THESIS

CHARACTERIZATION OF TWO 1-AMINOCYCLOPROPANE-1-CARBOXYLATE (ACC) OXIDASE I AND II FLANKING REGIONS AND DETERMINATION OF THEIR PROMOTER ACTIVITIES USING AGROBACTERIUM TRANSIENT EXPRESSION

PIYAPAT PIN-ON

GRADUATE SCHOOL, KASETSART UNIVERSITY

2008

THESIS APPROVAL

GRADUATE SCHOOL, KASETSART UNIVERSITY

Master of Science (Genetic Engineering)

DEGREE

Genetic Engineering — **FIELD**

Interdisciplinary Graduate Program — **PROGRAM**

TITLE: Characterization of Two 1-Aminocyclopropane-1-Carboxylate (ACC) Oxidase I and II Flanking Regions and Determination of Their Promoter Activities using Agrobacterium Transient Expression

NAME: Mr. Piyapat Pin-on

THIS THESIS HAS BEEN ACCEPTED BY

(Assistant Professor Ratchanee Hongprayoon, Ph.D.) **THESIS ADVISOR**

(Mrs. Parichart Burns, Ph.D.) **COMMITTEE MEMBER**

(Professor Jingtair Siriphanich, Ph.D.) **COMMITTEE MEMBER**

(Assistant Professor Siriwan Prapong, Ph.D.) **GRADUATE COMMITTEE CHAIRMAN**

APPROVED BY THE GRADUATE SCHOOL ON 5/6/2008

(Associate Professor Gunjana Theeragool, D.Agr.) **DEAN**

THESIS

CHARACTERIZATION OF TWO 1-AMINOCYCLOPROPANE- 1-CARBOXYLATE (ACC) OXIDASE I AND II FLANKING REGIONS AND DETERMINATION OF THEIR PROMOTER ACTIVITIES USING AGROBACTERIUM TRANSIENT EXPRESSION

PIYAPAT PIN-ON

A Thesis Submitted in Partial Fulfillment of
the Requirements for the Degree of
Master of Science (Genetic Engineering)
Graduate School, Kasetsart University
2008

Piyapat Pin-on 2008: Characterization of Two 1-Aminocyclopropane-1-Carboxylate (ACC) Oxidase I and II Flanking Regions and Determination of Their Promoter Activities using Agrobacterium Transient Expression. Master of Science (Genetic Engineering), Major Field: Genetic Engineering, Interdisciplinary Graduate Program. Thesis Advisor: Assistant Professor Ratchanee Hongprayoon, Ph.D. 180 pages.

The *CP-ACO II* 5'flanking region (591 bp) was isolated from Khaek Nuan papaya by ligation-mediated PCR. Sequence comparison of this region and other *ACO* promoters was done using BLASTN program (NCBI). The *CP-ACO I* (1044 bp) and *CP-ACO II* (554 bp) 5'flanking regions were analyzed by PLACE PlantCARE and Plant Prom databases via bioinformatics approach. Several important elements corresponding to seed and endosperm (DOFCOREZM), light responsive (GT1CONSENSUS), hormone responsive (NTBBF1ARROLB), temperature responsive (MYCCONSENSUSAT), dehydration responsive (MYBCORE) and wounding responsive (WRKY71OS) were found in both *CP-ACO I* and *II* 5'flanking regions. However, Antioxidant responsive (ARE1) was found only in *CP-ACO II* 5'flanking region. In order to investigate promoter activity of these flanking regions in plant tissues, seven recombinant expression plasmids were constructed by replacing 35S promoter of pCAMBIA1304 with size and location variation of the flanking regions. They are named CPACOI-SPP, CPACOI-LPP, CPACOI-DOFCORN, CPACOII-SP1, CPACOII-NO-SEboxN, CPACOII-SP3 and CPACOII-SP4. The constructs were transiently transformed using *Agrobacterium* infiltration into flower, leaf, root and fruit tissues. The result indicated that *cis*-acting elements of CPACOI-DOFCORN and CPACOII-SP3 involved with root specific promoter. The 5'flanking region of CPACOI-SPP showed specific expression in flowers. All the constructs drove *GUS* expression in fruit tissue at colour break and ripening stages. The CPACOII-NO-SEboxN showed high *GUS* activity in all tested tissues.

Piyapat Pin-on
Student's signature

Thesis Advisor's signature

29, 05, 08

ACKNOWLEDGEMENTS

I would like to truthfully thank and deeply indebted to Assistant Professor Dr. Ratchanee Hongprayoon for valuable suggestion, dedicated effort encouragement for completely writing of this thesis. I would like to heartfelt thank to Dr. Parichart Burns and Professor Dr. Jingtair Siriphanich, for valuable comments.

I would like to gratefully thank Ms. Saifone Chuaboonmee for technical laboratory methods. I am deeply thank to my companions in Plant Engineering Unit and my friends for assistance and meaningful friendship.

I would like to gratefully thank Tropical Vegetable Research Center (TVRC) for kindly providing tomato fruits for using in this research.

This research was supported by The Graduate School, Kasetsart University.

Finally, I am especially appreciated my parents, my relatives and my intimate friend for consistent encouragement and understanding.

