieved by deletion mutation.

REFERENCES

Briggle, L.W. (1980). Hydroxamic acid content of triticum species. *Euphytica*, *37* (3), 289-293.

Chen, X. M. (2005). Review / Synthèse Epidemiology and control of stripe rust [*Puccinia striiformis* f . sp . *tritici*] on wheat. *Canadian Journal of Plant Pathology*, *27*, 314-337.

Cohn, J., Sessa, G., & Martin, G. B. (2001). Innate immunity in plants. *Current Opinion in Immunology*, *13*, (1), 55 – 62. doi:10.1016/S0952-7915(00)00182-5.

Godfrey, D., Böhlenius, H., Pedersen, C., Zhang, Z., Emmersen, J., & Thordal-Christensen, H. (2010). Powdery mildew fungal effector candidates share N-terminal Y/F/WxC-motif. *BMC genomics*, *11*, 317. doi:10.1186/1471-2164-11-317.

Göhre, V., & Robatzek, S. (2008). Breaking the Barriers: Microbial Effector Molecules Subvert Plant Immunity. *Annu. Rev. Phytopathol.*, *46*, 189-215.doi:10.1146 / annurev.phyto.46.120407.110050.

Hartley, J. L. (2000) DNA Cloning Using In Vitro Site-Specific Recombination. *Genome Research*, *10*(11), 1788-1795. doi:10.1101/gr.143000.

Heun, M. (1997). Site of Einkorn Wheat Domestication Identified by DNA Fingerprinting. *Science*, *278*(5341), 1312 – 1314. doi:10.1126/science.278.5341.1312.

https://plantscientist.wordpress.com/2013/05/04/organism-of-the-week-aequorea-victoria-jellyfish/ [last accessed on 18.02.2015]

Karimi, M., Depicker, A., & Hilson, P. (2007). Recombinational Cloning with Plant Gateway Vectors 1. *Plant Physiology*, *145*(December), 1144 – 1154. doi:10.1104/pp.107. 106989.

Maddison, A. C., & Manners, J. G. (1972) Sunlight and viability of cereal rust uredospores. *Transaction of British Mycological Society*. *59*, 429 – 443.

Morel, J. B., & Dangl, J. L. (1997). The hypersensitive response and the induction of cell death in plants. *Cell Death and Differentiation,* 4, 671 – 683.

Opabode, J. (2006) Agrobacterium-mediated transformation of plants: emerging factors that influence efficiency. *Biotechnology and Molecular Biology Review*, *1*, 12 – 20.

Pieterse, C. M. J., Reyes, A. L., Ent, S. V. & Wees, S. C. M. V. (2009). Networking by small-molecule hormones in plant immunity. *Nature Chemical Biology*, *5*, 308 – 316.

Pitzschke, A., & Hirt, H. (2010) New insights into an old story: Agrobacterium-induced tumour formation in plants by plant transformation. *The EMBO Journal*, *29*(6), 1021-32. Nature Publishing Group. doi:10.1038/emboj.2010.8.

Primorse, S. B., Twyman, R. M., & Old, R. W. (2001). *Principles of gene manipulation*. 6th ed. Oxford: Blackwell publishing company. 132 – 138.

Rapilly, F. (1979). Yellow rust epidemiology. *Ann. Re. Phytopathol.*, *17*(68), 59 – 73.

Scheuer, P. M., Francisco, A. D., Miranda, M. Z. D., Ogliari, P. J., Torres, G., Limberger, V., Montenegro, F. M., Ruffi, C. R., Biondi, S. (2011). Characterization of Brazilian wheat cultivars for specific technological applications. *Ciência e Tecnologia de Alimentos*, *31*(3), 816-826. doi:10.1590/S0101-20612011000300041

Schröder, J., and K. Hassebrauk. 1964. Untersuchungen über die Keimung der Uredosporen des Gelbrostes (*Puccina striiformis* West). *Zentrablatt für Bakteriologie, Parasitenkunde, Infiktionskrankheiten und Hygiene*, II. 118:622-657

Shewry, P. R. (2009). Darwin review, *60*(6), 1537 – 1553. doi:10.1093/jxb/erp058

Stergiopoulos I., de Wit P. J. (2009). Fungal effector proteins. *Annu. Rev. Phytopathol.,* 47 233 – 263 10.1146/annurev.phyto.112408.132637

Wirthmueller, L., Maqbool, A., & Banfield, M. J. (2013). On the front line: structural insights into plant-pathogen interactions. *Nature Reviews Microbiology*, *11*, 761–776. doi:10.1038/nrmicro3118. Retrieved from: http://www.nature.com/nrmicro/journal/v11/n11/images/nrmicro3118-f1.jpg [last accessed on 20.01.2015]

Wydro, M., Kozubek, E., & Lehmann, P. (2006) Optimization of transient *Agrobacterium* mediated gene expression system in leaves of *Nicotiana benthamiana*. *Acta Biochimica Polonica*, *53*(2), 289-98.

Xu, R., & Li, Q. Q. (2008). Protocol: Streamline cloning of genes into binary vectors in Agrobacterium *via* the Gateway(R) TOPO vector system. *Plant methods*, *4*, 4. doi:10.1186/1746-4811-4-4

Yin, C., Chen, X., Wang, X., Han, Q., Kang, Z., & Hulbert, S. H. (2009). Generation and analysis of expression sequence tags from haustoria of the wheat stripe rust fungus Puccinia striiformis f. sp. Tritici. *BMC Genomics*, *10*, 626-634. doi:10.1186/1471-2164-10-626

APPENDIX-A

PSTHA2A5 GENE SEQUENCE

(Obtained from NCBI databases)

```
LOCUS       GH737102   488 bp   mRNA   linear   EST 05-JAN-2010
DEFINITION  PSTha2a5 Puccinia striiformis f. sp. tritici haustoria cDNA library
            Puccinia striiformis f. sp. tritici cDNA clone 2a5 5', mRNA
            sequence.
ACCESSION   GH737102
VERSION     GH737102.1  GI:222429011
DBLINK      BioSample: LIBEST 024285
KEYWORDS    EST.
SOURCE      Puccinia striiformis f. sp. tritici
  ORGANISM  Puccinia striiformis f. sp. tritici
            Eukaryota; Fungi; Dikarya; Basidiomycota; Pucciniomycotina;
            Pucciniomycetes; Pucciniales; Pucciniaceae; Puccinia.
REFERENCE   1  (bases 1 to 488)
  AUTHORS   Yin,C., Chen,X., Wang,X., Han,Q., Kang,Z. and Hulbert,S.
  TITLE     Generation and analysis of expression sequence tags from haustoria
            of the wheat stripe rust fungus Puccinia striiformis f. sp. Tritici
  JOURNAL   BMC Genomics 10 (1), 626 (2009)
   PUBMED   20028560
COMMENT     Contact: Hulbert SH
            Department of Plant Pathology
            Washington State University
            345 Johnson Hall, Pullman, WA 99164, USA
            Tel: 509 335 3722
            Fax: 509 335 9581
            Email: scot_hulbert@wsu.edu
            Seq primer: M13 Forward.
FEATURES             Location/Qualifiers
```

```
     source          1..488
                     /organism="Puccinia striiformis f. sp. tritici"
                     /mol_type="mRNA"
                     /strain="PST-78"
                     /db_xref="taxon:168172"
                     /clone="2a5"
                     /dev_stage="haustoria"
                     /clone_lib="LIBEST_024285 Puccinia striiformis f. sp.
                     tritici haustoria cDNA library"
                     /note="Organ: haustoria; Vector: pDNR-LIB; Site_1: Sfi I;
                     Site_2: Sfi I Forma specialis: tritici;"
ORIGIN
        1 gggcaggtac aacaactttt ccaaccactc ctcaaaaatt aaattcacat tatttgttta
       61 gaattacaca agatgcaaag cttcaacttc ttcatcgtat tcgcagtgtt gttgatcaac
      121 actcaattca tttctgtgaa gtcgttcaag tgtcccggtt tgcatggaac gccaagccaa
      181 acacatggtt attgcaccag atcaatcacc gatgaagaac gaaaggcaaa aaagattggc
      241 aaggagttca ccatgtggaa ggaagaaatc aagacagtcg acgggaaatt ctcgtgtgat
      301 aaagtggact tgaatgggtc ggttgccaca gatagcttct gttgtgacgt tgcaggtaga
      361 attggtgaag ttgagaaaag taaacaagct atgtggacaa acaactgctc caaagcatct
      421 tagggaatac caacattacc ttgctctgag gctggcccat gtcttcaaat tcgaccattc
      481 ccttttgg
```

