Universitetet i Stavanger

Faculty of Science and Technology

MASTER'S THESIS

Study program/specialization: Biological Chemistry	Spring semester, 2016 Open
Writer: Sindre Skjæveland	…… (Writer's signature)
Faculty supervisor: Cathrine Lillo External supervisor(s):	
Thesis title: Protein phosphatase 4 recombinant protein – transformation of bacteria and *Arabidopsis thaliana*	
Credits (ECTS): 60	
Key words: PP4-2, PP4-1, Protein phosphatase, *Arabidopsis thaliana*, recombinant protein	Pages: 68 + enclosure: Appendix (30 pages), CD Stavanger, 15.06.2016 Date/year

Protein phosphatase 4 recombinant protein – transformation of bacteria and *Arabidopsis thaliana*

University of Stavanger
Faculty of Science and Technology
Master's degree in Biological Chemistry

June, 2016

Sindre Skjæveland

Abstract

Protein phosphatase 4 (PP4) is well conserved in eukaryote organisms. The protein has been studied in yeast and mammalian cells, but little is known about PP4 in plants. The function of PP4 in plants is unknown and it will not have the same function as in the other studied organisms because the processes does not exist in plants. It is likely that the PP4 proteins has been adapted for other uses in plants.

In this thesis project, I worked toward making *Arabidopsis thaliana* plants with the two PP4 proteins PP4-1 and PP4-2 using a variety of recombination technologies. Plants were successfully transformed with PP4-2 and PP4-1. The plants with PP4-2 made seeds that have a chance of having the PP4-2 gene themselves. These plants can be used for further study of PP4-2 in plants.

As the practical part of the thesis work ended, the plants transformed with the PP4-1 gene were starting to produce seeds that later can be selected for the gene and grow to produce more seeds. These can potentially be used in further studies.

I also demonstrated that the PP4-2 protein could be produced with *E. coli* Rosetta strain. The protein was visualized on a polyacrylamide gel.

The PP4-1 has been successfully cloned into a pMAL-C2X vector that can be used for further studies.

Acknowledgements

I would like to thank my supervisor Cathrine Lillo for her guidance and advice.

I thank Maria Therese Chreighton and Behzad Heidari for their guidance and help in the lab.

I am grateful for all the help and advice I have received at CORE by the people there.

Abbreviations

A. tumefaciens	*Agrobacterium tumefaciens*
bp	base pair(s)
DNA	Deoxyribonucleic acid
E. coli	*Escherichia coli*
IPTG	Isopropyl β-D-1-thiogalactopyranoside
LB	Lysogeny broth
MBP	Maltose-binding protein
PCR	Polymerase Chain Reaction
PP4	Protein phosphatase 4

Table of contents

List of figures

List of tables

Appendix list

Introduction

Protein phosphorylation

Protein phosphorylation is a post-translational modification of a protein and is a regulatory mechanism in many metabolic processes. Phosphorylation is the addition of a phosphate group to a molecule. Kinases are enzymes that catalyze phosphorylation reactions. In eukaryote organisms, it is most common that the amino acid residues serine, threonine and tyrosine are phosphorylated. (Berg, Tymoczko, Stryer, 2012, p 308).

Protein phosphatases

Protein phosphatases are enzymes that reverses phosphorylation by catalyzing the removal of the phosphoryl groups (Berg et al., 2012, p 310).

PP4

Protein phosphatase 4 is a group consisting of PP4-1 and PP4-2. They are serine/threonine phosphatases belonging to the type 2A phosphatase subfamily (Pujol, et al., 2000) These catalytic subunits are found as dimeric or trimeric complexes. Their activity is regulated by scaffolding and regulatory subunits. PP4 highly conserved in eukaryotes, but have different function. For mammals, PP4 is shown to be essential for the development of a group of white blood cells. PP4 can not have the same function in plants as in mammalian cells, and because the genes are highly conserved it is likely have an important function (Lillo et al., 2014)

PP4-1

The sequence of PP4-1 is listed in Appendix A-1.

PP4-2

The sequence of PP4-2 is listed in Appendix A-2.

Materials and methods

Materials

LB agar

Petri dishes with LB agar was used for growth of *E. coli* and *A. tumefaciens*. The LB agar was prepared by mixing dry LB agar with water and then autoclaving it. Any antibiotic needed for an experiment was added once the LB agar had cooled so the heat would not destroy the antibiotic. The agar was poured into Petri dishes and left to solidify.

LB broth

LB broth was used to grow over-night cultures of *E. coli* and *A. tumefaciens*. Dry LB broth was mixed with water and autoclaved. Antibiotic was added together with the bacteria to be incubated.

List of antibiotics

- Kanamycin
- Chloramphenicol
- Ampicillin
- Hygromycin B

Vectors

pETMBP1a

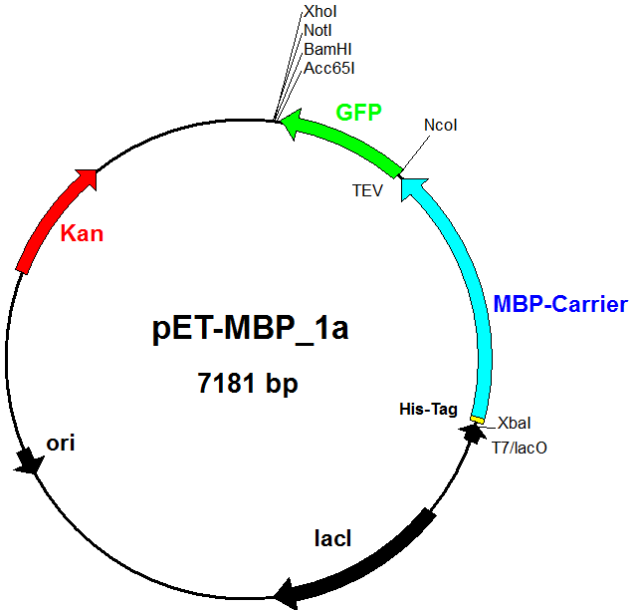

Figure 1:pETMBP1a

pENTR™/D-TOPO®

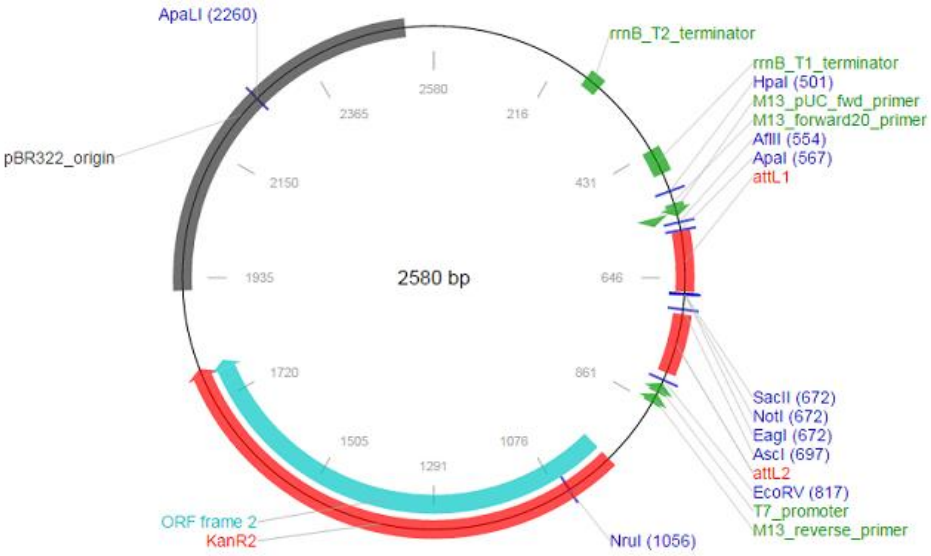

Figure 2: pENTR/D-TOPO vector (addgene.org/vector-database/2519/)

pMAL-C2X

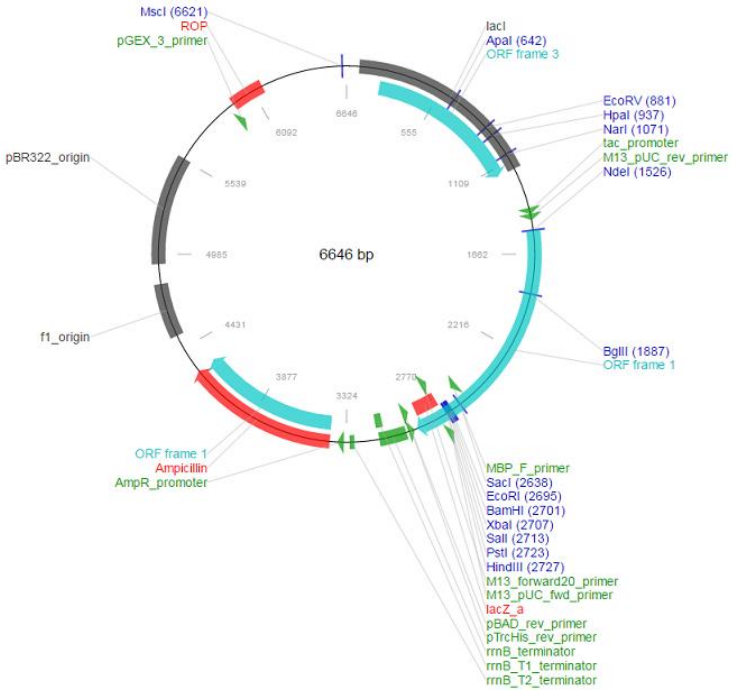

Figure 3: pMAL-C2X (addgene.org/vector-database/3509/)

pSF

The pSF vector is made from pGWB2 vector with the gateway cassette cut out from pDEST/N-SF-TAP and inserted into the pGWB2. pSF have Strep-FLAG-tag. pSF has kanamycin and hygromycin resistance.

Table 1: Primers, 5'-3' direction, F=Forward, R=Reverse

Primer name	Sequence	Restriction enzyme	Destination vector
PP4.2_MBP_F	CACCATGGCAATGTCAGACCTAGACAAGCA	NcoI	pET-MBP1a
PP4.2_MBP_R	CAAGGTACCTCACAGGAAATAATCAGGTGC	Acc65I	pET-MBP1a
PP4.1_MBP_F	CAGGATCCATGTCAGACCTAGATCGGCAAATAG	BamHI	pMAL-C2X
PP4.1_MBP_R	CAAAGCTTTTATAGGAAGTAATCAGGGGC	HindIII	pMAL-C2X
PP4.2_dTOPO_F	CACCATGTCAGACCTAGACAAGCA		pENTR-D-TOPO
PP4.2_dTOPO_R	TCACAGGAAATAATCAGGTGCA		pENTR-D-TOPO
PP4.1_dTOPO_F	CACCATGTCAGACCTAGATCGGCA		pENTR-D-TOPO
PP4.1_dTOPO_R	TTATAGGAAGTAATCAGGGGCC		pENTR-D-TOPO
M13 F	GTAAAACGACGGCCAG		
M13 R	CAGGAAACAGCTATGAC		
pSF F	ATGGATTATAAAGATGATGATG		pSF
pSF R	TTGCGGGACTCTAATCATAAAAA		pSF

Table 2: Restriction enzymes and their cut sites

Restriction enzyme	Cut sites
NcoI	5'…C^CATGG…3' 3'…GGTAC^C…5'
Acc65I	5'…G^GTACC…3' 3'…CCATG^G…5'
BamHI	5'…G^GATCC…3' 3'…CCTAG^G…5'
HindIII	5'…A^AGCTT…3' 3'…TTCGA^A…5'
ApaI	5'…GGGCC^C…3' 3'…C^CCGGG…5'

Table 3: Bacteria

Bacterial strain	Key features
JM109 cells (*E. coli*)	Minimizes recombination, higher quality of plasmid DNA, sensitive to all common antibiotics
Rosetta (*E. coli*)	Used for expression of eukaryotic protein, has chloramphenicol resistance
DH5α (*E. coli*)	Increases insert stability
ABI (*A. tumefaciens*)	Can transform plants

Glycerol

To preserve bacterial solutions long-term, 20 % glycerol was mixed in a 50/50 ratio with the bacterial solution. The mix was then frozen in -80° C.

Methods

Transformation of *E. coli*

To incorporate vectors into cells, transformation was done. Each vector (0.5 – 1 µl) were mixed gently with an Eppendorf tube with JM109 competent cells/Rosetta competent cells/DH5 α competent cells recently thawed from a freezer. The cells were heat shocked at 42° C for 1.5 min and put on ice for 5 min. LB broth (400 µl) was added and the tubes with the cells were placed on a shaker for 1 h. The cells were spread on plates with antibiotic and incubated at 37° C over night.

Transformation for *A. tumefaciens*

An Eppendorf tube with *Agrobacterium tumefaciens* was retrieved from a freezer and thawed on ice for 10 min. DNA (1 µg) was added, gently mixed and kept on ice for 10 min. The solution was submerged in liquid nitrogen to freeze. The bacteria were heat shocked in a water bath at 37° C for 15 min. The tube with bacteria was put briefly on ice. The solution was added to LB broth (500 µl) and shaken at 28° C for 3 h.

Polymerase Chain Reactions

To amplify genes, polymerase chain reaction (PCR) was used. Expand High Fidelity^{PLUS} PCR System (Roche) was used to amplify genes that needed to be amplified accurately. It is a

blend of Taq DNA Polymerase and a proofreading protein that gives it a sixfold greater replicational accuracy than with only Taq DNA Polymerase.

Table 4: Reagents, volumes and final concentrations for one PCR using the Expand High Fidelity^{PLUS} PCR System

Reagents	Volume	Final concentration
PCR-grade water	Up to 50 µl total volume	
Expand HiFi^{PLUS} Reaction Buffer, 5x (+Mg)	10 µl	1x
Nucleotides (10 mM)	1 µl	200 µM
Forward primer (10 µM)	2 µl	0.4 µM
Reverse primer (10 µM)	2 µl	0.4 µM
Template DNA	1 µl	
Expand HiFi^{PLUS} Enzyme Blend	0.5 µl	

Table 5: Thermal cycler program used for the Expand High Fidelity^{PLUS} PCR System

Step	Temperature	Time
Initial denaturation	94° C	2 min
Denaturation	94° C	30 s
Annealing	55° C	30 s
Elongation	72° C	3 min
Final elongation	72° C	7 min

For amplification of genes with a blunt-end that could be used in TOPO-cloning, amplification with the Phusion system (Fermentas) was attempted. The polymerase in the Phusion system can proofread. A Pfu system (Fermentas) was also used because the Phusion system did not work well. The polymerase in the Pfu system can proofread.

Table 6: Reagents and volumes for one PCR using the Phusion system

Reagents	Volume	Final concentration
PCR-grade water	Up to 50 µl total volume	
Phusion HF Buffer, 5x (+ Mg)	10 µl	1x
Nucleotides (10 mM)	1 µl	200 µM
Forward primer (10 µM)	2 µl	0.4 µM
Reverse primer (10 µM)	2 µl	0.4 µM
Template DNA	1-2 µl	
Phusion DNA Polymerase	0.5 µl	

Table 7: Thermal cycler program used for PCR with the Phusion system

Step	Temperature	Time
Initial denaturation	98° C	30 s
Denaturation	98° C	10 s
Annealing	55° C	30 s
Elongation	72° C	1.5 min
Final elongation	72° C	3 min

Table 8: Reagents and volumes for one PCR using the Pfu system

Reagents	Volume	Final concentration
PCR-grade water	Up to 50 µl total volume	
Pfu Buffer, 10x (+$MgSO_4$)	5 µl	1x
Nucleotides (10 mM)	1 µl	200 µM
Forward primer (10 µM)	2 µl	0.4 µM
Reverse primer (10 µM)	2 µl	0.4 µM
Template DNA	2 µl	
Pfu DNA Polymerase	0.5 µl	

Table 9: Thermal cycler program used for PCR with the Pfu system

Step	Temperature	Time
Initial denaturation	95° C	2 min
Denaturation	95° C	45 s
Annealing	55° C	30 s
Elongation	72° C	3 min
Final elongation	72° C	15 min

To verify that various steps had been successful and to do colony PCR, Taq DNA Polymerase (Thermo Fisher Scientific) was used.

Table 10: Reagents and volumes for one PCR using Taq DNA Polymerase

Reagents	Volumes for 20 µl	Final concentration	Volumes for 50 µl	Final concentration
PCR-grade water	Up to 20 µl		Up to 50 µl	
10x Rxn PCR buffer	2 µl	1x	5 µl	1x
Nucleotides (10 mM)	1 µl	0.5 µM	1 µl	0.5 µM
$MgCl_2$ (50 mM)	1.5 µl	3750 µM	1.5 µl	1500 µM
Forward primer (10 µM)	1 µl	0.5 µM	2 µl	0.4 µM
Reverse primer (10 µM)	1 µl	0.5 µM	2 µl	0.4 µM
Template DNA	1 µl		Various	
Taq DNA Polymerase	0.1 µl		0.5 µl	

Table 11: Thermal cycler program used for PCR with Taq DNA Polymerase

Step	Temperature	Time
Initial denaturation	94° C	3 min
Denaturation	94° C	45 s
Annealing	55° C	30 s
Elongation	72° C	~2 min (various depending on gene length)
Final elongation	72° C	7 min

When the Taq DNA Polymerase was no longer available, DreamTaq DNA Polymerase (Thermo Fisher Scientific) was used.

Table 12: Reagents and volumes for one PCR using DreamTaq DNA Polymerase

Reagents	Volumes	Final concentration
PCR-grade water	Up to 10 µl	
10x DreamTaq Buffer (+Mg 20 mM)	1 µl	1x
Nucleotides (10 mM)	1 µl	10 mM
Forward primer (10 µM)	0.5 µl	0.5 µM
Reverse primer (10 µM)	0.5 µl	0.5 µM
Template DNA	Various	
DreamTaq DNA Polymerase	0.1 µl	

Table 13: Thermal cycler program used for PCR with DreamTaq DNA Polymerase

Step	Temperature	Time
Initial denaturation	94° C	3 min
Denaturation	94° C	30 s
Annealing	60° C	30 s
Elongation	72° C	~2 min (various depending on gene length)
Final elongation	72° C	15 min

Gel electrophoresis

Gel electrophoresis is a method used to separate molecules based on their size and charge. To select or verify DNA fragments of one size and separate them from other chemicals and DNA of, they were run on a 1 % agarose gel using electrophoresis. The agarose gels were made by heating 0.5 g agarose and 50 ml 1x TAE buffer in a microwave oven until the agarose was dissolved in the buffer. The gels were cast and allowed to solidify for approximately 30 min in room temperature. All gels were used within one hour before use. HyperLadder I or GelPilot 1 kb PLUS Ladder was used as a marker to compare DNA fragment sizes. Electrophoresis was run at 90 V for 40 min.

Plasmid isolation

Plasmids from bacteria with vectors were isolated using the PLN70 GenElute^{TM} Plasmid Miniprep Kit (Sigma). A pellet from an over-night culture was resuspended with Resuspention Solution (200 µl). Lysis Solution (200 µl) was added and gently mixed. Before 5 min had past, Neutralization Solution (350 µl) was added, mixed and centrifuged at 21000 x g for 10 min. A column from the kit was prepared with Column Preparation Solution (500 µl) and centrifuged for 1 min at 12000 x g. The flow-through was discarded. The lysate was added to the column and centrifuged at 1.5 min. The flow-through was discarded. Wash Solution (750 µl) was added and centrifuged for 1.5 min. The flow-through was discarded. The column was centrifuged again for 1 min to dry it. Pure water (100 µl) was added and centrifuged to elute the plasmids.

Concentration measurement with NanoDrop 2000

Concentrations of plasmids were measured using the NanoDrop 2000 (Thermo Scientific) spectrophotometer. The NanoDrop 2000 also measures the purity of the sample. Samples of DNA (1 µl) were placed on the sample pedestal and analyzed. The 260/280 ratio of absorbance and the 260/230 ratio of absorbance should be ~1.8 and ~2.0, respectively for a DNA sample to be considered pure.

Gel extraction

After the PCR product of the Expand High Fidelity^{PLUS} PCR System or the Pfu system was run and observed on a gel, it was cut out and put into a pre-weighed Eppendorf tube. The tube with the gel piece was weighed and the weight of the gel was determined. The NucleoSpin Gel and PCR Clean-up (Macherey-Nagel) kit was used to extract the DNA from the gel. For

every 0.01 g of gel, 20 µl NTI was added to the tube with the gel piece and heated up at 50° C until it was melted. The melted gel was transferred to a column and centrifuged at 11000 x g for 30 s. The flow-through was discarded. NT3 with ethanol (700 µl) was added and centrifuged for 1 min. The flow-through was discarded. It was centrifuged again for 1 min. The flow-through was discarded. It was centrifuged again for 1 min and the column was placed in a new tube. 15 µl water was added. After 1 min another 15 µl water was added and then centrifuged for 1 min. The Wizard® SV Gel and PCR Clean-Up System (Promega) was also used for gel extraction. The gel pieces were weighed. For every 0.01 g gel, 10 µl Membrane binding Solution was added and incubated until it was dissolved. The dissolved gel was transferred to a column in a collection tube. The assembly with the dissolved gel was centrifuged at 16000 x g for 1 min. The flow-through was discarded. Membrane Wash Solution (700 µl) and centrifuged at 16000 x g for 1 min. The flow-through was discarded. Membrane Wash Solution (500 µl) and centrifuged at 16000 x g for 5 min. The column was transferred to a new tube. Water (40 µl) was used to elute and centrifuged at 16000 x g for 1 min.

Digestion

To cut the vectors and outside genes at specific sites so that they can be fused together later, restriction enzymes were used. To solutions with vectors or genes there was added 2 µl FastDigest buffer, 1 µl of each restriction enzyme. The solutions were mixed well and incubated over-night at 37° C.

PCR Clean up

The NucleoSpin® Gel and PCR Clean-up was used to purify a product after digestion. 1 volume of sample was mixed with 2 volumes of Buffer NTI. A column was placed in a collection tube and the sample/buffer solution was centrifuged at 11000 x g for 30 s. The flow-through was discarded and the column was placed back in the collection tube. Buffer NT3 (700 µl) was added to the column and centrifuged at 11000 x g for 30 s. The flow-through was discarded and the column was placed back in the collection tube. The column was centrifuged again to get out any remaining liquid. The column was placed in a new tube and 15 µ - 30 µl water was used to elute the DNA. The column was centrifuged at 11000 x g for 1 min.

Ligation

To catalyze the fusion of two DNA ends with complementary cohesive ends, ligation was performed. Digested vector, digested gene, 5x Ligase buffer (4 µl) and T4 Ligase (1 µl). The mixture was mixed well. The ligase product was transformed into competent cells, incubated and spread on plates.

Protein production

An Erlenmeyer flask was filled with LB broth (50 ml). *E. coli* Rosetta strain (chloramphenicol resistant) transformed with pET-MBP_1a containing a kanamycin resistance gene and the gene of interest was added. Kanamycin (50 µl, 50 mg/ml) and glucose (5 ml, 20 %) was added and the flask was incubated on a shaker for 2 h at 37° C. IPTG (25 µl, 0.1 M) was added and the flask was incubated on a shaker over-night at 20° C to induce protein expression. The solution was put on ice.

Protein visualization on polyacrylamide gel

An incubated solution of *E.coli* Rosetta (25 ml) which had produced proteins was centrifuged and the supernatant was discarded. Lysis solution (1 ml) and lysosome (100 µl) was added and mixed. The solution was put on ice for 15 min.

To prepare the control, which had not been added IPTG, incubated control solution with *E.coli* Rosetta (1 ml) was centrifuged and the supernatant was discarded. Lysis solution (40 µl) and lysosome (4 µl) was added and mixed and put on ice for 15 min.

Laemmli Sample Buffer (950 µl) and 2-Mercapoethanol (50 µl) were mixed. Laemmli Sample buffer-2-Mercapoethanol (40 µl) solution was mixed with 40 µl of the lysed protein solution and the lysed control solution. The solutions were boiled at 95° C for 10 min and centrifuged for 1 min. The solutions (20 µl) were loaded into a pre-made polyacrylamide gel. The samples were run at 200 V for ~30 min. The polyacrylamide gel was washed in with water. The gel was covered in Coomassie Brilliant Blue and placed on a slow shaker for 30 min. The gel was washed with water and it was destained with distaining solution for 2 days.

TOPO Cloning reaction

The pENTR™/D-TOPO® Cloning kit (Invitrogen) was used for the TOPO Cloning reactions. PCR product (3 µl) and Salt Solution (1 µl) was mixed. Water was added up to 5 µl. TOPO® vector (1 µl) was added. The reaction sat in room temperature for 1.5 h to yield more colonies.

Gateway LR Clonase reaction

To transfer a gene of interest that is in the pENTR™/D-TOPO® vector to a final destination vector like pSF, purified pENTR™/D-TOPO® vector was heated to 65° C for 10 min. The pENTR™/D-TOPO® vector was mixed with pSF and put briefly on ice. The Gateway™ LR Clonase™ II Enzyme Mix (Invitrogen) was used. LR Clonase II Enzyme Mix (2 µl) was added and the reaction was incubated at 25° C over-night. To stop the reaction, Protinase K (1 µl) was added and mixed gently. The mixture was incubated at 37° C for 15 min. The solution was used to transform DH5α. The Gateway® LR Clonase™ Enzyme Mix (Invitrogen) was also used. The major difference in procedure was that 5X LR Clonase Reaction Buffer and LR Clonase were added separately.

Sequencing

Samples were set to Seqlab-Microsynth for sequencing using their Barcode Economy Run Service. The samples were sent together with a primer.

Floral dipping

A 150 ml culture of *A. tumefaciens* with the pSF plasmid with the gene of interest was centrifuged and the supernatant containing kanamycin and hygromycin was discarded. A buffer was made: 50 g sucrose (EMD Millipore Corporation, Darmstat), 2.03 g $MgCl_2$-6-hydrate (Riedel-de Haën, Seelze) and 50 µl Silwet L-77 (Lehle Seeds, Round Rock) in 1 L water. The pelleted *A. tumefaciens* was resuspended in the buffer (240 ml). Prior to the floral dipping, the siliques of the *A. thaliana* were cut off. The flowers and the stem of the plant was submerged in the buffer for 20 min. Figure 4 shows the setup. The plants were kept in plastic bags on its side to keep the moisture in. The plants were kept in the bag for 24 h.

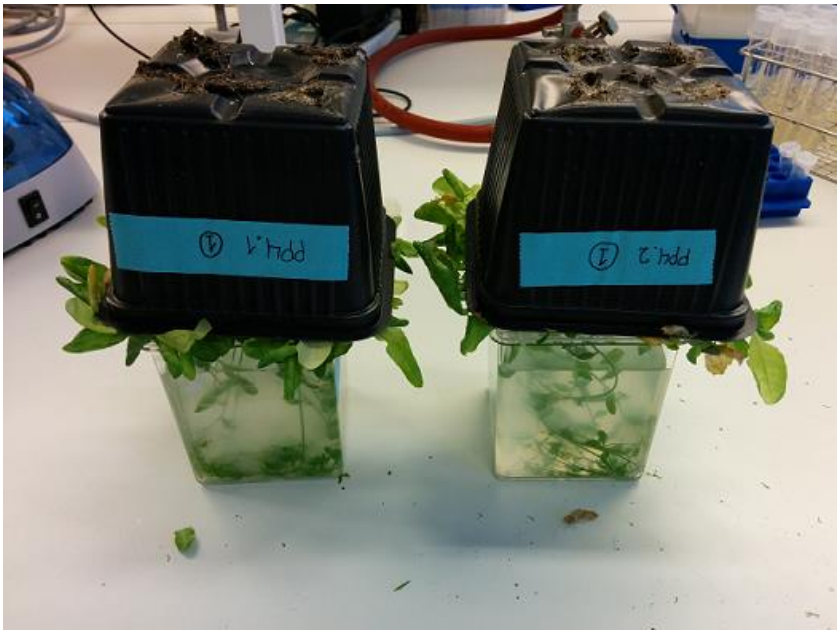

Figure 4: Floral dipping setup. Flowering *A. thaliana* submerged in bacterial buffer solution

Results

PP4-2 protein production

Vectors pETGST with kanamycin resistance and pETMBP with kanamycin resistance were transformed into JM109 competent cells to amplify the respective plasmids. The cells were grown in Petri dishes with LB agar and kanamycin. Colonies were incubated in LB broth with kanamycin.

PP4-2 (and PSY2L and PP42R which another student was working with) available in the pCAT vector, were amplified with Expand High Fidelity^{PLUS} PCR System from (Roche).

Cells with pETGST and pETMBP were lysed and the plasmids were purified using the GenElute Plasmid Miniprep Kit. The plasmids were run on agarose gel together with PSY2L, PP4-2 and PP4R2 amplified by PCR. The pETGST and pETMBP could not be observed (Figure 5).

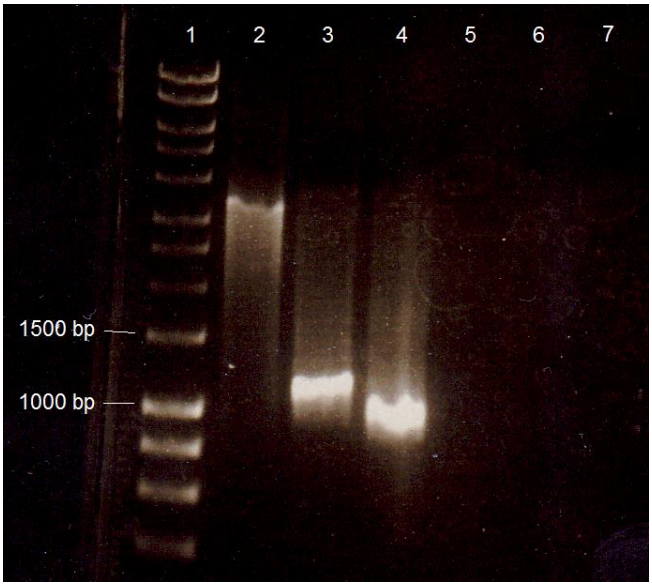

Figure 5: Gel electrophoresis of Hyperladder I and PCR products. Lane 1: Hyperladder I, lane 2: PSY2L, lane 3: PP4-2, lane 4: PP4R2, lane 5: empty, lane 6: pETGST, lane 7: pETMBP

Cells with pETGST and pETMBP were lysed and the plasmids purified and run again together with PSY2L, PP4-2 and PP4R2. All expected bands were observed (Figure 6), PP4-2 is 918 bp long, PSY2L is 2661 bp long, PP4R2L is 834 bp long, pETGST is 6722 bp long and pETMBP is 7181 bp long.

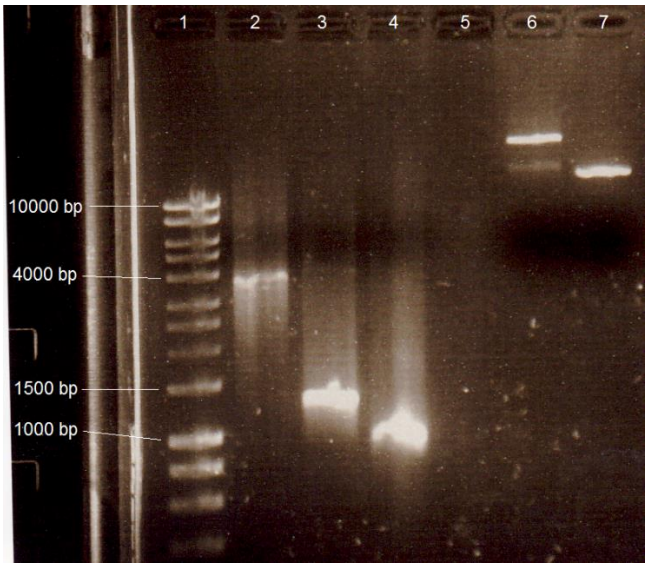

Figure 6: Gel electrophoresis of Hyperladder I, PCR products and the plasmids. Lane 1: Hyperladder I, lane 2: PSY2L, lane 3: PP4-2, lane 4: PP4R2, lane 5: empty, lane 6: pETMBP, lane 7: pETGST

PCR bands of PSY2L, PP4.2 and PP4R2 were cut out of the gel and cleaned up using gel extraction kit NucleoSpin Gel and PCR Clean-up (Macherey-Nagel). Samples of the PSY2L, PP4-2 and PP4R2 from the gel extraction were run on a gel. PP4-2 and PP4R2 band could be observed, but not the PSY2L (Figure 7).

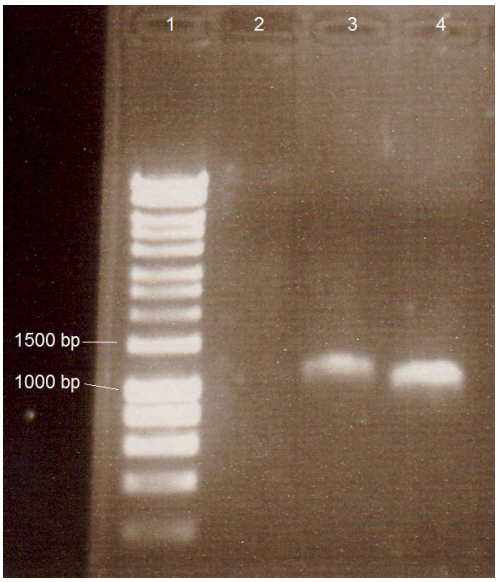

Figure 7: Gel electrophoresis of Hyperladder I and gene samples from the gel extraction. Lane 1: Hyperladde I, lane 2: PSY2L, lane 3: PP4-2, lane 4: PP4R2

A PSY2L sample was run on a gel again, but could not be observed. PCR of PSY2L was performed. The PCR product was run on a gel. PSY2L band failed to appear on the gel. Digestion of pETGST, pETMBP, PP4-2 and PP4R2 was performed using the restriction enzymes NcoI and Acc651 to cut them for ligation.

PCR of PSY2L was performed. Digested PP4-2, PP4R2, PSY2L, pETMBP and pETGST were run on a gel and extracted using DNA gel extraction. The samples extracted from the gel pieces were run on a gel. Only the PSY2L could be observed on the gel, and the experiment was considered a failure. DNA ligation of the unconfirmed digested PP4-2 and unconfirmed digested pETMBP was performed for training, and then discarded. The focus on PSY2L and PP4R2 was stopped. The pETGST was not used any further.

High Fidelity PCR of the pCAT vector with PP4-2 was performed. The PCR product was run on a gel (Figure 8). The gel piece was cut out and DNA gel extraction was performed. A sample of gel extraction product was run on a gel and it showed the extraction was successful. The extraction product was digested with NcoI and Acc65I.

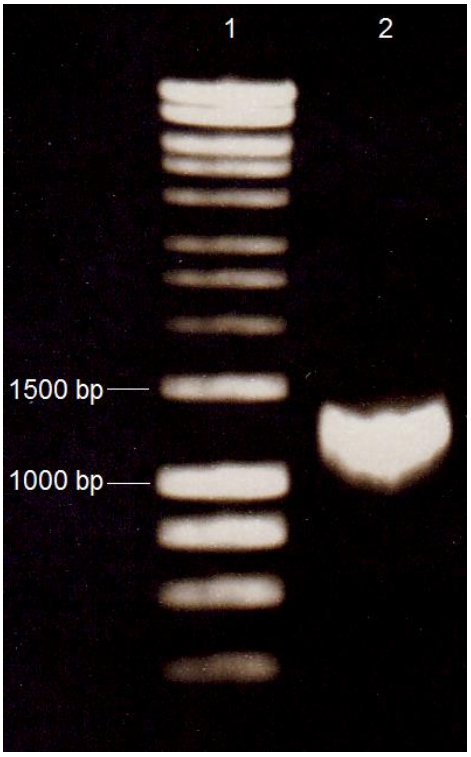

Figure 8: Gel electrophoresis of Hyperladder I and PCR product from PP4-2 pCAT plasmid.

PCR-clean up was performed to remove enzymes, etc. Some of the product was run on a gel to show that the clean-up was successful (Figure 9).

