Faculty of Resource Science and Technology

Heterologous Expression of Recombinant Alcohol Dehydrogenase Gene from Sago Palm in Bacterial and Plant System

Mastura Binti Sani

Master of Science

(Biotechnology)

2016

HETEROLOGOUS EXPRESSION OF RECOMBINANT ALCOHOL DEHYDROGENASE GENE FROM SAGO PALM IN BACTERIAL AND PLANT SYSTEM

Mastura Binti Sani

A thesis submitted
in fulfilment of the requirement for the degree of
Master of Science (Biotechnology)

Faculty of Resource Science and Technology
UNIVERSITI MALAYSIA SARAWAK
2016

AUTHOR'S DECLARATION

I declare that the work in this thesis was carried out in accordance with the regulations of Universiti Malaysia Sarawak. It is original and is the result of my work, unless otherwise indicated or acknowledged as referenced work. This thesis has not been submitted at Universiti Malaysia Sarawak or to any other academic institution or non-academic institution for any other degree or qualification.

Name of Student	: Mastura Binti Sani
Student ID	: 11021733
Programme Degree	: Master of Science (Biotechnology)
Faculty	: Faculty of Resource Science and Technology
Thesis Title	: Heterologous Expression of Recombinant Alcohol Dehydrogenase Gene from Sago Palm in Bacterial and Plant System
Signature of Student	:
Date	:

AKNOWLEDGEMENTS

Alhamdulillah, all praises to Allah for the strength and His blessing in completing this thesis. While my name may be alone on the front cover of this thesis, I am by no means its sole contributor. Rather, there are number of people behind this piece of work who deserved to be both acknowledged and thanked here. I am forever indebted to my academic supervisors, Associate Professor Madya Dr. Hairul Azman Roslan and Associate Profesor Madya Dr. Awang Ahmad Sallehin Awang Husaini, for their enthusiasm, guidance and unrelenting support throughout my study. I am also would like to thanks to my examiners, Associate Prof. Dr. Md Hasnain Md Hussain and Dr. Tengku Haziyamin Tengku Abdul Hamid for their advices and suggestions during viva voce session. Without their persistence help, this dissertation would not have been possible.

Special appreciation goes to my parents, Aminah Hj Jinggal and Sani Marais for giving me strength to reach for the stars and chase my dream. Thank you very much for the financial help you have extended to me in time of dire need. Both of you are the most wonderful parents in this world. Most importantly of all, I would like to show extensive gratitude to all of the past and present Genetic Engineering laboratory, who warmly contributed their experiences, support and commitment: Dr. Md. Anowar Hossain, Ms. Nabella Holling, Mr. Jerry Gerunsin, Ms. Chai Suk Phin, Mr. Saiful Shakirin Rosli, Mr. Mohd Farith Kota, Mdm. Nurul 'Izzati Chik, Ms. Nurul Hasniza Zaini, Ms. Arfazida Anuar, Ms. Siti Ba'yah Ramlan, Ms. Felicia Chin Yien Yin, Mr. Ngieng Ngui Sing and Mr. Nikson Chong Fatt Ming. Lastly, I would like to thank Tunku Abdul Rahman Sarawak scholarship foundation which supported me. I am much honoured to be the recipient of this scholarship.

ABSTRACT

Alcohol dehydrogenase (ADH) is an enzyme involved in pathways that respond to various environmental stresses such as osmotic, wound and anaerobic condition. In this study, the recombinant *Metroxylon sagu Adh1* cDNA (r-*msAdh1*) was cloned into an expression vector; pET-41a(+) and expressed in *Escherichia coli* (*E. coli*) strain BL21 (DE3). SDS-PAGE analysis of the *E. coli* lysate revealed that large amount of expressed protein formed insoluble aggregation, non-active protein. However, soluble r-msAdh1 protein was successfully produced at low-growth temperature and showed catalytic activity when spectrophotometrically assayed. Second part of this study involves the *in-planta* transformation of r-*msAdh1* cDNA into tomato seeds using *Agrobacterium tumefaciens* (*A. tumefaciens*) strain LBA4404. The result showed that, out of 15 putative transformed plants (T_0) were analysed, five putative transgenic plant lines were determined to carry r-*msAdh1* cDNA in their genome. Further analysis also showed that r-*msAdh1* cDNA have passed into all T_1 transgenic lines and expression analysis at transcript level confirmed the presence of r-*msAdh1* in transgenic tomato genome.

Keywords: Alcohol dehydrogenase, Heterologous expression, Insoluble and soluble protein, *Agrobacterium tumefaciens*, *In planta* transformation.

Pengekspresan Heterologus Gen Rekombinan Alkohol Dehidrogenase dari Pokok Sagu ke dalam Sistem Bakteria dan Tumbuhan

ABSTRAK

Alkohol dehidrogenase (ADH) merupakan enzim yang terlibat dalam laluan yang bertindak balas kepada pelbagai tekanan alam sekitar seperti osmosis, luka dan keadaan anaerobik. Dalam kajian ini, rekombinan Metroxylon sagu Adh1 cDNA (r-*msAdh1*) diklon ke dalam *vektor pET-41a(+) dan diekpres ke dalam Escherichia coli* (*E. coli*) *strain BL21 (DE3). Analisis SDS-PAGE menunjukkan sebahagian besar protein yang diekpres membentuk protein agregat tidak larut yang tidak berfungsi. Walaubagaimanapun, protein larut r-msAdh1 berjaya diekpres pada suhu pertumbuhan yang rendah dan menunjukkan aktiviti tindak balas apabila dianalisis mengunakan spektrofotometer. Manakala, bahagian kedua kajian ini melibatkan transformasi in planta r-msAdh1 cDNA ke dalam genom tomato dengan mengunakan Agrobacterium tumefaciens strain LBA4404. Keputusan menunjukkan bahawa, daripada 15 putatif tumbuhan transformasi; (T_0) yang dianalisis, terdapat lima putatif pokok transgenik adalah positif r-msAdh1 cDNA. Analisis selanjutnya menunjukkan r-msAdh1 telah dipindah kepada semua generasi T_1 dan analisis pengekpresan pada peringkat transkrip juga telah mengesahkan kehadiran r-msAdh1.*

Kata Kunci: Gen alkohol dehidrogenase, Pengekspresan heterologus, Protein tidak larut dan larut, BL21 (DE3), Agrobacterium tumefaciens, Transformasi in-planta.

TABLE OF CONTENTS

LIST OF TABLES

Page

LIST OF FIGURES

Page

Page

LIST OF ABBREVATIONS

ADH	Alcohol dehydrogenase
AS	Acetosyringone
BAR	Bialaphos Resistance
cDNA	Complementary Deoxyribonucleic Acid
$CaCl_2$	Calcium Cloride
CaMV	Cauliflower Mosaic Virus
EtBr	Ethidium Bromide
E. coli	*Escherichia coli*
IPTG	Isopropyl-β thiogalactopyranoside
LA	Luria Agar
LB	Luria Bertani
MES	2-(N-Morpholino)ethanesulfonic Acid, Monohydrate
MS	Murashige and Skoog basal salt mixture
NAD	Nicotinamide Adenine Dinucleotide
OD	Optical Density
PCR	Polymerase Chain Reaction
PMSF	Phenylmethane Sulfonyl Fluoride
r-*msAdh1*	Recombinant *Metroxylon sagu* Alcohol dehydrogenase 1
RNA	Ribonucleic Acid
RT-PCR	Reverse Transcriptase Polymerase Chain Reaction
SDS PAGE	Sodium Deodecy Sulphate Polyacrylamide Gel Electrophoresis

LIST OF APPENDICES

CHAPTER 1

INTRODUCTION

Plants frequently encounter unfavourable growth conditions such as low temperature, dehydration, phytohormone, abcisic acid and anaerobic stress (Roslan *et al.,* 2010) which delay growth and development, reduce the productivity, and in extreme cases can cause plant death (Krasensky and Jonak, 2012). However, by changing their cell structure, biochemistry pathways and evolving adaptive physiological changes, plant are able to adapt and response to these various environmental stresses (Roslan *et al.,* 2010). Meanwhile, at the molecular level; many plants also alters the pattern of gene expression upon stress (Kim *et al.,* 2003; Krasensky and Jonak, 2012). According to Krasensky and Jonak (2012), stress-inducible genes comprise of genes that involved in direct protection from stress, including the synthesis of osmoprotectants, detoxifying enzymes, and transporters, as well as genes that encode regulatory proteins such as transcription factors, protein kinases, and phosphatase.

One of the many enzymes that is involved in pathways that response to environmental stresses is alcohol dehydrogenase enzyme (ADH). Recently, Roslan *et al.* (2010) has identified the activity of ADH enzyme in various sago palm tissues (*Metroxylon sagu*) such as in young and matured leaves, waterlogged and non-waterlogged roots. The expression of *Adh* gene were reported to be highest in young leaves due to various physiological and biochemical changes as young leaf matures (Roslan *et al.,* 2010). By using the rapid amplification of cDNA ends technique (RACE), they had successfully isolated the full length of the *Adh1* cDNA from both young and matured leaf of sago palm with the size of approximately 1.3 kb in length corresponding to a predicted 380 amino acids. However, a start site of recombinant *Metroxylon sagu Adh1* cDNA (r-*msAdh1*) generated via RACE includes a 20 nucleotides that

was derived from *Adh1* cDNA sequence of oil palm (Roslan *et al.,* 2010). Nevertheless, the missing 5'-end of the gene in sago palm has been later identified via genomic walking (Wee and Roslan, 2012). The discovery of the full length of r-*msAdh1* cDNA from sago palm has prompted the work of r-*msAdh1* expression either in prokaryote and eukaryote system. This is because the ability to clone foreign gene into various heterologous host could make it possible to further investigate the function and activity of r-*msAdh1*. The first attempt to clone r-*msAdh1* in an expression vector pET-41a(+) was undertaken by Wee (2011). However, a single nucleotide mutagenesis (nucleotide addition) was discovered at the 5' of the r-*msAdh1* cDNA sequence which generated from the step to blunt-end the r-*msAdh1* cDNA during cloning procedure. The mutagenesis caused a frameshift to the open reading frame and consequently generated a stop codon approximately after 49 amino acids from start codon.

Nevertheless, further research by Miew (2011) in the reconstruction of the pET-41a(+) containing r-*msAdh1* cDNA had successfully cloned the gene in the correct frame. However, problem arose as the r-*msAdh1* cDNA sequence contained seven stop codons prior to the histidine sequence in pET-41a(+), thus rendering non-usability of his-tag column that is required in the purification step. This work therefore aims at rectifying the cloning mistakes and overexpress the r-*msAdh1* cDNA in *Escherichia coli* (*E. coli)* strain BL21 (DE3), so that the expressed protein can be later purified and use in enzyme assay. However, as the expression of eukaryotic gene in *E. coli* had reported to be significantly challenging due to the codon bias, protein folding, mRNA stability and limited eukaryotic post-translational machinery function; therefore, the expression of r-*msAdh1* cDNA in eukaryote system was also studied. The aim of the second part of this work is to construct binary vector; pGSA1131 to contains the r-*msAdh1* cDNA and subsequently used for *Agrobacterium tumefaciens*-mediated *in planta* seed transformation without involvement of tissue culture procedure. After

transformation, integration of r-*msAdh1* cDNA in T_0 and T_1 generation was analysed using polymerase chain reaction (PCR) and confirmed by nucleotide sequencing. The phenotype changes of transgenic lines over generation (T_0 and T_1) also observed and discussed. Therefore, the objectives of this study are:

1. To remove the stop codons in r-*msAdh1* cDNA sequence and clone into the expression vector, pET-41a(+).
2. To express of the pET-41a(+)/r-msAdh1 construct in *E. coli* strain BL21 (DE3).
3. To purify the recombinant protein and assess the activity of r-msAdh1.
4. To construct a binary vector, pGSA1131 that contains r-*msAdh1* cDNA.
5. To transform *Lycopersicum esculentum* cultivar MT1 with the newly construct, pGSA1131/r-msAdh1.
6. To analyse the integration and expression of r-*msAdh1* cDNA in transgenic tomato.
7. To observe the phenotype changes of transgenic lines; T_0 and T_1.

CHAPTER TWO

LITERATURE REVIEW

2.1 Alcohol dehydrogenase

Alcohol dehydrogenase (ADH, alcohol:NAD^+ oxidoreductase, EC 1.1.1.1) is a Zn-binding enzyme that acts as a dimer and relies on an NAD(P) co-factor to interconvert ethanol and acetaldehyde (and other short linear alcohol/aldehyde pairs) (Strommer, 2011). It is a member of a well-studied medium-length dehydrogenase/reductase (MDR) protein superfamily (Strommer, 2011). ADH enzyme presents in all organisms, ranging from prokaryotes to fungi, plants and animals (Machielsen *et al.,* 2006) however its metabolic function varies (Dolferus *et al.,* 1997).

The *Adh* gene have been characterized at the molecular level in wide range of flowering plants. Generally, the flowering plants possess two or three isoenzymes with exception in *Arabidopsis* that appears to have a single locus (Chang and Meyerowitz, 1986). The activity of ADH has been detected in a wide range of higher plants such as maize, rice, tomato, wheat and sago palm (Roslan *et al.,* 2010). It plays in many important roles which include;

- Facilitates cell survival during episodes of low-oxygen stress in water-logged roots (Garabagi *et al.,* 2005) and anoxic or hypoxic conditions (Roslan *et al.,* 2010).
- Responsible for adaptive functions for plant defence against environmental stresses such as osmotic and wound (Kato-noguchi, 2000 ; Kato-noguchi, 2001), cool stress and also dehydration (Dolferus *et al.,* 1994). Low temperature and osmotic stress cause

a rapid increase in activity of *Adh1* in *Arabidpsis* (Dolferus *et al.,* 1994) and also in maize and rice seedling (Christie *et al.,* 1991).

- Involved in germination of seed under anaerobic conditions (Dolferus *et al.,* 1994), in seed development (MacNicol and Jacobsen, 2001), in aerobic metabolisme in pollen (Bucher *et al.,* 1995) and expressed in dry seeds and in anaerobically treated seeds, roots and shoots (Chang and Meyerowitz, 1986).
- Detoxification of acetaldehyde to ethanol during seed storage of seed in tobacco, carbon reutilization (Garabagi *et al.*, 2005).
- Interconversion of aldehydes and alcohols in tomato, thus affects the flavour of the tomato fruit (Speirs *et al.,* 1998).
- Ethanol removal and carbon scavenging in pollinated pistils of potato (Van Eldik *et al.,* 1998).

2.2 Recombinant Alcohol dehydrogenase cDNA from *Metroxylon sagu* (r-*msAdh1*)

The recombinant alcohol dehydrogenase cDNA (r-*msAdh1*) of sago palm showed highest identity (91%) with oil palm *Adh*. This was followed by 87% identity to rice and maize *Adh1*, 85% identity to *Washingtonia robusta AdhB,* and 82% identity to *Adh* from *Arabidopsis*. The r-*msAdh1* cDNA has 1.34 kb in length and contains a predicted 380 amino acids (Wee and Roslan, 2012).

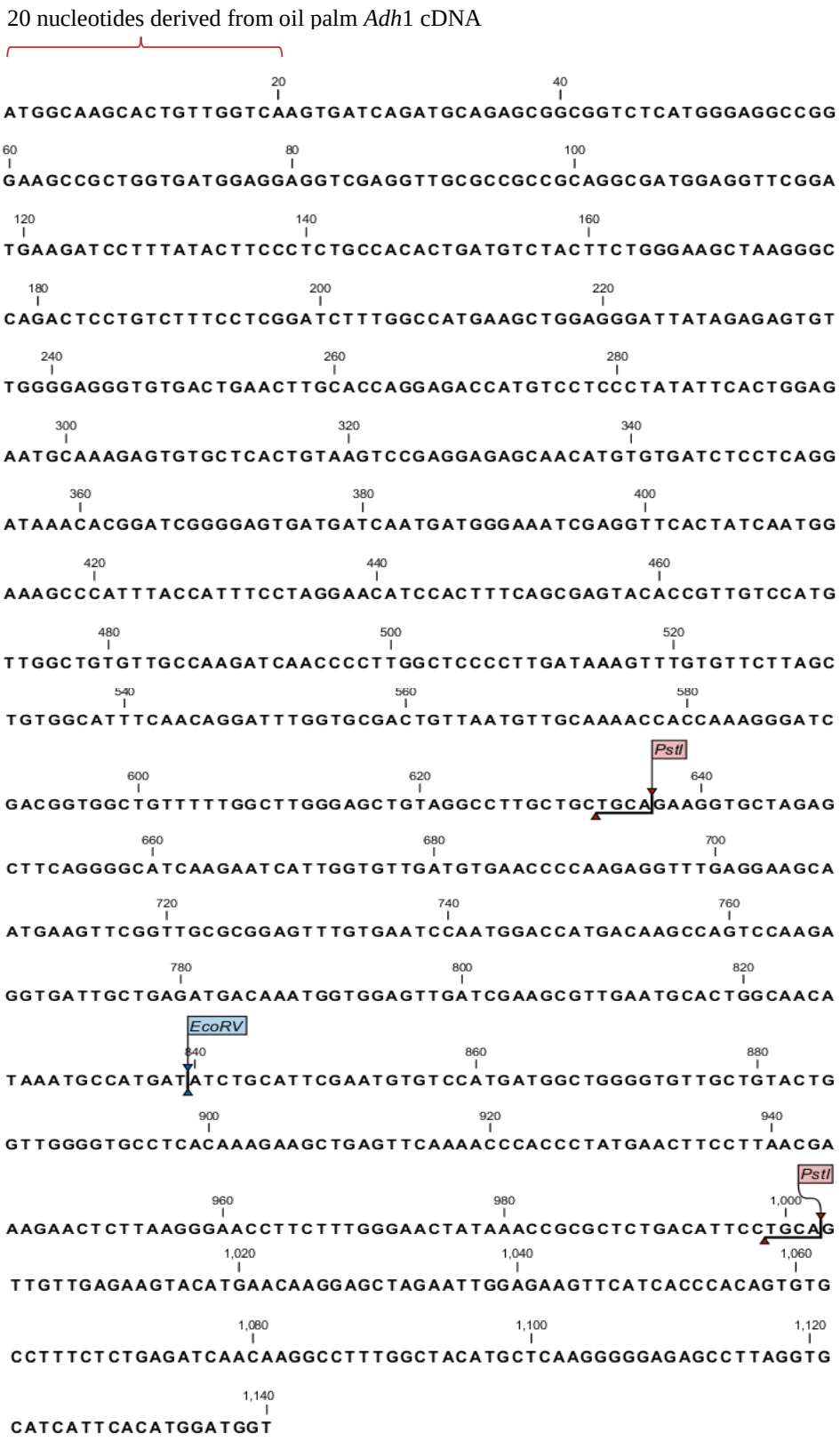


Figure 1: The full length of r-*msAdh1* cDNA generated via RACE was includes a 20 nucleotide sequence of the start site that derived from oil palm *Adh1* cDNA sequence.

2.3 Heterologous expression

Heterologous expression involves the identification of genes and transferring of the corresponding DNA fragments into other host other than its original source for synthesis of the encoded protein (Yesilirmak and Sayers, 2009). This technique is applied with the aim to overexpress the protein thus allowing large-scale purification as well as providing insights into the functions of proteins in complex pathways (Yesilirmak and Sayers, 2009).

2.4 Cloning and expression of foreign gene in *Escherichia coli*

Bacterial expression is the most common expression system for the production of recombinant proteins. For this purpose, *Escherichia coli* (*E. coli*) have long been the favourite choice because of its well-known genetics, high transformation efficiency, cultivation simplicity, rapidity and cheap (Saïda *et al.*, 2006).

The genotype of *E. coli* strains widely used for cloning and expression with the pET system is BL21 and its derivatives, which have the advantage of being deficient in both *lon* protease and lacks the *ompT* outer membrane protease (Jeong *et al.*, 2015) that can degrade proteins during purification (Grodberg and Dunn, 1988). Thus, this confers stability to some target proteins in BL21 than in host strains that contain these proteases (Novagen, 2005).

In this study, BL21 (DE3) was chosen as host of expression to express r-*msAdh1* cDNA. BL21 (DE3) is lysogens of bacteriophage DE3, a lambda derivative that has the immunity region of phage 21, carries a DNA fragment containing the ***l****acI* gene, the *T7* gene 1 (encoding T7 RNA polymerase) and the UV5 *lac* promoter (Studier and Moffatt, 1986). The addition of IPTG to growing culture (lDE3 lysogen) induces the UV5 *lac* promoter to transcribe T7 RNA polymerase which in turn transcribes the target DNA in the pET vector. Once the system is fully induced, almost all of the cell's resources are shifted towards

expressing the target protein. Therefore, the desired product can comprise more than 50% of the total protein after a few hours of induction (Novagen, 2005).

2.5 Challenges of expressing eukaryotic proteins in bacterial expression system

Expression in bacterial system is not without its challenges. Since it is a prokaryotic-based system, many heterologous expressed eukaryotic proteins are not folded properly and form insoluble aggregates, called inclusion bodies, when expressed in large amount (Khow and Suntrarachun, 2012). Apart from that, the lysis process to recover cytoplasmic protein often result in the release of endotoxin (Rai and Padh, 2001).

The insoluble and inactive proteins are co-produced in a bacterial expression system due to protein misfolding, codon bias, mRNA stability (seconds to 20 minutes) and limited eukaryotic post-translational machinery function such as glycosylation and phosphorylation (Khow and Suntrarachun, 2012). Nevertheless, successful heterologous expression of eukaryotic protein in bacterial host such as BL21 (DE3) has been reported as well. For example, the recombinant his-tagged *Bomboryx cori Adh* gene (*BmAdh*) that cloned into pET-30 (+) was expressed in BL21 (DE3) under low temperature at 15 °C (Wang *et al.*, 2011). Meanwhile, Tesniere and Verries (2000) has reported successful expression of the full length cDNA of *Adh2* and *Adh3* from the grape berries in BL21 (DE3) under the induction of short period (2 hours) and low concentration of IPTG (0.4 mM).

2.6 Fusion tag

The use of fusion tags can facilitate the detection and subsequent purification of the target protein, increase the probability of biological activity by affecting solubility in the cytoplasm or exporting to the periplasm (Novagen, 2005), protect the protein from intracellular proteolysis (Martinez *et al.,* 1995) and can be used as a specific expression reporter (Waldo *et al.,* 1999)**.**

The histidine tag sequence in pET-41a(+) encodes for a string of eight histidine residues (Refer to Appendix D). The advantageous in the use of histidine tag is that the tagged protein can be purified using commercial kits under fully denaturing condition which is particularly convenient for proteins expressed as inclusion bodies (Novagen, 2005).

2.7 *Agrobacterium tumefaciens* strain LBA4404

The *A. tumefaciens* strain used for transformation in this study is an octopine type strain, LBA4404. The strain contains *vir* helper plasmid; LBA4404, and has the TiAch5 chromosomal background (Hellens and Mullineaux, 2000). This strain does not show b-lactamase activity well and therefore, can be easily killed at low concentration of either carbenicillin or timentin (Hooykaas, 1988). According to Yarizade *et al.* (2012) the carbenicillin shows little or no detrimental effect on eukaryotic plant cells. It kills the bacteria by inhibiting the bacterial enzymes; transpeptidase and carboxypeptidase (commonly called penicillin-binding-protein) that catalyzes the reaction of prokaryotic peptidoglycan synthesis of the cell wall (Yarizade *et al.*, 2012).

Table 1 below shows the *A. tumefaciens* strains that are grouped according to the opine catabolism of the original progenitor wild-type strain and/or non-disarmed parental *Ti*

plasmid. This generally accepted classification of *Agrobacterium* strains does not necessarily imply that their disarmed counterpart still make opines (Hellens and Mullineaux, 2000).

Table 1: Disarmed *Agrobacterium tumefaciens* strains defined by the *Agrobacterium* chromosomal background and *Ti* plasmid. (Table adapted from Hellens and Mullineaux, 2000).

***Agrobacterium* strain**	**Chromosomal**		***Ti* plasmid**		**Opine**
	Back-ground	**Marker gene**		**Marker Gene**	
LBA4404	TiAch5	rif	pAL4404	strep spec	Octopine
GV2260	C58	rif	pGV2260 (pTiB6S3DT-DNA)	Carb	Octopine
C58C1	C58	-	Cured	-	Nopaline
GV3100	C58	-	Cured	-	Nopaline
A136	C58	rif and nal	Cured	-	Nopaline
GV3101	C58	rif	Cured	-	Nopaline
GV3850	C58	rif	pGV3850 (pTiC58Donc. genes	carb	Nopaline
GV3101::pMP90	C58	rif	pMP90 (pTiC58DT-DNA	gent	Nopaline
GV3101::pMP90RK	C58	rif	pMP90RK (pTiC58DT-DNA)	gent kan	Nopaline
EHA101	C58	rif	pEHA101 (pTiBo542DT-DNA	kan	Nopaline
EHA105	C58	rif	pEHA105 (pTiBo542DT-DNA	-	Succinamopine
AGL-1	C58, RecA	rif	pTiBo542DT-DNA	-	Succinamopine

Abbreviations:
carb, carbenicillin resistance; gent, gentamicin resistance; kan, kanamycin resistance; nal, nalidixic acid resistance; rif, rifampicin resistance; spec, spectinomycin resistance; strep, streptomycin resistance.

2.8 Tomato (*Lycopersicon esculentum* Mill.) as a model plant for genetic transformation.

Tomato (*Lycopersicon esculentum* Mill.) is a member of the Solanaceae family, which is dicot plant. For genetic and genomic studies, tomato has many advantages such as relatively small genome; 12 chromosomes (Hasan *et al.,* 2008), having numerous mapped traits (Tanksley *et al.*, 1992), developed DNA marker (Tanksley *et al.,* 1993), abundant collections of germplasm and mutants (Menda *et al.,* 2004), and increasing number of expressed tag (ESTs) (Van der Hoeven *et al.,* 2002; Yamamoto *et al.,* 2005). In this study, for the transformation and expression of r-*msAdh1* cDNA in plant system, tomato cultivar MT1 was chosen and seeds were purchased from Malaysia Agriculture Research and Development Institute (MARDI).

2.9 Expression of foreign genes in plants

Plants are among promising and suitable bio-platform system for production of recombinant biopharmaceutical proteins (SoltanMohammadi *et al.,* 2014). This is due to several features such as safety, no need for fermentation, cheaper investment, fast, easy to scale-up for mass production, high expression, and glycosylation ability (SoltanMohammadi *et al.,* 2014; Daniell *et al.,* 2001). The most commonly used methods for plant genetic transformation are the *Agrobacterium*-mediated transformation and particle bombardment (Ibrahim *et al.,* 2014). Transformation utilizing *A. tumefaciens* is the method of choice because it allows stable integration of foreign DNA into the plant genome and generally results in a lower copy number integrant, fewer rearrangements and improved stability of expression over generations than other DNA delivery methods (Smith and Hood 1995; Dai *et al.,* 2001; Hu *et al.,* 2003). Hence, this method produced fewer problems such as transgene instability, gene silencing, and/or co-suppression (Koncz *et al.,* 1994, Hansen *et al.,* 1997). Another advantage of *Agrobacterium*-mediated transformation is mosaicism can be avoided, which is more

frequently occurs when intact organs are transform by direct methods (Enriquez-Obregon *et al.* 1998).

The *Agrobacterium*-mediated transformation mechanism works well with dicotyledonous plants; however, monocotyledonous plants are recalcitrant towards gene transfer using *A. tumefaciens* (Ibrahim *et al.,* 2014). According to Sood *et al.* (2011), T-DNA fails to target the specific monocots meristemic cells that are competent to differentiate. In other hand, the meristamic cells lose the ability to dedifferentiate at a very early stage of development (Graves, 1988). Unlike dicot plants, monocot plants also have extremely weak wound response and have an absence or low levels of *vir*-inducing excudates (Hooykaas, 1989). However, later it was shown that monocots do produces *vir* inducing compounds but in mixtures rather than in individual compound which is stronger in activity (Xu *et al.,* 1989). Monocots were also incapable of responding to either auxins and/cytokinins in the culture medium. Besides, low efficiency of the transformation in monocots is also because of the transgene inactivation due to methylation of the T-DNA sequence (Matzke, 1991). Nevertheless, it is now possible to transform even difficult monocots using tailor-made gene construct and promoter, suitable *A. tumefaciens* strains and a proper understanding of the entire process (Sood *et al.,* 2011).

Tomato is considered as one of the most important vegetable crops for the genetic engineers because it serves as a model for introduction of agronomically important genes into dicotyledonous crop plants (Wing *et al.,* 1994; McCormick, 1986). The first report of *Agrobacterium*-mediated tomato transformation was made by McCormick and his colleagues in 1986. Since then, there have been many reports of tomato being engineered for a variety of purposes; including characterization of gene function (Janssen *et al.,* 1998) production of insect- and disease-resistant plants (Lin *et al.,* 2004), herbicide tolerance (Fillatti *et al.,* 1987),

improved fruit quality (Davuluri *et al.,* 2005; Bramley, 2002), delay in fruit ripening (Gupta *et al.,* 2013), production of foreign proteins (Youm *et al.,* 2008) and in improving the transformation protocol (Park *et al.,* 2003). The most protocol used for tomato transformation are based on shoot regeneration from leaf disk/cotyledon tissue that is co-cultivated with disarmed *A. tumefaciens* harboring a binary vector (Sharma *et al.,* 2009 ; Farzaneh *et al.,* 2013). Though genetic transformation with in vitro regeneration has been successfully used for genetic improvement of tomato (Lindsey, 1992), the drawback using tissue culturing method is the formation of somaclonal variation in transgenic plants due to the long exposure with growth hormones (Larkin and Scowcroft, 1981; Lee *et al.,* 2013), the possibility of increased contamination in rich media (Lee *et al.,* 2013), time consuming, genotype specificity, recalcitrant and transgenic plants may fail to acclimatize (Mayavan *et al.,* 2013). Thus, considerable refinements of current transformation systems are required to achieve commercial application of transgenics.

In planta transformation is an alternative method which does not involve in vitro culture of plant cells or tissues, thereby reducing time labour cost and the most importantly is somaclonal variation that is encountered during in vitro culture-mediated genetic transformation and regeneration can be avoided (Mayavan *et al.,* 2013). Recently, *in planta* transformation of dicotyledonous plants has been established for many commercial valuable crops such as citrus (Ahmad and Mirza, 2005), cotton (TianZi *et al.,* 2010), *Arabidopsis thaliana* (Bent, 2000), peanut (Rohini and Rao, 2000), apple, pear, peach, strawberry (Spolaore *et al.,* 2001). The *in planta Agrobacterium*-mediated tomato transformation by fruit injection (Yasmeen *et al.,* 2009; Hasan *et al.,* 2008) and floral dip (Yasmeen *et al.,* 2008) also have been reported. According to Yasmeen *et al*. (2009), fruit injection method gave the best transformation results compared to floral dip transformation. The finding was consistent with

those of Hasan *et al*. (2008) that reported *Agrobacterium*-infiltration of ripened fruits of tomato presented an excellent protocol for transformation. Meanwhile, the successful *in planta* transformation of dicotyledonous plants targeting on germinating seeds had been reported for radish (Park *et al.,* 2005), cotton (Keshamma *et al.,* 2008), *Brassica napus* (Song *et al.,* 2009) and sugarcane (Mayavan *et al.,* 2013). An efficient strategy of *in planta* seeds transformation developed by Mayavan *et al.* (2013) had produced stable integration of the transgene into five Indian sugarcane genome with highest transformation efficiency of 45.5 % for genotype CoC671. Up to date, however; *in planta* transformation directly using seed has not yet been reported in any tomato cultivars.

CHAPTER THREE

CONSTRUCTION OF RECOMBINANT ALCOHOL DEHYDROGENASE 1 (r-*msAdh1*) INTO EXPRESSION VECTOR AND EXPRESSION IN BACTERIAL SYSTEM

3.0 Overview

This chapter will focus on the heterologous expression of r-*msAdh1* cDNA in bacterial system. Among host systems that are available for expression of gene interest, the Gram-negative *Escherichia coli* (*E. coli*) is still the first host to try (Larsen, 2009). Though there are some major drawback in *E. coli* for eukaryotic protein expression; such as lack of post-translational modification and the limited ability to deal with disulphide bridge; the successful heterologous expression of eukaryote *Adh* cDNA in *E. coli* strain BL21 (DE3) has been reported by Wang *et al.* (2011) and Tesniere and Verries (2000). Therefore, to achieve the goals, the reconstruction of r-*msAdh1* cDNA into pET-41a(+) was done and the newly construct then transformed into *E. coli* strain BL21 (DE3). In the attempt to obtained soluble r-msAdh1 protein, expression condition was optimized by growing the induced culture at different growth temperatures and periods. Subsequently, the expressed protein then was purified using NI-NTA Spin Column and catalytic activity of r-msAdh was assayed using spectrophotometry by monitoring the increase of absorbance at 340 nm. The objectives of this chapter includes;

1. To remove the stop codons in r-*msAdh1* cDNA sequence and clone into the expression vector, pET-41a(+).
2. To express of the pET-41a(+)/r-msAdh1 construct in *E. coli* strain BL21 (DE3).
3. To purify the recombinant protein and assess the activity of r-msAdh1.