APPENDIX-B

AMINO ACID SEQUENCE OF PSTHA2A5 GENE
(GH737102.1)

MQSFNFFIVFAVLLINTQFISVKSFKCPGLHGTPSQTHGYCTRSITDEERKAKKIGKEF
TMWKEEIKTVDGKFSCDKVDLNGSVATDSFCCDVAGRIGEVEKSKQAMWTNNCSKAS

APPENDIX-C

SEQUENCED PSTHA2A5-C27S MUTANT

>PstHa2a5-C27-forward; 684 bp.
TGCGCCCCGGGCAAGAAGGAATCCGCAATGTTGTTGATCACACTCATTCATTTCTGTGAAGTCG
TTCAGTCTCCCGGTTTGCATGGAACGCCAAGCCAAACACATGGTTATTGCACCAGATCAATCAC
CGATGAAGAACGAAAGGCAAAAAAGATTGGCAAGGAGTTCACCATGTGGAAGGAAGAAATCAAG
ACAGTCGACGGGAAATTCTCGTGTGATAAAGTGGACTTGAATGGGTCGGTTGCCACAGATAGCT
TCTGTTGTGACGTTGCAGGTAGAATTGGTGAAGTTGAGAAAAGTAAACAAGCTATGTGGACAAA
CAACTGCTCCAAAGCATCTTAGGCGGCGGCCGCGGAGCCTGCTTTTTTGTACAAACTTGTGATA
TCCTTGTACAGCTCGTCCATGCCGAGAGTGATCCCGGCGGCGGTCACGAACTCCAGCAGGACCA
TGTGATCGCGCTTCTCGTTGGGGTCTTTGCTCAGGGCGGACTGGGTGCTCAGGTAGTGGTTGTC
GGGCAGCAGCACGGGGCCGTCGCCGATGGGGGTGTTCTGCTGGTAGTGGTCGGCGAGCTGCACG
CTGCCGTCCTCGATGTTGTGGCGGATCTTGAAGTTCACCTTGATGCCGTTCTTCTGCTTGTCGG
CCATGATATAGACGTTGTGGCTGTTGTAGTTGTACTCCAGCTTG

> PstHa2a5-C27-reverse; 674 bp.
TGACTACATGGAAGTCATTGTTGTCCATAGCTTGTTTACTTTTCTCACTTCACCATTTCTCCAC
CTGCACGTCACAACAGAAGCTATCTGTGGCAACCGACCCATTCAAGTCCACTTTATCACACGAG
AATTTCCCGTCGACTGTCTTGATTTCTTCCTTCCACATGGTGAACTCCTTGCCAATCTTTTTTG
CCTTTCGTTCTTCATCGGTGATTGATCTGGTGCAATAACCATGTGTTTGGCTTGGCGTTCCATG
CAAACCGGGAGACTTGAACGACTTCACAGAAATGAATTGAGTGTTGATCAACAACACTGCGAAT
ACGATGAAGAAGTTGAAGCTTTGCATGGTGAAGGGTGGGCGCGCCGACCCAGCTTTCTTGTACA
AAGTGGTGATATCCCGCGGCCATGCTAGAGTCCGCAAAAATCACCAGTCTCTCTCTACAAATCT
ATCTCTCTCTATTTTTCTCCAGAATAATGTGTGAGTAGTTCCCAGATAAGGGAATTAGGGTTCT
TATAGGGTTTCGCTCATGTGTTGAGCATATAAGAAACCCTTAGTATGTATTTGTATTTGTAAAA
TACTTCTATCAATAAAATTTCTAATTCCTAAAACCAAAATCCAGTGACCTGCAGGCATGCGACG
TCGGGCCCTCTAGAGGATCCCCGGGTACCGCGAA

APPENDIX-D

pK7WGF2 GATEWAY DESTINATION VECTOR SEQUENCE

(Karimi *et al*., 2002)

>pK7WGF2,0 standard; circular DNA; 11876 BP.
CGACGTCGCATGCCTGCAGGTCACTGGATTTTGGTTTTAGGAATTAGAAATTTTATTGATA
GAAGTATTTTACAAATACAAATACATACTAAGGGTTTCTTATATGCTCAACACATGAG
CGAAACCCTATAAGAACCCTAATTCCCTTATCTGGGAACTACTCACACATTATTCTGG
AGAAAAATAGAGAGAGATAGATTTGTAGAGAGAGACTGGTGATTTTTGCGGACTCTAG
CATGGCCGCGGGATATCACCACTTTGTACAAGAAAGCTGAACGAGAAACGTAAAATGA
TATAAATATCAATATATTAAATTAGATTTTGCATAAAAAACAGACTACATAATACTGT
AAAACACAACATATCCAGTCACTATGGTCGACCTGCAGACTGGCTGTGTATAAGGGAG
CCTGACATTTATATTCCCCAGAACATCAGGTTAATGGCGTTTTTGATGTCATTTTCGC
GGTGGCTGAGATCAGCCACTTCTTCCCCGATAACGGAGACCGGCACACTGGCCATATC
GGTGGTCATCATGCGCCAGCTTTCATCCCCGATATGCACCACCGGGTAAAGTTCACGG
GAGACTTTATCTGACAGCAGACGTGCACTGGCCAGGGGGATCACCATCCGTCGCCCGG
GCGTGTCAATAATATCACTCTGTACATCCACAAACAGACGATAACGGCTCTCTCTTTT
ATAGGTGTAAACCTTAAACTGCATTTCACCAGTCCCTGTTCTCGTCAGCAAAAGAGCC
GTTCATTTCAATAAACCGGGCGACCTCAGCCATCCCTTCCTGATTTTCCGCTTTCCAG
CGTTCGGCACGCAGACGACGGGCTTCATTCTGCATGGTTGTGCTTACCAGACCGGAGA
TATTGACATCATATATGCCTTGAGCAACTGATAGCTGTCGCTGTCAACTGTCACTGTA
ATACGCTGCTTCATAGCACACCTCTTTTTGACATACTTCGGGTATACATATCAGTATA
TATTCTTATACCGCAAAAATCAGCGCGCAAATACGCATACTGTTATCTGGCTTTTAGT
AAGCCGGATCCACGCGTTTACGCCCCGCCCTGCCACTCATCGCAGTACTGTTGTAATT
CATTAAGCATTCTGCCGACATGGAAGCCATCACAGACGGCATGATGAACCTGAATCGC
CAGCGGCATCAGCACCTTGTCGCCTTGCGTATAATATTTGCCCATGGTGAAAACGGGG
GCGAAGAAGTTGTCCATATTGGCCACGTTTAAATCAAAACTGGTGAAACTCACCCAGG
GATTGGCTGAGACGAAAAACATATTCTCAATAAACCCTTTAGGGAAATAGGCCAGGTT
TTCACCGTAACACGCCACATCTTGCGAATATATGTGTAGAAACTGCCGGAAATCGTCG