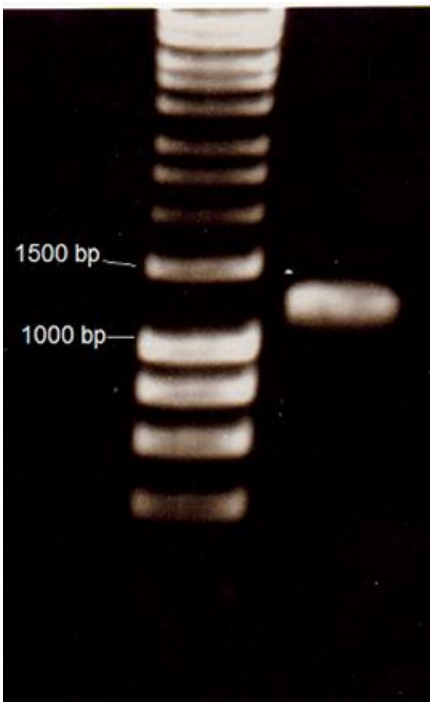

Figure 9: Gel electrophoresis of Hyperladder I and PCR clean-up product (PP4-2)

Ligation of PP4-2 and pETMBP was performed. The ligase product was transformed into JM109 competent cells. The bacteria were incubated with LB broth. The incubated bacteria were spread on plates with kanamycin and incubated. Colony PCR was performed with eight different colonies. Seven of the samples were run on a gel. All colonies except colony 2 had bands with correct size (Figure 10).

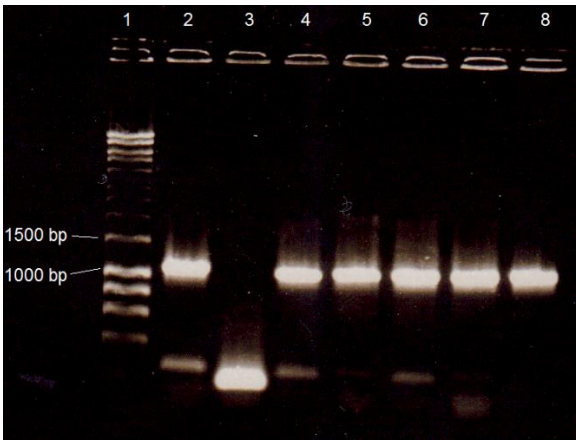

Figure 10: Gel electrophoresis of Hyperladder I and PCR products from colony 1-7 in lanes 2-8

Bacteria from samples 1,3,4,5,6,7 were added to tubes with LB broth and kanamycin and incubated.

Plasmid isolation was performed. The plasmid samples were run on a gel (Figure 11).

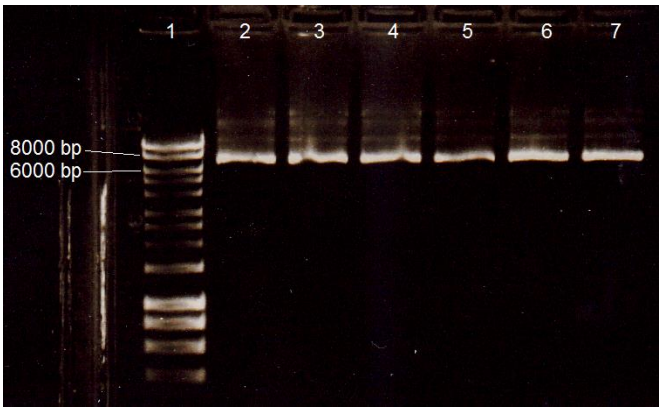

Figure 11: Gel electrophoresis of plasmids from colonies 1 (lane 2), 3-7 (lanes 3-7)

Glycerol stock solutions were made with some of the bacteria for long-term storage.

PCR was run with the samples 1,3,4,5,6,7,a pETMBP vector control and PP4.2_MBP F/R primers. The PCR products were run on a gel (Figure 12).

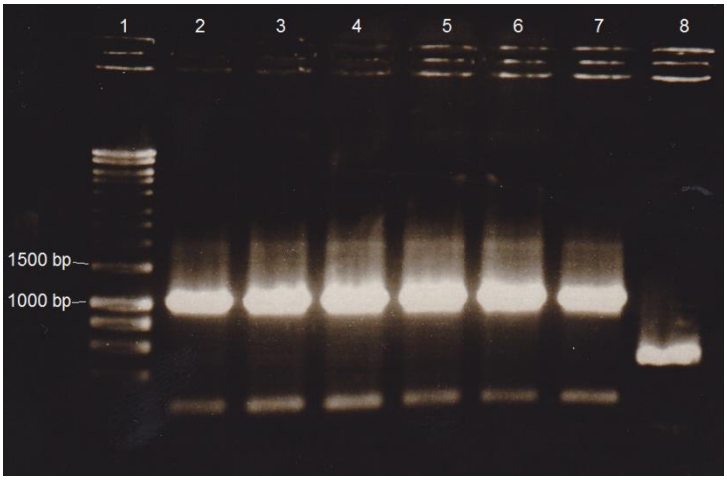

Figure 12: Gel electrophoresis of samples 1 (lane 2), 3-7 (lanes 3-7) and an pETMBP vector control (Lane 8)

Competent Rosetta bacteria were transformed with the isolated plasmids samples 1 and 3, spread on two plates with kanamycin and chloramphenicol and incubated. Colonies grew on the plate with bacteria transformed with plasmid sample 1. Four colonies were selected and a colony PCR was run using PP4.2_MBP F/R primers. The PCR products were run on a gel (Figure 13). All four bands in figure 9 have a size around 1000 bp, indicating the presence of PP4-2.

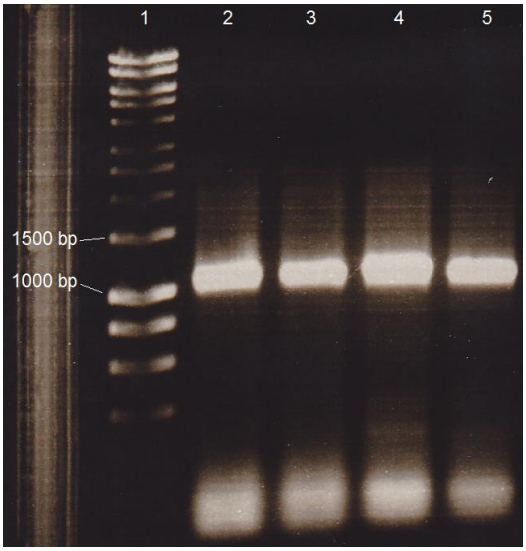

Figure 13: Gel electrophoresis of colony PCR products from transformed Rosetta in lanes 2-5

Colonies 1 and 2 from the plate was incubated in LB broth with kanamycin and chloramphenicol.

Glycerol stock solutions of the incubated bacteria were made for long-term storage.

The plasmids of the incubated Rosetta solutions 1 and 2 were isolated and run on a gel (Figure 14). From figure 10, it is apparent that Rosetta plasmid solution 2 had the best amount of plasmid, but there is also a weak band from Rosetta plasmid solution 1, indicating presence of plasmids.

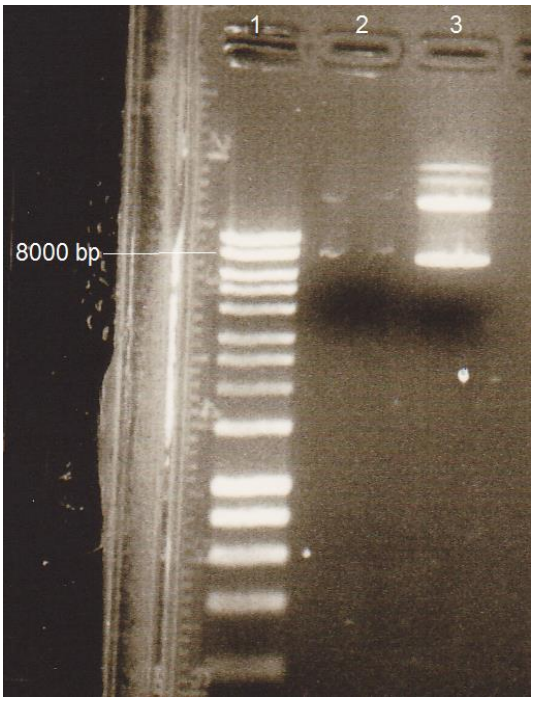

Figure 14: Gel electrophoresis of plasmids of Rosetta plasmid solutions 1 (lane 2) and 2 (lane 3)

Protein production was performed with Rosetta solution 1 and 2 in two Erlenmeyer flasks. The absorbance was not measured, as it was decided that visually observing the increased opacity level was sufficient to proceed with the experiment. IPTG was added after 2 h. Two controls from each solution were made without adding any IPTG. All the solutions were incubated at 20° C over-night.

The incubated bacteria were lysed and added lysosome. Lysed samples from flask 1 and 2 were run on a polyacrylamide gel together with a control that had not been added IPTG

(Figure 15). The red arrow in figure 11 points to a band in sample 1, which has the expected size of the protein. It is probable that PP4-2-MBP has been expressed in sample 1.

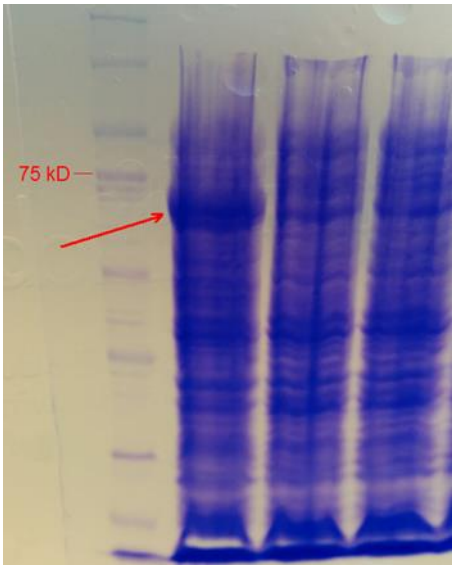

Figure 15: SDS-PAGE-gel electrophoresis of proteins stained with Coomassie Brilliant Blue. Lane 1: Ladder, lane 2: sample 1, lane 3: control, lane 4: sample 2

A sample of each of the isolated plasmids from JM109 colonies 1 and 3 (plasmids from colony 1 was used to transform the Rosetta previously) were sent for sequencing by Seqlab. The sequencing alignment between PP4-2 and the PP4-2 in plasmid 1(Appendix A-3) shows one apparent mutation. The alignment of PP4-2 in the other plasmid showed the same apparent mutation. It was learned that this error was present in the pCAT source vector.

PP4-2 pSF preparation

PP4-2 available in an EYFP-PP4-2 plasmid was amplified with PCR using the Phusion system and PP4.2_dTOPO_F and PP4.2_dTOPO_R to make a blunt-end DNA fragment for use in TOPO cloning. A sample was run on a gel, but no band was visible. The experiment was repeated with double the amount of plasmid and double the amount of DNA polymerase, but it was not visible on the gel. The experiment was repeated a third time using a different EYFP-PP4-2 plasmid and newly diluted primers, but there was no visible band on the gel. The Phusion system was considered unfit for the experiment. The experiment was repeated with the Pfu system. A PCR with the same plasmid and primers, and the Taq DNA polymerase system was run simultaneous with the Pfu system. The bands that appeared on the gel were of poor quality. The Pfu system was tried again, but there was no band on the gel. A supervisor

helped in performing the PCR again using the Pfu system. The amplification was successful and a band appeared on the gel. The DNA was cut out from the gel (Figure 16)

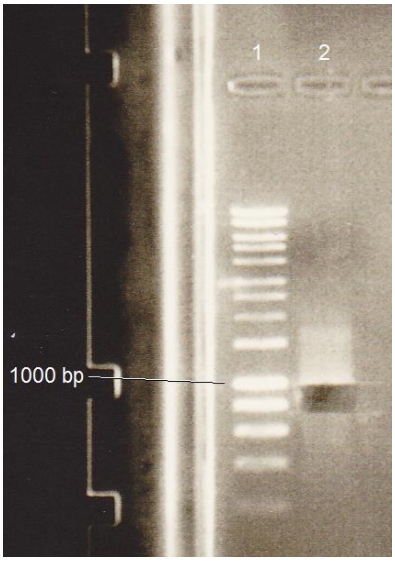

Figure 16: Gel electrophoresis of PCR product using PP4-2_dTOPO F/R primers. The gel piece with the DNA has been cut out from lane 2.

DNA from the cut out gel piece was extracted by gel extraction. A sample of the extract was run on a gel (Figure 17).

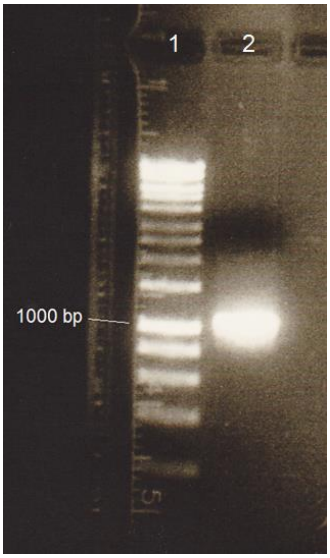

Figure 17: Gel electrophoresis of a sample of the gel extract of PP4-2 dTOPO

A TOPO cloning reaction was performed using 3 µl of the extracted PCR product. The reaction sat in room temperature for 1.5 h. JM109 competent cells were transformed with the

reaction product. The bacteria were plated on LB agar with kanamycin and incubated overnight.

Colonies (A1-A8) were selected on the plate and colony PCR for them was performed using M13 forward and reverse primers. The PCR products were run on a gel (Figure 18). Bands with size about 1000 bp are visible in Figure 18 and indicates that the bacteria contains the vector with the PP4-2 dTOPO, but multiple bands are visible, making the result unreliable.

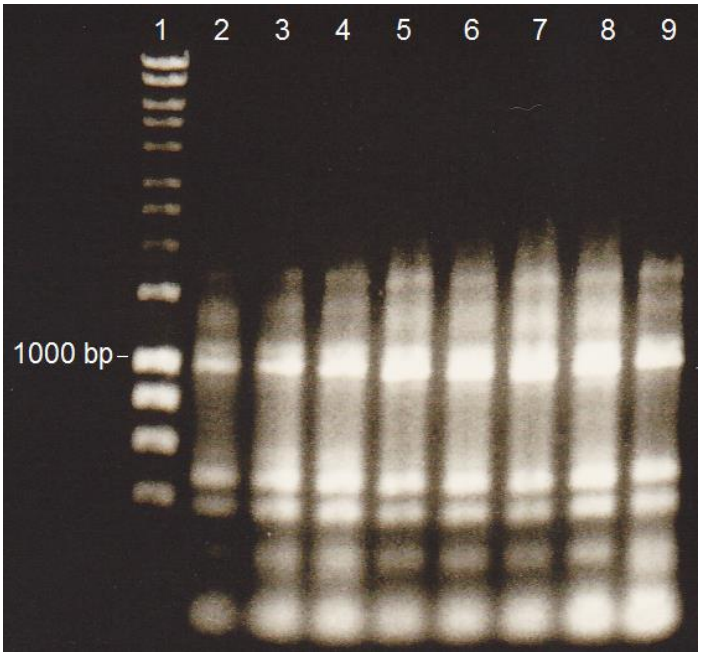

Figure 18: Gel electrophoresis of colony PCR of PP4-2 dTOPO colonies A1-A8 in lanes 2-9

New colonies (B1 – B8) were selected from the PP4-2 dTOPO plate. A colony PCR was run using the PP4-2_dTOPO forward and M13 reverse primers. The products were run on a gel (Figure 19). Bands in Figure 19 in lane 7 and 8 (Colony B5 and B6) have the size around 1000 bp. This indicates that PP4-2 dTOPO has been successfully cloned.

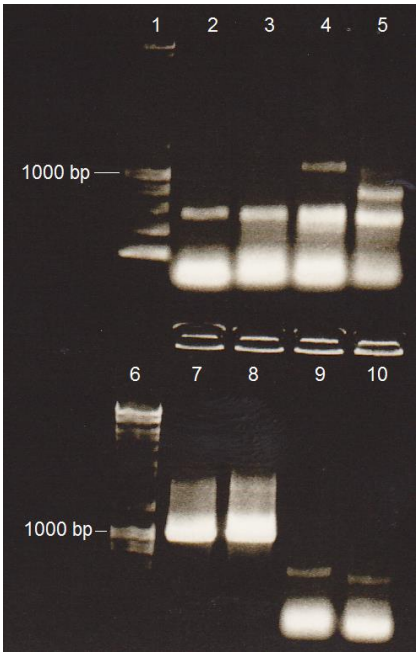

Figure 19: Gel electrophoresis of colony PCR of colonies B1-B8 in lanes 2-5, 7-10

More colonies (C1 – C8) were selected to in case B5 and B6 had wrong DNA sequences. Colony PCR using PP4-2_dTOPO forward and M13 reverse primers was performed and the products were run on a gel (Figure 20). In Figure 20, colonies C2, C3, C4, C7 have bands about 1000 bp, indicating that the colonies have the PP4-2 dTOPO in the pENTR vector.

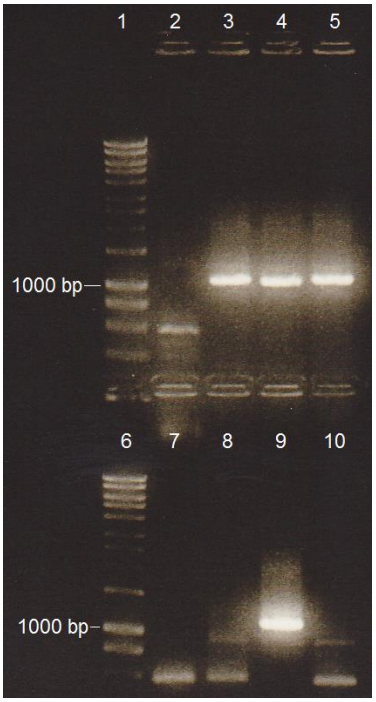

Figure 20: Gel electrophoresis of colony PCR of colonies C1-C8 in lanes 2-5, 7-10

Over-night cultures with kanamycin of B5, B8, C2, C3, C4 and C7 were made. Glycerol stock solutions were made for each of the cultures for long-term storage.

Plasmids were isolated from B5, B8, C2, C3, C4 and C7. A PCR of the isolated plasmids were run using PP4.2 dTOPO forward and M13 reverse primers. The PCR products were run on gels (Figure 21, Figure 22)

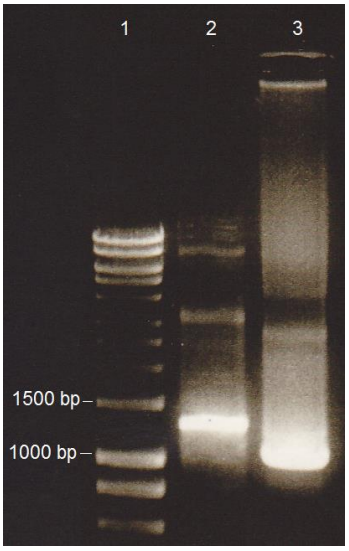

Figure 21: Gel electrophoresis of PCR products of plasmids of C2 and C3 in lane 2 and 3

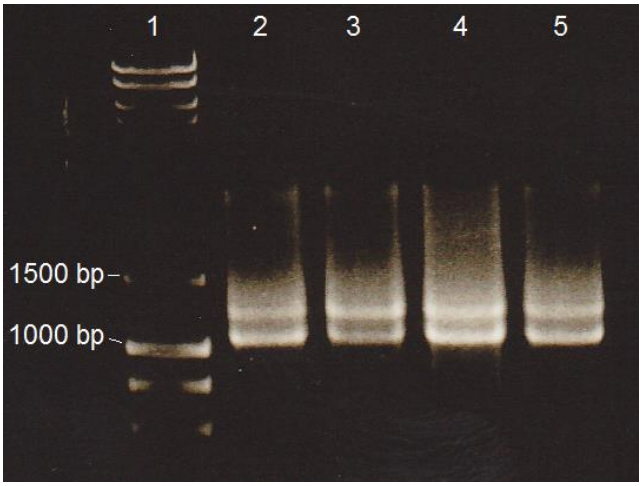

Figure 22: Gel electrophoresis of PCR products of plasmids C4, C7, B5 and B6 in lanes 2-5

Concentrations of plasmids C2, C3 and C4 were measured with the NanoDrop 2000.

Table 14: Concentrations measured with the NanoDrop 2000

Sample	Measured concentration (ng/µl)	260/280	260/230
PP4-2 dTOPO C2	50.4	1.90	2.22
PP4-2 dTOPO C3	200.7	1.90	2.20
PP4-2 dTOPO C4	71.6	1.88	2.26

A sample of PP4-2 dTOPO C2, PP4-2 dTOPO C3 and PP4-2 dTOPO C4 was sent for sequencing with the M13 F/R primers. The sequencing result (Appendix A-4) showed that the sequences were correct, but all the sequences had the known error as the one found in the pCAT vector. The multiple sequence alignment between PP4-2 and PP4-2 dTOPO C2 in Appendix A-4 is representative for all the plasmids.

The restriction enzyme ApaI was found to cut outside the *att*-region of pSF and not in the gene. ApaI destroys the kanamycin resistance so that bacteria with the final destination vector pSF can be selected.

Digestion of PP4-2 dTOPO C3 plasmid was performed using ApaI. The product was run on a gel and cut out. Gel extraction was performed. The concentration was measured to be 25.4 ng/µl. Plasmid isolation of the cut gel extract PP4-2 dTOPO C3 was performed.

PP4-2 dTOPO C3 was heated at 65° C for 10 min. pSF vector (3 µl, 80 ng/µl) was mixed with PP4-2 dTOPO C3 plasmid (8 µl, 25.4 ng/µl) and briefly put on ice. LR Clonase II enzyme mix (2 µl) was added and mixed gently. The solution was incubated at 25° C over-night. Proteinase K (1 µl) was added to the PP4 dTOPO C3-pSF LR reaction to stop it. The solution was incubated at 37° C for 15 min. DH5α competent cells were transformed with the product, spread on agar plates with kanamycin and incubated at 37° C over-night. DH5α competent cells were transformed with uncut PP4-2 dTOPO C3 as a control and grown the same way.

Three colonies: PP4-2-pSF 1, PP4-2-pSF 2, PP4-2-pSF 3 were picked from the plate with DH5α transformed with PP4-2-pSF. Three colonies with uncut PP4 dTOPO C3 were picked. A colony PCR was performed. A pSF vector control was included. The products were run on a gel (Figure 23).

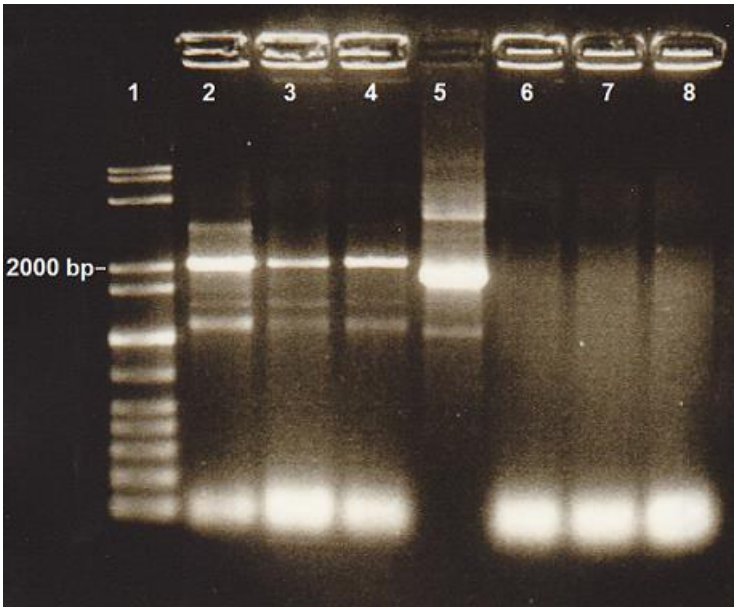

Figure 23: Gel electrophoresis of colony PCR products. Lane 1: GelPilot 1 kb Plus Ladder, lane: 2-4: PP4-2-pSF 1, 2, 3, lane 5: pSF vector control, lane 6-8: PP4-2 dTOPO C3

Figure 23 indicates that the pSF has been transformed into the PP4-2-pSF colonies picked from the LR plate. Over-night cultures were made from PP4-2-pSF 1 and PP4-2-pSF 2 with kanamycin.

Glycerol stock solutions were made from the over-night cultures. The plasmids of PP4-2-pSF 1 and PP4-2-pSF 2 were isolated. The concentrations of the plasmid solutions were measured (Table 15)

Table 15: Concentration of isolated plasmids of PP4-2-pSF 1 and PP4-2-pSF 2 measured with the NanoDrop2000

Sample	Measured concentration (ng/µl)	260/280	260/230
PP4-2-pSF 1	241	1.94	2.17
PP4-2-pSF 2	160	1.95	2.22

A PCR of the isolated plasmids of PP4-2-pSF 1 and PP4-2-pSF were performed using pSF F/R primers, and PP4.2_dTOPO F/R primers. pSF vector controls were included. A negative control was included. The products were run on a gel (Figure 24).

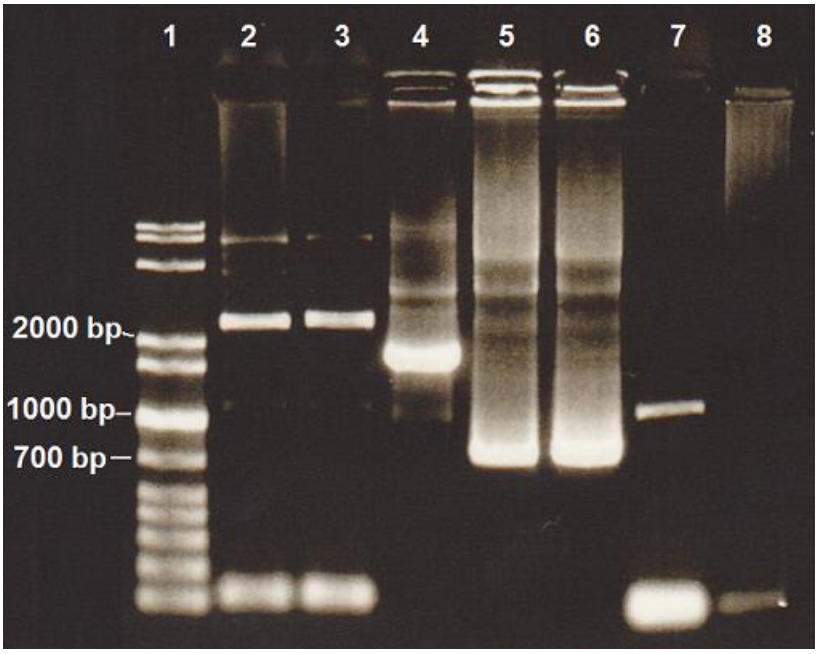

Figure 24: Gel electrophoresis of PCR products. Lane 1: GelPilot 1 kb Plus Ladder, lane 2: PP4-2-pSF 1 using pSF F/R, lane 3: PP4-2-pSF 2 using pSF F/R, lane 4: pSF vector control using pSF F/R, lane 5: PP4-2-pSF 1 using PP4.2_dTOPO F/R, lane 6: PP4-2-pSF 2 using PP4.2_dTOPO F/R, lane 7: pSF vector control using PP4.2_dTOPO F/R, lane 8: negative control using pSF F/R

Lane 2 and 3 in Figure 24 indicates that PP4-2-pSF 1 and PP4-2-pSF 2 both has the pSF vectors. They are larger than the pSF vector control indicating a longer sequence has been amplified. PP4-2-pSF 1 and PP4-2-pSF 2 in lane 5 and 6 indicate the presence of PP4-2.

The plasmids of PP4-2-pSF 1 and PP4-2-pSF 2 were sent for sequencing using PP4.2_dTOPO F primer and pSF F primer. The sequencing results for the plasmids sent with pSF F primer were uninterpretable. The sequencing failed to read the nucleotides. The sequencing of PP4-2-pSF 1 with PP4.2_dTOPO F primer (Appendix A-16) shows that the sequencing have not read the start of the gene. There are also apparent mutations in the after the start region. The sequencing of PP4-2-pSF 2 with PP4.2_dTOPO F primer (Appendix A-17) shows that the sequencing have not read the start of the gene. There are some nucleotides that it failed to be read and a nucleotide has been deleted in the end of the gene.

It was theorized that pSF primers could bind in to the pENTR/D-TOPO vector, giving a false positive result. It was suggested that M13 primers could be used to screen for colonies by finding colonies that had no band on an electrophoresis gel after amplification, even though

the M13 primers should bind in the pSF. This indication came from a single result by a co-student, but it was attempted in the next PCR and gel run.

Bacteria transformed with PP4 dTOPO C3-pSF were transferred to an agar plate with kanamycin and hygromycin using toothpicks. The plate had a field where three colonies that had been transformed with the digested PP4-2 dTOPO C3. The plates were incubated at 37° C over-night. Colonies grew on the kanamycin/hygromycin plate. There were also growth in the field which had bacteria transformed with the digested PP4-2 dTOPO C3.

Colonies (1-10) were selected from the plate in addition to one colony from the field with bacteria transformed with digested PP4-2 dTOPO C3. A colony PCR was performed using M13 F/R primers. A pSF vector control, a PP4-2 dTOPO C2 plasmid control and a negative control was included. The PCR products were run on a gel (Figure 25).

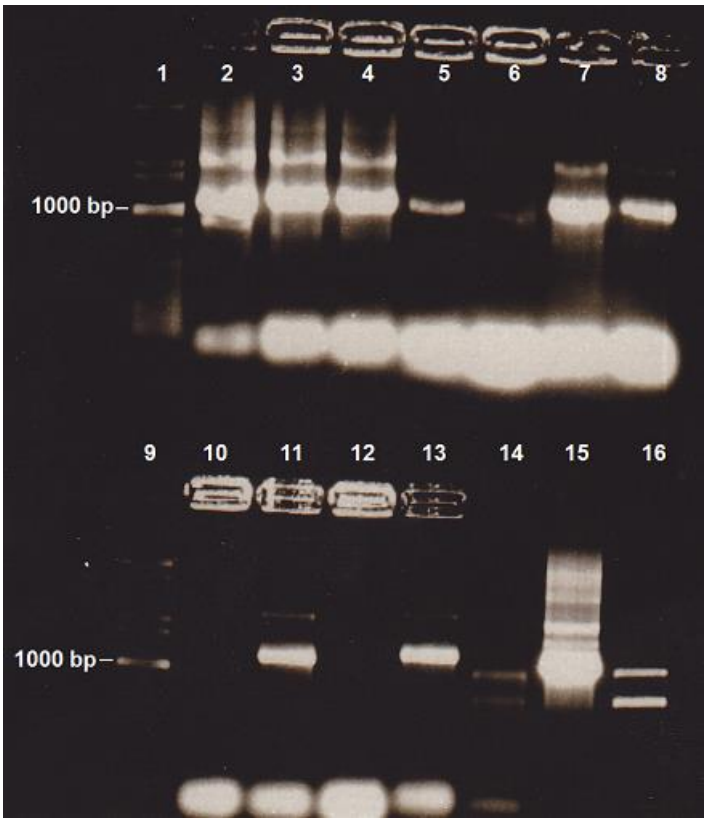

Figure 25: Gel electrophoresis of colony PCR products amplified using M13 F/R primers. Lane 1 and 9: GelPilot 1 kb Plus Ladder, lane 2-8, 10-12: PP4 dTOPO C3-pSF colonies 1-10, lane 13: PP4-2 dTOPO C3 colony, lane 14: pSF vector, lane 15: PP4-2 dTOPO C2 plasmid control, lane 16: negative control

From Figure 25, it is apparent that colony 5 (lane 6), 8 (lane 10) and 10 (lane 12) had not been

amplified with the M13 primers, indicating they might be positive. Over-night cultures were made from colony 5, 6 and 10 with kanamycin and hygromycin.

There was growth in culture 10 (PP4-2-(pSF)). The plasmids of PP4-2-(pSF) were isolated. The concentration was measured to be 43.9 ng/µl.

A PCR of the isolated plasmids of PP4-2-(pSF) was performed. A pSF vector control, PP4-2 dTOPO C2 plasmid control and a water control were included. The product were run on a gel (Figure 26).

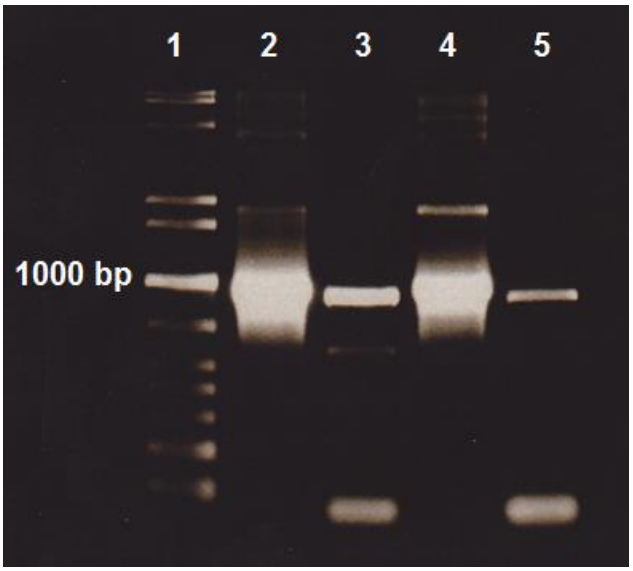

Figure 26: Gel electrophoresis of PCR products. Lane 1: GelPilot 1 kb Plus Ladder, lane 2: PP4-2-(pSF), lane 3: pSF vector control, lane 4: PP4-2 dTOPO C2, lane 5: water control

There were bands in the pSF vector control lane (lane 3) and negative control lane (lane 5). The PCR was repeated with the addition of PP4.2_dTOPO F/R and pSF F/R primers. The products were run on a gel (Figure 27).

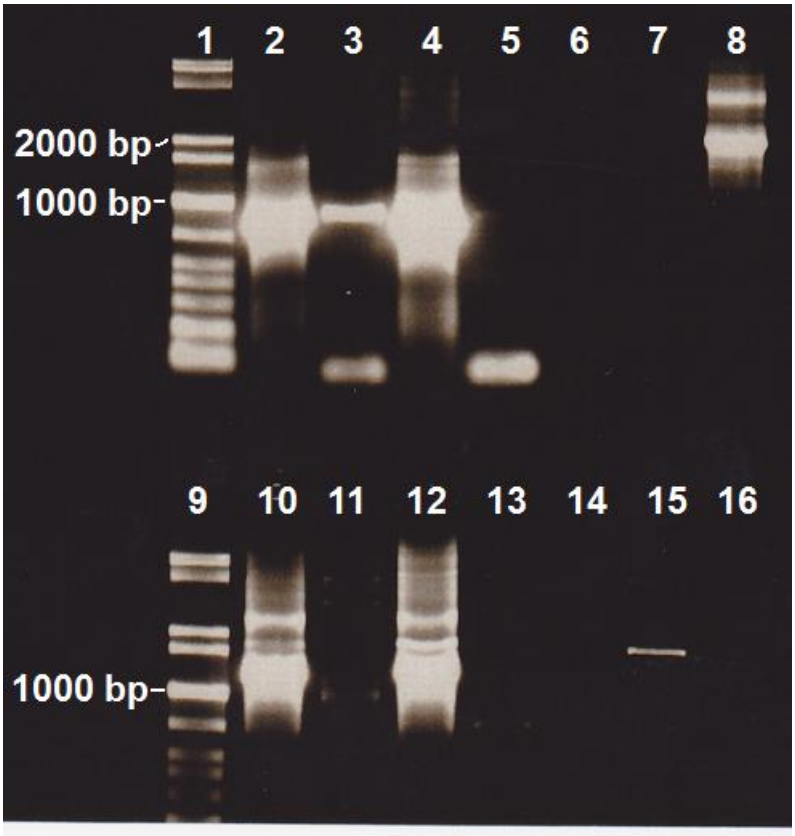

Figure 27: Gel electrophorese of PCR products. Lane 1 and 9: GelPilot 1 kb Plus Ladder, lane 2: PP4-2-(pSF) (PP4.2_dTOPO F/R), lane 3:,pSF vector control (PP4.2_dTOPO F/R) lane 4: PP4-2 dTOPO C2 (PP4.2_dTOPO F/R), lane 5: water control (PP4.2_dTOPO F/R), lane 6: empty, lane 7: PP4-2-(pSF) a(pSF F/R), lane 8: pSF vector control (pSF F/R), lane 10: PP4-2-(pSF) (M13 F/R), lane 11: pSF vector control (M13 F/R), lane 12: PP4-2 dTOPO C2 (M13 F/R), lane 13: water control (M13 F/R), lane 14: empty, lane 15: PP4-2 dTOPO C2 (pSF F/R), lane 16: water control (pSF F/R).