The workflow is briefly illustrated in Figure 2 below.

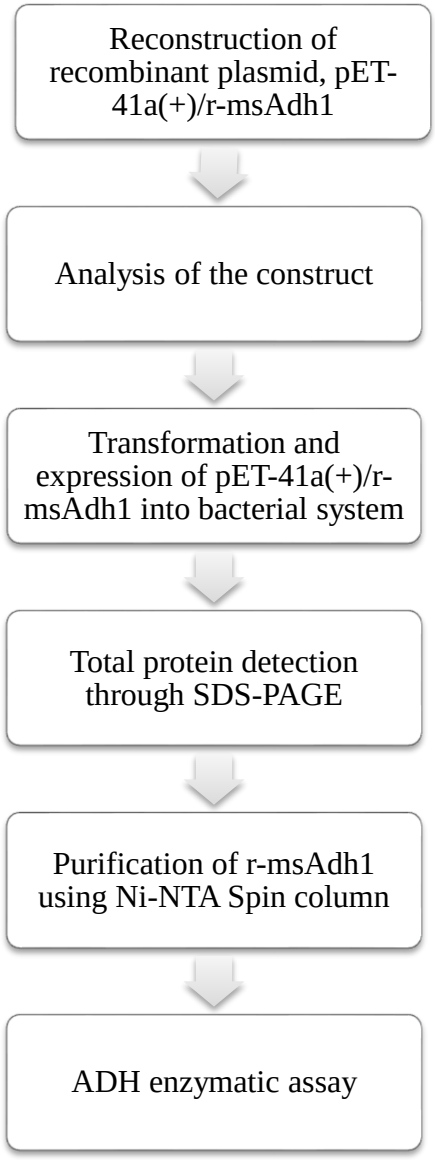


Figure 2: Workflow for construction of r-*msAdh1* into expression vector and their expression in bacterial system.

3.1 Materials and Methods

3.1.1 Construction of r-*msAdh1* cDNA in expression vector, pET-41a(+)

3.1.1.1 Directional cloning of PCR product using 5'NdeI_adaptor and 3'XhoI_r-msAdh1 specific primers

The open reading frame of the r-*msAdh1* coding region was obtained through amplification of pET-41a(+)/r-msAdh1 plasmid that constructed by Miew (2011) using the following primers: 5'NdeI_adaptor (5'-GGAATTTATGGCAAGCAGTGTTGG-3') and 3_r-msAdh1_XhoI (5'-GCTAACTCGAGACCATCCATGTGAATGATGCAC-3'). The PCR reaction profile used is shown in Table 2. Meanwhile, the composition of the PCR reaction mixture is shown in Table 3.

Table 2: PCR reaction profile to generate the full length of r-*msAdh1* cDNA. PCR was undertaken using the 5'NdeI_adaptor and 3_r-msAdh1_XhoI primers combination.

Segment	Temperature (°C)	Duration	Number of cycles
Initial denature	94	3 m	1x
Denaturation	94	30 s	35x
Annealing	62	1 m	
Elongation	72	2 m	
Initial denature	72	10 m	1x

Table 3: The composition of 1X PCR mixture to generate the full length of r-*msAdh1* cDNA.

Reagents	**1X (µL)**
10x High Fidelity buffer with 15mM $MgCl_2$	2.5
dNTP (2.5 mM)	2.5
5_NcoI_msAdh1 (10 µM)	1
3'_BamHI_msAdh1(10 µM)	1
Extracted plasmid, pET-41a(+)/r-msAdh1 (1.0 µg/µL)	0.5
High Fidelity enzyme (0.625 U/µL)	0.5
Nuclease-free water	17
Total reaction	**25**

Next, the PCR product was purified using the PCR Cleanup Kit (Vivantis, USA) and subsequently digested with *Nde*I (Fermentas, Lithuania) and *Xho*I (Fermentas, Lithuania) restriction enzymes for 6 hours at 37 °C. The composition for the reaction is shown in Table 4. Thermal inactivation was done at 80 °C for 20 minutes. After the double digestion, the r-*msAdh1* cDNA was then purified from the digestion mixture using the PCR Cleanup Kit (Vivantis, USA).

Table 4: The composition of restriction enzyme double digestion using *Nde*I and *Xho*I to produce sticky ended r-*msAdh1* fragment.

Reagents	**Volume (µL)**
Purified r-*msAdh1* fragment (0.5 µg/µL)	20
Orange Buffer (10X)	4
*Nde*I (10 U/µL)	5
*Xho*I (10 U/µL)	5
Total reaction	**34**

At the same time, the isolation of pET-41a(+) was done using GeneMATRIX™ Plasmid Miniprep DNA Purification Kit (Molecular Biology Resources, USA) and subsequently subjected to double restriction enzyme digestion using *Nde*I (Fermentas, Lithuania) and *Xho*I (Fermentas, Lithuania) for 3 hours at 37 °C. The composition for the reaction is shown in Table 5 and heat inactivation was done at 80 °C.

Table 5: The composition of restriction enzyme double digestion using *Nde*I and *Xho*I to produce sticky ended pET-41a(+).

Reagents	**Volume (µL)**
Extracted plasmid, pET-41a(+), (1.0 µg/µL)	20
Orange Buffer (10X)	3
*Nde*I (10 U/µL)	3
*Xho*I (10 U/µL)	3
Nuclease-free water	1
Total reaction	**30**

The fragments were separated by running an agarose gel electrophoresis at 100 V with expected band sizes of approximately 5.013 kb and 920 bp. The desired fragment (~ 5.013 kb) was then extracted from the gel using GF-Gel Recovery Kit (Vivantis, USA). Subsequently, the r-*msAdh1* cDNA fragment was ligated into the vector using T4 DNA Ligase (Fermentas, Lithuania) with incubation at 22 °C for 2 hours using composition shown in Table 6.

Table 6: The composition of ligation mixture to clone r-*msAdh1* into pET-41a(+).

Reagents	**Volume (µL)**
Linearized vector, pET41-a(+), (0.5 µg/µL)	9.5
Purified r-*msAdh1* fragment, (0.5 µg/µL)	20
T4 Ligase Buffer (10X)	3.5

T4 Ligase (1 weiss U/µL)	2
Total reaction	**35**

3.1.1.2 Preparation of *Escherichia coli* strain XL-1 Blue competent cells

Ten microliters of thawed frozen glycerol stock of *E. coli* strain XL-1 Blue was inoculated into 10 mL of Luria Bertani (LB) and grown for overnight at 37 °C with shaking at 150 rpm. On the next day, 500 µL of the overnight culture was added into 10 mL LB without any antibiotics and allowed to grow at 37 °C with shaking at 150 rpm until the OD_{600} reached approximately 0.5-0.6. Upon reaching the desired density, the tube was cooled on ice for 20 minutes and centrifuged at 3500 at 4 °C for 5 minutes. The supernantant was discarded and the cell pellet was then gently resuspended with 15 mL iced-cold 100 mM calcium chloride ($CaCl_2$). The cell suspensions were kept on ice for 10 minutes and centrifuged again as above. After the centrifugation, the supernantant was discarded and the cells pellet resuspended with 1 mL of cold sterile 100 mM $CaCl_2$. For long period storage, the cell suspension was added with 20 % of pure glycerol, mixed well and aliquoted into several microcentrifuge tubes and stored at -80 °C.

3.1.1.3 Transformation of the ligation reaction into *E. coli* strain XL-1 Blue via heat shock

The ligation mixture was mixed with 50 µL competent cells in a pre-chilled microcentrifuge tube and incubating for 20 minutes on ice. The cells were then heat-shocked for 2 minutes at 42 °C pre-set water bath and quickly returned onto ice for 10 minutes. Next, 1 mL of LB was added into the tube and incubated at 37 °C, with gentle agitation at 150 rpm in an incubator-shaker for an hour. The culture was then spread onto Luria agar (LA) plates containing 100 µg/mL of kanamycin. All plates were incubated overnight at 37 °C.

3.1.1.4 Plasmid isolation

A single colony was selected from the selection plate and inoculated into a 50 mL falcon tube containing 10 mL of LB supplemented with 100 µg/mL kanamycin. The tube was then cultured overnight at 37 °C with shaking at 150 rpm. On the following day, 4 mL of overnight culture was used for plasmid isolation using GeneMATRIX™ Plasmid Miniprep DNA Purification Kit (Molecular Biology Resources, USA).

3.1.1.5 Verification of reading frame, restriction enzyme digestion analysis and nucleotide sequencing

Verification of nucleotide sequence of the recombinant gene was undertaken to ensure the proper open reading frame. The extracted plasmid was verified via PCR using specific r-*msAdh1* primers. The PCR reaction profile is shown in Table 7 and the composition of PCR mixture is shown in Table 8.

Table 7: PCR reaction profile to determine the presence of r-*msAdh1* in pET-41a(+) using the 5'NdeI_adaptor and 3_r-msAdh1_XhoI primers combination.

Segment	Temperature (°C)	Duration	Number of cycles
Initial denature	94	3 m	1x
Denaturation	94	30 s	35x
Annealing	62	1 m	
Elongation	72	2 m	
Initial denature	72	10 m	1x

Table 8: The composition of 1X PCR mixture used to determine the presence of r-*msAdh1* in pET-41a(+) using the 5'NdeI_adaptor and 3_r-msAdh1_XhoI primers combination.

Reagents	**1X (µL)**
2X Green Go-Taq Master Mix (Promega)	7.5
5'NdeI_adaptor (10 µM)	1
3_r-msAdh1_XhoI (10 µM)	1
Extracted plasmid, pET-41a(+)/r-msAdh1(1.0 µg/µL)	1
Nuclease-free water	4.5
Total reaction	**15**

Verification of the open reading frame of the new construct was also carried out by restriction enzyme digestion analysis. The double and single digestions were performed using composition as shown in Table 9 for an overnight period at 37 °C. The fragments were separated by running agarose gel electrophoresis, 1 % at 100 V for 30 minutes with expected sizes of approximately 5.013 kb and 1.140 kb for double digestion of pET-41a(+)/r-msAdh1, approximately 6.153 kb in size for single digestion of pET-41a(+)/r-msAdh1 and approximately 5.933 kb in size for single digestion of pET-41a(+)/Empty (control).

Table 9: The composition of restriction enzyme digestion to verify open reading frame of pET-41a(+)/r-msAdh1.

Reagents	**Single digestion of pET-41a(+)/r-msAdh1 and pET-41a(+)/Empty (µL)**	**Double digestion of pET-41a(+)/r-msAdh1 (µL)**
Plasmid (1.0 µg/µL)	7	11.5
Orange Buffer (10X)	1	1.5
*Nde*I (10 U/ µL)	1	1
*Xho*I (10 U/ µL)	-	1
Nuclease-free water	1	-
Total Volume	**10**	**15**

For further verification, the new construct was also sequenced using T7 Promoter: 5'-TAATACGACTCACTATAGGG-3', T7 Terminator: 5'-GCTAGTTATTGCTCAGCGG-3', Adhmor8_F: 5'-CTAGAGCTTCAGGGGCATCA-3' and 5'_msAdh1_R: 5'-AACACAGCCAACATGGACAA-3'.

3.1.1.6 Transformation of pET-41a(+)/r-msAdh1 into an expression host, *Escherichia coli* strain BL21 (DE3)

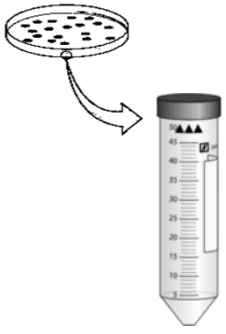
Preparation of *E. coli* competent cells strain BL21 (DE3) and bacterial transformation were carried out as described in segment 3.1.1.3.

3.1.2 Heterologous expression of r-*msAdh1* in *Escherichia coli* strain BL21 (DE3)

3.1.2.1 Induction and expression of r-*msAdh1* in BL21 (DE3)

The procedure for expression of pET-41a(+)/r-msAdh1 in BL21 (DE3) is illustrated in the figure 3. To verify the result, the induction and expression of pET-41a(+)/Empty and pET41a(+)/r-msAdh1 in BL21 (DE3) was repeated three times.

Day 1



Freshly transformed colony BL21(DE3) harbouring pET-41a(+)/r-msAdh1 was inoculated into 50 mL falcon tube containing 6 mL LB and kanamycin 100 µg/mL.

Culture incubated at 37 °C with shaking 250 rpm to OD_{600}: 0.6-1.0. Upon reached OD_{600} 0.6-1.0, culture kept at 4°C for overnight.

Day 2

Two millilitres culture was centrifuged for 30 seconds at 3500 rpm and supernatant discarded. Then, pellet was resuspended with 1 mL fresh LB media and used to inoculate 49 mL LB media + 100 µg/mL kanamycin in 250 mL Erlenmeyer flask.

Culture was incubated at 37 °C with shaking 250 rpm to OD_{600}: 0.5-0.6. Upon reached the desired density, 1.5 mL culture was aliquot to serve as un-induced control.

Remaining culture were divided into four 100 mL Erlenmeyer flask; A, B, C and D (each flask containing 10 mL culture). All flasks were added with IPTG to final concentration of 0.4 mM. Flask A and B were incubated for 4 and 6 hours at temperature 15 °C. Meanwhile, flask C and D were incubated for 4 and 6 hours at temperature 27 °C.

A: T_4 **B: T_6** **C: T_4** **D: T_6**

Incubated at temperature 15 °C Incubated at temperature 27 °C

The OD_{600} was then measured from each flask after period T_4 and T_6. At the same time, 1.5 mL culture was aliquot and centrifuged at 13, 200 rpm for 4 minutes. The supernantant was discarded and pellet was stored at -20 °C until use.

Figure 3: Figure showed the procedure for the expression of pET-41a (+)/r-msAdh1 in BL21 (DE3). The procedure of expression was followed as recommended by Novagen (2005) with minor modifications.

3.1.3 Total protein extraction and detection

3.1.3.1 Cell lysis and total protein extraction

The pellets stored at -20 °C were thawed and resuspended with SDS sample buffer. The pellet from non-induced culture (T_0) was resuspended with 75 µL of SDS sample buffer. To standardize the total protein concentration, the pellets from induced culture (T_4 and T_6) were resuspended with the volume that corresponds to the cell density (OD at 600). After that, samples were incubated at 94 °C for 4 minutes on a heat block. The soluble and insoluble fractions were separated by centrifugation at 35, 000 rpm for 10 minutes. Soluble fraction was transferred into new microcentrifuge tube and kept in -20 °C.

3.1.3.2 Total protein detection through denaturing Polyacrylamide Gel Electrophoresis (PAGE)

The protein samples were analysed on denaturing sodium deodecy sulphate polyacrylamide gel electrophoresis (SDS-PAGE). PAGE was prepared by pouring 12.0 % resolving gel into gel cast until it reached 1 cm below the end of the notched plate. The gel was layered with distilled water and left to polymerize for approximately 30 minutes. After the gel had polymerized, the water was poured off and dried with Whatman paper. Then, 4 % stacking gel was loaded into the gel cast until it reached the end of the notched plate. The comb was then placed slowly and the gel was left to polymerize for 15 minutes. Prior to electrophoretic run, the comb was carefully removed and the gel was placed into the electrophoresis chamber (Owl Scientific, USA). The 1X running buffer was poured to the inner and outer of the reservoir until the gel was fully immersed. Meanwhile, the protein sample was mixed with sample buffer at a ratio of 3:1 (15 µL protein sample to 5 µL loading buffer) and loaded into the well.

The preparation of denaturing PAGE was followed as recommended by Mini-Protean Tetra Cell (BIORAD, USA).

Figure 4: Picture showed denaturing polyacrylamide gel electrophoresis (PAGE) was carried out at a constant voltage of 120 V. The electrophoresis run until the tracking dye reached approximately 1 cm from the bottom of the gel.

Once electrophoresis was completed, gel was carefully removed from the glass plate and transferred into fixing solution [45 % methanol, 10 % acetic acid / 45 % water] for 30 minutes. Then, the gel was stained in staining solution [50 % methanol / 10 % acetic acid / 40 % water / 0.25 % Coomassie Brilliant Blue R250 (Merck, USA)] for an overnight. On the following morning, gel was then transferred into destaining solution [30 % methanol / 10 % acetic acid / 60 % water] until protein band can be observed.

3.1.4 Protein purification using NI-NTA spin column under native conditions from *E. coli* lysates

Protein purification was done using Ni-NTA Spin Kits (Qiagen, Germany). The kit provide a simple method for rapid screening and purification of 6xHistagged r-msAdh1 protein from small-scale expression culture. To purify histagged r-msAdh1 protein, a pellet derived from 20 mL culture of BL21 (DE3) expressing r-*msAdh1* cDNA was resuspended in 1400 µL Lysis Buffer (NP1-10) containing 1 mg/mL lysozyme, 0.5 mM PMSF, 60 µL DNAse 1 (1 u/µL) and subsequently incubated for 30 minutes on ice. Then, 1.6 g glass beads was added and vortex at 13, 200 rpm for 2 minutes. After that, soluble and insoluble fractions were separated by centrifuging at 12, 000 x g for 25 minutes. The soluble fraction was then loaded into pre - equilibrate column and centrifuged at 1600 rpm for 5 minutes. To remove unbind protein (contaminants), column was washed three times by adding 600 µL Buffer NP1-20. After that, 200 µL Buffer NPI-500 was added into column and subsequently centrifuged at 2900 rpm for 2 minutes to elute pure his-tagged protein. The procedure is followed as recommended by Ni-NTA Spin Kit Handbook (2008) with minor modifications.

3.1.5 Enzymatic assay of ADH enzyme

The catalyzing activity of r-msAdh1 was assayed spectrophotometrically by measuring the increase in absorbance at 340 nm, following the reduction of NAD^+ to NADH. Briefly, 1 mL of reaction buffer; Tris-HCl 100 mM pH 8.3 containing 0.08 mL ethanol and 2 mM NAD^+ were incubated at 25 °C, and the reaction was initiated by adding 0.02 mL of crude extract msAdh1 (Bergmeyer, 1983). The rate of increase at A_{340} in the first 15 minutes was recorded.

3.2 Results and Discussion

3.2.1 Construction of pET-41a(+)/r-msAdh1

3.2.1.1 Verification of pET-41a(+)/r-msAdh1

The pET-41a(+)/r-msAdh1 construct from Miew (2011), was verified via polymerase chain reaction (PCR) utilizing specific r-*msAdh1* cDNA primers; Adhmor8-F and Adhmor8-R. The PCR produced approximately 650 bp of fragment on 1 % agarose gel electrophoresis and the photograph of agarose gel electrophoresis is shown in Figure 5.

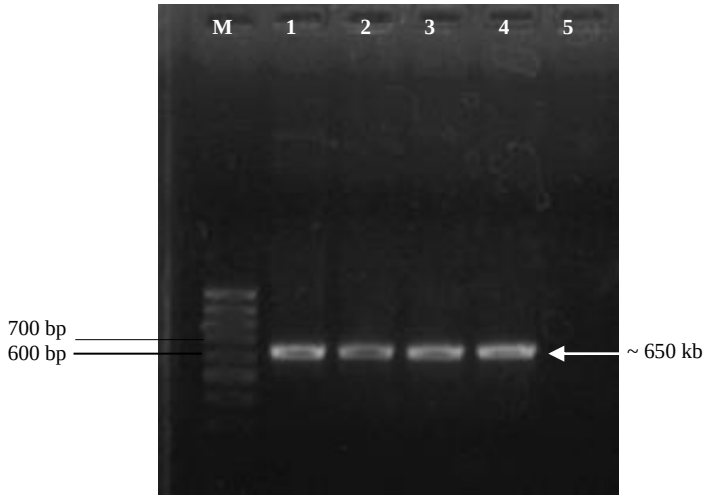


Figure 5: Gel electrophoresis of PCR product visualized on 1 % agarose gel stained with ethidium bromide (EtBr). The PCR product is partial r-*msAdh1* cDNA fragment that amplified using the Adhmor8-F and Adhmor8-R primers combination. Lane 1: GeneRuler™ 100 bp DNA Ladder (Fermentas, Lithuania), Lane 1-4: PCR product, Lane 5: Negative control.

The result indicates that plasmid extracted from bacterial glycerol stock is containing r-*msAdh1* cDNA. To further verify the nucleotide sequence, purified plasmid was sequenced (1st Base Sdn Bhd, Malaysia). The different fragments resulted from sequencing were assembled by using overlapping consensus sequence. The full length of r-*msAdh1* cDNA with size approximately of 1.441 kb in pET-41a(+) is shown in Figure 6.

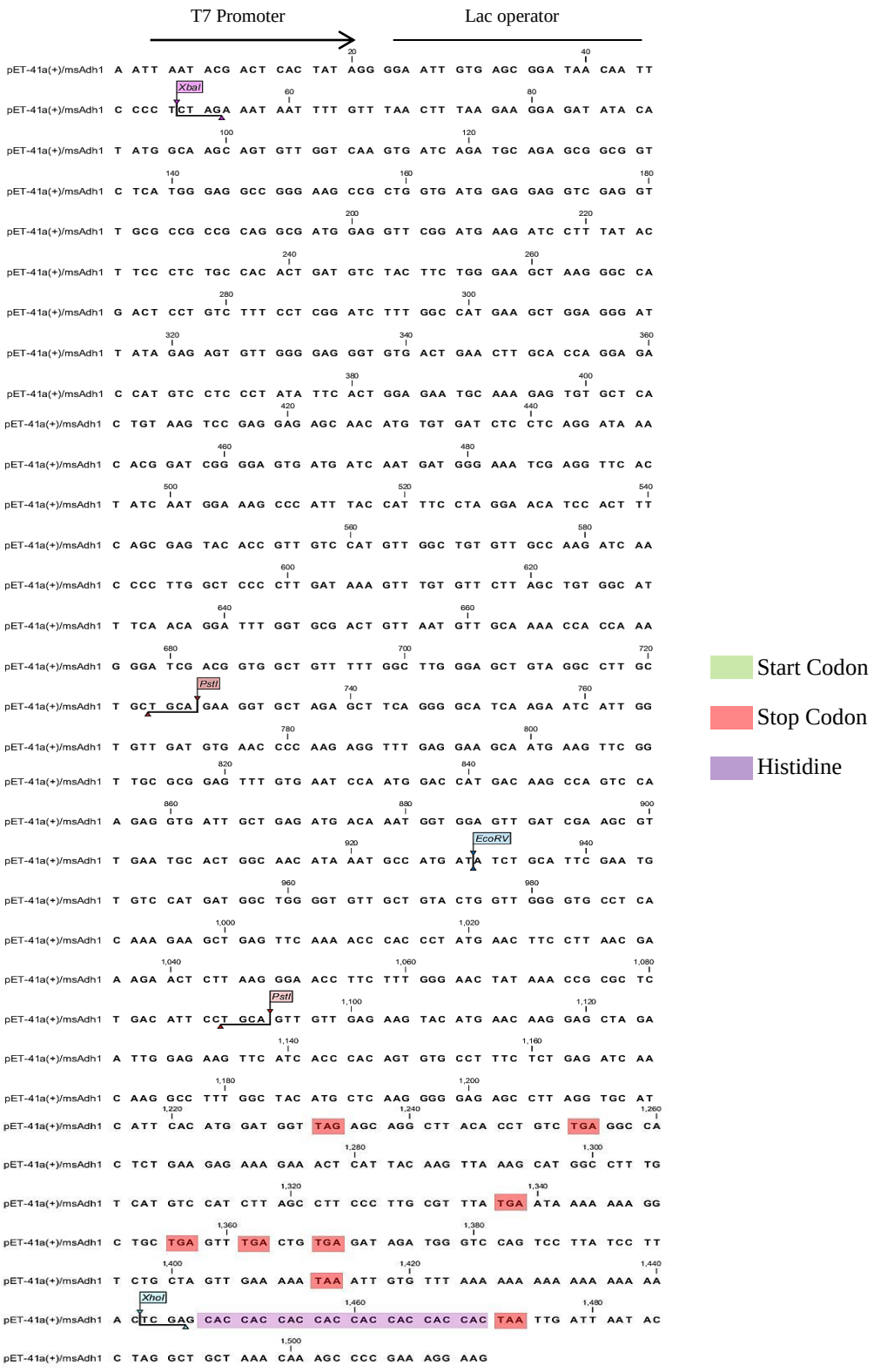


Figure 6: Figure showed the r-*msAdh1* cDNA sequence with seven stop codons highlighted in red.

Based on cDNA r-*msAdh1* sequence (Figure 6), it can be seen that methionine (start codon) in r-*msAdh1* sequence is located after a few nucleotides of ribosomal binding site in pET-41a(+). Downstream of the sequence, seven stop codons was identified (highlighted in red) prior to the his-tag sequence (highlighted in purple). This created a problem because the his-tag sequence is required for r-msAdh1 protein purification; therefore the stop codons were removed from the r-*msAdh1* cDNA sequence by PCR using 5'NdeI_adaptor and 3'XhoI_r-msAdh1 primers combination.

3.2.1.2 Directional cloning of PCR product using 5'NdeI_adaptor and 3'XhoI_r-msAdh1 specific primers to produce non-stop r-*msAdh1* cDNA

The use of primers combination; 3'XhoI_r-msAdh1 and 5'NdeI_adaptor in PCR to amplify non-stop r-*msAdh1* cDNA resulted amplification of approximately 1.147 kb fragment on 1 % agarose gel electrophoresis. The amplified r-*msAdh1* cDNA fragment containing endonuclease restriction sites (*Nde*I and *Xho*I) on 5' and 3' respectively. The photograph of agarose gel electrophoresis is shown in Figure 7.

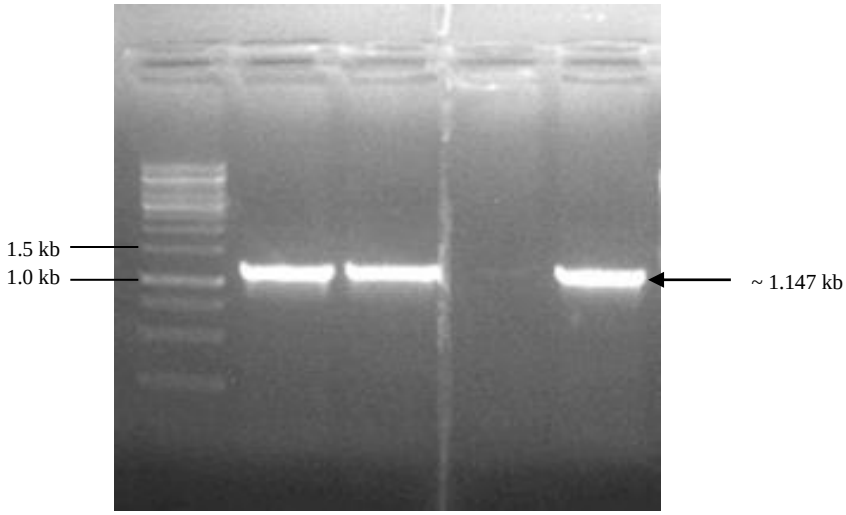


Figure 7 : Gel electrophoresis of PCR product visualized on 1 % agarose gel stained with EtBr. The PCR product is the full length of r-*msAdh1* cDNA fragment that amplified using the 3'XhoI_r-msAdh1 and 5'NdeI_adaptor primers combination. Lane M: GeneRuler™ 1 kb DNA Ladder (Fermentas, Lithuania), Lane 1, 2 and 4: PCR products, Lane 3: Negative control.

The purified PCR product subsequently digested using restriction enzymes; *Nde*I (Fermentas, Lithuania) and *Xho*I (Fermentas, Lithuania). The pET-41a(+) vector was also digested using the same restriction enzyme as well. The excision of the vector using *Nde*I together with *Xho*I in orange buffer for three hours was sufficient to digest the vector. After that, the r-*msAdh1* cDNA fragment was ligated into pET-41a(+) using T4 DNA Ligase (Fermentas, Lithuania). For the purpose of cloning, the ligation mixture was transformed into *E. coli* strain XL-1 Blue.

3.2.1.3 Plasmid extraction

Plasmids were successfully extracted from bacterial culture, as indicated by the two bands on the agarose gel (Figure 8). The two bands corresponded to the different migration patterns of the two forms of plasmid; supercoiled plasmid and nicked circular plasmid. Five microliters of isolated plasmid, pET-41a(+)/Empty (control) and pET-41a(+)/r-msAdh1 respectively were loaded into agarose gel.

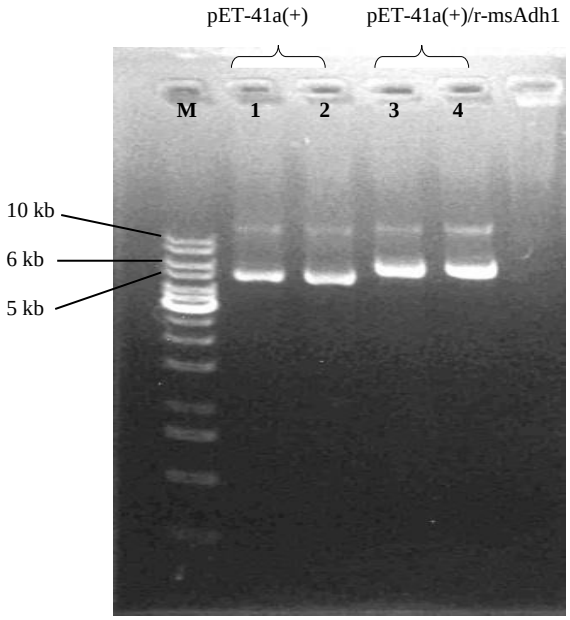


Figure 8: Gel electrophoresis result of plasmid extraction. Lane M: GeneRuler™ 1 kb DNA Ladder (Fermentas, Lithuania), Lane 1 and 2: Isolated plasmid, pET-41a(+)/Empty, Lane 3 and 4: Isolated plasmid, pET-41a(+)/r-msAdh1.

3.2.1.4 Verification of pET-41a(+)/r-msAdh1 reading frame via restriction enzyme digestion analysis and nucleotide sequencing.

Figure 9 showed the photograph of gel electrophoresis for restriction digestions of both plasmid; pET-41a(+)/Empty and pET-41a(+)/r-msAdh1 to prove the inclusion of r-*msAdh1* cDNA in pET vector. Single digestion using *Nde*I only cut the supercoiled plasmid; pET-41a(+)/r-msAdh1 and pET-41a(+)/Empty at only one site, thus converting the supercoiled plasmid into a linear form with size approximately of 6.153 kb (Lane 4) and 5.933 kb (Lane 2), respectively. This step was conducted to estimate the exact molecular weight of both plasmids.

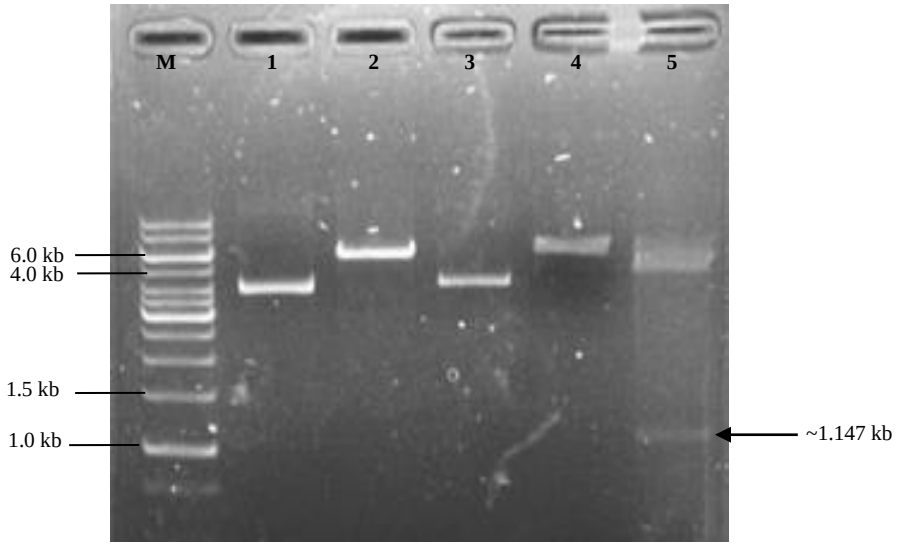


Figure 9: Restriction enzyme analysis of pET-41a(+)/r-msAdh1 visualized on 1% of agarose gel stained with EtBr. Lane M: GeneRuler™ 1 kb DNA Ladder (Fermentas, Lithuania). Lane 1: Uncut supercoiled pET-41a(+)/Empty, Lane 2: Linearized pET-41a(+)/Empty – 5.933 kb, Lane 3: Uncut supercoiled pET-41a(+)/r-msAdh1 , Lane 4: Linearized pET-41a(+)/r-msAdh1 – 6.153 kb, Lane 5: Double digestion of pET-41a(+)/r-msAdh1 using *Nde1* and *Xho1*- 5.013 kb, 1.147 kb.

Meanwhile, double digestion of pET-41a(+)/r-msAdh1 using *Nde*I and *Xho*I had cut this plasmid at two sites (Lane 5). Result from gel electrophoresis of the restriction digestion showed that the sizes of the restricted fragments corresponded to the expected sizes calculated from plasmid restriction maps; 5.013 kb (pET vector) and 1.147 kb (r-*msAdh1* cDNA). The results confirmed that the sizes of the restricted fragments are approximately the same to the predicted sizes.