TGGTATTCACTCCAGAGCGATGAAAACGTTTCAGTTTGCTCATGGAAAACGGTGTAAC
AAGGGTGAACACTATCCCATATCACCAGCTCACCGTCTTTCATTGCCATACGGAATTC
CGGATGAGCATTCATCAGGCGGGCAAGAATGTGAATAAAGGCCGGATAAAACTTGTGC
TTATTTTTCTTTACGGTCTTTAAAAAGGCCGTAATATCCAGCTGAACGGTCTGGTTAT
AGGTACATTGAGCAACTGACTGAAATGCCTCAAAATGTTCTTTACGATGCCATTGGGA
TATATCAACGGTGGTATATCCAGTGATTTTTTTCTCCATTTTAGCTTCCTTAGCTCCT
GAAAATCTCGCCGGATCCTAACTCAAAATCCACACATTATACGAGCCGGAAGCATAAA
GTGTAAAGCCTGGGGTGCCTAATGCGGCCGCCATAGTGACTGGATATGTTGTGTTTTA
CAGTATTATGTAGTCTGTTTTTTATGCAAAATCTAATTTAATATATTGATATTTATAT
CATTTTACGTTTCTCGTTCAGCTTTTTTGTACAAACTTGTGATATCCTTGTACAGCTC
GTCCATGCCGAGAGTGATCCCGGCGGCGGTCACGAACTCCAGCAGGACCATGTGATCG
CGCTTCTCGTTGGGGTCTTTGCTCAGGGCGGACTGGGTGCTCAGGTAGTGGTTGTCGG
GCAGCAGCACGGGGCCGTCGCCGATGGGGGGTGTTCTGCTGGTAGTGGTCGGCGAGCTG
CACGCTGCCGTCCTCGATGTTGTGGCGGATCTTGAAGTTCACCTTGATGCCGTTCTTC
TGCTTGTCGGCCATGATATAGACGTTGTGGCTGTTGTAGTTGTACTCCAGCTTGTGCC
CCAGGATGTTGCCGTCCTCCTTGAAGTCGATGCCCTTCAGCTCGATGCGGTTCACCAG
GGTGTCGCCCTCGAACTTCACCTCGGCGCGGGTCTTGTAGTTGCCGTCGTCCTTGAAG
AAGATGGTGCGCTCCTGGACGTAGCCTTCGGGCATGGCGGACTTGAAGAAGTCGTGCT
GCTTCATGTGGTCGGGGTAGCGGCTGAAGCACTGCACGCCGTAGGTCAGGGTGGTCAC
GAGGGTGGGCCAGGGCACGGGCAGCTTGCCGGTGGTGCAGATGAACTTCAGGGTCAGC
TTGCCGTAGGTGGCATCGCCCTCGCCCTCGCCGGACACGCTGAACTTGTGGCCGTTTA
CGTCGCCGTCCAGCTCGACCAGGATGGGCACCACCCCGGTGAACAGCTCCTCGCCCTT
GCTCACCATACTAGTGCGGCCGCCTGCAGGTCGACTAGAATAGTAAATTGTAATGTTG
TTTGTTGTTTGTTTTGTTGTGGTATTGTTGTAAAAATACCGGAGTCCTCTCCAAATGA
AATGAACTTCCTTATATAGAGGAAGGGTCTTGCGAAGGATAGTGGGATTGTGCGTCAT
CCCTTACGTCAGTGGAGATATCACATCAATCCACTTGCTTTGAAGACGTGGTTGGAAC
GTCTTCTTTTTCCACGATGCTCCTCGTGGGTGGGGGTCCATCTTTGGGACCACTGTCG
GCAGAGGCATCTTGAACGATAGCCTTTCCTTTATCGCAATGATGGCATTTGTAGGTGC
CACCTTCCTTTTCTACTGTCCTTTTGATGAAGTGACAGATAGCTGGGCAATGGAATCC
GAGGAGGTTTCCCGATATTACCCTTTGTTGAAAAGTCTCAATAGCCCTTTGGTCTTCT
GAGACTGTATCTTTGATATTCTTGGAGTAGACGAGAGTGTCGTGCTCCACCATGTTGA
CGAAGATTTTCTTCTTGTCATTGAGTCGTAAAAGACTCTGTATGAACTGTTCGCCAGT
CTTCACGGCGAGTTCTGTTAGATCCTCGATCTGAATTTTTGACTCCATGGCCTTTGAT
TCAGTAGGAACTACTTTCTTAGAGACTCCAATCTCTATTACTTGCCTTGGTTTATGAA
GCAAGCCTTGAATCGTCCATACTGGAATAGTACTTCTGATCTTGAGAAATATATCTTT
CTCTGTGTTCTTGATGCAGTTAGTCCTGAATCTTTTGACTGCATCTTTAACCTTCTTG
GGAAGGTATTTGATCTCCTGGAGATTATTACTCGGGTAGATCGTCTTGATGAGACCTG
CCGCGTAGGCCTCTCTAACCATCTGTGGGTCAGCATTCTTTCTGAAATTGAAGAGGCT
AATCTTCTCATTATCGGTGGTGAACATGGTATCGTCACCTTCTCCGTCGAACTTTCTT
CCTAGATCGTAGAGATAGAGAAAGTCGTCCATGGTGATCTCCGGGGCAAAGGAGATCA
GCTTGGCTCTAGTCGACCATATGGGAGAGCTCAAGCTTAGCTTGAGCTTGGATCAGAT
TGTCGTTTCCCGCCTTCAGTTTAAACTATCAGTGTTTGACAGGATATATTGGCGGGTA

AACCTAAGAGAAAAGAGCGTTTATTAGAATAACGGATATTTAAAAGGGCGTGAAAAGG
TTTATCCGTTCGTCCATTTGTATGTGCATGCCAACCACAGGGTTCCCCTCGGGATCAA
AGTACTTTGATCCAACCCCTCCGCTGCTATAGTGCAGTCGGCTTCTGACGTTCAGTGC
AGCCGTCTTCTGAAAACGACATGTCGCACAAGTCCTAAGTTACGCGACAGGCTGCCGC
CCTGCCCTTTTCCTGGCGTTTTCTTGTCGCGTGTTTTAGTCGCATAAAGTAGAATACT
TGCGACTAGAACCGGAGACATTACGCCATGAACAAGAGCGCCGCCGCTGGCCTGCTGG
GCTATGCCCGCGTCAGCACCGACGACCAGGACTTGACCAACCAACGGGCCGAACTGCA
CGCGGCCGGCTGCACCAAGCTGTTTTCCGAGAAGATCACCGGCACCAGGCGCGACCGC
CCGGAGCTGGCCAGGATGCTTGACCACCTACGCCCTGGCGACGTTGTGACAGTGACCA
GGCTAGACCGCCTGGCCCGCAGCACCCGCGACCTACTGGACATTGCCGAGCGCATCCA
GGAGGCCGGCGCGGGCCTGCGTAGCCTGGCAGAGCCGTGGGCCGACACCACCACGCCG
GCCGGCCGCATGGTGTTGACCGTGTTCGCCGGCATTGCCGAGTTCGAGCGTTCCCTAA
TCATCGACCGCACCCGGAGCGGGCGCGAGGCCGCCAAGGCCCGAGGCGTGAAGTTTGG
CCCCCGCCCTACCCTCACCCCGGCACAGATCGCGCACGCCCGCGAGCTGATCGACCAG
GAAGGCCGCACCGTGAAAGAGGCGGCTGCACTGCTTGGCGTGCATCGCTCGACCCTGT
ACCGCGCACTTGAGCGCAGCGAGGAAGTGACGCCCACCGAGGCCAGGCGGCGCGGTGC
CTTCCGTGAGGACGCATTGACCGAGGCCGACGCCCTGGCGGCCGCCGAGAATGAACGC
CAAGAGGAACAAGCATGAAACCGCACCAGGACGGCCAGGACGAACCGTTTTTCATTAC
CGAAGAGATCGAGGCGGAGATGATCGCGGCCGGGTACGTGTTCGAGCCGCCCGCGCAC
GTCTCAACCGTGCGGCTGCATGAAATCCTGGCCGGTTTGTCTGATGCCAAGCTGGCGG
CCTGGCCGGCCAGCTTGGCCGCTGAAGAAACCGAGCGCCGCCGTCTAAAAAGGTGATG
TGTATTTGAGTAAAACAGCTTGCGTCATGCGGTCGCTGCGTATATGATGCGATGAGTA
AATAAACAAATACGCAAGGGGAACGCATGAAGGTTATCGCTGTACTTAACCAGAAAGG
CGGGTCAGGCAAGACGACCATCGCAACCCATCTAGCCCGCGCCCTGCAACTCGCCGGG
GCCGATGTTCTGTTAGTCGATTCCGATCCCCAGGGCAGTGCCCGCGATTGGGCGGCCG
TGCGGGAAGATCAACCGCTAACCGTTGTCGGCATCGACCGCCCGACGATTGACCGCGA
CGTGAAGGCCATCGGCCGGCGCGACTTCGTAGTGATCGACGGAGCGCCCCAGGCGGCG
GACTTGGCTGTGTCCGCGATCAAGGCAGCCGACTTCGTGCTGATTCCGGTGCAGCCAA
GCCCTTACGACATATGGGCCACCGCCGACCTGGTGGAGCTGGTTAAGCAGCGCATTGA
GGTCACGGATGGAAGGCTACAAGCGGCCTTTGTCGTGTCGCGGGCGATCAAAGGCACG
CGCATCGGCGGTGAGGTTGCCGAGGCGCTGGCCGGGTACGAGCTGCCCATTCTTGAGT
CCCGTATCACGCAGCGCGTGAGCTACCCAGGCACTGCCGCCGCCGGCACAACCGTTCT
TGAATCAGAACCCGAGGGCGACGCTGCCCGCGAGGTCCAGGCGCTGGCCGCTGAAATT
AAATCAAAACTCATTTGAGTTAATGAGGTAAAGAGAAAATGAGCAAAAGCACAAACAC
GCTAAGTGCCGGCCGTCCGAGCGCACGCAGCAGCAAGGCTGCAACGTTGGCCAGCCTG
GCAGACACGCCAGCCATGAAGCGGGTCAACTTTCAGTTGCCGGCGGAGGATCACACCA
AGCTGAAGATGTACGCGGTACGCCAAGGCAAGACCATTACCGAGCTGCTATCTGAATA
CATCGCGCAGCTACCAGAGTAAATGAGCAAATGAATAAATGAGTAGATGAATTTTAGC
GGCTAAAGGAGGCGGCATGGAAAATCAAGAACAACCAGGCACCGACGCCGTGGAATGC
CCCATGTGTGGAGGAACGGGCGGTTGGCCAGGCGTAAGCGGCTGGGTTGTCTGCCGGC
CCTGCAATGGCACTGGAACCCCCAAGCCCGAGGAATCGGCGTGACGGTCGCAAACCAT
CCGGCCCGGTACAAATCGGCGCGGCGCTGGGTGATGACCTGGTGGAGAAGTTGAAGGC

