CONSTRUCTION OF *NptII-^{mt}sfGFP* FUSION GENE CONSTRUCT BY OVERLAP EXTENSION PCR

By

CHENG YANG YAN

A project report submitted to the Department of

Biological Science

Faculty of Science

Universiti Tunku Abdul Rahman

In partial fulfilment of the requirements for the degree of

Bachelor of Science (Honours) Biotechnology

OCTOBER 2025

Abstract

CONSTRUCTION OF *Npt*II-*mtsfGFP* FUSION GENE CONSTRUCT BY OVERLAP EXTENSION PCR

CHENG YANG YAN

In plant transformation, the *Agrobacterium*-mediated system (AMT) is widely used due to deliver transgenes into multi-layered plant cells without causing damage. Previously, the Gateway^{TM}-compatible binary vectors, pG103 and pG104 were built to enable convenient transgene insertion. For the selection of transformed cells, plasmid vectors pG103 and pG104 harbor the selectable markers *neomycin phosphotransferase* II (*Npt*II) and mitochondrial recoded *neomycin phosphotransferase* II (*^{mt}Npt*II*)* genes, respectively, which confer resistance to aminoglycoside antibiotics for nuclear and mitochondrial transformation. To facilitate visualization of the protein expression of the gene of interest (GOI), a reporter gene is often translationally fused it. The green fluorescent protein (*GFP*) gene is widely used as a reporter gene. This project used mitochondrial recoded superfolder GFP (*^{mt}sfGFP)* gene*, which encode the same superfolder GFP protein, is designed for expression in either the host's nucleus and mitochondria. This project aimed to fuse the *Npt*II and *^{mt}Npt*II with *^{mt}sfGFP* using overlap extension PCR (OE-PCR). A 3-step OE-PCR approach using megaprimers was employed. However, due to the formation of secondary structures of the megaprimer during the PCR reaction, non-specific amplicons

were generated. To eliminate these non-specific amplicons and obtain the desired amplicons, parameters such as primer and template DNA concentrations, number of thermal cycles, and annealing temperature, were performed for each step of the OE-PCR. Consequently, amplicons carrying the *Npt*II- $^{mt}sfGFP$ and ^{mt}NptII- $^{mt}sfGFP$ fusion genes were successfully generated. In the future, these amplicons will be cloned into the *Agrobacterium* binary vectors for host transformation. They will facilitate qualitative and quantitative analysis of transgenes in the host's nucleus and mitochondria.

ACKNOWLEDGEMENTS

First, I would like to express my special thanks of gratitude to my supervisor, Prof. Dr. Wong Hann Ling. Without his support, priceless recommendations, and guidance throughout my final year project, I would not have been able to complete it. Prof. Wong is professional, patient, and kind; he always helped me solve problems whenever I faced difficulties. I am also grateful to my postgraduate seniors, Mrs. Yap Kiao Huio and Mr. Jordan Ooi, as well as my PhD senior, Mr. Toh Wai Keat, for everything. They assisted me whenever I encountered challenges in my project, discussed solutions with me, and supported me during my benchwork. I would like to thank my parents, sister, friends, Lim Wei, and Jeon Won Woo. Lastly, I want to thank my benchmates, Victoria Law and Lam Yee Ling.

DECLARATION

I hereby declare that the project report is based on my original work except for quotation and citations, which have been duly acknowledged. I also declare that it has not been previously or concurrently submitted for any other degree at Universiti Tunku Abdul Rahman (UTAR) or other institutions.

<div style="text-align:right">
_____
</div>

<div style="text-align:right">
Student's signature
</div>

TABLE OF CONTENT

LIST OF TABLES

LIST OF FIGURES

LIST OF ABBREVIATIONS

AMT	*Agrobacterium*-mediated system
A. tumefaciens	*Agrobacterium tumefaciens*
bp	Base pair
^{o}C	Degree Celsius
*COX*1	Cytochrome c oxidase subunit 1
*COX*2	Cytochrome c oxidase subunit 2
dH$_2$O	Distilled water
DMSO	Dimethyl sulfoxide
E. coli	Escherichia coli
EDTA	Ethylenediaminetetraacetic acid
EtOH	Ethyl alcohol
$\times g$	Times gravity
G418	Geneticin
GFP	Green fluorescent protein
Gly67	Glycine in 67th position
GOI	Gene of interest
h	Hour
HR	Homologous recombination
kb	Kilobase pair
kDNA	Kinetoplast DNA
L	Litre
min	Minute
mL	Millilitre
μg	Microgram

μL	Microlitre
μM	Micromolar
mtDNA	Mitochondrial genome
^{mt}NptII	Mitochondrial recoded *neomycin phosphotransferase* II
MTS	Mitochondrial targeting sequences
^{mt}sfGFP	Mitochondrial recoded superfolder GFP
^{mt}sfGFP(*^{mt}Npt*II)	Mitochondrial recoded superfolder GFP overlapping with Mitochondrial recoded *neomycin phosphotransferase* II
^{mt}sfGFP(*Npt*II)	Mitochondrial recoded superfolder GFP overlapping with *neomycin phosphotransferase* II
NaOH	Sodium Hydroxide
NaOAc	Sodium acetate
nDNA	Nuclear genome
Neo	Neomycin
nm	Nanometers
*Npt*II	*Neomycin phosphotransferase* II
OE-PCR	Overlap extension PCR
PCI	Phenol-chloroform-isoamyl alcohol
PCR	Polymerase Chain Reaction
pDEST	Destination vector
pDONR	Donor vector

pENTR	Entry plasmid
pEXPR	Expression vector
rRNA	Ribosomal RNA
rpm	Revolutions per minute
s	Second
Ser65	Serine in 65th position
sfGFP	Superfolder green fluorescent protein
TBE	Tris/Borate/EDTA
T-DNA	Transferred DNA
*TEF*1	*Translation elongation factor 1α*
Ti	Tumor-inducing
T$_m$	Melting temperature
tRNAs	Transfer ribonucleic acid
Tyr66	Tyrosine in 66$_{th}$ position
UV	Ultraviolet
V	Volt
vir	Virulence

CHAPTER 1

INTRODUCTION

1.1 Plant transformation

In the field of plant biotechnology, plant transformation is a fundamental tool, allowing the introduction of foreign genes into plant genomes. This technique contributes to the agricultural field by improving plant traits, such as productivity, resistance to trash environments, and increasing nutritional value. To generate the transformed plant, the foreign gene is delivered, integrated, and expressed into the plant cells through different methods, such as *Agrobacterium*-mediated transformation, biolistics, electroporation, and microinjection. (Kumar and Ling, 2021; Law, Miyamoto and Numata, 2023). *Agrobacterium*-mediated transformation is the most common method for plant genetic engineering because it efficiently delivers DNA through the multilayered plant cell wall without causing significant damage. However, the efficiency of forming transformed cells is low, even in the most successful transfer system (Hwang, Yu and Lai, 2017). Therefore, the selective methods are vital to identify and select the transformed cell in the plant transformation process. One of the most significant selective methods to distinguish transformed from non-transformed cells under selective conditions is the selectable marker.

1.2 *Neomycin phosphotransferase* **II (*Npt*II) and mitochondrial *neomycin phosphotransferase* II (*^{mt}Npt*II): Selectable markers**

Selectable marker genes are essential tools for plant transformation and are typically involved in the same DNA construct as the gene of interest. Most commonly, these genes make transformed cells resistant to antibiotics and survive when exposed to a selective medium. The *Npt*II gene serves as an antibiotic selectable marker, and is usually inserted into plasmid vectors used for genetic transformation. *Neomycin phosphotransferase* II (*Npt*II), also known as *aminoglycoside* 3'-*phosphotransferase* II, is a type of transferase isolated from transposon Tn5 found in *Escherichia coli* (*E. coli*), which encodes *Npt*II (E.C. 2.7.1.95). *Npt*II can phosphorylate aminoglycoside antibiotics to inactivate their function. Under the selection medium, the transformed colonies can survive and grow due to resistance to antibiotics (Das et al., 2020).

The *^{mt}Npt*II gene is a modified version of the *Npt*II gene, recoded according to the yeast mitochondrial genetic code. When expressed from the yeast mitochondrial genome, it produces the same protein as the ordinary *Npt*II gene and confers resistance to aminoglycoside antibiotics, just like the nuclear-encoded version (Yap, 2025).

1.3 Mitochondrial superfolder green fluorescent protein (*^{mt}sfGFP*): A reliable reporter gene

In plant transformation, the reporter gene is co-introduced with the gene of interest to provide a measurable or visual signal that proves the successful gene transfer and expression in the plant cell. The green fluorescent protein (*GFP*) is a widely used reporter gene. *GFP* was originally derived from the jellyfish (*Aequorea victoria*). Subsequently, a similar green fluorescent protein (*GFP*) was discovered in the sea pansies, *Renilla reniformis* (Soboleski, Oaks and Halford, 2005). Unlike most fluorescent proteins, which require an external cofactor for visualization, *GFP* can be detected directly by fluorescence microscopy without an exogenous substrate. It consists of 238 amino acids; and cyclization of three residues (positions 65-67) forms a unique chromophore. When exposed to blue or ultraviolet light, the chromophore will be triggered and emit green fluorescence (Pedelacq and Cabantous, 2019).

However, there are some limitations to traditional *GFP*, such as poor folding efficiency or weaker fluorescence in challenging cellular environments, which affect fluorescence intensity. Hence, superfolder green fluorescent protein (*sfGFP*) was used in this project because of its folding kinetics and stability. *sfGFP* is an engineered variant of *GFP* that contains multiple mutations, has a higher folding yield, and causes brighter and more stable fluorescence. The fluorescence capability of *sfGFP* was stable even after fusion with other proteins or misfolding. These features allow *sfGFP* to function in different

biological applications, such as a reporter for protein localization or protein-protein interaction detection. Mitochondrial superfolder green fluorescent protein ($^{mt}sfGFP$) is an engineered variant of *sfGFP* expressed in the nucleus and mitochondria. A few nucleotide sequences in *sfGFP* were modified to remove the stop codon and enable protein synthesis when integrated with the mitochondrial genome (Andrews *et al.,* 2007; Kumar and Pal, 2016).

1.4 Overlap extension PCR: An effective fusion method

In the late 20th century, fusion protein technology enabled the formation of chimeric proteins that combined functional domains from different sources. By joining two or more genes, the modified protein can be constructed with the desired traits, such as improving protein stability, valuability, or adding detection tags. Overlap extension PCR (OE-PCR) is a widely used fusion gene method that does not rely on restriction enzymes and ligation. It condenses the fusion process into three RCR reactions, allowing seamless junctions compared with the traditional restriction enzyme technique. First, each gene fragment was amplified separately by PCR using primers that added complementary overlapping sequences to the ends of the fragments. The PCR products were purified and added to the same reaction mixture as the calculated ratio. This fusion step forms the fusion DNA molecule through DNA fragments fused with other fragments in the correct orientation by annealing the overlapping regions. After fusion, the end-to-end primers were added to the PCR mixture to amplify

the full-length fusion product. Finally, a single continuous DNA sequence is formed that can encode the fusion protein.

OE-PCR involves the gene of interest at any desired junction, adding linker sequences, or inserting specific mutations or tags using well-designed primers. The high flexibility, cost efficiency, and simplicity of the fusion gene construction make it suitable for this project (Bryksin and Matsumura, 2010; Hashemabadi *et al*., 2025).

1.5 Dual-function plasmid: The fusion of the selectable marker and reporter gene

By fusing the selectable marker and reporter gene, we constructed a dual-function plasmid that combines antibiotic resistance with fluorescent reporting. Dual-function plasmids are useful for genetic engineering and functional studies in a variety of organisms, such as bacteria and plants (Strathdee, McLeod and Underhill, 2000). The main purpose of fusing *Npt*II and [mt]*sfGFP* genes is for simultaneous selection and real-time monitoring of the gene expression in the transformation process. Under the selective medium, only the cells that take up and express the fusion gene will survive antibiotic selection and emit fluorescence. The *Npt*II gene allows resistance to aminoglycoside antibiotics by phosphorylating them, whereas [mt]*sfGFP* can directly visualize living transformed cells by emitting strong green fluorescence under UV or blue light (Das *et al*., 2020; Andrews *et al*., 2007).

By linking two genes, the occurrence of false-positive results is reduced and the reliability of transformation studies is increased. The researchers allow rapid identification of the true transformation, increasing the screening efficiency and saving time and resources. In addition, the dual-function plasmid achieved simultaneous qualitative and quantitative processes through survival in the selectable medium and the gene expression levels were assessed by measuring the intensity of *sfGFP* fluorescence. Based on this criterion, a comparison of the expression efficiency, transformation rates, or promoter strength across samples can be performed (Strathdee, McLeod and Underhill, 2000).

1.6 Objectives of this study

The main objective of this project was to construct a $NptII$-$^{mt}sfGFP$ fusion gene using overlap extension PCR.

The specific objectives of this project are as follows:

1. To fuse the target gene $mtNptII$ in the pENTR-P$_{COX1}$:: $NptII$-T7TER plasmid and $^{mt}sfGFP$ in the pENTR-P$_{COX2}$-$^{mt}sfGFP$ plasmid.

2. To fuse the target gene $NptII$ in the pENTR-P$_{TEF1}$-$NptII$ plasmid and $^{mt}sfGFP$ in the pENTR-P$_{COX2}$-$^{mt}sfGFP$ plasmid.

CHAPTER 2

LITERATURE REVIEW

2.1 *Agrobacterium*-mediated transformation (AMT)

Agrobacterium-mediated transformation (AMT) is a fundamental technique in plant genetic engineering that utilizes the natural ability of the Gram-negative soil-borne phytopathogen *Agrobacterium tumefaciens* (*A. tumefaciens*), allowing the stable or transient insertion of a gene of interest (GOI) into various species. The transferred DNA (T-DNA) is a tumor-inducing (Ti) plasmid that is integrated into the host genome. It is the predominant tool for generating transgenic plants and contributes to the development of functional genomic studies in genetic engineering (Hwang, Yu, and Lai, 2017; De Saeger *et al*., 2021). A significant breakthrough in genetic engineering occurred with the first successful AMT for plant transformation. However, a major challenge was the transfer of native T-DNA, which contains oncogenes, triggered crown gall disease in transformed plants through the formation of tumors. This problem was solved by subsequent research, in which disarmed vectors were engineered by removing these oncogenic genes while retaining the ability to transfer T-DNA. This crucial modification prevented tumorigenesis and

established AMT as an essential and practical tool for plant genetic engineering (Bourras, Rouxel and Meyer, 2015).

Compared with other plant transformation methods, it overcomes this obstacle and is able to pass through the impenetrable multi-layered plant cell, solving the most challenging barrier in plant transformation (Hwang, Yu, and Lai, 2017). Other methods, such as biolistic methods, electroporation, and microinjection, rely on mechanical forces to deliver DNA into plant cells. This method often causes irreparable cell damage and reduces transformation efficiency compared with AMT (Kumar and Ling, 2021).

AMT allows the delivery of large DNA fragments with low transgene copy numbers and minimal rearrangements, resulting in greater genetic stability. It is an artificially prepared Ti plasmid in the transfer DNA, or T-DNA region. Then, the helper plasmid carrying the virulence (*vir*) genes can transfer T-DNA into plant cells (De Saeger *et al*., 2021).

2.1.1 Advances in AMT Methodology

In the beginning stage, native sequences from the T-DNA region, such as oncogenes and opine biosynthetic genes, were removed. The removal process eliminates its pathogenicity, but the ability of transgene delivery remains, forming the "disarmed" *A. tumefaciens* (Barton *et al*., 1983).

In the past, researchers integrated a GOI into the disarmed Ti plasmid of *A. tumefaciens through a* modified *E. coli* plasmid, which serves as a shuttle vector via homologous recombination (HR) in the cointegration vector system. However, this system was shifted to a binary vector system because of its labor and complexity. This requires two cointegrating plasmids in *A. tumefaciens* (Barton *et al*., 1983).

Over the years, researchers have discovered that the process of transferring T-DNA from A. tumefaciens will not be affected after separating the T-DNA region from the Ti plasmid, as long as the T-DNA and the Ti plasmid exist in the same *Agrobacterium* cell. Binary vector systems arose with the development of these studies. Based on the ability of GOI to be cloned into a smaller binary vector with a wide host range, this simplified molecular cloning system has become the standard method for AMT. Various *A. tumefaciens* strains and plasmid vectors have been developed to meet different research requirements (Tzfira and Citovsky, 2008).

2.1.2 Binary vector system

In AMT, the binary vector system acts as a modular plasmid system and is widely used because of its efficiency in delivering transgenes into the target host genomes. It includes two plasmids: a disarmed Ti plasmid containing the virulence (*vir*) genes that are essential for T-DNA transfer, and a separate small

binary vector containing the T-DNA region with the GOI flanked by left and right border sequences (Becker *et al*., 1992).

The engineered binary vector can easily transfer foreign DNA into a variety of hosts, including plant or fungal cells, and is stably maintained by *Agrobacterium*. Selectable markers and multiple cloning sites exist for GOI insertions. The utility of this system was increased by constructing a Gateway-compatible binary vector with integrated Gateway cloning cassettes, such as *att* recombination sites and *ccd*B negative-selectable markers (Munaweera *et al*., 2022).

Traditionally, the insertion of a GOI into a binary vector relies on conventional cloning methods, such as restriction enzymes and ligation, which are time-consuming and laborious. Gateway cloning is a site-specific recombination technology that enables seamless transfer of DNA fragments by recombination at *att* sites between vectors without requiring restriction sites, allowing bypassing of restriction-based cloning methods. The plasmid was introduced into *A. tumefaciens* after the recombinant binary vector was constructed in *E. coli*. Subsequently, *A. tumefaciens* mediates the transfer of the assembled T-DNA region into the host genome and undergoes stable integration and expression (Hwang, Yu and Lai, 2017)

In summary, binary vectors serve as the core T-DNA delivery and expression platform, and at the same time, Gateway cloning serves as an efficient cloning method that simplifies complex gene construction within the system. This

method reduces the cost, time, effort, and errors in cloning for ATM. Therefore, the binary vector system combined with Gateway cloning enhances the speed and flexibility of assembling and delivering transgenes into various hosts through *Agrobacterium* transformation (Karimi, Inzé and Depicker, 2002). This combination provides a powerful platform for research on genetic functions, expression, and subcellular localization to enhance genetic and molecular research in biological systems.

2.1.3 Engineered *Agrobacterium*-mediated transformation

In the AMT study, the destination vector carrying the GOI was expressed in the nucleus of the host cell. Based on my previous studies, an engineered AMT system was constructed that was able to transfer the GOI into the host's cell mitochondria. For transformation, a suitable entry vector capable of carrying the GOI and transforming it into mitochondria was built. Based on previous studies, the promoter of the plasmid was modified to initiate and express GOI, similar to the selectable marker. ^{mt}NptII is the modified version of NptII, which is only able to express inside the mitochondria, not inside the nucleus. This action efficiency eliminates the possibility that the GOI is accidentally expressed inside the nucleus, proving that the desired products are generated from the host's cell mitochondria. The fusion of ^{mt}NptII with the *sfGFP* gene allows the false positive result to be removed through observation of gene expression via growth on the selectable marker and visualization of the desired protein synthesis by fluorescent microscopy (Yap, 2025).

2.2 Gateway cloning

In recent years, the study of plant biological systems has relied on transgenic research by analyzing recombinant genes to a large extent using DNA cloning technologies. Unlike classical restriction-enzyme-based cloning methods, recombination cloning is an efficient and common cloning technology based on site-specific recombination. Without the requirement of a restriction site to insert the DNA sequence of interest, it allows rapid and efficient parallel transfer of the DNA of interest into diverse expression systems. One of the most popular recombination cloning technologies, Gateway cloning, allows the assembly of DNA fragments regardless of their sequence (Curtis and Grossniklaus, 2003).

Gateway cloning systems allow the joining of DNA fragments into a desired orientation and order and maintain their reading frame. This protocol relies on two main steps: a BP clonase reaction to insert the gene of interest (GOI) into an entry vector with an *att*L site. Lastly, the recombination of *att*R sites in the destination vector and *att*L sites in pENTR transfers the GOI into the expression vector via the LR reaction. For gene functional assay studies, a suitable destination vector can be recombined with a compatible entry clone

with an inserted GOI. The resulting expression clones were used for gene function testing, such as transformation into plants (Karimi, Depicker and Hilson, 2007).