DNA sequencing also was performed using several different primers to ensure that r-*msAdh1* was successfully cloned in correct orientation into pET-41a(+) vector. The result was assembled by using overlapping consensus sequence. The junction of nucleotide sequence of r-*msAdh1* in pET-41a(+) vector is illustrated in graphical as shown in Figure 10.

Figure 10: The junction of nucleotide sequence of r-*msAdh1* cDNA in pET-41a(+) vector. Graphic showing the r-*msAdh1* sequence (in red box) located between *Nde*I and *Xho*I restriction site in pET-41a(+). Eight histidine codons (in orange box) are located before stop codon (TAA).

The r-*msAdh1* sequence analysis using CLC Sequence Viewer (version 6.9.1) has generated 380 amino acids. Meanwhile, the standard protein BLASTp (NCBI) has identified the amino acids as *Adh* gene from plant (Figure 11).

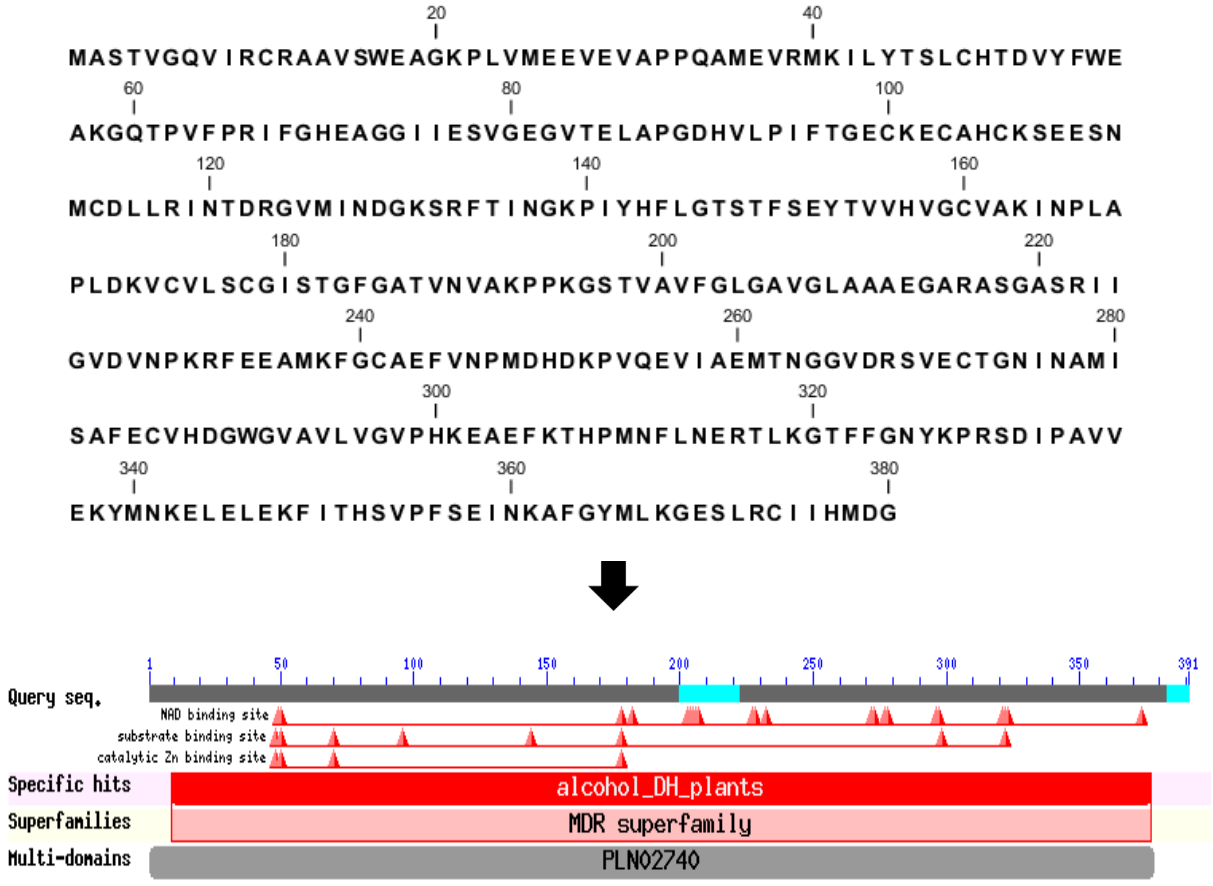


Figure 11: Analysis of r-*msAdh1* cDNA using standard protein blast (NCBI). The r-*msAdh1* cDNA of approximately 1.147 kb corresponded to the 380 amino acids and identified as *Adh* gene from plant.

3.2.1.5 Transformation of pET-41a(+)/r-msAdh1 into *E. coli* strain BL21 (DE3)

Both plasmid; pET-41a(+)/Empty (control) and pET-41a(+)/r-msAdh1 were transformed into *E. coli* strain BL21 (DE3). Compared to the XL-1 Blue, BL21 (DE3) transformed colonies have big morphology and the number of transformant is less even though the concentration of cells and DNA used for heat shocked transformation is same (Figure 12). BL21 (DE3) is not recommended as host for primary cloning or ligation because it is naturally lack of Dcm, therefore its endonuclease 1 activity may degrade plasmid after isolation (NEB UK Expression, 2012).

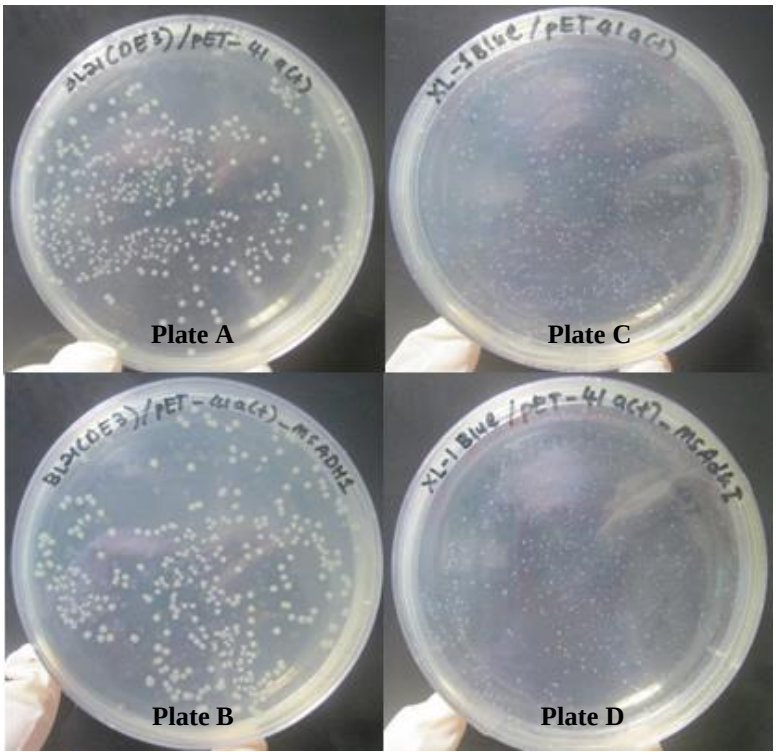


Figure 12: Transformant selection on Luria agar (LA) plates supplemented with 100 µg/mL of kanamycin. Plate A: Colonies BL21 (DE3) harbouring pET-41a(+)/Empty, Plate B: Colonies BL21 (DE3) harbouring pET-41a(+)/r-msAdh1. Plate C: Colonies XL-1 Blue harbouring pET-41a (+)/Empty, Plate D: Colonies XL-1 Blue harbouring pET-41a(+)/r-msAdh1.

3.2.2 Preliminary analysis of the expression of r-*msAdh1* cDNA in BL21 (DE3

The general practice in lab is to make starting culture by growing an overnight culture using a rich medium, such as LB, at 37 °C and using 1 mM IPTG for induction. However, in case of expression of r-*msAdh1* cDNA in BL21 (DE3), no distinct or extra band was seen in total protein lysate when visualized on 12 % SDS PAGE; indicating that r-*msAdh1* cDNA could not express under this condition. Because many researchers have recommended using a freshly transformed colony to prepare a starter culture, therefore, in this study; the expression of r-*msAdh1* cDNA in BL21 (DE3) using a freshly transformed colony and low concentration of IPTG (0.4 mM) was attempted.

Nevertheless, for high level of protein production, it is important to optimize expression conditions, such as temperature and the time after IPTG-induction (Sivashanmugam *et al.*, 2009). Therefore, time courses and temperature optimizations were carried out. The level of expression was checked by resuspending the cells pellet in sample buffer and heating at 95 °C for 5 minutes. The total protein then analysed on 12 % SDS PAGE (Figure 13).

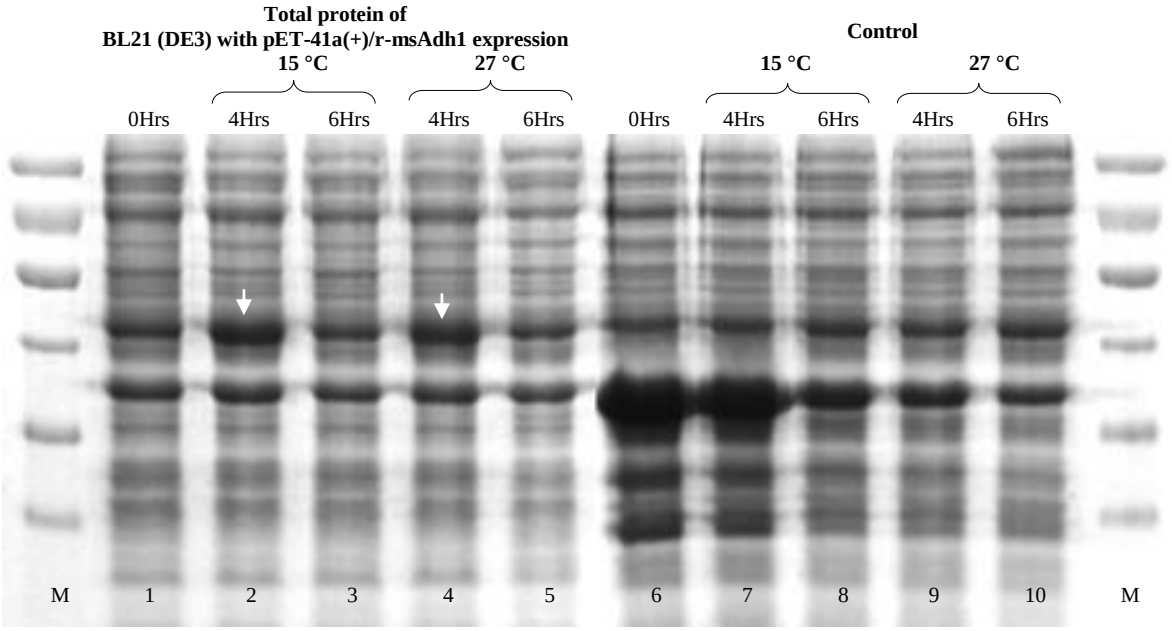


Figure 13: Total protein of uninduced (Lane 1) and induced culture (Lane 2-5) of BL21 (DE3) with r-*msAdh1* expression visualized on 12 % SDS- PAGE stained with commassie blue. The expected r-msAdh1 protein band is shown by white arrow. Lane M: EZ Run Pre-stained *Rec* Protein Ladder. Lane 6-10: BL21 (DE3) with pET-41a(+)/Empty expression (control).

Figure 13 showed the total protein of BL21(DE3) with r-*msAdh1* expression visualized on 12 % SDS-PAGE. The concentration of total protein loaded in each land is standardized, so that the r-*msAdh1* expression could be analysed on the stained gel by identify for any increase in the protein band intensity compared to the control.

Based on the figure 13, it can be seen that a distinct band with moderate intensity at molecular weight around 43 kD - 48 kD was observed in Lane 2 and 4 (showed by white arrow). Wee and Roslan (2012) reported that the r-msAdh1 protein size is almost similar to the 380 deduced amino acid of *Adh1* gene from, *Zea mays, Oryza sativa, Mus musculus, Homo sapiens, Saccharomyces cerevisiae* and *Bacillus cereus* which is the same to the expected r-msAdh1 protein band size produced on SDS-PAGE in Figure 13 above. Even though the same

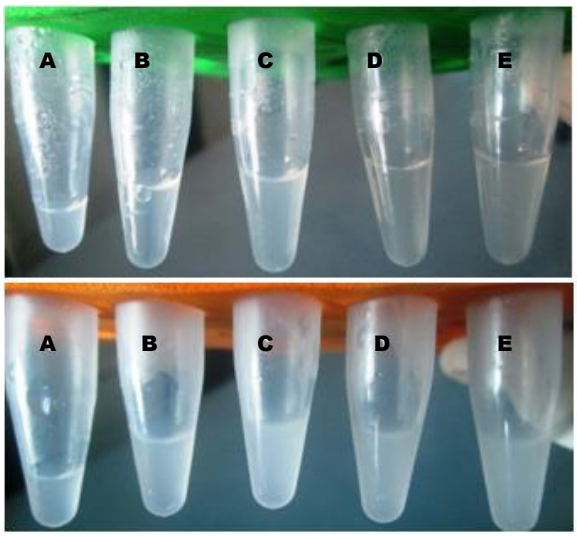
band size was identified in the control (Lane 6-10), the band intensity is much consistent at any temperature and time after IPTG induction, therefore it is highly likely it is an *E. coli* cellular protein. Due to this, the presence of the r-msAdh1 band is quite difficult to accurately distinguish through SDS PAGE alone as r-msAdh1 band might be masked by other cellular protein.

3.2.3 Determination of r-msAdh1 protein solubility

Before proceed to the downstream application such as protein purification and enzyme assay, it is important to determine the solubility of r-msAdh1 protein. Figure 14a and 14b showed the crude lysate in microcentrifuge tube after cells pellet being lysed using lysis buffer for purification of r-msAdh1 protein under native condition. The volume of sample buffer used to lyse the cells pellet is in accordance to cells culture density (OD_{340}: 0.6 = 75 µL of NP10 lysis buffer), thus the total protein concentration can be standardized.

Based on figure 14a, it can be clearly seen that crude lysate of BL21 (DE3) with an empty pET-41a(+) expression appeared translucent after lysis upon induction. Vice versa with crude lysate from BL21 (DE3) expressing r-*msAdh1* cDNA, only uninduced lysate appears translucent, while others appeared cloudy even after the addition of extra volume of lysis buffer. The crude lysate became cloudier as induction temperature was increased to 27 °C (Figure 14b).

When the crude lysate was centrifuged, the fractions separated as a clear supernatant (soluble fraction) and yellowish pellet (insoluble fraction) at the bottom of the centrifuge tube. To further investigate the insoluble fraction, it was resuspended using NP10 lysis buffer and analysed on SDS-PAGE. The result is shown in Figure 15.



a) Crude lysate of BL21 (DE3) with pET-41a(+)/Empty expression (control).

b) Crude lysate of BL21 (DE3) with pET-41a(+)/r-*msAdh1* expression.

Figure 14: Crude lysate from 1.5 mL culture of BL21 (DE3) (a) from pET-41a(+)/Empty expression (control), (b) from pET-41a(+)/r-msAdh1 expression. Tube A: Crude lysate of uninduced culture, Tube B: Crude lysate of induced culture (4 hours induction at 15 °C), Tube C: Crude lysate of induced culture (6 hours induction at 15 °C) Tube D: Crude lysate of induced culture (4 hours induction at 27 °C) Tube E: Crude lysate of induced culture (6 hours induction at 27 °C).

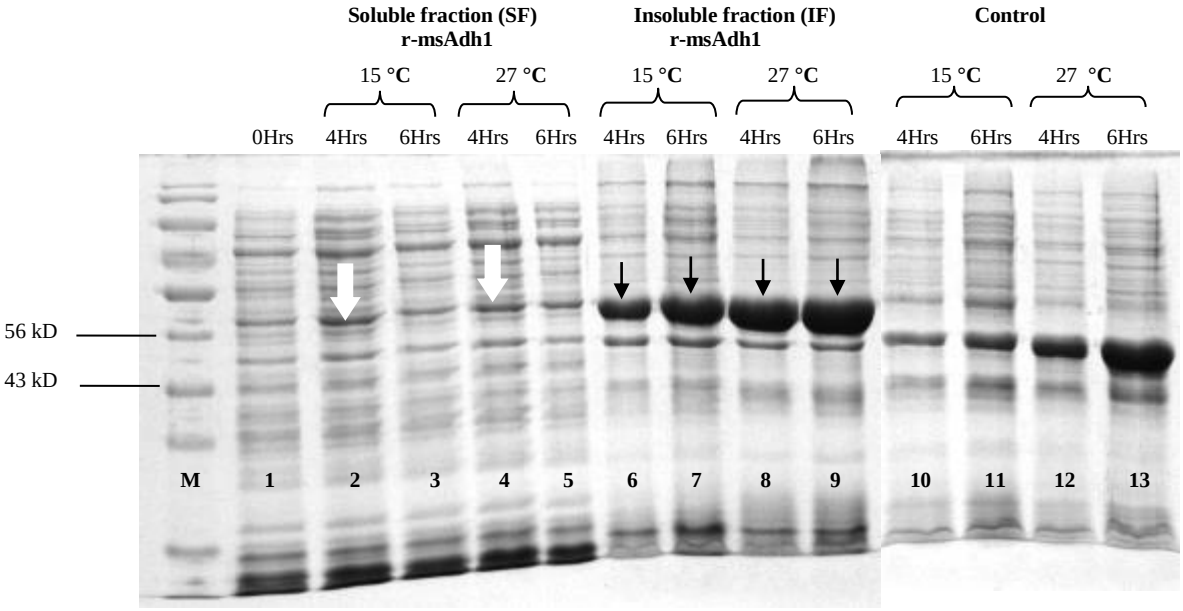


Figure 15: Soluble and insoluble fraction of BL21 (DE3) with pET-41a(+)/r-msAdh1 expression and BL21 (DE3) with pET-41a(+)/Empty expression (control) were visualized on 12 % SDS PAGE stained with commassie blue. Lane M: EZ Run Pre-stained *Rec* Protein Ladder. White arrow shows the expected soluble r-msAdh1 protein band, meanwhile black arrow shows the expected insoluble r-msAdh1 protein band.

Stained gels were analysed for any increased solubility of the target protein; this was judged by increase in the protein band intensity on the SDS-PAGE gel. Based on Figure 15, it can be seen that the intensity of expected r-msAdh1 band is more in Lane 2 and 4. Lane 2 and 4 are soluble fraction from BL21 (DE3) with pET-41a(+)/r-msAdh1 expression that was induced at 15 °C, T: 4 hours and 27 °C, T: 4 hours, respectively. This indicates that BL21 (DE3) grown at lower temperature was able to improve the solubility of r-msAdh1 protein. Expression at low temperature conditions leads to an increase of stability and correct folding patterns due to the fact that hydrophobic interactions determining inclusion body formation are temperature dependent (Lakshmi *et al.*, 2014). Sørensen and Mortensen (2005) also reported that this strategy has proven effective in improving the solubility of a number of difficult proteins.

Meanwhile, as seen in Figure 15 at lane 6-9, bands with very high intensity can clearly be differentiated from control (pET-41a(+)/Empty), and these bands correspond to the r-msAdh1 band size that are seen in the soluble fraction. This indicated that a large amount of r-msAdh1 protein produced were actually present in insoluble fraction. The expected r-msAdh1 band become more intense as the induction temperature was increased to 27 °C and time after induction continued to 6 hours.

Due to the phenomenon observed, two suggestions offered to explain it:

1. The r-msAdh1 protein expressed in BL21 (DE3) formed aggregates called inclusion bodies in BL21 (DE3). The inclusion body formation is common when proteins are overexpressed (Fink, 1998) and many recombinant proteins when overexpressed become insoluble because of misfolding (Trimpin and Brizzard, 2009). According to Dyson *et al.* (2004), protein misfolding occurs when the translation rate exceeds the rate of protein folding. When this happens, *in vivo* equilibrium favours protein aggregation, rather than solubilization (Sorensen and Mortensen, 2005).

2. The soluble r-msAdh1 protein formed aggregation after cells lysis. Leibly *et al.* (2012) hypothesized that a significant fraction of proteins are not found in inclusion bodies but rather are expressed as soluble proteins in *E. coli* and aggregates after cell lysis. In addition, *in silico* characterization has shown r-msAdh1 protein to be hydrophobic, GRAVY value = 0.006 (Appendix F). But GRAVY simply calculates overall hydrophobicity of the linear polypeptide sequence with increasing positive score indicating greater hydrophobicity, but no account is taken of the way the protein folds

in three dimensions or the percentage of residues buried in the hydrophobic core of the protein (Dyson *et al.*, 2004).

3.2.4 Protein purification using NI-NTA Spin Column under native conditions from *E. coli* lysates.

Purification of r-msAdh1 using NI-NTA Spin Column was partially successful. The r-msAdh1 was able to bind to the resin (Figure 16, lane 2, shown by black arrow), but only partial purity was achieved as several contaminant bands were also eluted from purification column (lane 3 and 4). The purification of native protein is challenging as the potential for unrelated, non-tagged proteins to interact with the Ni-NTA resin is usually higher under native than under denaturing conditions (QIAExpressionist, 2003).

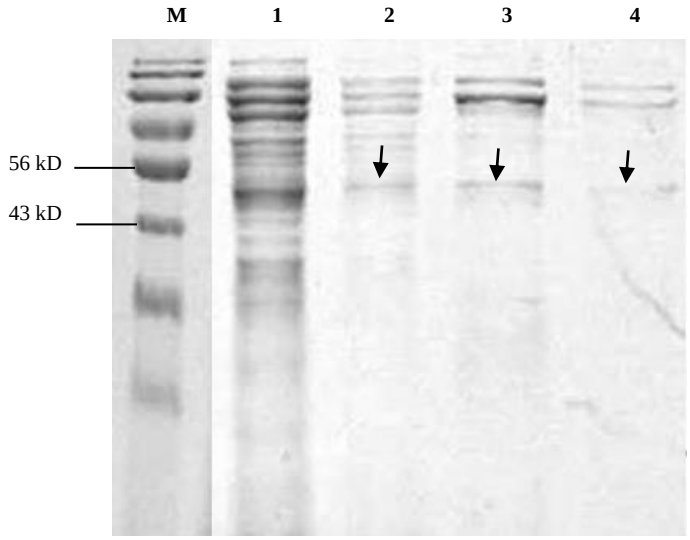


Figure 16: Purification of r-msAdh1 proteins under native conditions using Ni-NTA Spin Columns from *E. coli* cell lysates derived from 20 mL LB cultures. Lane M: EZ Run Pre-stained *Rec* Protein Ladder, Lane 1: Clear lysate, Lane 2: Binding, Lane 3: First elution using 500 mM imidazole, Lane 4: Second elution using 500 mM imidazole. Expected r-msAdh1 protein band shown by black arrow.

3.2.5 Enzymatic assay of ADH enzyme

The activity of r-msAdh1 was assayed using UV spectrophotometry for 15 minutes. Soluble fraction of BL21 (DE3) with r-*msAdh1* expression obtained from section 3.2.3 was used for the assay.

As shown in Figure 17, the ADH enzyme breaks down the ethanol to acetaldehyde in the presence of NAD^+ as a coenzyme to be reacted. Because reduced NAD (NADH) exhibits strong UV absorption at 340 nm while the oxidized form has virtually no absorption at this wavelength, therefore, the catalytic reaction of r-msAdh1 was followed by monitoring the increase in absorbance at 340 nm (Refer to Appendix G).

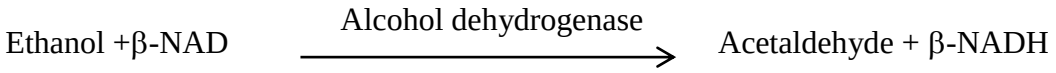


Figure 17: The catalytic reaction of ADH

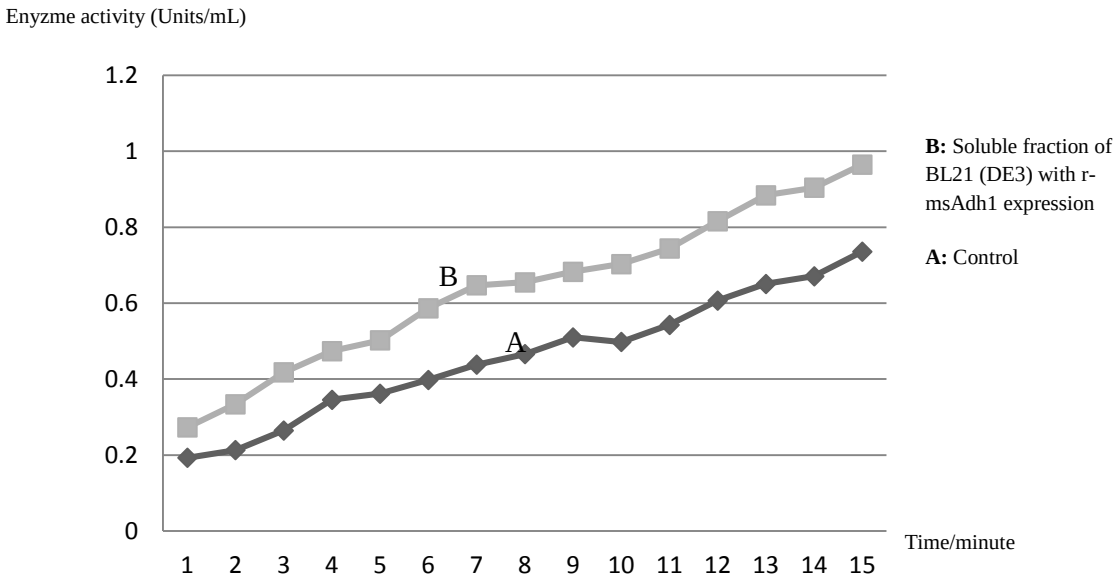
According to the Lambert-Beer law, the extinction of NAD^+ is proportional to ADH enzyme concentration. Therefore, the change in concentration of NAD^+ over time will allow the calculation of the activity of alcohol dehydrogenase in the solution, using the following unit definition for alcohol dehydrogenase: one unit will convert 1.0 μmole of ethanol to acetaldehyde per minute at pH 8.8 at 25 °C. Therefore, one unit will also convert 1.0 μmole of NAD^+ to NADH per minute. By follow this principle, amount of units of r-msAdh1 enzyme present in cell lysate can be estimated using the formula below. The data obtained presented in Table 10.

$$\text{Units/ml enzyme} = \frac{(\Delta A_{340nm}/\text{min Test} - \Delta A_{340nm}/\text{min Blank})\ (1.0)\ (df)}{(6.22)\ (V)}$$

1.0 = Total volume (in mL) of assay
df = Dilution factor
6.22 = Millimolar extinction coefficient of β-NADH at 340nm
V = Volume (in mL) of enzyme used

Table 10: ADH enzyme activity (Units/mL) obtained from (A) soluble fraction of BL21 (DE3) transformed with pET-41a(+)/Empty and (B) soluble fraction of BL21 (DE3) with r-*msAdh1* expression.

Time/minute	Enzyme activity (Units/mL)	
	A	**B**
1	0.193	0.273
2	0.213	0.334
3	0.265	0.418
4	0.346	0.474
5	0.362	0.502
6	0.398	0.587
7	0.438	0.647
8	0.466	0.655
9	0.510	0.683
10	0.498	0.703
11	0.543	0.744
12	0.607	0.816
13	0.651	0.884
14	0.671	0.904
15	0.736	0.965



Graph 1: Graph showed catalytic activity of ADH enzyme present in soluble fraction of BL21 (DE3) with r-*msAdh1* expression and BL21 (DE3) with pET-41a(+)/Empty expression (control).

The data obtained in Table 10 is illustrated in Graph 1 as shown above. Based on the result, catalytic activity of ADH enzyme was detected not only in soluble fraction of BL21 (DE3) with r-*msAdh1* expression but also in soluble fraction of BL21 (DE3) with pET-41a(+) expression (control). The catalytic activity of ADH enzyme that present in control is expected because endogenous ADH3 with molecular weight approximately 39 kD is known to be present in cellular BL21 (DE3) (Retrieved on: http://www.ncbi.nlm.nih.gov/protein/YP_003052980.1).

Nevertheless, the activity of r-msAdh1 enzyme can be differentiated from *E. coli* ADH3 as it consistently showed higher absorbance over time as NADH formed (Appendix L). This is proven by statistical analysis using Student‘s Paired t-test (http://www.physics.csbsju.edu/cgi-bin/stats/Paired_t-test) (Appendix H). A p-value of 0.000 < 0.05 indicates there was significant difference between the catalytic activity of ADH

enzyme present in soluble fraction of BL21 (DE3) with r-*msAdh1* expression and BL21 (DE3) transformed with empty pET-41a(+) vector (control).

3.3 Conclusion

First objective of this study was successfully achieved. The open reading frame (ORF) of recombinant alcohol dehydrogenase cDNA from sago palm denominated as r-*msAdh1* cDNA has been successfully cloned into the expression vector; pET-41a(+) and expressed in *E. coli* strain BL21 (DE3). Verification using restriction enzyme analysis and nucleotide sequencing confirmed correct r-*msAdh1* ORF fusion to the C-terminal 8xHistag of pET-41a(+).

Second objective is to express the r-*msAdh1* cDNA in *E. coli* strain BL21 (DE3). Protein quantification using SDS-PAGE analysis revealed that a moderate amount of soluble r-msAdh1 protein band obtained when r-*msAdh1* cDNA was induced using 0.4 mM IPTG and low induction temperature condition (27 °C and 15 °C) for 4 hours incubation period. Meanwhile, third objective is partially achieved as purification of r-msAdh1 protein using NI-NTA Spin column is co-purified with *E. coli* cellular protein. However, the changes of absorbance at wavelength 340 nm indicated that r-msAdh1 cDNA expressed in BL21 (DE3) has the ability to reduce NAD^+ to NADH; thus proving that r-*msAdh1* cDNA expressed in BL21 (DE3) can produces functional r-msAdh1 enzyme.

Nevertheless, the amount of soluble protein obtained was poor as large portion of expressed protein present in insoluble fraction. This might be due to the hydrophobicity of the r-msAdh1 protein (GRAVY: 0.006) and/or the expressed protein formed aggregation known as inclusion bodies. The SDS-PAGE analysis revealed that the solubility of r-msAdh1 protein did not much increase even when the expression temperature reduced to 15 °C. In fact, the

intensity of insoluble protein band appeared thicker when time after induction was continued to 6 hours and expression temperature increased to room temperature; thereby indicating that formation of inclusion bodies is temperature dependent. From these findings, it can be concluded that expression of eukaryote protein in *E. coli* is challenging because it is prokaryotic based system which has several advantages: (1) inability to carry out post-translational modification and disulfide bond formation, (2) expressed protein formed aggregation known as inclusion bodies due to the protein misfolding, (3) instability of mRNA and (4) codon bias.

However, since most of genes can be expressed in many different systems, it is important to determine which system offers the most efficient for the production of the recombinant protein. Expressing the eukaryote gene in plant host is promising since this system offers several advantages over prokaryotic system such as the ability to synthesis of large and complex protein compound while retaining the recombinant protein activities (post-translational modifications). The ideal expression system would be the one that can produces safe and biologically active desired product at the lowest cost. For these reasons, expression of r-*msAdh1* cDNA in plant system was studied and the result is discussed in chapter 4.

3.4 Future Research

Even though the expression of eukaryote gene in prokaryote system encounters many difficulties; it could promise large production of eukaryote protein in short time by adjusting the parameters that govern efficient expression of r-*msAdh1* in *E. coli* such as inducer concentration, temperature and changing the bacterial strain.

Since large amount of r-msAdh1 can be obtained in inclusion bodies, therefore purification under fully denaturing condition using Ni-NTA Spin Column is suggested. After

purifying the inclusion body by dissolving in a protein denaturant or by various concentrations of chaotropic agents such as urea or guanidinium hydrochloride then it can be refolded into its native structure (Yamaguchi and Miyazaki, 2014). Inclusion body proteins that are solubilizing under mild denaturing condition are better in refolding yields and retaining its biological activities (Sahdev *et al.*, 2008). On the other hand, highly efficient production of soluble proteins from insoluble inclusion bodies has been reported by Yang *et al.* (2011) by using two-step-denaturing and refolding method. Meanwhile, if r-msAdh1 protein tends to precipitate in lysis buffer, this protein would appear in the soluble fraction if the cell lysis buffer conditions were adjusted, whether by pH, ionic strength or presence of an additive (Leibly *et al.*, 2012).