CGCGCAGGCCGCCCAGCGGCAACGCATCGAGGCAGAAGCACGCCCCGGTGAATCGTGG
CAAGCGGCCGCTGATCGAATCCGCAAAGAATCCCGGCAACCGCCGGCAGCCGGTGCGC
CGTCGATTAGGAAGCCGCCCAAGGGCGACGAGCAACCAGATTTTTTCGTTCCGATGCT
CTATGACGTGGGCACCCGCGATAGTCGCAGCATCATGGACGTGGCCGTTTTCCGTCTG
TCGAAGCGTGACCGACGAGCTGGCGAGGTGATCCGCTACGAGCTTCCAGACGGGCACG
TAGAGGTTTCCGCAGGGCCGGCCGGCATGGCCAGTGTGTGGGATTACGACCTGGTACT
GATGGCGGTTTCCCATCTAACCGAATCCATGAACCGATACCGGGAAGGGAAGGGAGAC
AAGCCCGGCCGCGTGTTCCGTCCACACGTTGCGGACGTACTCAAGTTCTGCCGGCGAG
CCGATGGCGGAAAGCAGAAAGACGACCTGGTAGAAACCTGCATTCGGTTAAACACCAC
GCACGTTGCCATGCAGCGTACGAAGAAGGCCAAGAACGGCCGCCTGGTGACGGTATCC
GAGGGTGAAGCCTTGATTAGCCGCTACAAGATCGTAAAGAGCGAAACCGGGCGGCCGG
AGTACATCGAGATCGAGCTAGCTGATTGGATGTACCGCGAGATCACAGAAGGCAAGAA
CCCGGACGTGCTGACGGTTCACCCCGATTACTTTTTGATCGATCCCGGCATCGGCCGT
TTTCTCTACCGCCTGGCACGCCGCGCCGCAGGCAAGGCAGAAGCCAGATGGTTGTTCA
AGACGATCTACGAACGCAGTGGCAGCGCCGGAGAGTTCAAGAAGTTCTGTTTCACCGT
GCGCAAGCTGATCGGGTCAAATGACCTGCCGGAGTACGATTTGAAGGAGGAGGCGGGG
CAGGCTGGCCCGATCCTAGTCATGCGCTACCGCAACCTGATCGAGGGCGAAGCATCCG
CCGGTTCCTAATGTACGGAGCAGATGCTAGGGCAAATTGCCCTAGCAGGGGAAAAAGG
TCGAAAAGGTCTCTTTCCTGTGGATAGCACGTACATTGGGAACCCAAAGCCGTACATT
GGGAACCGGAACCCGTACATTGGGAACCCAAAGCCGTACATTGGGAACCGGTCACACA
TGTAAGTGACTGATATAAAAGAGAAAAAAGGCGATTTTTCCGCCTAAAACTCTTTAAA
ACTTATTAAAACTCTTAAAACCCGCCTGGCCTGTGCATAACTGTCTGGCCAGCGCACA
GCCGAAGAGCTGCAAAAAGCGCCTACCCTTCGGTCGCTGCGCTCCCTACGCCCCGCCG
CTTCGCGTCGGCCTATCGCGGCCGCTGGCCGCTCAAAAATGGCTGGCCTACGGCCAGG
CAATCTACCAGGGCGCGGACAAGCCGCGCCGTCGCCACTCGACCGCCGGCGCCCACAT
CAAGGCACCCTGCCTCGCGCGTTTCGGTGATGACGGTGAAAACCTCTGACACATGCAG
CTCCCGGAGACGGTCACAGCTTGTCTGTAAGCGGATGCCGGGAGCAGACAAGCCCGTC
AGGGCGCGTCAGCGGGTGTTGGCGGGTGTCGGGGCGCAGCCATGACCCAGTCACGTAG
CGATAGCGGAGTGTATACTGGCTTAACTATGCGGCATCAGAGCAGATTGTACTGAGAG
TGCACCATATGCGGTGTGAAATACCGCACAGATGCGTAAGGAGAAAATACCGCATCAG
GCGCTCTTCCGCTTCCTCGCTCACTGACTCGCTGCGCTCGGTCGTTCGGCTGCGGCGA
GCGGTATCAGCTCACTCAAAGGCGGTAATACGGTTATCCACAGAATCAGGGGATAACG
CAGGAAAGAACATGTGAGCAAAAGGCCAGCAAAAGGCCAGGAACCGTAAAAAGGCCGC
GTTGCTGGCGTTTTTCCATAGGCTCCGCCCCCCTGACGAGCATCACAAAAATCGACGC
TCAAGTCAGAGGTGGCGAAACCCGACAGGACTATAAAGATACCAGGCGTTTCCCCCTG
GAAGCTCCCTCGTGCGCTCTCCTGTTCCGACCCTGCCGCTTACCGGATACCTGTCCGC
CTTTCTCCCTTCGGGAAGCGTGGCGCTTTCTCATAGCTCACGCTGTAGGTATCTCAGT
TCGGTGTAGGTCGTTCGCTCCAAGCTGGGCTGTGTGCACGAACCCCCCGTTCAGCCCG
ACCGCTGCGCCTTATCCGGTAACTATCGTCTTGAGTCCAACCCGGTAAGACACGACTT
ATCGCCACTGGCAGCAGCCACTGGTAACAGGATTAGCAGAGCGAGGTATGTAGGCGGT
GCTACAGAGTTCTTGAAGTGGTGGCCTAACTACGGCTACACTAGAAGGACAGTATTTG
GTATCTGCGCTCTGCTGAAGCCAGTTACCTTCGGAAAAAGAGTTGGTAGCTCTTGATC