In Figure 27, there should not be a band in lane 3 where the amplified pSF vector control (PP4.2_dTOPO F/R primers). A similar band appeared in Figure 26, suggesting PP4.2_dTOPO F/R can bind somewhere in the pSF vector. In lane 15, there is a weak band where the amplified PP4-2 dTOPO C2 (pSF F/R) have run. There may be a chance pSF primers can bind to the pENTR/D-TOPO vector or it may bay caused by an impurity. PP4-2-(pSF) is considered to not have the pSF plasmid because there is no band in lane 7.It should have reacted with the pSF primers.

A new LR reaction was performed with PP4-2 dTOPO C3 plasmid (1.5 µl, 200 ng/µl) and pSF plasmid (3 µl, 80 ng/µl) using the Gateway® LR Clonase™ Enzyme Mix (Invirtogen). The two volumes of plasmid solution were mixed and incubated at 65° C for 10 min. 5X LR Clonase Reaction Buffer was added. The solution was filled with water up to 8 µl. The LR

Clonase enzyme mix was added and the reaction was incubated in 25° C over-night. Proteinase K (1 µl) was added and the solution was incubated at 37° C for 15 min. DH5α completent cells were transformed with the LR reaction product. The bacteria were spread on an agar plate with kanamycin and hygromycin and incubated at 37° C over-night. There was no growth on the plate.

The LR reaction was repeated and new DH5α competent cells were transformed with the LR reaction product. The bacteria were spread on an agar plate with kanamycin and hygromycin and incubated at 37° C over-night. Two colonies grew on the plate. The colonies (PP4-2-pSF A and PP4-2-pSF B) were incubated in LB broth, kanamycin and hygromycin at 37° C over night. The plasmids were isolated from the cells and the concentration was measured (Table 16)

Table 16: Concentration of isolated plasmids of PP4-2-pSF A and PP4-2-pSF B measured by the NanoDrop 2000

Sample	Measured concentration (ng/µl)	260/280	260/230
PP4-2-pSF A	86.2	1.97	1.66
PP4-2-pSF B	60.5	2.06	2.43

PP4-2-pSF A and PP4-2-pSF B were sent for sequencing with pSF R, PP4.2_dTOPO F and PP4.2_dTOPO R primers (Appendix A-18, A-19). PP4-2-pSF A sequences in A-18 and A-19 overlaps and aligns with the PP4-2 sequence and PP4-2 is considered to be successfully cloned. PP4-2-pSF B sequences in A-18 and A-19 also overlaps and aligns with the PP4-2 sequence and PP4-2 is considered to be successfully cloned. The sequence of PP4-2-pSF A sequenced with pSF R primer was aligned with the sequence of pGWB2 vector (Appendix A-20), which the pSF vector has been made from. It indicates that the pSF R primer has bond to the pSF vector. It does not confirm that PP4-2 dTOPO is in the vector. The sequence of PP4-2-pSF B sequenced with pSF R primer showed a similar result, but was ~100 bp shorter.

The next step was transformation of *A. tumefaciens*.

PP4-1 pSF preparation

A PCR using a PP4-1-pGEMT DNA source, PP4.1_dTOPO forward and PP4.1_dTOPO reverse primers and the Pfu system was run. The PCR product was run on a gel (Figure 28). A sample was run in lane 2 to document the size of the product. The same PCR products in higher volumes were in lane 5 and 6, but were cut out before the photograph was taken to minimize potential DNA damage.

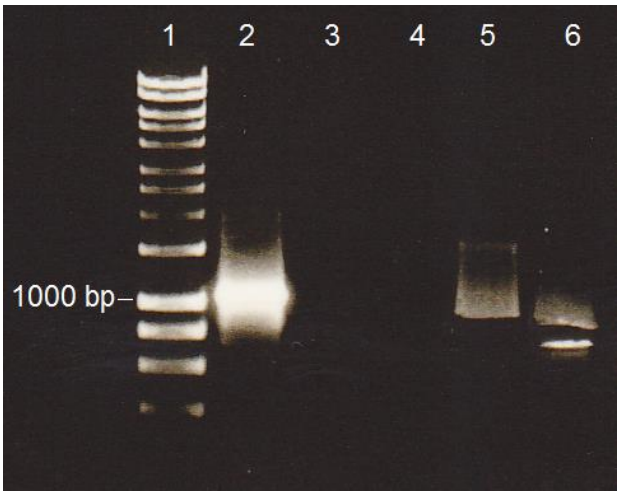

Figure 28: Gel electrophoresis PCR products of the PP4-1-pGEMT DNA source and PP4-1 dTOPO F/R primers. Lane 2 contains the run PCR product. The run PCR products in lane 5 and 6 has been cut out.

The two gel pieces of PP4-1 dTOPO were collected in a tube and gel extraction was performed using the Promega kit. The concentration was measured.

Table 17: Concentration of PP4-1 dTOPO measured by the NanoDrop 2000

Sample	Measured concentration (ng/µl)	260/280	260/230
PP4-1 dTOPO	27.9	1.87	0.91

A TOPO cloning reaction was performed with PP4-1 dTOPO and was incubated in room temperature over-night.

JM109 competent cells were transformed with the TOPO cloning reaction product. The bacteria were spread on two plates with kanamycin. The plates were incubated at 37° C over-

night. Eight colonies were selected from one of the plates and colony PCR was performed with M13 F/R primers. The PCR products were run on a gel (Figure 29).

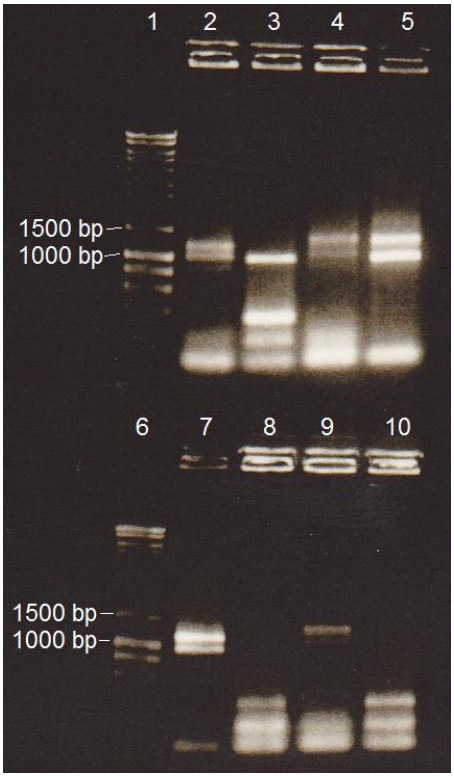

Figure 29: Gel electrophoresis of colony PCR of colonies 1-8 PP4-1 dTOPO in lanes 2-5, 7-10.

Over-night cultures with kanamycin were made from colonies 1 (lane 2), 4 (lane 5) and 5 (lane 7). The plasmids of culture 1, 4 and 5 were isolated. The concentration of the isolated plasmids were measured with the NanoDrop 2000.

Table 18: Concentrations of PP4-1 dTOPO plasmids measured by the NanoDrop 2000

Sample	Measured concentration (ng/µl)	260/280	260/230
PP4-1 dTOPO 1	106	1.91	2.34
PP4-1 dTOPO 4	116	1.87	2.22
PP4-1 dTOPO 5	21	1.97	2.41

Samples of isolated PP4-1 dTOPO plasmids 1 and 4 were diluted to ~20 ng/µl.

A PCR was preformed using M13 F/R primers, the diluted plasmid solutions PP4-1 dTOPO 1 and 4, PP4-1 dTOPO 5, PP4-1 original vector control and TOPO LCMT control. The PCR also included PP4-1 dTOPO 1 and 4, PP4-1 dTOPO 5, PP4-1 original vector control and TOPO LCMT control run with PP4.1_dTOPO F/R primers. The PCR products were run on a gel (Figure 30).

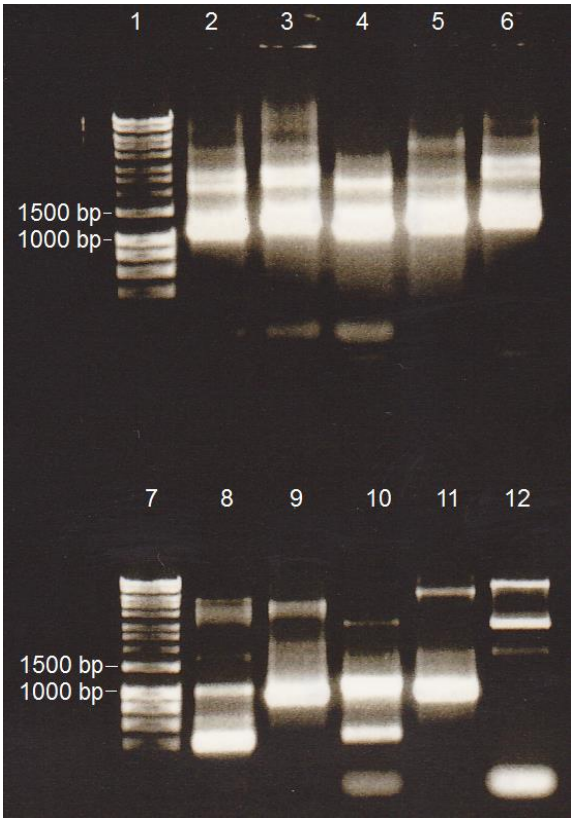

Figure 30: Gel electrophoresis of plasmid solutions PCR products. Lane 1 and 7: Hyperladder I, lane 2, 3, 4: PP4-1 dTOPO 1, 4, 5 using M13 F/R primers. Lane 5: PP4-1 original vector using M13 F/R primers, lane 6: TOPO LCMT using M13 F/R primers. Lane 8, 9, 10: PP4-1 dTOPO 1, 4, 5 using PP4.1_dTOPO F/R primers. Lane 11: PP4-1 original vector using PP4.1_dTOPO F/R primers, lane 12: TOPO LCMT using PP4.1_dTOPO F/R primers

In Figure 30 it is apparent from lanes 2, 3, 4, 5 and 6 that an M13 F/R compatible vector is present and that it contains a sequence ~1000 bp long. The bands in lane 8 indicate that there is some PP4-1 dTOPO fragments and a higher amount of product that is shorter. The band in lane 9 indicate that there is PP4-1 dTOPO successfully amplified. The band in lane 10 indicate that there is successfully amplified PP4-1 dTOPO, but also a shorter product. The band in lane 11 shows amplified PP4-1 dTOPO form the original PP4-1 dTOPO vector. It is similar to the band in lane 9. The bands in lane 12 indicate that the PP4.1_dTOPO F/R

primers does not help amplify a product that is ~1000 bp in a vector that does not contain the PP4-1 dTOPO sequence.

From Figure 30, PP4-1 dTOPO plasmids 4 and 5 were the best candidates to continue with. The plasmids were sent for sequencing. The multiple sequencing alignment between PP4-1 and the sequenced PP4-1 dTOPO 4 (Appendix A-5) shows that the PP4-1 dTOPO 4 has a mutation in the stop codon. The reverse complimentary alignment of it shows the same error. The sequence of PP4-1 dTOPO 5 had many differences from PP4-1 and an alignment between the sequences was not applicable.

New colonies (9-16) from the two plates that had JM109 transformed with the TOPO cloning reaction product (PP4-1 dTOPO) were picked and a colony PCR using PP4.1_dTOPO F/R primers and a colony PCR using M13 F/R primers were performed. Negative controls were included with each primer sets using water instead of bacteria solution. The PCR products were run on a gel (Figure 31).

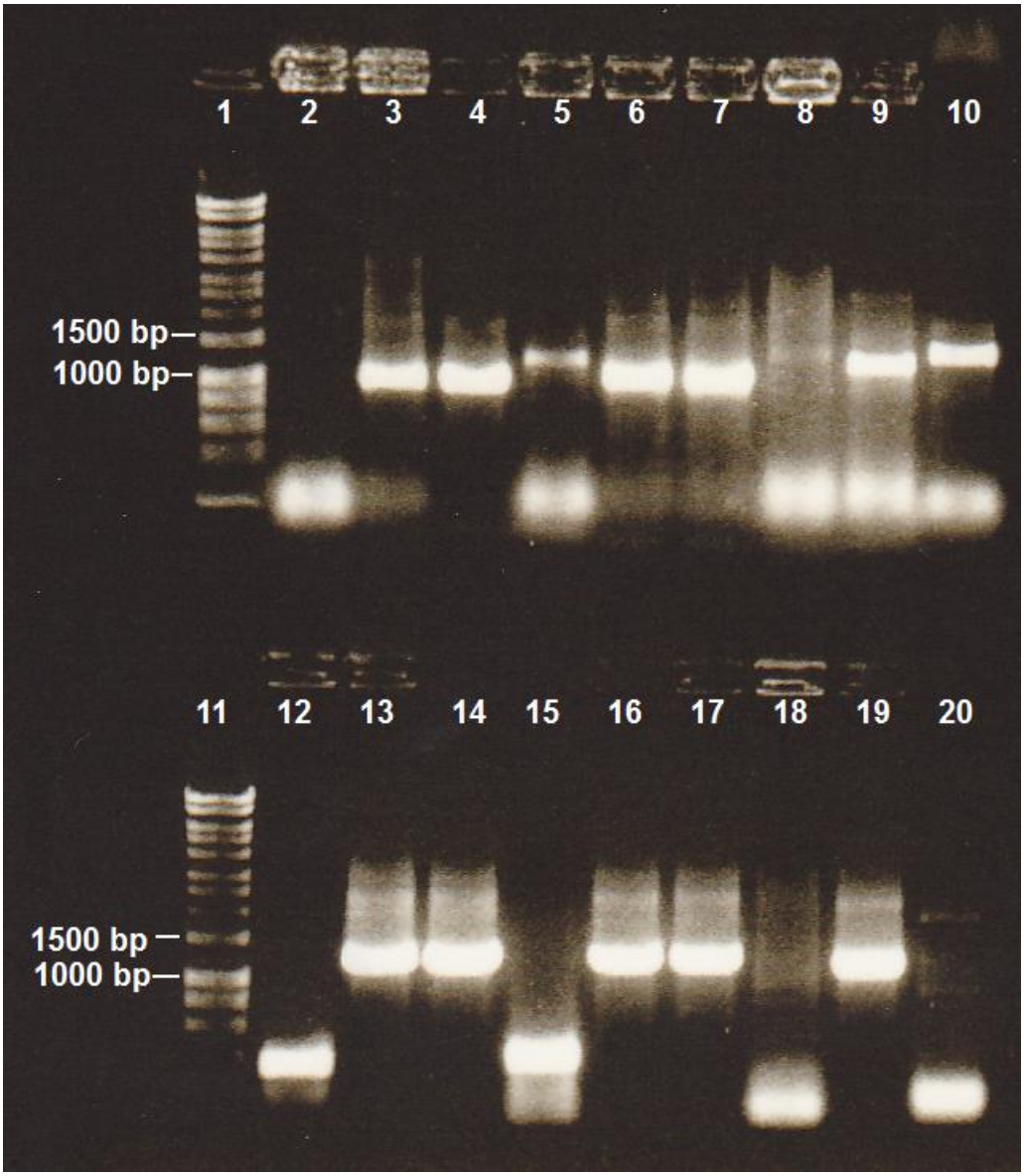

Figure 31: Gel electrophoresis of colony PCR products from colonies PP4-1 dTOPO 9-16. Lane 1 and 11: Hyperladder I, lane 2-9: PP4-1 dTOPO colonies 9-16 using PP4.1_dTOPO F/R primers, lane 10: negative control using PP4.1_dTOPO F/R primers, lane 12-19: PP4-1 dTOPO colonies 9-16 using M13 F/R primers, lane 20: negative control using M13 F/R primers

From the observations of Figure 31, it was concluded that PP4-1 dTOPO colonies 10 (lane 3 and 13), 11 (lane 4 and 14), 13 (lane 6 and 16), 14 (lane 7 and 17) and 16 (lane 9 and 19) had a good chance of having the PP4-1 gene in the pENTR-D-TOPO vector because they all had bands ~1000 bp. The negative control in lane 10 shows a band between 1500 bp and 1000 bp long. It is unknown why this band has appeared. A mistake may have been made or one of the reactant solutions had impurities. It was decided that cultures should be grown and their plasmids should be sent for sequencing despite this potential error.

Over-night cultures of PP4-1 dTOPO 10, 11, 13 and 14 were made and the plasmids were isolated. The concentration of the plasmid solutions were measured (Table 19) and sent for sequencing using M13 F/R primers.

Table 19: Concentrations of plasmids PP4-1 dTOPO 10, 11, 13 and 14 measured by the NanoDrop 2000

Sample	Measured concentration (ng/µl)	260/280	260/230
PP4-1 dTOPO 10	400.7	1.93	2.23
PP4-1 dTOPO 11	136.5	1.93	2.09
PP4-1 dTOPO 13	190.0	1.94	2.31
PP4-1 dTOPO 14	236.3	1.94	2.30

The sequencing results of PP4-1 dTOPO 10 and PP4-1 dTOPO 11 indicated that they had a mutation in the stop codon. The sequencing results of PP4-1 dTOPO 13 (Appendix A-6) and PP4-1 dTOPO 14 (Appendix A-7) did not have the mutation in the stop codon.

The restriction enzyme ApaI was found to cut outside the *att*-region and not in the gene. ApaI destroys the kanamycin resistance so that bacteria with the final destination vector pSF can be selected.

Digestion of PP4-1 dTOPO 13 plasmid was performed using ApaI. The product was run on a gel and cut out (Figure 32).

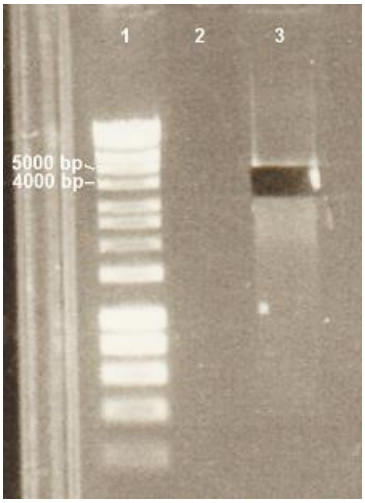

Figure 32: Gel electrophoresis of digested PP4-1 dTOPO 13 plasmid that has been cut out. Lane 1: Hyperladde I, lane 2: empty, lane 3: digested PP4-1 dTOPO 13 plasmid

Gel extraction of the cut out digested PP4-1 dTOPO 13 plasmid was performed and the concentration was measured to be 13.1 ng/µl.

The extracted PP4-1 dTOPO 13 plasmid was heated at 65° C for 10 min. LR Clonase was prepared by vortexing the solution. pSF vector (2 µl, 30 ng/µl) was mixed with PP4-1 dTOPO 13 plasmid (7 µl, 13.1 ng/µl) and put briefly on ice. LR Clonase (2µl) was added. The reaction was set to incubate at 25° C over-night. Proteinase K (1 µl) was added to the LR PP4-1 dTOPO 13-pSF and gently mixed. The solution was incubated at 37° C for 15 min. DH5α competent cells were transformed with the solution and plated with kanamycin. A control/comparison plate was made by transforming DH5α competent cells with digested PP4-1 dTOPO 13 plasmid. The plates were incubated at 37° C over-night. There was about equal amount of colonies on the plates and it was the decided that the experiment was to be repeated.

Digestion of PP4-1 dTOPO 13 plasmid was performed using ApaI. The product was run on a gel and cut out (Figure 33)

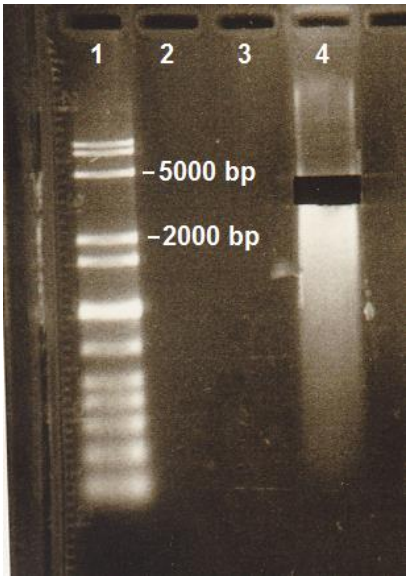

Figure 33: Gel electrophoresis of digested PP4-1 dTOPO 13 plasmid that has been cut out. Lane 1: GelPilot 1 kb Plus Ladder, lane 2 and 3: empty, lane 4: digested PP4-1 dTOPO 13 plasmid

Gel extraction of the gel piece was performed and the concentration was measured to be 6.4 ng/µl). An LR reaction was performed using pSF (3 µl, 60 ng/µl), extracted PP4-1 dTOPO 13 plasmid and LR Clonase (2.5 µl). The reaction was stopped after a day with proteinase K (1 µl), incubated at 37° C for 15 min and put on ice. DH5α competent cells were transformed with the solution. A control/comparison plate was made by transforming DH5α competent cells with digested PP4-1 dTOPO 13 plasmid. The solutions were plated with kanamycin. The plates was incubated at 37° C over-night.

A colony PCR of three colonies from the LR plate with PP4-1 dTOPO 13-pSF, three colonies from the control plate with PP4-1 dTOPO 13 and a pSF vector control was performed using the DreamTaq DNA polymerase kit and pSF F/R primers. The PCR products were run on a gel (Figure 34).

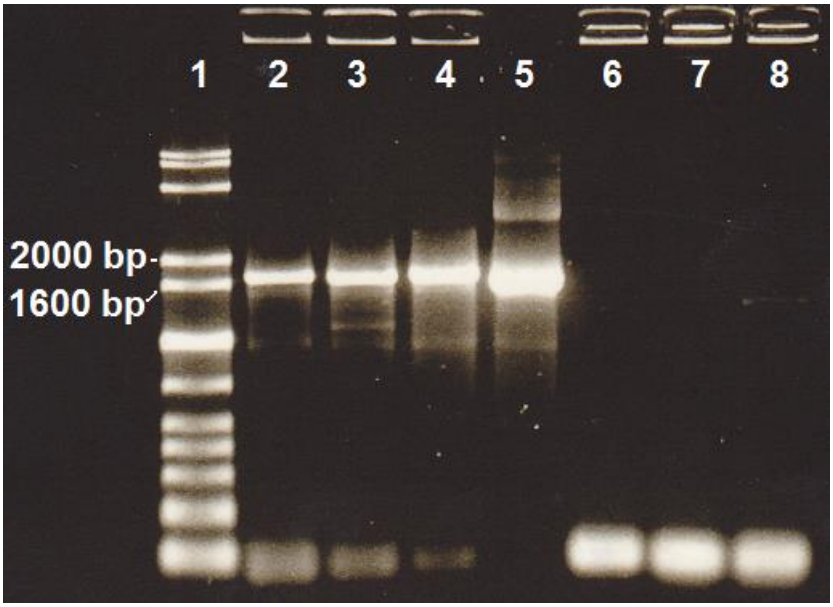

Figure 34: Gel electrophoresis of colony PCR products. Lane 1: GelPilot 1 kb Plus Ladder, lane: 2-4: PP4-1 dTOPO 13-pSF, lane 5: pSF vector control, lane 6-8: PP4-1 dTOPO 13

Figure 34 indicates that pSF has been transformed into the three colonies picked from the LR plate. Over-night cultures were made with colony 1 (PP4-1-pSF 1)(lane 2) and colony 2 (PP4-1-pSF 2)(lane 3) with kanamycin. Glycerol stock solutions were made from each bacteria solutions for long-term storage.

The plasmids of PP4-1-pSF 1 and PP4-1-pSF 2 were isolated. The concentration of the isolated plasmids were measured (Table 20).

Table 20: Concentrations of plasmids PP4-1-pSF 1 and PP4-1-pSF 2 measured by the NanoDrop2000

Sample	Measured concentration (ng/µl)	260/280	260/230
PP4-1-pSF 1	181.2	1.95	2.22
PP4-1-pSF 2	172.1	1.94	2.28

PP4-1-pSF 1 and PP4-1-pSF 2 plasmids were sent for sequencing using pSF F and PP4.1_dTOPO F primers for each. The sequencing failed to sequence using the pSF F primers and only the PP4.1_dTOPO F primers yielded readable results. The sequence with PP4-1-pSF 1 using the PP4.1_dTOPO F primer failed to sequence the start of the gene and there appears to be a mutation in the end of the gene (Appendix A-14). The sequence with PP4-1-pSF 2

using the PP4.1_dTOPO F primer has a similar result (Appendix A-15).

A new LR reaction was performed with PP4-1 dTOPO 13 plasmid (1.5 µl, 190 ng/µl) and pSF plasmid (3 µl, 80 ng/µl) using the Gateway® LR Clonase™ Enzyme Mix (Invirtogen). The two volumes of plasmid solutions were mixed and incubated at 65° C for 10 min. 5X LR Clonase Reaction Buffer was added. The solution was filled with water up to 8 µl. The LR Clonase enzyme mix was added and the reaction was incubated in 25° C over-night. Proteinase K (1 µl) was added and the solution was incubated at 37° C for 15 min. DH5α completent cells were transformed with the LR reaction product. The bacteria were spread on an agar plate with kanamycin and hygromycin and incubated at 37° C over-night. There was no growth on the plate.

The LR reaction was repeated using PP4-1 dTOPO 13 plasmid (2.5 µl, 190 ng/µl) and pSF plasmid (1.5 µl, 80 ng/µl). DH5α competent cells were transformed with the LR reaction product. The bacteria were spread on an agar plate with kanamycin and hygromycin and incubated at 37° C over-night. Colonies grew on the plate.

Colonies (1-4) were picked and a colony PCR was performed using the primersets pSF F/R, PP4.1_dTOPO F/R and M13 F/R. A pSF plasmid control, a PP4-1 dTOPO 13 plasmid control and a negative water control were included. The PCR products were run on a gel (Figure 35)

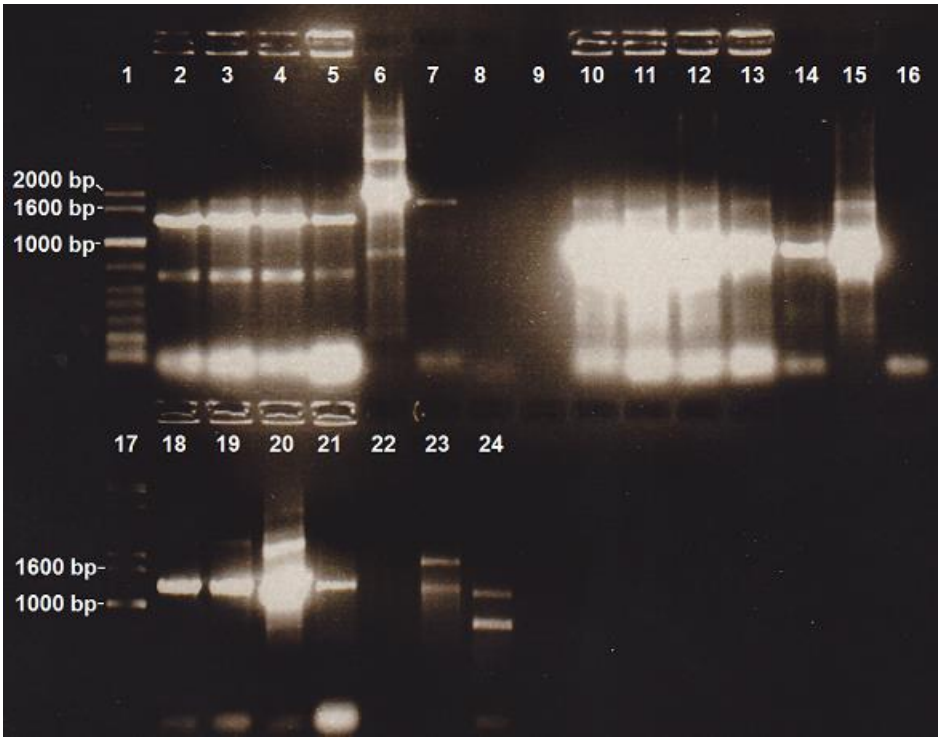

Figure 35: Gel electrophoresis of colony PCR products. Lane 1 and 17: GelPilot 1 kb Plus Ladder, lane 2-5: Colony 1-4 (pSF F/R), lane 6: pSF plasmid control (pSF F/R), lane 7: PP4-1 dTOPO 13 plasmid control (pSF F/R), lane 8: water control (pSF F/R), lane 9: empty, lane 10-13: Colony 1-4 (PP4.1_dTOPO F/R), lane 14: pSF plasmid control (PP4.1_dTOPO F/R), lane 15: PP4-1 dTOPO 13 plasmid control (PP4.1_dTOPO F/R), lane 16: water control (PP4.1_dTOPO F/R), lane 18-21: Colony 1-4 (M13 F/R), lane 22: pSF plasmid control (M13 F/R), lane 23: PP4-1 dTOPO 13 plasmid control (M13 F/R), lane 24: water control (M13 F/R)

Figure 35 indicate that the four colonies have pSF plasmids because bands can be observed in lane 2-5, but they are smaller than the pSF plasmid control band (lane 6). Lane 10-13 indicate that the colonies has the PP4-1 dTOPO gene in some form. The genes may be present in the pSF vector or PP4-1 dTOPO 13 plasmid originating from the LR reaction. The M13 primers could bind in either pENTR/D-TOPO vector or the pSF vector.

Over-night cultures of colonies 1-4 were made with kanamycin and hygromycin and incubated at 37° C. The plasmids (PP4-1-pSF A1, PP4-1-pSF A2, PP4-1-pSF A3, PP4-1-pSF A4) were isolated and the concentrations were measured (Table 21)

Table 21: Concentration of plasmids PP4-1-pSF A1 – A4 measured with the NanoDrop 2000

Sample	Measured concentration (ng/µl)	260/280	260/230
PP4-1-pSF A1	76.5	2.00	1.99
PP4-1-pSF A2	95.4	2.02	2.08
PP4-1-pSF A3	362.7	1.93	2.23
PP4-1-pSF A4	58.2	1.97	1.74

Because of the *Arabidopsis* was ready to be transformed and there was limited time, it was decided that the newly isolated plasmids of PP4-1-pSF were to be transformed into *A. tumefaciens* before the sequencing result was ready.

The sequence of PP4-1-pSF A3 (PP4-1-pSF A3 was used to transform *A. tumefaciens* and *A. thaliana*) did not have the start of the gene and there was a mutation in the stop codon region (Appendix A-21).

New PP4-1-pSF cultures were made from PP4-1-pSF A1, PP4-1-pSF A2 and PP4-1-pSF A4. The plasmids were isolated and the samples of the plasmids were set for sequencing. The sequencing results indicated that the PP4-1-pSF A1 had a good chance of being correct, but it was not conclusive. The multiple sequence alignment between PP4-1 and PP4-1-pSF A1 with PP4.1_dTOPO F is in Appendix A-24. PP4-1-pSF A1 bacterial culture was made again and the plasmids were isolated. The concentration was measured: 66.8 ng/µl. Samples were sent for sequencing with PP4.1_dTOPO R primer and PP4.1_MBP_R primer. The sequence alignment between PP4-1 and PP4-1-pSF A1 with PP4.1_dTOPO R is in Appendix A-23. PP4-1-pSF A1 with PP4.1_dTOPO F and PP4-1-pSF A1 with PP4.1_dTOPO R overlaps and indicate that the PP4-1 gene is in the plasmid.

The plasmid was used to transform *A. tumefaciens*.

PP4-1 pMAL-C2X plasmid

PP4-1 from the PP4-1 source vector was amplified with the Expand High Fidelity^{PLUS} PCR System and PP4.1_MBP_F and PP4.1_MBP_R primers. The PRC product was run on a gel and the bands with high concentration of PP4-1 were cut out before a photograph was taken (Figure 36).

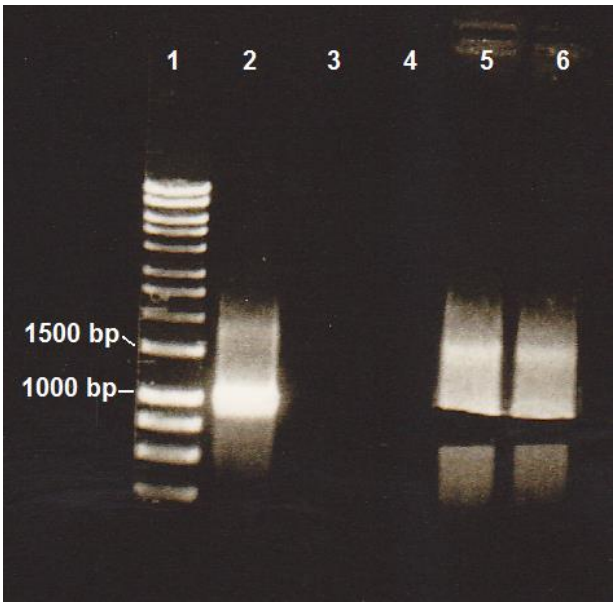

Figure 36: Gel electrophoresis of PP4-1 PCR product. Lane 1: Hyperladder I, lane 2: low concentration PP4-1, lane 5 and 6: PP4-1 cut out

Gel extraction of the cut out PP4-1 was performed using the Promega kit.

The extracted PP4-1 was digested with BamHI and HindIII restriction enzymes. pMAL-C2X in a solution was also digested with BamHI and HindIII. The solutions were incubated at 37° C over-night. PCR clean-up was performed with the two solutions. The concentration of the clean-up products were measured (Table 22). The low 260/230-ratios in table 22 indicate that there are impurities in the solution. The concentration was low, so it was decided that further purification was not beneficial.

Table 22: Concentrations of the PP4-1 solution and pMAL-C2X solution after PCR clean-up

Sample	Measured concentration (ng/µl)	260/280	260/230
PP4-1	32.4	1.86	0.27
pMAL-C2X	21.7	1.91	0.24

Ligation between PP4-1 and pMAL-C2X was performed for 1 h. The reaction was put in a 4° C refrigerator over-night.

JM109 competent cells were transformed with the ligation product. LB broth was added and the solution was incubated on a shaker at 37° C for 1 h. The bacterial solution was spread in two agar plates with ampicillin and incubated at 37° C over-night.

Colonies (1-8) were picked from the two incubated plates with the PP4-1-pMAL-C2X bacteria. A colony PCR was performed with the eight colonies, the PP4-1 source vector as a positive control and water instead of plasmid using PP4.1_MBP F/R primers. The PCR products were run on a gel (Figure 37).