2.2.1 Principle of Gateway cloning

First, the GOI, for instance, a PCR product flanked by two *att*B sites, is transferred into a donor vector (pDONR) by the BP *Clonase* II enzyme. It includes the integration host factor and phage integrase, which catalyze the BP reaction. The pDONR that carries two *att*P sites will recombine with *att*B sites on the GOI, allowing the GOI gene to be inserted into the donor backbone and form an entry plasmid (pENTR). pENTR is flanked by *att*L sites and is a key substrate in LR reactions. LR reaction is catalyzed by LR *Clonase* II enzymes, including phage excision, integration host factor, and integrase. It can transfer a GOI into an expression vector (pEXPR), which is flanked by *att*B sites, by recombination of *att*L sites on the entry vector and the destination vector (pDEST) carrying two *att*R sites. The researchers used pDONR, pENTR, pDEST, and pEXPR to differentiate the input and output plasmids in the cloning process (Curtis and Grossniklaus, 2003; Depicker and Hilson, 2007).

To maintain the open reading frame of the GOI, the engineered Gateway system alters the original *att*B, *att*P, *att*L, and *att*R sites to ensure specificity in

site-specific recombination. For example, *att*B1 reacts only with *att*P1, not *att*P2, to ensure directional cloning in this method. The entry vector is the key substrate in Gateway cloning and is generated by inserting the GOI into the pDONR flanked by *att*L sites. However, *att*L sites are relatively long (96 bp) and can interfere with applications that require minimal sequence addition, such as protein fusion or regulatory element spacing. In contrast, shorter *att*B sites (21-25 bp), specifically designed to lack translation initiation or stop codons, are more suitable for seamless cloning and expression constructs with minimal extra sequences. These sites function in the LR recombination step to generate an expression vector with *att*B sites in order to maintain the open reading frame of the GOI (Depicker and Hilson, 2007).
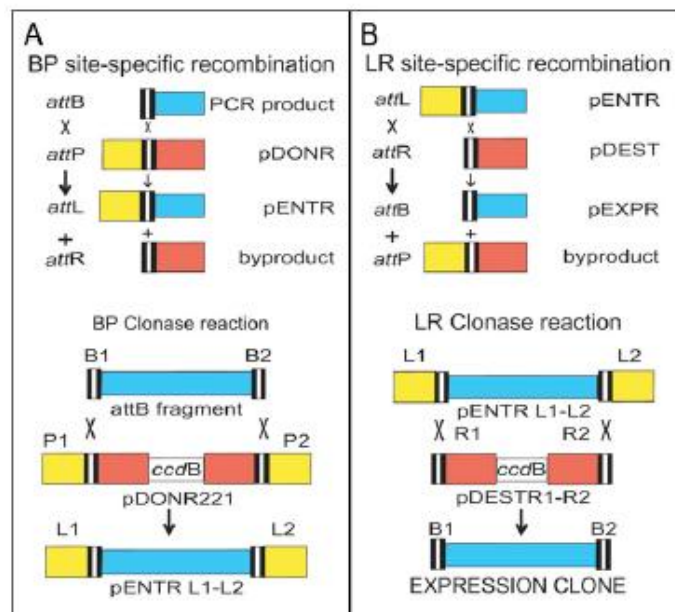
Figure 2.1: Schematic showing the flowchart of the Gateway cloning reaction. In part A, the BP clonase reaction, attB sites recombine with adaption *att*P sites of the donor vector (pDONR). Then, a novel entry vector (pENTR) was generated and flanked by *att*L sites and the attR sides in the byproducts. In part B, the LR Clonase reaction, *att*R sites in the destination vector (pDEST) recombine with matching *att*L sites in the pENTR, generating

a novel expression vector (pEXTR) and *att*P as a byproduct (Depicker and Hilson, 2007).

Typically, ideal plasmids are constructed by in vitro recombination and transformed into *E. coli* strains. Through antibiotic selection and *ccd*B gene counterselection, the plasmid that successfully takes up the GOI can be segregated from the input vector and the reaction byproducts. Importantly, to avoid false-positive results in selection, different types of bacterial antibiotic resistance markers must be included in the plant binary T-DNA destination vector from the donor and entry vectors. Because of the kanamycin resistance that exists in most donor vectors used in the BP reaction, expression clones generated from the LR reaction should carry other selectable markers (Reece-Hoyes and Walhout, 2018).

In fact, ccdB is the toxic byproduct generated by Gateway cloning. It encodes a toxin that affects DNA replication by interfering with the bacterial DNA gyrase. DNA replication is inhibited and causes cell death when *ccd*B is expressed in *E. coli*. In the cloning process, the *ccd*B cassette will be replaced by a GOI through site-specific recombination. Only bacteria that successfully take up the replacement are able to survive and grow under selection conditions. Because ccdB kills the host cells without removing it, the background colonies containing non-recombinant or empty plasmids will be unable to survive (Lund *et,al*, 2014). Through negative selection, the screening burden was reduced and the recovery rate containing the desired recombinant construct increased. However, ccdB is toxic to ordinary *E. coli* strains. Hence, the engineered E.

coli strains that have mutated DNA gyrase are immune to the toxin, maintaining the stability of the vector containing ccdB before recombination. After recombination, plasmids are transformed into standard *E. coli* strains that are sensitive to ccdB toxicity (Reece-Hoyes and Walhout, 2018). Hence, this property is utilized in Gateway cloning of the cloning cassette in both donor and destination vectors within the *ccd*B gene.

2.2.2 Entry vector (pENTR)

Entry vectors (pENTR) are specially designed to serve as initial entrance for GOI in the Gateway cloning workflow. The GOI fragment flanked by *att*B sites recombines with the donor vector flanked by *att*P sites to form the pENTR vector with GOI and the toxic ccdB byproduct. The inserted DNA, flanked by the *att*L1 and *att*L2 recombination sites, can undergo further recombination into different destination vectors through the LR reaction. pENTR is a key component of gateway cloning because it is highly efficient and adaptable for high-throughput applications. Without requiring traditional restriction enzymes and ligation, GOI can easily shuttle between vectors for different downstream applications (Depicker and Hilson, 2007).

2.3 *Neomycin phosphotransferase* II (*Npt*II): A Reliable Selectable Marker

Neomycin phosphotransferase II (*Npt*II), which is isolated from the Tn5 transposon of *E. coli*, encodes an aminoglycoside phosphotransferase that resists aminoglycoside antibiotics such as kanamycin, G418, and neomycin. By phosphorylating aminoglycoside antibiotics, these antibiotics become inactive and lose the function that causes cell death (Das *et al*., 2020). Thus, during the transformation process, the *Npt*II selectable marker is transferred along with the GOI into the target host cell, thereby conferring antibiotic resistance to the transformants under selection (Numata *et al*., 2016; Das *et al*., 2020).

The presence of *Npt*II in the aminoglycoside antibiotic G418 selectable medium affected the elongation step of translation to inhibit protein synthesis. In the transformed cells, the expression of the *Npt*II gene inhibited the cytotoxic effect of G418; only the transformed cells survived. When *Npt*II combines with G418, it specifically binds to the 30S ribosomal subunit in prokaryotes and organelles in eukaryotic cells (Miki and McHugh, 2004).

As a selectable marker, *Npt*II is highly effective in screening transformed cells from non-transformed cells, thereby increasing the transformation selection speed. Even at low antibiotic concentrations, the *Npt*II/G418 system undergoes effective selection, making this antibiotic resistance system highly reproducible and dependable. As the dominant selection system, *Npt*II /G418 provides another option when auxotrophic selection is not available for wild-type and non-model organisms. It is widely used as a selectable marker gene in various organisms involved in prokaryotic and eukaryotic systems, from bacteria and

yeasts to plants and mammalian cells, with functions independent of the host metabolic background (Miki and McHugh, 2004; Mosey *et al*., 2021).

2.3.1 Mitochondrial *Npt*II (*mtNpt*II) : A modified version from *Npt*II

Two distinct genomes exist in eukaryotic cells: the nuclear genome (nDNA) and mitochondrial genome (mtDNA). The nuclear genome contains thousands of genes, encodes most cellular proteins, and stores the genetic information of the organism. In contrast, mitochondrial DNA is a smaller circular DNA found within the mitochondria. It encodes essential proteins, tRNAs, and rRNAs. Most mitochondrially required proteins are imported after being encoded by the nucleus (Wiese and Bannister, 2020). Due to the difference in nuclear and mitochondrial genomes, *Npt*II that expresses inside the nucleus after it integrates into the host's genome will not integrate into mitochondria and express. To overcome this problem, a modified *mtNpt*II gene was successfully constructed from *Npt*II. Based on previous studies, *mtNpt*II was reliably integrated into the host mitochondrial genome and successfully expressed (Yap, 2025).

2.4 Green fluorescent protein (*GFP*)

Ormo et al. (1996) discovered a green fluorescent protein (*GFP*) in jellyfish (*Aequorea victoria*). He found that jellyfish emit green fluorescence but not

blue fluorescence. Theoretically, aequorin emits blue light when interacting with calcium ions. Based on this observation, he realized that another protein was involved. Through protein isolation, he obtained a second protein that emitted green fluorescence under UV light, which was named Green Fluorescent Protein (*GFP*). They also found that *GFP* can emit green fluorescence in the absence of a substrate or in the jellyfish cellular environment. Due to its non-invasive nature, *GFP* can be used as a reporter gene to visualize molecular activities within living cells.

2.4.1 GFP's three-dimensional structure

The GFP structure was characterized by Ormo et al. (1996) using X-ray crystallography; it is composed of 238 amino acids and has an approximate molecular weight of 27 kDNA. GFP has a distinctive beta-barrel fold in its tertiary structure, consisting of 11 beta strands and only one alpha strand. The β-strands are mostly arranged in an antiparallel orientation; each strand contains 9 to 13 amino acid residues and is stabilized by an extensive network of hydrogen bonds. The β-strands surround a central alpha helix, forming a cylindrical barrel conformation in the presence of hydrogen bonds. The barrel enclosed a highly protected internal environment for the GFP chromophore.

The alpha helix in the center of the barrel consists of only three amino acids: Ser65, Tyr66, and Gly67. These three residues produce a p-

hydroxybenzylideneimidazolidone chromophore through a series of post-translational autocatalytic modifications, including cyclization, dehydration, and oxidation. GFP's fluorescence is determined by folding the barrel structure. Almost all the sequences in the primary structure take part in the formation of the beta and alpha strands; therefore, it is difficult to remove them (Andrews *et al.*, 2007).

2.4.2 Mechanism of the chromophore

In GFP, green fluorescence occurs only when the protein folds into a tertiary beta-barrel structure. The primary amino acid sequences alone without folding do not emit any fluorescence. During folding, the alpha-strand amino acid residues Ser65, Tyr66, and Gly67 are incorporated into the barrel structure. Initially, the Ser65's carbonyl group undergoes nucleophilic attack by the amide nitrogen of Gly67. This process proceeds via cyclization to form a five-membered imidazolidone ring structure. Then, the water molecules were removed by dehydration. Lastly, the alpha-beta bond of Tyr66 is an oxidase that couples the aromatic group of Tyr66 with the imidazolidone ring to create a conjugated system. After post-translational chemical reactions, a mature chromophore for fluorescence was formed. Moreover, the unique environment of the beta-barrel structure protects the chromophore from solvent exposure and facilitates the correct conformational changes necessary through oxidation to confirm the maturation of the chromophore (Ormo et al., 1996).

The mature chromophore can absorb two excitation peaks and emit visible light radiation. The major peak was at approximately 395 nm (UV light), and a minor peak was observed near 475 nm (blue light). Both excitation peaks, which are absorbed by the chromophore, cause an emission peak at approximately 508 nm, which corresponds to green fluorescence. This dual excitation feature allows *GFP* to be visualized under both blue and UV light, making it a convenient reporter gene (Barondeau *et al*., 2003).

2.5 Mitochondrial superfolder green fluorescent protein (*sfGFP*)

The folding of GFP is an essential step for chromophore maturation and fluorescent emission. If the folding process is lacking, cyclization and oxidation reactions will not proceed and the chromophore will not be formed. The folding efficiency affects the fluorescence intensity of GFP. To address the limitations of wild-type GFP, superfolder GFP (sfGFP) was engineered to enhance folding robustness and fluorescence intensity. The mutated *sfGFP* was generated through Cycle3 mutations F99S, M153T, and V163A; the enhanced *GFP* mutations F64L and S65T; and six additional substitutions: S30R, Y39N, N105T, Y145F, I171V, and A206V. These mutations improve the performance of GFP by increasing its folding kinetics, solubility, and stability. The rate of protein misfolding and aggregation was reduced by Cycle3 mutations (Hsu,

Blaser and Jackson, 2009). However, the refolding efficiency of denatured protein reached a maximum of approximately 80%. However, based on the enhanced mutation and six additional mutations with Cycle3 mutations, sfGFP surpassed this limitation by achieving nearly 100% refolding efficiency, showing exceptional kinetic refolding and chemical resilience in various experiments. The excitation peak of sfGFP is at 450 nm, giving an emission peak at 508 nm, allowing bright green fluorescence to emit under blue light illumination. Importantly, sfGFP emits two-fold greater fluorescence intensity than wild-type GFP, enhancing the sensitivity for visualization and imaging applications (Pédelacq *et al*., 2006).
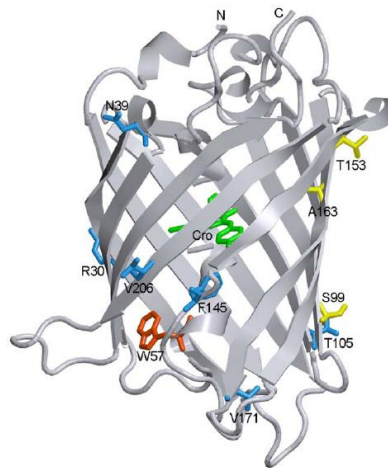
Figure 2.2: sfGFP's spatial structure. The green stick union is the location of the chromophore, and the red stick union shows the single tryptophan residue Trp57. Six mutation sites (S30R, Y39N, N105T, Y145F, I171V, and A206V) in cycle-3 are shown as blue stick unions. The mutations F99S, M153T, and V163A are shown in yellow stick unions (Pédelacq *et al*., 2006).

CHAPTER 3

MATERIALS AND METHODS

3.1 Experimental workflow of the project

This section describes a method for constructing two sets of fusion genes. Both sets included the same core gene (*mtsfGFP*) and different partner genes (*Npt*II or *mtNpt*II). This study used overlap-extension PCR to join the target gene form into a single, continuous reading frame. The experimental workflow is shown as part A and part B for constructing the *Npt*II-*mtsfGFP* and *mtNpt*II-*mtsfGFP* fusion genes, respectively.

Part A: Construct *Npt*II-*mtsfGFP* fusion gene
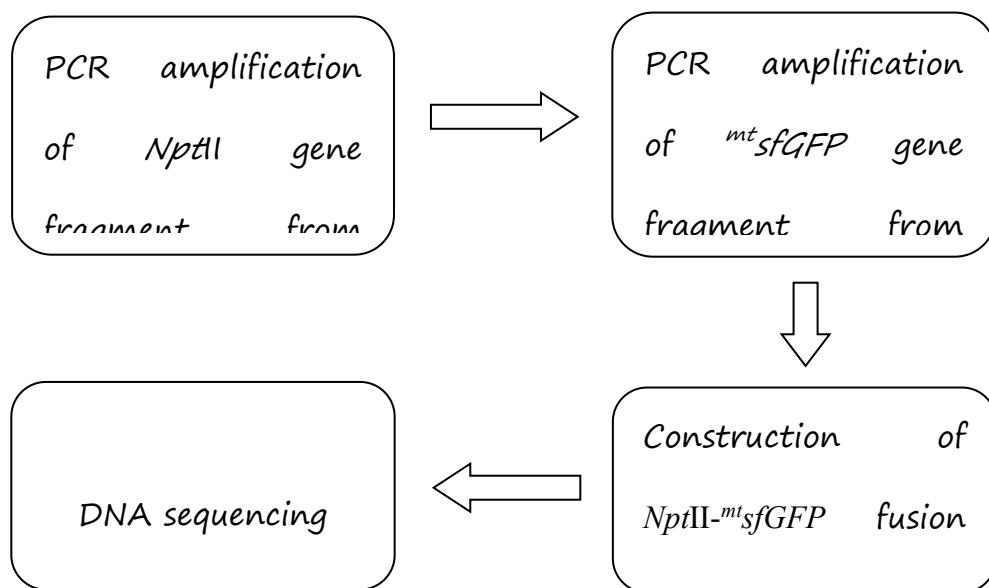
Figure 3.1: Experimental workflow of construct *Npt*II-*mtsfGFP* fusion gene

Part B: Construct *mtNpt*II-*mtsfGFP* fusion gene
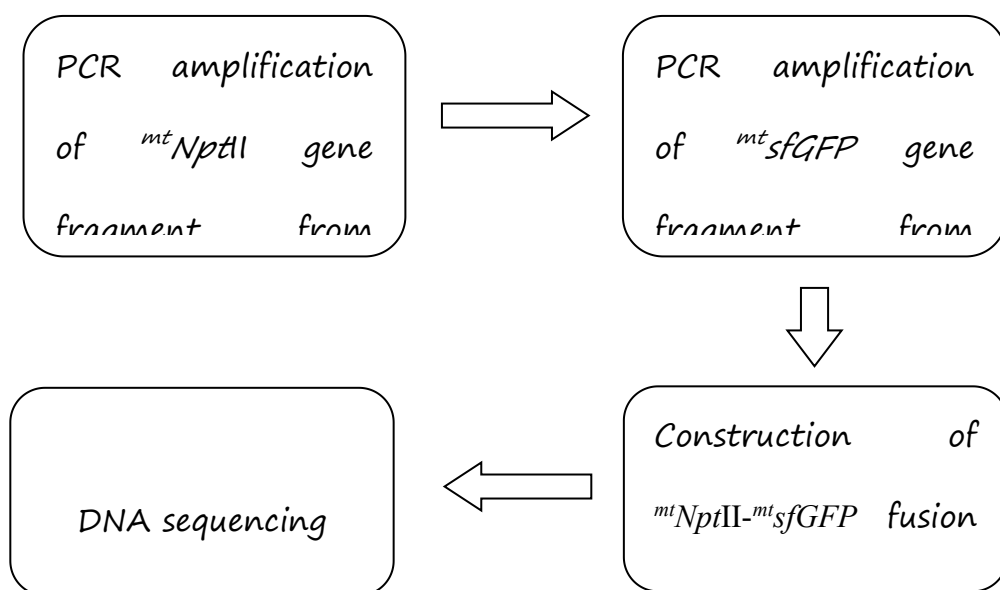
Figure 3.2: Experimental workflow of construct *mtNpt*II-*mtsfGFP* fusion gene

3.2 Preparation of buffer, reagents and apparatus

In this final year project, all the chemical reagents and equipment were provided by Department of Biological Science, Faculty of Science of Universiti Tunku Abdul Rahman (UTAR) and my supervisor Prof. Dr. Wong Hann Ling, as listed in Appendices A.

3.2.1 Preparation of 5X Tris/Borate/EDTA (TBE) buffer

0.5 M of EDTA was prepared in a 5X TBE buffer. An amount of 18.61 g EDTA disodium salt (FW: 372.2 g/mol) was dissolved in 80 mL of dH_2O. The solution was colorless when adjusted to pH8 using 5M NaOH. For the preparation of the 1 L 5X TBE buffer, 5.4 g of tris base (MW: 121.14 g/mol) and 27.5 g of boric acid (FW: 61.83 g/mol.) was dissolved in 800 mL dH_2O. The solution was topped up to 1 L after adding 20 mL of 0.5 M EDTA solution.

3.2.2 Preparation of 3 M sodium acetate (NaOAc) solution

Fifty mL of 3M NaOAc solution was used as the essential solution for ethanol precipitation. 12.30 g of sodium acetate (MW: 82.03 g/mol) was dissolved in a beaker containing 30 mL of dH_2O. Then, the pH of the solution was adjusted to pH 5.2 with glacial acetic acid. Finally, the volume of the solution was increased to fifty mL using dH_2O.

3.3 Agarose gel electrophoresis

The PCR products were verified by agarose gel electrophoresis. The 5X TBE was diluted to $0.5 \times$ TBE buffer with dH_2O, which served as the running buffer, and the agarose gel was cast. The percentage of agarose gel electrophoresis was determined based on the desired size of the DNA fragment. A larger DNA fragment size required a lower percentage of agarose gel. To verify the PCR product from single primer amplification and target gene amplification below 1500 bp, a 1.2% agarose gel was cast. Agarose gel powder (0.48 g) was added to 40 mL of $0.5 \times$ TBE and completely dissolved in the solution after being heated in a microwave. Then, gel stain (Dye All^{TM}) was added to the solution when it slightly cooled. After that, the mixed solution was poured into the gel casting set and left to settle and solidify. This took approximately 40 min. Next, the gel was run in $0.5 \times$ TBE buffer at 100 V for 35 min and viewed under a UV transilluminator.