CGGCAAACAAACCACCGCTGGTAGCGGTGGTTTTTTTGTTTGCAAGCAGCAGATTACG
CGCAGAAAAAAAGGATCTCAAGAAGATCCTTTGATCTTTTCTACGGGGTCTGACGCTC
AGTGGAACGAAAACTCACGTTAAGGGATTTTGGTCATGCATGATATATCTCCCAATTT
GTGTAGGGCTTATTATGCACGCTTAAAAATAATAAAAGCAGACTTGACCTGATAGTTT
GGCTGTGAGCAATTATGTGCTTAGTGCATCTAATCGCTTGAGTTAACGCCGGCGAAGC
GGCGTCGGCTTGAACGAATTTCTAGCTAGACATTATTTGCCGACTACCTTGGTGATCT
CGCCTTTCACGTAGTGGACAAATTCTTCCAACTGATCTGCGCGCGAGGCCAAGCGATC
TTCTTCTTGTCCAAGATAAGCCTGTCTAGCTTCAAGTATGACGGGCTGATACTGGGCC
GGCAGGCGCTCCATTGCCCAGTCGGCAGCGACATCCTTCGGCGCGATTTTGCCGGTTA
CTGCGCTGTACCAAATGCGGGACAACGTAAGCACTACATTTCGCTCATCGCCAGCCCA
GTCGGGCGGCGAGTTCCATAGCGTTAAGGTTTCATTTAGCGCCTCAAATAGATCCTGT
TCAGGAACCGGATCAAAGAGTTCCTCCGCCGCTGGACCTACCAAGGCAACGCTATGTT
CTCTTGCTTTTGTCAGCAAGATAGCCAGATCAATGTCGATCGTGGCTGGCTCGAAGAT
ACCTGCAAGAATGTCATTGCGCTGCCATTCTCCAAATTGCAGTTCGCGCTTAGCTGGA
TAACGCCACGGAATGATGTCGTCGTGCACAACAATGGTGACTTCTACAGCGCGGAGAA
TCTCGCTCTCTCCAGGGGAAGCCGAAGTTTCCAAAAGGTCGTTGATCAAAGCTCGCCG
CGTTGTTTCATCAAGCCTTACGGTCACCGTAACCAGCAAATCAATATCACTGTGTGGC
TTCAGGCCGCCATCCACTGCGGAGCCGTACAAATGTACGGCCAGCAACGTCGGTTCGA
GATGGCGCTCGATGACGCCAACTACCTCTGATAGTTGAGTCGATACTTCGGCGATCAC
CGCTTCCCCCATGATGTTTAACTTTGTTTTAGGGCGACTGCCCTGCTGCGTAACATCG
TTGCTGCTCCATAACATCAAACATCGACCCACGGCGTAACGCGCTTGCTGCTTGGATG
CCCGAGGCATAGACTGTACCCCAAAAAAACATGTCATAACAAGAAGCCATGAAAACCG
CCACTGCGCCGTTACCACCGCTGCGTTCGGTCAAGGTTCTGGACCAGTTGCGTGACGG
CAGTTACGCTACTTGCATTACAGCTTACGAACCGAACGAGGCTTATGTCCACTGGGTT
CGTGCCCGAATTGATCACAGGCAGCAACGCTCTGTCATCGTTACAATCAACATGCTAC
CCTCCGCGAGATCATCCGTGTTTCAAACCCGGCAGCTTAGTTGCCGTTCTTCCGAATA
GCATCGGTAACATGAGCAAAGTCTGCCGCCTTACAACGGCTCTCCCGCTGACGCCGTC
CCGGACTGATGGGCTGCCTGTATCGAGTGGTGATTTTGTGCCGAGCTGCCGGTCGGGG
AGCTGTTGGCTGGCTGGTGGCAGGATATATTGTGGTGTAAACAAATTGACGCTTAGAC
AACTTAATAACACATTGCGGACGTTTTTAATGTACTGAATTAACGCCGAATTGAATTA
TCAGCTTGCATGCCGGTCGATCTAGTAACATAGATGACACCGCGCGCGATAATTTATC
CTAGTTTGCGCGCTATATTTTGTTTTCTATCGCGTATTAAATGTATAATTGCGGGACT
CTAATCATAAAAACCCATCTCATAAATAACGTCATGCATTACATGTTAATTATTACAT
GCTTAACGTAATTCAACAGAAATTATATGATAATCATCGCAAGACCGGCAACAGGATT
CAATCTTAAGAAACTTTATTGCCAAATGTTTGAACGATCTGCTTGACTCTAGCTAGAG
TCCGAACCCCAGAGTCCCGCTCAGAAGAACTCGTCAAGAAGGCGATAGAAGGCGATGC
GCTGCGAATCGGGAGCGGCGATACCGTAAAGCACGAGGAAGCGGTCAGCCCATTCGCC
GCCAAGCTCTTCAGCAATATCACGGGTAGCCAACGCTATGTCCTGATAGCGGTCCGCC
ACACCCAGCCGGCCACAGTCGATGAATCCAGAAAAGCGGCCATTTTCCACCATGATAT
TCGGCAAGCAGGCATCGCCGTGGGTCACGACGAGATCCTCGCCGTCGGGCATCCGCGC
CTTGAGCCTGGCGAACAGTTCGGCTGGCGCGAGCCCCTGATGCTCTTCGTCCAGATCA
TCCTGATCGACAAGACCGGCTTCCATCCGAGTACGTGCTCGCTCGATGCGATGTTTCG

CTTGGTGGTCGAATGGGCAGGTAGCCGGATCAAGCGTATGCAGCCGCCGCATTGCATC
AGCCATGATGGATACTTTCTCGGCAGGAGCAAGGTGAGATGACAGGAGATCCTGCCCC
GGCACTTCGCCCAATAGCAGCCAGTCCCTTCCCGCTTCAGTGACAACGTCGAGCACAG
CTGCGCAAGGAACGCCCGTCGTGGCCAGCCACGATAGCCGCGCTGCCTCGTCTTGGAG
TTCATTCAGGGCACCGGACAGGTCGGTCTTGACAAAAAGAACCGGGCGCCCCTGCGCT
GACAGCCGGAACACGGCGGCATCAGAGCAGCCGATTGTCTGTTGTGCCCAGTCATAGC
CGAATAGCCTCTCCACCCAAGCGGCCGGAGAACCTGCGTGCAATCCATCTTGTTCAAT
CATGCCTCGATCGAGTTGAGAGTGAATATGAGACTCTAATTGGATACCGAGGGGAATT
TATGGAACGTCAGTGGAGCATTTTTGACAAGAAATATTTGCTAGCTGATAGTGACCTT
AGGCGACTTTTGAACGCGCAATAATGGTTTCTGACGTATGTGCTTAGCTCATTAAACT
CCAGAAACCCGCGGCTGAGTGGCTCCTTCAACGTTGCGGTTCTGTCAGTTCCAAACGT
AAAACGGCTTGTCCCGCGTCATCGGCGGGGGTCATAACGTGACTCCCTTAATTCTCAT
GTATGATAATTCGAGGGTACCCGGGGATCCTCTAGAGGGCC

APPENDIX-E

pK7FWG2 GATEWAY DESTINATION VECTOR SEQUENCE

(Karimi *et al*., 2002)

pK7FWG2,0 standard; circular DNA; 11880 BP

CGACGTCGCATGCCTGCAGGTCACTGGATTTTGGTTTTAGGAATTAGAAATTTTATTG
ATAGAAGTATTTTACAAATACAAATACATACTAAGGGTTTCTTATATGCTCAACACAT
GAGCGAAACCCTATAAGAACCCTAATTCCCTTATCTGGGAACTACTCACACATTATTC
TGGAGAAAAATAGAGAGAGATAGATTTGTAGAGAGAGACTGGTGATTTTTGCGGACTC
TAGCATGGCCGCGGTTACTTGTACAGCTCGTCCATGCCGAGAGTGATCCCGGCGGCGG
TCACGAACTCCAGCAGGACCATGTGATCGCGCTTCTCGTTGGGGTCTTTGCTCAGGGC
GGACTGGGTGCTCAGGTAGTGGTTGTCGGGCAGCAGCACGGGGCCGTCGCCGATGGGG
GTGTTCTGCTGGTAGTGGTCGGCGAGCTGCACGCTGCCGTCCTCGATGTTGTGGCGGA
TCTTGAAGTTCACCTTGATGCCGTTCTTCTGCTTGTCGGCCATGATATAGACGTTGTG
GCTGTTGTAGTTGTACTCCAGCTTGTGCCCCAGGATGTTGCCGTCCTCCTTGAAGTCG
ATGCCCTTCAGCTCGATGCGGTTCACCAGGGTGTCGCCCTCGAACTTCACCTCGGCGC
GGGTCTTGTAGTTGCCGTCGTCCTTGAAGAAGATGGTGCGCTCCTGGACGTAGCCTTC
GGGCATGGCGGACTTGAAGAAGTCGTGCTGCTTCATGTGGTCGGGGTAGCGGCTGAAG
CACTGCACGCCGTAGGTCAGGGTGGTCACGAGGGTGGGCCAGGGCACGGGCAGCTTGC
CGGTGGTGCAGATGAACTTCAGGGTCAGCTTGCCGTAGGTGGCATCGCCCTCGCCCTC
GCCGGACACGCTGAACTTGTGGCCGTTTACGTCGCCGTCCAGCTCGACCAGGATGGGC
ACCACCCCGGTGAACAGCTCCTCGCCCTTGCTCACCATTGATATCACCACTTTGTACA
AGAAAGCTGAACGAGAAACGTAAAATGATATAAATATCAATATATTAAATTAGATTTT
GCATAAAAAACAGACTACATAATACTGTAAAACACAACATATCCAGTCACTATGGTCG
ACCTGCAGACTGGCTGTGTATAAGGGAGCCTGACATTTATATTCCCCAGAACATCAGG
TTAATGGCGTTTTTGATGTCATTTTCGCGGTGGCTGAGATCAGCCACTTCTTCCCCGA
TAACGGAGACCGGCACACTGGCCATATCGGTGGTCATCATGCGCCAGCTTTCATCCCC
GATATGCACCACCGGGTAAAGTTCACGGGAGACTTTATCTGACAGCAGACGTGCACTG
GCCAGGGGGATCACCATCCGTCGCCCGGGCGTGTCAATAATATCACTCTGTACATCCA