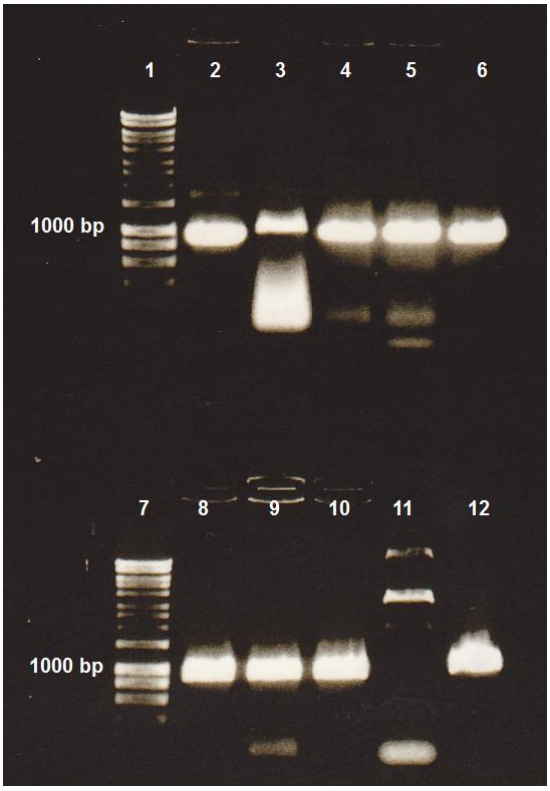

Figure 37: Gel electrophoresis of PCR products. Lane 1 and 7: Hyperladde I, lane 2-6, 8-10: PP4-1-pMAL-C2X colonies, lane 11: PP4-1 positive control, lane 12: negative control

PP4-1-pMAL-C2X 1 (lane 2), 3 (lane 4) , 4 (lane 5) , 5 (lane 6) , 6 (lane 8) , 7 (lane 9) and 8 (lane 10) in Figure 37 have clear bands ~1000 bp, indicating the presence of PP4-1. The positive control in lane 11 does not have a band with the expected size. The negative control

in lane 12 had a band ~1000 bp, indicating the presence of PP4-1. A mistake may have been made and the positive control and the negative control has been switched. It can also indicate that there were impurities in the water.

The colony PCR was repeated with the same PP4-1-pMAL-C2X colonies, one newly diluted PP4-1-pGEMT source vector positive control, negative control using the previously used water and a negative control using water from another bottle. The PP4.1_MBP F/R primers were used. The PCR products were run on a gel (Figure 38).

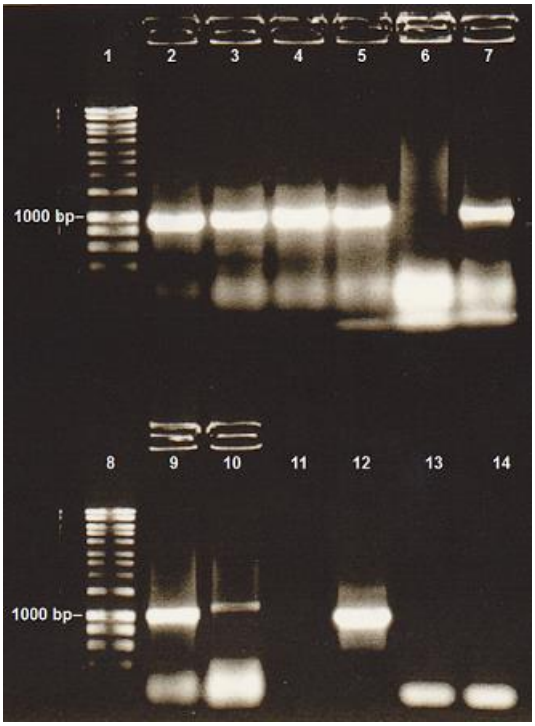

Figure 38: Gel electrophoresis of PCR products. Lane 1 and 8: Hyperladde I, lane 2-7, 9, 10: PP4-1-pMAL-C2X colonies, lane 11: empty, lane 12: PP4-1 postive control, lane 13: negative control 1, lane 14: negative control 2

In Figure 38, the positive PP4-1 control has a band ~1000 bp, which is expected. Both the negative controls with water show no sign of contamination. Over-night cultures of colonies 1-4 (lanes 2-5) were made with ampicillin and incubated on a shaker at 37° C.

The plasmids of culture 1-4 were isolated and the plasmid concentrations were measured (Table 23).

Table 23: Concentration of isolated plasmids from PP4-1-pMAL-C2X cultures 1-4

Sample	Measured concentration (ng/µl)	260/280	260/230
PP4-1-pMAL-C2X 1	71.3	1.99	1.94
PP4-1-pMAL-C2X 2	68.2	1.99	2.27
PP4-1-pMAL-C2X 3	78.0	1.99	2.10
PP4-1-pMAL-C2X 4	80.5	1.98	2.25

A PCR was performed with PP4-1-pMAL-C2X 1, 2, 3, 4 plasmids using PP4.1_MBP F and M13 F primers. A negative control was included. The PCR product were run on a gel (Figure 39).

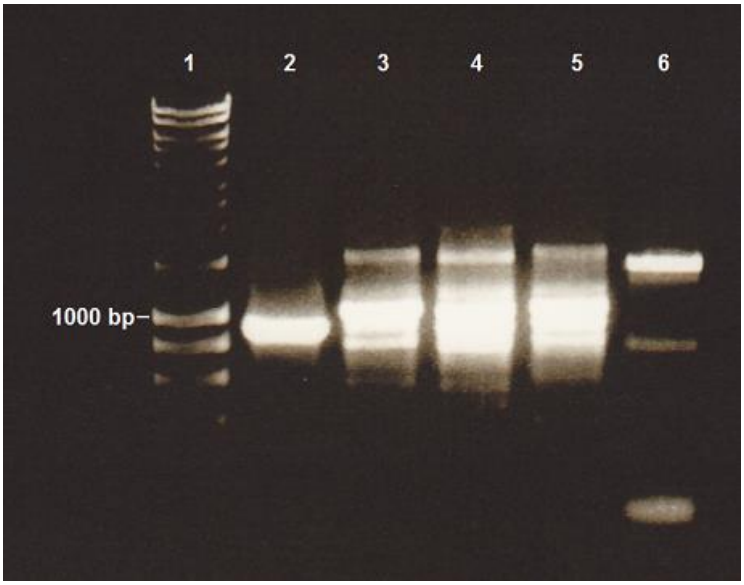

Figure 39: Gel electrophoresis of PCR products. Lane 1: Hyperladder I, lane 2-5: PP4-1-pMAL-C2X plasmids, lane 6: negative control

In Figure 39, PP4-1-C2X 1 plasmid has a clear band with a size ~1000 bp. The concentration of the plasmid solution was measured to 71.3 ng/µl. Two samples of it was sent for sequencing, both with the M13 F primer. The sequencing failed to sequence the last of the nucleotides (Appendix A-8)

New colonies (9-16) were picked from a PP4-1 pMAL-C2X plate. A colony PCR was performed using PP4.1_MBP F and M13 F primers. A negative control using water instead of colony solution was included. The PCR products were run on a gel (Figure 40).

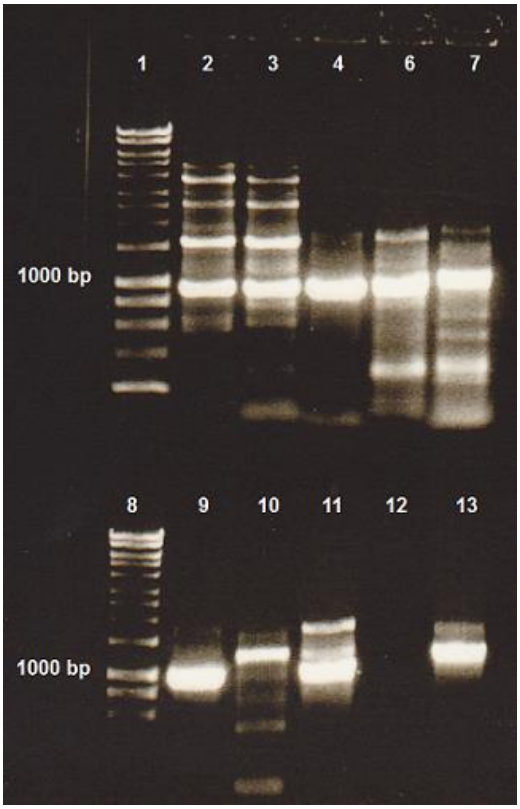

Figure 40: Gel electrophoresis of PCR products. Lane 1 and 8: Hyperladder I, lane 2-7, 9-11: PP4-1-pMAL-C2X 9-13, 14-16 , lane 12: empty, lane 13: negative control

From Figure 40, colonies 11 and 14 (lane 4 and 9) appeared to be the best because of the single band ~1000 bp long. The band in the negative control was noted, but the experiment continued.

Over-night cultures of PP4-1-pMAL-C2X 11 and 14 were made and the plasmids were isolated. The concentration of the plasmids were measured (Table 24).

Table 24: Concentration of PP4-1-pMAL-C2X 11 and 14 plasmids

Sample	Measured concentration (ng/µl)	260/280	260/230
PP4-1-pMAL-C2X 11	48.5	1.94	2.22
PP4-1-pMAL-C2X 14	39.5	1.90	2.06

A PCR was run to confirm the plasmids. PP4-1-pMAL-C2X 11 and PP4-1-pMAL-C2X 14 were run with M13 F and PP4.1_MBP F primers. A negative control was included. Different volumes of the PCR products were run on a gel (Figure 41).

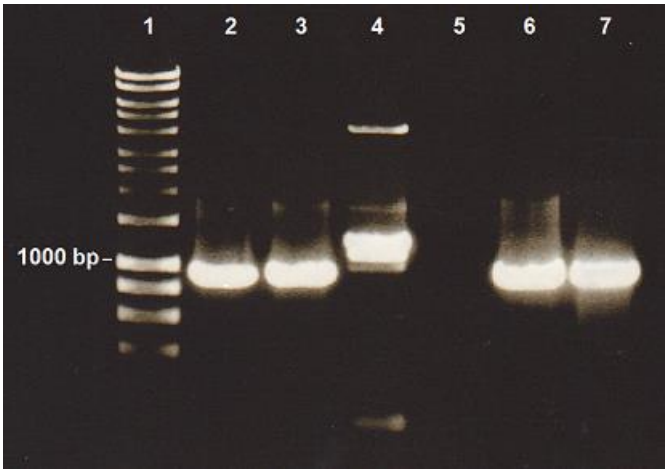

Figure 41: Gel electrophoresis of PCR products. Lane 1: Hyperladder I, lane 2: 2 µl PP4-1-pMAL-C2X 11, lane 3: 2 µl PP4-1-pMAL-C2X 14, lane 4: negative control, lane 5: empty, lane 6: 1 µl PP4-1-pMAL-C2X 11, 1 µl PP4-1-pMAL-C2X 14

Figure 41 indicate that PP4-1-pMAL-C2X 11 and PP4-1-pMAL-C2X 14 has good quality.

Several plasmid samples were sent to be sequenced with the following primers:

- PP4-1- pMAL-C2X 1, PP4.1_MBP F primer
- PP4-1- pMAL-C2X 2, PP4.1_MBP F primer
- PP4-1-pMAL-C2X 11, M13 F and PP4.1_MBP F primers
- PP4-1-pMAL-C2X 14, M13 F and PP4.1_MBP F primers

The sequencing of PP4-1- pMAL-C2X 1 with the PP4.1_MBP F primer (Appendix A-9) shows the end of the gene matches the PP4-1 gene, but the sequence lacks the beginning of the gene. The sequencing of PP4-1- pMAL-C2X 1 with the M13 F primer in A-8 has the start of the gene. The sequences overlap each other indicating PP4-1- pMAL-C2X 1 is positive and has the PP4-1 gene inside the pMAL-C2X vector.

The sequencing of PP4-1- pMAL-C2X 2 with the PP4.1_MBP F primer was uninterpretable.

The sequencing of PP4-1- pMAL-C2X 11 with the PP4.1_MBP F primer (Appendix A-10) overlaps with the sequencing with M13 F (Appendix A-11). Together they match the sequence of PP4-1, indicating that it is positive with the PP4-1 gene inside the pMAL-C2X

vector.

The sequencing of PP4-1- pMAL-C2X 14 with the PP4.1_MBP F primer (Appendix A-12) overlaps with the sequencing with M13 F (Appendix A-13). Together they match the sequence of PP4-1, indicating that it is positive with the PP4-1 gene inside the pMAL-C2X vector.

Agrobacterium tumefaciens transformation with PP4.1-pSF and PP4-2-pSF

Tubes with *Agrobacterium tumefaciens* were thawed on ice and transformed with PP4-1-pSF A2 (10 µl, 95.4 ng/µl), PP4-1-pSF A3 (2.7 µl, 362.7 ng/µl), PP4-2-pSF A (11 µl, 86.2 ng/µl) and PP4-2-pSF B (16 µl, 60 ng/µl).

The bacterial solutions were spread on agar plates with kanamycin and hygromycin and incubated at 28° C for 48 h. Over-night cultures were made from the bacteria on the four plates:

- PP4-1-pSF A2 → PP4-1_1 A, B, C, D
- PP4-1-pSF A3 → PP4-1_2 A, B, C, D
- PP4-2-pSF A → PP4-2_1 A, B, C, D
- PP4-2-pSF B → PP4-2_2 A, B, C, D

The cultures were incubated at 28° C with kanamycin and hygromycin over-night.

Growth was observed in PP4-1_2 A, B, C, D, PP4-2_1 A, B, C, D and PP4-2_2 A, B, C, D. The cultures (500 µl) were pelleted and the supernatant was pipetted out. The pellets were resuspended in 20 µl water and the solutions were boiled at 95° C for 5 min.

Glycerol stock solutions of each culture were made for long-term storage.

A colony PCR was run with the 12 boiled cultures using gene specific primers (PP4.1_dTOPO F/R and PP4.2_dTOPO F/R) and pSF F/R primers. pSF vector controls and reaction mix controls were included.

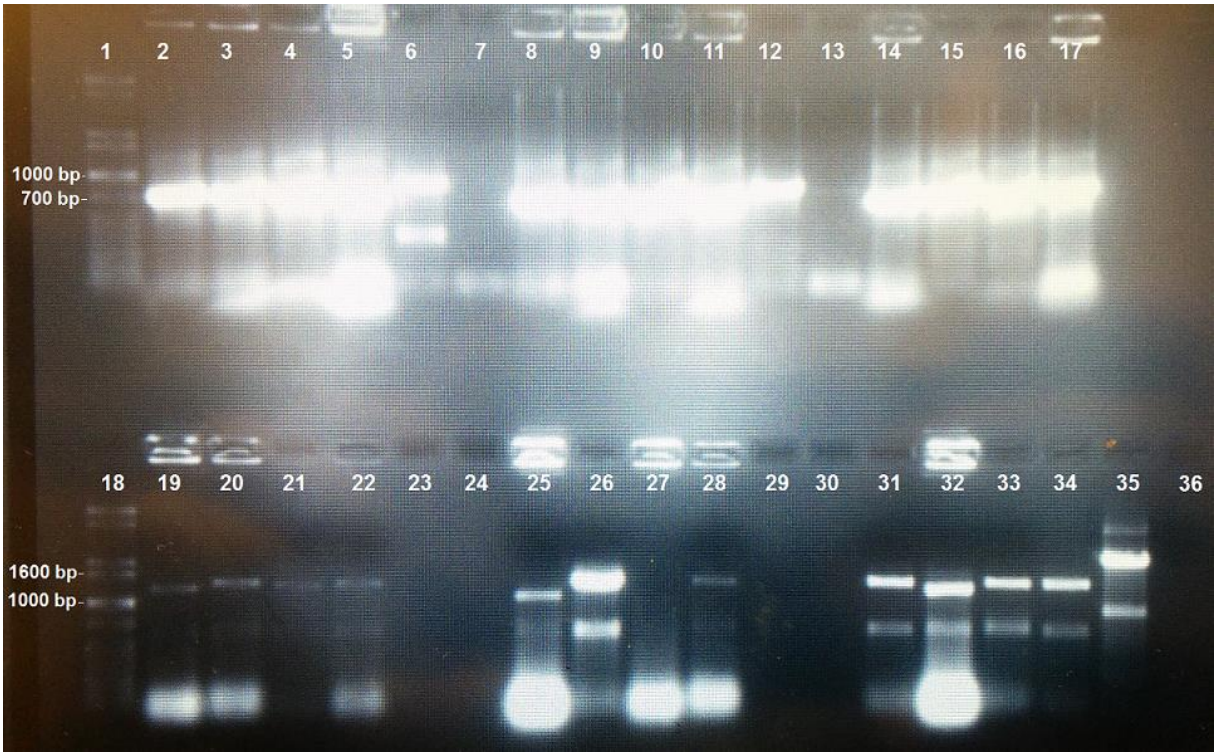

Figure 42: Gel electrophoresis of colony PCR products of transformed *A. tumefaciens* cultures. Explanation of lanes in table 25

Table 25: Explanation of lanes in Figure 42, primers in parenthesis

Lane 1	GelPilot 1 kb Plus Ladder	Lane 13	Reaction mix control (PP4.2_dTOPO F/R)	Lane 25	PP4-2_1 A (pSF F/R)
Lane 2	PP4-1_2 A (PP4.1_dTOPO F/R)	Lane 14	PP4-2_2 A (PP4.2_dTOPO F/R)	Lane 26	PP4-2_1 B (pSF F/R)
Lane 3	PP4-1_2 B (PP4.1_dTOPO F/R)	Lane 15	PP4-2_2 B (PP4.2_dTOPO F/R)	Lane 27	PP4-2_1 C (pSF F/R)
Lane 4	PP4-1_2 C (PP4.1_dTOPO F/R)	Lane 16	PP4-2_2 C (PP4.2_dTOPO F/R)	Lane 28	PP4-2_1 D (pSF F/R)
Lane 5	PP4-1_2 D (PP4.1_dTOPO F/R)	Lane 17	PP4-2_2 D (PP4.2_dTOPO F/R)	Lane 29	Empty
Lane 6	pSF vector control (PP4.1_dTOPO F/R)	Lane 18	GelPilot 1 kb Plus Ladder	Lane 30	Empty
Lane 7	Reaction mix control (PP4.1_dTOPO F/R)	Lane 19	PP4-1_2 A (pSF F/R)	Lane 31	PP4-2_2 A (pSF F/R)
Lane 8	PP4-2_1 A (PP4.2_dTOPO F/R)	Lane 20	PP4-1_2 B (pSF F/R)	Lane 32	PP4-2_2 B (pSF F/R)
Lane 9	PP4-2_1 B (PP4.2_dTOPO F/R)	Lane 21	PP4-1_2 C (pSF F/R)	Lane 33	PP4-2_2 C (pSF F/R)
Lane 10	PP4-2_1 C (PP4.2_dTOPO F/R)	Lane 22	PP4-1_2 D (pSF F/R)	Lane 34	PP4-2_2 D (pSF F/R)
Lane 11	PP4-2_1 D (PP4.2_dTOPO F/R)	Lane 23	Empty	Lane 35	pSF vector control (pSF F/R)
Lane 12	pSF vector control (PP4.1_dTOPO F/R)	Lane 24	Empty	Lane 36	Reaction mix control (pSF F/R)

PP4-1_2 A (lane 2/19) was chosen for *Arabidopsis* transformation because the band in lane 2 in Figure 42 was the clearest of the four PP4-1 bands between 1000 bp and 700 bp range. PP4-1_2 A had a weak band in lane 19, indicating that the pSF might be present, but it can also be an error. pSF vector control in lane 6 has two bands, indicating that the PP4.1_dTOPO primers bind in the pSF.

PP4-2_2 C (lane 16/33) was chosen for *Arabidopsis* transformation because the band in lane 16 was between the 1000 bp and 700 bp range, indicating the presence of the gene. The band in lane 33 was clear and a good indication that the pSF is present.

Colony PP4-1_2 A (glycerol stock solution) and colony PP4-2_2 C (glycerol stock solution) was added to two flasks of LB broth (150 ml). Kanamycin (150 µl, 50 ng/µl) and hygromycin (120 µl, 40 ng/µl) were added to the flasks. The flasks were incubated at 28° C over-night.

Floral dipping was performed, transforming *A. thaliana* with PP4-1_2 A and PP4-2_2 C (Two pots, each with five plants per gene).

The plants were grown and the plants with PP4-1 were discarded when it was learned the sequence was wrong.

A. tumefaciens was transformed with PP4-1-pSF A1 plasmid shown to have the correct sequence. *A. thaliana* was transformed by floral dipping.

Plant selection

The *A. thaliana* transformed with PP4-2_2 C produced seeds. The seeds were harvested and sterilized. Agar plates with 1/2 MS-media 1 % sucrose and kanamycin (200 µl, 50 ng/µl in 400 ml 1/2 MS 1 % sucrose) were made. Seeds from the PP4-2 plants were spread on the agar plates. Wild type seed were spread on a plate with kanamycin and one without (Figure 43) to observe the effect of the kanamycin. The plates were put in a cold room for stratification. The plants were grown for ~10 days. Isolated green sprouts indicate they have received the transfer DNA with the kanamycin resistance gene and the PP4-2 gene from the *Agrobacterium* (Figure 44). The green sprouts were transferred to pots and grown (Figure 45)

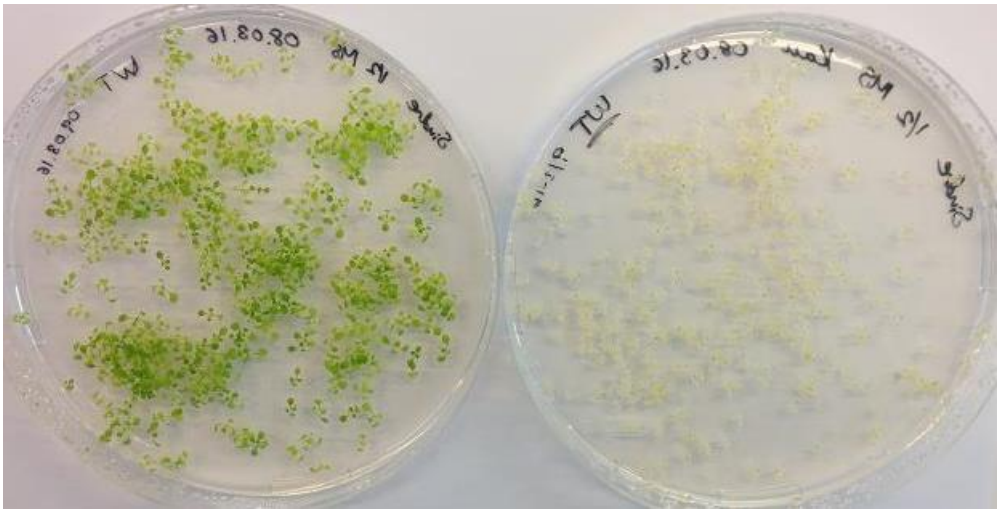

Figure 43: Petri dishes with 1/2 MS 1 % sucrose with A. thaliana (WT). Left plate has no kanamycin, right plate has kanamycin

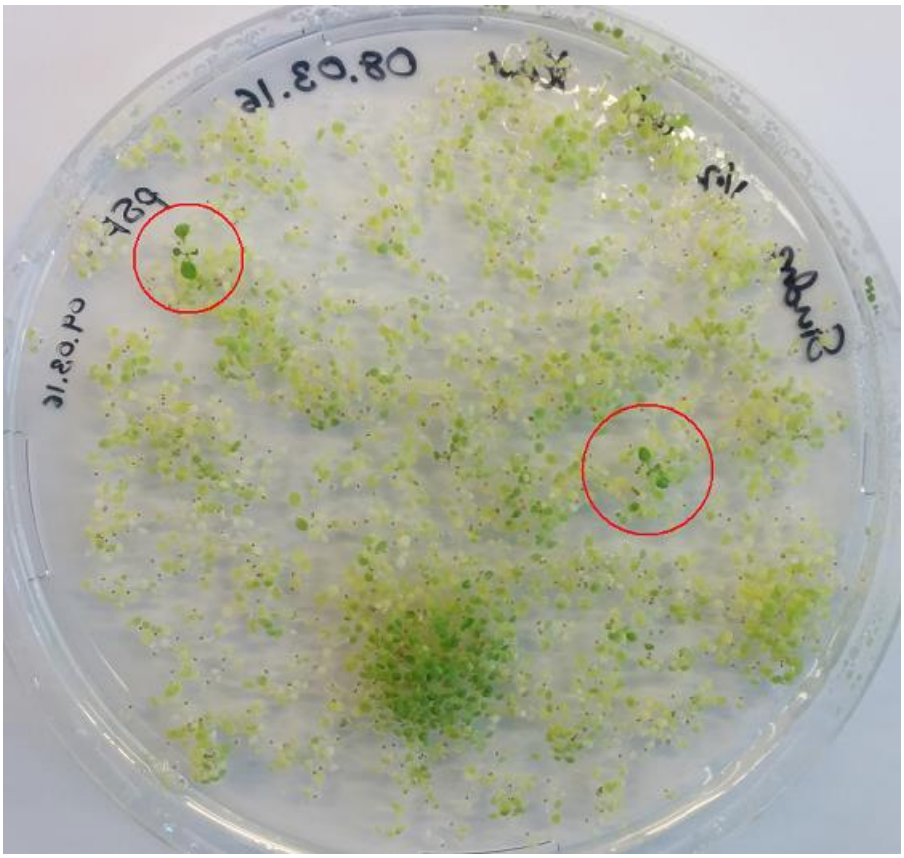

Figure 44: Sprouts grown on 1/2 MS 1 % sucrose and kanamycin.

Figure 45: Sprouts with PP4-2 plants growing. Two sprouts with pSF plants also grows.

Discussion

There is a good chance the PP4-2 protein was successfully produced and visualized on the polyacrylamide gel, demonstrating the method and that it can be used in similar experiments in the future.

The PP4-2 seems have been successfully transformed into the *Arabidopsis* plants. The plants produced seeds that grew on kanamycin. The plants were transferred to soil, but further testing is required to confirm that they produce seeds that have PP4-2.

The PP4-2 had a known mutation from the pCAT source vector. The mutation causes an alanine to be translated instead of a valine. Alanine is not very different from valine, so it will not make not cause major changes in the protein or its function.

The PP4-1 was transformed into the *Arabidopsis*. It remains to be seen if it can produce seeds that will survive on kanamycin and thus indicate the experiment has been a success.

The PP4-1-pMAL-C2X plasmid has been made and can be used in further studies. For example protein production with Rosetta.

References

Berg, J. M., Tymoczko, J. L., Stryer, L., 2012. *Biochemistry* New York: W. H. Freeman and Company

Pujol, G., Baskin, T. I., Casamayor, A., Cortadellas, N., Ferrer, A., Ariño, J., 2000 The Arabidopsis thaliana PPX/PP4 phosphatases: molecular cloning and structural organization of the genes and immunolocalization of the proteins to plastids. *Plant Molecular Biology*. Volume 44, Issue 4 pp 499-511

Lillo, C., Kataya, A. M. R., Heidari, B., Creighton, M. T., Nemie-Feyissa, D., Gimbot, Z., Jonassen, E. M., 2014 Protein phosphatases PP2A, PP4 and PP6: mediators and regulators in development and responses to environmental cues. *Plant, Cell and Environment*

Appendix

A-1: PP4-1 sequence, 5'-3' direction

```
ATGTCAGACCTAGATCGGCAAATAGGGCAGCTTAAGCGATGCGAACCATTGAGCGAATCGGAGGTGAAGGCTCTT
TGCCTCAAAGCCATGGAAATTCTTGTTGAAGAGAGTAATGTTCAGAGAGTTGATGCCCCTGTCACTTTATGTGGT
GACATCCATGGGCAGTTCTATGATATGATGGAGCTTTTCAAAGTTGGGGGTGATTGTCCTAAGACCAACTATTTG
TTTATGGGAGATTTTGTTGATCGTGGATATTATTCGGTTGAGACATTTCTACTTCTACTCGCACTCAAGGTTAGA
TATCCAGACCGCATAACTCTCATCAGAGGAAACCATGAAAGCAGGCAAATCACACAGGTTTATGGATTTTATGAT
GAGTGTTTGCGTAAATATGGCTCTTCAAATGTCTGGAGATACTGCACCGACATTTTTGACTACATGAGTCTTTCA
GCTGTTGTGGAGAACAAGATATTCTGTGTTCATGGTGGTCTTTCTCCAGCTATTATGACTCTTGATCAGATTAGG
ACAATTGACCGGAAGCAAGAAGTACCACATGATGGTGCCATGTGTGATCTCCTATGGTCTGATCCTGAAGATATT
GTTGATGGCTGGGGATTGAGCCCTCGTGGTGCCGGATTCCTTTTTGGTGGCAGTGTTGTCACGTCTTTTAACCAC
TCAAACAACATAGACTACATAGCCCGTGCCCATCAACTAGTTATGGAGGGTTACAAATGGATGTTTGATAGCCAG
ATTGTGACAGTGTGGTCAGCCCCAAATTACTGTTACAGATGCGGTAATGTGGCTTCAATTCTAGAGCTTGACGAG
AATCTAAATAAAGAATTCCGTGTGTTTGATGCAGCCCAGCAGGACTCGAGAGGGCCTCCCGCCAAAAAGCCGGCC
CCTGATTACTTCCTATAA
```

A-2: PP4-2 sequence, 5'-3' direction

```
ATGTCAGACCTAGACAAGCAAATAGAGCAGCTTAAACGCTGCGAGGCTTTGAAGGAATCAGAAGTGAAGGCTCTT
TGTCTTAAAGCTATGGAGATTCTAGTTGAAGAGAGCAATGTTCAAAGAGTCGATGCTCCTGTCACTATATGTGGC
GACATTCATGGACAGTTCTATGACATGAAAGAGCTTTTCAAAGTTGGGGGTGATTGCCCTAAGACCAATTATTTG
TTTCTTGGAGATTTTGTTGACCGAGGTTTTTATTCGGTTGAGACATTTCTACTTCTTCTAGCTCTCAAGGTTAGA
TATCCAGACCGTATAACTCTCATTAGAGGGAACCACGAGAGCCGGCAGATTACGCAGGTATATGGATTTTATGAT
GAGTGTCTGCGTAAATATGGCTCTGTAAATGTTTGGAGATACTGCACAGATATCTTTGACTACTTGAGTCTTTCA
GCTCTTGTCGAGAACAAGATATTTTGTGTTCATGGAGGTCTCTCTCCAGCTATTATGACTCTAGACCAGATCAGG
GCTATTGATCGAAAGCAAGAAGTACCACATGATGGTGCTATGTGTGATCTTTTATGGTCTGATCCAGAAGATATT
GTCGATGGTTGGGGATTGAGTCCCCGTGGTGCCGGATTCCTTTTCGGCGGCAGTGTTGTTACGTCTTTTAACCAC
TCAAACAACATTGATTACATATGTCGAGCTCATCAGCTAGTGATGGAAGGTTACAAATGGATGTTCAATAGCCAG
ATAGTCACTGTTTGGTCTGCCCCAAATTACTGTTATAGATGCGGTAATGTAGCTGCAATTCTAGAGCTCGATGAG
AATCTAAACAAAGAGTTTCGTGTCTTCGATGCAGCCCCACAAGAATCGAGAGGAGCTCTAGCCAAGAAACCTGCA
CCTGATTATTTCCTGTGA
```

A-3: Multiple sequence alignment between PP4-2 sequence and PP4-2 from isolated plasmids from transformed JM109 colony 1 (used to transform Rosetta), one apparent mutation

```
PP4-2           ----------------------------------ATGTCAGACCTAGACAAGCAAATAGA
PP4-2_1_FP      TTCTGAGAATCTTTATTTTCAGGGCGCCATGGCAATGTCAGACCTAGACAAGCAAATAGA
                                                   **************************

PP4-2           GCAGCTTAAACGCTGCGAGGCTTTGAAGGAATCAGAAGTGAAGGCTCTTTGTCTTAAAGC
PP4-2_1_FP      GCAGCTTAAACGCTGCGAGGCTTTGAAGGAATCAGAAGTGAAGGCTCTTTGTCTTAAAGC
                ************************************************************

PP4-2           TATGGAGATTCTAGTTGAAGAGAGCAATGTTCAAAGAGTCGATGCTCCTGTCACTATATG
PP4-2_1_FP      TATGGAGATTCTAGTTGAAGAGAGCAATGTTCAAAGAGTCGATGCTCCTGCCACTATATG
                ********************************************** *********

PP4-2           TGGCGACATTCATGGACAGTTCTATGACATGAAAGAGCTTTTCAAAGTTGGGGGTGATTG
PP4-2_1_FP      TGGCGACATTCATGGACAGTTCTATGACATGAAAGAGCTTTTCAAAGTTGGGGGTGATTG
                ************************************************************

PP4-2           CCCTAAGACCAATTATTTGTTTCTTGGAGATTTTGTTGACCGAGGTTTTTATTCGGTTGA
PP4-2_1_FP      CCCTAAGACCAATTATTTGTTTCTTGGAGATTTTGTTGACCGAGGTTTTTATTCGGTTGA
                ************************************************************

PP4-2           GACATTTCTACTTCTTCTAGCTCTCAAGGTTAGATATCCAGACCGTATAACTCTCATTAG
PP4-2_1_FP      GACATTTCTACTTCTTCTAGCTCTCAAGGTTAGATATCCAGACCGTATAACTCTCATTAG
                ************************************************************
```

```
PP4-2           AGGGAACCACGAGAGCCGGCAGATTACGCAGGTATATGGATTTTATGATGAGTGTCTGCG
PP4-2_1_FP      AGGGAACCACGAGAGCCGGCAGATTACGCAGGTATATGGATTTTATGATGAGTGTCTGCG
                ************************************************************

PP4-2           TAAATATGGCTCTGTAAATGTTTGGAGATACTGCACAGATATCTTTGACTACTTGAGTCT
PP4-2_1_FP      TAAATATGGCTCTGTAAATGTTTGGAGATACTGCACAGATATCTTTGACTACTTGAGTCT
                ************************************************************

PP4-2           TTCAGCTCTTGTCGAGAACAAGATATTTTGTGTTCATGGAGGTCTCTCTCCAGCTATTAT
PP4-2_1_FP      TTCAGCTCTTGTCGAGAACAAGATATTTTGTGTTCATGGAGGTCTCTCTCCAGCTATTAT
                ************************************************************

PP4-2           GACTCTAGACCAGATCAGGGCTATTGATCGAAAGCAAGAAGTACCACATGATGGTGCTAT
PP4-2_1_FP      GACTCTAGACCAGATCAGGGCTATTGATCGAAAGCAAGAAGTACCACATGATGGTGCTAT
                ************************************************************

PP4-2           GTGTGATCTTTTATGGTCTGATCCAGAAGATATTGTCGATGGTTGGGGATTGAGTCCCCG
PP4-2_1_FP      GTGTGATCTTTTATGGTCTGATCCAGAAGATATTGTCGATGGTTGGGGATTGAGTCCCCG
                ************************************************************

PP4-2           TGGTGCCGGATTCCTTTTCGGCGGCAGTGTTGTTACGTCTTTTAACCACTCAAACAACAT
PP4-2_1_FP      TGGTGCCGGATTCCTTTTCGGCGGCAGTGTTGTTACGTCTTTTAACCACTCAAACAACAT
                ************************************************************

PP4-2           TGATTACATATGTCGAGCTCATCAGCTAGTGATGGAAGGTTACAAATGGATGTTCAATAG
PP4-2_1_FP      TGATTACATATGTCGAGCTCATCAGCTAGTGATGGAAGGTTACAAATGGATGTTCAATAG
                ************************************************************

PP4-2           CCAGATAGTCACTGTTTGGTCTGCCCCAAATTACTGTTATAGATGCGGTAATGTAGCTGC
PP4-2_1_FP      CCAGATAGTCACTGTTTGGTCTGCCCCAAATTACTGTTATAGATGCGGTAATGTAGCTGC
                ************************************************************

PP4-2           AATTCTAGAGCTCGATGAGAATCTAAACAAAGAGTTTCGTGTCTTCGATGCAGCCCCACA
PP4-2_1_FP      AATTCTAGAGCTCGATGAGAATCTAAACAAAGAGTTTCGTGTCTTCGATGCAGCCCCACA
                ************************************************************

PP4-2           AGAATCGAGAGGAGCTCTAGCCAAGAAACCTGCACCTGATTATTTCCTGTGA--------
PP4-2_1_FP      AGAATCGAGAGGAGCTCTAGCCAAGAAACCTGCACCTGATTATTTCCTGTGAGGTACCGG
                *************************************************

PP4-2           ------------------------------------------------------------
PP4-2_1_FP      ATCCGAATTCGAGCTCCGTCGACAAGCTTGCGGCCGCACTCGAGCACCACCACCACCACC


PP4-2           ------------------------------------------------------------
PP4-2_1_FP      ACTGAGATCCGGCTGCTAACAAAGCCCGAAAGGAAGCTGAATTGGCTGCTGCCACCGCTG


PP4-2           ------------------------------------------------------------
PP4-2_1_FP      AACAATAACTAACATAACCCCTTGGGGCCTCTAAACGGGTCTTGAGGGGTTTTTTGCTGA


PP4-2           -----------
PP4-2_1_FP      AAGGARGAACT
```