To verify the larger amplicon (> 1500 bp) from overlap extension PCR, a 0.8% agarose gel was cast. The agarose gel electrophoresis process was the same as that used for 1.2% agarose gel electrophoresis, except that the amount of agarose powder was decreased to 0.32 g.

3.4 Purification of PCR fragment using EZ-10 Spin Column DNA Gel Extraction Kit

The PCR product was purified using the EZ-10 spin column DNA gel extraction kit. After the target band location was confirmed by comparison with the DNA ladder under a UV transilluminator, gel excision was performed. The

target band was cut from the agarose gel by using a sterile scalpel. The weight of the gel slide was then placed into a 1.5 mL microcentrifuge tube and measured using an electronic weighting scale. Next, the binding buffer was added to the same 1.5 mL microcentrifuge tube. The volume of binding buffer added followed the rules, in which 200 µL of binding buffer was added to each 100 mg of agarose gel. The sample was then incubated in a preheated heat block at 50 °C until the gel slide completely melted. The sample was vortexed every 3 min to evenly mix the binding buffer and gel slide, and the gel slide melting process was accelerated. The EZ-10 column was placed in a 2.0 mL collection tube first when cooling down the mixture at room temperature. A volume of 700 µL of the mixture was transferred to the column and centrifuged at 11,000 × g for 30 s. The flow-through was discarded, and 700 µL of washing buffer was added to the column. Then, the flow-through was discarded after centrifugation at 11,000 × g for 30 s. The washing steps were repeated to minimize the chaotropic salt and improve the A_{260}/A_{230} values. The column containing the collection tube was centrifuged at 11,000 × g for 1 min to completely remove the washing buffer. The column was transferred to a new 1.5 mL microcentrifuge tube. Subsequently, 30 µL of preheated elution buffer was pipetted directly onto the center of the membrane and fully covered the membrane surface to improve the elution efficiency of the bound DNA. The column was incubated at room temperature for 1 min before centrifugation at 11,000 × g for 1 min. The bound DNA was eluted and the concentration was measured using a NanoDrop. Finally, the extracted DNA was stored at -20°C freezer until further use.

3.5 Construction of *Npt*II/*mtNpt*II-*mtsfGFP* fusion gene

3.5.1 Graphical overview of fusion gene construction

Step 1: Amplification of primary fragments

PCR reaction for *Npt*II/*mtNpt*II

Primer 1 (P1)   Primer 1-P2

*Npt*II/*mtNpt*II DNA template

*Npt*II/*mtNpt*II

PCR reaction for *GFP*

Primer 2-P1   Primer 2 (P2)

GFP DNA template

GFP
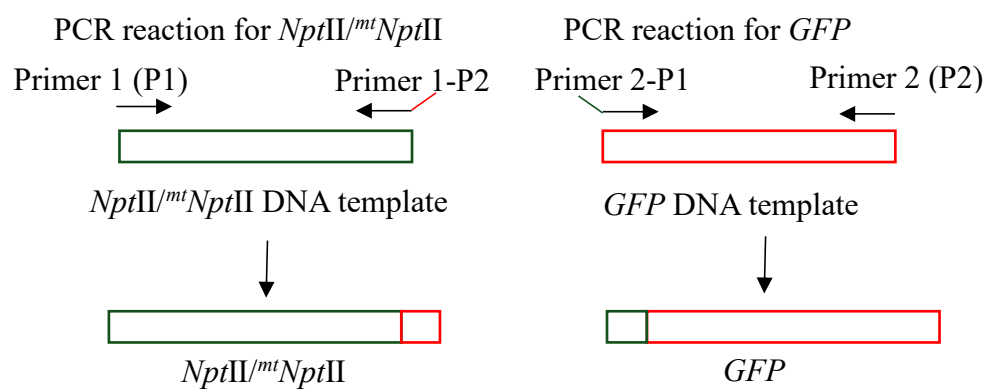
Step 2: Overlap PCR reaction

*Npt*II/*mtNpt*II

GFP

15 cycles PCR without primers

*Npt*II/*mtNpt*II-*GFP*

28
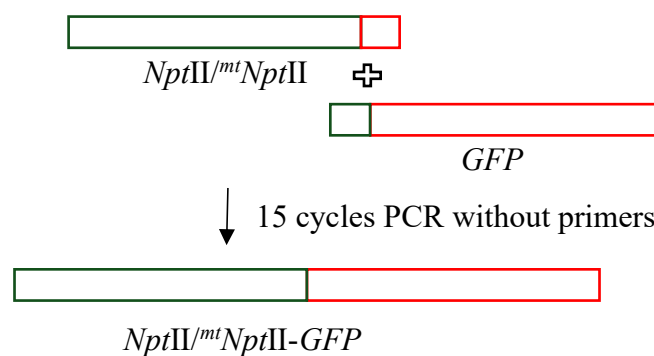
Step 3: Final fusion gene PCR amplification

Final amplified fusion product
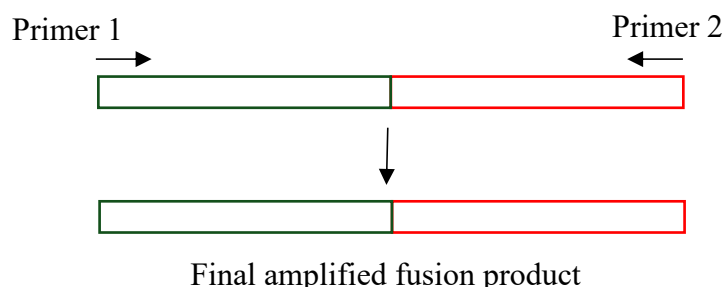
Figure 3.1: Graphical overview of constructing the *Npt*II/*^{mt}Npt*II-*^{mt}sfGFP* fusion gene by overlap extension PCR. In step 1 (top panel), target gene fragments containing overlapping regions were amplified by PCR. Step 2 (middle panel) uses the *Npt*II/*^{mt}Npt*II and *^{mt}sfGFP* gene fragments to generate the full-length *Npt*II/*^{mt}Npt*II-*^{mt}sfGFP* fusion gene by overlap extension PCR. Step 3: Final PCR amplification of the full-length *Npt*II/*^{mt}Npt*II-*^{mt}sfGFP* fusion gene as the final product with the outermost primer.

3.5.2 Megaprimer Design

Megaprimers were designed to generate the *Npt*II/*^{mt}Npt*II-*^{mt}sfGFP* fusion gene. The target fusion gene that contains the overlap region was integrated into a single, continuous reading frame through overlap PCR. For successful fusion, the stop codon at the end of *Npt*II/*^{mt}Npt*II was removed and only located at the end of *^{mt}sfGFP*. Both megaprimers shown in the overlap region contained fused sequences that were complementary to each other, ensuring precise amplification and incorporation of the overlapping segment necessary for efficient gene fusion. To construct the *Npt*II/*^{mt}Npt*II-*^{mt}sfGFP* fusion gene, the reverse megaprimer (R-OPCR-*Npt*II-*^{mt}sfGFP*/R-OPCR-*^{mt}Npt*II-*^{mt}sfGFP*) for amplifying the *Npt*II/*^{mt}Npt*II gene contains the *^{mt}sfGFP* gene sequence. The forward primer (F-OPCR-*Npt*II-*^{mt}sfGFP* /F-OPCR-*^{mt}Npt*II-*^{mt}sfGFP*) contains

the end of *Npt*II/*ᵐᵗNpt*II gene sequence before starting to amplify the *ᵐᵗsfGFP*

gene sequence. The primer sequences are listed in Tables 3.1, 3.2, and 3.3.
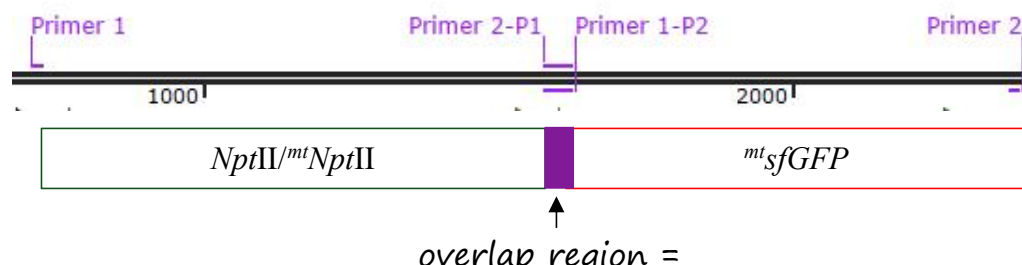
Figure 3.2: Schematic diagram showing megaprimer design to construct the *Npt*II/*ᵐᵗNpt*II-*ᵐᵗsfGFP* fusion gene. The designed megaprimer introduces an overlap region that contain a 55 bp sequence, labeled as a purple region, for seamless fusion by overlap extension PCR. Both primers shown in the figure were positioned to flank their respective target sequences. The reverse primer was designated as primer 1-P2 and the forward primer was labeled as primer 2-P1.

Table 3.1: Nucleotide sequence of primers used to amplify the *Npt*II gene fragment from the pENTR-P*TEF1*-*Npt*II plasmid.

Primer	Sequences (5'…3')
F-*att*L-*Not*I	CAACTTTGTACAAAAAAGCAGGCT
R-OPCR-*Npt*II-*ᵐᵗsfGFP*	GAATAATTCTTCACCTTTAGACATGGTCTCG GATCCGAAAAACTCATCGAGCATC

* Underlined sequences indicate overhanging regions.

Table 3.2: Nucleotide sequences of primers used to amplify *ᵐᵗNpt*II gene fragments from the pENTR-P*Cox1*-*ᵐᵗNpt*II plasmid.

Primer	Sequences (5'…3')
F-*att*L-*Not*I	CAACTTTGTACAAAAAAGCAGGCT
R-OPCR-*ᵐᵗNpt*II-*ᵐᵗsfGFP*	GAATAATTCTTCACCTTTAGACATGGTCTC

<u>GGATCC</u>GAAGAACTCGTCCAACATC

* Underlined sequences indicate overhanging regions.

Table 3.3: Nucleotide sequences of primers used to amplify the $^{mt}sfGFP$ gene fragment from pENTR-P$_{Cox2}$- $^{mt}sfGFP$ plasmid.

Primer	Sequences (5'…3')
F-OPCR-*Npt*II-$^{mt}sfGFP$	<u>GATGCTCGATGAGTTTTTC</u>GGATCCGAGA CCATGTCTAAAGGTGAAGAATTATTC
F-OPCR-^{mt}NptII-$^{mt}sfGFP$	<u>GATGTTGGACGAGTTCTTC</u>GGATCCGAGA <u>CC</u>ATGTCTAAAGGTGAAGAATTATTC
R-T7-Ter-AscI	GGCGCGCCCACCCTTCAAAAAAC

* Underlined sequences indicate overhanging regions.

3.5.3 Single-Primer Amplification Test with DNA Template

To eliminate false positive results, a single primer was tested by PCR with the related DNA template. Due to the $^{mt}sfGFP$ gene fragment fuses with different gene fragments, it was tested with different pairs of primers. The data from this section will contribute to further modifications of the PCR conditions. The PCR reagent components and thermocycling parameters are listed in Tables 3.3, 3.4, and 3.5, respectively.

Table 3.4: PCR master mix and reaction components with their corresponding volumes.

Component	Stock concentration	Final concentration	Volume (µL)

Vayzme 2X Taq master mix	2X	1X	5
Forward primer/ Reverse primer	10 μM	0.3 μM	0.3
DNA template	50 ng/ μL	1 ng/ μL	0.2
Sterile distilled water	-	-	4.2
Total volume	-	-	10

Table 3.5: PCR thermocycling parameters for ^{mt}NptII and $^{mt}sfGFP$ single primer testing with the extracted pENTR-P$_{Cox1}$-^{mt}NptII and pENTR-P$_{Cox2}$-$^{mt}sfGFP$ plasmids, respectively.

Reaction	Temperature (ºC)	Time	Numbers of cycles
Initial denaturation	95	3 min	1
Denaturation	95	15 sec	
*Annealing	56	15 sec	25
	53		
Extension	72	1 min	
Final extension	72	5 min	1
Hold	4	∞	-

* For ^{mt}NptII and $^{mt}sfGFP$(^{mt}NptII) single-primer testing, the annealing temperature was set at 56°C. For $^{mt}sfGFP$(NptII) single-primer testing, the annealing temperature was set to 53°C.

Table 3.6: PCR thermocycling parameters for NptII single-primer amplification with the extracted pENTR-p$_{TEF1}$-*Npt*II plasmid.

Reaction	Temperature (°C)	Time	Numbers of cycles
Initial denaturation	95	3 min	1
Denaturation	95	15 sec	
Annealing	56	15 sec	20
Extension	72	1:30 min	
Final extension	72	5 min	1
Hold	4	∞	-

3.5.4 PCR amplification of *Npt*II, ^{mt}*Npt*II and ^{mt}*sfGFP* gene fragment

The first PCR amplification generates amplicons that contain overlapping sequences for further overlap extension PCR, as described in Section 3.5.5. To increase the efficiency of PCR, the thermocycling parameters for fragment amplification were modified based on the results of the single-primer amplification test. The primer sets, PCR components, and thermocycling parameters are summarized in Tables 3.1, 3.2, 3.3, 3.6, 3.7, and 3.8, respectively.

Table 3.7: PCR master mix and reaction components with their corresponding volumes.

Component	Stock	Final	Volume (µL)

	concentration	concentration	
Vayzme 2X Taq master mix	2X	1X	25
Forward primer	10 µM	0.3 µM	1.5
Reverse primer	10 µM	0.3 µM	1.5
DNA template	50 ng/ µL	1 ng/ µL	1
Sterile distilled water	-	-	21
Total volume	-	-	50

Table 3.8: PCR thermocycling parameters for amplifying *Npt*II with the extracted pENTR-P$_{TEF1}$-*Npt*II plasmid.

Reaction	Temperature (ºC)	Time	Numbers of cycles
Initial denaturation	95	3 min	1
Denaturation	95	15 sec	
Annealing	56	15 sec	20
Extension	72	1:30 min	
Final extension	72	5 min	1
Hold	4	∞	-

Table 3.9: PCR thermocycling parameters for amplifying ^{mt}NptII/$^{mt}sfGFP$ with the extracted pENTR-P$_{Cox1}$-^{mt}NptII/pENTR-P$_{Cox2}$-$^{mt}sfGFP$ plasmid.

Reaction	Temperature (ºC)	Time	Numbers of cycles
Initial denaturation	95	3 min	1
Denaturation	95	15 sec	
	53	15 sec	
*Annealing	56		25
	60		
Extension	72	1 min	
Final extension	72	5 min	1
Hold	4	∞	-

* For the amplified $^{mt}sfGFP$(NptII) the annealing temperature was set to 53°C.

* For the amplified ^{mt}NptII, the annealing temperature was set to 56°C.

* For the amplified $^{mt}sfGFP$(^{mt}NptII), the annealing temperature was set to 60°C.

3.5.5 Overlap extension PCR for NptII/^{mt}NptII-$^{mt}sfGFP$ fusion gene construction

The purpose of the overlap extension PCR was to obtain the full-length NptII/^{mt}NptII-$^{mt}sfGFP$ fusion gene. In this PCR reaction, no primer was used, and it relied on the purified PCR product generated in Section 3.5.4. First, the PCR products were verified by agarose gel electrophoresis (Section 3.3) and then purified as described in Section 3.4. The purified NptII/^{mt}NptII and $^{mt}sfGFP$ genes were diluted and the molar ratio of the DNA template was calculated using the Vayzme calculator, followed by each fragment bp size. The larger the fragment size, the larger the amount of input DNA. The PCR

components and thermocycling parameters are presented in Tables 3.9, 3.10, and 3.11, respectively.

Table 3.10: PCR master mix and reaction components with their corresponding volumes for the *Npt*II-*^{mt}sfGFP* fusion gene.

Component	Stock concentration	Final concentration	Volume (μL)
Vayzme 2X Taq master mix	2X	1X	25
DNA template - *Npt*II	1 ng/ μL	0.15 ng/ μL	7.5
DNA template - *^{mt}sfGFP*	1 ng/ μL	0.1 ng/ μL	5

Component	Stock concentration	Final concentration	Volume (µL)
Sterile distilled water	-	-	9.5
Total volume	-	-	47

Table 3.11: PCR master mix and reaction components with their corresponding volumes for ^{mt}NptII-$^{mt}sfGFP$ fusion gene.

Component	Stock concentration	Final concentration	Volume (µL)
Vayzme 2X Taq master mix	2X	1X	25
DNA template - ^{mt}NptII	1 ng/ µL	0.15 ng/ µL	7.5
DNA template - $^{mt}sfGFP$	1 ng/ µL	0.135 ng/ µL	6.75

Sterile distilled water	-	-	7.75
Total volume	-	-	47

Table 3.12: PCR thermocycling parameters for constructing the *Npt*II/*^{mt}Npt*II-*$^{mt}sfGFP* fusion gene.

Reaction	Temperature (ºC)	Time	Numbers of cycles
Initial denaturation	95	3 min	1
Denaturation	95	15 sec	
Annealing	60	15 sec	15
Extension	72	30 sec	
Final extension	72	5 min	1
Hold	4	∞	-

3.5.6 Final *Npt*II/*^{mt}Npt*II-*$^{mt}sfGFP* fusion gene PCR amplification

Full-length *Npt*II/*^{mt}Npt*II-*$^{mt}sfGFP* fusion gene was amplified for further DNA cloning. Due to both amplified fragments have an overlap region, they fuse together in section 3.5.5. Subsequently, the primers were added directly to the same PCR mixture for amplification. The *Npt*II/*^{mt}Npt*II-*$^{mt}sfGFP* fusion gene was amplified using the same outer primer set. The primer set, number of added primers, and thermocycling parameters are listed in Tables 3.12, 3.13, 3.14, and 3.15, respectively.

Table 3.13: Nucleotide sequences of the outer primers used to amplify the *Npt*II/*ᵐᵗNpt*II-*ᵐᵗsfGFP* fusion gene fragment.

Primer	Sequences (5'…3')
F-*att*L-*Not*I	CAACTTTGTACAAAAAAGCAGGCT
R-T7-Ter-*Asc*I	GGCGCGCCCACCCTTCAAAAAAC

Table 3.14: PCR mixtures and primers added with their corresponding volumes.

Component	Stock concentration	Final concentration	Volume (µL)
PCR mixture	-	-	47
Forward primer	10 µM	0.3 µM	1.5
Reverse primer	10 µM	0.3 µM	1.5
Total volume	-	-	50

Table 3.15: PCR thermocycling parameters for amplifying the *Npt*II-*ᵐᵗsfGFP* fusion gene.

Reaction	Temperature (ºC)	Time	Numbers of cycles
Initial denaturation	95	3 min	1
Denaturation	95	15 sec	
Annealing	59	15 sec	25
Extension	72	2:30 min	
Final extension	72	5 min	1

| | Hold | 4 | ∞ | - |

Table 3.15: PCR thermocycling parameters for amplifying the ^{mt}NptII-$^{mt}sfGFP$ fusion gene.

Reaction	Temperature (ºC)	Time	Numbers of cycles
Initial denaturation	95	3 min	1
Denaturation	95	15 sec	
Annealing	60	15 sec	25
Extension	72	2 min	
Final extension	72	5 min	1
Hold	4	∞	-

3.6 Purification of NptII-$^{mt}sfGFP$ fusion gene fragment

For further DNA cloning, purification is required to remove impurities in the PCR products, such as salt, residual primers, and non-specific amplification products. DNA purity increases the efficiency of cloning. First, the PCR product was verified using 0.8% agarose gel electrophoresis, as described in Section 3.3. The target PCR product was purified following the steps described in Section 3.4. After purification, the purified DNA was stored at − 20°C for further use.

3.7 Purification of ^{mt}NptII-$^{mt}sfGFP$ fusion gene fragment

3.7.1 Low-melting agarose gel electrophoresis

To easily extract the desired DNA fragment from the agarose gel, a low-melting agarose gel was used. The criteria for this type of agarose gel were utilized to make it easier to dissolve in dH_2O at moderate temperatures (~60°C). First, 0.2 g of low-melting agarose powder was weighed and dissolved in 0.5X TBE buffer. The subsequent process preparation of 1.0% low-melting agarose gel was identical to that of normal agarose gel preparation in section 3.4. The gel was run in 0.5 ×TBE buffer at 100 V for 35 min and viewed under a UV transilluminator.