CAAACAGACGATAACGGCTCTCTCTTTTATAGGTGTAAACCTTAAACTGCATTTCACC
AGTCCCTGTTCTCGTCAGCAAAAGAGCCGTTCATTTCAATAAACCGGGCGACCTCAGC
CATCCCTTCCTGATTTTCCGCTTTCCAGCGTTCGGCACGCAGACGACGGGCTTCATTC
TGCATGGTTGTGCTTACCAGACCGGAGATATTGACATCATATATGCCTTGAGCAACTG
ATAGCTGTCGCTGTCAACTGTCACTGTAATACGCTGCTTCATAGCACACCTCTTTTTG
ACATACTTCGGGTATACATATCAGTATATATTCTTATACCGCAAAAATCAGCGCGCAA
ATACGCATACTGTTATCTGGCTTTTAGTAAGCCGGATCCACGCGTTTACGCCCCGCCC
TGCCACTCATCGCAGTACTGTTGTAATTCATTAAGCATTCTGCCGACATGGAAGCCAT
CACAGACGGCATGATGAACCTGAATCGCCAGCGGCATCAGCACCTTGTCGCCTTGCGT
ATAATATTTGCCCATGGTGAAAACGGGGGCGAAGAAGTTGTCCATATTGGCCACGTTT
AAATCAAAACTGGTGAAACTCACCCAGGGATTGGCTGAGACGAAAAACATATTCTCAA
TAAACCCTTTAGGGAAATAGGCCAGGTTTTCACCGTAACACGCCACATCTTGCGAATA
TATGTGTAGAAACTGCCGGAAATCGTCGTGGTATTCACTCCAGAGCGATGAAAACGTT
TCAGTTTGCTCATGGAAAACGGTGTAACAAGGGTGAACACTATCCCATATCACCAGCT
CACCGTCTTTCATTGCCATACGGAATTCCGGATGAGCATTCATCAGGCGGGCAAGAAT
GTGAATAAAGGCCGGATAAAACTTGTGCTTATTTTTCTTTACGGTCTTTAAAAAGGCC
GTAATATCCAGCTGAACGGTCTGGTTATAGGTACATTGAGCAACTGACTGAAATGCCT
CAAAATGTTCTTTACGATGCCATTGGGATATATCAACGGTGGTATATCCAGTGATTTT
TTTCTCCATTTTAGCTTCCTTAGCTCCTGAAAATCTCGCCGGATCCTAACTCAAAATC
CACACATTATACGAGCCGGAAGCATAAAGTGTAAAGCCTGGGGTGCCTAATGCGGCCG
CCATAGTGACTGGATATGTTGTGTTTTACAGTATTATGTAGTCTGTTTTTTATGCAAA
ATCTAATTTAATATATTGATATTTATATCATTTTACGTTTCTCGTTCAGCTTTTTTGT
ACAAACTTGTGATATCACTAGTGCGGCCGCCTGCAGGTCGACTAGAATAGTAAATTGT
AATGTTGTTTGTTGTTTGTTTTGTTGTGGTATTGTTGTAAAAATACCGGAGTCCTCTC
CAAATGAAATGAACTTCCTTATATAGAGGAAGGGTCTTGCGAAGGATAGTGGGATTGT
GCGTCATCCCTTACGTCAGTGGAGATATCACATCAATCCACTTGCTTTGAAGACGTGG
TTGGAACGTCTTCTTTTTCCACGATGCTCCTCGTGGGTGGGGGTCCATCTTTGGGACC
ACTGTCGGCAGAGGCATCTTGAACGATAGCCTTTCCTTTATCGCAATGATGGCATTTG
TAGGTGCCACCTTCCTTTTCTACTGTCCTTTTGATGAAGTGACAGATAGCTGGGCAAT
GGAATCCGAGGAGGTTTCCCGATATTACCCTTTGTTGAAAAGTCTCAATAGCCCTTTG
GTCTTCTGAGACTGTATCTTTGATATTCTTGGAGTAGACGAGAGTGTCGTGCTCCACC
ATGTTGACGAAGATTTTCTTCTTGTCATTGAGTCGTAAAAGACTCTGTATGAACTGTT
CGCCAGTCTTCACGGCGAGTTCTGTTAGATCCTCGATCTGAATTTTTGACTCCATGGC
CTTTGATTCAGTAGGAACTACTTTCTTAGAGACTCCAATCTCTATTACTTGCCTTGGT
TTATGAAGCAAGCCTTGAATCGTCCATACTGGAATAGTACTTCTGATCTTGAGAAATA
TATCTTTCTCTGTGTTCTTGATGCAGTTAGTCCTGAATCTTTTGACTGCATCTTTAAC
CTTCTTGGGAAGGTATTTGATCTCCTGGAGATTATTACTCGGGTAGATCGTCTTGATG
AGACCTGCCGCGTAGGCCTCTCTAACCATCTGTGGGTCAGCATTCTTTCTGAAATTGA
AGAGGCTAATCTTCTCATTATCGGTGGTGAACATGGTATCGTCACCTTCTCCGTCGAA
CTTTCTTCCTAGATCGTAGAGATAGAGAAAGTCGTCCATGGTGATCTCCGGGGCAAAG
GAGATCAGCTTGGCTCTAGTCGACCATATGGGAGAGCTCAAGCTTAGCTTGAGCTTGG
ATCAGATTGTCGTTTCCCGCCTTCAGTTTAAACTATCAGTGTTTGACAGGATATATTG