A-4: Multiple sequence alignment between PP4-2 sequence and PP4-2 dTOPO C2, one apparent mutation

```
PP4-2             ------------------------------------------------------------
PP4-2_C2_M13F     TGACCTGTTCGTTGCAACAAATTGATGAGCAATGCTTTTTTATAATGCCAACTTTGTACA


PP4-2             ---------------------------------ATGTCAGACCTAGACAAGCAAATAGAGC
PP4-2_C2_M13F     AAAAAGCAGGCTCCGCGGCCGCCCCCTTCACCATGTCAGACCTAGACAAGCAAATAGAGC
```

```
                                                          ****************************

PP4-2                 AGCTTAAACGCTGCGAGGCTTTGAAGGAATCAGAAGTGAAGGCTCTTTGTCTTAAAGCTA
PP4-2_C2_M13F         AGCTTAAACGCTGCGAGGCTTTGAAGGAATCAGAAGTGAAGGCTCTTTGTCTTAAAGCTA
                      ************************************************************

PP4-2                 TGGAGATTCTAGTTGAAGAGAGCAATGTTCAAAGAGTCGATGCTCCTGTCACTATATGTG
PP4-2_C2_M13F         TGGAGATTCTAGTTGAAGAGAGCAATGTTCAAAGAGTCGATGCTCCTGCCACTATATGTG
                      ********************************************** ***********

PP4-2                 GCGACATTCATGGACAGTTCTATGACATGAAAGAGCTTTTCAAAGTTGGGGGTGATTGCC
PP4-2_C2_M13F         GCGACATTCATGGACAGTTCTATGACATGAAAGAGCTTTTCAAAGTTGGGGGTGATTGCC
                      ************************************************************

PP4-2                 CTAAGACCAATTATTTGTTTCTTGGAGATTTTGTTGACCGAGGTTTTTATTCGGTTGAGA
PP4-2_C2_M13F         CTAAGACCAATTATTTGTTTCTTGGAGATTTTGTTGACCGAGGTTTTTATTCGGTTGAGA
                      ************************************************************

PP4-2                 CATTTCTACTTCTTCTAGCTCTCAAGGTTAGATATCCAGACCGTATAACTCTCATTAGAG
PP4-2_C2_M13F         CATTTCTACTTCTTCTAGCTCTCAAGGTTAGATATCCAGACCGTATAACTCTCATTAGAG
                      ************************************************************

PP4-2                 GGAACCACGAGAGCCGGCAGATTACGCAGGTATATGGATTTTATGATGAGTGTCTGCGTA
PP4-2_C2_M13F         GGAACCACGAGAGCCGGCAGATTACGCAGGTATATGGATTTTATGATGAGTGTCTGCGTA
                      ************************************************************

PP4-2                 AATATGGCTCTGTAAATGTTTGGAGATACTGCACAGATATCTTTGACTACTTGAGTCTTT
PP4-2_C2_M13F         AATATGGCTCTGTAAATGTTTGGAGATACTGCACAGATATCTTTGACTACTTGAGTCTTT
                      ************************************************************

PP4-2                 CAGCTCTTGTCGAGAACAAGATATTTTGTGTTCATGGAGGTCTCTCTCCAGCTATTATGA
PP4-2_C2_M13F         CAGCTCTTGTCGAGAACAAGATATTTTGTGTTCATGGAGGTCTCTCTCCAGCTATTATGA
                      ************************************************************

PP4-2                 CTCTAGACCAGATCAGGGCTATTGATCGAAAGCAAGAAGTACCACATGATGGTGCTATGT
PP4-2_C2_M13F         CTCTAGACCAGATCAGGGCTATTGATCGAAAGCAAGAAGTACCACATGATGGTGCTATGT
                      ************************************************************

PP4-2                 GTGATCTTTTATGGTCTGATCCAGAAGATATTGTCGATGGTTGGGGATTGAGTCCCCGTG
PP4-2_C2_M13F         GTGATCTTTTATGGTCTGATCCAGAAGATATTGTCGATGGTTGGGGATTGAGTCCCCGTG
                      ************************************************************

PP4-2                 GTGCCGGATTCCTTTTCGGCGGCAGTGTTGTTACGTCTTTTAACCACTCAAACAACATTG
PP4-2_C2_M13F         GTGCCGGATTCCTTTTCGGCGGCAGTGTTGTTACGTCTTTTAACCACTCAAACAACATTG
                      ************************************************************

PP4-2                 ATTACATATGTCGAGCTCATCAGCTAGTGATGGAAGGTTACAAATGGATGTTCAATAGCC
PP4-2_C2_M13F         ATTACATATGTCGAGCTCATCAGCTAGTGATGGAAGGTTACAAATGGATGTTCAATAGCC
                      ************************************************************

PP4-2                 AGATAGTCACTGTTTGGTCTGCCCCAAATTACTGTTATAGATGCGGTAATGTAGCTGCAA
PP4-2_C2_M13F         AGATAGTCACTGTTTGGTCTGCCCCAAATTACTGTTATAGATGCGGTAATGTAGCTGCAA
                      ************************************************************

PP4-2                 TTCTAGAGCTCGATGAGAATCTAAACAAAGAGTTTCGTGTCTTCGATGCAGCCCCACAAG
PP4-2_C2_M13F         TTCTAGAGCTCGATGAGAATCTAAACAAAGAGTTTCGTGTCTTCGATGCAGCCCCACAAG
                      ************************************************************

PP4-2                 AATCGAGAGGAGCTCTAGCCAAGAAACCTGCACCTGATTATTTCCTGTGA----------
PP4-2_C2_M13F         AATCGAGAGGAGCTCTAGCCAAGAAACCTGCACCTGATTATTTCCTGTGAAAGGGTGGGC
                      ************************************************

PP4-2                 ------------------------------------------------------------
PP4-2_C2_M13F         GCGCCGACCCAGCTTTCTTGTACAAAGTTGGCATTATAAGAAAGCATTGCTTATCAATTT


PP4-2                 --------------------------------------------------
PP4-2_C2_M13F         GTTGCAACGAACAGGTCMCTATCAGTCAAAATAAAATCATTATTTGSCAT
```

A-5: Multiple sequence alignment between PP4-1 sequence and PP4-1 dTOPO 4, mutation of stop codon

```
PP4-1             ------------------------------------------------------------
PP4-1_4_M13R      AAATAATGATTTTATTTTGACTGATAGTGACCTGTTCGTTGCAACAAATTGATAAGCAAT


PP4-1             ------------------------------------------------------------A
PP4-1_4_M13R      GCTTTCTTATAATGCCAACTTTGTACAAGAAAGCTGGGTCGGCGCGCCCACCCTTCACCA
                                                                             *

PP4-1             TGTCAGACCTAGATCGGCAAATAGGGCAGCTTAAGCGATGCGAACCATTGAGCGAATCGG
PP4-1_4_M13R      TGTCAGACCTAGATCGGCAAATAGGGCAGCTTAAGCGATGCGAACCATTGAGCGAATCGG
                  ************************************************************

PP4-1             AGGTGAAGGCTCTTTGCCTCAAAGCCATGGAAATTCTTGTTGAAGAGAGTAATGTTCAGA
PP4-1_4_M13R      AGGTGAAGGCTCTTTGCCTCAAAGCCATGGAAATTCTTGTTGAAGAGAGTAATGTTCAGA
                  ************************************************************

PP4-1             GAGTTGATGCCCCTGTCACTTTATGTGGTGACATCCATGGGCAGTTCTATGATATGATGG
PP4-1_4_M13R      GAGTTGATGCCCCTGTCACTTTATGTGGTGACATCCATGGGCAGTTCTATGATATGATGG
                  ************************************************************

PP4-1             AGCTTTTCAAAGTTGGGGGGTGATTGTCCTAAGACCAACTATTTGTTTATGGGAGATTTTG
PP4-1_4_M13R      AGCTTTTCAAAGTTGGGGGGTGATTGTCCTAAGACCAACTATTTGTTTATGGGAGATTTTG
                  ************************************************************

PP4-1             TTGATCGTGGATATTATTCGGTTGAGACATTTCTACTTCTACTCGCACTCAAGGTTAGAT
PP4-1_4_M13R      TTGATCGTGGATATTATTCGGTTGAGACATTTCTACTTCTACTCGCACTCAAGGTTAGAT
                  ************************************************************

PP4-1             ATCCAGACCGCATAACTCTCATCAGAGGAAACCATGAAAGCAGGCAAATCACACAGGTTT
PP4-1_4_M13R      ATCCAGACCGCATAACTCTCATCAGAGGAAACCATGAAAGCAGGCAAATCACACAGGTTT
                  ************************************************************

PP4-1             ATGGATTTTATGATGAGTGTTTGCGTAAATATGGCTCTTCAAATGTCTGGAGATACTGCA
PP4-1_4_M13R      ATGGATTTTATGATGAGTGTTTGCGTAAATATGGCTCTTCAAATGTCTGGAGATACTGCA
                  ************************************************************

PP4-1             CCGACATTTTTGACTACATGAGTCTTTCAGCTGTTGTGGAGAACAAGATATTCTGTGTTC
PP4-1_4_M13R      CCGACATTTTTGACTACATGAGTCTTTCAGCTGTTGTGGAGAACAAGATATTCTGTGTTC
                  ************************************************************

PP4-1             ATGGTGGTCTTTCTCCAGCTATTATGACTCTTGATCAGATTAGGACAATTGACCGGAAGC
PP4-1_4_M13R      ATGGTGGTCTTTCTCCAGCTATTATGACTCTTGATCAGATTAGGACAATTGACCGGAAGC
                  ************************************************************

PP4-1             AAGAAGTACCACATGATGGTGCCATGTGTGATCTCCTATGGTCTGATCCTGAAGATATTG
PP4-1_4_M13R      AAGAAGTACCACATGATGGTGCCATGTGTGATCTCCTATGGTCTGATCCTGAAGATATTG
                  ************************************************************

PP4-1             TTGATGGCTGGGGATTGAGCCCTCGTGGTGCCGGATTCCTTTTTGGTGGCAGTGTTGTCA
PP4-1_4_M13R      TTGATGGCTGGGGATTGAGCCCTCGTGGTGCCGGATTCCTTTTTGGTGGCAGTGTTGTCA
                  ************************************************************

PP4-1             CGTCTTTTAACCACTCAAACAACATAGACTACATAGCCCGTGCCCATCAACTAGTTATGG
PP4-1_4_M13R      CGTCTTTTAACCACTCAAACAACATAGACTACATAGCCCGTGCCCATCAACTAGTTATGG
                  ************************************************************

PP4-1             AGGGTTACAAATGGATGTTTGATAGCCAGATTGTGACAGTGTGGTCAGCCCCAAATTACT
PP4-1_4_M13R      AGGGTTACAAATGGATGTTTGATAGCCAGATTGTGACAGTGTGGTCAGCCCCAAATTACT
                  ************************************************************

PP4-1             GTTACAGATGCGGTAATGTGGCTTCAATTCTAGAGCTTGACGAGAATCTAAATAAAGAAT
PP4-1_4_M13R      GTTACAGATGCGGTAATGTGGCTTCAATTCTAGAGCTTGACGAGAATCTAAATAAAGAAT
                  ************************************************************

PP4-1             TCCGTGTGTTTGATGCAGCCCAGCAGGACTCGAGAGGGCCTCCCGCCAAAAAGCCGGCCC
```

```
PP4-1_4_M13R      TCCGTGTGTTTGATGCAGCCCAGCAGGACTCGAGAGGGCCTCCCGCCAAAAAGCCGGCCC
                  ************************************************************

PP4-1             CTGATTACTTCCTATAA-------------------------------------------
PP4-1_4_M13R      CTGATTACTTCCTATGGGGCGGCCGCGGAGCCTGCTTTTTTGTACAAAGTTGGCATTATA
                  ***************

PP4-1             ------------------------------------------------------------
PP4-1_4_M13R      AAAAAGCATTGCTCATCAATTTGTTGCAACGAACAGGTCMCTATCAGTCAAAAWAAAATC


PP4-1             --
PP4-1_4_M13R      AT
```

A-6: Multiple sequence alignment between PP4-1 sequence and PP4-1 dTOPO 13, completely aligned

```
PP4-1              ------------------------------------------------------------
PP4-1_13_M13F      TGACTGATAGTGACCTGTTCGTTGCAACAAATTGATGAGCAATGCTTTTTTATAATGCCA


PP4-1              -------------------------------------------ATGTCAGACCTAGATCGG
PP4-1_13_M13F      ACTTTGTACAAAAAAGCAGGCTCCGCGGCCGCCCCCTTCACCATGTCAGACCTAGATCGG
                                                               ******************

PP4-1              CAAATAGGGCAGCTTAAGCGATGCGAACCATTGAGCGAATCGGAGGTGAAGGCTCTTTGC
PP4-1_13_M13F      CAAATAGGGCAGCTTAAGCGATGCGAACCATTGAGCGAATCGGAGGTGAAGGCTCTTTGC
                   ************************************************************

PP4-1              CTCAAAGCCATGGAAATTCTTGTTGAAGAGAGTAATGTTCAGAGAGTTGATGCCCCTGTC
PP4-1_13_M13F      CTCAAAGCCATGGAAATTCTTGTTGAAGAGAGTAATGTTCAGAGAGTTGATGCCCCTGTC
                   ************************************************************

PP4-1              ACTTTATGTGGTGACATCCATGGGCAGTTCTATGATATGATGGAGCTTTTCAAAGTTGGG
PP4-1_13_M13F      ACTTTATGTGGTGACATCCATGGGCAGTTCTATGATATGATGGAGCTTTTCAAAGTTGGG
                   ************************************************************

PP4-1              GGTGATTGTCCTAAGACCAACTATTTGTTTATGGGAGATTTTGTTGATCGTGGATATTAT
PP4-1_13_M13F      GGTGATTGTCCTAAGACCAACTATTTGTTTATGGGAGATTTTGTTGATCGTGGATATTAT
                   ************************************************************

PP4-1              TCGGTTGAGACATTTCTACTTCTACTCGCACTCAAGGTTAGATATCCAGACCGCATAACT
PP4-1_13_M13F      TCGGTTGAGACATTTCTACTTCTACTCGCACTCAAGGTTAGATATCCAGACCGCATAACT
                   ************************************************************

PP4-1              CTCATCAGAGGAAACCATGAAAGCAGGCAAATCACACAGGTTTATGGATTTTATGATGAG
PP4-1_13_M13F      CTCATCAGAGGAAACCATGAAAGCAGGCAAATCACACAGGTTTATGGATTTTATGATGAG
                   ************************************************************

PP4-1              TGTTTGCGTAAATATGGCTCTTCAAATGTCTGGAGATACTGCACCGACATTTTTGACTAC
PP4-1_13_M13F      TGTTTGCGTAAATATGGCTCTTCAAATGTCTGGAGATACTGCACCGACATTTTTGACTAC
                   ************************************************************

PP4-1              ATGAGTCTTTCAGCTGTTGTGGAGAACAAGATATTCTGTGTTCATGGTGGTCTTTCTCCA
PP4-1_13_M13F      ATGAGTCTTTCAGCTGTTGTGGAGAACAAGATATTCTGTGTTCATGGTGGTCTTTCTCCA
                   ************************************************************

PP4-1              GCTATTATGACTCTTGATCAGATTAGGACAATTGACCGGAAGCAAGAAGTACCACATGAT
PP4-1_13_M13F      GCTATTATGACTCTTGATCAGATTAGGACAATTGACCGGAAGCAAGAAGTACCACATGAT
                   ************************************************************

PP4-1              GGTGCCATGTGTGATCTCCTATGGTCTGATCCTGAAGATATTGTTGATGGCTGGGGATTG
PP4-1_13_M13F      GGTGCCATGTGTGATCTCCTATGGTCTGATCCTGAAGATATTGTTGATGGCTGGGGATTG
                   ************************************************************

PP4-1              AGCCCTCGTGGTGCCGGATTCCTTTTTGGTGGCAGTGTTGTCACGTCTTTTAACCACTCA
PP4-1_13_M13F      AGCCCTCGTGGTGCCGGATTCCTTTTTGGTGGCAGTGTTGTCACGTCTTTTAACCACTCA
```

```
                  ************************************************************

PP4-1             AACAACATAGACTACATAGCCCGTGCCCATCAACTAGTTATGGAGGGTTACAAATGGATG
PP4-1_13_M13F     AACAACATAGACTACATAGCCCGTGCCCATCAACTAGTTATGGAGGGTTACAAATGGATG
                  ************************************************************

PP4-1             TTTGATAGCCAGATTGTGACAGTGTGGTCAGCCCCAAATTACTGTTACAGATGCGGTAAT
PP4-1_13_M13F     TTTGATAGCCAGATTGTGACAGTGTGGTCAGCCCCAAATTACTGTTACAGATGCGGTAAT
                  ************************************************************

PP4-1             GTGGCTTCAATTCTAGAGCTTGACGAGAATCTAAATAAAGAATTCCGTGTGTTTGATGCA
PP4-1_13_M13F     GTGGCTTCAATTCTAGAGCTTGACGAGAATCTAAATAAAGAATTCCGTGTGTTTGATGCA
                  ************************************************************

PP4-1             GCCCAGCAGGACTCGAGAGGGCCTCCCGCCAAAAAGCCGGCCCCTGATTACTTCCTATAA
PP4-1_13_M13F     GCCCAGCAGGACTCGAGAGGGCCTCCCGCCAAAAAGCCGGCCCCTGATTACTTCCTATAA
                  ************************************************************

PP4-1             ------------------------------------------------------------
PP4-1_13_M13F     AAGGGTGGGCGCGCCGACCCAGCTTTCTTGTACAAAGTTGGCATTATAAGAAAGCATTGC


PP4-1             ------------------------------------------------------------
PP4-1_13_M13F     TTATCAATTTGTTGCAACGAACAGGTCACTATCAGTCAAAATAAAATCATTATTTGCCAT


PP4-1             --------------
PP4-1_13_M13F     CCAGCTGAWWCCCC
```

A-7: Multiple sequence alignment between PP4-1 sequence and PP4-1 dTOPO 14, completely aligned

```
PP4-1             ------------------------------------------------------------
PP4-1_14_M13F     CCTGTTCGTTGCAACAAATTGATGAGCAATGCTTTTTTATAATGCCAACTTTGTACAAAA


PP4-1             -----------------------------ATGTCAGACCTAGATCGGCAAATAGGGCAGC
PP4-1_14_M13F     AAGCAGGCTCCGCGGCCGCCCCCTTCACCATGTCAGACCTAGATCGGCAAATAGGGCAGC
                                               *******************************

PP4-1             TTAAGCGATGCGAACCATTGAGCGAATCGGAGGTGAAGGCTCTTTGCCTCAAAGCCATGG
PP4-1_14_M13F     TTAAGCGATGCGAACCATTGAGCGAATCGGAGGTGAAGGCTCTTTGCCTCAAAGCCATGG
                  ************************************************************

PP4-1             AAATTCTTGTTGAAGAGAGTAATGTTCAGAGAGTTGATGCCCCTGTCACTTTATGTGGTG
PP4-1_14_M13F     AAATTCTTGTTGAAGAGAGTAATGTTCAGAGAGTTGATGCCCCTGTCACTTTATGTGGTG
                  ************************************************************

PP4-1             ACATCCATGGGCAGTTCTATGATATGATGGAGCTTTTCAAAGTTGGGGGTGATTGTCCTA
PP4-1_14_M13F     ACATCCATGGGCAGTTCTATGATATGATGGAGCTTTTCAAAGTTGGGGGTGATTGTCCTA
                  ************************************************************

PP4-1             AGACCAACTATTTGTTTATGGGAGATTTTGTTGATCGTGGATATTATTCGGTTGAGACAT
PP4-1_14_M13F     AGACCAACTATTTGTTTATGGGAGATTTTGTTGATCGTGGATATTATTCGGTTGAGACAT
                  ************************************************************

PP4-1             TTCTACTTCTACTCGCACTCAAGGTTAGATATCCAGACCGCATAACTCTCATCAGAGGAA
PP4-1_14_M13F     TTCTACTTCTACTCGCACTCAAGGTTAGATATCCAGACCGCATAACTCTCATCAGAGGAA
                  ************************************************************

PP4-1             ACCATGAAAGCAGGCAAATCACACAGGTTTATGGATTTTATGATGAGTGTTTGCGTAAAT
PP4-1_14_M13F     ACCATGAAAGCAGGCAAATCACACAGGTTTATGGATTTTATGATGAGTGTTTGCGTAAAT
                  ************************************************************

PP4-1             ATGGCTCTTCAAATGTCTGGAGATACTGCACCGACATTTTTGACTACATGAGTCTTTCAG
PP4-1_14_M13F     ATGGCTCTTCAAATGTCTGGAGATACTGCACCGACATTTTTGACTACATGAGTCTTTCAG
                  ************************************************************
```

```
PP4-1              CTGTTGTGGAGAACAAGATATTCTGTGTTCATGGTGGTCTTTCTCCAGCTATTATGACTC
PP4-1_14_M13F      CTGTTGTGGAGAACAAGATATTCTGTGTTCATGGTGGTCTTTCTCCAGCTATTATGACTC
                   ************************************************************

PP4-1              TTGATCAGATTAGGACAATTGACCGGAAGCAAGAAGTACCACATGATGGTGCCATGTGTG
PP4-1_14_M13F      TTGATCAGATTAGGACAATTGACCGGAAGCAAGAAGTACCACATGATGGTGCCATGTGTG
                   ************************************************************

PP4-1              ATCTCCTATGGTCTGATCCTGAAGATATTGTTGATGGCTGGGGATTGAGCCCTCGTGGTG
PP4-1_14_M13F      ATCTCCTATGGTCTGATCCTGAAGATATTGTTGATGGCTGGGGATTGAGCCCTCGTGGTG
                   ************************************************************

PP4-1              CCGGATTCCTTTTTGGTGGCAGTGTTGTCACGTCTTTTAACCACTCAAACAACATAGACT
PP4-1_14_M13F      CCGGATTCCTTTTTGGTGGCAGTGTTGTCACGTCTTTTAACCACTCAAACAACATAGACT
                   ************************************************************

PP4-1              ACATAGCCCGTGCCCATCAACTAGTTATGGAGGGTTACAAATGGATGTTTGATAGCCAGA
PP4-1_14_M13F      ACATAGCCCGTGCCCATCAACTAGTTATGGAGGGTTACAAATGGATGTTTGATAGCCAGA
                   ************************************************************

PP4-1              TTGTGACAGTGTGGTCAGCCCCAAATTACTGTTACAGATGCGGTAATGTGGCTTCAATTC
PP4-1_14_M13F      TTGTGACAGTGTGGTCAGCCCCAAATTACTGTTACAGATGCGGTAATGTGGCTTCAATTC
                   ************************************************************

PP4-1              TAGAGCTTGACGAGAATCTAAATAAAGAATTCCGTGTGTTTGATGCAGCCCAGCAGGACT
PP4-1_14_M13F      TAGAGCTTGACGAGAATCTAAATAAAGAATTCCGTGTGTTTGATGCAGCCCAGCAGGACT
                   ************************************************************

PP4-1              CGAGAGGGCCTCCCGCCAAAAAGCCGGCCCCTGATTACTTCCTATAA-------------
PP4-1_14_M13F      CGAGAGGGCCTCCCGCCAAAAAGCCGGCCCCTGATTACTTCCTATAAAAGGGTGGGCGCG
                   *********************************************

PP4-1              ------------------------------------------------------------
PP4-1_14_M13F      CCGACCCAGCTTTCTTGTACAAAGTTGGCATTATAAGAAAGCATTGCTTATCAATTTGTT


PP4-1              ------------------------------------------------------------
PP4-1_14_M13F      GCAACGAACAGGTCACTATCAGTCAAAATAAAATCATTATTTGCCATCCAGCTGAWW
```

A-8: Multiple sequence alignment between PP4-1 sequence and the two PP4-1-pMAL-C2X 1 samples with M13 F primer, sequencing failed to sequence the last nucleotides

```
PP4-1                  ------------------------------------------------------------
PP4-1_pMAL1.2M13F      GGGCCATGTTAAGGGCGGTTTTTTCCTGTTTGGTCACTGATGCCTCCGKGTAAGGGGGAT
PP4-1_pMAL1.1M13F      GGGCCATGTTAAGGGCGGTTTTTTCCTGTTTGGTCACTGATGCCTCCGKGTAAGGGGGAT


PP4-1                  ------------------------------------------------------------
PP4-1_pMAL1.2M13F      TTCTGTTCATGGGGGTAATGATACCGATGAAACGAGAGAGGATGCTCACGATACGGGTTA
PP4-1_pMAL1.1M13F      TTCTGTTCATGGGGGTAATGATACCGATGAAACGAGAGAGGATGCTCACGATACGGGTTA


PP4-1                  ------------------------------------------------------------
PP4-1_pMAL1.2M13F      CTGATGATGAACATGCCCGGTTACTGGAACGTTGTGAGGGTAAACAACTGGCGGTATGGA
PP4-1_pMAL1.1M13F      CTGATGATGAACATGCCCGGTTACTGGAACGTTGTGAGGGTAAACAACTGGCGGTATGGA


PP4-1                  ------------------------------------------------------------
PP4-1_pMAL1.2M13F      TGCGGCGGGACCAGAGAAAAATCACTCAGGGTCAATGCCAGCGCTTCGTTAATACAGATG
PP4-1_pMAL1.1M13F      TGCGGCGGGACCAGAGAAAAATCACTCAGGGTCAATGCCAGCGCTTCGTTAATACAGATG


PP4-1                  ----------------------------------------------ATGTCAGACCTAGAT
PP4-1_pMAL1.2M13F      TAGGTGTTCCACAGGGTAGCCAGCAGCATCCTGCGATGCAGATCCATGTCAGACCTAGAT
PP4-1_pMAL1.1M13F      TAGGTGTTCCACAGGGTAGCCAGCAGCATCCTGCGATGCAGATCCATGTCAGACCTAGAT
                                                              **************
```

```
PP4-1                   CGGCAAATAGGGCAGCTTAAGCGATGCGAACCATTGAGCGAATCGGAGGTGAAGGCTCTT
PP4-1_pMAL1.2M13F       CGGCAAATAGGGCAGCTTAAGCGATGCGAACCATTGAGCGAATCGGAGGTGAAGGCTCTT
PP4-1_pMAL1.1M13F       CGGCAAATAGGGCAGCTTAAGCGATGCGAACCATTGAGCGAATCGGAGGTGAAGGCTCTT
                        ************************************************************

PP4-1                   TGCCTCAAAGCCATGGAAATTCTTGTTGAAGAGAGTAATGTTCAGAGAGTTGATGCCCCT
PP4-1_pMAL1.2M13F       TGCCTCAAAGCCATGGAAATTCTTGTTGAAGAGAGTAATGTTCAGAGAGTTGATGCCCCT
PP4-1_pMAL1.1M13F       TGCCTCAAAGCCATGGAAATTCTTGTTGAAGAGAGTAATGTTCAGAGAGTTGATGCCCCT
                        ************************************************************

PP4-1                   GTCACTTTATGTGGTGACATCCATGGGCAGTTCTATGATATGATGGAGCTTTTCAAAGTT
PP4-1_pMAL1.2M13F       GTCACTTTATGTGGTGACATCCATGGGCAGTTCTATGATATGATGGAGCTTTTCAAAGTT
PP4-1_pMAL1.1M13F       GTCACTTTATGTGGTGACATCCATGGGCAGTTCTATGATATGATGGAGCTTTTCAAAGTT
                        ************************************************************

PP4-1                   GGGGGTGATTGTCCTAAGACCAACTATTTGTTTATGGGAGATTTTGTTGATCGTGGATAT
PP4-1_pMAL1.2M13F       GGGGGTGATTGTCCTAAGACCAACTATTTGTTTATGGGAGATTTTGTTGATCGTGGATAT
PP4-1_pMAL1.1M13F       GGGGGTGATTGTCCTAAGACCAACTATTTGTTTATGGGAGATTTTGTTGATCGTGGATAT
                        ************************************************************

PP4-1                   TATTCGGTTGAGACATTTCTACTTCTACTCGCACTCAAGGTTAGATATCCAGACCGCATA
PP4-1_pMAL1.2M13F       TATTCGGTTGAGACATTTCTACTTCTACTCGCACTCAAGGTTAGATATCCAGACCGCATA
PP4-1_pMAL1.1M13F       TATTCGGTTGAGACATTTCTACTTCTACTCGCACTCAAGGTTAGATATCCAGACCGCATA
                        ************************************************************

PP4-1                   ACTCTCATCAGAGGAAACCATGAAAGCAGGCAAATCACACAGGTTTATGGATTTTATGAT
PP4-1_pMAL1.2M13F       ACTCTCATCAGAGGAAACCATGAAAGCAGGCAAATCACACAGGTTTATGGATTTTATGAT
PP4-1_pMAL1.1M13F       ACTCTCATCAGAGGAAACCATGAAAGCAGGCAAATCACACAGGTTTATGGATTTTATGAT
                        ************************************************************

PP4-1                   GAGTGTTTGCGTAAATATGGCTCTTCAAATGTCTGGAGATACTGCACCGACATTTTTGAC
PP4-1_pMAL1.2M13F       GAGTGTTTGCGTAAATATGGCTCTTCAAATGTCTGGAGATACTGCACCGACATTTTTGAC
PP4-1_pMAL1.1M13F       GAGTGTTTGCGTAAATATGGCTCTTCAAATGTCTGGAGATACTGCACCGACATTTTTGAC
                        ************************************************************

PP4-1                   TACATGAGTCTTTCAGCTGTTGTGGAGAACAAGATATTCTGTGTTCATGGTGGTCTTTCT
PP4-1_pMAL1.2M13F       TACATGAGTCTTTCAGCTGTTGTGGAGAACAAGATATTCTGTGTTCATGGTGGTCTTTCT
PP4-1_pMAL1.1M13F       TACATGAGTCTTTCAGCTGTTGTGGAGAACAAGATATTCTGTGTTCATGGTGGTCTTTCT
                        ************************************************************

PP4-1                   CCAGCTATTATGACTCTTGATCAGATTAGGACAATTGACCGGAAGCAAGAAGTACCACAT
PP4-1_pMAL1.2M13F       CCAGCTATTATGACTCTTGATCAGATTAGGACAATTGACCGGAAGCAAGAAGTACCACAT
PP4-1_pMAL1.1M13F       CCAGCTATTATGACTCTTGATCAGATTAGGACAATTGACCGGAAGCAAGAAGTACCACAT
                        ************************************************************

PP4-1                   GATGGTGCCATGTGTGATCTCCTATGGTCTGATCCTGAAGATATTGTTGATGGCTGGGGA
PP4-1_pMAL1.2M13F       GATGGTGCCATGTGTGATCTCCTATGGTCTGATCCTGAAGATATTGTTGATGGCTGGGGA
PP4-1_pMAL1.1M13F       GATGGTGCCATGTGTGATCTCCTATGGTCTGATCCTGAAGATATTGTTGATGGCTGGGGA
                        ************************************************************

PP4-1                   TTGAGCCCTCGTGGTGCCGGATTCCTTTTTGGTGGCAGTGTTGTCACGTCTTTTAACCAC
PP4-1_pMAL1.2M13F       TTGAGCCCTCGTGGTGCCGGATTCCTTTTTGGTGGCAGTGTTGTCACGTCTTTTAACCAC
PP4-1_pMAL1.1M13F       TTGAGCCCTCGTGGTGCCGGATTCCTTTTTGGTGGCAGTGTTGTCACGTCTTTTAACCAC
                        ************************************************************

PP4-1                   TCAAACAACATAGACTACATAGCCCGTGCCCATCAACTAGTTATGGAGGGTTACAAATGG
PP4-1_pMAL1.2M13F       TCAAACAACATAGACTACATAGCCCGTGCCCATCAACTAGTTATGGAGGGTTACAAATGG
PP4-1_pMAL1.1M13F       TCAAACAACATAGACTACATAGCCCGTGCCCATCAACTAGTTATGGAGGGTTACAAATGG
                        ************************************************************

PP4-1                   ATGTTTGATAGCCAGATTGTGACAGTGTGGTCAGCCCCAAATTACTGTTACAGATGCGGT
PP4-1_pMAL1.2M13F       ATGTTTGATAGCCAGATTGTGACAGTGTGGTCAGCCCCAAATTACTGTTACAGATGCGGT
PP4-1_pMAL1.1M13F       ATGTTTGATAGCCAGATTGTGACAGTGTGGTCAGCCCCAAATTACTGTTACAGATGCGGT
                        ************************************************************

PP4-1                   AATGTGGCTTCAATTCTAGAGCTTGACGAGAATCTAAATAAAGAATTCCGTGTGTTTGAT
PP4-1_pMAL1.2M13F       AATGTGGCTTCAATTCTAGAGCTTGACGAGAATCTAAATAAAGAATTCCGTGTGTTTGAT
```

```
PP4-1_pMAL1.1M13F      AATGTGGCTTCAATTCTAGAGCTTGACGAGAATCTAAATAAAGAATTCCGTGTGTTTGAT
                       ************************************************************

PP4-1                  GCAGCCCAGCAGGACTCGAGAGGGCCTCCCGCCAAAAAGCCGGCCCCTGATTACTTCCTA
PP4-1_pMAL1.2M13F      GCAGCCCAGCAGGACTCGAGAGGGCCTCCCGNCN--------------------------
PP4-1_pMAL1.1M13F      GCAGCCCAGCAGGACTCGAGAGGGCCTCCCGCCNNAAAAAG------------------
                       ****************************** *

PP4-1                  TAA
PP4-1_pMAL1.2M13F      ---
PP4-1_pMAL1.1M13F      ---
```