3.7.2 Ethanol precipitation for gel purification

The amplified ^{mt}NptII-$^{mt}sfGFP$ fusion gene product was purified by ethanol precipitation. The target band on the gel was cut using a sterile scalpel after the target band location was confirmed using a UV transilluminator. Next, the gel slide was collected into a 1.5 mL microcentrifuge tube with 500 µL of dH_2O. The microcentrifuge tube was incubated in a preheated heat block at 50 Â °C. To completely dissolve the gel slide, the microcentrifuge tube was vortexed

every 2 minutes. The supernatant (500 μL) was transferred into a new 1.5 mL microcentrifuge tube. 1 mL of phenol-chloroform-isoamyl alcohol (PCI) mixture was added inside, and the mixture was vortexed for 2 minutes. The microcentrifuge tube was then centrifuged at 12,500 rpm and 4°C for 10 min in a pre-cold low-temperature centrifuge. Five hundred microliters of supernatant were transferred into a new 1.5 mL microcentrifuge tube; the buffy coat was carefully avoided during collection to improve DNA purity and minimize contamination. Next, 50 μL of 3 M NaOAc solution and 500 μL of ice-cold 100% EtOH were added. To improve DNA sedimentation, the solution was gently mixed and placed in a -20°C freezer overnight. The next day, the solution was centrifuged at 12,500 rpm at 4°C for 30 min in a pre-cooled low-temperature centrifuge to obtain the DNA pellet. The supernatant was discarded and 1 mL of ice-cold 100% EtOH was pipetted into a microcentrifuge tube for pellet washing. Next, the supernatant was discarded and the remaining supernatant was pipetted. The whitish pellet became transparent after being fully air-dried. Finally, the pellet was fully dissolved in 30 μL of dH$_2$O, and the concentration of extracted DNA was measured using a NanoDrop spectrophotometer. For long-term storage, the extracted DNA was stored at -20°C until further use.

3.8 DNA sequencing

The fusion gene product was purified and stored as extracted DNA in a -20°C freezer. For both fusion gene products, *Npt*II-*^{mt}sfGFP* and *^{mt}Npt*II-*^{mt}sfGFP*, the size was over 1500 bp. For 1st BASE company DNA sequencing, the required concentration of DNA, which was over 1500 bp, was 30 ng/μL. Both fusion genes reached the concentration requirement, but the purity of the *^{mt}Npt*II-

mtsfGFP fusion gene did not meet this requirement. Twenty microliters of the extracted DNA and 10 µL of forward and reverse primers were used for sequencing.

CHAPTER 4

RESULT

4.1 Construction of *Npt*II/*ᵐᵗNpt*II-*ᵐᵗsfGFP* fusion gene

4.1.1 Megaprimer Design

```
                                                              F-OPCR-NptII-mtsfGFP
                                                              gatgctcgatgagttttcGGATCCGAGACCATGTCTAAAGGTGAAG
tttcaaaaatatggtattgataatcctgatatgaataaattgcagtttcatttgatgctcgatgagttttcGGATCCGAGACCATGTCTAAAGGTGAAG
+++++++++|+++++++++|+++++++++|+++++++++|+++++++++|+++++++++|+++++++++|+++++++++|+++++++++|+++++++++|
aaagtttttataccataactattaggactatacttatttaacgtcaaagtaaactacgagctactcaaaaagCCTAGGCTCTGGTACAGATTTCCACTTC
         250          255          260          265      1    G  S  E  T   M  S  K  G  E
 F  Q  K  Y  G  I  D  N  P  D  M  N  K  L  Q  F  H  L  M  L  D  E  F  F
 ◄─────────────────────────── NptII ───────────────────────────►   Linker     mtsfGFP
                                                              ctacgagctactcaaaaagCCTAGGCTCTGGTACAGATTTCCACTTC
                                                              R-OPCR-NptII-mtsfGFP

II-mtsfGFP
AATTATTC
AATTATTCACTGGTGTTGTACCAATTTTAGTTGAATTAGATGGTGATGTTAATGGTCATAAATTTTCTGTAAGAGGTGAAGGTGAAGGTGATGCTACAAA
+++++++++|+++++++++|+++++++++|+++++++++|+++++++++|+++++++++|+++++++++|+++++++++|+++++++++|+++++++++|
TTAATAAGTGACCACAACATGGTTAAAATCAACTTAATCTACCACTACAATTACCAGTATTTAAAAGACATTCTCCACTTCCACTTCCACTACGATGTTT
   15        20        25        30        35        40        45
 E  L  F  T  G  V  V  P  I  L  V  E  L  D  G  D  V  N  G  H  K  F  S  V  R  G  E  G  E  G  D  A  T  N
 ◄────────────────────────────── mtsfGFP ──────────────────────────────────►
                                                              GATGTTGGACGAGTTCTTCGGATCCGA
CTTCCTTGCAGAAGCGTTTGTTCCAAAAGTATGGTATCGATAACCCTGACATGAATAAATTACAGTTCCACTTGATGTTGGACGAGTTCTTCGGATCCGA
+++++++++|+++++++++|+++++++++|+++++++++|+++++++++|+++++++++|+++++++++|+++++++++|+++++++++|+++++++++|
GAAGGAACGTCTTCGCAAACAAGGTTTTCATACCATAGCTATTGGGACTGTACTTATTTAATGTCAAGGTGAACTACAACCTGCTCAAGAAGCCTAGGCT
    240       245       250       255       260       265      1
 P  S  L  Q  K  R  L  F  Q  K  Y  G  I  D  N  P  D  M  N  K  L  Q  F  H  L  M  L  D  E  F  F  G  S  E
 ◄────────────────────────────── mtNPTII ──────────────────────────────────►   Linker
                                                              CTACAACCTGCTCAAGAAGCCTAGGCT
                                                              R-OPCR-mtNptII-mtsfGFP

F-OPCR-mtNptII-mtsfGFP
GACCATGTCTAAAGGTGAAGAATTATTC
GACCATGTCTAAAGGTGAAGAATTATTCACTGGTGTTGTACCAATTTTAGTTGAATTAGATGGTGATGTTAATGGTCATAAATTTTCTGTAAGAGGTGAA
+++++++++|+++++++++|+++++++++|+++++++++|+++++++++|+++++++++|+++++++++|+++++++++|+++++++++|+++++++++|
CTGGTACAGATTTCCACTTCTTAATAAGTGACCACAACATGGTTAAAATCAACTTAATCTACCACTACAATTACCAGTATTTAAAAGACATTCTCCACTT
 1       10        15        20        25        30        35        40
 T  M  S  K  G  E  E  L  F  T  G  V  V  P  I  L  V  E  L  D  G  D  V  N  G  H  K  F  S  V  R  G  E
 Linker  ◄────────────────────────────── mtsfGFP ──────────────────────────────────►
CTGGTACAGATTTCCACTTCTTAATAAG
```

Figure 4.1: Schematic showing the sequence of the *Npt*II/*ᵐᵗNpt*II-*ᵐᵗsfGFP* fusion gene. Megaprimers were designed to cross the *Npt*II/*ᵐᵗNpt*II and *ᵐᵗsfGFP* gene fragments. The overlapping regions were complementary to each other. The stop codon in *Npt*II/*ᵐᵗNpt*II was removed. The full length of both forward and reverse megaprimers was 55 bp, and Tm was 69°C.

A) Potential hairpin formation :

5' GATGCTCGATGAGTTTTTCGGATCCGAGACCATGTCTAAAGGTGAAGAATTATTC 3'

3' Complementarity:
None !

All potential self-annealing sites are marked in red (allowing 1 mis-match):

```
5'          GATGCTCGATGAGTTTTTCGGATCCGAGACCATGTCTAAAGGTGAAGAATTATTC    3'
3' CTTATTAAGAAGTGGAAATCTGTACCAGAGCCTAGGCTTTTTGAGTAGCTCGTAG             5'
```

Potential hairpin formation :

5' GATGTTGGACGAGTTCTTCGGATCCGAGACCATGTCTAAAGGTGAAGAATTATTC 3'

3' Complementarity:
None !

All potential self-annealing sites are marked in red (allowing 1 mis-match):

```
5'          GATGTTGGACGAGTTCTTCGGATCCGAGACCATGTCTAAAGGTGAAGAATTATTC    3'
3' CTTATTAAGAAGTGGAAATCTGTACCAGAGCCTAGGCTTCTTGAGCAGGTTGTAG             5'

5' GATGTTGGACGAGTTCTTCGGATCCGAGACCATGTCTAAAGGTGAAGAATTATTC    3'
3'          CTTATTAAGAAGTGGAAATCTGTACCAGAGCCTAGGCTTCTTGAGCAGGTTGTAG    5'

5' GATGTTGGACGAGTTCTTCGGATCCGAGACCATGTCTAAAGGTGAAGAATTATTC    3'
3'                    CTTATTAAGAAGTGGAAATCTGTACCAGAGCCTAGGCTTCTTGAGCAGGTTGTAG 5'
```

B)
```
5'          GATGTTGGACGAGTTCTTCGGATCCGAGACCATGTCTAAAGGTGAAGAATTATTC    3'
3' CTTATTAAGAAGTGGAAATCTGTACCAGAGCCTAGGCTTCTTGAGCAGGTTGTAG             5'

5'          GATGTTGGACGAGTTCTTCGGATCCGAGACCATGTCTAAAGGTGAAGAATTATTC    3'
3' CTTATTAAGAAGTGGAAATCTGTACCAGAGCCTAGGCTTCTTGAGCAGGTTGTAG             5'
```

C)
```
5' GATGTTGGACGAGTTCTTCGGATCCGAGACCATGTCTAAAGGTGAAGAATTATTC    3'
3'          CTTATTAAGAAGTGGAAATCTGTACCAGAGCCTAGGCTTCTTGAGCAGGTTGTAG    5'
```

D) Potential hairpin formation :

```
5' GAATAATTCTTCACCTTTAGACATGGTCTCGGATCCGAAGAACTCGTCCAACATC 3'
5' GAATAATTCTTCACCTTTAGACATGGTCTCGGATCCGAAGAACTCGTCCAACATC 3'
5' GAATAATTCTTCACCTTTAGACATGGTCTCGGATCCGAAGAACTCGTCCAACATC 3'
```

3' Complementarity:
None !

All potential self-annealing sites are marked in red (allowing 1 mis-match):

```
5' GAATAATTCTTCACCTTTAGACATGGTCTCGGATCCGAAGAACTCGTCCAACATC                                    3'
3'           CTACAACCTGCTCAAGAAGCCTAGGCTCTGGTACAGATTTCCACTTCTTAATAAG                            5'

5'      GAATAATTCTTCACCTTTAGACATGGTCTCGGATCCGAAGAACTCGTCCAACATC                                 3'
3' CTACAACCTGCTCAAGAAGCCTAGGCTCTGGTACAGATTTCCACTTCTTAATAAG                                      5'

5'                                      GAATAATTCTTCACCTTTAGACATGGTCTCGGATCCGAAGAACTCGTCCAACATC 3'
3' CTACAACCTGCTCAAGAAGCCTAGGCTCTGGTACAGATTTCCACTTCTTAATAAG                                      5'

5'      GAATAATTCTTCACCTTTAGACATGGTCTCGGATCCGAAGAACTCGTCCAACATC                                 3'
3' CTACAACCTGCTCAAGAAGCCTAGGCTCTGGTACAGATTTCCACTTCTTAATAAG                                      5'

5' GAATAATTCTTCACCTTTAGACATGGTCTCGGATCCGAAGAACTCGTCCAACATC                                      3'
3'           CTACAACCTGCTCAAGAAGCCTAGGCTCTGGTACAGATTTCCACTTCTTAATAAG                            5'

5' GAATAATTCTTCACCTTTAGACATGGTCTCGGATCCGAAGAACTCGTCCAACATC                                      3'
3'           CTACAACCTGCTCAAGAAGCCTAGGCTCTGGTACAGATTTCCACTTCTTAATAAG                            5'
```

Figure 4.2: Analysis of megaprimer secondary structure using an Oligonucleotide Properties Calculator. In the megaprimer design for the *Npt*II-*mtsfGFP* fusion gene: A) Forward megaprimer: F-OPCR-*Npt*II-*mtsfGFP* shows the potential for one hairpin formation and five self-annealing sites; B) Reverse megaprimer: R-OPCR-*Npt*II-*mtsfGFP* shows the potential for two hairpin formation and five self-annealing sites. In the megaprimer design for the *mtNpt*II-*mtsfGFP* fusion gene: C) Forward primer: F-OPCR- *mtNpt*II-*mtsfGFP* shows the potential for one hairpin formation and six self-annealing sites; D) Reverse megaprimer: R-OPCR- *mtNpt*II-*mtsfGFP* shows the potential for three hairpin formations and six self-annealing sites.

4.2

A) Lane   1   2   3   4   5   6   7   8   45

1 kb→   ←1 kb

←500 bp

Single primer amplification test with DNA template

500 bp

100 bp ➤   ⬅100 bp

B)

Lane   1   2   3

1 kb ➤

500 bp ➤

100 bp ➤
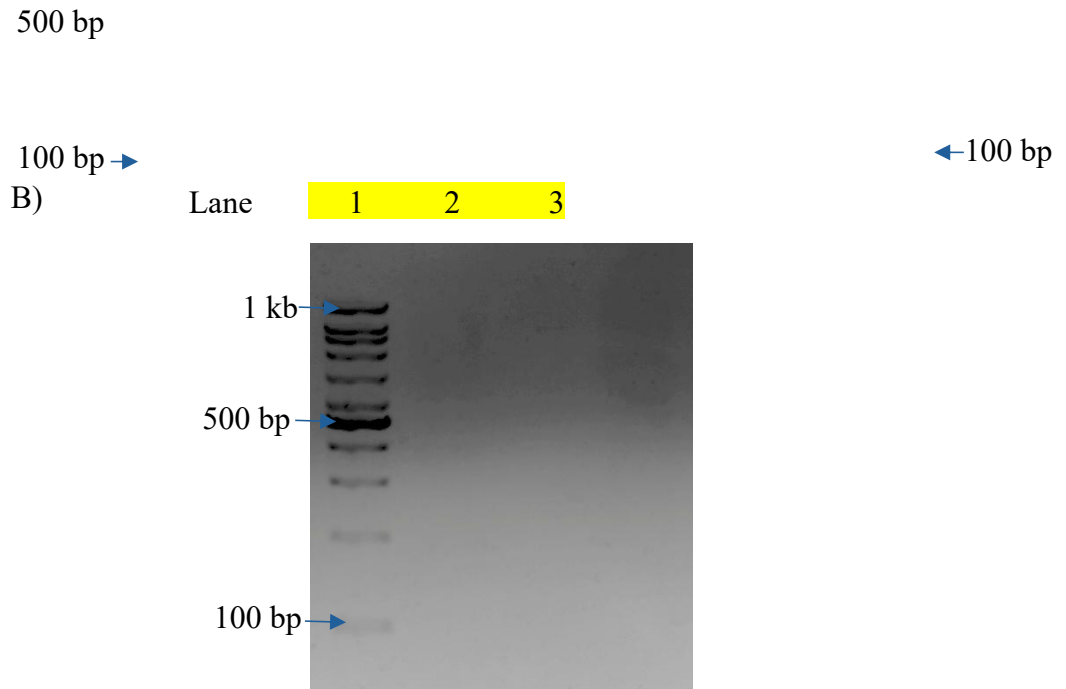
Figure 4.3: **Inverted 1.2% agarose gel image from the single-primer amplification test with the corresponding DNA template**. The gel images were digitally inverted to enhance band contrast for better visualization and assessment of product absence. In part A, three sets of single primer testing are illustrated: lanes 2 and 3 correspond to reactions using primers F-*att*L-*Not*I and R-OPCR- ^{mt}NptII-$^{mt}sfGFP$; lanes 4 and 5 show F-OPCR-^{mt}NptII-$^{mt}sfGFP$ and R-T7-Ter-*Asc*I; and lanes 6 and 7 display F-OPCR- *Npt*II-$^{mt}sfGFP$ and R-T7-Ter-*Asc*I. Lanes 1 and 8 contain the 100 bp DNA ladder as a molecular weight marker. In part B, a 1 kb DNA ladder was used in lane 1, suitable for the 1402 bp *Npt*II gene fragment. The PCR conditions were optimized to ensure no detectable bands in any of the single primer reactions, confirming the specificity of the amplification conditions shown in the figure.

4.3. PCR amplification of gene fragment

4.3.1 *Npt*II, *ᵐᵗNpt*II, and *ᵐᵗsfGFP* gene fragments
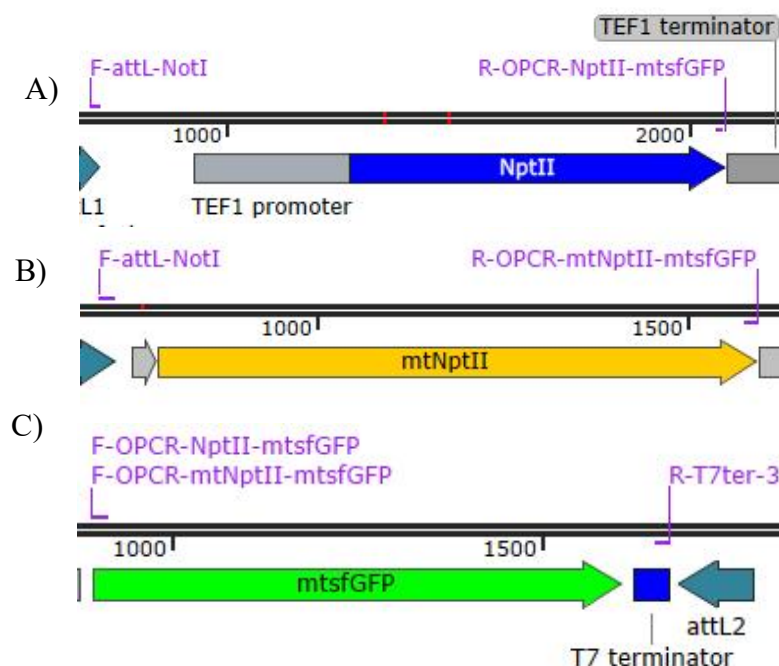
Figure 4.4: Schematic showing the sequence of the *Npt*II, *ᵐᵗNpt*II, and *ᵐᵗsfGFP* gene fragments with their primers. In part A, *Npt*II sequences was amplified by F-*att*L-*Not*I and R-OPCR- *Npt*II-*ᵐᵗsfGFP* primer set, R-OPCR-*Npt*II-*ᵐᵗsfGFP* contain the overhanging region complementary with *ᵐᵗsfGFP*. In part B, *ᵐᵗNpt*II sequences was amplified by F-*att*L-*Not*I and R-OPCR-*ᵐᵗNpt*II-*ᵐᵗsfGFP* primer set, R-OPCR-*ᵐᵗNpt*II-*ᵐᵗsfGFP* contain the overhanging region complementary with *ᵐᵗsfGFP*. In part C, *ᵐᵗsfGFP* was amplified with difference forward primer, which were F-OPCR-*Npt*II-*ᵐᵗsfGFP* and F-OPCR-*ᵐᵗNpt*II-*ᵐᵗsfGFP* to create different overhanging region for further fusion step, reverse primer was R-T7-Ter-*Asc*I.

4.3.2 Result of agarose gel electrophoresis
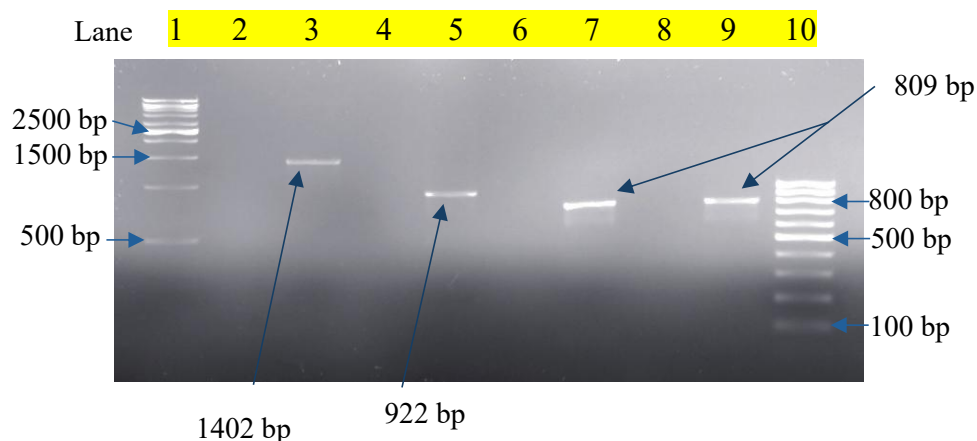
Figure 4.5: PCR amplification for *Npt*II, *mtNpt*II, and *mtsfGFP* gene fragments was electrophoresed on a 1.2% (w/v) agarose gel. Lane 1 contains the 1 kb DNA ladder as a molecular weight marker, followed by lane 2 containing the negative control for the *Npt*II gene fragment. Lane 3 shows a successfully amplified *Npt*II fragment at 1402 bp. Lane 4 contains the negative control for the *mtNpt*II gene fragment, whereas lane 5 presents the amplified *mtNpt*II gene fragment at 922 bp. Lane 6 contains the negative control for the *mtsfGFP* (*Npt*II) gene fragment and lane 7 displays its amplified product at 809 bp. Lane 8 contains the negative control for the *mtsfGFP* (*mtNpt*II) gene fragment, followed by lane 9, showing its amplified product at 809 bp. Lane 10 contained a 100 bp DNA ladder. The distinct bands shown in the amplified lanes prove the successful and specific amplification of the targeted gene fragments under the applied PCR conditions.