CHAPTER FOUR

CONSTRUCTION OF RECOMBINANT ALCOHOL DEHYDROGENASE 1 (r-*msAdh1*) INTO BINARY VECTOR AND EXPRESSION IN PLANT SYSTEM

4.0 Overview

In the previous chapter, expression of r-*msAdh1* cDNA in the prokaryote system; *E. coli* had been discussed. Meanwhile, this chapter will focus on expression of r-*msAdh1* cDNA in plant system. Amongst several plants use for genetic transformation for variety of purposes, tomato is one of the plants that have been extensively and widely used. This is because, tomato has a relatively small diploid genome; 12 chromosomes (Hasan *et al.,* 2008) with hundreds of mapped traits and comprised more than 1000 molecular marker (Tanksley, 1993) and it is transformable (McCormick *et al.,* 1986). In this study, the transformation using *Agrobacterium tumefaciens* (*A. tumefaciens)* has been chosen for introduction of r-*msAdh1* cDNA in tomato genome because of its simplicity and efficiency (Somayaji *et al.,* 2014). To achieve the goal, the r-*msAdh1* was firstly cloned into the binary vector, pGSA1131 that is driven by CAMV 35S promoter. Then, the newly construct; pGSA1131/r-msAdh1 was transformed into *A. tumefaciens* strain LBA4404 and subsequently used for the *in planta* tomato seeds transformation. The integration of r-*msAdh1* cDNA in transformed plant was confirmed by PCR. The objectives of this chapter are includes;

1. To construct a binary vector that contains r-*msAdh1* cDNA.
2. To transform *Lycopersicum esculentum* cultivar MT1 with the binary vector.
3. To analyse the integration and expression of r-*msAdh1* in transgenic tomato.

The workflow is briefly illustrated in Figure 18 below.

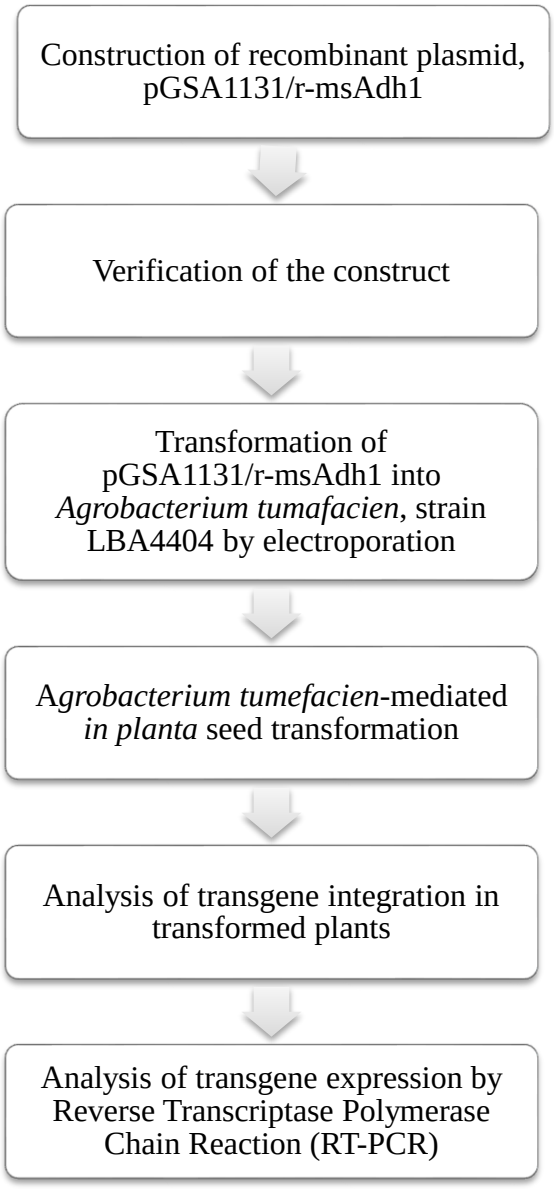


Figure 18: Workflow for construction of r-*msAdh1* in binary vector and their expression in plant system.

4.1 Materials and Methods

4.1.1 Construction of r-*msAdh1* cDNA in the binary vector; pGSA1131

4.1.1.1 Directional cloning of r-*msAdh1* using specific primers

The sago palm r-*msAdh1* cDNA was modified by specific PCR amplification primers to introduce a 5' *Nde*I site at the translational start codon, and a 3' *Bam*HI site at the stop codon. The forward and reverse primers used for the PCR amplifications were the following: 5_NcoI_msAdh1 (5'-GGAATTCCATGGCAAGCAGTGTTGGTCAA-3') and 3_BamHI_msAdh1 (5'-ACCAAGGATCCTTAGTGGTGGTGGTG-3'). The PCR reaction profile used is shown in Table 11. Meanwhile the composition of the PCR reaction mixture is shown in Table 12.

Table 11: The PCR reaction profile to generate the full length of r-*msAdh1* cDNA. PCR was undertaken using the 5_NcoI_msAdh1 and 3'_BamHI_msAdh1 primers combination

Segment	Temperature (°C)	Duration	Number of cycles
Initial denature	94	3 m	1x
Denaturation	94	30 s	35x
Annealing	64	1 m	
Elongation	72	1 m 30 s	
Initial denature	72	10 m	1x

Table 12: The composition of 1X PCR mixture to generate the full length of r-*msAdh1* cDNA using the 5_NcoI_msAdh1 and 3'_BamHI_msAdh1 primers combination

Reagents	**1X (µL)**
10x High Fidelity buffer with 15mM $MgCl_2$	2.5
dNTP (2.5 mM)	2.5
5_NcoI_msAdh1 (10 µM)	1
3'_BamHI_msAdh1(10 µM)	1
Plasmid pET-41a(+)/r-msAdh1 (1.0 µg/µL)	0.5
High Fidelity enzyme (0.625 U/µL)	1
Nuclease-free water	16
Total reaction	**25**

Next, the PCR product was purified using the PCR Cleanup Kit (Vivantis, USA) and subsequently digested with *Bam*HI (Fermentas, Lithuania) restriction enzyme for 5 hours at 37 °C. The composition for the reaction is shown in Table 13. Thermal inactivation was done at 80 °C for 20 minutes. After the digestion, r-*msAdh1* cDNA was purified from the digestion mixture using the PCR Cleanup Kit (Vivantis, USA).

The purified r-*msAdh1* then subjected to the *Nco*I (Fermentas, Lithuania) restriction enzyme digestion using composition as shown in Table 14 for 5 hours at 37 °C and thermal inactivation was done for 20 minutes at 65 °C. After that, the r-*msAdh1* cDNA purified using PCR Clean-up Kit (Vivantis, USA) and kept in - 20 °C.

Table 13: The composition of restriction enzyme digestion to produce 3' of *Bam*HI-sticky ended of r-*msAdh1* fragment. The mixture was prepared triplicates.

Reagents	**Volume (µL)**
Purified r-*msAdh1* fragment (1.0 µg/µL)	14
Buffer *Bam*HI (10X)	2
*Bam*HI (10 U/µL)	4
Total reaction	**20**

Table 14: The composition of restriction enzyme digestion to produce r-*msAdh1* fragment with 5' *Nco*I-sticky ended. The mixture was prepared triplicates.

Reagents	**Volume (µL)**
Purified r-*msAdh1* fragment (0.5 µg/µL)	14
Tango buffer (10X)	2
*Nco*I (10 U/µL)	4
Total reaction	**20**

Meanwhile, isolation of the binary vector; pGSA1131 was done using GF-1 Plasmid DNA Extraction Kit (Vivantis, USA) and subsequently subjected to *Bam*HI (Fermentas, Lithuania) restriction enzyme digestion using composition shown in Table 15 for 8 hours at 37 °C. Thermal inactivation was done at 80 °C. The linearized pGSA1131 then purified using PCR-Clean up Kit (Vivantis, USA).

Table 15: The composition of restriction enzyme digestion to produce pGSA1131 with 3' *BamH*I-sticky ended. The mixture was prepared triplicates.

Reagents	**Volume (µL)**
Extracted plasmid DNA, pGSA1131 (1.0 µg/µL)	14
Buffer *BamH*I (10X)	2
*BamH*I (10 U/µL)	4
Total reaction	**20**

Next, purified linearized pGSA1131 was subjected to digestion of *Nco*I (Fermentas, Lithuania) restriction enzyme using composition as shown in Table 16 for overnight period at 37 °C. Thermal inactivation was done at 65 °C for 20 minutes.

Table 16: The composition of restriction enzyme digestion to produce pGSA1131 fragment with 5' *Nco*I-sticky ended. The mixture was prepared triplicates.

Reagents	**Volume (µL)**
Linearized pGSA1131 (0.5 µg/µL)	14
Tango buffer (10X)	2.5
*Nco*1 (10 U/µL)	4
Nuclease-free water	4.5
Total reaction	**25**

After that, the fragments were separated by running an agarose gel electrophoresis at 100 V with the expected band sizes of approximately 395 bp and 8.999 kb. The desired fragment (~ 8.999 kb) was extracted from the gel using GF-Gel Recovery Kit (Vivantis, USA). Subsequently, the r-*msAdh1* fragments that recovered by *Nde*I/*Bam*HI digestion was ligated into the *Nde*I/*Bam*HI restricted vector pGSA1131 using T4 DNA Ligase (Fermentas, Lithuania) with incubation at 22 °C for 2 hours using composition shown in Table 17.

Table 17: The composition of ligation mixture to clone r-*msAdh1* into the binary vector, pGSA1131

Reagents	**Volume (µL)**
Linearized vector pGSA1131 (0.5 µg/µL)	5
Insert (Purified r-*msAdh1* fragment) (0.5 µg/µL)	9
T4 DNA Ligase Buffer (10X)	2
T4 Ligase (1 weiss U/µL).	1
Nuclease-free water	3
Total reaction	**20**

Next, competent *E. coli* strain XL1-Blue cells were transformed with the ligation mixture by heat shocked method (section 3.1.3, Chapter 3). Subsequently, the transformation culture spread on LA plates supplemented with 30 µg/mL chloramphenicol.

4.1.1.2 Plasmid isolation

A single colony was selected from transformation plate and inoculated into a 50 mL falcon tube containing 10 mL of LB supplemented with 30 µg/mL chloramphenicol. The tube was then cultured overnight at 37 °C with shaking at 150 rpm. Subsequently, 6 mL of overnight culture was used for plasmid isolation using GeneMATRIX™ Plasmid Miniprep DNA Purification Kit (Eurox, USA).

4.1.1.3 Verification of reading frame via PCR, restriction enzyme digestion analysis and nucleotide sequencing

Verification of r-*msAdh1* nucleotide sequence in pGSA1131 was undertaken to ensure the proper open reading frame. The verification was done via PCR using specific r-*msAdh1* primers. The PCR reaction profile is shown in Table 18 and the composition of PCR mixture is shown in Table 19.

Table 18: The PCR reaction profile to screen for the presence of r-*msAdh1* in the pGSA1131

Segment	Temperature (°C)	Duration	Number of cycles
Initial denature	94	3 m	1x
Denaturation	94	30 s	35x
Annealing	62	30 s	
Elongation	72	45 s	
Initial denature	72	5 m	1x

Table 19: The composition of 1X PCR mixture to screen for the presence of r-*msAdh1* in the pGSA1131. The PCR was undertaken using the 5_NcoI_msAdh1 and 3_BamHI_msAdh1 primers combination.

Reagents	**1X (µL)**
2X Green Go-Taq Master Mix (Promega)	7.5
5_NcoI_msAdh1 (10 µM)	1
3_BamHI_msAdh1 (10 µM)	1
Extracted plasmid, pGSA1131/r-*msAdh1* (1.0 µg/µL)	1
Nuclease-free water	4.5
Total volume	15

Verification of the open reading frame of the new construct was also carried out by restriction enzyme digestion analysis. Single and double digestions were performed using composition shown in Table 20. The reaction mixture was incubated overnight at 37 °C. The fragments were separated by running an agarose gel electrophoresis, 1 % at 100 V for 30 minutes with expected sizes of approximately 8.999 kb and 1.174 kb for double digestion of pGSA1131/r-msAdh1, approximately 10.173 kb for single digestion of pGSA1131/r-msAdh1 and approximately 9.394 kb for single digestion of pGSA1131 (control).

Table 20: The composition of restriction enzyme digestion to verify the open reading frame of pGSA1131/r-msAdh1. Single enzyme digestion was using *Nco*I, meanwhile double enzyme digestion was using *Nco*I and *BamH*I.

Reagents	Single digestion of pGSA1131/r-msAdh1 and pGSA1131 (µL)	Double digestion of pGSA1131/r-msAdh1 (µL)
Plasmid (1.0 µg/µL)	7	13
Tango Buffer (10X)	2	4
*Bam*HI (10 U/µL)	-	1.5
*Nco*I (10 U/µL)	1	1.5
Nuclease-free water	-	-
Total Volume	**10**	**20**

For further verification, the new construct also sent to First BASE Laboratories Sdn Bhd (Selangor, Malaysia) for the nucleotide sequencing using two set of forward and reverse primers as following:

5_Nco1_msAdh1 (5'-GGAATTCCATGGCAAGCAGTGTTGGTCAA-3),

3_BamH1_msAdh1 (5'-ACCAAGGATCCTTAGTGGTGGTGGTG-3'),

Adhmor8_F (5'-CTAGAGCTTCAGGGGCATCA-3'),

5'_msAdh1_R (5'-AACACAGCCAACATGGACAA-3').

4.1.2 Transformation of pGSA1131/r-msAdh1 into tomato seeds

4.1.2.1 Preparation of *A. tumefaciens* strain LBA4404 electro-competent cells

Ten microliters of thawed frozen glycerol stock of *A. tumefaciens* strain LBA4404 (donated by Evra Raunie Ibrahim from Craun Research Sdn Bhd, Sarawak, Malaysia) was streaked onto LA plate supplemented with 100 µg/mL rifampicin and allowed to grow at room temperature for 2-3 days. Then, a single colony was selected from streaking plate and inoculated into a 50 mL falcon tube containing 10 mL LB supplemented with 100 µg/mL rifampicin.

The culture was then grown for 2 days at room temperature with shaking at 180 rpm. Next, 300 µL of culture was aliquot and inoculated into a 50 mL falcon tube containing 10 mL LB supplemented with 100 µg/mL rifampicin. The tube was then incubated at room temperature with shaking at 180 rpm to $O.D_{500}$ of 1.0. Upon reaching the desired density, cells were pelleted by centrifugation at 3000 x g for 5 minutes at 4 °C. The broth was carefully poured off and the cells pellet resuspended with 3 mL of sterilized, ice-cold 10 % glycerol. Next, the resuspended cells were centrifuged at 3000 x g for 3 minutes at 4 °C and the supernatant discarded. After that, the cells pellet was resuspended again with 1.5 mL ice-cold 10 % glycerol and centrifuged at 3000 x g for 3 minutes at 4 °C. Lastly, the supernatant was discarded and the cells pellet was resuspended with 1 mL ice-cold 10 % glycerol and stored at -80 °C for long storage. The preparation of electro competent cell was followed as recommended by BIORAD (USA) with minor modifications.

4.1.2.2 Transformation of pGSA1131/r-msAdh1 into *A. tumefaciens* strain LBA4404 via electroporation

A pre-cooled 1.5 mL microcentrifuge tube was added with 1.5 µL of recombinant plasmid, pGSA1131/r-msAdh1 and 50 µL *A. tumefaciens* competent cells. The cells-DNA mixture was then transferred into a pre-chilled electroporation cuvette and gently tapped until the mixtures settled evenly at the bottom. The cuvette was slide into the electroporation chamber and then pulse once. After that, the cuvette was quickly removed and immediately added with 1 mL of pre-warmed LB to transfer the cells from cuvette into a 1.5 mL microcentrifuge tube. Next, the tube was incubated at room temperature, with gentle agitation at 100 rpm for overnight. The following day, 100 µL of overnight culture was aliquot and spread onto LA supplemented

with 100 µg/mL rifampicin and 30 µg/mL chloramphenicol. The transformation plate was incubated at room temperature for 2-3 days.

4.1.2.3 Colony PCR

Three colonies were randomly selected from transformation plates and inoculated into a 50 mL falcon tube containing 10 mL LB supplemented with 100 µg/mL rifampicin and 30 µg/mL chloramphenicol, respectively. All tubes then cultured for 2 days at room temperature with shaking at 180 rpm. To quickly screen for the plasmid containing a desired insert directly from the bacterial culture, 50 µL of two days culture was aliquoted and centrifuged at 3500 rpm. After that, supernatant was discarded and cells pellet resuspended with 300 µL sterilized distilled water. Subsequently, 1 µL of suspended cells was used as a template in colony PCR. The PCR result was analysed by running an agarose gel electrophoresis constantly at 100 V for 30 minutes.

4.1.2.4 Plasmid isolation

For further verification, 6 mL of *A. tumefaciens* culture harbouring pGSA1131/r-msAdh1 was used for plasmid isolation using GF-1 Plasmid DNA Extraction Kit (Vivantis, USA). The remaining culture left was added with 20 % of glycerol and mixed well before kept at -80 °C for long storage. The extracted plasmid was subsequently sent to First BASE Laboratories Sdn Bhd (Selangor, Malaysia) for nucleotide sequencing using r-*msAdh1* specific forward and reverse primers.

4.1.2.5 Preparation of *A. tumefaciens* culture for tomato seed transformation

Hundred microliters of thawed frozen glycerol stock of *A. tumefaciens* stock culture harbouring pGSA1131/r-msAdh1 was inoculated into a 50 mL falcon tube containing 10 mL LB supplemented with 100 µg/mL rifampicin and 30 µg/mL chloramphenicol. They were then grown at room temperature with shaking at 180 rpm for 2 days. After that, 600 µL was aliquot and added into a 50 mL falcon tube containing 30 mL fresh LB supplemented with 100 µg/mL rifampicin and 30 µg/mL chloramphenicol. The culture was then brought to an $O.D_{600}$ of 0.5-0.6. Upon reaching the desired O.D. value, the cells were centrifuged at 3500 rpm for 5 minutes at room temperature. Cells were then resuspended in an infiltration media [0.5X MS; 3 % sucrose; 0.5 g/L MES]. Subsequently, 0.003 % Silwet L-77 and 200 µM acetosyringone (*Phyto* Technologies Laboratories™, USA) were added.

4.1.2.6 Tomato seeds preparation

The tomato (*Solanum lycopersicum*) variety used for experiments was the cultivar MT1, which was purchased from Malaysia Research Institute (MARDI). About 50 seeds were washed with sterilized distilled water for several times. Then, the seeds were immersed in sterilized distilled water in a 100 mL Erlenmeyer flask and incubated overnight at 4 °C.

4.1.2.7 Infiltration

The following day, the seeds which are in the water solution in a 100 mL Erlenmeyer flask were sonicated for 10 minutes. Later, the seeds were dabbed dry and added into the infiltration broth with *A. tumefaciens* harbouring pGSA1131/r-msAdh1 that prepared in section 4.1.2.5. They were then incubated at room temperature with gentle shaking at 120 rpm for 3 hours.

4.1.2.8 Post-infiltration

After 3 hours, the infiltration broth was drained from the seeds. Then, the seeds were rinsed with distilled water several times and then treated with 500 μg/mL carbenicillin for an hour. After that, seeds were sown into the ground. The seeds were germinated after 2-4 days soil onto the ground.

4.1.3 Analysis of r-*msAdh1* and *bar* gene integration in tomato genome

4.1.3.1 Genomic DNA extraction

Young leaves of putative transformed seedling, T_0 were surfaced sterilized for 2 minutes in 75 % ethanol, then 2 minutes in 10 % Clorox and followed by three washes in sterilized distilled water. Genomic DNA was isolated using GF-1 Plant DNA Extraction Kit (Vivantis, USA). Then, quantity and purity of the extracted genomic DNA (gDNA) was measured using UltrospecR 1100 *pro* (Amersham Pharmacia Biotech, USA).

4.1.3.2 Analysis of r-*msAdh1* integration in putative transformed seedlings (T_0 generation) and transgenic progeny seedling (T_1 generation)

The integration of r-*msAdh1* in putative transformed seedlings, $T_0$1-$T_0$15 and transgenic progeny seedling, $T_1$1-$T_1$11 were screened by PCR using r-*msAdh1* specific primers: 5_Comseq_Adh (5'-ATGGCAAGCAGTGTTGGTCAAGTGATC-3') and 3_Comseq_Adh (5'- ACCATCCATGTGAATGATGCACCTAAGGC-3'). The PCR reaction profile shown in Table 21 and the composition of the reaction mixture is shown in Table 22.

Table 21: The PCR reaction profile to screen the integration of r-*msAdh1* in in putative transformed seedling, T_0 and transgenic progeny seedling, T_1.

Segment	Temperature (°C)	Duration	Number of cycles
Initial denature	94	3 m	1x
Denaturation	94	30 s	35x
Annealing	57	45 s	
Elongation	72	1 m 30 s	
Initial denature	72	10 m	1x

Table 22: The composition of 1X PCR reaction mixture to screen the integration of r-*msAdh1* in putative transformed seedlings, T_0 and transgenic progeny seedling, T_1. PCR was undertaken using the 5_comseq_Adh and 3'_Comseq_Adh primers combination.

Reagents	1X (µL)
2X Green Go-Taq Master Mix (Promega)	7.5
5_comseq_Adh (10 µM)	1
3'_comseq_Adh (10 µM)	1
Genomic DNA (0.5 µg/µL)	1
Nuclease-free water	4.5
Total volume	**15**

4.1.3.3 Analysis of *bar* gene integration in putative transformed seedlings (T_0 generation)

Molecular analysis of *bar* gene was only carried out for the generation of putative transformed seedling; T_0 as the priority of this study is r-*msAdh1*. Gradient PCR was performed using annealing temperature ranging from 55 °C – 68 °C to amplify GC-rich *bar* sequence from the gDNA template of $T_0$1. The amplification of *bar* gene was attempted using the *bar* gene specific primers: Bar3_F (5'-ATGAGCCCAGAACGACGCC-3') and Bar3_R (5'-ATCTCGGTGACGGGCAGG-3'). The PCR reaction profile shown in Table 23 and the composition of the reaction mixture is shown in Table 24.

Table 23: The gradient PCR reaction profile to determine the annealing temperature for amplification of *bar* gene in putative transformed seedling, $T_0$1.

Segment	Temperature (°C)	Duration	Number of cycles
Initial denature	94	3 m	1x
Denaturation	94	30 s	35x
Annealing	55-68	45 s	
Elongation	72	30 s	
Initial denature	72	10 m	1x

Table 24: The composition of 1X PCR to determine the annealing temperature for amplification of *bar* gene in putative transformed seedling, $T_0$1. PCR was undertaken using Bar3_F and Bar3_R primers combination.

Reagents	1X (µL)
10x High Fidelity buffer with 15mM $MgCl_2$	2.5
dNTP (2.5 mM)	2.5
Bar3_F (10 µM)	1
Bar3_R (10 µM)	1
Genomic DNA (0.5 µg/µL)	1
High Fidelity enzyme (0.625 U/µL)	1
Nuclease-free water	16
Total reaction	25

4.1.4 Analysis of r-*msAdh1* expression in transformed seedling, T_0 and transgenic progeny seedling, T_1

4.1.4.1 Total RNA extraction

The transformed T_0 seedling: $T_0$2 and transgenic progeny seedling: $T_1$8-$T_1$11 was selected for expression analysis. Young leaves at the 4-6 stage was used as a sample. Before RNA extraction, the leaves were surfaced sterilized with 75 % ethanol, followed by 10 % Clorox, and subsequently rinsed thrice with sterilized distilled water with continuous shaking for 2

minutes. Next, in a pre-cooled mortar, sample was ground to fine powder with liquid nitrogen. The powder tissue was then transferred into a 1.5 mL microcentrifuge tube. Total RNA extraction was performed using Total RNA Mini Kit Plant (Geneaid, Taiwan).

4.1.4.2 DNAse treatment

Ten microliters of the total RNA was treated with DNAse I (Promega) to removes traces of genomic DNA. The composition for DNase I treatment is shown in Table 25. The reaction mixture was incubated at 37 °C for 30 minutes. Thermal inactivation was done by adding 1 µL of 10 mM EDTA (Fermentas, Lithuania) into reaction mixture and subsequently incubated at 65 °C for 10 minutes. The quantity and purity of the total RNA obtained then measured using UltrospecR 1100 *pro* (Amersham Pharmacia Biotech, USA). Meanwhile, to assess the integrity of total RNA extracted, an aliquot of RNA was running on 1.5 % of agarose gel.

Table 25: The composition of reagent used for DNase treatment

Reagents	**Volume (µL)**
Extracted total RNA (0.5 µg/µL)	10
10X reaction buffer with $MgCl_2$	1
DNase 1, RNase-free (1 u/uL)	1
Total reaction	**12**

4.1.4.3 First strand cDNA synthesis

First strand cDNA synthesis of r-*msAdh1* was carried out according to RevertAid First strand cDNA synthesis kit protocol (Fermentas, Lithuania). Before started, all reagents were thawed on ice and then briefly centrifuged. Approximately 2 µL of 3_Comseq_Adh (5'-ACCATCCATGTGAATGATGCACCTAAGGC-3'), 2 µL 3_BamH1_msAdh1 (5'-

ACCAAGGATCCTTAGTGGTGGTGGTG-3') and 2 µL of adaptor $(dt)_{17}$ (GACTCGAGTCGACATCGATTTTTTTTTTTTTTTT-3') was added respectively into a 1.5 mL RNAase-free microcentrifuge tube that contains 10 µL of RNA. All tubes were then incubated at 65 °C for 5 minutes, then briefly centrifuged and placed on ice for 5 minutes. After that, the reagent for cDNA synthesis was added into the tube in indicated order as shown in Table 26. The mixture was then briefly centrifuged and incubated at 42 °C for 1 hour and half. Thermal inactivation was done at 75 °C for 10 minutes to stop the reaction.

Table 26: The composition of reagent mixture to synthesis first strand cDNA

Reagents	**Volume (µL)**
Total RNA plus primers	12
5X reaction buffer	4
Ribolock RNase inhibitor (20 U/ µL)	0.5
dNTP mix, 10 mM each	2
DEPC treated water	0.5
RT M-MuLV Reverse Transcriptase 200U/ µL	1
Total reaction	**20**

First strand cDNA synthesis of the transgenic progeny seedling; $T_1$8-$T_1$11 were also carried out. The treatment is same as described above except the first strand cDNA synthesis was priming by r-*msAdh1* specific internal primer; 5_msAdh1_R (5'-AACACAGCCAACATGGACAA-3').

4.1.4.4 Amplification of r-*msAdh1* by Reverse-Transcriptase Polymerase Chain Reaction (RT-PCR)

The presence of r-*msAdh1* expression in putative transformants, $T_0$2 and transgenic progeny seedling, ($T_1$8-$T_1$11) were determined by RT-PCR using combination of r-*msAdh1* cDNA specific forward and reverse primers: 5_Comseq_Adh (5'-ATGGCAAGCAGTGTTGGTCAAGTGATC-3') and 5'_msAdh1_R (5'-AACACAGCCAACATGGACAA-3'). The PCR reaction profile is shown in Table 27 and the composition of PCR mixture is shown in Table 28.

Table 27: RT-PCR reaction profile to determine the expression of r-*msAdh1*

Segment	Temperature (°C)	Duration	Number of cycles
Initial denature	94	3 m	1x
Denaturation	94	30 s	35x
Annealing	55	30 s	
Elongation	72	45 s	
Initial denature	72	5 m	1x

Table 28: The composition of 1X RT-PCR mixture to determine the expression of r-*msAdh1* using the 5_Comseq_Adh and 5'_msAdh1_R primers combination.

Reagents	1X (µL)
2X Dream-Taq Green PCR Master Mix (Promega)	12.5
5_comseq_Adh (10 µM)	1
5'_msAdh1_R (10 µM)	1
cDNA	2
Nuclease-free water	9.5
Total volume	**25**

4.2 Results and Discussion

4.2.1 Construction of pGSA1131/r-msAdh1

4.2.1.1 Directional cloning of r-*msAdh1* PCR product using specific primers

The used of primers combination; 5'NcoI_msAdh1 and 3_BamHI_msAdh1 in PCR to amplify full length of r-*msAdh1* cDNA had resulted amplification of approximately 1.184 kb fragment on 1 % agarose gel electrophoresis. The amplified r-*msAdh1* cDNA fragments containing endonuclease restriction sites (*Nde*I and *Xho*I-8x histidine nucleotides) on 5' and 3', respectively. The photograph of agarose gel electrophoresis result is shown in Figure 19.

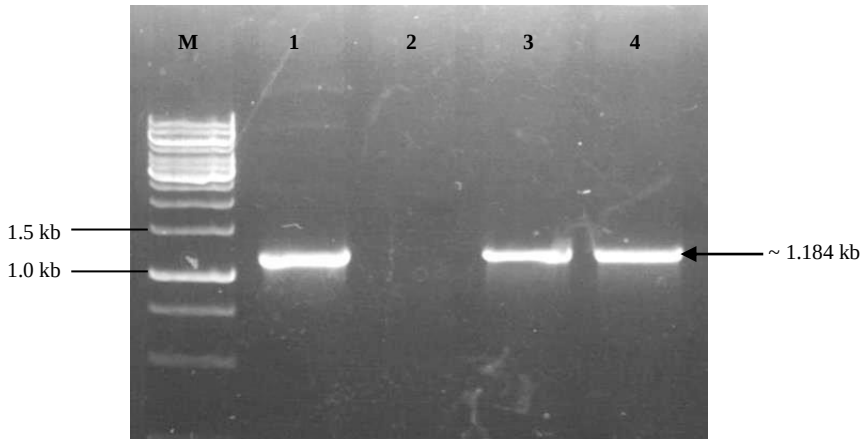


Figure 19: Gel electrophoresis of PCR product visualized on 1 % of agarose gel stained with EtBr. The used of 5'NcoI_msAdh1 and 3_BamHI_msAdh1 primers in PCR had produced r-*msAdh1* band with expected size of 1.1 kb. Lane M: GeneRuler™ 1 kb DNA Ladder (Fermentas, Lithuania), Lane 2: Negative control, Lane 1, 3 and 4: PCR product.

The PCR product was purified and subsequently digested using restriction enzymes; *Nco*I (Fermentas, Lithuania) and *BamH*I (Fermentas, Lithuania). There was no difference in r-*msAdh1* size before and after digestion of the endonuclease as these enzymes removed only a few nucleotides at the 5' and 3' end of the r-*msAdh1* sequence. Meanwhile, the vector; pGSA1131 was digested using the same restriction enzyme as well, thus removing the *gus*

gene from vector which then replaced with r-*msAdh1* cDNA by ligation. The ligation of r-*msAdh1* cDNA into the binary vector and transformation of this newly construct into XL-Blue competence cells through heat shocked method was successful as many colonies were observed on selection media.

4.2.1.2 Plasmid isolation

The isolated plasmid was analysed by running an agarose gel at 100 V for 30 minutes. The electrophoresis result of agarose gel showed a single band with size smaller than expected (Figure not shown). This indicates that the extracted plasmid was in supercoiled form. The exact molecular weight of plasmid only can determine if they were in linearized form as shown in Figure 20 (Page 69).

4.2.1.3 Verification of pET-41a(+)/r-msAdh1 reading frame via restriction enzyme digestion analysis and nucleotide sequencing.

The restriction enzyme analysis was carried out to prove the inclusion of r-*msAdh1* cDNA in the binary vector, pGSA1131. Figure 20 showed the result of gel electrophoresis for restriction digestions of both plasmids; pGSA1131 and pGSA1131/r-msAdh1. Single digestion using *Nco*I only cut the supercoiled plasmid; pGSA1131/r-msAdh1 and pGSA1131 at only one site, thus converting the supercoiled plasmid into a linear form with size approximately of 10.204 kb and 9.394 kb respectively (Lane 2 and Lane 4). This step was conducted to estimate the exact size of both plasmids.

Meanwhile, the double digestion of pGSA1131/r-msAdh1 using *Nco*I and *BamH*I cut this plasmid at two sites (Lane 5). Result obtained showed that the sizes of the restricted fragments corresponded to the expected sizes calculated from plasmid restriction maps;

~ 9.025 kb (pGSA1131 backbone) and ~ 1.179 kb (r-*msAdh1* cDNA). Based on this result, it proofed and confirmed that the sizes of the restricted fragments obtained approximately the same to the predicted sizes from the calculation thus indicating that the construction of the recombinant vector pGSA1131/r-msAdh1 was successful.

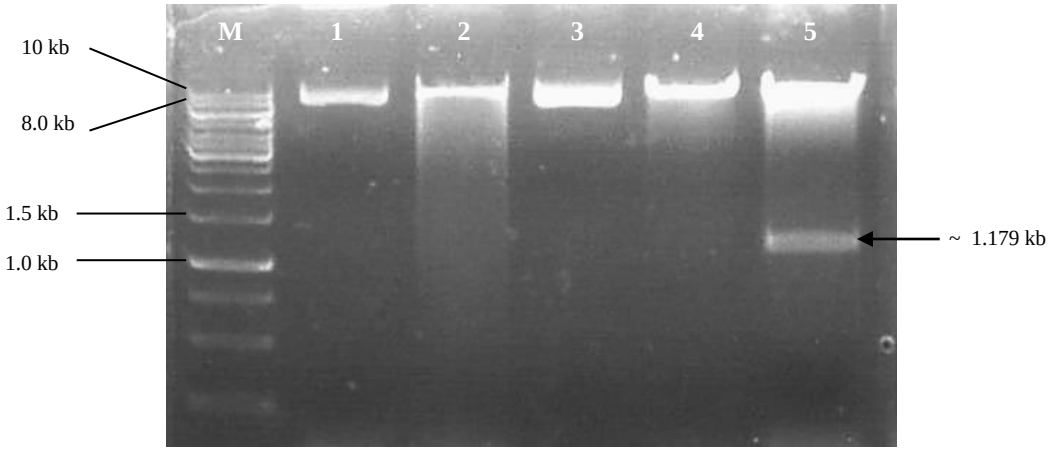


Figure 20: Restriction enzyme analysis of pGSA1131/r-msAdh1. The digestion mixture was visualized on 1 % of agarose gel stained with EtBr. Lane M: GeneRuler™ 1 kb DNA Ladder (Fermentas, Lithuania). Lane 1: Uncut supercoiled pGSA131, Lane 2: Linearized pGSA1131**,** Lane 3: Uncut supercoiled pGSA1131/r-msAdh1, Lane 4: Linearized pGSA1131/r-msAdh1, Lane 5: Double digestion of pGSA1131/r-msAdh.