A-9: Multiple sequence alignment between PP4-1 sequence and PP4-1-pMAL-C2X 1 samples with PP4-1_MBP F primer, beginning of gene not sequenced

```
PP4-1                  ATGTCAGACCTAGATCGGCAAATAGGGCAGCTTAAGCGATGCGAACCATTGAGCGAATCG
PP4-1_pMAL1_MBPF       ------------------------------------------------------------TCG
                                                                                   ***

PP4-1                  GAGGTGAAGGCTCTTTGCCTCAAAGCCATGGAAATTCTTGTTGAAGAGAGTAATGTTCAG
PP4-1_pMAL1_MBPF       GAGGTGAAGGCTCTTTGCCTCAAAGCCATGGAAATTCTTGTTGAAGAGAGTAATGTTCAG
                       ************************************************************

PP4-1                  AGAGTTGATGCCCCTGTCACTTTATGTGGTGACATCCATGGGCAGTTCTATGATATGATG
PP4-1_pMAL1_MBPF       AGAGTTGATGCCCCTGTCACTTTATGTGGTGACATCCATGGGCAGTTCTATGATATGATG
                       ************************************************************

PP4-1                  GAGCTTTTCAAAGTTGGGGGTGATTGTCCTAAGACCAACTATTTGTTTATGGGAGATTTT
PP4-1_pMAL1_MBPF       GAGCTTTTCAAAGTTGGGGGTGATTGTCCTAAGACCAACTATTTGTTTATGGGAGATTTT
                       ************************************************************

PP4-1                  GTTGATCGTGGATATTATTCGGTTGAGACATTTCTACTTCTACTCGCACTCAAGGTTAGA
PP4-1_pMAL1_MBPF       GTTGATCGTGGATATTATTCGGTTGAGACATTTCTACTTCTACTCGCACTCAAGGTTAGA
                       ************************************************************

PP4-1                  TATCCAGACCGCATAACTCTCATCAGAGGAAACCATGAAAGCAGGCAAATCACACAGGTT
PP4-1_pMAL1_MBPF       TATCCAGACCGCATAACTCTCATCAGAGGAAACCATGAAAGCAGGCAAATCACACAGGTT
                       ************************************************************

PP4-1                  TATGGATTTTATGATGAGTGTTTGCGTAAATATGGCTCTTCAAATGTCTGGAGATACTGC
PP4-1_pMAL1_MBPF       TATGGATTTTATGATGAGTGTTTGCGTAAATATGGCTCTTCAAATGTCTGGAGATACTGC
                       ************************************************************

PP4-1                  ACCGACATTTTTGACTACATGAGTCTTTCAGCTGTTGTGGAGAACAAGATATTCTGTGTT
PP4-1_pMAL1_MBPF       ACCGACATTTTTGACTACATGAGTCTTTCAGCTGTTGTGGAGAACAAGATATTCTGTGTT
                       ************************************************************

PP4-1                  CATGGTGGTCTTTCTCCAGCTATTATGACTCTTGATCAGATTAGGACAATTGACCGGAAG
PP4-1_pMAL1_MBPF       CATGGTGGTCTTTCTCCAGCTATTATGACTCTTGATCAGATTAGGACAATTGACCGGAAG
                       ************************************************************

PP4-1                  CAAGAAGTACCACATGATGGTGCCATGTGTGATCTCCTATGGTCTGATCCTGAAGATATT
PP4-1_pMAL1_MBPF       CAAGAAGTACCACATGATGGTGCCATGTGTGATCTCCTATGGTCTGATCCTGAAGATATT
                       ************************************************************

PP4-1                  GTTGATGGCTGGGGATTGAGCCCTCGTGGTGCCGGATTCCTTTTTGGTGGCAGTGTTGTC
PP4-1_pMAL1_MBPF       GTTGATGGCTGGGGATTGAGCCCTCGTGGTGCCGGATTCCTTTTTGGTGGCAGTGTTGTC
                       ************************************************************

PP4-1                  ACGTCTTTTAACCACTCAAACAACATAGACTACATAGCCCGTGCCCATCAACTAGTTATG
PP4-1_pMAL1_MBPF       ACGTCTTTTAACCACTCAAACAACATAGACTACATAGCCCGTGCCCATCAACTAGTTATG
                       ************************************************************

PP4-1                  GAGGGTTACAAATGGATGTTTGATAGCCAGATTGTGACAGTGTGGTCAGCCCCAAATTAC
PP4-1_pMAL1_MBPF       GAGGGTTACAAATGGATGTTTGATAGCCAGATTGTGACAGTGTGGTCAGCCCCAAATTAC
                       ************************************************************
```

```
PP4-1                 TGTTACAGATGCGGTAATGTGGCTTCAATTCTAGAGCTTGACGAGAATCTAAATAAAGAA
PP4-1_pMAL1_MBPF      TGTTACAGATGCGGTAATGTGGCTTCAATTCTAGAGCTTGACGAGAATCTAAATAAAGAA
                      ************************************************************

PP4-1                 TTCCGTGTGTTTGATGCAGCCCAGCAGGACTCGAGAGGGCCTCCCGCCAAAAAGCCGGCC
PP4-1_pMAL1_MBPF      TTCCGTGTGTTTGATGCAGCCCAGCAGGACTCGAGAGGGCCTCCCGCCAAAAAGCCGGCC
                      ************************************************************

PP4-1                 CCTGATTACTTCCTATAA------------------------------------------
PP4-1_pMAL1_MBPF      CCTGATTACTTCCTATAAAAGCTTGGCACTGGCCGTCGTTTTACAACGTCGTGACTGGGA
                      ******************

PP4-1                 ------------------------------------------------------------
PP4-1_pMAL1_MBPF      AAACCCTGGCGTTACCCAACTTAATCGCCTTGCAGCACATCCCCCTTTCGCCAGCTGGCG


PP4-1                 ------------------------------------------------------------
PP4-1_pMAL1_MBPF      TAATAGCGAAGAGGCCCGCACCGATCGCCCTTCCCAACAGTTGCGCAGCCTGAATGGCGA


PP4-1                 ------------------------------------------------------------
PP4-1_pMAL1_MBPF      ATGGCAGCTTGGCTGTTTTGGCGGATGAGATAAGATTTTCAGCCTGATACAGATTAAATC


PP4-1                 ----------------------
PP4-1_pMAL1_MBPF      AGAAACGCAGAAAGCGGTCTGA
```

A-10: Multiple sequence alignment between PP4-1 sequence and PP4-1-pMAL-C2X 11 sample with PP4-1_MBP F primer, beginning of gene not sequenced

```
PP4-1                  ATGTCAGACCTAGATCGGCAAATAGGGCAGCTTAAGCGATGCGAACCATTGAGCGAATCG
PP4-1_pMAL11_MBPF      ------------------------------------------------------------TCG
                                                                                   ***

PP4-1                  GAGGTGAAGGCTCTTTGCCTCAAAGCCATGGAAATTCTTGTTGAAGAGAGTAATGTTCAG
PP4-1_pMAL11_MBPF      GAGGTGAAGGCTCTTTGCCTCAAAGCCATGGAAATTCTTGTTGAAGAGAGTAATGTTCAG
                       ************************************************************

PP4-1                  AGAGTTGATGCCCCTGTCACTTTATGTGGTGACATCCATGGGCAGTTCTATGATATGATG
PP4-1_pMAL11_MBPF      AGAGTTGATGCCCCTGTCACTTTATGTGGTGACATCCATGGGCAGTTCTATGATATGATG
                       ************************************************************

PP4-1                  GAGCTTTTCAAAGTTGGGGGTGATTGTCCTAAGACCAACTATTTGTTTATGGGAGATTTT
PP4-1_pMAL11_MBPF      GAGCTTTTCAAAGTTGGGGGTGATTGTCCTAAGACCAACTATTTGTTTATGGGAGATTTT
                       ************************************************************

PP4-1                  GTTGATCGTGGATATTATTCGGTTGAGACATTTCTACTTCTACTCGCACTCAAGGTTAGA
PP4-1_pMAL11_MBPF      GTTGATCGTGGATATTATTCGGTTGAGACATTTCTACTTCTACTCGCACTCAAGGTTAGA
                       ************************************************************

PP4-1                  TATCCAGACCGCATAACTCTCATCAGAGGAAACCATGAAAGCAGGCAAATCACACAGGTT
PP4-1_pMAL11_MBPF      TATCCAGACCGCATAACTCTCATCAGAGGAAACCATGAAAGCAGGCAAATCACACAGGTT
                       ************************************************************

PP4-1                  TATGGATTTTATGATGAGTGTTTGCGTAAATATGGCTCTTCAAATGTCTGGAGATACTGC
PP4-1_pMAL11_MBPF      TATGGATTTTATGATGAGTGTTTGCGTAAATATGGCTCTTCAAATGTCTGGAGATACTGC
                       ************************************************************

PP4-1                  ACCGACATTTTTGACTACATGAGTCTTTCAGCTGTTGTGGAGAACAAGATATTCTGTGTT
PP4-1_pMAL11_MBPF      ACCGACATTTTTGACTACATGAGTCTTTCAGCTGTTGTGGAGAACAAGATATTCTGTGTT
                       ************************************************************

PP4-1                  CATGGTGGTCTTTCTCCAGCTATTATGACTCTTGATCAGATTAGGACAATTGACCGGAAG
PP4-1_pMAL11_MBPF      CATGGTGGTCTTTCTCCAGCTATTATGACTCTTGATCAGATTAGGACAATTGACCGGAAG
                       ************************************************************

PP4-1                  CAAGAAGTACCACATGATGGTGCCATGTGTGATCTCCTATGGTCTGATCCTGAAGATATT
```

```
PP4-1_pMAL11_MBPF      CAAGAAGTACCACATGATGGTGCCATGTGTGATCTCCTATGGTCTGATCCTGAAGATATT
                       ************************************************************

PP4-1                  GTTGATGGCTGGGGATTGAGCCCTCGTGGTGCCGGATTCCTTTTTGGTGGCAGTGTTGTC
PP4-1_pMAL11_MBPF      GTTGATGGCTGGGGATTGAGCCCTCGTGGTGCCGGATTCCTTTTTGGTGGCAGTGTTGTC
                       ************************************************************

PP4-1                  ACGTCTTTTAACCACTCAAACAACATAGACTACATAGCCCGTGCCCATCAACTAGTTATG
PP4-1_pMAL11_MBPF      ACGTCTTTTAACCACTCAAACAACATAGACTACATAGCCCGTGCCCATCAACTAGTTATG
                       ************************************************************

PP4-1                  GAGGGTTACAAATGGATGTTTGATAGCCAGATTGTGACAGTGTGGTCAGCCCCAAATTAC
PP4-1_pMAL11_MBPF      GAGGGTTACAAATGGATGTTTGATAGCCAGATTGTGACAGTGTGGTCAGCCCCAAATTAC
                       ************************************************************

PP4-1                  TGTTACAGATGCGGTAATGTGGCTTCAATTCTAGAGCTTGACGAGAATCTAAATAAAGAA
PP4-1_pMAL11_MBPF      TGTTACAGATGCGGTAATGTGGCTTCAATTCTAGAGCTTGACGAGAATCTAAATAAAGAA
                       ************************************************************

PP4-1                  TTCCGTGTGTTTGATGCAGCCCAGCAGGACTCGAGAGGGCCTCCCGCCAAAAAGCCGGCC
PP4-1_pMAL11_MBPF      TTCCGTGTGTTTGATGCAGCCCAGCAGGACTCGAGAGGGCCTCCCGCCAAAAAGCCGGCC
                       ************************************************************

PP4-1                  CCTGATTACTTCCTATAA------------------------------------------
PP4-1_pMAL11_MBPF      CCTGATTACTTCCTATAAAAGCTTGGCACTGGCCGTCGTTTTACAACGTCGTGACTGGGA
                       ******************

PP4-1                  ------------------------------------------------------------
PP4-1_pMAL11_MBPF      AAACCCTGGCGTTACCCAACTTAATCGCCTTGCAGCACATCCCCCTTTCGCCAGCTGGCG

PP4-1                  ------------------------------------------------------------
PP4-1_pMAL11_MBPF      TAATAGCGAAGAGGCCCGCACCGATCGCCCTTCCCAACAGTTGCGCAGCCTGAATGGCGA

PP4-1                  ------------------------------------------------------------
PP4-1_pMAL11_MBPF      ATGGCAGCTTGGCTGTTTTGGCGGATGAGATAAGATTTTCAGCCTGATACAGATTAAATC

PP4-1                  ------------------------------------------------------------
PP4-1_pMAL11_MBPF      AGAACGCAGAAGCGGTCTGATAAAACAGAATTTGCCTGGCGGCAGTAGCGCGGTGGTCCC

PP4-1                  ----------------
PP4-1_pMAL11_MBPF      MCCTGACCCCATGCCGA
```

A-11: Multiple sequence alignment between PP4-1 sequence and reverse complimentary PP4-1-pMAL-C2X 11 with M13 F, end of gene not sequenced

```
PP4-1                  ------------------------------------------------------------
PP4-1_pMAL11_M13F      CACTGATGCCTCCGKGTAAGGGGGATTTCTGTTCATGGGGGTAATGATACCGATGAAACG

PP4-1                  ------------------------------------------------------------
PP4-1_pMAL11_M13F      AGAGAGGATGCTCACGATACGGGTTACTGATGATGAACATGCCCGGTTACTGGAACGTTG

PP4-1                  ------------------------------------------------------------
PP4-1_pMAL11_M13F      TGAGGGTAAACAACTGGCGGTATGGATGCGGCGGGACCAGAGAAAAATCACTCAGGGTCA

PP4-1                  ------------------------------------------------------------
PP4-1_pMAL11_M13F      ATGCCAGCGCTTCGTTAATACAGATGTAGGTGTTCCACAGGGTAGCCAGCAGCATCCTGC

PP4-1                  -----------ATGTCAGACCTAGATCGGCAAATAGGGCAGCTTAAGCGATGCGAACCAT
PP4-1_pMAL11_M13F      GATGCAGATCCATGTCAGACCTAGATCGGCAAATAGGGCAGCTTAAGCGATGCGAACCAT
```

```
                                    **********************************************

PP4-1                   TGAGCGAATCGGAGGTGAAGGCTCTTTGCCTCAAAGCCATGGAAATTCTTGTTGAAGAGA
PP4-1_pMAL11_M13F       TGAGCGAATCGGAGGTGAAGGCTCTTTGCCTCAAAGCCATGGAAATTCTTGTTGAAGAGA
                        ************************************************************

PP4-1                   GTAATGTTCAGAGAGTTGATGCCCCTGTCACTTTATGTGGTGACATCCATGGGCAGTTCT
PP4-1_pMAL11_M13F       GTAATGTTCAGAGAGTTGATGCCCCTGTCACTTTATGTGGTGACATCCATGGGCAGTTCT
                        ************************************************************

PP4-1                   ATGATATGATGGAGCTTTTCAAAGTTGGGGGTGATTGTCCTAAGACCAACTATTTGTTTA
PP4-1_pMAL11_M13F       ATGATATGATGGAGCTTTTCAAAGTTGGGGGTGATTGTCCTAAGACCAACTATTTGTTTA
                        ************************************************************

PP4-1                   TGGGAGATTTTGTTGATCGTGGATATTATTCGGTTGAGACATTTCTACTTCTACTCGCAC
PP4-1_pMAL11_M13F       TGGGAGATTTTGTTGATCGTGGATATTATTCGGTTGAGACATTTCTACTTCTACTCGCAC
                        ************************************************************

PP4-1                   TCAAGGTTAGATATCCAGACCGCATAACTCTCATCAGAGGAAACCATGAAAGCAGGCAAA
PP4-1_pMAL11_M13F       TCAAGGTTAGATATCCAGACCGCATAACTCTCATCAGAGGAAACCATGAAAGCAGGCAAA
                        ************************************************************

PP4-1                   TCACACAGGTTTATGGATTTTATGATGAGTGTTTGCGTAAATATGGCTCTTCAAATGTCT
PP4-1_pMAL11_M13F       TCACACAGGTTTATGGATTTTATGATGAGTGTTTGCGTAAATATGGCTCTTCAAATGTCT
                        ************************************************************

PP4-1                   GGAGATACTGCACCGACATTTTTGACTACATGAGTCTTTCAGCTGTTGTGGAGAACAAGA
PP4-1_pMAL11_M13F       GGAGATACTGCACCGACATTTTTGACTACATGAGTCTTTCAGCTGTTGTGGAGAACAAGA
                        ************************************************************

PP4-1                   TATTCTGTGTTCATGGTGGTCTTTCTCCAGCTATTATGACTCTTGATCAGATTAGGACAA
PP4-1_pMAL11_M13F       TATTCTGTGTTCATGGTGGTCTTTCTCCAGCTATTATGACTCTTGATCAGATTAGGACAA
                        ************************************************************

PP4-1                   TTGACCGGAAGCAAGAAGTACCACATGATGGTGCCATGTGTGATCTCCTATGGTCTGATC
PP4-1_pMAL11_M13F       TTGACCGGAAGCAAGAAGTACCACATGATGGTGCCATGTGTGATCTCCTATGGTCTGATC
                        ************************************************************

PP4-1                   CTGAAGATATTGTTGATGGCTGGGGATTGAGCCCTCGTGGTGCCGGATTCCTTTTTGGTG
PP4-1_pMAL11_M13F       CTGAAGATATTGTTGATGGCTGGGGATTGAGCCCTCGTGGTGCCGGATTCCTTTTTGGTG
                        ************************************************************

PP4-1                   GCAGTGTTGTCACGTCTTTTAACCACTCAAACAACATAGACTACATAGCCCGTGCCCATC
PP4-1_pMAL11_M13F       GCAGTGTTGTCACGTCTTTTAACCACTCAAACAACATAGACTACATAGCCCGTGCCCATC
                        ************************************************************

PP4-1                   AACTAGTTATGGAGGGTTACAAATGGATGTTTGATAGCCAGATTGTGACAGTGTGGTCAG
PP4-1_pMAL11_M13F       AACTAGTTATGGAGGGTTACAAATGGATGTTTGATAGCCAGATTGTGACAGTGTGGTCAG
                        ************************************************************

PP4-1                   CCCCAAATTACTGTTACAGATGCGGTAATGTGGCTTCAATTCTAGAGCTTGACGAGAATC
PP4-1_pMAL11_M13F       CCCCAAATTACTGTTACAGATGCGGTAATGTGGCTTCAATTCTAGAGCTTGACGAGAATC
                        ************************************************************

PP4-1                   TAAATAAAGAATTCCGTGTGTTTGATGCAGCCCAGCAGGACTCGAGAGGGCCTCCCGCCA
PP4-1_pMAL11_M13F       TAAATAAAGAATTCCGTGTGTTTGATGCAGCCCAGCAGGACTCGAGAGGGCCTCCCGCCA
                        ************************************************************

PP4-1                   AAAAGCCGGCCCCTGATTACTTCCTATAA
PP4-1_pMAL11_M13F       AAAAG------------------------
                        *****
```

A-12: Multiple sequence alignment between PP4-1 sequence and PP4-1-pMAL-C2X 14 sample with PP4-1_MBP F primer, beginning of gene not sequenced

```
PP4-1                   ATGTCAGACCTAGATCGGCAAATAGGGCAGCTTAAGCGATGCGAACCATTGAGCGAATCG
PP4-1_pMAL14_MBPF       -------------------------------------------CATTGAGCGAATCG
                                                                   **************
```

```
PP4-1                   GAGGTGAAGGCTCTTTGCCTCAAAGCCATGGAAATTCTTGTTGAAGAGAGTAATGTTCAG
PP4-1_pMAL14_MBPF       GAGGTGAAGGCTCTTTGCCTCAAAGCCATGGAAATTCTTGTTGAAGAGAGTAATGTTCAG
                        ************************************************************

PP4-1                   AGAGTTGATGCCCCTGTCACTTTATGTGGTGACATCCATGGGCAGTTCTATGATATGATG
PP4-1_pMAL14_MBPF       AGAGTTGATGCCCCTGTCACTTTATGTGGTGACATCCATGGGCAGTTCTATGATATGATG
                        ************************************************************

PP4-1                   GAGCTTTTCAAAGTTGGGGGTGATTGTCCTAAGACCAACTATTTGTTTATGGGAGATTTT
PP4-1_pMAL14_MBPF       GAGCTTTTCAAAGTTGGGGGTGATTGTCCTAAGACCAACTATTTGTTTATGGGAGATTTT
                        ************************************************************

PP4-1                   GTTGATCGTGGATATTATTCGGTTGAGACATTTCTACTTCTACTCGCACTCAAGGTTAGA
PP4-1_pMAL14_MBPF       GTTGATCGTGGATATTATTCGGTTGAGACATTTCTACTTCTACTCGCACTCAAGGTTAGA
                        ************************************************************

PP4-1                   TATCCAGACCGCATAACTCTCATCAGAGGAAACCATGAAAGCAGGCAAATCACACAGGTT
PP4-1_pMAL14_MBPF       TATCCAGACCGCATAACTCTCATCAGAGGAAACCATGAAAGCAGGCAAATCACACAGGTT
                        ************************************************************

PP4-1                   TATGGATTTTATGATGAGTGTTTGCGTAAATATGGCTCTTCAAATGTCTGGAGATACTGC
PP4-1_pMAL14_MBPF       TATGGATTTTATGATGAGTGTTTGCGTAAATATGGCTCTTCAAATGTCTGGAGATACTGC
                        ************************************************************

PP4-1                   ACCGACATTTTTGACTACATGAGTCTTTCAGCTGTTGTGGAGAACAAGATATTCTGTGTT
PP4-1_pMAL14_MBPF       ACCGACATTTTTGACTACATGAGTCTTTCAGCTGTTGTGGAGAACAAGATATTCTGTGTT
                        ************************************************************

PP4-1                   CATGGTGGTCTTTCTCCAGCTATTATGACTCTTGATCAGATTAGGACAATTGACCGGAAG
PP4-1_pMAL14_MBPF       CATGGTGGTCTTTCTCCAGCTATTATGACTCTTGATCAGATTAGGACAATTGACCGGAAG
                        ************************************************************

PP4-1                   CAAGAAGTACCACATGATGGTGCCATGTGTGATCTCCTATGGTCTGATCCTGAAGATATT
PP4-1_pMAL14_MBPF       CAAGAAGTACCACATGATGGTGCCATGTGTGATCTCCTATGGTCTGATCCTGAAGATATT
                        ************************************************************

PP4-1                   GTTGATGGCTGGGGATTGAGCCCTCGTGGTGCCGGATTCCTTTTTGGTGGCAGTGTTGTC
PP4-1_pMAL14_MBPF       GTTGATGGCTGGGGATTGAGCCCTCGTGGTGCCGGATTCCTTTTTGGTGGCAGTGTTGTC
                        ************************************************************

PP4-1                   ACGTCTTTTAACCACTCAAACAACATAGACTACATAGCCCGTGCCCATCAACTAGTTATG
PP4-1_pMAL14_MBPF       ACGTCTTTTAACCACTCAAACAACATAGACTACATAGCCCGTGCCCATCAACTAGTTATG
                        ************************************************************

PP4-1                   GAGGGTTACAAATGGATGTTTGATAGCCAGATTGTGACAGTGTGGTCAGCCCCAAATTAC
PP4-1_pMAL14_MBPF       GAGGGTTACAAATGGATGTTTGATAGCCAGATTGTGACAGTGTGGTCAGCCCCAAATTAC
                        ************************************************************

PP4-1                   TGTTACAGATGCGGTAATGTGGCTTCAATTCTAGAGCTTGACGAGAATCTAAATAAAGAA
PP4-1_pMAL14_MBPF       TGTTACAGATGCGGTAATGTGGCTTCAATTCTAGAGCTTGACGAGAATCTAAATAAAGAA
                        ************************************************************

PP4-1                   TTCCGTGTGTTTGATGCAGCCCAGCAGGACTCGAGAGGGCCTCCCGCCAAAAAGCCGGCC
PP4-1_pMAL14_MBPF       TTCCGTGTGTTTGATGCAGCCCAGCAGGACTCGAGAGGGCCTCCCGCCAAAAAGCCGGCC
                        ************************************************************

PP4-1                   CCTGATTACTTCCTATAA------------------------------------------
PP4-1_pMAL14_MBPF       CCTGATTACTTCCTATAAAAAGCTTGGCACTGGCCGTCGTTTTACAACGTCGTGACTGGGA
                        ******************

PP4-1                   ------------------------------------------------------------
PP4-1_pMAL14_MBPF       AAACCCTGGCGTTACCCAACTTAATCGCCTTGCAGCACATCCCCCTTTCGCCAGCTGGCG


PP4-1                   ------------------------------------------------------------
PP4-1_pMAL14_MBPF       TAATAGCGAAGAGGCCCGCACCGATCGCCCTTCCCAACAGTTGCGCAGCCTGAATGGCGA
```

```
PP4-1                   ------------------------------------------------------------
PP4-1_pMAL14_MBPF       ATGGCAGCTTGGCTGTTTTGGCGGATGAAATAAGATTTTCAGCCTGATACAGATTAAATC

PP4-1                   --------------------------------------
PP4-1_pMAL14_MBPF       AGAACGCAAAAGCGGTCTGAWAAAACARAATTTGCCTG
```

A-13: Multiple sequence alignment between PP4-1 sequence and reverse complimentary PP4-1-pMAL-C2X 14 with M13 F, end of gene not sequenced

```
PP4-1                   ------------------------------------------------------------
PP4-1_pMAL14_M13F       CCCGATCAACTGGGTKCCCAGCGTGGTGGTTTCGATGGTAGAACGAAGCGGCGTCGAAGC

PP4-1                   ------------------------------------------------------------
PP4-1_pMAL14_M13F       CTGTAAAGCGGCGGTGCACAATCTTCTCGCGCAACGCGTCAGTGGGCTGATCATTAACTA

PP4-1                   ------------------------------------------------------------
PP4-1_pMAL14_M13F       TCCGCTGGATGACCAGGATGCCATTGCTGTGGAAGCTGCCTGCACTAATGTTCCGGCGTT

PP4-1                   ------------------------------------------------------------
PP4-1_pMAL14_M13F       ATTTCTTGATGTCTCTGACCAGACACCCATCAACAGTATTATTTTCTCCCATGAAGACGG

PP4-1                   ----------------------ATGTCAGACCTAGATCGGCAAATAGGGCAGCTTAAGCG
PP4-1_pMAL14_M13F       TACGCGACTGGGCGCAGGATCCATGTCAGACCTAGATCGGCAAATAGGGCAGCTTAAGCG
                                              ************************************

PP4-1                   ATGCGAACCATTGAGCGAATCGGAGGTGAAGGCTCTTTGCCTCAAAGCCATGGAAATTCT
PP4-1_pMAL14_M13F       ATGCGAACCATTGAGCGAATCGGAGGTGAAGGCTCTTTGCCTCAAAGCCATGGAAATTCT
                        ************************************************************

PP4-1                   TGTTGAAGAGAGTAATGTTCAGAGAGTTGATGCCCCTGTCACTTTATGTGGTGACATCCA
PP4-1_pMAL14_M13F       TGTTGAAGAGAGTAATGTTCAGAGAGTTGATGCCCCTGTCACTTTATGTGGTGACATCCA
                        ************************************************************

PP4-1                   TGGGCAGTTCTATGATATGATGGAGCTTTTCAAAGTTGGGGGTGATTGTCCTAAGACCAA
PP4-1_pMAL14_M13F       TGGGCAGTTCTATGATATGATGGAGCTTTTCAAAGTTGGGGGTGATTGTCCTAAGACCAA
                        ************************************************************

PP4-1                   CTATTTGTTTATGGGAGATTTTGTTGATCGTGGATATTATTCGGTTGAGACATTTCTACT
PP4-1_pMAL14_M13F       CTATTTGTTTATGGGAGATTTTGTTGATCGTGGATATTATTCGGTTGAGACATTTCTACT
                        ************************************************************

PP4-1                   TCTACTCGCACTCAAGGTTAGATATCCAGACCGCATAACTCTCATCAGAGGAAACCATGA
PP4-1_pMAL14_M13F       TCTACTCGCACTCAAGGTTAGATATCCAGACCGCATAACTCTCATCAGAGGAAACCATGA
                        ************************************************************

PP4-1                   AAGCAGGCAAATCACACAGGTTTATGGATTTTATGATGAGTGTTTGCGTAAATATGGCTC
PP4-1_pMAL14_M13F       AAGCAGGCAAATCACACAGGTTTATGGATTTTATGATGAGTGTTTGCGTAAATATGGCTC
                        ************************************************************

PP4-1                   TTCAAATGTCTGGAGATACTGCACCGACATTTTTGACTACATGAGTCTTTCAGCTGTTGT
PP4-1_pMAL14_M13F       TTCAAATGTCTGGAGATACTGCACCGACATTTTTGACTACATGAGTCTTTCAGCTGTTGT
                        ************************************************************

PP4-1                   GGAGAACAAGATATTCTGTGTTCATGGTGGTCTTTCTCCAGCTATTATGACTCTTGATCA
PP4-1_pMAL14_M13F       GGAGAACAAGATATTCTGTGTTCATGGTGGTCTTTCTCCAGCTATTATGACTCTTGATCA
                        ************************************************************

PP4-1                   GATTAGGACAATTGACCGGAAGCAAGAAGTACCACATGATGGTGCCATGTGTGATCTCCT
PP4-1_pMAL14_M13F       GATTAGGACAATTGACCGGAAGCAAGAAGTACCACATGATGGTGCCATGTGTGATCTCCT
                        ************************************************************

PP4-1                   ATGGTCTGATCCTGAAGATATTGTTGATGGCTGGGGATTGAGCCCTCGTGGTGCCGGATT
```

```
PP4-1_pMAL14_M13F      ATGGTCTGATCCTGAAGATATTGTTGATGGCTGGGGATTGAGCCCTCGTGGTGCCGGATT
                       ************************************************************

PP4-1                  CCTTTTTGGTGGCAGTGTTGTCACGTCTTTTAACCACTCAAACAACATAGACTACATAGC
PP4-1_pMAL14_M13F      CCTTTTTGGTGGCAGTGTTGTCACGTCTTTTAACCACTCAAACAACATAGACTACATAGC
                       ************************************************************

PP4-1                  CCGTGCCCATCAACTAGTTATGGAGGGTTACAAATGGATGTTTGATAGCCAGATTGTGAC
PP4-1_pMAL14_M13F      CCGTGCCCATCAACTAGTTATGGAGGGTTACAAATGGATGTTTGATAGCCAGATTGTGAC
                       ************************************************************

PP4-1                  AGTGTGGTCAGCCCCAAATTACTGTTACAGATGCGGTAATGTGGCTTCAATTCTAGAGCT
PP4-1_pMAL14_M13F      AGTGTGGTCAGCCCCAAATTACTGTTACAGATGCGGTAATGTGGCTTCAATTCTAGAGCT
                       ************************************************************

PP4-1                  TGACGAGAATCTAAATAAAGAATTCCGTGTGTTTGATGCAGCCCAGCAGGACTCGAGAGG
PP4-1_pMAL14_M13F      TGACGAGAATCTAAATAAAGAATTCCGTGTGTTTGATGCAGCCCAGCAGGACTCGAGAGG
                       ************************************************************

PP4-1                  GCCTCCCGCCAAAAAGCCGGCCCCTGATTACTTCCTATAA
PP4-1_pMAL14_M13F      GCCTCCCGCC------------------------------
                       **********
```

A-14: Multiple sequence alignment between PP4-1 sequence and PP4-1-pSF 1 using PP4.1_dTOPO F primer, missing the start of the gene and mutations before the stop codon

```
PP4-1                        ATGTCAGACCTAGATCGGCAAATAGGGCAGCTTAAGCGATGCGAACCATT
PP4-1-pFS_1_PP41_TOPO_F      --------------------------------------------------


PP4-1                        GAGCGAATCGGAGGTGAAGGCTCTTTGCCTCAAAGCCATGGAAATTCTTG
PP4-1-pFS_1_PP41_TOPO_F      --GCGAATCGGAGGTGAAGGCTCTTTGCCTCAAAGCCATGGAAATTCTTG
                               **********************************************

PP4-1                        TTGAAGAGAGTAATGTTCAGAGAGTTGATGCCCCTGTCACTTTATGTGGT
PP4-1-pFS_1_PP41_TOPO_F      TTGAAGAGAGTAATGTTCAGAGAGTTGATGCCCCTGTCACTTTATGTGGT
                             **************************************************

PP4-1                        GACATCCATGGGCAGTTCTATGATATGATGGAGCTTTTCAAAGTTGGGGG
PP4-1-pFS_1_PP41_TOPO_F      GACATCCATGGGCAGTTCTATGATATGATGGAGCTTTTCAAAGTTGGGGG
                             **************************************************

PP4-1                        TGATTGTCCTAAGACCAACTATTTGTTTATGGGAGATTTTGTTGATCGTG
PP4-1-pFS_1_PP41_TOPO_F      TGATTGTCCTAAGACCAACTATTTGTTTATGGGAGATTTTGTTGATCGTG
                             **************************************************

PP4-1                        GATATTATTCGGTTGAGACATTTCTACTTCTACTCGCACTCAAGGTTAGA
PP4-1-pFS_1_PP41_TOPO_F      GATATTATTCGGTTGAGACATTTCTACTTCTACTCGCACTCAAGGTTAGA
                             **************************************************

PP4-1                        TATCCAGACCGCATAACTCTCATCAGAGGAAACCATGAAAGCAGGCAAAT
PP4-1-pFS_1_PP41_TOPO_F      TATCCAGACCGCATAACTCTCATCAGAGGAAACCATGAAAGCAGGCAAAT
                             **************************************************

PP4-1                        CACACAGGTTTATGGATTTTATGATGAGTGTTTGCGTAAATATGGCTCTT
PP4-1-pFS_1_PP41_TOPO_F      CACACAGGTTTATGGATTTTATGATGAGTGTTTGCGTAAATATGGCTCTT
                             **************************************************

PP4-1                        CAAATGTCTGGAGATACTGCACCGACATTTTTGACTACATGAGTCTTTCA
PP4-1-pFS_1_PP41_TOPO_F      CAAATGTCTGGAGATACTGCACCGACATTTTTGACTACATGAGTCTTTCA
                             **************************************************

PP4-1                        GCTGTTGTGGAGAACAAGATATTCTGTGTTCATGGTGGTCTTTCTCCAGC
PP4-1-pFS_1_PP41_TOPO_F      GCTGTTGTGGAGAACAAGATATTCTGTGTTCATGGTGGTCTTTCTCCAGC
                             **************************************************

PP4-1                        TATTATGACTCTTGATCAGATTAGGACAATTGACCGGAAGCAAGAAGTAC
```

```
PP4-1-pFS_1_PP41_TOPO_F      TATTATGACTCTTGATCAGATTAGGACAATTGACCGGAAGCAAGAAGTAC
                             **************************************************

PP4-1                        CACATGATGGTGCCATGTGTGATCTCCTATGGTCTGATCCTGAAGATATT
PP4-1-pFS_1_PP41_TOPO_F      CACATGATGGTGCCATGTGTGATCTCCTATGGTCTGATCCTGAAGATATT
                             **************************************************

PP4-1                        GTTGATGGCTGGGGATTGAGCCCTCGTGGTGCCGGATTCCTTTTTGGTGG
PP4-1-pFS_1_PP41_TOPO_F      GTTGATGGCTGGGGATTGAGCCCTCGTGGTGCCGGATTCCTTTTTGGTGG
                             **************************************************

PP4-1                        CAGTGTTGTCACGTCTTTTAACCACTCAAACAACATAGACTACATAGCCC
PP4-1-pFS_1_PP41_TOPO_F      CAGTGTTGTCACGTCTTTTAACCACTCAAACAACATAGACTACATAGCCC
                             **************************************************

PP4-1                        GTGCCCATCAACTAGTTATGGAGGGTTACAAATGGATGTTTGATAGCCAG
PP4-1-pFS_1_PP41_TOPO_F      GTGCCCATCAACTAGTTATGGAGGGTTACAAATGGATGTTTGATAGCCAG
                             **************************************************

PP4-1                        ATTGTGACAGTGTGGTCAGCCCCAAATTACTGTTACAGATGCGGTAATGT
PP4-1-pFS_1_PP41_TOPO_F      ATTGTGACAGTGTGGTCAGCCCCAAATTACTGTTACAGATGCGGTAATGT
                             **************************************************

PP4-1                        GGCTTCAATTCTAGAGCTTGACGAGAATCTAAATAAAGAATTCCGTGTGT
PP4-1-pFS_1_PP41_TOPO_F      GGCTTCAATTCTAGAGCTTGACGAGAATCTAAATAAAGAATTCCGTGTGT
                             **************************************************

PP4-1                        TTGATGCAGCCCAGCAGGACTCGAGAGGGCCTCCCGCCAAAAAGCCGGCC
PP4-1-pFS_1_PP41_TOPO_F      TTGATGCAGCCCAGCAGGACTCGAGAGGGCCTCCCGCCAAAAAGCCGGCC
                             **************************************************

PP4-1                        CCTGATTACTTCCTATAA--------------------------------
PP4-1-pFS_1_PP41_TOPO_F      CCTGATTACTTCTAWAAAAAGGGGTGGGCGCGCCGACCCAGCTTTCTGGW
                             ************    **

PP4-1                        --------------------------------------------------
PP4-1-pFS_1_PP41_TOPO_F      ACAAAGTTGGCATTATAAGAAAGCATTGCTTATCAATTTGTTGCAACGAA

PP4-1                        --------------------------------------------------
PP4-1-pFS_1_PP41_TOPO_F      CAGGTCACTATCAGTCAAAATAAAATCATTATTTGCCATCCAGCTGATAT

PP4-1                        --------------------------------------------------
PP4-1-pFS_1_PP41_TOPO_F      CCCCTATAGTGAGTCGTATTACATGGTCATAGCTGTTTCCTGGCAGCTCT

PP4-1                        --------------------------------------------------
PP4-1-pFS_1_PP41_TOPO_F      GGCCCGTGTCTCAAAATCTCTGATGTTACATTGCMCAAGATAAAAATWTA

PP4-1                        -----------------------------------
PP4-1-pFS_1_PP41_TOPO_F      TCATCCTGAACAATAAAACTGTCTGCTTACATAAAC
```