4.3.3 Quantification of purified product

The amplified *Npt*II, ^{mt}*Npt*II and ^{mt}*sfGFP* gene products were purified using an EZ-10 Spin Column DNA Gel Extraction Kit. Table 4.1 shows the concentration and purity of the purified PCR product, measured using a NanoDrop spectrophotometer.

Table 4.1: Concentrations and purities of purified *Npt*II, ^{mt}*Npt*II and ^{mt}*sfGFP* fragments.

Purified PCR product	Concentration (ng/ μL)	Primary purity (A260/A280)	Secondary purity (A260/A230)
*Npt*II gene fragment	82.3	1.88	2.20
^{mt}*Npt*II gene fragment	105.9	1.85	2.20
^{mt}*sfGFP* (*Npt*II) gene fragment	60.3	1.85	2.19
^{mt}*sfGFP* (^{mt}*Npt*II) gene fragment	176.3	1.85	2.23

4.4 Overlap extension PCR for construct *Npt*II-*^{mt}sfGFP* and *^{mt}Npt*II-*^{mt}sfGFP* fusion gene

4.4.1 *Npt*II-*^{mt}sfGFP* and *^{mt}Npt*II-*^{mt}sfGFP* fusion gene

A)

B)
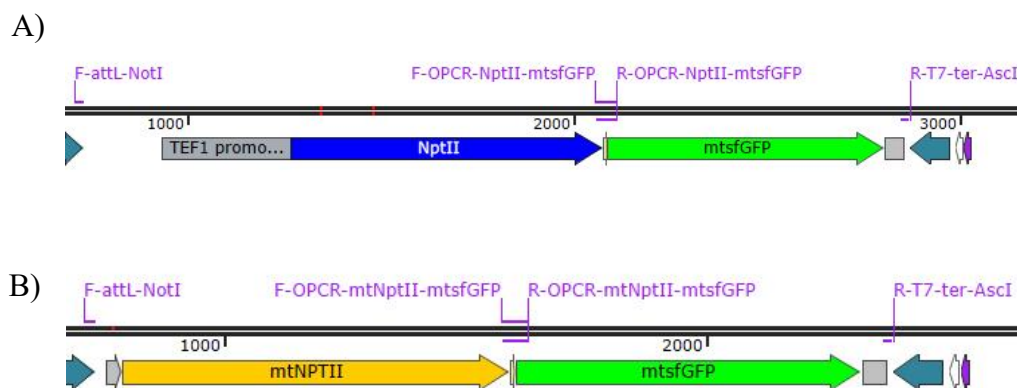
Figure 4.6: Schematic showing the sequence of the construction of the *Npt*II-*^{mt}sfGFP* and *^{mt}Npt*II-*^{mt}sfGFP* fusion genes. In part A, the *Npt*II-*^{mt}sfGFP* was fused by the overhanging regions, whole fusion gene sequences were amplified by end-to-end primers which is F-*att*L-*Not*I and R-T7-Ter-*Asc*I. In part B, *^{mt}Npt*II-*^{mt}sfGFP* was also fused by the overhanging regions, whole fusion gene sequences were amplified by end-to-end primers which is F-*att*L-*Not*I and R-T7-Ter-*Asc*I.

4.4.2 Result of agarose gel electrophoresis
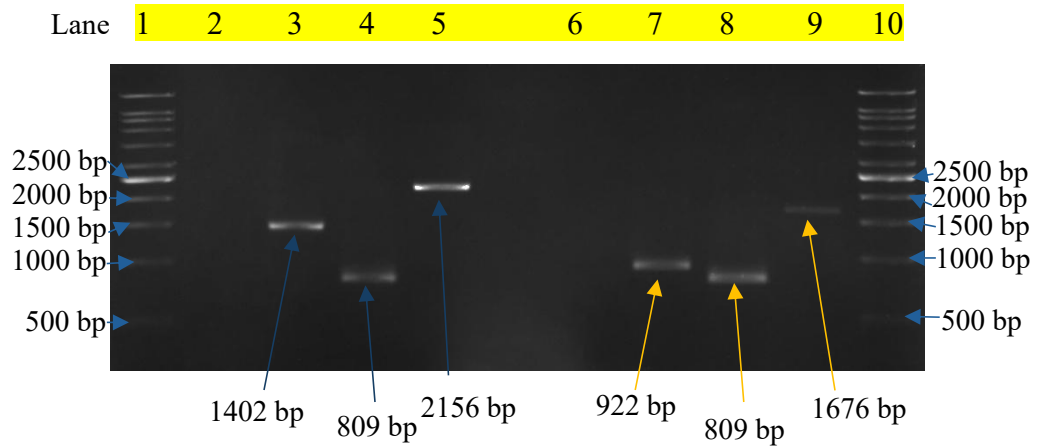
Figure 4.7: **Overlap extension PCR for *Npt*II-*^{mt}sfGFP* and *^{mt}Npt*II-*^{mt}sfGFP* fusion genes electrophoresed on 0.8% (w/v) agarose gel.** Lane 1 contains the 100 bp DNA ladder as a molecular weight marker, and lane 2 contains the negative control for the *Npt*II-*^{mt}sfGFP* fusion gene. Lanes 3 and 4 show positive controls for the *Npt*II gene fragment (1402 bp) and *^{mt}sfGFP*(*Npt*II) gene fragments (809 bp), respectively. Lane 5 shows the amplified *Npt*II-*^{mt}sfGFP* fusion gene product at 2156 bp. Lane 6 contains the negative control for the *^{mt}Npt*II-*^{mt}sfGFP* fusion gene; at the same time, lanes 7 and 8 show the positive controls for the *^{mt}Npt*II gene fragment (922 bp) and *^{mt}sfGFP* (*^{mt}Npt*II) gene fragment (809 bp) accordingly. Lane 9 shows the successfully amplified *^{mt}Npt*II-*^{mt}sfGFP* fusion gene at 1676 bp, while lane 10 contained the 1 kb DNA ladder. The presence of single, clearly fused gene bands in lanes 5 and 9 proves the successful and specific fusion under the optimized overlap extension PCR conditions, with the absence of amplification in negative controls.

4.4.3 Quantification of purified product

Two PCR fusion gene products were purified using different methodologies for DNA sequencing. *Npt*II-*^{mt}sfGFP* fusion gene product was purified using an EZ-10 Spin Column DNA Gel Extraction Kit. To reduce DNA loss during gel electrophoresis, ethanol precipitation was used to purify the *^{mt}Npt*II-*^{mt}sfGFP* fusion gene product. Table 4.2 shows the concentration and purity of the purified PCR product measured using a NanoDrop spectrophotometer.

Table 4.2: Concentrations and purites of purified *Npt*II-*^{mt}sfGFP* and *^{mt}Npt*II-*^{mt}sfGFP* fusion genes.

Purified PCR product	Concentration (ng/ μL)	Primary purity (A260/A280)	Secondary purity (A260/A230)
*Npt*II-*^{mt}sfGFP* fusion gene	109.7	1.86	2.16
*^{mt}Npt*II-*^{mt}sfGFP* fusion gene	624.7	1.41	1.82

4.5 Verification of the fusion gene product by DNA sequencing

The *Npt*II-*^{mt}sfGFP* fusion gene product was successfully verified by DNA sequencing. The results were analyzed with the *Npt*II and *^{mt}sfGFP* gene fragments, which were already verified by DNA sequencing by BLASTN. The alignment sequences are displayed in Appendices E and G.

CHAPTER 5

DISCUSSION

5.1 Justification for using overlap extension PCR in generate the fusion gene construction

In this project, overlap extension PCR (OE-PCR) was utilized to construct the *Npt*II-*mtsfGFP* and *mtNpt*II-*mtsfGFP* fusion gene constructs. Using custom-made overlapping primers, OE-PCR enables the precise and seamless fusion of two or more DNA fragments by creating the complementary overlap regions between the target genes (Bryksin and Matsumura, 2010). This method provides a highly specific fusion junction that remains in the open reading frame of the target fusion gene without adding unwanted sequences such as restriction enzyme recognition sites or linker residues (Hilgarth and Lanigan, 2020).

Compared with the traditional restriction enzyme-based cloning method, OE-PCR has significant advantages. Restriction enzyme cloning is highly dependent on the availability of suitable restriction sites in the vector and the insertion of DNA. This requires the introduction of an extra sequence that may affect protein function, without suitable restriction sites (Hashemabadi *et al*., 2025). In addition, multiple labor-intensive steps were included in the

restriction enzyme cloning such as digestion with restriction enzymes, fragment purification, ligation, and screening with colony PCR. The whole process requires several days and is costly, especially when rare restriction enzymes are chosen.

In contrast, OE-PCR is more rapid and cost-effective, without the need for preexisting or engineered restriction sites. Only three rounds of PCR are required for OE-PCR, including the individual fragment amplification and fusion step. Without the requirement of restriction sites, OE-PCR is highly versatile, allowing seamless joining of any DNA fragment by designing overlapping primers. Moreover, the linearized fusion product can also be directly cloned into the plasmid through recombination methods (Nelson and Fitch, 2012), thereby simplifying further workflow.

Furthermore, the troubleshooting of OE-PCR has focused on optimizing the PCR conditions and primer design; these issues are often straightforward to address. In comparison, restriction enzyme cloning method is more challenging due to multiple problems such as incomplete digestion or inefficient ligation, which are kore difficult and time-consuming to resolve (Makam *et al.*, 2018). Therefore, these factors led the selection of OE-PCR as the preferred method for construct *Npt*II/*^{mt}Npt*II-*^{mt}sfGFP* fusion gene**.**

5.2 Megaprimer design

To generate the *Npt*II/*ᵐᵗNpt*II-*ᵐᵗsfGFP* fusion gene by OE-PCR, overlapping regions were designed between the two DNA fragments, allowing them to serve as primers for each other in subsequent reactions using fragments with complementary overlapping ends. Megaprimers include overlapping sequences and target fragment sequences to facilitate gene fusion with specific binding and extension steps.

One of the most important design features of OE-PCR is that only the second gene fragment contains a stop codon (Hilgarth and Lanigan, 2020). By removing the stop codon in the first fusion gene fragment, the polymerase allowed to fully extend the overlap region without premature dissociation and incomplete fusion.

In this project, megaprimers with a length of 55 bp were designed to carry overlapping regions. Longer than the 18-24 bp typical primer. The increased length of the primer increases the melting temperature (T_m) and enhances binding specificity, providing sufficient homology to minimize nonspecific binding and increase the annealing specificity. According to Tyagi, Lai, and Duggleby (2004), some megaprimers exceeded 300 bp in some OE-PCR applications; therefore, the primer length is not strictly limiting, but good hybridization and extension need to be ensured to maximize fusion efficiency. The basis of the megaprimer is to contain the overlapping region between two

gene fragments and be long enough to generate a higher-specificity overlap region for further overlapping steps.

5.3 Challenges in single-primer testing

In figure 4.2, all megaprimers in this project were prone to secondary structure formation, including hairpin formation and self-annealing from the Oligonucleotide Properties Calculator analysis. The increased length of the megaprimer increases the possibility of an intramolecular secondary structure because of inverted repeats with a higher occurrence rate. These secondary structures can hinder the proper annealing of the target DNA and cause premature termination or reduce the efficiency of DNA polymerase extension (Hilgarth and Lanigan, 2020). Additionally, another potential problem reducing the amplification specificity is primer-dimer formation.

During single-primer testing with the corresponding DNA template, target-sized bands were clearly observed on agarose gel. This is the most significant issue and reveals a risk of false positives because there is a fake band at the target size. Hence, the main function of single-primer testing is to modify the PCR conditions until the fake bands disappear for further steps.

To mitigate these issues, the PCR conditions were modified step-by-step. The annealing temperatures were initially calculated using the Vayzme T_m calculator and optimized as required because of the high T_m (~69°C) of the

megaprimers. Annealing temperatures above 60°C yielded cleaner products but lower overall amplification efficiency. This makes the purification challenging. Thus, a suitable annealing temperature to compromise between specificity and yield was determined using PCR (Forloni, Liu and Wajapeyee, 2019).

The number of PCR thermocycles was optimized. Excessive cycles increase non-specific amplification and mutation rates. For most fragments, cycling was reduced from 35 to 25 cycles to eliminate false-positive results in single primer testing. Some cases need to be reduced to 20 cycles. The most challenging problem in cycling-number optimization is the comparison of the balance between productivity and specificity in the PCR amplification step.

The addition of DMSO was tested to destabilize the secondary structure by weakening the hydrogen bonds between complementary base pairs. The addition of DMSO to the PCR mixture effectively reduced non-specific amplification (Varadharajan and Parani, 2021) and interfered with target amplification in the later stages. Although T_m was lowered, the target PCR product could not be formed properly due to the presence of DMSO.

Then, based on the previous adjustment, the concentration of primer in the PCR mixture was reduced from 0.5 μM to 0.3 μM. Primer-dimer formation and non-specific amplification were successfully minimized, except for the *Npt*II gene fragment, which required further adjustment of the cycling number to achieve a clean band.

This part required substantial time and repeated adjustments before moving forward. Although sometimes the results of single-primer testing yielded clean agarose gel without any band formation, the modified parameter often caused a significant drop in PCR product yield. The low concentration of DNA makes purification challenging. Hence, the PCR product must be reduced because the concentration of the DNA template is less than a certain level. In addition, resolving this balance between specificity and yield was challenging with the modification of the annealing temperature, primer concentration, thermocycling number, and additives. This iterative process is time-consuming and laborious, indicating that high specificity does not guarantee smooth downstream amplification and purification.

5.4 Amplification of the target gene fragments

Based on the PCR conditions modified by single-primer testing, each gene fragment was amplified with the appropriate primer pairs listed in Tables 3.1, 3.2, and 3.3. Modifications to the thermocycling parameters and PCR mixtures were necessary for several fragments. except for the $^{mt}sfGFP$ gene fragment that fused to NptII, which required only a few adjustments.

For the NptII gene fragment, the significant differences in T_m between the forward (54.8°C) and reverse primers (66.2°C) resulted in a higher annealing temperature (58°C). This is the maximum annealing temperature with the target band size after PCR testing to balance specificity and yield. Thus, annealing

temperatures of 56°C and 58°C were maintained as a compromise between the specificity and yield. In this case, both PCR products showed nonspecific amplicons; thus, a lower annealing temperature was chosen. Based on the data from single-primer testing and further PCR condition modification, a light band was observed on the agarose gel. Although the target band was successfully obtained, the low DNA concentration product required re-amplification using the PCR product as the template.

^{mt}NptII gene fragment amplification was similar to NptII optimization with an annealing temperature of approximately 56°C by a series of gradient PCR tests. In contrast, the annealing temperature of the $^{mt}sfGFP$ gene fragment containing the ^{mt}NptII gene overlap region was approximately 60°C, without reducing the amplicon concentration and non-specific appearance. Because the concentration of the target DNA product was drastically reduced by the gel purification kit, the $^{mt}sfGFP$ gene fragment was reamplified using the purified product as the DNA template.

The elongation time was set according to the length of the gene fragment and efficiency of the Vayzme Taq polymerase. With 1 min per 1 kb DNA synthesis extension speed recommendation, 1:30 min for the NptII gene fragment and other gene fragments below 1 kb only required 1 min to synthesize.

5.5 Overlap extension PCR

The final OE-PCR step fused the purified individual gene fragments containing overlapping regions to construct the full-length *Npt*II-*^{mt}sfGFP* and *^{mt}Npt*II-*^{mt}sfGFP* fusion genes. Before the fusion step, the PCR products were purified to remove salts, enzymes, and buffers that may inhibit further reactions, and the concentration was measured by NanoDrop spectrophotometry.

Although the input DNA ratio for fusion was initially calculated using bioinformatics tools, such as the Vayzme Input DNA calculator (https://www.vazymeglobal.com/cetool/restructure.html), empirical modifications are still required based on product quality. The primary challenge was non-specific bands, which can be caused by suboptimal PCR conditions, including lower annealing temperature, excessive cycle number, or primers with sequence complementarity, leading to self- or cross-annealing problems.

To solve this problem, the PCR conditions and the ratio of the input DNA were modified. The annealing temperature for the fusion reactions was set to match the T_m of the overlap region. Fusion without primers at 60°C showed adequate annealing but may have reduced the efficiency; therefore, the temperature was carefully optimized. Although in single primer testing, the primer will not anneal with the template, in end-to-end PCR amplification, the primer may anneal with the residual template fragments and cause non-specific amplification. Based on this hypothesis, the DNA ratio was modified. The amount of the longer gene fragment in the fusion gene decreased, whereas the shorter gene fragment remained. The results showed that it efficiently reduced the non-specific bands, especially the residual template fragments. In end-to-

end PCR amplification with external primers followed by fusion, the annealing temperature were chosen as 59°C for the *Npt*II-*ᵐᵗsfGFP* fusion gene and 60°C for the *ᵐᵗNpt*II-*ᵐᵗsfGFP* fusion gene in the balance between yield and specificity to reduce nonspecific products.

5.6 Purification of *Npt*II-*ᵐᵗsfGFP* and *ᵐᵗNpt*II-*ᵐᵗsfGFP* fusion genes products

Two strategies were used to purify the fusion gene products. Ethanol precipitation was preferred for the *ᵐᵗNpt*II-*ᵐᵗsfGFP* fusion gene to minimize DNA loss compared with gel extraction, especially with the faint amplification band. In the absence of agarose gel electrophoresis, the DNA content of the PCR products was preserved.

To send the sequencing to 1ˢᵗ BASE, the quality of the DNA concentrations required exceeded 30 ng/µL for fragments longer than 1500 bp. The *Npt*II-*ᵐᵗsfGFP* fusion gene met this requirement, whereas *ᵐᵗNpt*II-*ᵐᵗsfGFP* did not result from gel residue contamination. The residual agarose gel in the phenol-chloroform-isoamyl alcohol supernatant extraction step led to a lower purity of the *ᵐᵗNpt*II-*ᵐᵗsfGFP* fusion gene product. This problem can be eliminated by repeating this extraction step until the buffy coat between the supernatant and phenol-chloroform-isoamyl alcohol is clear. This action maximizes DNA separation from the melting agarose gel.

5.7 DNA sequencing

Sequencing of the *Npt*II-*mtsfGFP* fusion gene product confirmed the successful fusion gene construction. The alignment results that matched the control sequence were confirmed by a senior. Owing to insufficient purity, sequencing of the *mtNpt*II-*mtsfGFP* fusion gene product was not feasible. Instead, it was cloned into *E. coli* and sequenced after plasmid extraction for correct fusion. The sequencing data and alignments are provided in the Appendix C.

5.8 The fusion pf selectable marker and reporter gene

In previous human cell studies, selectable marker and reporter fusion genes have been constructed. The green fluorescent protein (GFP) was successfully fused with the *neomycin* (*Neo*) resistance gene (Zhang, 2014). This study successfully overcame the common challenges in maintaining both gene functions by simple in-frame linking, without causing misfiling or steric hindrance. The native *GFP* stop codon was removed, enabling fusion with Neo, and the fusion genes were initiated by a mutated start codon. Approximately 17 amino acid linkers were inserted between the fusion genes to ensure the proper folding and activity of both protein domains. Through fusion gene testing, the researchers found that the fusion gene products exhibited green fluorescence 30% better than wild-type GFP, while the Neo function still retained resistance to aminoglycoside antibiotics. With a specific targeting sequence, the synthesis protein generated by the nucleus is transferred and functions in the mitochondria of the host. In related studies, GFP-Neo proteins, which fuse with the mitochondrial targeting leader sequence from mouse mitochondrial

transcription factor A, were able to target and localize inside the mitochondrial matrix of human cells (HeLa cells) (Yoon and Koob, 2008). This functional fusion protein proves that the fusion of the selectable marker and reporter gene is available and the localization can be confirmed by fluorescence microscopy and biochemical fractionation while maintaining the antibiotic resistance within mitochondria and showing mitochondrial network co-localization without cytosolic presence.

In studies of the ^{mt}NptII gene, researchers fused it with mitochondrial targeting sequences (MTS), which are derived from mitochondrial proteins and allow the import of the fusion protein from the cytosol into the mitochondria. The synthesis protein can be expressed in mitochondria (Boob et al., 2024). The specific localization of ^{mt}NptII was confirmed using fluorescence microscopy and biochemical fractionation.