Apart of that, to further verify, the DNA sequencing was also performed using several different primers to ensure the r-*msAdh1* was successfully cloned in correct orientation into the pGSA1131 vector. The different fragments resulted from sequencing were assembled by using overlapping consensus sequence. DNA sequence analysis revealed clearly that r-*msAdh1* clone is approximately 1.140 kb in length and contains a reading frame of 380 amino acids. The junction of the nucleotide sequence of r-*msAdh1* in the pGSA1131 vector is illustrated in Figure 21.

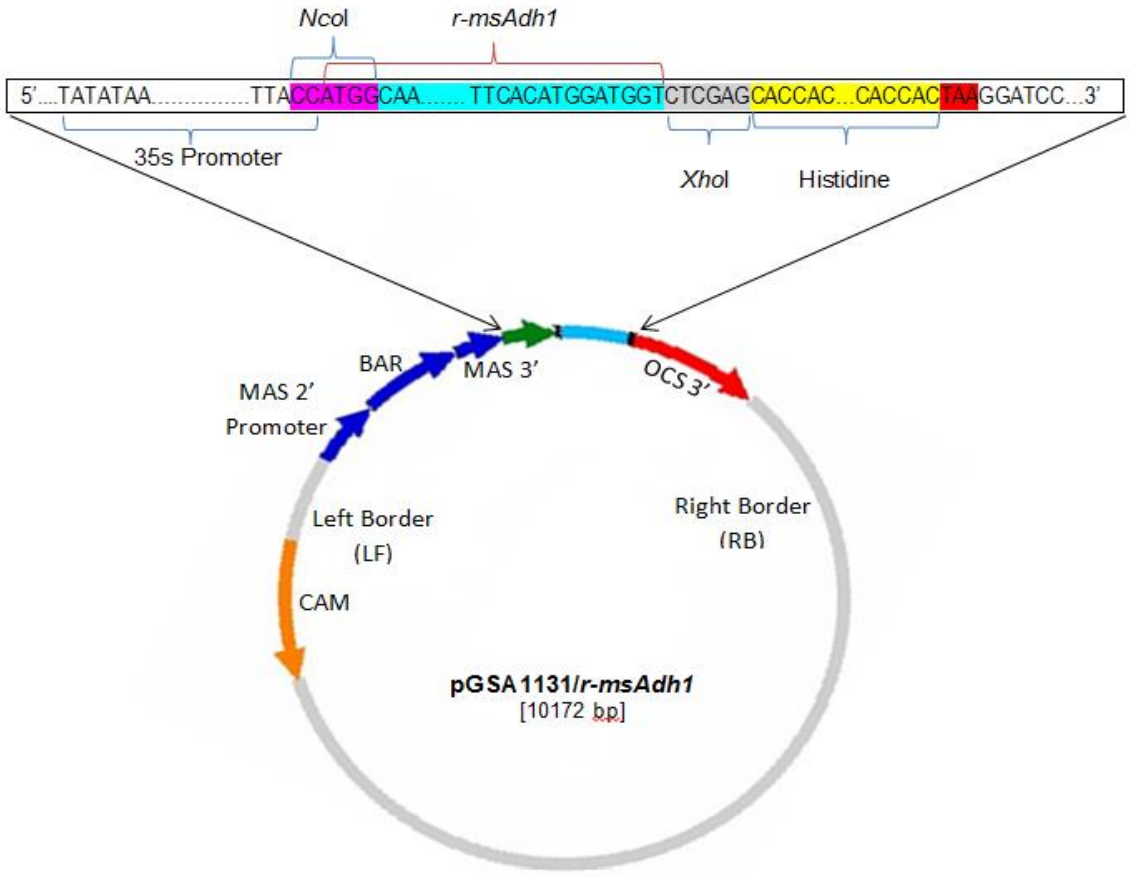


Figure 21: Plasmid map of the newly construct; pGSA1131/r-msAdh1. Map showing left (LB) and right borders (RB) flanking the plant selection genes (*Bar*), r-*msAdh1* cDNA, and a selection genes (CAM) used in bacteria.

4.2.2 Establishment of the tomato genetic transformation system

4.2.2.1 Transformation of pGSA1131/r-msAdh1 into *A. tumefaciens* via electroporation

The verification by restriction enzyme analysis and nucleotide sequencing had proven that pGSA1131/r-msAdh1 was in correct orientation to proceed to the *A. tumefaciens* transformation via electroporation method. The efficiency of transformation through this method was high, however the growth of *A. tumefaciens* is slow as transformed colonies only can be observed on selection plates after 3 days. Three colonies were selected, cultured and their cells pellet were then resuspended with sterilized distilled water. The resuspended cells

were then directly used as a template in the colony PCR. All the selected colonies showed positive PCR result for r-*msAdh1* with size approximately of 1.184 kb when primers 5_ NcoI_msAdh1 and 3_ NcoI_msAdh1 were used. The photograph of gel electrophoresis result is shown in Figure 22. The slight smearing appeared in gel because whole *E. coli* lysate was used as template. Thus, contaminating genomic DNA, plasmid DNA and cellular debris which interfered with DNA migration were present.

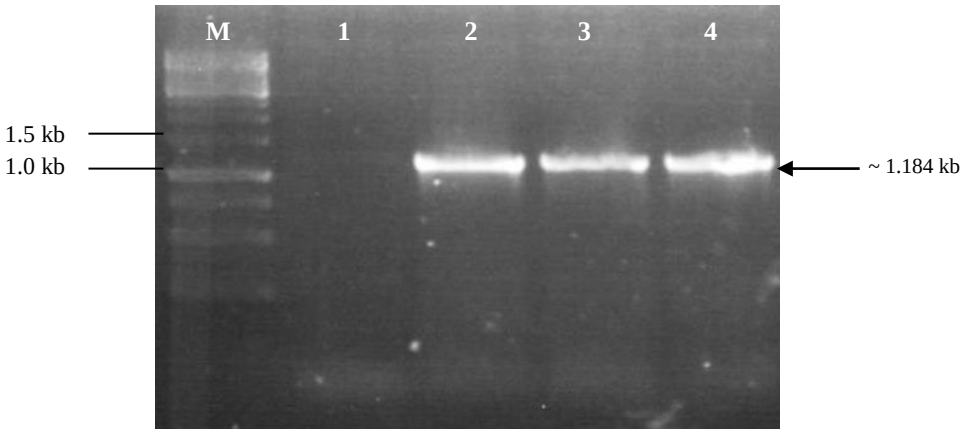


Figure 22: Gel electrophoresis of colony PCR product visualized on 1 % agarose gel stained with EtBr to screen the positive colony harbouring pGSA1131/r-msAdh1. A single band estimated at ~ 1.184 kb was observed for all selected colony**.** Lane M: GeneRuler™ 1 kb DNA Ladder (Fermentas, Lithuania), Lane 1: Negative control (DNA template exclude in PCR reaction), Lane 2-4: PCR result for colony 1, 2 and 3.

4.2.2.2 *Agrobacterium*-mediated transformation using tomato seeds

In this study, the r-*msAdh1* cDNA has been cloned downstream to the 35s promoter. Pratheesh *et al.* (2014) reported that, 35s promoter effectively puts its downstream gene outside virtually any regulatory control by the host genome and expresses the gene at approximately two to three orders of magnitude higher, thus allowing a strong positive selection. A successful *in planta* transformation had been reported by Lee *et al.* (2013) when pGSA1131 containing β –

glucuronidase gene (GUS) that located on downstream to the 35s promoter was used to transform GUS gene into plant *Morinda citrifolia.*

Initially, A. *tumefaciens* harboured pGSA1131/r-msAdh1 was grown to OD_{500nm}: 0.6 in which at this point, the cells were active and at the optimum condition to carry out the infiltration task. Upon reaching the desired density, cells were pelleted and re-suspended into infiltration media containing Silwet L-77 and acetosyringone. Meanwhile, about 50 of tomato seeds were sonicated for 10 minutes to introduce wound. Trick and Finer (1997) reported that sonicating had introduced thousand micro wound on immature soybean cotyledons and greatly enhanced the levels of transient expression. The ruptures on the surface of the plant material are clearly large enough for *Agrobacterium* to invade the wounded cells or tissues (Trick and Finer, 1997). When cells are wounded, it will secrete low molecular weight molecules; acetosyringone and hydroxy-acetosyringone (Chaudhry and Rashid, 2010). The *Agrobacterium* VirA/VirG two-components system will recognize the acetosyringone as a host specific signal and activate *vir* gene expression (Winans, 1990). Further addition of acetosyringone will act as chemical attractant in vitro and may act as a chemotactic agent in nature, thus it was used to enhance the transformation efficiency (Chaudhry and Rashid, 2010). Meanwhile, addition of surfactant such as Silwet L-77 are known to function either as enhancer of cuticle penetration by making the plant cuticular membrane susceptible to solute transfer or acting as co-solvents (Madhou *et al.,* 2006). In addition, application of surfactants will enhances transformation frequency by aiding *A. tumefaciens* attachment or eliminating substances that inhibit bacterial attachment (Curtis and Nam 2001; Huang and Wei 2005).

4.2.2.3 Germination rate

Fifty transformed seeds were sown and from that amount, only 26 seeds germinated giving 52 % germination rate for *Solanum lycopersicum* var. MT1 (MARDI). This might be because; the seeds were not kept in the proper condition to preserve the fertility. Even so, the analysis of transformation was preceded using the remaining germinated seeds. Out of 26 germinated seeds, only 15 plants were available for further analysis.

4.2.3 Analysis of r-*msAdh1* and *bar* gene integration in tomato genome

4.2.3.1 Genomic DNA (gDNA) extraction from putative transformed seedling, T_0

Young leaf with size approximately 0.5 cm^3 in diameter was used for genomic DNA extraction. An aliquot of gDNA was then analysed on 1 % of agarose gel (Figure 23). As can be seen in the figure, high molecular weight of band was observed in all wells, indicating successful extraction. However, smearing towards the end of the gel suggested that degradation of DNA had occurred during extraction. The genomic DNA were quantified at 260 nm and the purity was calculated using 260/280 nm absorbance ratio (UltrospecR 1100 *pro*, Amersham Pharmacia Biotech, USA), (Table 32). The genomic DNA obtained subsequently used as a template for PCR amplification to screen and identify positive plants.

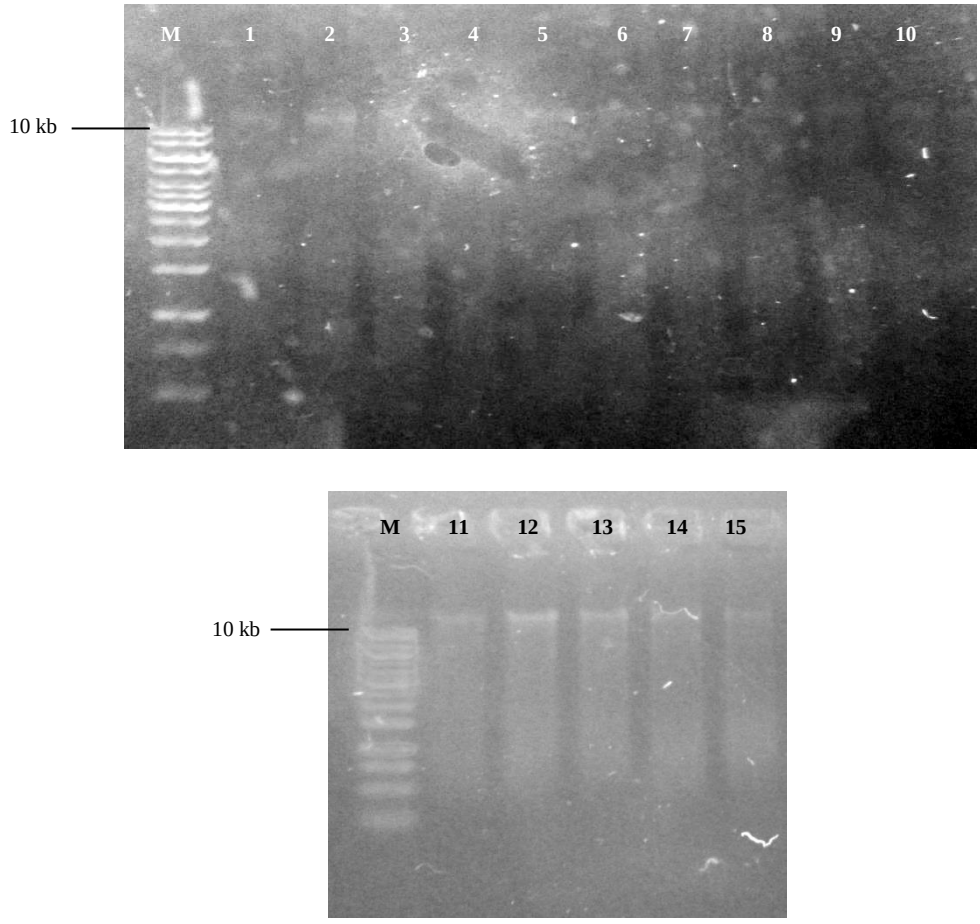


Figure 23: Gel electrophoresis of genomic DNA visualized on 1 % agarose gel stained with EtBr. The gDNA extracted from putative transformed seedling; $T_0$1-$T_0$15. Lane M: GeneRuler™ 1 kb DNA Ladder (Fermentas, Lithuania), Lane 1- 15: The genomic DNA extracted from $T_0$1-$T_0$15.

Table 29: DNA quantitative measurement of putative transformed seedling, T_0

Samples	Absorbance reading (A260/280)	Quantity of DNA concentration (µg/µL)
$T_0$1	1.50	0.30
$T_0$2	2.00	0.20
$T_0$3	1.67	0.25
$T_0$4	2.00	0.30
$T_0$5	1.67	0.25
$T_0$6	2.00	0.20
$T_0$7	1.75	0.35
$T_0$8	2.50	0.25
$T_0$9	1.75	0.35
$T_0$10	2.33	0.35
$T_0$11	2.00	0.30
$T_0$12	2.25	0.45
$T_0$13	2.00	0.3
$T_0$14	1.67	0.25
$T_0$15	1.67	0.25

4.2.3.2 Screening for the presence of r-*msAdh1* cDNA in putative transformed seedlings, T_0

The putative transformed seedlings were subjected to molecular analysis via PCR. The genomic DNA obtained from section 4.2.3.1 was used as a template to screen and identify the positive plants. Genomic DNA from tomato that was not transformed with the pGSA1131/r-msAdh1 (wild type) was used as a negative control meanwhile the plasmid pGSA1131/r-msAdh1 was used as positive control.

The first PCR attempt was using r-*msAdh1* internal specific forward and reverse primers; Inter_Adh_F (5'-TGCAAAGAGTGTGCTCACTGT-3') and Inter_Adh_R; (5'-CTCCACCATTTGTCATCTCAGC-3'). However, a band with expected size was observed in all tested plants including the wild type. When the band was sequenced and analysed via BLAST tool search against the nucleotide database in Genebank NCBI, the result obtained showed that the band is *Adh2* of wild type tomato (Result not shown). This indicates that the internal primer used was not specific to detect the r-*msAdh1* cDNA. The analysis via BLAST also showed that r-*msAdh1* has a high degree of similarity (78 %) with *Adh2* of wild type tomato (Appendix I).

Therefore, a new set of diagnostic primers were designed; the 5_Comseq_Adh and 3_Comseq_Adh, and combination of these primers was able to amplify full length of r-*msAdh1* cDNA from transformed tomato seedling. Out of 15 putative transformed plants analysed, five putative transformed plant lines ($T_0$1, $T_0$2, $T_0$4, $T_0$12, $T_0$13) were determined to contain r-*msAdh1* cDNA with band size of approximately 1.140 kb. Meanwhile, none of the wild type seedlings showed this band (Figure 24a and 24b). Further verification by nucleotide sequencing also confirmed and showed that the band was r-*msAdh1* cDNA (Appendix Ji).

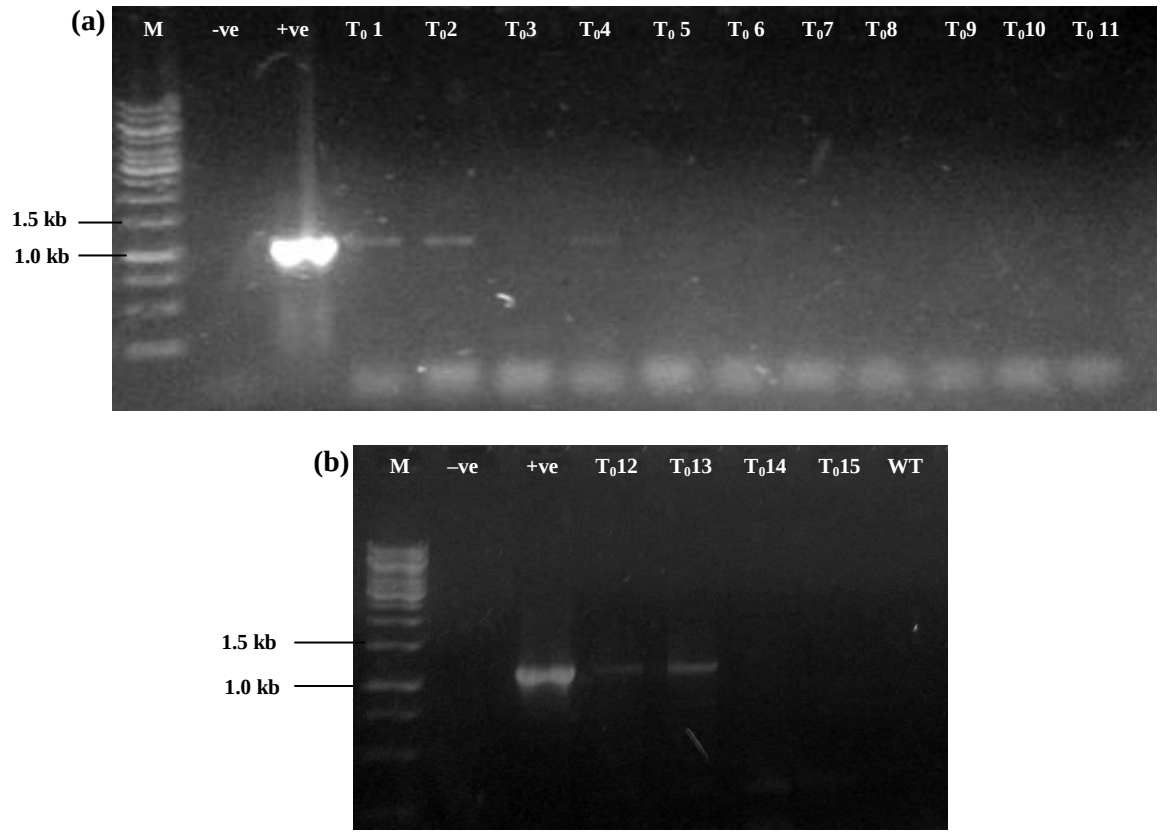


Figure 24(a)(b): Gel electrophoresis of PCR product visualized on 1.5 % agarose gel stained with EtBr to screen the integration of r-*msAdh1* cDNA in the putative transformed seedling, T_0. Lane M: GeneRuler™ 1 kb DNA Ladder (Fermentas, Lithuania), Lane –ve: Negative control (PCR without DNA template), Lane +ve: Positive control, Lane $T_0$1- $T_0$15: Putative transformed seedling 1-15. Lane WT: Wild type (control).

Up to date, the efficiency of stable *in planta* transformation of tomato had been reported from range 12 % -23 % using floral dip method, 17 % - 21 % using in vitro fruit injection method, 35 % - 42 % using in vivo fruit injection method (Yasmeen *et al.,* 2008). Meanwhile, higher stable *in planta* transformation efficiency up to 54 % - 68 % had been reported by Hassan *et al*. (2008) through *Agrobacterium* infiltration of ripened fruits.

In this study, the efficiency of transformation is lower (33.3 %) compared with method that developed by Hasan *et al.* (2008) and Yasmeen *et al.* (2008). However, method developed

in this study is faster as the putative transformed seedling can be analyzed after two weeks of transformation.

Every experiment produced different transformation efficiency as different *A. tumefaciens* strain, binary vector, target tissue and plant species were used. Bakhsh *et al.* (2014) reported that the transformation efficiencies of five *Agrobacterium* strain in *Nicotiana tabacum* L. cultivar Samsun were significantly different except between LBA4404 and GV2260. The efficiency of *Agrobacterium*-mediated plant transformation depends on the successful interaction between plant cell and *Agrobacterium*, which involves numerous genes from both *Agrobacterium* and plants (Gelvin, 2000).

4.2.3.3 Screening for the presence of *bar* gene in putative transformed seedlings, T_0

No amplification of *bar* gene was seen in all tested plants when PCR was carried out using Green GoTaq (Promega) at calculated annealing temperature of 59 °C. An attempt to amplify the GC-rich *bar* gene sequence from gDNA by adding the organic molecules such as dimethyl sulfoxide (DMSO) and bovine serum albumin (BSA) into PCR reaction mixture also failed. The difficulty has been experienced in achieving *bar* amplification possibly due to the high GC content (68.3 %) and the presence of blocks of continuous GC sequence of this gene (Vicker *et al.,* 1996). High GC content of gene generate complication during primer designing because of the secondary structure, and the most prominent problem associated is hairpin loop; which directly interferes during annealing of primers on difficult DNA template that leads to no amplification (Kumar and Kaur, 2014).

Malabadi and Nataraja (2007) had reported the successful amplification of *bar* gene from genomic template by using Expand^{TM} high Fidelity *Taq* Polymerase as described by Vicker *et al.* (1996). Meanwhile, Mammedov *et al.* (2008) had suggested that for primers with

high GC content, higher annealing temperatures may be necessary. Therefore, gradient was performed using High Fidelity enzyme mix (Fermentas, Lithuania) to test the efficiency and to determine the suitable annealing temperature for *bar* gene amplification.

From Figure 25, it can be seen that the amplification was obtained at high annealing temperature: 62.4 °C (Lane 7), 64.0 °C (Lane 8), 65.6 °C (Lane 9), and 66.7 °C (Lane 10). As a single bright band seen at annealing temperature 64.0 °C (Lane 8) and 65.6 °C (Lane 9); therefore, the decision to use annealing temperature: 65.6 °C to screen the positive seedling carrying the *bar* gene was made.

Out of the 15 plants tested, seven putative transformed plant lines: $T_0$1, $T_0$3, $T_0$4, $T_0$5, $T_0$9, $T_0$12 and $T_0$13 are positive *bar* gene (Figure 26a and 26b); thus resulting higher transformation efficiency (46.7 %) than r-*msAdh1* (33.33 %). The result indicating that co-integration of both r-*msAdh1* cDNA and the herbicide resistant *bar* gene in putative transformed plant did not always occur; perhaps the smaller gene was easier to incorporate and expressed at high rate during selection (Malabadi and Nataraja, 2007). Similarly, Afolabi (2004) found that the number of copies of transgenes; *bar* versus *gus* genes that originally present in the same T-DNA were often different in individually transformed plant lines. They also found that nonintact T-DNAs were present in >70 % of transgenic rice lines, in most cases reflecting loss of the mid to right border portion of the T-DNA. These disparities could result from rearranged or truncated T-DNA (Sallaud *et al.*, 2003) and this has been demonstrated directly by fiber-FISH in potato (Wolters *et al.* 1998). Kohli *et al.* (2010) reported that, the rearrangement may be induced by specific recombinogenic sequences such as CaMV 35S promoter.

Even though PCR can show the presence or absence of a transgene and provide a dependable copy number estimate, it provides little in the way of transformation about the

structure of a transgenic locus unless the genomic flanking sequences are already known (Kohli *et al.*, 2010). The structure of a transgene locus can have major influence on the level and stability of transgene expression (Kohli *et al.*, 2010).

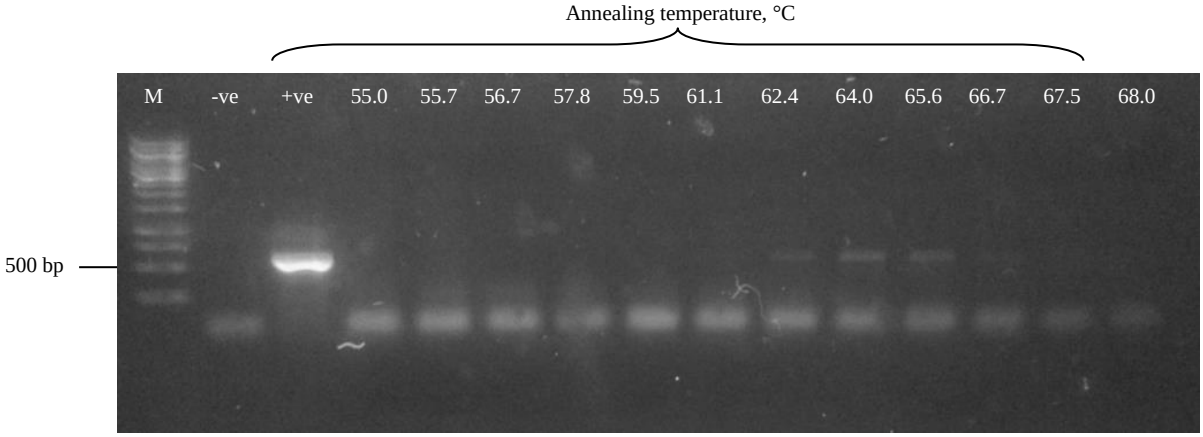


Figure 25: Gel electrophoresis of gradient PCR product visualized on 1.5 % agarose gel stained with EtBr. The PCR was performed to determine the annealing temperature for the amplification of *bar* gene from putative transformed seedling; $T_0$1. Lane M: GeneRuler™ 1 kb DNA Ladder (Fermentas, Lithuania), Lane –ve: Negative control (PCR without DNA template), Lane +ve: Positive control.

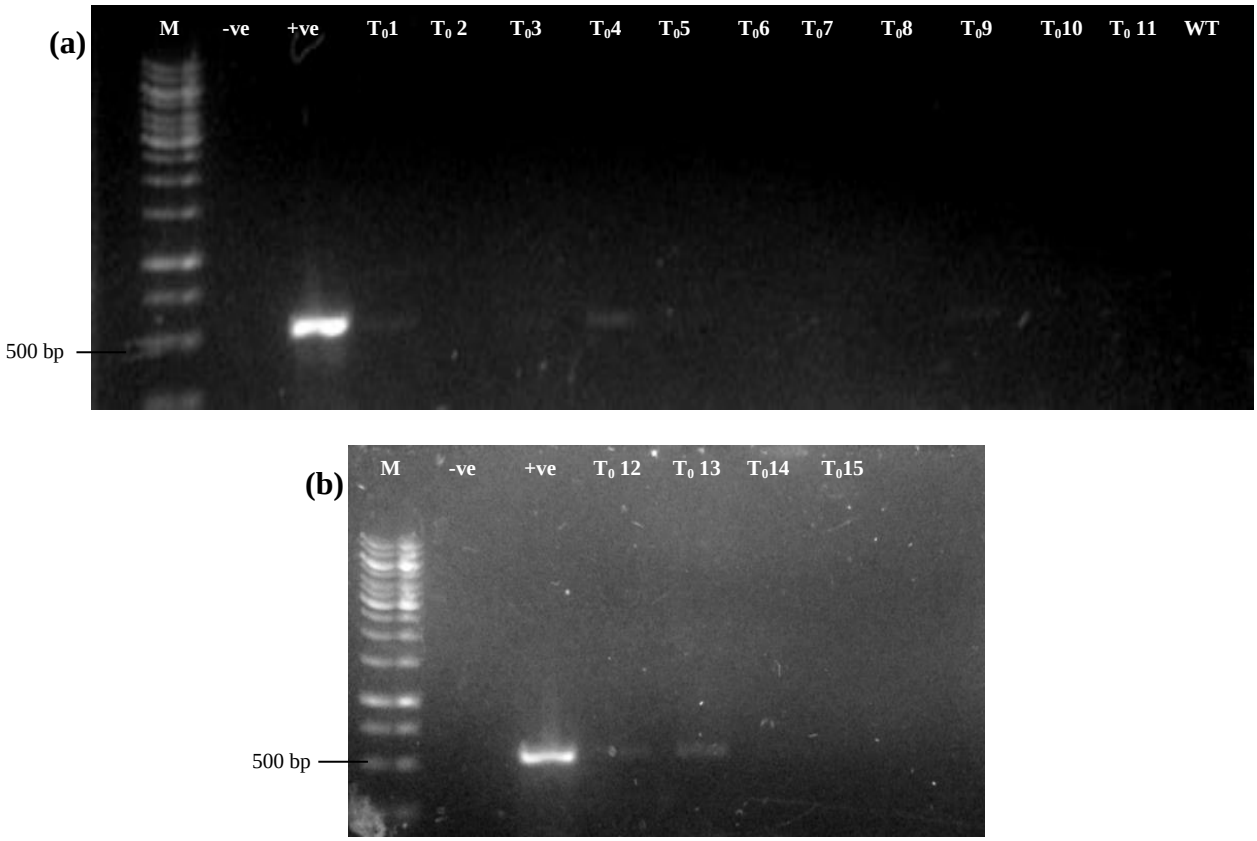


Figure 26(a)(b): Gel electrophoresis of PCR product visualized on 1.5 % agarose gel stained with EtBr to screen the integration of *bar* gene in the putative transformed seedling; T_0. Lane M: GeneRuler™ 1 kb DNA Ladder (Fermentas, Lithuania), Lane –ve: Negative control (PCR without DNA template), Lane +ve: Positive control (PCR using pGSA1131/r-msAdh1 plasmid as a template), Lane $T_0$1- $T_0$15: Putative transformed seedling, 1-15. Lane WT: Wild type (control).

4.2.3.4 Screening for the r-*msAdh1* cDNA in T_1 progeny

Fully mature red tomatoes were harvested from three putative transformed plant lines: $T_0$5, $T_0$12 and $T_0$13. Seeds were collected and germinated to obtain T_1 progeny. There were one hundred seedling obtained; however only 11 of them were randomly selected for further analysis.

The PCR analysis detected approximately 1.14 kb band in 9 out of 11 progeny seedlings (Figure 27-30). No amplification was obtained from $T_1$4 (Figure 28) and $T_1$11 (Result not shown). Even though the same concentration of gDNA was used in PCR, however different intensity of r-*msAdh1* band obtained among the transgenic lines probably indicates the different level of expression. This phenomenon has been generally described to different integration sites of the transgenes into the plant genome in each independent transformant (position effect) (Li *et al*, 2000).

Further analysis of $T_1$4 using gradient PCR had detected r-*msAdh1* band at high annealing temperature: 60.1 °C, 61.1 °C and 61.7 °C (Figure 29). Nucleotide sequencing had proven that it was full sequence of r-*msAdh1* cDNA. Meanwhile, molecular analysis of $T_1$11 through RT-PCR showed positive result therefore proved that $T_1$11 is carrying r-*msAdh1* in their genome (Refer Figure 32 and Appendix Jiii). From these results, it can be concluded that the r-*msAdh1* cDNA was inherited successfully in all T_1 progeny.

T_1 generation from $T_0$12

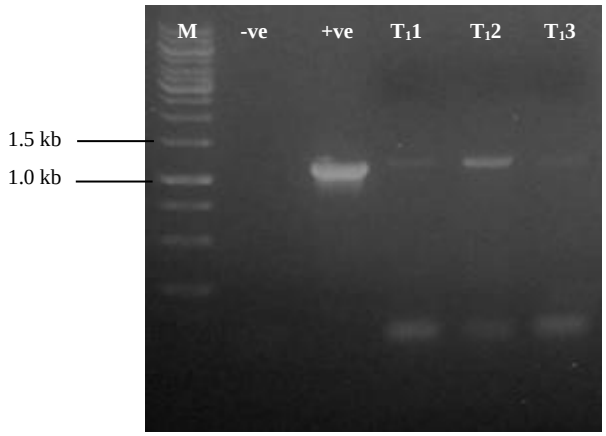


Figure 27: Gel electrophoresis of PCR product visualized on 1.5 % agarose gel stained with EtBr to analyse the integration of r-*msAdh1* in T_1 generation from $T_0$12. Lane M: GeneRuler™ 1 kb DNA Ladder (Fermentas, Lithuania), Lane –ve: Negative control (PCR without DNA template), Lane +ve: Positive control (PCR using pGSA1131/r-msAdh1 plasmid as template), Lane $T_1$1-$T_1$3: T_1 generation from $T_0$2.