A-15: Multiple sequence alignment between PP4-1 sequence and PP4-1-pSF 2 using PP4.1_dTOPO F primer, missing the start of the gene and one error around the stop codon

```
PP4-1                        ATGTCAGACCTAGATCGGCAAATAGGGCAGCTTAAGCGATGCGAACCATT
PP4-1-pFS_1_PP41_TOPO_F      ------------------------------------------------ATT
                                                                              ***

PP4-1                        GAGCGAATCGGAGGTGAAGGCTCTTTGCCTCAAAGCCATGGAAATTCTTG
PP4-1-pFS_1_PP41_TOPO_F      GAGCGAATCGGAGGTGAAGGCTCTTTGCCTCAAAGCCATGGAAATTCTTG
                             **************************************************

PP4-1                        TTGAAGAGAGTAATGTTCAGAGAGTTGATGCCCCTGTCACTTTATGTGGT
PP4-1-pFS_1_PP41_TOPO_F      TTGAAGAGAGTAATGTTCAGAGAGTTGATGCCCCTGTCACTTTATGTGGT
```

```
                          **************************************************

PP4-1                     GACATCCATGGGCAGTTCTATGATATGATGGAGCTTTTCAAAGTTGGGGG
PP4-1-pFS_1_PP41_TOPO_F   GACATCCATGGGCAGTTCTATGATATGATGGAGCTTTTCAAAGTTGGGGG
                          **************************************************

PP4-1                     TGATTGTCCTAAGACCAACTATTTGTTTATGGGAGATTTTGTTGATCGTG
PP4-1-pFS_1_PP41_TOPO_F   TGATTGTCCTAAGACCAACTATTTGTTTATGGGAGATTTTGTTGATCGTG
                          **************************************************

PP4-1                     GATATTATTCGGTTGAGACATTTCTACTTCTACTCGCACTCAAGGTTAGA
PP4-1-pFS_1_PP41_TOPO_F   GATATTATTCGGTTGAGACATTTCTACTTCTACTCGCACTCAAGGTTAGA
                          **************************************************

PP4-1                     TATCCAGACCGCATAACTCTCATCAGAGGAAACCATGAAAGCAGGCAAAT
PP4-1-pFS_1_PP41_TOPO_F   TATCCAGACCGCATAACTCTCATCAGAGGAAACCATGAAAGCAGGCAAAT
                          **************************************************

PP4-1                     CACACAGGTTTATGGATTTTATGATGAGTGTTTGCGTAAATATGGCTCTT
PP4-1-pFS_1_PP41_TOPO_F   CACACAGGTTTATGGATTTTATGATGAGTGTTTGCGTAAATATGGCTCTT
                          **************************************************

PP4-1                     CAAATGTCTGGAGATACTGCACCGACATTTTTGACTACATGAGTCTTTCA
PP4-1-pFS_1_PP41_TOPO_F   CAAATGTCTGGAGATACTGCACCGACATTTTTGACTACATGAGTCTTTCA
                          **************************************************

PP4-1                     GCTGTTGTGGAGAACAAGATATTCTGTGTTCATGGTGGTCTTTCTCCAGC
PP4-1-pFS_1_PP41_TOPO_F   GCTGTTGTGGAGAACAAGATATTCTGTGTTCATGGTGGTCTTTCTCCAGC
                          **************************************************

PP4-1                     TATTATGACTCTTGATCAGATTAGGACAATTGACCGGAAGCAAGAAGTAC
PP4-1-pFS_1_PP41_TOPO_F   TATTATGACTCTTGATCAGATTAGGACAATTGACCGGAAGCAAGAAGTAC
                          **************************************************

PP4-1                     CACATGATGGTGCCATGTGTGATCTCCTATGGTCTGATCCTGAAGATATT
PP4-1-pFS_1_PP41_TOPO_F   CACATGATGGTGCCATGTGTGATCTCCTATGGTCTGATCCTGAAGATATT
                          **************************************************

PP4-1                     GTTGATGGCTGGGGATTGAGCCCTCGTGGTGCCGGATTCCTTTTTGGTGG
PP4-1-pFS_1_PP41_TOPO_F   GTTGATGGCTGGGGATTGAGCCCTCGTGGTGCCGGATTCCTTTTTGGTGG
                          **************************************************

PP4-1                     CAGTGTTGTCACGTCTTTTAACCACTCAAACAACATAGACTACATAGCCC
PP4-1-pFS_1_PP41_TOPO_F   CAGTGTTGTCACGTCTTTTAACCACTCAAACAACATAGACTACATAGCCC
                          **************************************************

PP4-1                     GTGCCCATCAACTAGTTATGGAGGGTTACAAATGGATGTTTGATAGCCAG
PP4-1-pFS_1_PP41_TOPO_F   GTGCCCATCAACTAGTTATGGAGGGTTACAAATGGATGTTTGATAGCCAG
                          **************************************************

PP4-1                     ATTGTGACAGTGTGGTCAGCCCCAAATTACTGTTACAGATGCGGTAATGT
PP4-1-pFS_1_PP41_TOPO_F   ATTGTGACAGTGTGGTCAGCCCCAAATTACTGTTACAGATGCGGTAATGT
                          **************************************************

PP4-1                     GGCTTCAATTCTAGAGCTTGACGAGAATCTAAATAAAGAATTCCGTGTGT
PP4-1-pFS_1_PP41_TOPO_F   GGCTTCAATTCTAGAGCTTGACGAGAATCTAAATAAAGAATTCCGTGTGT
                          **************************************************

PP4-1                     TTGATGCAGCCCAGCAGGACTCGAGAGGGCCTCCCGCCAAAAAGCCGGCC
PP4-1-pFS_1_PP41_TOPO_F   TTGATGCAGCCCAGCAGGACTCGAGAGGGCCTCCCGCCAAAAAGCCGGCC
                          **************************************************

PP4-1                     CCTGATTACTTCCTATAA---------------------------------
PP4-1-pFS_1_PP41_TOPO_F   CCTGATTACTTC-TATAAAAAARGGGTGGGCGCGCCGACCCAGCTTTCTTG
                          ************ *****

PP4-1                     --------------------------------------------------
PP4-1-pFS_1_PP41_TOPO_F   TACAAAGTTGGCATTATAAGAAAGCATTGCTTATCAATTTGTTGCAACGA
```

```
PP4-1                          --------------------------------------------------
PP4-1-pFS_1_PP41_TOPO_F        ACAGGTCACTATCAGTCAAAATAAAATCATTATTTGCCATCCAGCTGATA

PP4-1                          --------------------------------------------------
PP4-1-pFS_1_PP41_TOPO_F        TCCCCTATAGTGAGTCGTATTACATGGTCATAGCTGTTTCCTGGCAGCTC

PP4-1                          --------------------------------------------------
PP4-1-pFS_1_PP41_TOPO_F        TGGCCCGTGTCTCAAAATCTCTGATGTTACATTGCMCAAGATAAAAATWT

PP4-1                          ------------------------------------------
PP4-1-pFS_1_PP41_TOPO_F        ATCATCATGAACAATAAAACTGTCTGCTTACWTAAACAGTAA
```

A-16: Multiple sequence alignment of PP4-2 sequence and PP4-2-pSF 1 using PP4.2_dTOPO F primer, start of gene not read, apparent mutations

```
PP4-2                          ATGTCAGACCTAGACAAGCAAATAGAGCAGCTTAAACGCTGCGAGGCTTT
PP4-2-pSF_1_PP42_TOPO_F        -------------------------------------------------TNN
                                                                                 *

PP4-2                          GAAGGAATCAGAAGTGAAGGCTCTTTGTCTTAAAGCTATGGAGATTCTAG
PP4-2-pSF_1_PP42_TOPO_F        TTGAGGATCAGAAGTGAAGGCTCTTTGTCTTAAAGCTATGGAGATTCTAG
                                   * ******************************************

PP4-2                          TTGAAGAGAGCAATGTTCAAAGAGTCGATGCTCCTGTCACTATATGTGGC
PP4-2-pSF_1_PP42_TOPO_F        TTGAAGAGAGCAATGTTCAAAGAGTCGATGCTCCTGCCACTATATGTGGC
                               ********************************** *************

PP4-2                          GACATTCATGGACAGTTCTATGACATGAAAGAGCTTTTCAAAGTTGGGGG
PP4-2-pSF_1_PP42_TOPO_F        GACATTCATGGACAGTTCTATGACATGAAAGAGCTTTTCAAAGTTGGGGG
                               **************************************************

PP4-2                          TGATTGCCCTAAGACCAATTATTTGTTTCTTGGAGATTTTGTTGACCGAG
PP4-2-pSF_1_PP42_TOPO_F        TGATTGCCCTAAGACCAATTATTTGTTTCTTGGAGATTTTGTTGACCGAG
                               **************************************************

PP4-2                          GTTTTTATTCGGTTGAGACATTTCTACTTCTTCTAGCTCTCAAGGTTAGA
PP4-2-pSF_1_PP42_TOPO_F        GTTTTTATTCGGTTGAGACATTTCTACTTCTTCTAGCTCTCAAGGTTAGA
                               **************************************************

PP4-2                          TATCCAGACCGTATAACTCTCATTAGAGGGAACCACGAGAGCCGGCAGAT
PP4-2-pSF_1_PP42_TOPO_F        TATCCAGACCGTATAACTCTCATTAGAGGGAACCACGAGAGCCGGCAGAT
                               **************************************************

PP4-2                          TACGCAGGTATATGGATTTTATGATGAGTGTCTGCGTAAATATGGCTCTG
PP4-2-pSF_1_PP42_TOPO_F        TACGCAGGTATATGGATTTTATGATGAGTGTCTGCGTAAATATGGCTCTG
                               **************************************************

PP4-2                          TAAATGTTTGGAGATACTGCACAGATATCTTTGACTACTTGAGTCTTTCA
PP4-2-pSF_1_PP42_TOPO_F        TAAATGTTTGGAGATACTGCACAGATATCTTTGACTACTTGAGTCTTTCA
                               **************************************************

PP4-2                          GCTCTTGTCGAGAACAAGATATTTTGTGTTCATGGAGGTCTCTCTCCAGC
PP4-2-pSF_1_PP42_TOPO_F        GCTCTTGTCGAGAACAAGATATTTTGTGTTCATGGAGGTCTCTCTCCAGC
                               **************************************************

PP4-2                          TATTATGACTCTAGACCAGATCAGGGCTATTGATCGAAAGCAAGAAGTAC
PP4-2-pSF_1_PP42_TOPO_F        TATTATGACTCTAGACCAGATCAGGGCTATTGATCGAAAGCAAGAAGTAC
                               **************************************************

PP4-2                          CACATGATGGTGCTATGTGTGATCTTTTATGGTCTGATCCAGAAGATATT
PP4-2-pSF_1_PP42_TOPO_F        CACATGATGGTGCTATGTGTGATCTTTTATGGTCTGATCCAGAAGATATT
                               **************************************************
```

```
PP4-2                       GTCGATGGTTGGGGATTGAGTCCCCGTGGTGCCGGATTCCTTTTCGGCGG
PP4-2-pSF_1_PP42_TOPO_F     GTCGATGGTTGGGGATTGAGTCCCCGTGGTGCCGGATTCCTTTTCGGCGG
                            **************************************************

PP4-2                       CAGTGTTGTTACGTCTTTTAACCACTCAAACAACATTGATTACATATGTC
PP4-2-pSF_1_PP42_TOPO_F     CAGTGTTGTTACGTCTTTTAACCACTCAAACAACATTGATTACATATGTC
                            **************************************************

PP4-2                       GAGCTCATCAGCTAGTGATGGAAGGTTACAAATGGATGTTCAATAGCCAG
PP4-2-pSF_1_PP42_TOPO_F     GAGCTCATCAGCTAGTGATGGAAGGTTACAAATGGATGTTCAATAGCCAG
                            **************************************************

PP4-2                       ATAGTCACTGTTTGGTCTGCCCCAAATTACTGTTATAGATGCGGTAATGT
PP4-2-pSF_1_PP42_TOPO_F     ATAGTCACTGTTTGGTCTGCCCCAAATTACTGTTATAGATGCGGTAATGT
                            **************************************************

PP4-2                       AGCTGCAATTCTAGAGCTCGATGAGAATCTAAACAAAGAGTTTCGTGTCT
PP4-2-pSF_1_PP42_TOPO_F     AGCTGCAATTCTAGAGCTCGATGAGAATCTAAACAAAGAGTTTCGTGTCT
                            **************************************************

PP4-2                       TCGATGCAGCCCCACAAGAATCGAGAGGAGCTCTAGCCAAGAAACCTGCA
PP4-2-pSF_1_PP42_TOPO_F     TCGATGCAGCCCCACAAGAATCGAGAGGAGCTCTAGCCAAGAAACCTGCA
                            **************************************************

PP4-2                       CCTGATTATTTCCTGTGA--------------------------------
PP4-2-pSF_1_PP42_TOPO_F     CCTGATTATTTCCTGTGAAAAGGGTGGGCGCGCCGACCCAGCTTTCTTGT
                            ******************

PP4-2                       --------------------------------------------------
PP4-2-pSF_1_PP42_TOPO_F     ACAAAGTTGGCATTATAAGAAAGCATTGCTTATCAATTTGTTGCAACGAA


PP4-2                       --------------------------------------------------
PP4-2-pSF_1_PP42_TOPO_F     CAGGTCACTATCAGTCAAAATAAAATCATTATTTGCCATCCAGCTGATAT


PP4-2                       --------------------------------------------------
PP4-2-pSF_1_PP42_TOPO_F     CCCCTATAGTGAGTCGTATTACATGGTCATAGCTGTTTCCTGGCAGCTCT


PP4-2                       --------------------------------------------------
PP4-2-pSF_1_PP42_TOPO_F     GGCCCGTGTCTCAAAATCTCTGATGTTACATTGCACAAGATAAAAATATA


PP4-2                       --------------------------------
PP4-2-pSF_1_PP42_TOPO_F     TCATCATGAACAATAAAACTGTCTGCTTACA
```

A-17: Multiple sequence alignment of PP4-2 sequence and PP4-2-pSF 1 using PP4.2_dTOPO F primer, start of gene not sequenced, some nucleotides were not sequenced correctly, possible mutation

```
PP4-2                       ATGTCAGACCTAGACAAGCAAATAGAGCAGCTTAAACGCTGCGAGGCTTT
PP4-2-pSF_2_PP42_TOPO_F     --------------------------------------------------


PP4-2                       GAAGGAATCAGAAGTGAAGGCTCTTTGTCTTAAAGCTATGGAGATTCTAG
PP4-2-pSF_2_PP42_TOPO_F     --------------------------------------------------


PP4-2                       TTGAAGAGAGCAATGTTCAAAGAGTCGATGCTCCTGTCACTATATGTGGC
PP4-2-pSF_2_PP42_TOPO_F     --------------------------------------------------


PP4-2                       GACATTCATGGACAGTTCTATGACATGAAAGAGCTTTTCAAAGTTGGGGG
PP4-2-pSF_2_PP42_TOPO_F     --------------------------------------------------
```

```
PP4-2                          TGATTGCCCTAAGACCAATTATTTGTTTCTTGGAGATTTTGTTGACCGAG
PP4-2-pSF_2_PP42_TOPO_F        ----------------------TGWTYCTTGGAGATTTTGTKGACCGAG
                                                     ** * ************** *******

PP4-2                          GTTTTTATTCGGTTGAGACATTTCTACTTCTTCTAGCTCTCAAGGTTAGA
PP4-2-pSF_2_PP42_TOPO_F        GTTTTTATTCGGTTGAGACATTTCTACTTCTTCTAGCTCTCAAGGTTAGA
                               **************************************************

PP4-2                          TATCCAGACCGTATAACTCTCATTAGAGGGAACCACGAGAGCCGGCAGAT
PP4-2-pSF_2_PP42_TOPO_F        TATCCAGACCGTATAACTCTCATTAGAGGGAACCACGAGAGCCGGCAGAT
                               **************************************************

PP4-2                          TACGCAGGTATATGGATTTTATGATGAGTGTCTGCGTAAATATGGCTCTG
PP4-2-pSF_2_PP42_TOPO_F        TACGCAGGTATATGGATTTTATGATGAGTGTCTGCGTAAATATGGCTCTG
                               **************************************************

PP4-2                          TAAATGTTTGGAGATACTGCACAGATATCTTTGACTACTTGAGTCTTTCA
PP4-2-pSF_2_PP42_TOPO_F        TAAATGTTTGGAGATACTGCACAGATATCTTTGACTACTTGAGTCTTTCA
                               **************************************************

PP4-2                          GCTCTTGTCGAGAACAAGATATTTTGTGTTCATGGAGGTCTCTCTCCAGC
PP4-2-pSF_2_PP42_TOPO_F        GCTCTTGTCGAGAACAAGATATTTTGTGTTCATGGAGGTCTCTCTCCAGC
                               **************************************************

PP4-2                          TATTATGACTCTAGACCAGATCAGGGCTATTGATCGAAAGCAAGAAGTAC
PP4-2-pSF_2_PP42_TOPO_F        TATTATGACTCTAGACCAGATCAGGGCTATTGATCGAAAGCAAGAAGTAC
                               **************************************************

PP4-2                          CACATGATGGTGCTATGTGTGATCTTTTATGGTCTGATCCAGAAGATATT
PP4-2-pSF_2_PP42_TOPO_F        CACATGATGGTGCTATGTGTGATCTTTTATGGTCTGATCCAGAAGATATT
                               **************************************************

PP4-2                          GTCGATGGTTGGGGATTGAGTCCCCGTGGTGCCGGATTCCTTTTCGGCGG
PP4-2-pSF_2_PP42_TOPO_F        GTCGATGGTTGGGGATTGAGTCCCCGTGGTGCCGGATTCCTTTTCGGCGG
                               **************************************************

PP4-2                          CAGTGTTGTTACGTCTTTTAACCACTCAAACAACATTGATTACATATGTC
PP4-2-pSF_2_PP42_TOPO_F        CAGTGTTGTTACGTCTTTTAACCACTCAAACAACATTGATTACATATGTC
                               **************************************************

PP4-2                          GAGCTCATCAGCTAGTGATGGAAGGTTACAAATGGATGTTCAATAGCCAG
PP4-2-pSF_2_PP42_TOPO_F        GAGCTCATCAGCTAGTGATGGAAGGTTACAAATGGATGTTCAATAGCCAG
                               **************************************************

PP4-2                          ATAGTCACTGTTTGGTCTGCCCCAAATTACTGTTATAGATGCGGTAATGT
PP4-2-pSF_2_PP42_TOPO_F        ATAGTCACTGTTTGGTCTGCCCCAAATTACTGTTATAGATGCGGTAATGT
                               **************************************************

PP4-2                          AGCTGCAATTCTAGAGCTCGATGAGAATCTAAACAAAGAGTTTCGTGTCT
PP4-2-pSF_2_PP42_TOPO_F        AGCTGCAATTCTAGAGCTCGATGAGAATCTAAACAAAGAGTTTCGTGTCT
                               **************************************************

PP4-2                          TCGATGCAGCCCCACAAGAATCGAGAGGAGCTCTAGCCAAGAAACCTGCA
PP4-2-pSF_2_PP42_TOPO_F        TCGATGCAGCCCCACAAGAATCGAGAGGAGCTCTAGCCAAGAAACCTGCA
                               **************************************************

PP4-2                          CCTGATTATTTCCTGTGA--------------------------------
PP4-2-pSF_2_PP42_TOPO_F        CCTGATTATT-CCTGNGAAAAGGGTGGGCGCGCCGACCCAGCTTTCTTGT
                               ********** **** **

PP4-2                          --------------------------------------------------
PP4-2-pSF_2_PP42_TOPO_F        ACAAAGTTGGCATTATAAGAAAGCATTGCTTATCAATTTGTTGCAACGAA


PP4-2                          --------------------------------------------------
PP4-2-pSF_2_PP42_TOPO_F        CAGGTCACTATCAGTCAAAATAAAATCATTATTTGCCATCCAGCTGATAT
```

```
PP4-2                          --------------------------------------------------
PP4-2-pSF_2_PP42_TOPO_F        CCCCTATAGTGAGTCGTATTACATGGTCATAGCTGTTTCCTGGCAGCTCT


PP4-2                          --------------------------------------------------
PP4-2-pSF_2_PP42_TOPO_F        GGCCCGTGTCTCAAAATCTCTGATGTTACATTGCMCAAGATAAAAATWTA


PP4-2                          ----------------------------------------
PP4-2-pSF_2_PP42_TOPO_F        TCATCATGAACAATAAAACTGTCTGCTTACWTAAACAGTAA
```

A-18: Multiple sequence alignment of PP4-2 sequence, PP4-2-pSF A and PP4-2-pSF B using PP4.2_dTOPO F primer, start of genes not sequenced, known mutation

```
PP4-2                          ------------ATGTCAGACCTAGACAAGCAAATAGAGCAGCTTAAACG
PP4-2-pSF_B_PP42_TOPOF         NNGNCTTTGAGGATGTCAGACCTAGACAAGCAAATAGAGCAGCTTAAACG
PP4-2-pSF_A_PP42_TOPOF         --------------------------------------------------


PP4-2                          CTGCGAGGCTTTGAAGGAATCAGAAGTGAAGGCTCTTTGTCTTAAAGCTA
PP4-2-pSF_B_PP42_TOPOF         CTGCGAGGCTTTGAAGGAATCAGAAGTGAAGGCTCTTTGTCTTAMAGCTA
PP4-2-pSF_A_PP42_TOPOF         ------------------NTNNGAAGTGAAGGCTCTTTGTCTTAAAGCTA
                                                 *  ********************** *****

PP4-2                          TGGAGATTCTAGTTGAAGAGAGCAATGTTCAAAGAGTCGATGCTCCTGTC
PP4-2-pSF_B_PP42_TOPOF         TGGAGATTCTAGTTGAAGAGAGCAATGTTCAAAGAGTCGATGCTCCTGCC
PP4-2-pSF_A_PP42_TOPOF         TGGAGATTCTAGTTGAAGAGAGCAATGTTCAAAGAGTCGATGCTCCTGCC
                               ************************************************* *

PP4-2                          ACTATATGTGGCGACATTCATGGACAGTTCTATGACATGAAAGAGCTTTT
PP4-2-pSF_B_PP42_TOPOF         ACTATATGTGGCGACATTCATGGACAGTTCTATGACATGAAAGAGCTTTT
PP4-2-pSF_A_PP42_TOPOF         ACTATATGTGGCGACATTCATGGACAGTTCTATGACATGAAAGAGCTTTT
                               **************************************************

PP4-2                          CAAAGTTGGGGGTGATTGCCCTAAGACCAATTATTTGTTTCTTGGAGATT
PP4-2-pSF_B_PP42_TOPOF         CAAAGTTGGGGGTGATTGCCCTAAGACCAATTATTTGTTTCTTGGAGATT
PP4-2-pSF_A_PP42_TOPOF         CAAAGTTGGGGGTGATTGCCCTAAGACCAATTATTTGTTTCTTGGAGATT
                               **************************************************

PP4-2                          TTGTTGACCGAGGTTTTTATTCGGTTGAGACATTTCTACTTCTTCTAGCT
PP4-2-pSF_B_PP42_TOPOF         TTGTTGACCGAGGTTTTTATTCGGTTGAGACATTTCTACTTCTTCTAGCT
PP4-2-pSF_A_PP42_TOPOF         TTGTTGACCGAGGTTTTTATTCGGTTGAGACATTTCTACTTCTTCTAGCT
                               **************************************************

PP4-2                          CTCAAGGTTAGATATCCAGACCGTATAACTCTCATTAGAGGGAACCACGA
PP4-2-pSF_B_PP42_TOPOF         CTCAAGGTTAGATATCCAGACCGTATAACTCTCATTAGAGGGAACCACGA
PP4-2-pSF_A_PP42_TOPOF         CTCAAGGTTAGATATCCAGACCGTATAACTCTCATTAGAGGGAACCACGA
                               **************************************************

PP4-2                          GAGCCGGCAGATTACGCAGGTATATGGATTTTATGATGAGTGTCTGCGTA
PP4-2-pSF_B_PP42_TOPOF         GAGCCGGCAGATTACGCAGGTATATGGATTTTATGATGAGTGTCTGCGTA
PP4-2-pSF_A_PP42_TOPOF         GAGCCGGCAGATTACGCAGGTATATGGATTTTATGATGAGTGTCTGCGTA
                               **************************************************

PP4-2                          AATATGGCTCTGTAAATGTTTGGAGATACTGCACAGATATCTTTGACTAC
PP4-2-pSF_B_PP42_TOPOF         AATATGGCTCTGTAAATGTTTGGAGATACTGCACAGATATCTTTGACTAC
PP4-2-pSF_A_PP42_TOPOF         AATATGGCTCTGTAAATGTTTGGAGATACTGCACAGATATCTTTGACTAC
                               **************************************************

PP4-2                          TTGAGTCTTTCAGCTCTTGTCGAGAACAAGATATTTTGTGTTCATGGAGG
PP4-2-pSF_B_PP42_TOPOF         TTGAGTCTTTCAGCTCTTGTCGAGAACAAGATATTTTGTGTTCATGGAGG
PP4-2-pSF_A_PP42_TOPOF         TTGAGTCTTTCAGCTCTTGTCGAGAACAAGATATTTTGTGTTCATGGAGG
                               **************************************************

PP4-2                          TCTCTCTCCAGCTATTATGACTCTAGACCAGATCAGGGCTATTGATCGAA
```

```
PP4-2-pSF_B_PP42_TOPOF      TCTCTCTCCAGCTATTATGACTCTAGACCAGATCAGGGCTATTGATCGAA
PP4-2-pSF_A_PP42_TOPOF      TCTCTCTCCAGCTATTATGACTCTAGACCAGATCAGGGCTATTGATCGAA
                            **************************************************

PP4-2                       AGCAAGAAGTACCACATGATGGTGCTATGTGTGATCTTTTATGGTCTGAT
PP4-2-pSF_B_PP42_TOPOF      AGCAAGAAGTACCACATGATGGTGCTATGTGTGATCTTTTATGGTCTGAT
PP4-2-pSF_A_PP42_TOPOF      AGCAAGAAGTACCACATGATGGTGCTATGTGTGATCTTTTATGGTCTGAT
                            **************************************************

PP4-2                       CCAGAAGATATTGTCGATGGTTGGGGATTGAGTCCCCGTGGTGCCGGATT
PP4-2-pSF_B_PP42_TOPOF      CCAGAAGATATTGTCGATGGTTGGGGATTGAGTCCCCGTGGTGCCGGATT
PP4-2-pSF_A_PP42_TOPOF      CCAGAAGATATTGTCGATGGTTGGGGATTGAGTCCCCGTGGTGCCGGATT
                            **************************************************

PP4-2                       CCTTTTCGGCGGCAGTGTTGTTACGTCTTTTAACCACTCAAACAACATTG
PP4-2-pSF_B_PP42_TOPOF      CCTTTTCGGCGGCAGTGTTGTTACGTCTTTTAACCACTCAAACAACATTG
PP4-2-pSF_A_PP42_TOPOF      CCTTTTCGGCGGCAGTGTTGTTACGTCTTTTAACCACTCAAACAACATTG
                            **************************************************

PP4-2                       ATTACATATGTCGAGCTCATCAGCTAGTGATGGAAGGTTACAAATGGATG
PP4-2-pSF_B_PP42_TOPOF      ATTACATATGTCGAGCTCATCAGCTAGTGATGGAAGGTTACAAATGGATG
PP4-2-pSF_A_PP42_TOPOF      ATTACATATGTCGAGCTCATCAGCTAGTGATGGAAGGTTACAAATGGATG
                            **************************************************

PP4-2                       TTCAATAGCCAGATAGTCACTGTTTGGTCTGCCCCAAATTACTGTTATAG
PP4-2-pSF_B_PP42_TOPOF      TTCAATAGCCAGATAGTCACTGTTTGGTCTGCCCCAAATTACTGTTATAG
PP4-2-pSF_A_PP42_TOPOF      TTCAATAGCCAGATAGTCACTGTTTGGTCTGCCCCAAATTACTGTTATAG
                            **************************************************

PP4-2                       ATGCGGTAATGTAGCTGCAATTCTAGAGCTCGATGAGAATCTAAACAAAG
PP4-2-pSF_B_PP42_TOPOF      ATGCGGTAATGTAGCTGCAATTCTAGAGCTCGATGAGAATCTAAACAAAG
PP4-2-pSF_A_PP42_TOPOF      ATGCGGTAATGTAGCTGCAATTCTAGAGCTCGATGAGAATCTAAACAAAG
                            **************************************************

PP4-2                       AGTTTCGTGTCTTCGATGCAGCCCCACAAGAATCGAGAGGAGCTCTAGCC
PP4-2-pSF_B_PP42_TOPOF      AGTTTCGTGTCTTCGATGCAGCCCCACAAGAATCGAGAGGAGCTCTAGCC
PP4-2-pSF_A_PP42_TOPOF      AGTTTCGTGTCTTCGATGCAGCCCCACAAGAATCGAGAGGAGCTCTAGCC
                            **************************************************

PP4-2                       AAGAAACCTGCACCTGATTATTTCCTGTGA--------------------
PP4-2-pSF_B_PP42_TOPOF      AAGAAACCTGCACCTGATTATTTCCTGTGAAAGGGTGGGCGCGCCGACCC
PP4-2-pSF_A_PP42_TOPOF      AAGAAACCTGCACCTGATTATTTCCTGTGAAAGGGTGGGCGCGCCGACCC
                            ******************************

PP4-2                       --------------------------------------------------
PP4-2-pSF_B_PP42_TOPOF      AGCTTTCTTGTACAAAGTGGTTGATAACAGCGCTTAGAGCTCGAATTTCC
PP4-2-pSF_A_PP42_TOPOF      AGCTTTCTTGTACAAAGTGGTTGATAACAGCGCTTAGAGCTCGAATTTCC


PP4-2                       ------------
PP4-2-pSF_B_PP42_TOPOF      CCGATCGTTCMA
PP4-2-pSF_A_PP42_TOPOF      CCGATCGTYCMA
```