Based on both studies, we discovered that the fusion of the selective markers and reporter genes is possible with proper folding and undergoes the function. In addition, the modified ^{mt}NptII can be expressed inside the host cell mitochondria using MTS. Thus, we planned to construct the NptII/^{mt}NptII-$^{mt}sfGFP$ fusion gene product based on this concept.

5.8.1 Advantages of selectable marker and reporter fusion genes

In this study, ^{mt}NptII genes served as dual-function genes. As an antibiotic resistance gene, it is possible to select transformed plasmids from non-

transformed plasmids. In addition, it serves as a transgene. Typically, the GOI is tagged with GFP to visualize its expression and level. In this project, ^{mt}NptII replaced the GOI and fused the $^{mt}sfGFP$ gene inside the pENTR cassette. With the engineered AMT system, which is able to transfer the GOI to integrate and express the host's mitochondria, it is able to generate functional genes. When the GOI is replaced by ^{mt}NptII and expressed inside the host's mitochondria, it can be visualized inside the host cell through fluorescence microscopy due to the criteria of subcellular localization. In addition, the synthesized protein can be verified using an anti-GFP tag to detect GFP with its fusion gene. Based on these methods, quantitative and qualitative methods were used. For example, we can observe that the result in the selectable medium is qualitative, and now quantitative is used to detect the fluorescent intensity. Through both sets of obtained data, we can eliminate the false positive result that is caused by the auxochrome of some host cells.

5.8.2 Disadvantages of selectable marker and reporter fusion genes

Due to the NptII/^{mt}NptII-$^{mt}sfGFP$ fusion gene, larger proteins require longer mRNA transcripts, increasing the energetic need and slowing down the transcriptional process. Similar to the translational process, longer mRNA require more time and resources to translate into proteins. Thus, larger fusion genes will decrease expression levels compared to smaller proteins (Lopes *et al*., 2021; Lemos, 2005).

In addition, larger proteins often have complex tertiary structures, creating a challenge for folding efficiency, and the misfolding or aggregation rate increases. Improper folding causes misfolded proteins to be degraded by the cellular quality control mechanism to reduce their net expression level (Lemos, 2005).

5.9 Trouble shooting

Secondary structures such as hairpin formation and primer dimers in PCR reactions decrease the amplification efficiency and cause non-specific bands to appear in the results. These structures were observed when the complementary primer sequences were self-annealing intramolecular or intermolecular. One effective method to solve this problem is to replace amino acids, which are encoded by problematic DNA sequences. These secondary structures can be removed by changing the nucleotide composition of megaprimers while maintaining protein function. For instance, the codons GAA and GAG encode the same amino acid glutamic acid. By substituting amino acids without changing the properties of the protein, the secondary structure can be destabilized and removed. Moreover, compared to chemical additives, such as DMSO, conservative amino acid substitutions provide a molecular solution to the problem beyond chemical mitigation (Bohórquez, Suárez and Patarroyo, 2017).

In addition, without modifying the megaprimers, we could increase the specificity of the primer in the PCR reaction. Before using the megaprimers to amplify each gene fragment, the DNA template was amplified using typical

primers. First, the gene fragments were amplified with normal primers to increase the specificity of the primer, and the PCR products were used as the DNA template with megaprimers containing the overhanging regions. Using the target gene fragments, PCR products as the DNA template instead of the original DNA fragments, increases the specificity of megaprimers and decreases the chance of forming secondary structures. This approach increases the rate of occurrence of specific band formation.

Furthermore, it is easier to construct fusion genes using traditional restriction enzyme methods. Due to the higher secondary structure formation rate using OE-PCR, using a restriction enzyme to recombine the target genes is a way to overcome this problem. In megaprimer design, we can add a linker between the two fusion genes. This linker includes the desired recognition sites, allowing the fusion of two genes using the traditional restriction enzyme method if OE-PCR fails. This method can reduce the time required to modify the optimum PCR reaction parameters based on these studies to remove secondary structures.

In summary, using synonymous codon changes in megaprimers, using target gene PCR products as the DNA template in the first amplification with megaprimers, and changing the fusion methods to restriction enzyme methods overcomes the secondary structure formation in PCR reactions.

5.10 Further studies

This project successfully fused two gene fragments using OE-PCR to create a dual-function fusion gene plasmid. The *Npt*II gene serves as a selectable marker and is able to resist antibiotics, such as kanamycin and neomycin (Numata *et al*., 2016). The ^{mt}NptII gene is modified for mitochondrial DNA expression compatibility. The $^{mt}sfGFP$ gene acts as a reporter gene, emitting green fluorescence that allows visualization of protein localization and dynamics inside the mitochondria (Yap,2025).

The fusion gene plasmid enables the selection of host cells and allows direct observation of the expressed fusion protein, thereby facilitating gene functional studies (Zhao *et al*., 2011). This construct can be cloned into an expression vector and transformed in yeast using a yeast-two-hybrid assay, aiding in interaction mapping and functional characterization. Finally, it can be transferred into agrobacteria using the yeast-two hybrid method to infect the target plant and perform functional gene studies. From this ^{mt}NptII-$^{mt}sfGFP$ fusion gene construct, the GOI, which is targeted to be expressed in the host's mitochondria, can be proven by observing protein synthesis and remaining inside the mitochondria by green fluorescence under UV or blue light.

CHAPTER 6

CONCLUSION

The purpose of this project was to fuse two sets of gene fragments to generate the *Npt*II-*^{mt}sfGFP* and *^{mt}Npt*II-*^{mt}sfGFP* fusion genes. Both fusion products were successfully verified by DNA sequencing, proving the project's aim. In this project, overlap extension PCR (OE-PCR) enabled precise and seamless fusion of DNA fragments through custom-designed overlapping primers and proper protocols. Compared to traditional restriction enzyme cloning, which requires multiple labor-intensive steps, such as digestion, ligation, and screening, OE-PCR condenses the process into a few PCR reactions. This technique reduces the time, cost, and simplicity of the workflow, while maintaining high specificity. However, the length of the megaprimers contributed to challenges, including secondary structure formation and single-primer amplification. Hence, single-primer testing was essential to prove critical in eliminating false positive results and to preliminarily optimize PCR reaction conditions. Parameters such as annealing temperature, primer concentration, cycle number, and additives such as DMSO make it difficult for the single primer to anneal to the template. For gene fragment amplification, PCR conditions were modified to accommodate differences in primer melting temperatures and template complexity. This balance was crucial in the specificity and amplification yield; increasing the specificity tended to complicate the subsequent purification step.

*Npt*II (1402 bp), ^{mt}*Npt*II (922 bp), and ^{mt}*sfGFP* (802 bp) gene fragments with overlapping regions were successfully amplified. Subsequent fusion and final amplification steps amplified the *Npt*II-^{mt}*sfGFP* (2156 bp) and ^{mt}*Npt*II-^{mt}*sfGFP* (1676 bp) fusion gene products by modifying the PCR conditions. Different purification methods were used to maximize DNA recovery, despite the low yield in different cases. Successful assembly of fusion gene products was verified by gel electrophoresis and DNA sequencing. The fusion gene product was cloned into *E. coli* by a senior lab member, and the selectable marker and reporter gene functions were combined into a plasmid. This dual-function plasmid is a valuable tool for future functional genetic studies. Lastly, this project highlights the essential role of the primer design and PCR condition optimization in achieving fusion gene construction by OE-PCR with technical demands.

REFERENCES

Andrews, B.T. *et al.* (2007) 'The rough energy landscape of superfolder GFP is linked to the chromophore', *Journal of Molecular Biology*, 373(2), pp. 476–490. Available at: https://doi.org/10.1016/j.jmb.2007.07.071.

Boob, A.G. *et al.* (2024) 'Design of diverse, functional mitochondrial targeting sequences across eukaryotic organisms using variational autoencoder', *bioRxiv (Cold Spring Harbor Laboratory)*. Available at: https://doi.org/10.1101/2024.08.28.610205.

Barondeau, D.P. *et al.* (2003) 'Mechanism and energetics of green fluorescent protein chromophore synthesis revealed by trapped intermediate structures', *Proceedings of the National Academy of Sciences*, 100(21), pp. 12111–12116. Available at: https://doi.org/10.1073/pnas.2133463100.

Barton, K.A. *et al.* (1983) 'Regeneration of intact tobacco plants containing full length copies of genetically engineered T-DNA, and transmission of T-DNA to R1 progeny', *Cell*, 32(4), pp. 1033–1043. Available at: https://doi.org/10.1016/0092-8674(83)90288-x.

Becker, D. *et al.* (1992) 'New plant binary vectors with selectable markers located proximal to the left T-DNA border', *Plant Molecular Biology*, 20(6), pp. 1195–1197. Available at: https://doi.org/10.1007/bf00028908.

Bernard, P. (1996) 'Positive selection of recombinant DNA by ccdB', *BioTechniques*, 21(2), pp. 320–323. Available at: https://doi.org/10.2144/96212pf01.

Bohórquez, H.J., Suárez, C.F. and Patarroyo, M.E. (2017) 'Mass & secondary structure propensity of amino acids explain their mutability and evolutionary replacements', *Scientific Reports*, 7(1). Available at: https://doi.org/10.1038/s41598-017-08041-7.

Bourras, S., Rouxel, T. and Meyer, M. (2015) '*Agrobacterium tumefaciens* gene transfer: how a plant pathogen hacks the nuclei of plant and nonplant organisms', *Phytopathology*, 105(10), pp. 1288–1301. Available at: https://doi.org/10.1094/phyto-12-14-0380-rvw.

Bryksin, A. and Matsumura, I. (2010) 'Overlap extension PCR cloning: a simple and reliable way to create recombinant plasmids', *BioTechniques*, 48(6), pp. 463–465. Available at: https://doi.org/10.2144/000113418.

Chang, S. *et al.* (2025) 'Malaysian journal of microbiology published by malaysian society for microbiology a vir binary vector system for *Agrobacterium*-mediated transient transformation in *Nicotiana benthamiana*', *Malaysian Journal of Microbiology*, 21(2). Available at: https://doi.org/10.21161/mjm.240632.

Curtis, M.D. and Grossniklaus, U. (2003) 'A gateway cloning vector set for high-throughput functional analysis of genes in planta', *Plant Physiology*, 133(2), pp. 462–469. Available at: https://doi.org/10.1104/pp.103.027979.

Das, A. *et al.* (2020) 'Estimation of *neomycin phosphotransferase*-II (*Npt*II) protein in vegetative and reproductive tissues of transgenic chickpea (Cicer arietinum L.) and biosafety perspectives', *Journal of Plant Biochemistry and Biotechnology*, 29(3), pp. 568–570. Available at: https://doi.org/10.1007/s13562-020-00562-z.

De Saeger, J. *et al.* (2021) '*Agrobacterium* strains and strain improvement: Present and outlook', *Biotechnology Advances*, 53, p.107677. Available at: https://doi.org/10.1016/j.biotechadv.2020.107677.

Forloni, M., Liu, A.Y. and Wajapeyee, N. (2019) 'Megaprimer polymerase chain reaction (PCR)-based mutagenesis', *Cold Spring Harbor Protocols*, 2019(6), p.097824. Available at: https://doi.org/10.1101/pdb.prot097824.

Gelvin, S.B. (2003) '*Agrobacterium*-Mediated plant transformation: the biology behind the "gene-jockeying" tool', *Microbiology and Molecular Biology Reviews*, 67(1), pp. 16–37. Available at: https://doi.org/10.1128/mmbr.67.1.16-37.2003.

Hashemabadi, M. *et al.* (2025) 'Restriction enzyme-free method for generating PCR product sticky ends in recombinant vector construction for cloning', *Discover Applied Sciences*, 7(5). Available at: https://doi.org/10.1007/s42452-025-06770-3.

Hilgarth, R.S. and Lanigan, T.M. (2020) 'Optimization of overlap extension PCR for efficient transgene construction', *MethodsX*, 7, p. 100759. Available at: https://doi.org/10.1016/j.mex.2019.12.001.

Hsu, S.-T.D., Blaser, G. and Jackson, S.E. (2009) 'The folding, stability and conformational dynamics of β-barrel fluorescent proteins', *Chemical Society Reviews*, 38(10), p. 2951. Available at: https://doi.org/10.1039/b908170b.

Hwang, H.-H., Yu, M. and Lai, E.-M. (2017) '*Agrobacterium*-mediated plant transformation: biology and applications', *The Arabidopsis Book*, 15(15), p. e0186. Available at: https://doi.org/10.1199/tab.0186.

Karimi, M., Depicker, A. and Hilson, P. (2007) 'Recombinational cloning with plant gateway vectors', *Plant Physiology*, 145(4), pp. 1144–1154. Available at: https://doi.org/10.1104/pp.107.106989.

Karimi, M., Inzé, D. and Depicker, A. (2002) 'GATEWAY™ vectors for *Agrobacterium*-mediated plant transformation', *Trends in Plant Science*, 7(5), pp. 193–195. Available at: https://doi.org/10.1016/s1360-1385(02)02251-3.

Kumar, A. and Pal, D. (2016) 'Green fluorescent protein and their applications in advance research', *Journal of Research in Engineering and Applied Sciences*, 01(01), pp. 42–46. Available at: https://doi.org/10.46565/jreas.2016.v01i01.007.

Kumar, A.U. and Ling, A.P.K. (2021) 'Gene introduction approaches in chloroplast transformation and its applications', *Journal of Genetic Engineering and Biotechnology*, 19(1). Available at: https://doi.org/10.1186/s43141-021-00255-7.

Law, S.S.Y. *et al.* (2022) 'Polymer-coated carbon nanotube hybrids with functional peptides for gene delivery into plant mitochondria', *Nature Communications*, 13(1), p. 2417. Available at: https://doi.org/10.1038/s41467-022-30185-y.

Lemos, B. (2005) 'Evolution of proteins and gene expression levels are coupled in Drosophila and are independently associated with mRNA abundance, protein length, and number of protein-protein interactions', *Molecular Biology and Evolution*, 22(5), pp. 1345–1354. Available at: https://doi.org/10.1093/molbev/msi122.

Lopes, I. *et al.* (2021) 'Gene size matters: an analysis of gene length in the human genome', *Frontiers in Genetics*, 12. Available at: https://doi.org/10.3389/fgene.2021.559998.

Lund, B.A., Leiros, H.-K.S. and Bjerga, G.K. (2014) 'A high-throughput, restriction-free cloning and screening strategy based on *ccd*B gene replacement,' *Microbial Cell Factories*, 13(1). https://doi.org/10.1186/1475-2859-13-38.

Makam, S. *et al.* (2018) 'An efficient method for integration of PCR fragments into adjacent or overlapping restriction sites during gene cloning', *3 Biotech*, 8(4), p. 197. Available at: https://doi.org/10.1007/s13205-018-1214-2.

Miki, B. and McHugh, S. (2004) 'Selectable marker genes in transgenic plants: applications, alternatives and biosafety', *Journal of Biotechnology*, 107(3), pp. 193–232. Available at: https://doi.org/10.1016/j.jbiotec.2003.10.011.

Mosey, M. *et al.* (2021) 'Methodological review of genetic engineering approaches for non-model algae', *Algal Research*, 54, p. 102221. Available at: https://doi.org/10.1016/j.algal.2021.102221.

Munaweera, T.I.K. *et al.* (2022) 'Modern plant biotechnology as a strategy in addressing climate change and attaining food security', *Agriculture & Food Security*, 11(1). Available at: https://doi.org/10.1186/s40066-022-00369-2.

Nelson, M.N. and David (2011) 'Overlap Extension PCR: an efficient method for transgene construction', *Methods in Molecular Biology*, pp. 459–470. Available at: https://doi.org/10.1007/978-1-61779-228-1_27.

Numata, K. *et al.* (2016) 'Direct introduction of *neomycin phosphotransferase* II protein into apple leaves to confer kanamycin resistance', *Plant Biotechnology*, 33(5), pp. 403–407. Available at: https://doi.org/10.5511/plantbiotechnology.16.0929a.

Orm , M. *et al.* (1996) 'Crystal structure of the *Aequorea victoria* green fluorescent protein', *Science*, 273(5280), pp. 1392–1395. Available at: https://doi.org/10.1126/science.273.5280.1392.

Pedelacq, J.-D. and Cabantous, S. (2019) 'Development and applications of superfolder and split fluorescent protein detection systems in biology', *International Journal of Molecular Sciences*, 20(14), p. 3479. Available at: https://doi.org/10.3390/ijms20143479.

Reece-Hoyes, J.S. and Walhout, A.J.M. (2018) 'Propagating gateway vectors', *Cold Spring Harbor Protocols*, 2018(1), p. pdb.prot094920. Available at: https://doi.org/10.1101/pdb.prot094920.

Soboleski, M.R., Oaks, J. and Halford, W.P. (2005) 'Green fluorescent protein is a quantitative reporter of gene expression in individual eukaryotic cells', *The FASEB Journal*, 19(3), pp. 440–442. Available at: https://doi.org/10.1096/fj.04-3180fje.

Strathdee, C.A., McLeod, M.R. and Underhill, T.M. (2000) 'Dominant positive and negative selection using luciferase, green fluorescent protein and B-Galactosidase Reporter gene fusions,' *BioTechniques*, 28(2), pp. 210–214. https://doi.org/10.2144/00282bm04.

Tyagi, R., Lai, R. and Duggleby, R.G. (2004) 'A new approach to "megaprimer" polymerase chain reaction mutagenesis without an intermediate gel purification step', *BMC Biotechnology*, 4(1), p. 2. Available at: https://doi.org/10.1186/1472-6750-4-2.

Tzfira, T. and Citovsky, V. (2008) *Agrobacterium: From Biology to Biotechnology*. Edited by T. Tzfira and V. Citovsky. New York, NY: Springer New York. Available at: https://doi.org/10.1007/978-0-387-72290-0.

Varadharajan, B. and Parani, M. (2021) 'DMSO and betaine significantly enhance the PCR amplification of ITS2 DNA barcodes from plants', *Genome*. Edited by J.S. Boatwright, 64(3), pp. 165–171. Available at: https://doi.org/10.1139/gen-2019-0221.

Wiese, M. and Bannister, A.J. (2020) 'Two genomes one cell: mitochondrial-nuclear coordination via epigenetic pathways', *Molecular Metabolism*, 38. Available at: https://doi.org/10.1016/j.molmet.2020.01.006.

Yap.K.H. (2025) *Construction of agrobacterium vector systems for nuclear and mitochondrial transformation in yeast*. Master thesis. Universiti Tunku Adbul Rahman.

Yoon, Y.G. and Koob, M.D. (2008) 'Selection by drug resistance proteins located in the mitochondria of mammalian cells', *Mitochondrion*, 8(5-6), pp. 345–351. Available at: https://doi.org/10.1016/j.mito.2008.07.004.

Zhang, M.L.J. (2014) 'Construction of functional *GFP-Neo* fusion protein by selection from a peptide library,' *Cloning & Transgenesis*, 03(03). https://doi.org/10.4172/2168-9849.1000131.

Zhao, H. *et al.* (2011) 'Fusion gene vectors allowing for simultaneous drug selection, cell labeling, and reporter assay in vitro and in vivo', *Analytical Chemistry*, 84(2), pp. 987–993. Available at: https://doi.org/10.1021/ac202541t.

Appendix A

Table A: The list of chemicals, reagents, and kits used in this final year project with their corresponding manufactures.