GCGGGTAAACCTAAGAGAAAAGAGCGTTTATTAGAATAACGGATATTTAAAAGGGCGT
GAAAAGGTTTATCCGTTCGTCCATTTGTATGTGCATGCCAACCACAGGGTTCCCCTCG
GGATCAAAGTACTTTGATCCAACCCCTCCGCTGCTATAGTGCAGTCGGCTTCTGACGT
TCAGTGCAGCCGTCTTCTGAAAACGACATGTCGCACAAGTCCTAAGTTACGCGACAGG
CTGCCGCCCTGCCCTTTTCCTGGCGTTTTCTTGTCGCGTGTTTTAGTCGCATAAAGTA
GAATACTTGCGACTAGAACCGGAGACATTACGCCATGAACAAGAGCGCCGCCGCTGGC
CTGCTGGGCTATGCCCGCGTCAGCACCGACGACCAGGACTTGACCAACCAACGGGCCG
AACTGCACGCGGCCGGCTGCACCAAGCTGTTTTCCGAGAAGATCACCGGCACCAGGCG
CGACCGCCCGGAGCTGGCCAGGATGCTTGACCACCTACGCCCTGGCGACGTTGTGACA
GTGACCAGGCTAGACCGCCTGGCCCGCAGCACCCGCGACCTACTGGACATTGCCGAGC
GCATCCAGGAGGCCGGCGCGGGCCTGCGTAGCCTGGCAGAGCCGTGGGCCGACACCAC
CACGCCGGCCGGCCGCATGGTGTTGACCGTGTTCGCCGGCATTGCCGAGTTCGAGCGT
TCCCTAATCATCGACCGCACCCGGAGCGGGCGCGAGGCCGCCAAGGCCCGAGGCGTGA
AGTTTGGCCCCCGCCCTACCCTCACCCCGGCACAGATCGCGCACGCCCGCGAGCTGAT
CGACCAGGAAGGCCGCACCGTGAAAGAGGCGGCTGCACTGCTTGGCGTGCATCGCTCG
ACCCTGTACCGCGCACTTGAGCGCAGCGAGGAAGTGACGCCCACCGAGGCCAGGCGGC
GCGGTGCCTTCCGTGAGGACGCATTGACCGAGGCCGACGCCCTGGCGGCCGCCGAGAA
TGAACGCCAAGAGGAACAAGCATGAAACCGCACCAGGACGGCCAGGACGAACCGTTTT
TCATTACCGAAGAGATCGAGGCGGAGATGATCGCGGCCGGGTACGTGTTCGAGCCGCC
CGCGCACGTCTCAACCGTGCGGCTGCATGAAATCCTGGCCGGTTTGTCTGATGCCAAG
CTGGCGGCCTGGCCGGCCAGCTTGGCCGCTGAAGAAACCGAGCGCCGCCGTCTAAAAA
GGTGATGTGTATTTGAGTAAAACAGCTTGCGTCATGCGGTCGCTGCGTATATGATGCG
ATGAGTAAATAAACAAATACGCAAGGGGAACGCATGAAGGTTATCGCTGTACTTAACC
AGAAAGGCGGGTCAGGCAAGACGACCATCGCAACCCATCTAGCCCGCGCCCTGCAACT
CGCCGGGGCCGATGTTCTGTTAGTCGATTCCGATCCCCAGGGCAGTGCCCGCGATTGG
GCGGCCGTGCGGGAAGATCAACCGCTAACCGTTGTCGGCATCGACCGCCCGACGATTG
ACCGCGACGTGAAGGCCATCGGCCGGCGCGACTTCGTAGTGATCGACGGAGCGCCCCA
GGCGGCGGACTTGGCTGTGTCCGCGATCAAGGCAGCCGACTTCGTGCTGATTCCGGTG
CAGCCAAGCCCTTACGACATATGGGCCACCGCCGACCTGGTGGAGCTGGTTAAGCAGC
GCATTGAGGTCACGGATGGAAGGCTACAAGCGGCCTTTGTCGTGTCGCGGGCGATCAA
AGGCACGCGCATCGGCGGTGAGGTTGCCGAGGCGCTGGCCGGGTACGAGCTGCCCATT
CTTGAGTCCCGTATCACGCAGCGCGTGAGCTACCCAGGCACTGCCGCCGCCGGCACAA
CCGTTCTTGAATCAGAACCCGAGGGCGACGCTGCCCGCGAGGTCCAGGCGCTGGCCGC
TGAAATTAAATCAAAACTCATTTGAGTTAATGAGGTAAAGAGAAAATGAGCAAAAGCA
CAAACACGCTAAGTGCCGGCCGTCCGAGCGCACGCAGCAGCAAGGCTGCAACGTTGGC
CAGCCTGGCAGACACGCCAGCCATGAAGCGGGTCAACTTTCAGTTGCCGGCGGAGGAT
CACACCAAGCTGAAGATGTACGCGGTACGCCAAGGCAAGACCATTACCGAGCTGCTAT
CTGAATACATCGCGCAGCTACCAGAGTAAATGAGCAAATGAATAAATGAGTAGATGAA
TTTTAGCGGCTAAAGGAGGCGGCATGGAAAATCAAGAACAACCAGGCACCGACGCCGT
GGAATGCCCCATGTGTGGAGGAACGGGCGGTTGGCCAGGCGTAAGCGGCTGGGTTGTC
TGCCGGCCCTGCAATGGCACTGGAACCCCCAAGCCCGAGGAATCGGCGTGACGGTCGC
AAACCATCCGGCCCGGTACAAATCGGCGCGGCGCTGGGTGATGACCTGGTGGAGAAGT

TGAAGGCCGCGCAGGCCGCCCAGCGGCAACGCATCGAGGCAGAAGCACGCCCCGGTGA
ATCGTGGCAAGCGGCCGCTGATCGAATCCGCAAAGAATCCCGGCAACCGCCGGCAGCC
GGTGCGCCGTCGATTAGGAAGCCGCCCAAGGGCGACGAGCAACCAGATTTTTTCGTTC
CGATGCTCTATGACGTGGGCACCCGCGATAGTCGCAGCATCATGGACGTGGCCGTTTT
CCGTCTGTCGAAGCGTGACCGACGAGCTGGCGAGGTGATCCGCTACGAGCTTCCAGAC
GGGCACGTAGAGGTTTCCGCAGGGCCGGCCGGCATGGCCAGTGTGTGGGATTACGACC
TGGTACTGATGGCGGTTTCCCATCTAACCGAATCCATGAACCGATACCGGGAAGGGAA
GGGAGACAAGCCCGGCCGCGTGTTCCGTCCACACGTTGCGGACGTACTCAAGTTCTGC
CGGCGAGCCGATGGCGGAAAGCAGAAAGACGACCTGGTAGAAACCTGCATTCGGTTAA
ACACCACGCACGTTGCCATGCAGCGTACGAAGAAGGCCAAGAACGGCCGCCTGGTGAC
GGTATCCGAGGGTGAAGCCTTGATTAGCCGCTACAAGATCGTAAAGAGCGAAACCGGG
CGGCCGGAGTACATCGAGATCGAGCTAGCTGATTGGATGTACCGCGAGATCACAGAAG
GCAAGAACCCGGACGTGCTGACGGTTCACCCCGATTACTTTTTGATCGATCCCGGCAT
CGGCCGTTTTCTCTACCGCCTGGCACGCCGCGCCGCAGGCAAGGCAGAAGCCAGATGG
TTGTTCAAGACGATCTACGAACGCAGTGGCAGCGCCGGAGAGTTCAAGAAGTTCTGTT
TCACCGTGCGCAAGCTGATCGGGTCAAATGACCTGCCGGAGTACGATTTGAAGGAGGA
GGCGGGGCAGGCTGGCCCGATCCTAGTCATGCGCTACCGCAACCTGATCGAGGGCGAA
GCATCCGCCGGTTCCTAATGTACGGAGCAGATGCTAGGGCAAATTGCCCTAGCAGGGG
AAAAAGGTCGAAAAGGTCTCTTTCCTGTGGATAGCACGTACATTGGGAACCCAAAGCC
GTACATTGGGAACCGGAACCCGTACATTGGGAACCCAAAGCCGTACATTGGGAACCGG
TCACACATGTAAGTGACTGATATAAAAGAGAAAAAAGGCGATTTTTCCGCCTAAAACT
CTTTAAAACTTATTAAAACTCTTAAAACCCGCCTGGCCTGTGCATAACTGTCTGGCCA
GCGCACAGCCGAAGAGCTGCAAAAAGCGCCTACCCTTCGGTCGCTGCGCTCCCTACGC
CCCGCCGCTTCGCGTCGGCCTATCGCGGCCGCTGGCCGCTCAAAAATGGCTGGCCTAC
GGCCAGGCAATCTACCAGGGCGCGGACAAGCCGCGCCGTCGCCACTCGACCGCCGGCG
CCCACATCAAGGCACCCTGCCTCGCGCGTTTCGGTGATGACGGTGAAAACCTCTGACA
CATGCAGCTCCCGGAGACGGTCACAGCTTGTCTGTAAGCGGATGCCGGGAGCAGACAA
GCCCGTCAGGGCGCGTCAGCGGGTGTTGGCGGGTGTCGGGGCGCAGCCATGACCCAGT
CACGTAGCGATAGCGGAGTGTATACTGGCTTAACTATGCGGCATCAGAGCAGATTGTA
CTGAGAGTGCACCATATGCGGTGTGAAATACCGCACAGATGCGTAAGGAGAAAATACC
GCATCAGGCGCTCTTCCGCTTCCTCGCTCACTGACTCGCTGCGCTCGGTCGTTCGGCT
GCGGCGAGCGGTATCAGCTCACTCAAAGGCGGTAATACGGTTATCCACAGAATCAGGG
GATAACGCAGGAAAGAACATGTGAGCAAAAGGCCAGCAAAAGGCCAGGAACCGTAAAA
AGGCCGCGTTGCTGGCGTTTTTCCATAGGCTCCGCCCCCCTGACGAGCATCACAAAAA
TCGACGCTCAAGTCAGAGGTGGCGAAACCCGACAGGACTATAAAGATACCAGGCGTTT
CCCCCTGGAAGCTCCCTCGTGCGCTCTCCTGTTCCGACCCTGCCGCTTACCGGATACC
TGTCCGCCTTTCTCCCTTCGGGAAGCGTGGCGCTTTCTCATAGCTCACGCTGTAGGTA
TCTCAGTTCGGTGTAGGTCGTTCGCTCCAAGCTGGGCTGTGTGCACGAACCCCCCGTT
CAGCCCGACCGCTGCGCCTTATCCGGTAACTATCGTCTTGAGTCCAACCCGGTAAGAC
ACGACTTATCGCCACTGGCAGCAGCCACTGGTAACAGGATTAGCAGAGCGAGGTATGT
AGGCGGTGCTACAGAGTTCTTGAAGTGGTGGCCTAACTACGGCTACACTAGAAGGACA
GTATTTGGTATCTGCGCTCTGCTGAAGCCAGTTACCTTCGGAAAAAGAGTTGGTAGCT