A-19: Multiple sequence alignment of PP4-2 sequence, PP4-2-pSF A and PP4-2-pSF B using PP4.2_dTOPO R primer, end of gene not sequenced, known mutation

```
PP4-2-pSF_A_PP42_TOPOR      --TTTCCACTGACGTAAGGGATGACGCCAATCCCNNTATCTTTGGCAAGA
PP4-2-pSF_B_PP42_TOPOR      GATTTCCACTGACGTAAGGGATGACGCCAATCCCAT-ATCCTTGGCAAGA
PP4-2                       --------------------------------------------------


PP4-2-pSF_A_PP42_TOPOR      CCCTTCCNTNNNTTAAGGANAGTTCATTTCATTTGNGAGAGAACACGGGG
PP4-2-pSF_B_PP42_TOPOR      CCCTTCCTCTAT---AAGGAAGTTCATTTCATTTGG-AGAGAACACGGGG
PP4-2                       --------------------------------------------------


PP4-2-pSF_A_PP42_TOPOR      GACTCTAGTAACGGCCGCCAGTGTGCTGGAATTCGCCGCCCACCATGGAT
```

```
PP4-2-pSF_B_PP42_TOPOR      GACTCTAGTAACGGCCGCCAGTGTGCTGGAATTCGCCGCC-ACCATGGAT
PP4-2                       ----------------------------------------------------

PP4-2-pSF_A_PP42_TOPOR      TATAAAGATGATGATGATAAAGGGTCGGCCGCCAGCTGGAGCCACCCTCA
PP4-2-pSF_B_PP42_TOPOR      TATAAAGATGATGATGATAAAGGGTCGGCCGCCAGCTGGAGCCACCCTCA
PP4-2                       ----------------------------------------------------

PP4-2-pSF_A_PP42_TOPOR      GTTCGAGAAGGGAGGAGGAAGCGGCGGAGGCAGCGGAGGAGGAAGGCTGG
PP4-2-pSF_B_PP42_TOPOR      GTTCGAGAAGGGAGGAGGAAGCGGCGGAGGCAGCGGAGGAGGAAGGCTGG
PP4-2                       ----------------------------------------------------

PP4-2-pSF_A_PP42_TOPOR      AGCCACCCGCAGTTCGAGAAAGGAGCTAGATCAACAAGTTTGTACAAAAA
PP4-2-pSF_B_PP42_TOPOR      AGCCACCCGCAGTTCGAGAAAGGAGCTAGATCAACAAGTTTGTACAAAAA
PP4-2                       ----------------------------------------------------

PP4-2-pSF_A_PP42_TOPOR      AGCAGGCTCCGCGGCCGCCCCCTTCACCATGTCAGACCTAGACAAGCAAA
PP4-2-pSF_B_PP42_TOPOR      AGCAGGCTCCGCGGCCGCCCCCTTCACCATGTCAGACCTAGACAAGCAAA
PP4-2                       ----------------------------ATGTCAGACCTAGACAAGCAAA
                                                         **********************

PP4-2-pSF_A_PP42_TOPOR      TAGAGCAGCTTAAACGCTGCGAGGCTTTGAAGGAATCAGAAGTGAAGGCT
PP4-2-pSF_B_PP42_TOPOR      TAGAGCAGCTTAAACGCTGCGAGGCTTTGAAGGAATCAGAAGTGAAGGCT
PP4-2                       TAGAGCAGCTTAAACGCTGCGAGGCTTTGAAGGAATCAGAAGTGAAGGCT
                            **************************************************

PP4-2-pSF_A_PP42_TOPOR      CTTTGTCTTAAAGCTATGGAGATTCTAGTTGAAGAGAGCAATGTTCAAAG
PP4-2-pSF_B_PP42_TOPOR      CTTTGTCTTAAAGCTATGGAGATTCTAGTTGAAGAGAGCAATGTTCAAAG
PP4-2                       CTTTGTCTTAAAGCTATGGAGATTCTAGTTGAAGAGAGCAATGTTCAAAG
                            **************************************************

PP4-2-pSF_A_PP42_TOPOR      AGTCGATGCTCCTGCCACTATATGTGGCGACATTCATGGACAGTTCTATG
PP4-2-pSF_B_PP42_TOPOR      AGTCGATGCTCCTGCCACTATATGTGGCGACATTCATGGACAGTTCTATG
PP4-2                       AGTCGATGCTCCTGTCACTATATGTGGCGACATTCATGGACAGTTCTATG
                            ************** ***********************************

PP4-2-pSF_A_PP42_TOPOR      ACATGAAAGAGCTTTTCAAAGTTGGGGGTGATTGCCCTAAGACCAATTAT
PP4-2-pSF_B_PP42_TOPOR      ACATGAAAGAGCTTTTCAAAGTTGGGGGTGATTGCCCTAAGACCAATTAT
PP4-2                       ACATGAAAGAGCTTTTCAAAGTTGGGGGTGATTGCCCTAAGACCAATTAT
                            **************************************************

PP4-2-pSF_A_PP42_TOPOR      TTGTTTCTTGGAGATTTTGTTGACCGAGGTTTTTATTCGGTTGAGACATT
PP4-2-pSF_B_PP42_TOPOR      TTGTTTCTTGGAGATTTTGTTGACCGAGGTTTTTATTCGGTTGAGACATT
PP4-2                       TTGTTTCTTGGAGATTTTGTTGACCGAGGTTTTTATTCGGTTGAGACATT
                            **************************************************

PP4-2-pSF_A_PP42_TOPOR      TCTACTTCTTCTAGCTCTCAAGGTTAGATATCCAGACCGTATAACTCTCA
PP4-2-pSF_B_PP42_TOPOR      TCTACTTCTTCTAGCTCTCAAGGTTAGATATCCAGACCGTATAACTCTCA
PP4-2                       TCTACTTCTTCTAGCTCTCAAGGTTAGATATCCAGACCGTATAACTCTCA
                            **************************************************

PP4-2-pSF_A_PP42_TOPOR      TTAGAGGGAACCACGAGAGCCGGCAGATTACGCAGGTATATGGATTTTAT
PP4-2-pSF_B_PP42_TOPOR      TTAGAGGGAACCACGAGAGCCGGCAGATTACGCAGGTATATGGATTTTAT
PP4-2                       TTAGAGGGAACCACGAGAGCCGGCAGATTACGCAGGTATATGGATTTTAT
                            **************************************************

PP4-2-pSF_A_PP42_TOPOR      GATGAGTGTCTGCGTAAATATGGCTCTGTAAATGTTTGGAGATACTGCAC
PP4-2-pSF_B_PP42_TOPOR      GATGAGTGTCTGCGTAAATATGGCTCTGTAAATGTTTGGAGATACTGCAC
PP4-2                       GATGAGTGTCTGCGTAAATATGGCTCTGTAAATGTTTGGAGATACTGCAC
                            **************************************************

PP4-2-pSF_A_PP42_TOPOR      AGATATCTTTGACTACTTGAGTCTTTCAGCTCTTGTCGAGAACAAGATAT
PP4-2-pSF_B_PP42_TOPOR      AGATATCTTTGACTACTTGAGTCTTTCAGCTCTTGTCGAGAACAAGATAT
PP4-2                       AGATATCTTTGACTACTTGAGTCTTTCAGCTCTTGTCGAGAACAAGATAT
                            **************************************************
```

```
PP4-2-pSF_A_PP42_TOPOR      TTTGTGTTCATGGAGGTCTCTCTCCAGCTATTATGACTCTAGACCAGATC
PP4-2-pSF_B_PP42_TOPOR      TTTGTGTTCATGGAGGTCTCTCTCCAGCTATTATGACTCTAGACCAGATC
PP4-2                       TTTGTGTTCATGGAGGTCTCTCTCCAGCTATTATGACTCTAGACCAGATC
                            **************************************************

PP4-2-pSF_A_PP42_TOPOR      AGGGCTATTGATCGAAAGCAAGAAGTACCACATGATGGTGCTATGTGTGA
PP4-2-pSF_B_PP42_TOPOR      AGGGCTATTGATCGAAAGCAAGAAGTACCACATGATGGTGCTATGTGTGA
PP4-2                       AGGGCTATTGATCGAAAGCAAGAAGTACCACATGATGGTGCTATGTGTGA
                            **************************************************

PP4-2-pSF_A_PP42_TOPOR      TCTTTTATGGTCTGATCCAGAAGATATTGTCGATGGTTGGGGATTGAGTC
PP4-2-pSF_B_PP42_TOPOR      TCTTTTATGGTCTGATCCAGAAGATATTGTCGATGGTTGGGGATTGAGTC
PP4-2                       TCTTTTATGGTCTGATCCAGAAGATATTGTCGATGGTTGGGGATTGAGTC
                            **************************************************

PP4-2-pSF_A_PP42_TOPOR      CCCGTGGTGCCGGATTCCTTTTCGGCGGCAGTGTTGTTACGTCTTTTAAC
PP4-2-pSF_B_PP42_TOPOR      CCCGTGGTGCCGGATTCCTTTTCGGCGGCAGTGTTGTTACGTCTTTTAAC
PP4-2                       CCCGTGGTGCCGGATTCCTTTTCGGCGGCAGTGTTGTTACGTCTTTTAAC
                            **************************************************

PP4-2-pSF_A_PP42_TOPOR      CACTCAAACAACATTGATTACATATGTCGAGCTCATCAGCTAGTGATGGA
PP4-2-pSF_B_PP42_TOPOR      CACTCAAACAACATTGATTACATATGTCGAGCTCATCAGCTAGTGATGGA
PP4-2                       CACTCAAACAACATTGATTACATATGTCGAGCTCATCAGCTAGTGATGGA
                            **************************************************

PP4-2-pSF_A_PP42_TOPOR      AGGTTACAAATGGATGTTCAATAGCCAGATAGTCACTGTTTGGTCTGCCC
PP4-2-pSF_B_PP42_TOPOR      AGGTTACAAATGGATGTTCAATAGCCAGATAGTCACTGTTTGGTCTGCCC
PP4-2                       AGGTTACAAATGGATGTTCAATAGCCAGATAGTCACTGTTTGGTCTGCCC
                            **************************************************

PP4-2-pSF_A_PP42_TOPOR      CAAATTACTGTTATAGATGCGGTAATGTAGCTGCAATTCTAGAGCTCGAT
PP4-2-pSF_B_PP42_TOPOR      CAAATTACTGTTATAGATGCGGTAATGTAGCTGCAATTCTAGAGCTCGAT
PP4-2                       CAAATTACTGTTATAGATGCGGTAATGTAGCTGCAATTCTAGAGCTCGAT
                            **************************************************

PP4-2-pSF_A_PP42_TOPOR      GAGAATCTAAACAAAGAGTTTCGTGTCNT----------------------
PP4-2-pSF_B_PP42_TOPOR      GAGAATCTAAACAAAGAGTTTCGTGTCTTCGATGCA--------------
PP4-2                       GAGAATCTAAACAAAGAGTTTCGTGTCTTCGATGCAGCCCCACAAGAATC
                            *************************** *

PP4-2-pSF_A_PP42_TOPOR      ---------------------------------------------
PP4-2-pSF_B_PP42_TOPOR      ---------------------------------------------
PP4-2                       GAGAGGAGCTCTAGCCAAGAAACCTGCACCTGATTATTTCCTGTGA
```

A-20: Multiple sequence alignment of a part of pGWB2 and the sequence of PP4-2-pSF A sequenced with pSF R primer

```
pGWB2                   CATAAATAACGTCATGCATTACATGTTAATTATTACATGCTTAACGTAATTCAACAGAAA
PP4-2-pSF_A_pSFR        ----------------------ATGTTAATTATTACATGCTTAACGTAATTCAACAGAAA
                                               **************************************

pGWB2                   TTATATGATAATCATCGCAAGACCGGCAACAGGATTCAATCTTAAGAAACTTTATTGCCA
PP4-2-pSF_A_pSFR        TTATATGATAATCATCGCAAGACCGGCAACAGGATTCAATCTTAAGAAACTTTATTGCCA
                        ************************************************************

pGWB2                   AATGTTTGAACGATCGGGGATCATCCGGGTCTGTGGCGGGAACTCCACGAAAATATCCGA
PP4-2-pSF_A_pSFR        AATGTTTGAACGATCGGGGATCATCCRGGTCTGTGGCGGGAACTCCACGAAAATATMCGA
                        ************************* ***************************** ***

pGWB2                   ACGCAGCAAGATATCGCGGTGCATCTCGGTCTTGCCTGGGCAGTCGCCGCCGACGCCGTT
PP4-2-pSF_A_pSFR        ACGCAGCGRGATRTCGCGGTGCATCTCGGTCTTGCCTGGGCAGTCGCCGCCGACGCCGTT
                        *******  *** ********************************************

pGWB2                   GATGTGGACGCCGGGCCCGATCATATTGTCGCTCAGGATCGTGGCGTTGTGCTTGTCGGC
PP4-2-pSF_A_pSFR        GATGTGGACGCCGGGCCCGATCATATTGTCGCTCAGGATCGTGGCGTTGTGCTTGTCGGC
                        ************************************************************
```

```
pGWB2                 CGTTGCTGTCGTAATGATATCGGCACCTTCGACCGCCTGTTCCGCAGAGATCCCGTGGGC
PP4-2-pSF_A_pSFR      CGTTGCTGTCGTAATGATATCGGCACCTTCGACCGCCTGTTCCGCAGAGATCCCGTGGGC
                      ************************************************************

pGWB2                 GAAGAACTCCAGCATGAGATCCCCGCGCTGGAGGATCATCCAGCCGGCGTCCCGGAAAAC
PP4-2-pSF_A_pSFR      GAAGAACTCCAGCATGAGATCCCCGCGCTGGAGGATCATCCAGCCGGCGTCCCGGAAAAC
                      ************************************************************

pGWB2                 GATTCCGAAGCCCAACCTTTCATAGAAGGCGGCGGTGGAATCGAAATCTCGTGATGGCAG
PP4-2-pSF_A_pSFR      GATTCCGAAGCCCAACCTTTCATAGAAGGCGGCGGTGGAATCGAAATCTCGTGATGGCAG
                      ************************************************************

pGWB2                 GTTGGGCGTCGCTTGGTCGGTCATTTCGAACCCCAGAGTCCCGCTCAGAAGAACTCGTCA
PP4-2-pSF_A_pSFR      GTTGGGCGTCGCTTGGTCGGTCATTTCGAACCCCAGAGTCCCGCTCAGAAGAACTCGTCA
                      ************************************************************

pGWB2                 AGAAGGCGATAGAAGGCGATGCGCTGCGAATCGGGAGCGGCGATACCGTAAAGCACGAGG
PP4-2-pSF_A_pSFR      AGAAGGCGATAGAAGGCGATGCGCTGCGAATCGGGAGCGGCGATACCGTAAAGCACGAGG
                      ************************************************************

pGWB2                 AAGCGGTCAGCCCATTCGCCGCCAAGCTCTTCAGCAATATCACGGGTAGCCAACGCTATG
PP4-2-pSF_A_pSFR      AAGCGGTCAGCCCATTCGCCGCCAAGCTCTTCAGCAATATCACGGGTAGCCAACGCTATG
                      ************************************************************

pGWB2                 TCCTGATAGCGGTCCGCCACACCCAGCCGGCCACAGTCGATGAATCCAGAAAAGCGGCCA
PP4-2-pSF_A_pSFR      TCCTGATAGCGGTCCGCCACACCCAGCCGGCCACAGTCGATGAATCCAGAAAAGCGGCCA
                      ************************************************************

pGWB2                 TTTTCCACCATGATATTCGGCAAGCAGGCATCGCCATGGGTCACGACGAGATCATCGCCG
PP4-2-pSF_A_pSFR      TTTTCCACCATGATATTCGGCAAGCAGGCATCGCCATGGGTCACGACGAGATCATCGCCG
                      ************************************************************

pGWB2                 TCGGGCATGCGCGCCTTGAGCCTGGCGAACAGTTCGGCTGGCGCGAGCCCCTGATGCTCT
PP4-2-pSF_A_pSFR      TCGGGCATGCGCGCCTTGAGCCTGGCGAACAGTTCGGCTGGCGCGAGCCCCTGATGCTCT
                      ************************************************************

pGWB2                 TCGTCCAGATCATCCTGATCGACAAGACCGGCTTCCATCCGAGTACGTGCTCGCTCGATG
PP4-2-pSF_A_pSFR      TCGTCCAGATCATCCTGATCGACAAGACCGGCTTCCATCCGAGTACGTGCTCGCTCGATG
                      ************************************************************

pGWB2                 CGATGTTTCGCTTGGTGGTCGAATGGGCAGGTAGCCGGATCAAGCGTATGCAGCCGCCGC
PP4-2-pSF_A_pSFR      CGATGTTTCGCTTGGTGGTCGAATGGGCAGGTAGCCGGATCAAGCGTATGCAGCCGCCGC
                      ************************************************************

pGWB2                 ATTGCATCAGCCATGATGGATACTTTCTCGGCAGGAGCAAGGTGAGATGACAGGAGATCC
PP4-2-pSF_A_pSFR      ATTGCATCAGCCATGATGGATACTTTCTCGGCAGGAKCAAGGTGAGATGACAGGAGATCC
                      ************************************ ************************

pGWB2                 TGCCCCGGCACTTCGCCCAATAGCAGCCAGTCCCTTCCCGCTTCAGTGACAACGTCGAGC
PP4-2-pSF_A_pSFR      TGCCCCGGCACTTCGCCCAATAGCAGCCAGTCCCTTCCCGCTTCAGTGACAACGTCGAGC
                      ************************************************************

pGWB2                 ACAGCTGCGCAAGGAACGCCCGTCGTGGCCAGCCACGATAGCCGCGCTGCCTCGTCCTGC
PP4-2-pSF_A_pSFR      ACAGCTGCGCAAGGAACGCCCGTCGTGGCCAGCCACGATAGCCGCGCTGCCTCGTCCTGC
                      ************************************************************

pGWB2                 AGTTCATTCAGGGCACCGGACAGGTCGGTCTTGACAAAAAGAACCGGGCGCCCCTGCGCT
PP4-2-pSF_A_pSFR      AGTTCATTCRGGGCACCGG----------------------------------------
                      ********* *********
```

A-21: Multiple sequence alignment of PP4-1 sequence and PP4-1-pSF A3 sequence, star region missing and mutation before the stop codon

```
PP4-1                  ATGTCAGACCTAGATCGGCAAATAGGGCAGCTTAAGCGATGCGAACCATTGAGCGAATCG
PP4-1-pSFA3_TOPOF      ----------------------------------------------ATTGAGCGAATCG
                                                                     *************
```

```
PP4-1                 GAGGTGAAGGCTCTTTGCCTCAAAGCCATGGAAATTCTTGTTGAAGAGAGTAATGTTCAG
PP4-1-pSFA3_TOPOF     GAGGTGAAGGCTCTTTGCCTCAAAGCCATGGAAATTCTTGTTGAAGAGAGTAATGTTCAG
                      ************************************************************

PP4-1                 AGAGTTGATGCCCCTGTCACTTTATGTGGTGACATCCATGGGCAGTTCTATGATATGATG
PP4-1-pSFA3_TOPOF     AGAGTTGATGCCCCTGTCACTTTATGTGGTGACATCCATGGGCAGTTCTATGATATGATG
                      ************************************************************

PP4-1                 GAGCTTTTCAAAGTTGGGGGTGATTGTCCTAAGACCAACTATTTGTTTATGGGAGATTTT
PP4-1-pSFA3_TOPOF     GAGCTTTTCAAAGTTGGGGGTGATTGTCCTAAGACCAACTATTTGTTTATGGGAGATTTT
                      ************************************************************

PP4-1                 GTTGATCGTGGATATTATTCGGTTGAGACATTTCTACTTCTACTCGCACTCAAGGTTAGA
PP4-1-pSFA3_TOPOF     GTTGATCGTGGATATTATTCGGTTGAGACATTTCTACTTCTACTCGCACTCAAGGTTAGA
                      ************************************************************

PP4-1                 TATCCAGACCGCATAACTCTCATCAGAGGAAACCATGAAAGCAGGCAAATCACACAGGTT
PP4-1-pSFA3_TOPOF     TATCCAGACCGCATAACTCTCATCAGAGGAAACCATGAAAGCAGGCAAATCACACAGGTT
                      ************************************************************

PP4-1                 TATGGATTTTATGATGAGTGTTTGCGTAAATATGGCTCTTCAAATGTCTGGAGATACTGC
PP4-1-pSFA3_TOPOF     TATGGATTTTATGATGAGTGTTTGCGTAAATATGGCTCTTCAAATGTCTGGAGATACTGC
                      ************************************************************

PP4-1                 ACCGACATTTTTGACTACATGAGTCTTTCAGCTGTTGTGGAGAACAAGATATTCTGTGTT
PP4-1-pSFA3_TOPOF     ACCGACATTTTTGACTACATGAGTCTTTCAGCTGTTGTGGAGAACAAGATATTCTGTGTT
                      ************************************************************

PP4-1                 CATGGTGGTCTTTCTCCAGCTATTATGACTCTTGATCAGATTAGGACAATTGACCGGAAG
PP4-1-pSFA3_TOPOF     CATGGTGGTCTTTCTCCAGCTATTATGACTCTTGATCAGATTAGGACAATTGACCGGAAG
                      ************************************************************

PP4-1                 CAAGAAGTACCACATGATGGTGCCATGTGTGATCTCCTATGGTCTGATCCTGAAGATATT
PP4-1-pSFA3_TOPOF     CAAGAAGTACCACATGATGGTGCCATGTGTGATCTCCTATGGTCTGATCCTGAAGATATT
                      ************************************************************

PP4-1                 GTTGATGGCTGGGGATTGAGCCCTCGTGGTGCCGGATTCCTTTTTGGTGGCAGTGTTGTC
PP4-1-pSFA3_TOPOF     GTTGATGGCTGGGGATTGAGCCCTCGTGGTGCCGGATTCCTTTTTGGTGGCAGTGTTGTC
                      ************************************************************

PP4-1                 ACGTCTTTTAACCACTCAAACAACATAGACTACATAGCCCGTGCCCATCAACTAGTTATG
PP4-1-pSFA3_TOPOF     ACGTCTTTTAACCACTCAAACAACATAGACTACATAGCCCGTGCCCATCAACTAGTTATG
                      ************************************************************

PP4-1                 GAGGGTTACAAATGGATGTTTGATAGCCAGATTGTGACAGTGTGGTCAGCCCCAAATTAC
PP4-1-pSFA3_TOPOF     GAGGGTTACAAATGGATGTTTGATAGCCAGATTGTGACAGTGTGGTCAGCCCCAAATTAC
                      ************************************************************

PP4-1                 TGTTACAGATGCGGTAATGTGGCTTCAATTCTAGAGCTTGACGAGAATCTAAATAAAGAA
PP4-1-pSFA3_TOPOF     TGTTACAGATGCGGTAATGTGGCTTCAATTCTAGAGCTTGACGAGAATCTAAATAAAGAA
                      ************************************************************

PP4-1                 TTCCGTGTGTTTGATGCAGCCCAGCAGGACTCGAGAGGGCCTCCCGCCAAAAAGCCGGCC
PP4-1-pSFA3_TOPOF     TTCCGTGTGTTTGATGCAGCCCAGCAGGACTCGAGAGGGCCTCCCGCCAAAAAGCCGGCC
                      ************************************************************

PP4-1                 CCTGATTACTTCCTATAA------------------------------------------
PP4-1-pSFA3_TOPOF     CCTGATTACTTC-TATAAAARGGGTGGGCGCGCCGACCCAGCTTTCTGTWCCAAAGTTGG
                      ************ *****

PP4-1                 ------------------------------------------------------------
PP4-1-pSFA3_TOPOF     CATTAWAAGAAAGCATTGCTTATCMATTTGTTGCACGAANNNGGTYACTWTCAGTCAAAA


PP4-1                 -------------
PP4-1-pSFA3_TOPOF     TTAAATCATTATT
```

A-22: Multiple sequence alignment of PP4-1 sequence and PP4-1-pSF A1 with PP4.1_MBP R

```
PP4-1                   ------------------------------------------------------------
PP4-1_pSF-A1_MBPR       RGAACACGGGGGACTCTAGTAACGGCCGCCCAGTGTGCTGGAATTCGCCGCCCACCATGG


PP4-1                   ------------------------------------------------------------
PP4-1_pSF-A1_MBPR       ATTATAAAGATGATGATGATAAAGGGTCGGCCGCCAGCTGGAGCCACCCTCAGTTCGAGA


PP4-1                   ------------------------------------------------------------
PP4-1_pSF-A1_MBPR       AGGGAGGAGGAAGCGGCGGAGGCAGCGGAGGAGGAAGCTGGAGCCACCCGCAGTTCGAGA


PP4-1                   ------------------------------------------------------------A
PP4-1_pSF-A1_MBPR       AAGGAGCTAGATCAACAAGTTTGTACAAAAAAGCAGGCTCCGCGGCCGCCCCCTTCACCA
                                                                                    *

PP4-1                   TGTCAGACCTAGATCGGCAAATAGGGCAGCTTAAGCGATGCGAACCATTGAGCGAATCGG
PP4-1_pSF-A1_MBPR       TGTCAGACCTAGATCGGCAAATAGGGCAGCTTAAGCGATGCGAACCATTGAGCGAATCGG
                        ************************************************************

PP4-1                   AGGTGAAGGCTCTTTGCCTCAAAGCCATGGAAATTCTTGTTGAAGAGAGTAATGTTCAGA
PP4-1_pSF-A1_MBPR       AGGTGAAGGCTCTTTGCCTCAAAGCCATGGAAATTCTTGTTGAAGAGAGTAATGTTCAGA
                        ************************************************************

PP4-1                   GAGTTGATGCCCCTGTCACTTTATGTGGTGACATCCATGGGCAGTTCTATGATATGATGG
PP4-1_pSF-A1_MBPR       GAGTTGATGCCCCTGTCACTTTATGTGGTGACATCCATGGGCAGTTCTATGATATGATGG
                        ************************************************************

PP4-1                   AGCTTTTCAAAGTTGGGGGTGATTGTCCTAAGACCAACTATTTGTTTATGGGAGATTTTG
PP4-1_pSF-A1_MBPR       AGCTTTTCAAAGTTGGGGGTGATTGTCCTAAGACCAACTATTTGTTTATGGGAGATTTTG
                        ************************************************************

PP4-1                   TTGATCGTGGATATTATTCGGTTGAGACATTTCTACTTCTACTCGCACTCAAGGTTAGAT
PP4-1_pSF-A1_MBPR       TTGATCGTGGATATTATTCGGTTGAGACATTTCTACTTCTACTCGCACTCAAGGTTAGAT
                        ************************************************************

PP4-1                   ATCCAGACCGCATAACTCTCATCAGAGGAAACCATGAAAGCAGGCAAATCACACAGGTTT
PP4-1_pSF-A1_MBPR       ATCCAGACCGCATAACTCTCATCAGAGGAAACCATGAAAGCAGGCAAATCACACAGGTTT
                        ************************************************************

PP4-1                   ATGGATTTTATGATGAGTGTTTGCGTAAATATGGCTCTTCAAATGTCTGGAGATACTGCA
PP4-1_pSF-A1_MBPR       ATGGATTTTATGATGAGTGTTTGCGTAAATATGGCTCTTCAAATGTCTGGAGATACTGCA
                        ************************************************************

PP4-1                   CCGACATTTTTGACTACATGAGTCTTTCAGCTGTTGTGGAGAACAAGATATTCTGTGTTC
PP4-1_pSF-A1_MBPR       CCGACATTTTTGACTACATGAGTCTTTCAGCTGTTGTGGAGAACAAGATATTCTGTGTTC
                        ************************************************************

PP4-1                   ATGGTGGTCTTTCTCCAGCTATTATGACTCTTGATCAGATTAGGACAATTGACCGGAAGC
PP4-1_pSF-A1_MBPR       ATGGTGGTCTTTCTCCAGCTATTATGACTCTTGATCAGATTAGGACAATTGACCGGAAGC
                        ************************************************************

PP4-1                   AAGAAGTACCACATGATGGTGCCATGTGTGATCTCCTATGGTCTGATCCTGAAGATATTG
PP4-1_pSF-A1_MBPR       AAGAAGTACCACATGATGGTGCCATGTGTGATCTCCTATGGTCTGATCCTGAAGATATTG
                        ************************************************************

PP4-1                   TTGATGGCTGGGGATTGAGCCCTCGTGGTGCCGGATTCCTTTTTGGTGGCAGTGTTGTCA
PP4-1_pSF-A1_MBPR       TTGATGGCTGGGGATTGAGCCCTCGTGGTGCCGGATTCCTTTTTGGTGGCAGTGTTGTCA
                        ************************************************************

PP4-1                   CGTCTTTTAACCACTCAAACAACATAGACTACATAGCCCGTGCCCATCAACTAGTTATGG
PP4-1_pSF-A1_MBPR       CGTCTTTTAACCACTCAAACAACATAGACTACATAGCCCGTGCCCATCAACTAGTTATGG
                        ************************************************************

PP4-1                   AGGGTTACAAATGGATGTTTGATAGCCAGATTGTGACAGTGTGGTCAGCCCCAAATTACT
```

```
PP4-1_pSF-A1_MBPR      AGGGTTACAAATGGATGTTTGATAGCCAGATTGTGACAGTGTGGTCAGCCCCAAATTACT
                       ************************************************************

PP4-1                  GTTACAGATGCGGTAATGTGGCTTCAATTCTAGAGCTTGACGAGAATCTAAATAAAGAAT
PP4-1_pSF-A1_MBPR      GTTACAGATGCGGTAATGTGGCTTCAATTCTAGAGCTTGACGAGAATCTAAATAAAGAAT
                       ************************************************************

PP4-1                  TCCGTGTGTTTGATGCAGCCCAGCAGGACTCGAGAGGGCCTCCCGCCAAAAAGCCGGCCC
PP4-1_pSF-A1_MBPR      TCCGTGTGTTTGATGCAGCCCAGCAGGACTCGAGNG-----------------------
                       ********************************* *

PP4-1                  CTGATTACTTCCTATAA
PP4-1_pSF-A1_MBPR      -----------------
```

A-23: Multiple sequence alignment of PP4-1 sequence and PP4-1-pSF A1 with PP4.1_dTOPO R

```
PP4-1                  ------------------------------------------------------------
PP4-1_pSF-A1_T_R       CACCCGCAGTTCGAGAAAGGAGCTAGATCAACAAGTTTGTACAAAAAAGCAGGCTCCGCG


PP4-1                  ----------------ATGTCAGACCTAGATCGGCAAATAGGGCAGCTTAAGCGATGCGA
PP4-1_pSF-A1_T_R       GNCCGCCCCCTTCACCATGTCAGACCTAGATCGGCAAATAGGGCAGCTTAAGCGATGCGA
                                       ****************************************

PP4-1                  ACCATTGAGCGAATCGGAGGTGAAGGCTCTTTGCCTCAAAGCCATGGAAATTCTTGTTGA
PP4-1_pSF-A1_T_R       ACCATTGAGCGAATCGGAGGTGAAGGCTCTTTGCCTCAAAGCCATGGAAATTCTTGTTGA
                       ************************************************************

PP4-1                  AGAGAGTAATGTTCAGAGAGTTGATGCCCCTGTCACTTT-ATGTGGTGACATCCATGGGC
PP4-1_pSF-A1_T_R       AGAGAGTAATGTTCAGAGAGTTGATGCCCCTGTCACTTTTATGTGGTGACATCCATGGGC
                       ************************************* ********************

PP4-1                  AGTTCTATGATATGATGGAGCTTTTCAAAGTTGGGGGTGATTGTCCTAAGACCAACTATT
PP4-1_pSF-A1_T_R       AGTTCTATGATATGATGGAGCTTTTCAAAGTTGGGGGTGATTGTCCTAAGACCAACTATT
                       ************************************************************

PP4-1                  TGTTTATGGGAGATTTTGTTGATCGTGGATATTATTCGGTTGAGACATTTCTACTTCTAC
PP4-1_pSF-A1_T_R       TGTTTATGGGAGATTTTGTTGATCGTGGATATTATTCGGTTGAGACATTTCTACTTCTAC
                       ************************************************************

PP4-1                  TCGCACTCAAGGTTAGATATCCAGACCGCATAACTCTCATCAGAGGAAACCATGAAAGCA
PP4-1_pSF-A1_T_R       TCGCACTCAAGGTTAGATATCCAGACCGCATAACTCTCATCAGAGGAAACCATGAAAGCA
                       ************************************************************

PP4-1                  GGCAAATCACACAGGTTTATGGATTTTATGATGAGTGTTTGCGTAAATATGGCTCTTCAA
PP4-1_pSF-A1_T_R       GGCAAATCACACAGGTTTATGGATTTTATGATGAGTGTTTGCGTAAATATGGCTCTTCAA
                       ************************************************************

PP4-1                  ATGTCTGGAGATACTGCACCGACATTTTTGACTACATGAGTCTTTCAGCTGTTGTGGAGA
PP4-1_pSF-A1_T_R       ATGTCTGGAGATACTGCACCGACATTTTTGACTACATGAGTCTTTCAGCTGTTGTGGAGA
                       ************************************************************

PP4-1                  ACAAGATATTCTGTGTTCATGGTGGTCTTTCTCCAGCTATTATGACTCTTGATCAGATTA
PP4-1_pSF-A1_T_R       ACAAGATATTCTGTGTTCATGGTGGTCTTTCTCCAGCTATTATGACTCTTGATCAGATTA
                       ************************************************************

PP4-1                  GGACAATTGACCGGAAGCAAGAAGTACCACATGATGGTGCCATGTGTGATCTCCTATGGT
PP4-1_pSF-A1_T_R       GGACAATTGACCGGAAGCAAGAAGTACCACATGATGGTGCCATGTGTGATCTCCTATGGT
                       ************************************************************

PP4-1                  CTGATCCTGAAGATATTGTTGATGGCTGGGGATTGAGCCCTCGTGGTGCCGGATTCCTTT
PP4-1_pSF-A1_T_R       CTGATCCTGAAGATATTGTTGATGGCTGGGGATTGAGCCCTCGTGGTGCCGGATTCCTTT
                       ************************************************************

PP4-1                  TTGGTGGCAGTGTTGTCACGTCTTTTAACCACTCAAACAACATAGACTACATAGCCCGTG
PP4-1_pSF-A1_T_R       TTGGTGGCAGTGTTGTCACGTCTTTTAACCACTCAAACAACATAGACTACATAGCCCGTG
```

```
                    ************************************************************

PP4-1               CCCATCAACTAGTTATGGAGGGTTACAAATGGATGTTTGATAGCCAGATTGTGACAGTGT
PP4-1_pSF-A1_T_R    CCCATCAACTAGTTATGGAGGGTTACAAATGGATGTTTGATAGCCAGATTGTGACAGTGT
                    ************************************************************

PP4-1               GGTCAGCCCCAAATTACTGTTACAGATGCGGTAATGTGGCTTCAATTCTAGAGCTTGACG
PP4-1_pSF-A1_T_R    GGTCAGCCCCAAATTACTGTTACAGATGCGGTAATGTGGCTTCAATTCTAGAGCTTGACG
                    ************************************************************

PP4-1               AGAATCTAAATAAAGAATTCCGTGTGTTTGATGCAGCCCAGCAGGACTCGAGAGGGCCTC
PP4-1_pSF-A1_T_R    AGAATCTAAATAAAGAATTCCGTGTGTTTGATGCAGCCCAGCAG-ACT------------
                    ******************************************** ***

PP4-1               CCGCCAAAAAGCCGGCCCCTGATTACTTCCTATAA
PP4-1_pSF-A1_T_R    -----------------------------------
```

A-24: Multiple sequence alignment of PP4-1 sequence and PP4-1-pSF A1 with PP4.1_dTOPO F

```
PP4-1               ATGTCAGACCTAGATCGGCAAATAGGGCAGCTTAAGCGATGCGAACCATTGAGCGAATCG
PP4-1_pSF-A1_T_F    ------------------------------------------------------------


PP4-1               GAGGTGAAGGCTCTTTGCCTCAAAGCCATGGAAATTCTTGTTGAAGAGAGTAATGTTCAG
PP4-1_pSF-A1_T_F    --GGTGAAGGCTCTTTGCCTCAAAGCCATGGAAATTCTTGTYGAAGAGAGTAATGTTCAG
                      ************************************ ******************

PP4-1               AGAGTTGATGCCCCTGTCACTTTATGTGGTGACATCCATGGGCAGTTCTATGATATGATG
PP4-1_pSF-A1_T_F    AGAGTTGATGCCCCTGTCACTTTATGTGGTGACATCCATGGGCAGTTCTATGATATGATG
                    ************************************************************

PP4-1               GAGCTTTTCAAAGTTGGGGGTGATTGTCCTAAGACCAACTATTTGTTTATGGGAGATTTT
PP4-1_pSF-A1_T_F    GAGCTTTTCAAAGTTGGGGGTGATTGTCCTAAGACCAACTATTTGTTTATGGGAGATTTT
                    ************************************************************

PP4-1               GTTGATCGTGGATATTATTCGGTTGAGACATTTCTACTTCTACTCGCACTCAAGGTTAGA
PP4-1_pSF-A1_T_F    GTTGATCGTGGATATTATTCGGTTGAGACATTTCTACTTCTACTCGCACTCAAGGTTAGA
                    ************************************************************

PP4-1               TATCCAGACCGCATAACTCTCATCAGAGGAAACCATGAAAGCAGGCAAATCACACAGGTT
PP4-1_pSF-A1_T_F    TATCCAGACCGCATAACTCTCATCAGAGGAAACCATGAAAGCAGGCAAATCACACAGGTT
                    ************************************************************

PP4-1               TATGGATTTTATGATGAGTGTTTGCGTAAATATGGCTCTTCAAATGTCTGGAGATACTGC
PP4-1_pSF-A1_T_F    TATGGATTTTATGATGAGTGTTTGCGTAAATATGGCTCTTCAAATGTCTGGAGATACTGC
                    ************************************************************

PP4-1               ACCGACATTTTTGACTACATGAGTCTTTCAGCTGTTGTGGAGAACAAGATATTCTGTGTT
PP4-1_pSF-A1_T_F    ACCGACATTTTTGACTACATGAGTCTTTCAGCTGTTGTGGAGAACAAGATATTCTGTGTT
                    ************************************************************

PP4-1               CATGGTGGTCTTTCTCCAGCTATTATGACTCTTGATCAGATTAGGACAATTGACCGGAAG
PP4-1_pSF-A1_T_F    CATGGTGGTCTTTCTCCAGCTATTATGACTCTTGATCAGATTAGGACAATTGACCGGAAG
                    ************************************************************

PP4-1               CAAGAAGTACCACATGATGGTGCCATGTGTGATCTCCTATGGTCTGATCCTGAAGATATT
PP4-1_pSF-A1_T_F    CAAGAAGTACCACATGATGGTGCCATGTGTGATCTCCTATGGTCTGATCCTGAAGATATT
                    ************************************************************

PP4-1               GTTGATGGCTGGGGATTGAGCCCTCGTGGTGCCGGATTCCTTTTTGGTGGCAGTGTTGTC
PP4-1_pSF-A1_T_F    GTTGATGGCTGGGGATTGAGCCCTCGTGGTGCCGGATTCCTTTTTGGTGGCAGTGTTGTC
                    ************************************************************

PP4-1               ACGTCTTTTAACCACTCAAACAACATAGACTACATAGCCCGTGCCCATCAACTAGTTATG
PP4-1_pSF-A1_T_F    ACGTCTTTTAACCACTCAAACAACATAGACTACATAGCCCGTGCCCATCAACTAGTTATG
                    ************************************************************
```

```
PP4-1                GAGGGTTACAAATGGATGTTTGATAGCCAGATTGTGACAGTGTGGTCAGCCCCAAATTAC
PP4-1_pSF-A1_T_F     GAGGGTTACAAATGGATGTTTGATAGCCAGATTGTGACAGTGTGGTCAGCCCCAAATTAC
                     ************************************************************

PP4-1                TGTTACAGATGCGGTAATGTGGCTTCAATTCTAGAGCTTGACGAGAATCTAAATAAAGAA
PP4-1_pSF-A1_T_F     TGTTACAGATGCGGTAATGTGGCTTCAATTCTAGAGCTTGACGAGAATCTAAATAAAGAA
                     ************************************************************

PP4-1                TTCCGTGTGTTTGATGCAGCCCAGCAGGACTCGAGAGGGCCTCCCGCCAAAAAGCCGGCC
PP4-1_pSF-A1_T_F     TTCCGTGTGTTTGATGCAGCCCAGCAGGACTCGAGAGGGCCTCCCGCCAAAAAGCCGGCC
                     ************************************************************

PP4-1                CCTGATTACTTCCTATAA------------------------------------------
PP4-1_pSF-A1_T_F     CCTGATTACTTCCTATAAAAGGGTGGGCGCGCCGACCCAGCTTTCTTGTACAAAGTGGTT
                     ******************

PP4-1                ------------------------------------------------------
PP4-1_pSF-A1_T_F     GATAACAGCGCTTAGAGCTCGAATTTCCCCGATCGTTCAAAACATTTGGCNNNN
```