Chemicals	Manufacturers
100 bp DNA ladder	Norgen Biotek Corp.
100 % ethanol	CIM
2 x Taq master mix	Vazyme
Agarose	1st BASE Pte Ltd
Boric acid	Chemos
Low-melting agarose	BM Bio
Sodium chloride	Chem Soln
DyeAll™ staining solution	Gene All
Ethylenediaminetetraacetic (EDTA)	SIME Scientific
EZ-10 Spin Column DNA Gel Extraction Kit	Bio Basic Canada Inc.
Glacial acetic acid	Sigma-Aldrich
Phenol-chloroform-isoamyl alcohol	Nacalai Tesque Inc.
Sodium acetate (NaOAc)	Sigma-Aldrich
Tris base	Vivantis

VC 1 kb DNA ladder   Vivantis
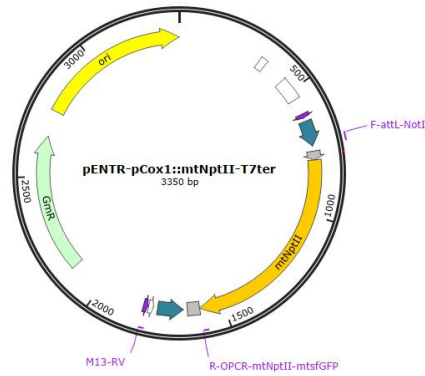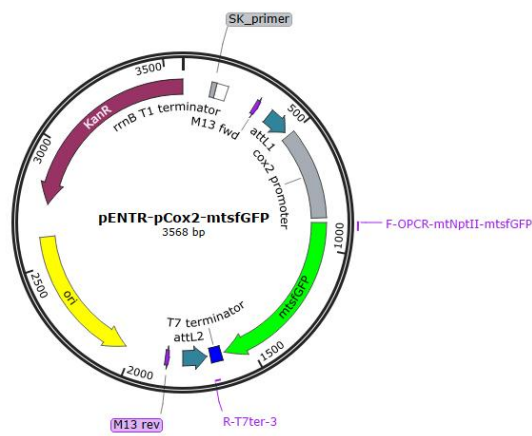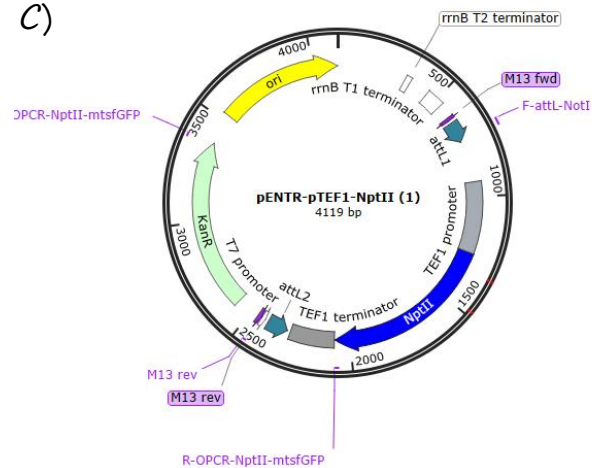
Figure B: The plasmid's map for each DNA template. In part A, this is the template for amplifying ^{mt}NptII, while part B, the plasmid is amplifying $^{mt}sfGFP$. In part C, the plasmid is used for amplifying the NptII gene.

Appendix C

A

	ore	Expect	Identities	Gaps	Strand
	29 bits(1315)	0.0	1315/1315(100%)	0/1315(0%)	Plus/Plus

```
Query  1     CGCCCCCTTCACCGGTGATGACGGTGAAAACCTCTGACACATGCAGCTCCCGGAGACGGT  60
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  1     CGCCCCCTTCACCGGTGATGACGGTGAAAACCTCTGACACATGCAGCTCCCGGAGACGGT  60

Query  61    CACAGCTTGTCTGTAAGCGGATGCCGGGAGCAGACAAGCCCGTCAGGGCGCGTCAGCGGG  120
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  61    CACAGCTTGTCTGTAAGCGGATGCCGGGAGCAGACAAGCCCGTCAGGGCGCGTCAGCGGG  120

Query  121   TGTTGGCGGGTGTCGGGGCTGGCTTAACTATGCGGCATCAGAGCAGATTGTACTGAGAGT  180
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  121   TGTTGGCGGGTGTCGGGGCTGGCTTAACTATGCGGCATCAGAGCAGATTGTACTGAGAGT  180

Query  181   GCACCATACCAAGGCCCAGAATACCCTCCTTGACAGTCTTGACGTGCGCAGCTCAGGGGC  240
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  181   GCACCATACCAAGGCCCAGAATACCCTCCTTGACAGTCTTGACGTGCGCAGCTCAGGGGC  240

Query  241   ATGATGTGACTGTCGCCCGTACATTTAGCCCATACATCCCCATGTATAATCATTTGCATC  300
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  241   ATGATGTGACTGTCGCCCGTACATTTAGCCCATACATCCCCATGTATAATCATTTGCATC  300

Query  301   CATACATTTTGATGGCCGCACGGCGCGAAGCAAAAATTACGGCTCCTCGCTGCAGACCTG  360
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  301   CATACATTTTGATGGCCGCACGGCGCGAAGCAAAAATTACGGCTCCTCGCTGCAGACCTG  360

Query  361   CGAGCAGGGAAACGCTCCCCTCACAGACGCGTTGAATTGTCCCCACGCCGCGCCCCTGTA  420
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  361   CGAGCAGGGAAACGCTCCCCTCACAGACGCGTTGAATTGTCCCCACGCCGCGCCCCTGTA  420

Query  421   GAGAAATATAAAAGGTTAGGATTTGCCACTGAGGTTCTTCTTTCATATACTTCCTTTTAA  480
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  421   GAGAAATATAAAAGGTTAGGATTTGCCACTGAGGTTCTTCTTTCATATACTTCCTTTTAA  480

Query  481   AATCTTGCTAGGATACAGTTCTCACATCACATCCGAACATAAACAACCATGGGTAAGGAA  540
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  481   AATCTTGCTAGGATACAGTTCTCACATCACATCCGAACATAAACAACCATGGGTAAGGAA  540

Query  541   AAGACTCACGTTTCGAGGCCGCGATTAAATTCCAACATGGATGCTGATTTATATGGGTAT  600
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  541   AAGACTCACGTTTCGAGGCCGCGATTAAATTCCAACATGGATGCTGATTTATATGGGTAT  600

Query  601   AAATGGGCTCGCGATAATGTCGGGCAATCAGGTGCGACAATCTATCGATTGTATGGGAAG  660
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  601   AAATGGGCTCGCGATAATGTCGGGCAATCAGGTGCGACAATCTATCGATTGTATGGGAAG  660

Query  661   CCCGATGCGCCAGAGTTGTTTCTGAAACATGGCAAAGGTAGCGTTGCCAATGATGTTACA  720
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  661   CCCGATGCGCCAGAGTTGTTTCTGAAACATGGCAAAGGTAGCGTTGCCAATGATGTTACA  720

Query  721   GATGAGATGGTCAGACTAAACTGGCTGACGGAATTTATGCCTCTTCCGACCATCAAGCAT  780
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  721   GATGAGATGGTCAGACTAAACTGGCTGACGGAATTTATGCCTCTTCCGACCATCAAGCAT  780

Query  781   TTTATCCGTACTCCTGATGATGCATGGTTACTCACCACTGCGATCCCCGGCAAAACAGCA  840
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  781   TTTATCCGTACTCCTGATGATGCATGGTTACTCACCACTGCGATCCCCGGCAAAACAGCA  840

Query  841   TTCCAGGTATTAGAAGAATATCCTGATTCAGGTGAAAATATTGTTGATGCGCTGGCAGTG  900
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  841   TTCCAGGTATTAGAAGAATATCCTGATTCAGGTGAAAATATTGTTGATGCGCTGGCAGTG  900

Query  901   TTCCTGCGCCGGTTGCATTCGATTCCTGTTTGTAATTGTCCTTTTAACAGCGATCGCGTA  960
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  901   TTCCTGCGCCGGTTGCATTCGATTCCTGTTTGTAATTGTCCTTTTAACAGCGATCGCGTA  960

Query  961   TTTCGTCTCGCTCAGGCGCAATCACGAATGAATAACGGTTTGGTTGATGCGAGTGATTTT  1020
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  961   TTTCGTCTCGCTCAGGCGCAATCACGAATGAATAACGGTTTGGTTGATGCGAGTGATTTT  1020

Query  1021  GATGACGAGCGTAATGGCTGGCCTGTTGAACAAGTCTGGAAAGAAATGCATAAGCTTTTG  1080
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  1021  GATGACGAGCGTAATGGCTGGCCTGTTGAACAAGTCTGGAAAGAAATGCATAAGCTTTTG  1080
```

```
Query  1021  GATGACGAGCGTAATGGCTGGCCTGTTGAACAAGTCTGGAAAGAAATGCATAAGCTTTTG  1080
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  1021  GATGACGAGCGTAATGGCTGGCCTGTTGAACAAGTCTGGAAAGAAATGCATAAGCTTTTG  1080

Query  1081  CCATTCTCACCGGATTCAGTCGTCACTCATGGTGATTTCTCACTTGATAACCTTATTTTT  1140
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  1081  CCATTCTCACCGGATTCAGTCGTCACTCATGGTGATTTCTCACTTGATAACCTTATTTTT  1140

Query  1141  GACGAGGGGAAATTAATAGGTTGTATTGATGTTGGACGAGTCGGAATCGCAGACCGATAC  1200
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  1141  GACGAGGGGAAATTAATAGGTTGTATTGATGTTGGACGAGTCGGAATCGCAGACCGATAC  1200

Query  1201  CAGGATCTTGCCATCCTATGGAACTGCCTCGGTGAGTTTTCTCCTTCATTACAGAAACGG  1260
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  1201  CAGGATCTTGCCATCCTATGGAACTGCCTCGGTGAGTTTTCTCCTTCATTACAGAAACGG  1260

Query  1261  CTTTTTCAAAAATATGGTATTGATAATCCTGATATGAATAAATTGCAGTTTCATT  1315
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  1261  CTTTTTCAAAAATATGGTATTGATAATCCTGATATGAATAAATTGCAGTTTCATT  1315
```

B)

Score		Expect	Identities	Gaps	Strand
1389 bits(752)		0.0	752/752(100%)	0/752(0%)	Plus/Plus

```
Query  1    AAGGTGAAGAATTATTCACTGGTGTTGTACCAATTTTAGTTGAATTAGATGGTGATGTTA  60
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  1    AAGGTGAAGAATTATTCACTGGTGTTGTACCAATTTTAGTTGAATTAGATGGTGATGTTA  60

Query  61   ATGGTCATAAATTTTCTGTAAGAGGTGAAGGTGAAGGTGATGCTACAAACGGAAAATTAA  120
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  61   ATGGTCATAAATTTTCTGTAAGAGGTGAAGGTGAAGGTGATGCTACAAACGGAAAATTAA  120

Query  121  CTTTAAAATTTATTTGTACTACTGGAAAATTACCTGTTCCATGACCAACATTAGTAACTA  180
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  121  CTTTAAAATTTATTTGTACTACTGGAAAATTACCTGTTCCATGACCAACATTAGTAACTA  180

Query  181  CTTTAACTTATGGTGTTCAATGTTTTTCAAGATATCCAGATCATATGAAAAGACATGATT  240
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  181  CTTTAACTTATGGTGTTCAATGTTTTTCAAGATATCCAGATCATATGAAAAGACATGATT  240

Query  241  TTTTCAAGAGTGCTATGCCTGAAGGTTATGTTCAAGAAAGAACAATTTCATTCAAAGATG  300
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  241  TTTTCAAGAGTGCTATGCCTGAAGGTTATGTTCAAGAAAGAACAATTTCATTCAAAGATG  300

Query  301  ATGGTACTTACAAGACTAGAGCTGAAGTAAAGTTTGAAGGTGATACTTTAGTTAATAGAA  360
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  301  ATGGTACTTACAAGACTAGAGCTGAAGTAAAGTTTGAAGGTGATACTTTAGTTAATAGAA  360

Query  361  TCGAATTAAAAGGTATTGATTTTAAAGAAGATGGAAACATTTTAGGACATAAATTAGAAT  420
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  361  TCGAATTAAAAGGTATTGATTTTAAAGAAGATGGAAACATTTTAGGACATAAATTAGAAT  420

Query  421  ACAACTTTAACTCACATAATGTTTATATTACAGCAGATAAACAAAAGAATGGAATCAAAG  480
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  421  ACAACTTTAACTCACATAATGTTTATATTACAGCAGATAAACAAAAGAATGGAATCAAAG  480

Query  481  CTAACTTCAAAATTAGACATAACGTAGAAGATGGTTCTGTTCAATTAGCAGATCATTATC  540
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  481  CTAACTTCAAAATTAGACATAACGTAGAAGATGGTTCTGTTCAATTAGCAGATCATTATC  540

Query  541  AACAAAATACTCCTATTGGTGATGGTCCTGTATTATTACCAGATAACCATTACTTATCTA  600
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  541  AACAAAATACTCCTATTGGTGATGGTCCTGTATTATTACCAGATAACCATTACTTATCTA  600

Query  601  CACAATCTGTATTATCTAAAGATCCTAACGAAAAGAGAGATCATATGGTATTATTAGAAT  660
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  601  CACAATCTGTATTATCTAAAGATCCTAACGAAAAGAGAGATCATATGGTATTATTAGAAT  660

Query  661  TTGTAACTGCTGCTGGTATTACACATGGTATGGATGAATTATACAAATAGTCCTAGCATA  720
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  661  TTGTAACTGCTGCTGGTATTACACATGGTATGGATGAATTATACAAATAGTCCTAGCATA  720

Query  721  ACCCCTTGGGGCCTCTAAACGGGTCTTGAGGG      752
            ||||||||||||||||||||||||||||||||
Sbjct  721  ACCCCTTGGGGCCTCTAAACGGGTCTTGAGGG      752
```

Figure C: Alignment of *Npt*II-*ᵐᵗsfGFP* fusion gene product. Part A was *Npt*II gene, part B was *ᵐᵗsfGFP* gene.
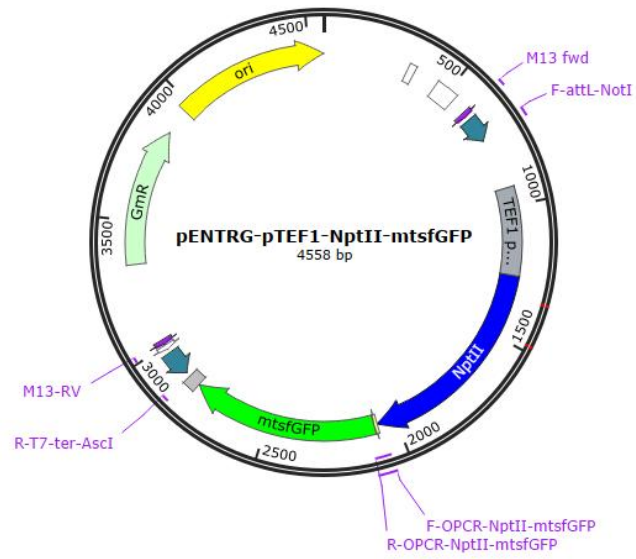
Figure D: Genetic map for *Npt*II-*mtsfGFP* fusion gene product. The fusion gene product had been successful clone inside *E. coli.*