T_1 generation from $T_0$13

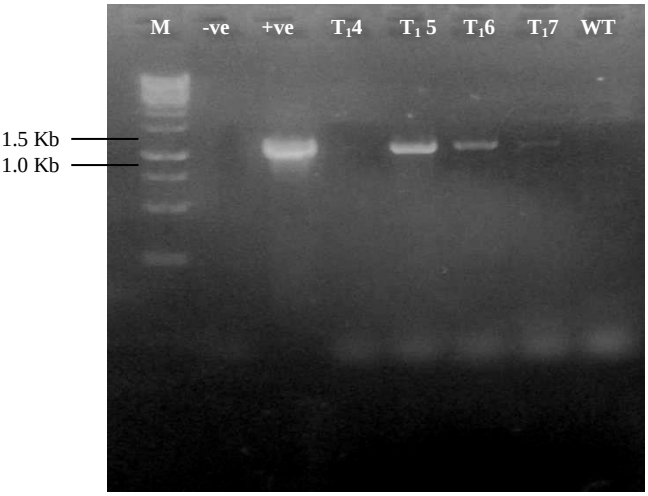


Figure 28: Gel electrophoresis of PCR product visualized on 1.5 % of agarose gel stained with EtBr to check the integration of r-*msAdh1* cDNA in T_1 progeny of $T_0$13. Lane M: GeneRuler™ 1 kb DNA Ladder (Fermentas, Lithuania), Lane –ve: Negative control (PCR without DNA template), Lane +ve: Positive control (PCR using pGSA1131/r-msAdh1 plasmid as template), Lane $T_1$4 - $T_1$7: T_1 generation from $T_0$13, Lane WT: Wild type (control).

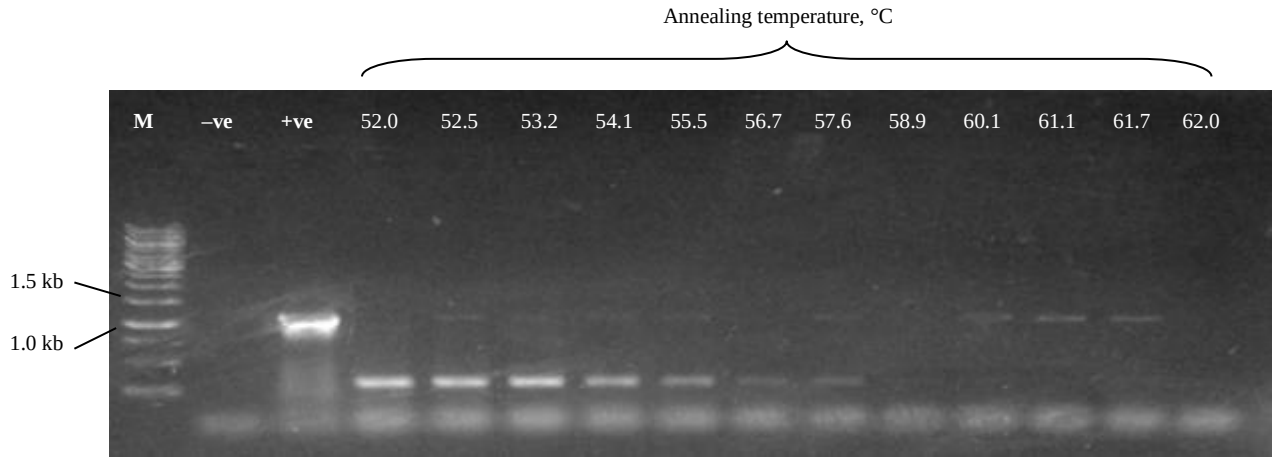


Figure 29: Gel electrophoresis of gradient PCR product visualized on 1.5 % agarose gel stained with EtBr to check the integration of r-*msAdh1* cDNA in progeny $T_1$4 of $T_0$13. M: GeneRuler™ 1 kb DNA Ladder (Fermentas, Lithuania), Lane –ve: Negative control (PCR without DNA template), Lane +ve: Positive control (PCR using pGSA1131/r-msAdh1 plasmid as template).

T_1 generation from $T_0$2

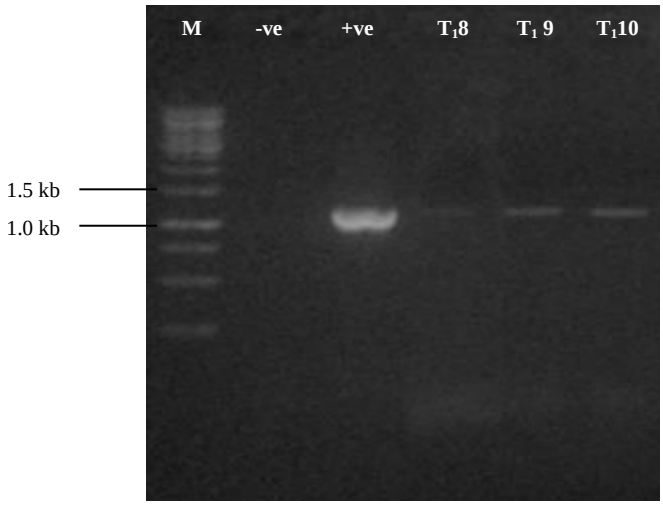


Figure 30: Gel electrophoresis of PCR product visualized on 1.5 % of agarose gel stained with EtBr to analyse the integration of r-*msAdh1* in T_1 progeny of $T_0$2. Lane M: GeneRuler™ 1 kb DNA Ladder (Fermentas, Lithuania), Lane –ve: Negative control (PCR without DNA template), Lane +ve: Positive control (PCR using pGSA1131/r-msAdh1 plasmid as template), Lane $T_1$8 -$T_1$10: T_1 generation from $T_0$2.

4.2.4 Expression analysis of r-*msAdh1*

4.2.4.1 Integrity and quality of total RNA

Purity and integrity of RNA are critical elements for the overall success of RNA-based analysis (Fleige and Pfaffl, 2006). Starting with low quality of RNA may strongly affected the result of downstream applications which are labour-intensive, time-consuming and highly expensive (Imbeaud *et al.,* 2005). Therefore, the integrity and quality of the used total RNA should be checked. In this study, total RNA was extracted from young leaves of four selected PCR positive lines; $T_1$8, $T_1$9, $T_1$10 and $T_1$11. The concentration and the purity of total RNA were assessed with UltrospecR 1100 *pro* (Amersham Pharmacia Biotech, USA), using diluted 1 uL aliquot of the total RNA solution. RNA purity was estimated from the 260/280 nm absorbance ratio, which is an estimation of contamination mainly from proteins and phenol. Meanwhile, in order to evaluate the degree of degradation, electrophoretic methods have been applied that separate the samples according to the size of the comprised molecules. RNA is considered of high quality when the ratio of 28S:18S bands is about 2.0 and higher (Schroeder *et al.,* 2006).

Based on the result that presented in Table 30, a 260/280 reading between 1.8-2.0 of samples $T_1$9 and $T_1$11 indicated good RNA quality. However, the bands were lacking sharp of 28S and 18S rRNA (Figure not shown), indicating that the RNA might be partially degraded by endogenous and environmental ribonucleases (RNAases) (Kingston, 2012). Degraded RNA may cause false negative results leading to delayed diagnosis or misdiagnosis (Schroeder *et al.,* 2006). Meanwhile, a low 260/280 reading obtained from samples $T_1$8 and $T_1$10 indicating that RNA is contaminated with protein or phenol. Regarding RNA integrity, RNA bands from these samples appeared smear (figure not shown) even though two distinct RNA fragments; the 28S and 18S rRNA was clearly observed on 1.5 % of agarose gel stained

with ethidium bromide. However, putative DNA contamination in total RNA solution was not directly assessed. Nevertheless, after DNAse treatment, no high molecular weight bands indicative of DNA presence are visible after gel electrophoresis of 5 uL of total RNA.

Table 30: RNA quantitative measurement of the selected T1 progeny seedlings

Samples	Absorbance reading A_{260}/A_{280}	Quantity of RNA (µg/µL)
$T_0$8	1.615	0.117
$T_0$9	1.860	0.032
$T_0$10	1.531	0.020
$T_0$11	1.981	0.057

4.2.4.2 Molecular analysis of r-*msAdh1* expression in transformed seedling through Reverse Transcription PCR

RT-PCR is a sensitive amplification procedure that has been used to detect the presence of a gene in a plant genome (Wang *et al.,* 2009). In this study, the expression of r-*msAdh1* cDNA in four PCR positive lines; $T_1$8, $T_1$9, $T_1$10 and $T_1$11 was detected by performing RT-PCR. First strand cDNA synthesis was carried out by reverse transcribing total RNA using ReverseAid^{TM} M-MuLV reverse transcriptase (Fermentas, Lithuania) primed with different antisense primers: 3_Comseq_Adh1, 3_BamH1_msAdh1 and oligo $_{dt}$17 (Refer to Figure 32). However, no amplification of r-*msAdh1* was obtained when the first cDNA generated by primers above is used as template in RT-PCR.

Surprisingly, the amplification of r-*msAdh1* was obtained when r-*msAdh1* specific antisense internal primer (5_msAdh1_R) was used to prime the cDNA synthesis reaction. The reason probably because of first strand cDNA generated either by 3_Comseq_Adh1 or

3_BamH1_msAdh1 priming was truncated; thus gave negative result in RT-PCR. In addition, secondary structures of mRNA can often cause difficulties for the reverse transcriptase, causing the enzyme to stall and end its synthesis well ahead of the 5' end (Brooks *et al.,* 1995 as cited in Resuehr and Spiess, 2003). Moreover, amplification of long product over 400 bp is strongly dependent on a good RNA quality (Fleige and Pfaffl, 2006). Therefore, a design of a primer to anneal an internal region of the gene of interest is useful if it is not possible to use intact RNA (Fleige and Pfaffl, 2006).

The RT-PCR using r-*msAdh1* specific primers (5_Comseq_Adh1 and 5_msAdh1_R) and subsequent agarose gel electrophoresis revealed a band of 500 bp in Lane 1 and Lane 2 (Figure 32), indicating that the r-*msAdh1* is presence in the transgenic progeny seedling $T_1$9 and $T_1$11. When the band was sequenced, result obtained confirmed and proven that the band was r-*msAdh1* cDNA (Appendix Jiii). Meanwhile, no amplification obtained from sample $T_1$8 and $T_1$10 were expected since the RNA quality is poor (Result not shown). According to Wang *et al.,* (2009), if RNA is a poor template for reverse transcription, it is very difficult or no chance to amplify a longer DNA fragment such as full length cDNA. Even intact RNA does not guarantee good results because RNA sample may contain inhibitors that can reduce reaction efficiency (Wong and Medrano, 2005). These factors include length of the amplicon, secondary structure and primer quality (Fleige and Pfaffl, 2006).

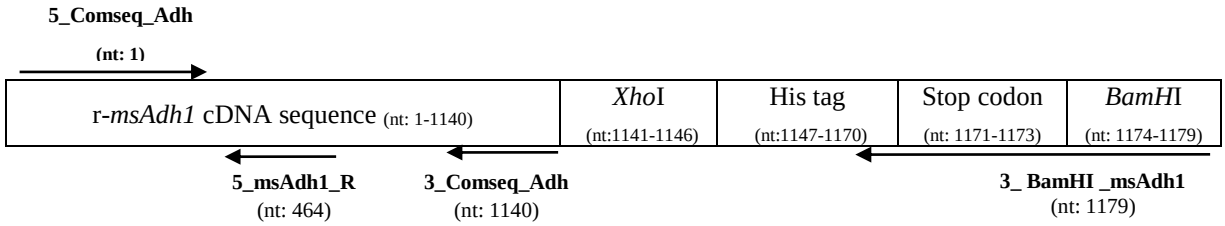


Figure 31: The orientation and position of primers used to prime the first strand cDNA synthesis reaction.

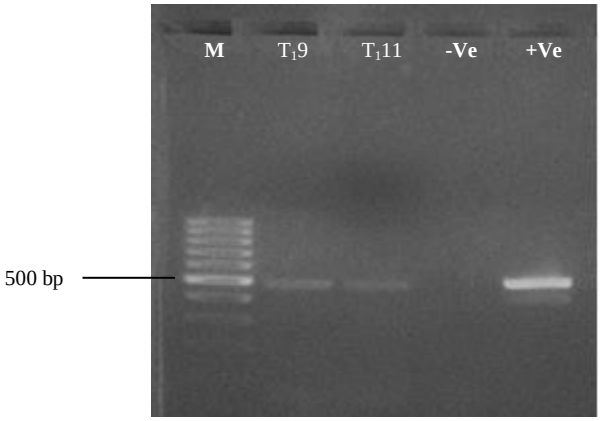


Figure 32: Gel electrophoresis of RT-PCR product visualized on 1.5 % agarose gel stained with EtBr to analyse the expression of r-*msAdh1* cDNA in transformed seedlings. Lane M: GeneRuler™ 1 kb DNA Ladder (Fermentas, Lithuania), Lane –ve: Negative control (PCR without DNA template), Lane +ve: Positive control (PCR using pGSA1131/r-msAdh1 plasmid as template), Lane 1: $T_1$9, Lane 2: $T_1$11.

4.2.5 Morphological observation on transgenic seedling

Several r-*msAdh1* expressing transgenic seedling (T_0 and T_1) had a change in their morphological cotyledon leaf when compared to wild type tomato seedling. The phenotype changed was not only observed in T_0 generation but also in T_1 generation. The unintended trait in genetic engineered plant is consequence of the change of the transformed plant's genome which is called as insertional effect (Konig *et al.* 2004). In case of this study, it can be seen that some T_0 transformed seedling has a mild wrinkled cotyledon leaf (Figure 33 and 34) when compared to wild type seedling that have smooth and plain cotyledon leaf (Figure 36). Besides, an abnormal shape of cotyledon leaf was always observed in T_0 and T_1 generation of transformed plants (Figure 35, 38 and 39) thus sometime renders easier transgenic seedling identification.

However, some r-*msAdh1* expressing transgenic seedling also exhibited normal and healthy cotyledon leaf (Figure not shown). The study conducted by Chern *et al.* (2007) on T-DNA insertional mutagenesis in rice (*Oryza sativa*) also found that not all transgenic line possess unintended traits as out of 22, 6665 field-grown T_1 lines examined, only 4, 065 lines was identified with visible mutant trait. Meanwhile, El Ouakfaoui and Miki (2005) demonstrated that genes can be inserted without altering the global gene expression profile. The unintended traits resulted when the insertion occurs within or near to an endogenous gene or regulatory element; thus either the expression of associated endogenous gene or the nature of the RNA and/or protein produced will be effected. However, the changes in their expression will not necessarily result in an unintended trait as plants are buffered against the consequences of genomic changes by the high level of gene redundancy in plant genome and by quality control systems active in plants (Schnell, 2015).

T_0 generation

Figure 33: Putative transformed seedling, T_0 has mild wrinkled of cotyledon leaf, as indicated by an arrow, when compared to wild type seedling.

Figure 34: Putative transformed seedling, T_0 has mild wrinkled of cotyledon leaf, as indicated by an arrow, when compared to wild type seedling. The growth of extra leaf also observed (showed by yellow arrow).

Figure 35: Putative transformed seedling, T_0 has abnormal cotyledon leaf shape, as indicated by arrow, when compared to wild type seedling.

Figure 36: Wild type seedling has smooth and plain cotyledon leaf.

T_1 generation

Figure 37: Wild type seedling has smooth and plain cotyledon leaf.

Figure 38: The progeny seedling**;** T_13 has abnormal shape of cotyledon leaf as compared to wild type seedling.

Figure 39: The progeny seedling; T_12 has an extra cotyledon leaf with abnormal shape as compared to wild type seedling.

4.3 Conclusion

First objective of this study was successfully achieved. The construct; pGSA1131/r-msAdh1 was successfully generated by cloning the complementary DNA (cDNA) of r-*msAdh1* into the *Nco*I and *Bam*HI restriction site of the T-DNA region of vector pGSA1131 under the control of the cauliflower mosaic virus 35S ribonucleic acid promoter (CaMV35S). Verification using restriction enzyme analysis and nucleotide sequencing proved the clone is in correct orientation in pGSA1131.

Second objective was successfully attempted. The transformation of the *msAdh1* cDNA into the tomato genome by *Agrobacterium*-mediated *in planta* transformation was carried out by infecting sonicated seeds with *A. tumefaciens* habouring pGSA1131/r-*msAdh1*. The developed protocol is easy, rapid, efficient, cost-effective, and can generate a relatively large number of transgenic plants in approximately one months. This is the first study that reports the *Agrobacterium tumefaciens*-mediated *in planta* strategy in tomato targeting on seed as gene transfer.

Third objective is to evaluate the integration of r-*msAdh1* cDNA and *bar* gene in tomato genome using PCR. Out of the 15 transformed plants, five (5) putative transgenic lines were determined to carry the r-*msAdh1* cDNA, and seven (7) putative transgenic lines were determined to carry *bar* gene in their genome resulting in an average of 33.3 % and 46.7 % transformation efficiency, respectively. Result obtained also showed that the r-*msAdh1* cDNA was inherited successfully in all T_1 progeny. Molecular analysis through RT-PCR also confirmed the presence of r-*msAdh1* cDNA in tomato genome.

4.4 Future Research

To validate the molecular transformation, the transgenic and control plants can be subjected to the swab test using various concentration of BASTA. Healthy leaves of plants should be selected, and swabs on the abaxial and adaxial leaves surface. Somayaji *et al.* (2014) had found that leaves of non-transformed tomato plants produced partial necrosis of leaves following the application of 2 mg/L BASTA and complete necrosis when 4 mg/L of BASTA was applied; indicating the plant is not transgenic. Meanwhile, leaves of transgenic plants remains unaffected to both concentration of BASTA. Further studies relating to level of gene expression is also required. Apart of that, observations of the physiological changes over generations are also important to be studied. The validation of gene expression in T_2 and T_3 generation of tomato transgenic for segregation patterns as well as stability and expression of r-*msAdh1* in tomato are also necessary.

CHAPTER FIVE

OVERALL CONCLUSION

The discovery of the full length of r-*msAdh1* cDNA from sago palm has prompted the work of r-*msAdh1* expression in both prokaryote and eukaryote system. The molecular and proteomic work of r-*msAdh1* expression study in prokaryote system had been extensively explained and discussed in chapter three. In that chapter, the reconstruction of r-*msAdh1* into expression vector, pET-41a(+) and transformation of the newly construct; pET-41a(+)/r-msAdh1 into expression host, *E. coli* strain BL21 (DE3) had been described. In an attempt to obtain soluble r-msAdh1 protein, on $O.D_{600}$ reaching 0.5-0.6, expression was induced with IPTG (final concentration 0.4 mM), and further cultured at different growth temperatures (15 °C and 27 °C) and periods (4 hours and 6 hours). SDS-PAGE analysis of the *E. coli* clear lysate revealed a moderate amount of soluble recombinant protein was successfully expressed at 27 °C and 15 °C (4 hours) and the molecular mass was about 43-48 kDa. When catalysing activity of r-msAdh1 was assayed spectrophotometrically, higher absorbance compared to control over time at 340 nm indicates that BL21 (DE3) was successfully expressed the r-msAdh1 enzyme that are able to reduce the NAD^+ to NADH. In spite of this successful expression, several difficulties were encountered. First, BL21 (DE3) is known to have endogenous ADH3 (3.9 kD) in their cellular with size almost the same with expressed r-msAdh1 protein (4.1 kD), therefore rendered difficult direct identification of the r-msAdh1 protein band on SDS-PAGE. Thus, after SDS-PAGE analysis, further verification need to be done such as western blotting to verify the presence of r-msAdh1 protein band. Second, since r-*msAdh1* cDNA is eukaryote gene, the insoluble and inactive r-msAdh1 proteins are co-produced in a BL21 (DE3). Though cytoplasmic folding is often enhance at low temperature, this is often accompanied by

misfolding and segregation into insoluble aggregates known as inclusion bodies. This is true when insoluble fraction of BL21 (DE3) with r-*msAdh1* expression was analysed on SDS-PAGE, very large amount of r-msAdh1 protein was presence in insoluble fraction; thereby indicating that most of the r-msAdh1 protein produced as inclusion bodies when overexpressed. Third, *in silico* analysis showed r-msAdh1 protein to be hydrophobic with Grand average of hydropathicity (GRAVY) of 0.006. The hydrophobicity of r-msAdh1 protein may cause aggregation after cell lysis. Based on these findings, it can be concluded that the expression of eukaryote gene in prokaryote system to obtain large amount of soluble r-msAdh1 protein is difficult. Insufficient amount of soluble r-msAdh1 protein may strongly compromise downstream applications such protein purification using Ni-NTA Spin Column. In this study, the purification of recombinant HIS-tagged r-msAdh1 was partially achieved as several bands were observed on SDS-PAGE. The significant challenge of using an organism such as *E. coli* for expression of r-*msAdh1* cDNA is that it is a prokaryote based system; therefore, lacks of post-translational machinery function and glycosylation. In addition, codon bias and mRNA stability are other factors that could contribute to the unsuccessful expression of eukaryote gene in prokaryote system.

Plant offers several advantages as host of expression such cheaper investment, high expression and glycosylation ability which is more promising bio-platform system for manufacture of recombinant protein compared to prokaryote system. Therefore, the expression of r-*msAdh1* was also studied in plant system. In this study, tomato was chosen as host of expression of r-*msAdh1* cDNA and *in planta* tomato seeds transformation using *A. tumefaciens* strain LBA4404 was described in chapter four. In order to develop an easy, rapid and efficient protocol for the tomato genetic transformation, without involving tissue culture, tomato seeds were sonicated and infiltrated with *A. tumefaciens* harbouring plasmid

pGSA1131 that contains r-*msAdh1* cDNA and *bar* gene as a selectable marker gene. Using this method, the transformation efficiency produced was determined to be 33.33 % for r-*msAdh1* and 46.67 % for *bar* gene. The results also showed that all the T_1 progeny carried the r-*msAdh1* cDNA. However, the integration analysis of r-*msAdh1* cDNA in tomato genome is quite difficult since the identity of *Adh2* of wild type tomato with r-*msAdh1* cDNA was high (78 %). Nevertheless, the full length of r-*msAdh1* from transgenic plants was successfully detected using 5_comseq_*Adh* and 3_comseq_*Adh* primers combination. Meanwhile, amplification of *bar* gene needed high annealing temperature ranging from 62.4 °C and 66.7 °C and by using High Fidelity enzyme mix. Nevertheless, the transformation protocol reported here is simple and reproducible and may be adapted to other tomato cultivars. For the future research, since the transformed plants are carrying bar gene in their genome, they can be subjected the swab test using 2-4 mg/L of BASTA. Alternatively, transformed plants also can be grown to maturity in the absence of any selection and then the progeny seeds can be collected and germinated on BASTA containing media to identify transformed plants. In addition, further studies relating to level of r-*msAdh1* expression in tomato genome as well as physiological changes of transformed plants over generation and segregation pattern are required. Besides, the stability expression of r-*msAdh1* cDNA in tomato is recommended to be studied in the future.

REFERENCES

Afolabi, A. S., Worland, B., Snape, J. W., & Vain, P. (2004). A large-scale study of rice plants transformed with different T-DNAs provides new insights into locus composition and T-DNA linkage configuration. *Theory and Applied Genetics, 109*(4), 815-826.

Ahmad, M., & Mirza, B. (2005). An efficient protocol for transient transformation of intact fruit and transgene expression in *Citrus*. *Plant Molecular Biology Reporter, 23*(4), *419-420.*

Assur, Z., Hendrickson, W. A., & Mancia, F. (2012). Tools for co-expressing multiple proteins in mammalian cells. *Methods Molecular Biology (Clifton, N. J), 801*, 173-187.

Bakhsh, A., Anayol, E., & Ozcan, S. F. (2014). Comparison of transformation efficiency of five *Agrobacterium tumefaciens* strains in *Nicotiana Tabacum* L. *Emirates Journal of Food and Agriculture, 26*(3), 259-264.

Bent, A. F. (2015). *Arabidopsis in planta* transformation. Uses, mechanisms, and prospects for transformation of other species. *Plant Physiology, 124*, 1541-1547.

Bergmeyer. H. U. (1986). *Methods of enzymatic analysis* (3rd ed.). Germany, Weinheim: Verlag Chemie.

Branden, C. I, Jornvall, H., Eklund, H., & Furugren, B. (1975). Alcohol dehydrogenases. In: P.D. Boyer (Ed.), *The enzymes XI* (pp. 103-190). New York, NY: Academic Press.

Brooks, E. M., & Sheflin, L. G., & Spaulding, S. W. (1995). Secondary structure in the 30-UTR of EGF and the choice of reverse transcriptase affect the detection of message diversity by RT-PCR. *BioTechniques, 19*(5), 806–814.

Bucher, M., Brander, K. A., Sbicego, S., Mandel, T., & Kuhlemeier, C. (1995). Aerobic fermentation in tobacco pollen. *Plant Molecular Biology, 28*(4), 739–750.

Chang, C., & Meyerowitz, E. M. (1986). Molecular cloning and DNA sequence of the Arabidopsis thaliana alcohol dehydrogenase gene. *Proceedings of the National Academy of Sciences, 83*(5), 1408–1412.

Chaudhry, Z., & Rashid, H. (2010). An improved *Agrobacterium* mediated transformation in tomato using hygromycin as a selective agent. *African Journal of Biotechnology, 9*(13), 1882–1891.

Chern, C. G., Fan, M. J., Yu, S. M., Hour, A. L., Lu, P. C., Lin, Y. C., Wei, F. J., Huang, S. C., Chen, S., Lai, M. H., Tseng, C. S., Yen, H. M., Jwo, W. S., Wu, C. C., Yang, T. L., Li, S. S., Kuo, Y, C., Li, S. M., Li, C. P., Wey, C. K., Trisiriroj, A., Lee, H. F., & Hsing,

Y. I. C. (2007). A rice phenomics study-phenotype scoring and seed propagation of a T-DNA insertion-induced rice mutant population.*Plant Molecular Biology, 65,* 427-438.

Christie, P. J., Hahn, M., & Walbot, V. (1991). Low-temperature accumulation of alcohol dehydrogenase-1 mRNA and protein activity in maize and rice seedlings. *Plant Physiology*, *95*(3), 699–706.

Curtis, I. S., & Nam, H. G. (2001). Transgenic radish (*Raphanus sativus* L. var. *longipinnatus* Bailey) by floral dip method: plant development and surfactant are important in opimizing transformation effeciency. *Transgenic Research, 10,* 363-371.

Dai, S., Zheng, P., Marmey, P., Zhang, S., Tian, W., Chen, S., Beachy, R. N., Fauquet, C. (2001). Comparative analysis of transgenic rice plants obtained by *Agrobacterium*-mediated transformation and particle bombardment. *Molecular Breeding, 7*(1): 25-33.

Daniell, H., Streatfield, S. J., & Wycoff, K. (2001). Medical molecular farming: production of antibodies, bipharmaceuticals and edible vaccines in plants. *Trends in Plant Science, 6*(5), 219-226.

Davuluri, G. R., Van Tuinen, A., Fraser, P. D., Manfredonia, A., Newman, R., & Burgess, D., Bowler, C. (2005). Fruit-specific RNAi-mediated suppression of DET1 enhances carotenoid and flavonoid content in tomatoes. *Nature Biotechnology, 23*(7), 890–895.

Dolferus, R., Jacobs, M., Peacock, W. J., & Dennis, E. S. (1994). Differential interactions of promoter elements in stress responses of the *Arabidopsis Adh* gene. *Plant Physiology*, *105*(4), 1075–1087.

Dolferus, R., Osterman, J. C., & Dennis, E. S. (1997). Cloning of the *Arabidopsis* and rice formaldehyde dehydrogenase genes: Implications for the origin of plant ADH enzymes. *Genetics*, *146*(3), 1131-1141.

Dyson, M. R., Shadbolt, S. P., Vincent, K. J., Perera, R. L., & McCafferty, J. (2004). Production of soluble mammalian proteins in *Escherichia coli*: identification of protein features that correlate with successful expression. *BMC Biotechnology, 4*(32).

El Ouakfaoui, S., & Miki, B. (2005). The stability of the *Arabidopsis* transcriptome in transgenic plants expressing the marker genes *nptII* and *uidA*. *The Plant Journal, 41*(6), 791-800**.**

Enriquez-Obregon, G. A., Vazquez-Padron, R. I., Prieto-Samsonov, D. L., De la Riva, G. A., & Selman-Housein, G. (1998). Herbicide-resistant sugarcane (*Saccharum officinarum* L.) plants by *Agrobacterium*-mediated transformation. *Planta, 206*(1), 20-27.

Fleige, S., & Pfaffl, M. W. (2006). RNA integrity and the effect on the real-time qRT-PCR performance. *Molecular Aspects of Medicine, 27*(2), 126-139.

Farzaneh, A., Adel, Y., Ali, N., & Younes, G. (2013). Determine effective concentrations of β-lactam antibiotics against three Strains of *Agrobacterium tumefaciens* and phytotoxicity on tomato and tobacco. *International Journal of Agronomy and Plant Production, 4*(11), 2919–2925.

Fillatti, J. J., Kiser, J., Rose, R., & Comai, L. (1987). Efficient transfer of a glyphosate tolerance gene into tomato using a binary *Agrobacterium tumefaciens* vector. *Nature Biotechnology, 5*(7), 726–730.

Fink, A. L. (1998). Protein aggregation: folding aggregates, inclusion bodies and amyloid. *Folding and Design, 3*(1), 9–23.

Garabagi, F., Duns, G., & Strommer, J. (2005). Selective recruitment of *Adh* genes for distinct enzymatic functions in *Petunia hybrida*. *Plant Molecular Biology, 58*(2), 283–294.

Gelvin, S. B. (2000). *Agrobacterium*-mediated plant transformation: the biology behind the "gene-jockeying" tool. *Microbiology and Molecular Biology Reviews, 67*(1), 16-37.

Graves, A. E., Goldman, S. L., Banks, S. W., & Graves, A. C. F. (1988). Scanning electron microscope studies of *Agrobacterium tumefaciens* attachment to *Zea mays, Gladiolus* sp. and *Triticum aestivum. Journal of Bacteriology, 170*, 2395-2400.

Grodberg, J., & Dunn, J. J. (1988). *ompT* encodes the *Escherichia coli* outer membrane protease that cleaves T7 RNA polymerase during purification. *Journal of Bacteriology, 170* (3), 1245–1253.

Gupta, A., Pal, R. K., & Rajam, M. V. (2013). Delayed ripening and improved fruit processing quality in tomato by RNA-i-mediated silencing of three homologs of 1-aminopropane-1-caboxylate synthase gene. *Journal of Plant Physiology, 170*(11), 987-995.

Hallens, R., & Mullineaux, P. (2000). A guide to *Agrobacterium* binary Ti vectors. *Trends in Plant Science, 5*(10), 446-451.

Hamilton, C. M., Frary, A., Lewis, C., & Tanksley, S. D. (1996). Stable transfer of intact high molecular weight DNA into plant chromosomes. *Proceedings of the National Academy of Sciences of the United States of America, 93*(18), 9975–9979.

Hamza, S., & Chupeau, Y. (1993). Re-evaluation of conditions for plant regeneration and *Agrobacterium*-mediated transformation from tomato (*Lycopersicon esculentum*). *Journal of Experimental Botany, 44*(12), 1837-1845.

Hasan, M., Khan, A. J., Khan, S., Shah, A. H., Khan, A. R., & Mirza, B. (2008). Transformation of tomato (*Lycopersicon esculentum* Mill.) with *Arabidopsis* early

flowering gene *apetalai* (*API*) through *Agrobacterium* infiltration of ripened fruits. *Pakistan Journal of Biotechnology, 40*(1), 161-173.

Hooykaas, P. J. J. (1988). *Agrobacterium* molecular genetics. In S. B. Gelvin & R. A. Schilperoort, (Eds.), *Plant Molecular Biology Manual* (pp: 83-87). Dordrecht, the Netherlands: *Kluwer Academic Publishers.*

Hu, T., Metz, S., Chay, C., Zhou, H. P., Biest, N., Chen, G., Cheng, M., Feng, X., Radionenko, M., Lu, F., & Fry, J. (2003). *Agrobacterium*-mediated large-scale transformation of wheat (*Triticum aestivum* L.) using glyphosate selection. *Plant Cell Reports*, *21*(10), 1010-1019.

Huang, X. Q., & Wei, Z. M. (2005). Successful *Agrobacterium*-mediated genetic transformation of maize elite inbred lines. *Plant Cell, Tissue and Organ Culture, 83*(2), 187-200.