CTTGATCCGGCAAACAAACCACCGCTGGTAGCGGTGGTTTTTTTGTTTGCAAGCAGCA
GATTACGCGCAGAAAAAAAGGATCTCAAGAAGATCCTTTGATCTTTTCTACGGGGTCT
GACGCTCAGTGGAACGAAAACTCACGTTAAGGGATTTTGGTCATGCATGATATATCTC
CCAATTTGTGTAGGGCTTATTATGCACGCTTAAAAATAATAAAAGCAGACTTGACCTG
ATAGTTTGGCTGTGAGCAATTATGTGCTTAGTGCATCTAATCGCTTGAGTTAACGCCG
GCGAAGCGGCGTCGGCTTGAACGAATTTCTAGCTAGACATTATTTGCCGACTACCTTG
GTGATCTCGCCTTTCACGTAGTGGACAAATTCTTCCAACTGATCTGCGCGCGAGGCCA
AGCGATCTTCTTCTTGTCCAAGATAAGCCTGTCTAGCTTCAAGTATGACGGGCTGATA
CTGGGCCGGCAGGCGCTCCATTGCCCAGTCGGCAGCGACATCCTTCGGCGCGATTTTG
CCGGTTACTGCGCTGTACCAAATGCGGGACAACGTAAGCACTACATTTCGCTCATCGC
CAGCCCAGTCGGGCGGCGAGTTCCATAGCGTTAAGGTTTCATTTAGCGCCTCAAATAG
ATCCTGTTCAGGAACCGGATCAAAGAGTTCCTCCGCCGCTGGACCTACCAAGGCAACG
CTATGTTCTCTTGCTTTTGTCAGCAAGATAGCCAGATCAATGTCGATCGTGGCTGGCT
CGAAGATACCTGCAAGAATGTCATTGCGCTGCCATTCTCCAAATTGCAGTTCGCGCTT
AGCTGGATAACGCCACGGAATGATGTCGTCGTGCACAACAATGGTGACTTCTACAGCG
CGGAGAATCTCGCTCTCTCCAGGGGAAGCCGAAGTTTCCAAAAGGTCGTTGATCAAAG
CTCGCCGCGTTGTTTCATCAAGCCTTACGGTCACCGTAACCAGCAAATCAATATCACT
GTGTGGCTTCAGGCCGCCATCCACTGCGGAGCCGTACAAATGTACGGCCAGCAACGTC
GGTTCGAGATGGCGCTCGATGACGCCAACTACCTCTGATAGTTGAGTCGATACTTCGG
CGATCACCGCTTCCCCCATGATGTTTAACTTTGTTTTAGGGCGACTGCCCTGCTGCGT
AACATCGTTGCTGCTCCATAACATCAAACATCGACCCACGGCGTAACGCGCTTGCTGC
TTGGATGCCCGAGGCATAGACTGTACCCCAAAAAAACATGTCATAACAAGAAGCCATG
AAAACCGCCACTGCGCCGTTACCACCGCTGCGTTCGGTCAAGGTTCTGGACCAGTTGC
GTGACGGCAGTTACGCTACTTGCATTACAGCTTACGAACCGAACGAGGCTTATGTCCA
CTGGGTTCGTGCCCGAATTGATCACAGGCAGCAACGCTCTGTCATCGTTACAATCAAC
ATGCTACCCTCCGCGAGATCATCCGTGTTTCAAACCCGGCAGCTTAGTTGCCGTTCTT
CCGAATAGCATCGGTAACATGAGCAAAGTCTGCCGCCTTACAACGGCTCTCCCGCTGA
CGCCGTCCCGGACTGATGGGCTGCCTGTATCGAGTGGTGATTTTGTGCCGAGCTGCCG
GTCGGGGAGCTGTTGGCTGGCTGGTGGCAGGATATATTGTGGTGTAAACAAATTGACG
CTTAGACAACTTAATAACACATTGCGGACGTTTTTAATGTACTGAATTAACGCCGAAT
TGAATTATCAGCTTGCATGCCGGTCGATCTAGTAACATAGATGACACCGCGCGCGATA
ATTTATCCTAGTTTGCGCGCTATATTTTGTTTTCTATCGCGTATTAAATGTATAATTG
CGGGACTCTAATCATAAAAACCCATCTCATAAATAACGTCATGCATTACATGTTAATT
ATTACATGCTTAACGTAATTCAACAGAAATTATATGATAATCATCGCAAGACCGGCAA
CAGGATTCAATCTTAAGAAACTTTATTGCCAAATGTTTGAACGATCTGCTTGACTCTA
GCTAGAGTCCGAACCCCAGAGTCCCGCTCAGAAGAACTCGTCAAGAAGGCGATAGAAG
GCGATGCGCTGCGAATCGGGAGCGGCGATACCGTAAAGCACGAGGAAGCGGTCAGCCC
ATTCGCCGCCAAGCTCTTCAGCAATATCACGGGTAGCCAACGCTATGTCCTGATAGCG
GTCCGCCACACCCAGCCGGCCACAGTCGATGAATCCAGAAAAGCGGCCATTTTCCACC
ATGATATTCGGCAAGCAGGCATCGCCGTGGGTCACGACGAGATCCTCGCCGTCGGGCA
TCCGCGCCTTGAGCCTGGCGAACAGTTCGGCTGGCGCGAGCCCCTGATGCTCTTCGTC
CAGATCATCCTGATCGACAAGACCGGCTTCCATCCGAGTACGTGCTCGCTCGATGCGA

TGTTTCGCTTGGTGGTCGAATGGGCAGGTAGCCGGATCAAGCGTATGCAGCCGCCGCA
TTGCATCAGCCATGATGGATACTTTCTCGGCAGGAGCAAGGTGAGATGACAGGAGATC
CTGCCCCGGCACTTCGCCCAATAGCAGCCAGTCCCTTCCCGCTTCAGTGACAACGTCG
AGCACAGCTGCGCAAGGAACGCCCGTCGTGGCCAGCCACGATAGCCGCGCTGCCTCGT
CTTGGAGTTCATTCAGGGCACCGGACAGGTCGGTCTTGACAAAAAGAACCGGGCGCCC
CTGCGCTGACAGCCGGAACACGGCGGCATCAGAGCAGCCGATTGTCTGTTGTGCCCAG
TCATAGCCGAATAGCCTCTCCACCCAAGCGGCCGGAGAACCTGCGTGCAATCCATCTT
GTTCAATCATGCCTCGATCGAGTTGAGAGTGAATATGAGACTCTAATTGGATACCGAG
GGGAATTTATGGAACGTCAGTGGAGCATTTTTGACAAGAAATATTTGCTAGCTGATAG
TGACCTTAGGCGACTTTTGAACGCGCAATAATGGTTTCTGACGTATGTGCTTAGCTCA
TTAAACTCCAGAAACCCGCGGCTGAGTGGCTCCTTCAACGTTGCGGTTCTGTCAGTTC
CAAACGTAAAACGGCTTGTCCCGCGTCATCGGCGGGGGTCATAACGTGACTCCCTTAA
TTCTCATGTATGATAATTCGAGGGTACCCGGGGATCCTCTAGAGGGCC