Appendix E

Score	Expect	Identities	Gaps	Strand
3989 bits(2160)	0.0	2160/2160(100%)	0/2160(0%)	Plus/Plus

```
Query  1     CAACTTTGTACAAAAAAGCAGGCTCCGCGGCCGCCCCCTTCACCGGTGATGACGGTGAAA  60
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  1     CAACTTTGTACAAAAAAGCAGGCTCCGCGGCCGCCCCCTTCACCGGTGATGACGGTGAAA  60

Query  61    ACCTCTGACACATGCAGCTCCCGGAGACGGTCACAGCTTGTCTGTAAGCGGATGCCGGGA  120
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  61    ACCTCTGACACATGCAGCTCCCGGAGACGGTCACAGCTTGTCTGTAAGCGGATGCCGGGA  120

Query  121   GCAGACAAGCCCGTCAGGGCGCGTCAGCGGGTGTTGGCGGGTGTCGGGGCTGGCTTAACT  180
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  121   GCAGACAAGCCCGTCAGGGCGCGTCAGCGGGTGTTGGCGGGTGTCGGGGCTGGCTTAACT  180

Query  181   ATGCGGCATCAGAGCAGATTGTACTGAGAGTGCACCATACCAAGGCCCAGAATACCCTCC  240
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  181   ATGCGGCATCAGAGCAGATTGTACTGAGAGTGCACCATACCAAGGCCCAGAATACCCTCC  240

Query  241   TTGACAGTCTTGACGTGCGCAGCTCAGGGGCATGATGTGACTGTCGCCCGTACATTTAGC  300
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  241   TTGACAGTCTTGACGTGCGCAGCTCAGGGGCATGATGTGACTGTCGCCCGTACATTTAGC  300

Query  301   CCATACATCCCCATGTATAATCATTTGCATCCATACATTTTGATGGCCGCACGGCGCGAA  360
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  301   CCATACATCCCCATGTATAATCATTTGCATCCATACATTTTGATGGCCGCACGGCGCGAA  360

Query  361   GCAAAAATTACGGCTCCTCGCTGCAGACCTGCGAGCAGGGAAACGCTCCCCTCACAGACG  420
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  361   GCAAAAATTACGGCTCCTCGCTGCAGACCTGCGAGCAGGGAAACGCTCCCCTCACAGACG  420

Query  421   CGTTGAATTGTCCCCACGCCGCGCCCCTGTAGAGAAATATAAAAGGTTAGGATTTGCCAC  480
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  421   CGTTGAATTGTCCCCACGCCGCGCCCCTGTAGAGAAATATAAAAGGTTAGGATTTGCCAC  480

Query  481   TGAGGTTCTTCTTTCATATACTTCCTTTTAAAATCTTGCTAGGATACAGTTCTCACATCA  540
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  481   TGAGGTTCTTCTTTCATATACTTCCTTTTAAAATCTTGCTAGGATACAGTTCTCACATCA  540

Query  541   CATCCGAACATAAACAACCATGGGTAAGGAAAAGACTCACGTTTCGAGGCCGCGATTAAA  600
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  541   CATCCGAACATAAACAACCATGGGTAAGGAAAAGACTCACGTTTCGAGGCCGCGATTAAA  600

Query  601   TTCCAACATGGATGCTGATTTATATGGGTATAAATGGGCTCGCGATAATGTCGGGCAATC  660
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  601   TTCCAACATGGATGCTGATTTATATGGGTATAAATGGGCTCGCGATAATGTCGGGCAATC  660

Query  661   AGGTGCGACAATCTATCGATTGTATGGGAAGCCCGATGCGCCAGAGTTGTTTCTGAAACA  720
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  661   AGGTGCGACAATCTATCGATTGTATGGGAAGCCCGATGCGCCAGAGTTGTTTCTGAAACA  720

Query  721   TGGCAAAGGTAGCGTTGCCAATGATGTTACAGATGAGATGGTCAGACTAAACTGGCTGAC  780
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  721   TGGCAAAGGTAGCGTTGCCAATGATGTTACAGATGAGATGGTCAGACTAAACTGGCTGAC  780

Query  781   GGAATTTATGCCTCTTCCGACCATCAAGCATTTTATCCGTACTCCTGATGATGCATGGTT  840
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  781   GGAATTTATGCCTCTTCCGACCATCAAGCATTTTATCCGTACTCCTGATGATGCATGGTT  840

Query  841   ACTCACCACTGCGATCCCCGGCAAAACAGCATTCCAGGTATTAGAAGAATATCCTGATTC  900
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  841   ACTCACCACTGCGATCCCCGGCAAAACAGCATTCCAGGTATTAGAAGAATATCCTGATTC  900

Query  901   AGGTGAAAATATTGTTGATGCGCTGGCAGTGTTCCTGCGCCGGTTGCATTCGATTCCTGT  960
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  901   AGGTGAAAATATTGTTGATGCGCTGGCAGTGTTCCTGCGCCGGTTGCATTCGATTCCTGT  960

Query  961   TTGTAATTGTCCTTTTAACAGCGATCGCGTATTTCGTCTCGCTCAGGCGCAATCACGAAT  1020
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  961   TTGTAATTGTCCTTTTAACAGCGATCGCGTATTTCGTCTCGCTCAGGCGCAATCACGAAT  1020
```

```
Query  1021  GAATAACGGTTTGGTTGATGCGAGTGATTTTGATGACGAGCGTAATGGCTGGCCTGTTGA  1080
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  1021  GAATAACGGTTTGGTTGATGCGAGTGATTTTGATGACGAGCGTAATGGCTGGCCTGTTGA  1080

Query  1081  ACAAGTCTGGAAAGAAATGCATAAGCTTTTGCCATTCTCACCGGATTCAGTCGTCACTCA  1140
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  1081  ACAAGTCTGGAAAGAAATGCATAAGCTTTTGCCATTCTCACCGGATTCAGTCGTCACTCA  1140

Query  1141  TGGTGATTTCTCACTTGATAACCTTATTTTTGACGAGGGGAAATTAATAGGTTGTATTGA  1200
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  1141  TGGTGATTTCTCACTTGATAACCTTATTTTTGACGAGGGGAAATTAATAGGTTGTATTGA  1200

Query  1201  TGTTGGACGAGTCGGAATCGCAGACCGATACCAGGATCTTGCCATCCTATGGAACTGCCT  1260
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  1201  TGTTGGACGAGTCGGAATCGCAGACCGATACCAGGATCTTGCCATCCTATGGAACTGCCT  1260

Query  1261  CGGTGAGTTTTCTCCTTCATTACAGAAACGGCTTTTTCAAAAATATGGTATTGATAATCC  1320
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  1261  CGGTGAGTTTTCTCCTTCATTACAGAAACGGCTTTTTCAAAAATATGGTATTGATAATCC  1320

Query  1321  TGATATGAATAAATTGCAGTTTCATTTGATGCTCGATGAGTTTTTCGGATCCGAGACCAT  1380
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  1321  TGATATGAATAAATTGCAGTTTCATTTGATGCTCGATGAGTTTTTCGGATCCGAGACCAT  1380

Query  1381  GTCTAAAGGTGAAGAATTATTCACTGGTGTTGTACCAATTTTAGTTGAATTAGATGGTGA  1440
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  1381  GTCTAAAGGTGAAGAATTATTCACTGGTGTTGTACCAATTTTAGTTGAATTAGATGGTGA  1440

Query  1441  TGTTAATGGTCATAAATTTTCTGTAAGAGGTGAAGGTGAAGGTGATGCTACAAACGGAAA  1500
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  1441  TGTTAATGGTCATAAATTTTCTGTAAGAGGTGAAGGTGAAGGTGATGCTACAAACGGAAA  1500

Query  1501  ATTAACTTTAAAATTTATTTGTACTACTGGAAAATTACCTGTTCCATGACCAACATTAGT  1560
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  1501  ATTAACTTTAAAATTTATTTGTACTACTGGAAAATTACCTGTTCCATGACCAACATTAGT  1560

Query  1021  GAATAACGGTTTGGTTGATGCGAGTGATTTTGATGACGAGCGTAATGGCTGGCCTGTTGA  1080
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  1021  GAATAACGGTTTGGTTGATGCGAGTGATTTTGATGACGAGCGTAATGGCTGGCCTGTTGA  1080

Query  1081  ACAAGTCTGGAAAGAAATGCATAAGCTTTTGCCATTCTCACCGGATTCAGTCGTCACTCA  1140
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  1081  ACAAGTCTGGAAAGAAATGCATAAGCTTTTGCCATTCTCACCGGATTCAGTCGTCACTCA  1140

Query  1141  TGGTGATTTCTCACTTGATAACCTTATTTTTGACGAGGGGAAATTAATAGGTTGTATTGA  1200
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  1141  TGGTGATTTCTCACTTGATAACCTTATTTTTGACGAGGGGAAATTAATAGGTTGTATTGA  1200

Query  1201  TGTTGGACGAGTCGGAATCGCAGACCGATACCAGGATCTTGCCATCCTATGGAACTGCCT  1260
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  1201  TGTTGGACGAGTCGGAATCGCAGACCGATACCAGGATCTTGCCATCCTATGGAACTGCCT  1260

Query  1261  CGGTGAGTTTTCTCCTTCATTACAGAAACGGCTTTTTCAAAAATATGGTATTGATAATCC  1320
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  1261  CGGTGAGTTTTCTCCTTCATTACAGAAACGGCTTTTTCAAAAATATGGTATTGATAATCC  1320

Query  1321  TGATATGAATAAATTGCAGTTTCATTTGATGCTCGATGAGTTTTTCGGATCCGAGACCAT  1380
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  1321  TGATATGAATAAATTGCAGTTTCATTTGATGCTCGATGAGTTTTTCGGATCCGAGACCAT  1380

Query  1381  GTCTAAAGGTGAAGAATTATTCACTGGTGTTGTACCAATTTTAGTTGAATTAGATGGTGA  1440
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  1381  GTCTAAAGGTGAAGAATTATTCACTGGTGTTGTACCAATTTTAGTTGAATTAGATGGTGA  1440

Query  1441  TGTTAATGGTCATAAATTTTCTGTAAGAGGTGAAGGTGAAGGTGATGCTACAAACGGAAA  1500
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  1441  TGTTAATGGTCATAAATTTTCTGTAAGAGGTGAAGGTGAAGGTGATGCTACAAACGGAAA  1500

Query  1501  ATTAACTTTAAAATTTATTTGTACTACTGGAAAATTACCTGTTCCATGACCAACATTAGT  1560
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  1501  ATTAACTTTAAAATTTATTTGTACTACTGGAAAATTACCTGTTCCATGACCAACATTAGT  1560
```

```
Query  1561  AACTACTTTAACTTATGGTGTTCAATGTTTTTCAAGATATCCAGATCATATGAAAAGACA  1620
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  1561  AACTACTTTAACTTATGGTGTTCAATGTTTTTCAAGATATCCAGATCATATGAAAAGACA  1620

Query  1621  TGATTTTTTCAAGAGTGCTATGCCTGAAGGTTATGTTCAAGAAAGAACAATTTCATTCAA  1680
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  1621  TGATTTTTTCAAGAGTGCTATGCCTGAAGGTTATGTTCAAGAAAGAACAATTTCATTCAA  1680

Query  1681  AGATGATGGTACTTACAAGACTAGAGCTGAAGTAAAGTTTGAAGGTGATACTTTAGTTAA  1740
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  1681  AGATGATGGTACTTACAAGACTAGAGCTGAAGTAAAGTTTGAAGGTGATACTTTAGTTAA  1740

Query  1741  TAGAATCGAATTAAAAGGTATTGATTTTAAAGAAGATGGAAACATTTTAGGACATAAATT  1800
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  1741  TAGAATCGAATTAAAAGGTATTGATTTTAAAGAAGATGGAAACATTTTAGGACATAAATT  1800

Query  1801  AGAATACAACTTTAACTCACATAATGTTTATATTACAGCAGATAAACAAAAGAATGGAAT  1860
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  1801  AGAATACAACTTTAACTCACATAATGTTTATATTACAGCAGATAAACAAAAGAATGGAAT  1860

Query  1861  CAAAGCTAACTTCAAAATTAGACATAACGTAGAAGATGGTTCTGTTCAATTAGCAGATCA  1920
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  1861  CAAAGCTAACTTCAAAATTAGACATAACGTAGAAGATGGTTCTGTTCAATTAGCAGATCA  1920

Query  1921  TTATCAACAAAATACTCCTATTGGTGATGGTCCTGTATTATTACCAGATAACCATTACTT  1980
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  1921  TTATCAACAAAATACTCCTATTGGTGATGGTCCTGTATTATTACCAGATAACCATTACTT  1980

Query  1981  ATCTACACAATCTGTATTATCTAAAGATCCTAACGAAAAGAGAGATCATATGGTATTATT  2040
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  1981  ATCTACACAATCTGTATTATCTAAAGATCCTAACGAAAAGAGAGATCATATGGTATTATT  2040

Query  2041  AGAATTTGTAACTGCTGCTGGTATTACACATGGTATGGATGAATTATACAAATAGTCCTA  2100
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  2041  AGAATTTGTAACTGCTGCTGGTATTACACATGGTATGGATGAATTATACAAATAGTCCTA  2100

Query  2101  GCATAACCCCTTGGGGCCTCTAAACGGGTCTTGAGGGGTTTTTTGAAGGGTGGGCGCGCC  2160
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  2101  GCATAACCCCTTGGGGCCTCTAAACGGGTCTTGAGGGGTTTTTTGAAGGGTGGGCGCGCC  2160
```

Figure E: Alignment of *Npt*II-*ᵐᵗsfGFP* fusion gene product.
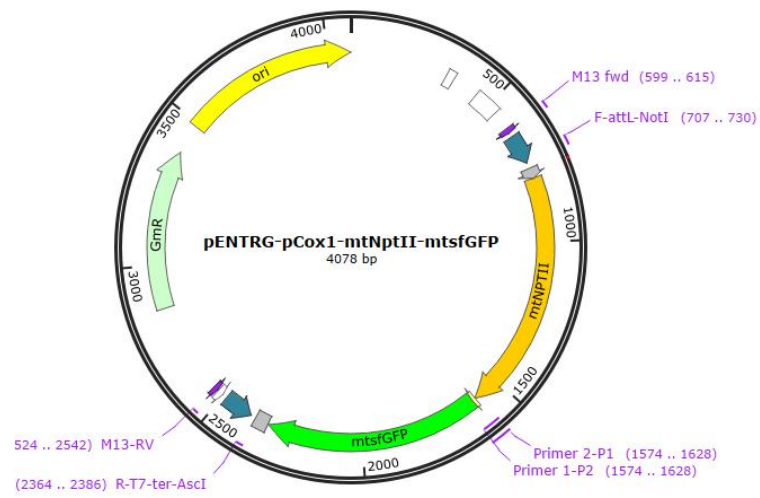
Figure F: Genetic map for ^{mt}*Npt*II-^{mt}*sfGFP* fusion gene product. The fusion gene product had been successful clone inside *E. coli.*

Appendix G

Score		Expect	Identities	Gaps	Strand
3103 bits(1680)		0.0	1680/1680(100%)	0/1680(0%)	Plus/Plus

```
Query  1     CAACTTTGTACAAAAAAGCAGGCTCCGCGGCCGCCCCCTTCACCGTATTGATATAAGTAA  60
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  1     CAACTTTGTACAAAAAAGCAGGCTCCGCGGCCGCCCCCTTCACCGTATTGATATAAGTAA  60

Query  61    TAGATATAATAATAATATTATGGGTAAGGAAAAGACTCACGTTTCTCGTCCTCGTTTAAA  120
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  61    TAGATATAATAATAATATTATGGGTAAGGAAAAGACTCACGTTTCTCGTCCTCGTTTAAA  120

Query  121   TTCTAATATGGATGCTGATTTATATGGTTATAAGTGAGCTCGTGATAATGTTGGTCAATC  180
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  121   TTCTAATATGGATGCTGATTTATATGGTTATAAGTGAGCTCGTGATAATGTTGGTCAATC  180

Query  181   TGGTGCTACTATTTATCGTTTATATGGTAAGCCTGACGCTCCTGAATTGTTTTTGAAGCA  240
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  181   TGGTGCTACTATTTATCGTTTATATGGTAAGCCTGACGCTCCTGAATTGTTTTTGAAGCA  240

Query  241   TGGTAAAGGTTCTGTTGCTAATGATGTTACTGATGAAATGGTCCGTTTAAATTGATTAAC  300
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  241   TGGTAAAGGTTCTGTTGCTAATGATGTTACTGATGAAATGGTCCGTTTAAATTGATTAAC  300

Query  301   TGAATTTATGCCTTTACCTACTATTAAACATTTTATTCGTACTCCTGATGATGCTTGATT  360
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  301   TGAATTTATGCCTTTACCTACTATTAAACATTTTATTCGTACTCCTGATGATGCTTGATT  360

Query  361   GTTAACTACTGCTATTCCTGGTAAGACTGCTTTTCAAGTCTTAGAAGAATATCCTGATTC  420
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  361   GTTAACTACTGCTATTCCTGGTAAGACTGCTTTTCAAGTCTTAGAAGAATATCCTGATTC  420

Query  421   TGGTGAAAATATTGTTGATGCTTTAGCTGTCTTTTTACGTCGTTTACATTCTATTCCTGT  480
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  421   TGGTGAAAATATTGTTGATGCTTTAGCTGTCTTTTTACGTCGTTTACATTCTATTCCTGT  480

Query  481   TTGTAATTGTCCTTTTAATTCTGATCGTGTTTTTCGTTTAGCTCAAGCTCAATCTCGTAT  540
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  481   TTGTAATTGTCCTTTTAATTCTGATCGTGTTTTTCGTTTAGCTCAAGCTCAATCTCGTAT  540

Query  541   GAACAATGGTTTAGTTGATGCTTCTGATTTCGACGATGACCGTAACGGTTGACCTGTTGA  600
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  541   GAACAATGGTTTAGTTGATGCTTCTGATTTCGACGATGACCGTAACGGTTGACCTGTTGA  600

Query  601   ACAAGTTTGAAAGGAAATGCATAAGTTATTACCTTTTTCTCCTGATTCTGTTGTTACTCA  660
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  601   ACAAGTTTGAAAGGAAATGCATAAGTTATTACCTTTTTCTCCTGATTCTGTTGTTACTCA  660

Query  661   TGGTGACTTCTCTTTAGACAATTTAATTTTCGATGAAGGTAAGTTGATCGGTTGTATCGA  720
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  661   TGGTGACTTCTCTTTAGACAATTTAATTTTCGATGAAGGTAAGTTGATCGGTTGTATCGA  720

Query  721   TGTCGGTCGTGTCGGTATCGCTGACCGTTACCAAGACTTGGCTATCTTGTGAAACTGTTT  780
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  721   TGTCGGTCGTGTCGGTATCGCTGACCGTTACCAAGACTTGGCTATCTTGTGAAACTGTTT  780

Query  781   AGGTGAGTTCTCTCCTTCCTTGCAGAAGCGTTTGTTCCAAAAGTATGGTATCGATAACCC  840
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  781   AGGTGAGTTCTCTCCTTCCTTGCAGAAGCGTTTGTTCCAAAAGTATGGTATCGATAACCC  840

Query  841   TGACATGAATAAATTACAGTTCCACTTGATGTTGGACGAGTTCTTCGGATCCGAGACCAT  900
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  841   TGACATGAATAAATTACAGTTCCACTTGATGTTGGACGAGTTCTTCGGATCCGAGACCAT  900

Query  901   GTCTAAAGGTGAAGAATTATTCACTGGTGTTGTACCAATTTTAGTTGAATTAGATGGTGA  960
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  901   GTCTAAAGGTGAAGAATTATTCACTGGTGTTGTACCAATTTTAGTTGAATTAGATGGTGA  960

Query  961   TGTTAATGGTCATAAATTTTCTGTAAGAGGTGAAGGTGAAGGTGATGCTACAAACGGAAA  1020
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  961   TGTTAATGGTCATAAATTTTCTGTAAGAGGTGAAGGTGAAGGTGATGCTACAAACGGAAA  1020

Query  1021  ATTAACTTTAAAATTTATTTGTACTACTGGAAAATTACCTGTTCCATGACCAACATTAGT  1080
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  1021  ATTAACTTTAAAATTTATTTGTACTACTGGAAAATTACCTGTTCCATGACCAACATTAGT  1080

Query  1081  AACTACTTTAACTTATGGTGTTCAATGTTTTTCAAGATATCCAGATCATATGAAAAGACA  1140
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  1081  AACTACTTTAACTTATGGTGTTCAATGTTTTTCAAGATATCCAGATCATATGAAAAGACA  1140
```

```
Query  1141  TGATTTTTTCAAGAGTGCTATGCCTGAAGGTTATGTTCAAGAAAGAACAATTTCATTCAA  1200
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  1141  TGATTTTTTCAAGAGTGCTATGCCTGAAGGTTATGTTCAAGAAAGAACAATTTCATTCAA  1200

Query  1201  AGATGATGGTACTTACAAGACTAGAGCTGAAGTAAAGTTTGAAGGTGATACTTTAGTTAA  1260
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  1201  AGATGATGGTACTTACAAGACTAGAGCTGAAGTAAAGTTTGAAGGTGATACTTTAGTTAA  1260

Query  1261  TAGAATCGAATTAAAAGGTATTGATTTTAAAGAAGATGGAAACATTTTAGGACATAAATT  1320
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  1261  TAGAATCGAATTAAAAGGTATTGATTTTAAAGAAGATGGAAACATTTTAGGACATAAATT  1320

Query  1321  AGAATACAACTTTAACTCACATAATGTTTATATTACAGCAGATAAACAAAAGAATGGAAT  1380
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  1321  AGAATACAACTTTAACTCACATAATGTTTATATTACAGCAGATAAACAAAAGAATGGAAT  1380

Query  1381  CAAAGCTAACTTCAAAATTAGACATAACGTAGAAGATGGTTCTGTTCAATTAGCAGATCA  1440
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  1381  CAAAGCTAACTTCAAAATTAGACATAACGTAGAAGATGGTTCTGTTCAATTAGCAGATCA  1440

Query  1441  TTATCAACAAAATACTCCTATTGGTGATGGTCCTGTATTATTACCAGATAACCATTACTT  1500
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  1441  TTATCAACAAAATACTCCTATTGGTGATGGTCCTGTATTATTACCAGATAACCATTACTT  1500

Query  1501  ATCTACACAATCTGTATTATCTAAAGATCCTAACGAAAAGAGAGATCATATGGTATTATT  1560
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  1501  ATCTACACAATCTGTATTATCTAAAGATCCTAACGAAAAGAGAGATCATATGGTATTATT  1560

Query  1561  AGAATTTGTAACTGCTGCTGGTATTACACATGGTATGGATGAATTATACAAATAGTCCTA  1620
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  1561  AGAATTTGTAACTGCTGCTGGTATTACACATGGTATGGATGAATTATACAAATAGTCCTA  1620

Query  1621  GCATAACCCCTTGGGGCCTCTAAACGGGTCTTGAGGGGTTTTTTGAAGGGTGGGCGCGCC  1680
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  1621  GCATAACCCCTTGGGGCCTCTAAACGGGTCTTGAGGGGTTTTTTGAAGGGTGGGCGCGCC  1680
```

Figure G: Alignment of *Npt*II-*mtsfGFP* fusion gene product.

Date:
_____

PERMISSION SHEET

It is hereby certified that CHENG YANG YAN (ID No: 22ADB07553) has completed this final year project report entitled "CONSTRUCTION OF *Npt*II-^{mt}*sfGFP* FUSION GENE CONSTRUCT BY OVERLAP EXTENSION PCR" under the supervision of Prof. Dr. Wong Hann Ling from the Department of Biological Science, Faculty of Science.

I hereby give permission to the University to upload the softcopy of my final year project dissertation in pdf format into the UTAR Institutional Repository, which may be made accessible to the UTAR community and public.

I hereby give permission to my supervisor to write and prepare manuscript of these research findings for publishing in any form, if I did not prepare it within six (6) months' time from this date provided that my name is included as one of the authors for this article. Arrangement of the name depends on my supervisor.

Yours truly,

_____

(CHENG YANG YAN)

Universiti Tunku Abdul Rahman			
Form Title : Supervisor's Comments on Originality Report Generated by Turnitin for Submission of Final Year Project Report (for Undergraduate Programmes)			
Form Number: FM-IAD-005	Rev No.: 1	Effective Date: 3/10/2019	Page No.: 1of 1

FACULTY OF SCIENCE

Full Name(s) of Candidate(s)	Chang Yang Yan
ID Number(s)	22ADB07353
Programme / Course	BACHELOR OF SCIENCE (HONOURS) BIOTECHNOLOGY
Title of Final Year Project	CONSTRUCTION OF NptII-mtsfGFP FUSION GENE CONSTRUCT BY OVERLAP EXTENSION PCR

Similarity	Supervisor's Comments (Compulsory if parameters of originality exceeds the limits approved by UTAR)
Overall similarity index: __14__ **% Similarity by source** Internet Sources: __8__ % Publications: __6__ % Student Papers: __11__ %	
Number of individual sources listed of more than 3% similarity: _0_	

Parameters of originality required and limits approved by UTAR are as follows::
(i) Overall similarity index is 20% and below, and
(ii) Matching of individual sources listed must be less than 3% each, and
(iii) Matching texts in continuous block must not exceed 8 words
Note: Parameters (i) – (ii) shall exclude quotes, bibliography and text matches which are less than 8 words.

Note Supervisor/Candidate(s) is/are required to provide softcopy of full set of the originality report to Faculty/Institute

Based on the above results, I hereby declare that I am satisfied with the originality of the Final Year Project Report submitted by my student(s) as named above.

Signature of Supervisor _____   Signature of Co-Supervisor

Name: _Wong Hann Ling_   Name: _____

Date: _03.09.2025_   Date: _____

Yang Yan Cheng

BT-ChengYangYan_FYP Project Report_commented

FYP2025.202

Document Details

Submission ID

trn:oid::3618:110623973

Submission Date

Sep 3, 2025, 1:19 AM GMT+8

Download Date

Sep 3, 2025, 2:18 AM GMT+8

File Name

BT-ChengYangYan_FYP Project Report_commented.docx

File Size

1.6 MB

72 Pages

12,224 Words

70,232 Characters

14% Overall Similarity

The combined total of all matches, including overlapping sources, for each database.

Filtered from the Report

> Small Matches (less than 8 words)

Match Groups

🔴 153 Not Cited or Quoted 13%
Matches with neither in-text citation nor quotation marks

🟠 11 Missing Quotations 1%
Matches that are still very similar to source material

🟡 0 Missing Citation 0%
Matches that have quotation marks, but no in-text citation

🔵 0 Cited and Quoted 0%
Matches with in-text citation present, but no quotation marks

Top Sources

8% 🌐 Internet sources

0% 📄 Publications

11% 👤 Submitted works (Student Papers)

Integrity Flags

0 Integrity Flags for Review

Our system's algorithms look deeply at a document for any inconsistencies that would set it apart from a normal submission. If we notice something strange, we flag it for you to review.

A Flag is not necessarily an indicator of a problem. However, we'd recommend you focus your attention there for further review.

94