Ibrahim, E. R., Hossain, A., & Roslan, H. A. (2014). Genetic transformation of *Metroxylon sagu* (*Rottb*.) cultures via *Agrobacterium*-mediated and particle bombardment. *BioMed Research International, 2014, 348140.*

Imbeaud, S., Graudens, E., Boulanger, V., Barlet, X., Zaborski, P., Eveno, E., Mueller, O., Schroeder, A., & Auffray, S. (2005). Toward standardization of RNA quality assessment using user-independent classifiers of microcapillary electrophoresis traces. *Nucleic Acids Research, 33*(6).

Janssen, B. J., Lund, L., & Sinha, N. (1998). Overexpression of a homeobox gene, *LeT6*, reveals indeterminate features in the tomato compound leaf. *Plant Physiology*, *117*(3), 771–786.

Jeong, H., Kim, J., & Lee, J. (2015). Complete genome sequence of *Escherichia coli* strain BL21. *Genome Announcements Journal*, *3*(2), 134-135.

Kato-Noguchi, H. (2000). Osmotic stress increases alcohol dehydrogenase activity in maize seedlings. *Biologia Plantarum, 43*(4), 621-624.

Kato-Noguchi, H. (2001). Wounding stress induces alcohol dehyrogenase in maize and lettuce seedlings. *Plant Growh Regulation, 35*(3), 285-288.

Kim, K. N., Cheong, Y. H., Grant, J. J., Pandey, G. K., & Luan, S. (2003). *CIPK3* , a calcium sensor–associated protein kinase that regulates abscisic acid and cold signal transduction in *Arabidopsis*. *The Plant Cells*, *15*(2), 411-423.

Kingston, R. E. (2012). Preparation and analysis of RNA. In Ausubel, F. M., Brent, R., Kingston, R. E., Moore, D. D., Seidman, J. G., Smith, J. A., Struhl, K. (Eds.), *Current protocol in molecular biology* (pp. 4.0.1-4.11.13). New York, NY: John Wiley & Sons.

Keshamma, E., Rohini, S., Rao, K. S., Madhusudhan, B., & Udaya Kumar, M. (2008). Tissue culture-independent *in planta* transformation strategy: an *Agrobacterium tumefaciens*-mediated gene transfer method to overcome recalcitrance in cotton (*Gossypium hirsutum* L.). *Journal Cotton Science*, *12*(3), 264-272.

Kransesky, J., & Jonak, C. (2012). Drought, salt, and temperature stress-induced metabolic rearrangements and regulatory networks. *Journal of Experimental Botany, 63*(4), 1593-1608.

Khow, O., & Suntrarachun, S. (2012). Strategies for production of active eukaryotic proteins in bacterial expression system. *Asian Pacific Journal of Tropical Biomedicine*, *2*(2), 159–162.

Kohli, A., Miro, B., & Twyman, R. M. (2010). Transgene integration, expression and stability in plants: strategies for improvements. In Kole, C., Michler, C. H., Abbot, A. G., & Hall, T. C. (Eds.), *Transgenic crop plants* (pp. 201-237). Berlin and Heidelberg: Springer Berlin Heidelberg.

Konig, A., Cockburn, A., Crevel, R. W., Debruyne, E., Grafstroem, R., Hammerling, U., Kimber, I., Knudsen, I., Kuiper, H. A., Peijnenburg, A. A., Penninks, A. H., Poulsen, M., Schauzu, M., & Wal, J. M. (2004). Assessment of the safety of food derived from genetically modified (GM) crops. *Food and Chemical Toxicology, 42*(7), 1047-1088.

Kumar, A., & Kaur, J. (2014). Primer based approach for PCR amplification of high GC content gene : *Mycobacterium* gene as a model. *Molecular Biology International, 2014, 93730.*

Larkin, P. J. & Scowcroft, W. R. (1981). Somaclonal variation - a novel source of variability from cell cultures for plant improvement, *Theoritical and Applied Genetics, 60*, 197-214.

Lakshmi, G. J., Kotra, S. R., Peravali, J. B., Kumar, P. P. B. S., & Rao, K. R. S. M. (2014). Molecular cloning, high level expression and activity analysis of constructed human interleukin–25 using industrially important IPTG inducible *Escherichia coli* BL21(DE3). *International Journal of Bio-Science and Bio-Technology*, *6*(3), 19-30.

Larsen, M. W. (2009). Expression of a lipase in prokaryote and eukaryote host systems allowing engineering (Doctoral dissertation). Retrieved on Febuary 6, 2015 from http://www.diva-portal.org/smash/get/diva2:236156/FULLTEXT01.pdf.

Lee, J. J., Ahmad, S., & Roslan, H. A. (2013). Transformation of *Morinda citrifolia* via simple mature seed imbibition method. *Pakistan Journal of Biological Sciences*, *16*(24), 1913-1921.

Leibly, D. J., Nguyen, T. N., Kao, L. T., Hewitt, S. N., Barrett, L. K., & Van Voorhis, W. C. (2012). Stabilizing additives added during cell lysis aid in the solubilization of recombinant proteins. *PLoS ONE*, *7*(12), e52482.

Li, Y., Fengqi, C., Liqun, D., Meihai, G., Hongjie, L., Jingsong, Z., Shouyi, C., & Zhiqing, Z. (2000). Genetic transformation of watercress with a gene encoding for betaine-aldehyde dehydrogenase (BADH). *Acta Botanica Sinica*, *42*(5), 480-484.

Lin, W. C., Lu, C. F., Wu, J. W., Cheng, M. L., Lin, Y. M., Yang, N. S., Black, L., Green, S. K., Wang, J. F., & Cheng, C. P. (2004). Transgenic tomato plants expressing the *Arabidopsis NPR1* gene display enhanced resistance to a spectrum of fungal and bacterial diseases. *Transgenic Reearch, 13*(6), 567–581.

Machielsen, R., Uria, A. R., Kengen, S. W. M., & Oost, J. Van Der. (2006). Production and characterization of a thermostable alcohol dehydrogenase that belongs to the aldo-keto reductase superfamily. *Applied and Environmental Microbiology, 72*(1), 233–238.

Madhou, P., Raghavan, C., Wells, A., & Stevenson, T. W. (2006). Genome wide microaaray analysis of the effect of a surfactant application in *Arabidopsis. Weed Research, 46*(4), 275-283.

MacNicol, P. K., & Jacobsen, J. V. (2001). Regulation of alcohol dehydrogenase gene expression in barley aleurone by gibberellin and abscisic acid. *Physiology Plant*, *111*(4), 533–539.

Malabadi, R. B., & Nataraja, K. (2007). A biolistic approach for the production of transgenic plants using embryogenic tissue in *Pinus kesiya* Royle Ex. Gord (Khasi pine). *Biotechnology*, *6*(1):86-92.

Mammedov, T. G., Pienaar, E., Whitney, S. E., TerMaat, J. R., Garvill, G., Goliath, R., Subramanian, A., & Viljoen, H. J. (2008). A fundamental study of the PCR amplification of GC-rich DNA templates. *Computional Biology and Chemistry, 32*(6), 452-457.

Martinez, A., Knappskog, P. M., Olafsdottir, S., Doskeland, A. P., Eiken, H. G., Svebak, R. M., Bozzini, M., Apold, J., & Flatmark, T. (1995). Expression of recombinant human phenylalanine hydroxylase as fusion protein in *Escherichia coli* circumvents proteolytic degradation by host cell proteases. *Biochemistry Journal*, *306*(2), 589-597.

Matsukura, C., Aoki, K., Fukuda, N., Mizoguchi, T., Asamizu, E., Saito, T., Shibata, D., & Ezura, H. (2008). Comprehensive resources for tomato functional genomics based on the miniature model tomato Micro-Tom. *Current Genomic*, *9*(7), 436–443.

Mayavan, S., Subramanyam, K., Arun, M., Rajesh, M., Dev, G. K., Sivanandhan, G., Jaganath, B., Manickavasagam, M., Selvaraj, N., & Ganapathi, A. (2013). *Agrobacterium tumefaciens*-mediated *in planta* seed transformation strategy in sugarcane. *Plant Cell Reports, 32*(10), 1557-1574.

Matzke, M. A., & Matzke, A. J. (1991). Differential inactivation and methylation of a transgene in plants by two suppressor loci containing homologous sequences. *Plant Molecular Biology, 16*(5), 821-830.

McCormic, S., Niedermeyer, J., Fry, J., Barnason, A., Horsch, R., & Fraley, R. (1986). Leaf disc transformation of cultivated tomato (L. *esculentum*) using *Agrobacterium tumefaciens*. *Plant Cells Report, 5*(2), 81-84.

Menda, N., Semel, Y., Peled, D., Eshed, Y., & Zamir, D. (2004). *In silico* screening of a saturated mutation library of tomato. *The Plant Journal, 38*(5), 861-872.

Miew, H. C. (2011). Reconstruction of *msAdh1* gene (Degree's thesis). Retrieved on February 20, 2012 from http://ir.unimas.my/314/1/Ho_Carl_Miew.pdf.

Novagen (11th ed.). (2005). pET system manual. Retrieved on August 5, 2012 from http://kirschner.med.harvard.edu/files/protocols/Novagen_petsystem.pdf.

Park, S. H., Morris, J. L., Park, J. E., Hirschi, K. D., & Smith, R. H. (2003). Efficient and genotype-independent *Agrobacterium*–mediated tomato transformation. *Journal of Plant Physiology, 160*(10), 1253–1257.

Park, B. J., Liu, Z., Konno, A., & Kameya, T. (2005). Transformation of radish (*Raphanus sativus* L.) via sonication and vacuum infiltration of germinated seeds with *Agrobacterium* habouring a group 3 LEA gene from *B. napus. Plant Cell Reports, 24*(8), 494-500.

Pratheesh, P. T., Vineetha, M., & Muraleedhara, K. G. (2014). Expression of recombinant human serum albumin in *Nicotiana Tabacum. World Journal of Pharmaceutical Research, 3*(2), 2600-2618.

Rai, M., & Padh, H. (2001). Expression systems for production of heterologous proteins. *Current Science, 80*(9), 1121-1128.

Resuehr, D., & Spiess, A. N. (2003). A real-time polymerase chain reaction-based evaluation of cDNA synthesis priming methods. *Analytical Biochemistry, 322*(2), 287–291.

Rohini, V. K., & Rao, K. S. (2000). Transformation of peanut (*Arachis hypogaea* L.) : a non-tissue culture based approach for generating transgenic plants. *Plant Science, 150*(1), 41-49.

Roslan, H. A., Ching, W. C., Ilyas, H., & Sundaraj, Y. (2010). Alcohol dehydrogenase gene in sago palm. *National Biotechnology Seminar,* Retrieved on September 31, 2010 from http://www.biotek.gov.my/nbs2010 .

Roslan. H. A. Sundaraj, Y., & Husaini, A. A. S. A. (2008). Multiple forms of alcohol dehydrogenase (*Adh*) genes in sago palm: A preliminary study. *Banwa, 5*(1), 29-40.

Sahdev, S., Khattar, S. K., & Saini, K. S. (2008). Production of active eukaryotic proteins through bacterial expression systems : a review of the existing biotechnology strategies. *Molecular Cell Biochemistry, 307*(1-2), 249–264.

Saïda, F., Uzan, M., Odaert, B., & Bontems, F. (2006). Expression of highly toxic genes in *E . coli* : special strategies and genetic tools. *Current Protein and Peptide Science, 7*(1), 47–56.

Sallaud, C., Meynard, D., Van Boxtel, J., Gay, C., Bes, M., Brizard, J. P., Larmande, P., Ortega, D., Raynal, M., Portefaix, M., Ouwerkerk, P. B., Rueb, S., Delseny, M., & Guiderdoni, E. (2003). Highly efficient production and characterizationof T-DNA plants for rice (*Oriza sativa* L.) functional genomics. *Theory and Applied Genetics, 106*(8), 1396-1408.

Schnell, J., Steele, M., Bean, J., Neuspiel, M., Girard, C., Dormann, N., Pearson, C., Savoie, A., Bourbonniere, L., & Macdonald, P. (2015). A comparative analysis of insertional effects in genetically engineered plants: considerations for pre-market assessments. *Transgenic Research, 24*(1), *1-17.*

Schroeder, A., Mueller, O., Stocker, S., Salowsky, R., Leiber, M., Gassmann, M., Lightfoot, S., Menzel, W., Granzow, M., & Ragg, T. (2006). The RIN: an RNA integrity number for assigning integrity values to RNA measurements. *BMC Molecular Biology, 7*(3).

Sharma, M. K., Solanke, A. U., jani, D., Singh, Y., & Sharma, A. K. (2009). A simple and effi cient *Agrobacterium*-mediated procedure for transformation of tomato. *Journal Bioscience, 34*(3), 423-433.

Sivashanmugam, A., Murray, V., Cui, C., Zhang, Y., Wang, J., & Li, Q. (2009). Practical protocols for production of very high yields of recombinant proteins using *Escherichia coli*. *Protein Science, 18*(1), 936–948.

Smith, R. H., & Hood, E. E. (1995), *Agrobacterium tumefaciens* transformation of monocotyledons. *Crop Science Society of America, 35*(2), 301-309.

SoltanMohammadi, B., Jalali-Javaran, M., Rajabi-Memari, H., & Mohebodini, M. (2014). Cloning, transformation and expression of proinsulin gene in tomato (*Lycopersicum esculentum* Mill.). *Jundishapur Journal of Natural Pharmaceutical Products, 9*(1), 9-15.

Somayaji, P., Doddananjappa, T. P., & Byrappa, S. M. (2014). *Agrobacterium*-mediated transformation for development of transgenic tomato with *ySAMdc* gene. *Indian Journal of Biotechnology*, *13*(1), 19–25.

Song, L., Zhao, D. G., Wu, Y. J., & Tian, X. (2009). A simplified seed transformation method for obtaining transgenic *Brassica napus* plants. *Agricultural Science in China, 8*(6), 658-663.

Sood, P., Bhattacharya, A., & Sood, A. (2011). Problems and possibilities of monocot transformation. *Biologia Plantarum, 55*(1), 1-15.

Sørensen, H. P., & Mortensen, K. K. (2005). Soluble expression of recombinant proteins in the cytoplasm of *Escherichia coli*. *Microbial Cell Factories, 4*(1).

Speirs, J., Lee, E., Holt, K., Yong-duk, K., Scott, N. S., Loveys, B., & Schuch, W. (1998). Genetic manipulation of alcohol dehydrogenase levels in ripening tomato fruit affects the balance of some flavor aldehydes and alcohols. *Plant Physiol*ogy, *117*(3), 1047–1058.

Spolaore, S., Trainotti, L., & Casadora, G. (2001). A simple protocol for transient gene expression in ripe flesy fruit mediated by *Agrobacterium. Journal of Experiemental Botany, 52*(357), 845-850.

Strommer, J. (2011). The plant ADH gene family. *The Plant Journal, 66*(1), 128–142.

Studier, F. W., & Moffattf, B. A. (1986). Use of bacteriophage T7 RNA polymerase to direct selective high-level expression of cloned genes. *Journal Molecular Biology*, *189*(1), 113–130.

Tanksley, S. D., Ganal, M. W., Prince, J. P., De Vicente, M. C., Bonierbale, M. W., Broun, P., Fulton, T. M., Giovannoni, J. J., Grandillo, S., Martin, G. B., Mesenguer, R. Miller, J. C., Miller, L., Paterson, A. H., Pineda, O., Roder, M. S., Wing, R. A., Wu. W., & Young, N. D. (1992). High density molecualr linkage maps of the tomato and potato genomes. *Genetic Society of America, 132*(4), 1141-1160.

Tanksley, S. D. (1993). Linkage map of tomato (*Lycopersicon esculentum*) (2N=24). In J. O'Brien (Ed.), *Genetic maps: Locus maps of complex genomes* (pp. 6.39-6.60). New York, NY: Cold Spring Harbor Laboratory Press.

Tesniere, C., & Verries, C. (2000). Molecular cloning and expression of of cDNAs encoding alcohol dehydrogenases from *Vitis vinifera L.* during berry development. *Plant Science, 157*(1), 77-88.

Tianzi, C., Shenjie, W., Jun, Z., WangZhen, G., & TianZhen, Z. (2010). Pistil drip following pollination: a simple *in planta Agrobacterium*-mediated transformation in cotton. *Biotechnology Letters. 32*(4), 547-555.

Trick, H. N. & Finer, J. J. (1997). SAAT: sonication-assisted *Agrobacterium*-mediated transformation. *Transgenic Research*, *6*(5), 329-336.

Trimpin, S., & Brizzard, B. (2009). Analysis of insoluble proteins. *Biotechnique*s, *46*(5),321-326.

Van Eldik, G. J., Ruiter, R. K., Van Herpen, M. M. A., Schrauwen, J. A. M., & Wullems, G. J. (2002). An alcohol dehydrogenase-like gene is specifically expressed in potato pistils. *Journal of Experimental Botany*, *49*(325), 1453-1454.

Vicker, J. E., Glenn, C., Graham, & Robert, J. H. (1996). A protocol for the efficient screening of putatively transformed plants for *bar*, the selectable marker gene, using the polymerase chain reaction. *Plant Molecular Biology Reporter*, *14*(4), 362-368.

Van der Hoeven, R., Ronning, C., Giovannoni, J., Martin, G., & Tanksley, S. (2002). Deductions about the number, organization, and evolution of genes in the tomato genome based on analysis of a large expressed sequence tag collection and selective genomic sequencing. *The Plant Cell*, *14*(7), 1441–1456.

Van Eldik, G. J., Ruiter, R. K., Van Herpen, M. M. A., Schrauwen, J. A. M., & Wullems, G. J. (1998). An alcohol dehydrogenase-like gene is specifically expressed in potato pistils. *Journal of Experimenal Botany*, *49*(325), 1453–1454.

Waldo, G. S., Standish, B. M., Berendzen, J., & Terwilliger, T. C. (1999). Rapid protein-folding assay using green fluorescent protein. *Nature Biotechnology, 17*(7), 691–695.

Wang, H. M., Yin, W. C., Wang, C. K., & To, K. Y. (2009). Isolation of functional RNA from different tissues of tomato suitable for developmental profiling by microarray analysis. *Botanical Studies, 50*(2), 115-125.

Wang, N., Shi, H., Yao, Q., Zhou, Y., Kang, L., Chen, H., & Chen, K. (2011). Cloning, expression and characterization of alcohol dehydrogenases in the silkworm *Bombyx mori*. *Genetic and Molecular Biology*, *34*(2), 240-243.

Wee, C. C., & Roslan, H. A. (2012). Isolation of alcohol dehydrogenase cDNA and basal regulatory region from *Metroxylon sagu*. *ISRN Molecular Biology, 2012*,1-10.

Wee. C. C. (2011). Generation of expressed sequence tags (ESTs) database and isolation of alcohol dehydrogenase gene from young leaf samples of *Metroxylon sagu* (Master 's thesis). Universiti Malaysia Sarawak.

Williams, D. C., Van Frank, R. M., Muth, W. L., & Burnett, J. P. (1982). Cytoplasmic inclusion bodies in *Escherichia coli* producing biosynthetic human insulin proteins. *Science, 215*(4533), *687-688.*

Winans, S. C. (1990). Transcriptional induction of an *Agrobacterium* regulatory gene tandem promoters by plant-released phenolic compounds, phosphate starvation, and acidic growth media. *Journal of Bacteriology, 172*(5), 2433–2438.

Wing, R.A., Zhang, H. B., & Tanksley, S. D. (1994). Map-based cloning in crop plants: tomato as a model system: i. genetic and physical mapping of jointless. *Molecular Genetics and Genomics, 242*(6), 681-688.

Wolters, A. M. A., Trindade, L. M., Jacobsen, E., & Visser, R. G. F. (1998). Fluorescence *in situ* hybridization on extended DNA fibers as a tool analyse complex T-DNA loci in potato. *The Plant Journal, 13*(6), 837-847.

Wong, M. L., & Medrano, J. F. (2005). Real-time PCR for mRNA quantification. *Biotechniques, 39*(1), 75-85.

Xu, Y., Jia, J. F., & Zheng, G. C. (1989). Phenolic compounds can promote efficient transformation of plants by *Agrobacterium. China Science Bulletin, 34*, 1902-1989

Yamamoto, N., Tsugane, T., Watanabe, M., Yano, K., & Maeda, F. (2005). Expressed sequence tags from the laboratory-grown miniature tomato (*Lycopersicon esculentum*) cultivar Micro-Tom and mining for single nucleotide polymorphisms and insertions / deletions in tomato cultivars. *Gene Section Functional Genomics, 15*(536), 127–134.

Yamaguchi, H., & Miyazaki, M. (2014). Refolding techniques for recovering biologically active recombinant proteins from inclusion bodies. *Biomolecules, 4*(1), 235-251.

Yang, Z., Zhang, L., Zhang, Y., Zhang, T., Feng, Y., Lu, X., & Wang, X. (2011). Highly efficient production of soluble proteins from insoluble inclusion bodies by a two-step-denaturing and refolding method. *PLoS ONE, 6*(7), e22981.

Yarizade, A., Aram, F., Niazi, A., & Ghasemi, Y. (2012). Evaluation of effect of β-lactam antibiotics on suppression of different strains of *Agrobacterium tumefaciens* and on wheat mature embryo culture. *Iranian Journal of Pharmaceutical Sciences, 8*(4), 267–276.

Yasmeen, A., Mirza, B., & Inayatullah, S. (2009). *In Planta* transformation of tomato. *Plant Molecular Biology Reporter, 27*(1), 20-28.

Yesilirmak, F., & Sayers, Z. (2009). Heterologous expression of plant genes. *International Journal of Plant Genomics, 2009,* 296482.

Youm, J. W., Jeon, J. H., & Kim, H. (2008). Transgenic tomatoes expressing human beta-amyloid for use as a vaccine against Alzheimer's disease. *Biotechnology Letters, 30*(10), 1839-1845.

APPENDICES

A. Buffer preparation

A1) Sample Buffer (SDS reducing buffer)

Deionized water	3.55 mL
0.5 M Tris-HCl, pH 6.8	1.25 mL
Glycerol	2.5 mL
10% (w/v) SDS	2.0 mL
Total volume	**9.3 mL**

A2) SDS loading buffer

The SDS loading buffer was prepared by adding 50 µL β-mercaptoethanol to 930 µL of sample buffer prior to use.

A3) Buffer for purification under native condition

NP1-10* (Binding/lysis buffer for native condition, 20 mL

50 mM NaH_2PO_4	0.069 g (MW 137.99 g/mol)
300 mM NaCl	0.175 g (MW 58.44 g/mol)
10 mM imidazole	0.068 g (MW 68.08 g/mol)

NP1-20 (Wash buffer for native condition, 20 mL)

50 mM NaH_2PO_4	0.069 g (MW 137.99 g/mol)
300 mM NaCl	0.175 g (MW 58.44 g/mol)
20 mM imidazole	0.014 g (MW 68.08 g/mol)

NP1-500 (Elution buffer for native conditions, 20 mL)

50 mM NaH_2PO_4	0.069 g (MW 137.99 g/mol)
300 mM NaCl	0.175 g (MW 58.44 g/mol)
500 mM imidazole	0.170 g (MW 68.08 g/mol)

All solution was adjusted to pH 8 using NaOH and filter sterilized (0.2 µm).

B. Media preparation

B1) Difco™ LB Broth, Miller (Luria-Bertani)

To prepare 100 mL, 2.5 g of the powder was dissolved in 100 mL distilled water and then mixed thoroughly. The solution was autoclaved at 121 for 20 minutes.

B2) Luria Agar (HIMEDIA)

To prepare 100 mL, 3.5 g of the powder was dissolved in 100 mL distilled water and then mixed thoroughly. The solution was autoclaved at 121^0C for 20 minutes.

B3) Infiltration media

Infiltration media was prepared by adding 0.5X Murashige and Skoog basal salt mixture (SIGMA), 3 % sucrose, 0.5 g/L MES. The solution was mixed thoroughly and autoclaved at 121^0C for 20 minutes.

C. Preparation of antibiotic stock concentration

C1) Kanamycin, 100 mg/mL (Amresco)

To prepare 5 mL stock concentration of 100 mg/mL kanamycin; 0.5 g of kanamycin powder was added in 5 mL sterile distilled water and mix thoroughly until dissolve. The stock solutions then filter sterilized (0.2 µm) and kept at -20°C.

C2) Rifampicin, 50 mg/mL (*Phyto* Technology Laboratories™)

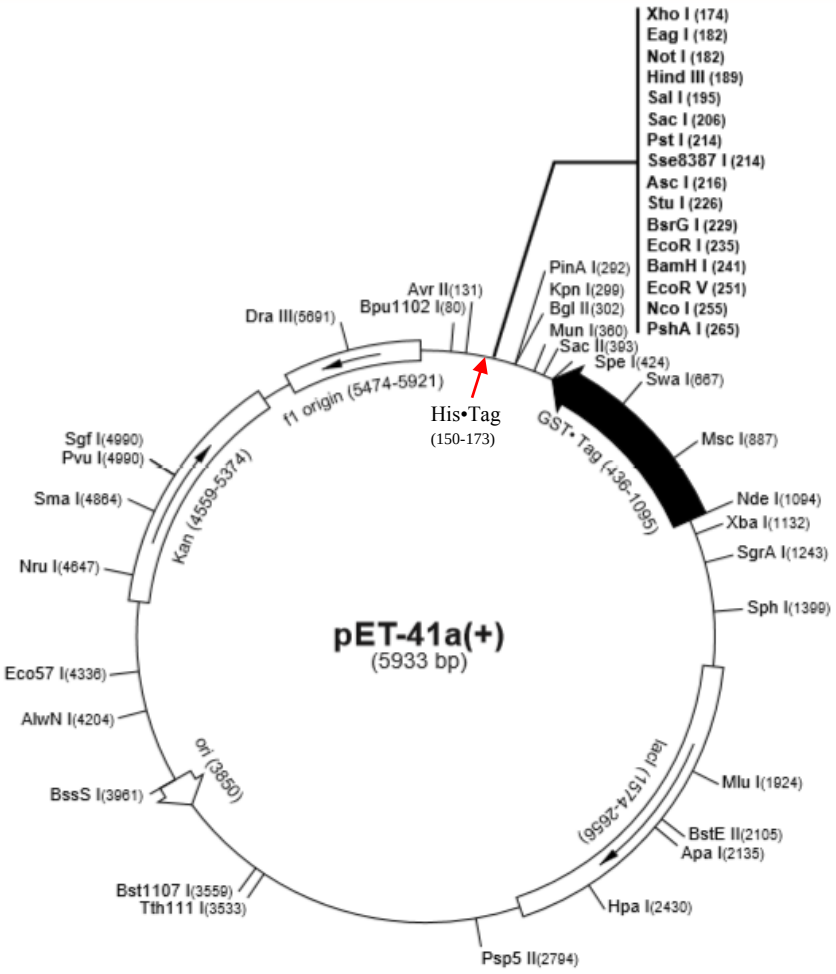
To prepare 1 mL stock concentration of 50 mg/mL rifampicin, 0.05 g of rifampicin powder was added in 1 mL methanol and mix thoroughly until dissolved. The stock solution was kept at -20°C.

C3) Chloramphenicol, 100 mg/mL (DUCHEFA)

To prepare 1 mL stock concentration of 100 mg/mL chloramphenicol, 0.1 g of chloramphenicol powder was added in 1 mL methanol and mix thoroughly until dissolved. The stock solution was kept at at -20°C.

D. Expression vector, pET-41a(+)

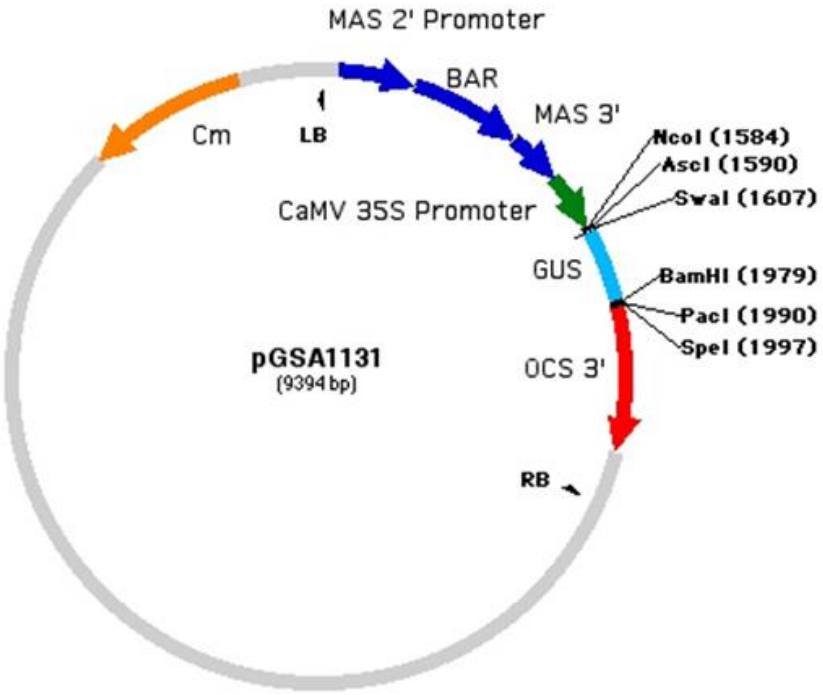
Plasmid map of pET-41a (+). Unique sites are shown on the circle map. Next to the *Xho*I restriction site is histidine sequence that encodes for a string of eight histidine residues (shown by red arrow). Image is provided by Novagen (EMD Milipore).



pET-41a(+) sequence landmarks	
T7 promoter: 1167–1183	His•Tag coding sequence: 150–173
T7 transcription start: 1166	T7 terminator: 26–72
GST•Tag coding sequence: 436–1095	*lac*I coding sequence: 1574–2656
His•Tag coding sequence: 397–414	pBR322 origin: 3850
S•Tag coding sequence: 310–354	Kan coding sequence: 4559–5374
Multiple cloning sites *(Psh*AI-*Xho*I):174–265	F1 origin: 5474–5921

E. Plant binary vector, pGSA1131

The pGSA1131 used in this study was obtained from The Arabidopsis Information Resource (TAIR). This binary vector is based on binary vector pCAMBIA1200 and harbours coding regions for β-glucuronidase (*GUS*) and herbicide resistance (*Bar*). Both of these genes are under transcriptional control of a 35S Cauliflower Mosaic Virus (CaMV 35s) promoter of approximately 200 bp in length. The map of binary vector pGSA1131 is showed below.



F. *In silico* analysis of r-msAdh1 protein

In silico analysis was performed using EXPASY ProtParam tool (http://web.expasy.org/protparam/). Results of the analysis are shown in table below.

Table A: Results of *in silico* analysis of r-msAdh1 using EXPASY ProtParam tool.

Elements	Details
Number of amino acids	380
Molecular weight:	41211.4 dalton
Theoretical pI	5.87
Estimated half-life	The N-terminal of the sequence considered is M (Met). The estimated half-life is: 30 hours (mammalian reticulocytes, in vitro). >20 hours (yeast, in vivo). >10 hours (Escherichia coli, in vivo).
Instability index	The instability index (II) is computed to be 30.76 This classifies the protein as stable.
Aliphatic index	81.03
Grand average of hydropathicity (GRAVY):	0.006

G. Absorbance reading at 340 nm

The ADH catalytic activity was detected by monitoring absorbance changes at 340 nm. The reaction was repeated for three times and data obtained is shown as below. Details about table**:** A: Absorbance reading obtained from total protein of BL21 (DE3) transformed with pET-41a(+), B: Absorbance reading obtained from total protein of BL21 (DE3) with r-*msAdh1*expression.

Time (minutes)	A (control)			Mean	B			Mean
	1	2	3		1	2	3	
1	0.030	0.012	0.036	0.026	0.011	0.027	0.041	0.026
2	0.045	0.016	0.037	0.033	0.033	0.038	0.045	0.039
3	0.048	0.027	0.039	0.038	0.056	0.049	0.055	0.053
4	0.060	0.041	0.045	0.049	0.062	0.058	0.060	0.060
5	0.068	0.044	0.046	0.053	0.075	0.063	0.062	0.067
6	0.078	0.054	0.045	0.059	0.079	0.071	0.075	0.075
7	0.080	0.061	0.048	0.063	0.092	0.076	0.085	0.084
8	0.082	0.062	0.054	0.066	0.098	0.081	0.082	0.087
9	0.089	0.068	0.059	0.072	0.102	0.086	0.084	0.091
10	0.095	0.070	0.054	0.073	0.108	0.089	0.086	0.094
11	0.095	0.078	0.057	0.077	0.110	0.090	0.095	0.098
12	0.100	0.087	0.064	0.084	0.118	0.100	0.103	0.107
13	0.103	0.093	0.069	0.088	0.121	0.110	0.110	0.113
14	0.107	0.095	0.072	0.091	0.129	0.111	0.114	0.118
15	0.113	0.100	0.083	0.099	0.131	0.120	0.120	0.124

H. Student's paired t-test

Is there significant difference between ADH catalytic activity obtained from sample of total protein of BL21 (DE3) with r-*msAdh1* expression and BL21(DE3) transformed with empty plasmid, pET-41a(+)?

Null hypothesis: There is no significant difference between ADH catalytic activity obtained from sample of total protein of Bl21 (DE3) with r-*msAdh1* expression and Bl21(DE3) transformed with empty plasmid, pET-41a(+).

Alternative hypothesis: There is significant difference between ADH activity obtained from sample of total protein of BL21 (DE3) with r-*msAdh1* expression and BL21(DE3) transformed with empty plasmid, pET-41a(+).

The results of a paired t-test performed at 02:38 on 27-JUN-2015

t= -15.1
degrees of freedom = 14

The probability of this result, assuming the null hypothesis, is 0.000

Group A: Number of items= 15
0.193 0.213 0.265 0.346 0.362 0.398 0.438 0.466 0.498 0.510 0.543 0.607 0.651 0.671 0.736

Mean = 0.460
95% confidence interval for Mean: 0.3677 thru 0.5519
Standard Deviation = 0.166
Hi = 0.736 Low = 0.193
Median = 0.466
Average Absolute Deviation from Median = 0.133

Group B: Number of items= 15
0.273 0.334 0.418 0.474 0.502 0.587 0.647 0.655 0.683 0.703 0.744 0.816 0.884 0.904 0.965

Mean = 0.639
95% confidence interval for Mean: 0.5237 thru 0.7548
Standard Deviation = 0.209
Hi = 0.965 Low = 0.273
Median = 0.655
Average Absolute Deviation from Median = 0.164

Group A-B: Number of items= 15
-0.233 -0.233 -0.229 -0.209 -0.209 -0.205 -0.201 -0.189 -0.189 -0.173 -0.153 -0.140 -0.128 -0.121 -8.000E-02

Mean = -0.179
95% confidence interval for Mean: -0.2049 thru -0.1541
Standard Deviation = 4.589E-02
Hi = -8.000E-02 Low = -0.233
Median = -0.189
Average Absolute Deviation from Median = 3.567E-02

Since *p*-value is 0.000 <0.05, null hypothesis is rejected and alternative hypothesis is accepted. There is significant difference between catalytic activity of ADH obtained from sample of total protein of Bl21 (DE3) with r-*msAdh1* expression and Bl21(DE3) transformed with empty plasmid, pET-41a(+).

I) Similarity analysis of sago palm r-*msAdh1* with tomato *Adh*

Sequence homology search of r-*msAdh1* using NCBI BLASTN against the nucleotide database of tomato (taxid:4081) in Genebank NCBI showed highest identity (78%) with *Adh2* tomato.

Description	Max score	Total score	Query cover	E value	Identity	Accession
Solanum lycopersicum, alcohol dehydrogenase (ADH2), mRNA	928	928	98%	0.0	78%	NM_001247170.1
Solanum lycopersicum, cDNA, clone: LEFL2007J14, HTC in fruit	922	922	98%	0.0	78%	AK326446.1
PREDICTED: *Solanum lycopersicum*, alcohol dehydrogenase 1 (LOC101261114), mRNA	755	755	97%	0.0	75%	XM_004237682.2
L.esculentum mRNA for alcohol dehydrogenase	324	324	33%	4e-86	79%	X60600.1
Solanum lycopersicum chromosome ch06, complete genome	307	1035	94%	3e-81	81%	HG975518.1
L.esculentum (de Ruiter 83G38) *Adh2* gene	307	946	94%	3e-81	81%	X77233.1

J. Sequencing results

(i) Multiple alignments were performed using ClustalW2 to confirm the amplicon obtained in PCR using gDNA of $T_0$1 as a template is *r-msAdh1* gene. The *r-msAdh1* specific forward and reverse primers (5_Comseq_Adh and 3_Comseq_Adh) were used for nucleotide sequencing. The result is shown below.

Forward primer: 5_Comseq_Adh

```
r-msAdh1                ATGGCAAGCAGTGTTGGTCAAGTGATCAGATGCAGAGCGGCGGTCTCATG 50
1st_BASE_5_comseq_Adh   -------------------------NNNNNGCAN-GCG-CGGTCTCATG 22
                                                    ***  *** **********

r-msAdh1                GGAGGCCGGGAAGCCGCTGGTGATGGAGGAGGTCGAGGTTGCGCCGCCGC 100
1st_BASE_5_comseq_Adh   GGAGGCCGGGA-GCCGCTGGTGATGGAGGAGGTCGAGGTTGCGCCGCCGC 71
                        *********** ************************************

r-msAdh1                AGGCGATGGAGGTTCGGATGAAGATCCTTTATACTTCCCTCTGCCACACT 150
1st_BASE_5_comseq_Adh   AGGCGATGGAGGTTCGGATGAAGATCCTTTATACTTCCCTCTGCCACACT 121
                        **************************************************

r-msAdh1                GATGTCTACTTCTGGGAAGCTAAGGGCCAGACTCCTGTCTTTCCTCGGAT 200
1st_BASE_5_comseq_Adh   GATGTCTACTTCTGGGAAGCTAAGGGCCAGACTCCTGTCTTTCCTCGGAT 171
                        **************************************************

r-msAdh1                CTTTGGCCATGAAGCTGGAGGGATTATAGAGAGTGTTGGGGAGGGTGTGA 250
1st_BASE_5_comseq_Adh   CTTTGGCCATGAAGCTGGAGGGATTATAGAGAGTGTTGGGGAGGGTGTGA 221
                        **************************************************

r-msAdh1                CTGAACTTGCACCAGGAGACCATGTCCTCCCTATATTCACTGGAGAATGC 300
1st_BASE_5_comseq_Adh   CTGAACTTGCACCAGGAGACCATGTCCTCCCTATATTCACTGGAGAATGC 271
                        **************************************************

r-msAdh1                AAAGAGTGTGCTCACTGTAAGTCCGAGGAGAGCAACATGTGTGATCTCCT 350
1st_BASE_5_comseq_Adh   AAAGAGTGTGCTCACTGTAAGTCCGAGGAGAGCAACATGTGTGATCTCCT 321
                        **************************************************

r-msAdh1                CAGGATAAACACGGATCGGGGAGTGATGATCAATGATGGGAAATCGAGGT 400
1st_BASE_5_comseq_Adh   CAGGATAAACACGGATCGGGGAGTGATGATCAATGATGGGAAATCGAGGT 371
                        **************************************************

r-msAdh1                TCACTATCAATGGAAAGCCCATTTACCATTTCCTAGGAACATCCACTTTC 450
1st_BASE_5_comseq_Adh   TCACTATCAATGGAAAGCCCATTTACCATTTCCTAGGAACATCCACTTTC 421
                        **************************************************

r-msAdh1                AGCGAGTACACCGTTGTCCATGTTGGCTGTGTTGCCAAGATCAACCCCTT 500
1st_BASE_5_comseq_Adh   AGCGAGTACACCGTTGTCCATGTTGGCTGTGTTGCCAAGATCAACCCCTT 471
                        **************************************************

r-msAdh1                GGCTCCCCTTGATAAAGTTTGTGTTCTTAGCTGTGGCATTTCAACAGGAT 550
1st_BASE_5_comseq_Adh   GGCTCCCCTTGATAAAGTTTGTGTTCTTAGCTGTGGCATTTCAACAGGAT 521
                        **************************************************
```

```
r-msAdh1                TTGGTGCGACTGTTAATGTTGCAAAACCACCAAAGGGATCGACGGTGGCT 600
1st_BASE_5_comseq_Adh   TTGGTGCGACTGTTAATGTTGCAAAACCACCAAAGGGATCGACGGTGGCT 571
                        **************************************************

r-msAdh1                GTTTTTGGCTTGGGAGCTGTAGGCCTTGCTGCTGCAGAAGGTGCTAGAGC 650
1st_BASE_5_comseq_Adh   GTTTTTGGCTTGGGAGCTGTAGGCCTTGCTGCTGCAGAAGGTGCTAGAGC 621
                        **************************************************

r-msAdh1                TTCAGGGGCATCAAGAATCATTGGTGTTGATGTGAACCCCAAGAGGTTTG 700
1st_BASE_5_comseq_Adh   TTCAGGGGCATCAAGAATCATTGGTGTTGATGTGAACCCCAAGAGGATTG 671
                        ********************************************** ***

r-msAdh1                AGGAAGCAATGAAGTTCGGTTGCGCGGAGTTTGTGAATCCAATGGACCAT 750
1st_BASE_5_comseq_Adh   AGGAAGCAATGAAGTTCGGTTGCGCGGAGTTTGTGAATCCAATGGACCAT 721
                        **************************************************

r-msAdh1                GACAAGCCAGTCCAAGAGGTGATTGCTGAGATGACAAATGGTGGAGTTGA 800
1st_BASE_5_comseq_Adh   GACAAGCCAGTCCAAGAGGTGATTGCTGAGATGACAGATGGTGGAGTTGA 771
                        ********************************** *************

r-msAdh1                TCGAAGCGTTGAATGCACTGGCAACATAAATGCCATGATATCTGCATTCG 850
1st_BASE_5_comseq_Adh   TCGAAGCGTTGAATGCACTGGCAACATCAATGCCATGATATCTGCATTCG 821
                        ************************** **********************

r-msAdh1                AATGTGTCCATGATGGCTGGGGTGTTGCTGTACTGGTTGGGG-TGCCTCA 899
1st_BASE_5_comseq_Adh   AATGTGTCCATGATGGGTGGGCTGGTGCTGTACTGGGTGGGGGCGCCCCC 871
                        **************** **** ** *********** *****  *** *

r-msAdh1                CAAA-GAAGCTGAGTTCAAAACCCACC-CTATGAACTTCCTTAAC-GAAA 946
1st_BASE_5_comseq_Adh   CAAAAGAATCTGAGATCAACACCCACCTCTAGGAACTTCCATCACAGGAA 921
                        **** *** ***** **** ******* *** ******** * ** * **

r-msAdh1                GAACTCTTAA-GGGAACCTTC--TTTGGGAACTA--TAAACCGCGCTCTG 991
1st_BASE_5_comseq_Adh   AAACTCGTAACGGGAACCCTTCATTTGCCAACTAATNAAACCGTCATCTG 971
                            ***** *** ******* *   ****  *****   ******   ****

r-msAdh1                ACATTCCTGCAGTTGTTGAGAAGTACATGAACAAGGAGCTAGAATTGGAG 1041
1st_BASE_5_comseq_Adh   A-------GCATT-----------------ACCGGAAGTTAGTAC--GAG 995
                        *       *** *                 **  * ** *** *   ***


r-msAdh1                AAGTTCATCACCCACAGTGTGCCTTTCTCTGAGATCAACAAGGCCTTTGG 1091
1st_BASE_5_comseq_Adh   AACTTC--CATGCAAAACCTGACATTC---GAA-------------TTGG 1027
                           ** ***  **  ** *   ** * ***   **             ****

r-msAdh1                CTACATGCTCAAGGGGGAGAGCCTTAGGTGCATCATTCACATGGATGGT 1140
1st_BASE_5_comseq_Adh   -------------GGAGGAAGCC-----------ACCCCC--------- 1043
                                     ** *  ****           *  * *
```

Reverse primer: 3_Comseq_Adh

```
r-msAdh1                ----ATGGCAAG----CAGTGTTGGTCAAGTGATCAGATGCAGAGC-GGC 41
1st_BASE_3_Comseq_Adh   NTTTAAGGCAAAGCACCGGTTGTGGTCAAGTGATCAGATGCAGAACCGGC 50
                            * *****     * **  ********************** * ***

r-msAdh1                GGTCT-CATGGGAGGCC-GGGAAGCCGC-TGGTGATGGAGGAGGTCGAGG 88
1st_BASE_3_Comseq_Adh   GGTTTTCATGGGAGGCCCGGGAAGCCGCCTGGTGATGGAGGAGTTCGAGG 100
                        *** * *********** ********** ************** ******

r-msAdh1                TTGCGCCGCC-GCAGGCGATGGAGGTT-CGGATGAAGATCCTTT-ATACT 135
1st_BASE_3_Comseq_Adh   TTGCGCCCCCCGCAGGCGATGGAGGTNTCGGATGAAGATCCTTTTATACT 150
                        ******* ** ***************  ***************** *****

r-msAdh1                TCCCTCTGCC-ACACTGATGTCTACTTCTGGGAAGCTAAGGGCCAGACTC 184
1st_BASE_3_Comseq_Adh   TCCCTCTGCCCACACTGATGTCTACTTCTGGGAAGCTAAGGGCCAGACTC 200
                        ********** ****************************************

r-msAdh1                CTGTCTTTCCTCGGATCTTTGGCCATGAAGCTGGAGGGATTATAGAGAGT 234
1st_BASE_3_Comseq_Adh   CTGTCTTTCCTCGGATCTTTGGCCATGAAGCTGGAGGGATTATAGAGAGT 250
                        **************************************************

r-msAdh1                GTTGGGGAGGGTGTGACTGAACTTGCACCAGGAGACCATGTCCTCCCTAT 284
1st_BASE_3_Comseq_Adh   GTTGGGGAGGGTGTGACTGAACTTGCACCAGGAGACCATGTCCTCCCTAT 300
                        **************************************************

r-msAdh1                ATTCACTGGAGAATGCAAAGAGTGTGCTCACTGTAAGTCCGAGGAGAGCA 334
1st_BASE_3_Comseq_Adh   ATTCACTGGAGAATGCAAAGAGTGTGCTCACTGTAAGTCCGAGGAGAGCA 350
                        **************************************************

r-msAdh1                ACATGTGTGATCTCCTCAGGATAAACACGGATCGGGGAGTGATGATCAAT 384
1st_BASE_3_Comseq_Adh   ACATGTGTGATCTCCTCAGGATAAACACGGATCGGGGAGTGATGATCAAT 400
                        **************************************************

r-msAdh1                GATGGGAAATCGAGGTTCACTATCAATGGAAAGCCCATTTACCATTTCCT 434
1st_BASE_3_Comseq_Adh   GATGGGAAATCGAGGTTCACTATCAATGGAAAGCCCATTTACCATTTCCT 450
                        **************************************************

r-msAdh1                AGGAACATCCACTTTCAGCGAGTACACCGTTGTCCATGTTGGCTGTGTTG 484
1st_BASE_3_Comseq_Adh   AGGAACATCCACTTTCAGCGAGTACACCGTTGTCCATGTTGGCTGTGTTG 500
                        **************************************************

r-msAdh1                CCAAGATCAACCCCTTGGCTCCCCTTGATAAAGTTTGTGTTCTTAGCTGT 534
1st_BASE_3_Comseq_Adh   CCAAGATCAACCCCTTGGCTCCCCTTGATAAAGTTTGTGTTCTTAGCTGT 550
                        **************************************************

r-msAdh1                GGCATTTCAACAGGATTTGGTGCGACTGTTAATGTTGCAAAACCACCAAA 584
1st_BASE_3_Comseq_Adh   GGCATTTCAACAGGATTTGGTGCGACTGTTAATGTTGCAAAACCACCAAA 600
                        **************************************************

r-msAdh1                GGGATCGACGGTGGCTGTTTTTGGCTTGGGAGCTGTAGGCCTTGCTGCTG 634
1st_BASE_3_Comseq_Adh   GGGATCGACGGTGGCTGTTTTTGGCTTGGGAGCTGTAGGCCTTGCTGCTG 650
                        **************************************************

r-msAdh1                CAGAAGGTGCTAGAGCTTCAGGGGCATCAAGAATCATTGGTGTTGATGTG 684
1st_BASE_3_Comseq_Adh   CAGAAGGTGCTAGAGCTTCAGGGGCATCAAGAATCATTGGTGTTGATGTG 700
                        **************************************************

r-msAdh1                AACCCCAAGAGGTTTGAGGAAGCAATGAAGTTCGGTTGCGCGGAGTTTGT 734
1st_BASE_3_Comseq_Adh   AACCCCAAGAGGTTTGAGGAAGCAATGAAGTTCGGTTGCGCGGAGTTTGT 750
                        **************************************************
```

```
r-msAdh1                GAATCCAATGGACCATGACAAGCCAGTCCAAGAGGTGATTGCTGAGATGA 784
1st_BASE_3_Comseq_Adh   GAATCCAATGGACCATGACAAGCCAGTCCAAGAGGTGATTGCTGAGATGA 800
                        **************************************************

r-msAdh1                CAAATGGTGGAGTTGATCGAAGCGTTGAATGCACTGGCAACATAAATGCC 834
1st_BASE_3_Comseq_Adh   CAAATGGTGGAGTTGATCGAAGCGTTGAATGCACTGGCAACATAAATGCC 850
                        **************************************************
r-msAdh1                ATGATATCTGCATTCGAATGTGTCCATGATGGCTGGGGTGTTGCTGTACT 884
1st_BASE_3_Comseq_Adh   ATGATATCTGCATTCGAATGTGTCCATGATGGCTGGGGTGTTGCTGTACT 900
                        **************************************************

r-msAdh1                GGTTGGGGTGCCTCACAAAGAAGCTGAGTTCAAAACCCACCCTATGAACT 934
1st_BASE_3_Comseq_Adh   GGTTGGGGTGCCTCACAAAGAAGCTGAGTTCAAAACCCACCCTATGAACT 950
                        **************************************************

r-msAdh1                TCCTTAACGAAAGAACTCTTAAGGGAACCTTCTTTGGGAACTATAAACCG 984
1st_BASE_3_Comseq_Adh   TCCTTAACGAAAGAACTCTTAAGGGAACCTTCTTTGGGAACTATAAACCG 1000
                        **************************************************

r-msAdh1                CGCTCTGACATTCCTGCAGTTGTTGAGAAGTACATGAACAAGGAGCTAGA 1034
1st_BASE_3_Comseq_Adh   CGCTCTGACATTCCTGCAGTTGTTGAGAAGTACATGAACAAGGAGCTAGA 1050
                        **************************************************

r-msAdh1                ATTGGAGAAGTTCATCACCCACAGTGTGCCTTTCTCTG-AGATCAACAAG 1083
1st_BASE_3_Comseq_Adh   ATTGGAGAAGTTCATCACCCACAGTGTGCCTTTCTCTGTAGATCAACAAG 1100
                        ************************************ ***********

r-msAdh1                GCCTTTGGCTACATGCTCAA-GGGGGAGAGCCTTAGGTGCATCATTCACA 1132
1st_BASE_3_Comseq_Adh   GCCTT-GGCTACATGCTCAACGGGGNNGNG-------------------- 1129
                        ***** ************** ****  * *

r-msAdh1                TGGATGGT 1140
1st_BASE_3_Comseq_Adh   -------- 
```

(ii) Multiple alignments were performed using ClustalW2 to confirm the amplicon obtained in PCR using gDNA of $T_0$1 as a template is *bar* gene. The *bar* specific forward and reverse primers (BAR3_F and BAR3_R) were used for nucleotide sequencing. The result is shown below.

Forward primer: BAR3_F

```
Bar                      ATGAGCCCAGAACGACGCCCGGCCGACATCCGCCGTGCCACCGAGGCGGA 50
1st_BASE_BAR3_F          -----NNNAACATGATTATTGTTCGAC--------------------GGA 25
                               *  * **     *  ****                    ***

Bar                      CATGCCGGCGG-TCTGCACCATCGTCAACCACTACATCGAGACAAGCACG 99
1st_BASE_ BAR3_F         -ATGCCGGCGGATCTGCACCATCGTCAACCACTACATCGAGACAAGCACG 74
                          ********** ************************************

Bar                     GTCAACTTCCGTACCGAGCCGCAGGAACCGCAGGAGTGGACGGACGACCT 149
1st_BASE_ BAR3_F        GTCAACTTCCGTACCGAGCCGCAGGAACCGCAGGAGTGGACGGACGACCT 124
                        **************************************************

Bar                     CGTCCGTCTGCGGGAGCGCTATCCCTGGCTCGTCGCCGAGGTGGACGGCG 199
1st_BASE_ BAR3_F        CGTCCGTCTGCGGGAGCGCTATCCCTGGCTCGTCGCCGAGGTGGACGGCG 174
                        **************************************************

Bar                     AGGTCGCCGGCATCGCCTACGCGGGCCCCTGGAAGGCACGCAACGCCTAC 249
1st_BASE_ BAR3_F        AGGTCGCCGGCATCGCCTACGCGGGCCCCTGGAAGGCACGCAACGCCTAC 224
                        **************************************************

Bar                     GACTGGACGGCCGAGTCGACCGTGTACGTCTCCCCCCGCCACCAGCGGAC 299
1st_BASE_ BAR3_F        GACTGGACGGCCGAGTCGACCGTGTACGTCTCCCCCCGCCACCAGCGGAC 274
                        **************************************************

Bar                     GGGACTGGGCTCCACGCTCTACACCCACCTGCTGAAGTCCCTGGAGGCAC 349
1st_BASE_ BAR3_F        GGGACTGGGCTCCACGCTCTACACCCACCTGCTGAAGTCCCTGGAGGCAC 324
                        **************************************************

Bar                     AGGGCTTCAAGAGCGTGGTCGCTGTCATCGGGCTGCCCAACGACCCGAGC 399
1st_BASE_ BAR3_F        AGGGCTTCAAGAGCGTGGTCGCTGTCATCGGGCTGCCCAACGACCCGAGC 374
                        **************************************************

Bar                     GTGCGCATGCACGAGGCGCTCGGATATGCCCCCCGCGGCATGCTGCGGGC 449
1st_BASE_ BAR3_F        GTGCGCATGCACGAGGCGCTCGGATATGCCCCCCGCGGCATGCTGCGGGC 424
                        **************************************************

Bar                     GGCCGGCTTCAAGCACGGGAACTGGCATGACGTGGGTTTCTGGCAGCTGG 499
1st_BASE_ BAR3_F        GGCCGGCTTCAAGCACGGGAACTGGCATGACGTGGGTTTCTGGCAGCTGG 474
                        **************************************************

Bar                     ACTTCAGCCTGCCGGTACCGCCCCGTCCGGTCCTGCCCGTCACCGAGAT- 548
1st_BASE_ BAR3_F        ACTTCAGCCTGCCGGTACCGCCCCGTCCGGTCCTGCCCGCCACCGAGATA 524
                        ************************************ *********
```

Reverse primer: BAR3_R

```
Bar                  ----ATGAGCCCAGAACGACGCCCGGCCGACATCCGCCGTGCCACCGAGG 46
1st_BASE_BAR3_R      TTTTATGAGCCCCAAACGACGCCCGGCCGACATCCGCCGTGCCACCGAGG 50
                         ********  ***********************************

Bar                  CGGACATGCCGGCGGTCTGCACCATCGTCAACCACTACATCGAGACAAGC 96
1st_BASE_ BAR3_R     CGGACATGCCGGCGGTCTGCACCATCGTCAACCACTACATCGAGACAAGC 100
                     ************************************************

Bar                  ACGGTCAACTTCCGTACCGAGCCGCAGGAACCGCAGGAGTGGACGGACGA 146
1st_BASE_ BAR3_R     ACGGTCAACTTCCGTACCGAGCCGCAGGAACCGCAGGAGTGGACGGACGA 150
                     ************************************************

Bar                  CCTCGTCCGTCTGCGGGAGCGCTATCCCTGGCTCGTCGCCGAGGTGGACG 196
1st_BASE_ BAR3_R     CCTCGTCCGTCTGCGGGAGCGCTATCCCTGGCTCGTCGCCGAGGTGGACG 200
                     ************************************************

Bar                  GCGAGGTCGCCGGCATCGCCTACGCGGGCCCCTGGAAGGCACGCAACGCC 246
1st_BASE_ BAR3_R     GCGAGGTCGCCGGCATCGCCTACGCGGGCCCCTGGAAGGCACGCAACGCC 250
                     ************************************************

Bar                  TACGACTGGACGGCCGAGTCGACCGTGTACGTCTCCCCCCGCCACCAGCG 296
1st_BASE_ BAR3_R     TACGACTGGACGGCCGAGTCGACCGTGTACGTCTCCCCCCGCCACCAGCG 300
                     ************************************************

Bar                  GACGGGACTGGGCTCCACGCTCTACACCCACCTGCTGAAGTCCCTGGAGG 346
1st_BASE_ BAR3_R     GACGGGACTGGGCTCCACGCTCTACACCCACCTGCTGAAGTCCCTGGAGG 350
                     ************************************************

Bar                  CACAGGGCTTCAAGAGCGTGGTCGCTGTCATCGGGCTGCCCAACGACCCG 396
1st_BASE_ BAR3_R     CACAGGGCTTCAAGAGCGTGGTCGCTGTCATCGGGCTGCCCAACGACCCG 400
                     ************************************************

Bar                  AGCGTGCGCATGCACGAGGCGCTCGGATATGCCCCCCGCGGCATGCTGCG 446
1st_BASE_ BAR3_R     AGCGTGCGCATGCACGAGGCGCTCGGATATGCCCCCCGCGGCATGCTGCG 450
                     ************************************************

Bar                  GGCGGCCGGCTTCAAGCACGGGAACTGGCATGACGTGGGTTTCTGGCAGC 496
1st_BASE_ BAR3_R     GGCGGCCGGCTTCAAGCACGGGAACTGGCATGACGTGGGTT-CTGGCAGC 499
                     ************************************ ********

Bar                  TGGACTTCAGCCTGCCGGTACCGCCCCGTCCGGTCCT-GCCCGTCACCGA 545
1st_BASE_ BAR3_R     TG-ACTCCA---TGCT-----CGACGCG---AATACTAGAACATNN---- 533
                     ** *** **   ***      ** * **     * ** *  * *

Bar                  GAT 548
1st_BASE_ BAR3_R     ---
```

(iii) Multiple alignments were performed using ClustalW2 to confirm the amplicon obtained in RT-PCR using first strand cDNA of T_1K as a template is *r-msAdh1* gene. The *r-msAdh1* specific forward and reverse primers (5_Comseq_Adh and 5_msAdh1_R) were used for nucleotide sequencing. The result is shown below.

Forward primer: 5_Comseq_Adh

```
r-msAdh1          ATGGCAAGCAGTGTTGGTCAAGTGATCAGATGCAGAGCGGCGGTCTCATGGGAGGCCGGG 60
1st_BASE_ADH_5cq ---------------------------NNANANACAACGGCGGTCTC-TGGGAGGCCGGG 32
                                             *   * * ********** ************

r-msAdh1          AAGCCGCTGGTGATGGAGGAGGTCGAGGTTGCGCCGCCGCAGGCGATGGAGGTTCGGATG 120
1st_BASE_ADH_5cq A-GCCGCTGGTGATGGAGGAGGTCGAGGTTGCGCCGCCGCAGGCGATGGAGGTTCGGATG 91
                 * ************************************************************

r-msAdh1          AAGATCCTTTATACTTCCCTCTGCCACACTGATGTCTACTTCTGGGAAGCTAAGGGCCAG 180
1st_BASE_ADH_5cq AAGATCCTTTATACTTCCCTCTGCCACACTGATGTCTACTTCTGGGAAGCTAAGGGCCAG 151
                 ************************************************************

r-msAdh1          ACTCCTGTCTTTCCTCGGATCTTTGGCCATGAAGCTGGAGGGATTATAGAGAGTGTTGGG 240
1st_BASE_ADH_5cq ACTCCTGTCTTTCCTCGGATCTTTGGCCATGAAGCTGGAGGGATTATAGAGAGTGTTGGG 211
                 ************************************************************

r-msAdh1          GAGGGTGTGACTGAACTTGCACCAGGAGACCATGTCCTCCCTATATTCACTGGAGAATGC 300
1st_BASE_ADH_5cq GAGGGTGTGACTGAACTTGCACCAGGAGACCATGTCCTCCCTATATTCACTGGAGAATGC 271
                 ************************************************************

r-msAdh1          AAAGAGTGTGCTCACTGTAAGTCCGAGGAGAGCAACATGTGTGATCTCCTCAGGATAAAC 360
1st_BASE_ADH_5cq AAAGAGTGTGCTCACTGTAAGTCCGAGGAGAGCAACATGTGTGATCTCCTCAGGATAAAC 331
                 ************************************************************

r-msAdh1          ACGGATCGGGGAGTGATGATCAATGATGGGAAATCGAGGTTCACTATCAATGGAAAGCCC 420
1st_BASE_ADH_5cq ACGGATCGGGGAGTGATGATCAATGATGGGAAATCGAGGTTCACTATCAATGGAAAGCCC 391
                 ************************************************************

r-msAdh1          ATTTACCATTTCCTAGGAACATCCACTTTCAGCGAGTACACCGTTGTCCATGTTGGCTGT 480
1st_BASE_ADH_5cq ATTTACCATTTCCTAGGAACATCCACTTTCAGCGAGTACACCGTTGTCCATGTTGGCTGT 451
                 ************************************************************

r-msAdh1          GTTGCCAAGATCAACCCCTTGGCTCCCCTTGATAAAGTTTGTGTTCTTAGCTGTGGCATT 540
1st_BASE_ADH_5cq GTT---AA---------------------------------------------------- 456
                 ***   **

r-msAdh1          TCAACAGGATTTGGTGCGACTGTTAATGTTGCAAAACCACCAAAGGGATCGACGGTGGCT 600
1st_BASE_ADH_5cq ------------------------------------------------------------

r-msAdh1          GTTTTTGGCTTGGGAGCTGTAGGCCTTGCTGCTGCAGAAGGTGCTAGAGCTTCAGGGGCA 660
1st_BASE_ADH_5cq ------------------------------------------------------------


r-msAdh1          TCAAGAATCATTGGTGTTGATGTGAACCCCAAGAGGTTTGAGGAAGCAATGAAGTTCGGT 720
1st_BASE_ADH_5cq ------------------------------------------------------------
```

```
r-msAdh1          TGCGCGGAGTTTGTGAATCCAATGGACCATGACAAGCCAGTCCAAGAGGTGATTGCTGAG 780
1st_BASE_ADH_5cq ------------------------------------------------------------
r-msAdh1          ATGACAAATGGTGGAGTTGATCGAAGCGTTGAATGCACTGGCAACATAAATGCCATGATA 840
1st_BASE_ADH_5cq ------------------------------------------------------------


r-msAdh1          TCTGCATTCGAATGTGTCCATGATGGCTGGGGTGTTGCTGTACTGGTTGGGGTGCCTCAC 900
1st_BASE_ADH_5cq ------------------------------------------------------------

r-msAdh1          AAAGAAGCTGAGTTCAAAACCCACCCTATGAACTTCCTTAACGAAAGAACTCTTAAGGGA 960
1st_BASE_ADH_5cq ------------------------------------------------------------
```

Reverse primer:5_msAdh1_R

```
r-msAdh1        --ATGGCAAGCAGTGTTGGTCAAGTGATCAGATGCAGAGCGGCGGTCTCATGGGAGGCCG 58
1st_BASE_ADH_R TTATGGCAAGCAGTGTTGGTCAAGTGATCAGATGCAGAGCGGCGGTCTCATGGGAGGCCG 60
               ************************************************************

r-msAdh1        GGAAGCCGCTGGTGATGGAGGAGGTCGAGGTTGCGCCGCCGCAGGCGATGGAGGTTCGGA 118
1st_BASE_ADH_R GGAAGCCGCTGGTGATGGAGGAGGTCGAGGTTGCGCCGCCGCAGGCGATGGAGGTTCGGA 120
               ************************************************************

r-msAdh1        TGAAGATCCTTTATACTTCCCTCTGCCACACTGATGTCTACTTCTGGGAAGCTAAGGGCC 178
1st_BASE_ADH_R TGAAGATCCTTTATACTTCCCTCTGCCACACTGATGTCTACTTCTGGGAAGCTAAGGGCC 180
               ************************************************************

r-msAdh1        AGACTCCTGTCTTTCCTCGGATCTTTGGCCATGAAGCTGGAGGGATTATAGAGAGTGTTG 238
1st_BASE_ADH_R AGACTCCTGTCTTTCCTCGGATCTTTGGCCATGAAGCTGGAGGGATTATAGAGAGTGTTG 240
               ************************************************************

r-msAdh1        GGGAGGGTGTGACTGAACTTGCACCAGGAGACCATGTCCTCCCTATATTCACTGGAGAAT 298
1st_BASE_ADH_R GGGAGGGTGTGACTGAACTTGCACCAGGAGACCATGTCCTCCCTATATTCACTGGAGAAT 300
               ************************************************************

r-msAdh1        GCAAAGAGTGTGCTCACTGTAAGTCCGAGGAGAGCAACATGTGTGATCTCCTCAGGATAA 358
1st_BASE_ADH_R GCAAAGAGTGTGCTCACTGTAAGTCCGAGGAGAGCAACATGTGTGATCTCCTCAGGATAA 360
               ************************************************************

r-msAdh1        ACACGGATCGGGGAGTGATGATCAATGATGGGAAATCGAGGTTCACTATCAATGGAAAGC 418
1st_BASE_ADH_R ACACGGATCGGGGAGTGATGATCAATGATGGGAAATCGAGGTTCACTATCAATGGAAAGC 420
               ************************************************************

r-msAdh1        CCATTTACCATTTCCTAGGAACATCCACTTTCAGCGAGTACACCGTTGTCCATGTTGGCT 478
1st_BASE_ADH_R CCATTTACCATT-CCTAG-AAC-TCCATTT--------------------AGGNAGGCN 456
               ************ ***** *** **** **                     * *  ***

r-msAdh1        GTGTTGCCAAGATCAACCCCTTGGCTCCCCTTGATAAAGTTTGTGTTCTTAGCTGTGGCA 538
1st_BASE_ADH_R NN------------------------------------------------------------ 458
```