Piyapat Pin-on

April 2008

TABLE OF CONTENTS

LIST OF TABLES

LIST OF TABLES (Continued)

LIST OF TABLES (Continued)

LIST OF FIGURES

LIST OF FIGURES (Continued)

LIST OF FIGURES (Continued)

LIST OF FIGURES (Continued)

LIST OF ABBREVIATIONS

ACC	=	1-Aminocyclopropane–1-carboxylic acid
bp	=	base pairs
BSA	=	bovine serum albumin
CaMV	=	Cauliflower mosaic virus
ºC	=	degree Celsius
cds	=	coding sequence
cm	=	centimeter
DDT	=	DL-dithiothreitol
dH_2O	=	distilled water
DNA	=	deoxyribonucleic acid
DNase	=	deoxyribonuclease
dNTPs	=	deoxynucleotide triphosphate (s)
E. coli	=	*Escherichia coli*
EDTA	=	ethylene diamine tetraacetic acid
g	=	gram
GUS	=	glucuronidase
h	=	hour
HCl	=	hydrogen chloride
IPTG	=	isopropyl-beta-D-thylogalactopyranosine
kb	=	kilobase
KCl	=	potassium chloride
kDa	=	kilodalton
kv	=	kilovoltage
M	=	molar
mg	=	milligram
$MgCl_2$	=	magnesium chloride
min	=	minute (s)
µF	=	microfallod
µg	=	microgram

LIST OF ABBREVIATIONS (Continued)

µl	=	microlitre
µmol	=	micromolar
ml	=	milliliter
mM	=	millimolar
mmol	=	millimolar
mRNA	=	messenger ribonucleic acid
MS	=	Murashige and Skoog medium
msec	=	millisecond
MW	=	molecular weight
NaCl	=	sodium chloride
NaOH	=	sodium hydroxide
ng	=	nanogram
nm	=	nanomater
OD	=	optical density
PAGE	=	polyacrylamide gel eletrophoresis
PCR	=	Polymerase chain reaction
PEG	=	polyethylene glycol
pmol	=	piccomolar
RNA	=	ribonucleic acid
RNase	=	ribonuclease
rpm	=	rotations per minute
RT	=	room temperature
SDS	=	sodium dodecyl sulphate
sec	=	second
TAE	=	Tris-acetate EDTA
UV	=	ultraviolet
Vol	=	volume
X-gal	=	5-bromo-4-chloro-3-indoyl-beta-D-galactopyranosine

LIST OF ABBREVIATIONS (Continued)

X-Gluc	=	5-bromo-4-chloro-3-indoly-ß-D-glucuronide
YAC	=	Yeast Artificial Chromosome

CHARACTERIZATION OF TWO 1-AMINOCYCLOPROPANE-1-CARBOXYLATE (ACC) OXIDASE I AND II FLANKING REGIONS AND DETERMINATION OF THEIR PROMOTER ACTIVITIES USING AGROBACTERIUM TRANSIENT EXPRESSION

INTRODUCTION

1-Aminocyclopropane-1-carboxylate (ACC) oxidase is an enzyme involving in conversion of ACC to ethylene. In climacteric fruits such as tomato and apple, this enzyme, hence its gene, is very important in controlling fruit ripening. ACC oxidase (*ACO*) gene is a multigene family. Five isoforms were reported in tomato. Despite their sequence similarity. *ACO* isoforms were shown to be tissue, stress response and developmental specific. It is, therefore, likely that their regulatory regions (promoter) could play a crucial role in their differential expression. Blume *et al*. (1997) reported the isolation of flanking regions of three tomato *ACOs* (*ACO1*, *ACO2* and *ACO3*). The isolation flanking regions of *ACOs* from other climacteric fruits including apple, banana, peach were later reported. Their promoter activities were also elucidated.

In papaya, the first report of papaya *ACO* was by Neupane *et al*. (1998) in solo variety. Later Chen *et al*. (2003) was also reported in Tainong 2 variety. Their flanking regions were also reported by Rodolfo *et al* (2004).

Two papaya *ACOs* (*CP-ACO I* and *CP-ACO II*) from Thai varieties were first reported by Kumdee *et al.* (2003). *CP-ACO I* expression was shown to correlate with fruit ripening and fruit tissue. *CP-ACO II*, on the other hand, was detected in various tissues. Chuaboome *et al*. (2004) reported the isolation of *CP-ACO I* 5'flanking region and demonstrated its promoter activity in fruit tissue under transient assay.

Although there are reports of flanking regions of papaya *ACOs*, the investigation in depth regarding its *cis*-acting elements and their roles in regulating *ACO* gene expression is lacking. The current tools of bioinformatics including

promoter databases, search engine for *cis*-acting element and transcription factor binding sites combining with the tissue specific transient assay developed in this study could enhance our understanding in *ACO* gene regulation and possibly find a good regulatory DNA region for molecular biology work.

The objectives of this study are 1) to isolate and characterize flanking regions of *CP-ACO II* 2) to develop expression vectors with *CP-ACO I* and *II* derived promoters and 3) to determine their promoter activities using transient expression assays in plant tissues including leaf, root, flower and fruit.

LITERATURE REVIEW

Promoters

1. Description of Promoter

A promoter is a regulatory region of DNA located 5' upstream region of a gene, providing a control point for regulated gene transcription. Eukaryotic promoter contains a TATA box (TATAAA) typically lies very close to the transcriptional start site +1 (often within 30 bases) and enhancers, elements increase the efficiency of transcription or activate the promoter, which can be found several kilobases (5' upstream and 3' downstream) from genes themselves that shown in Figure 1.

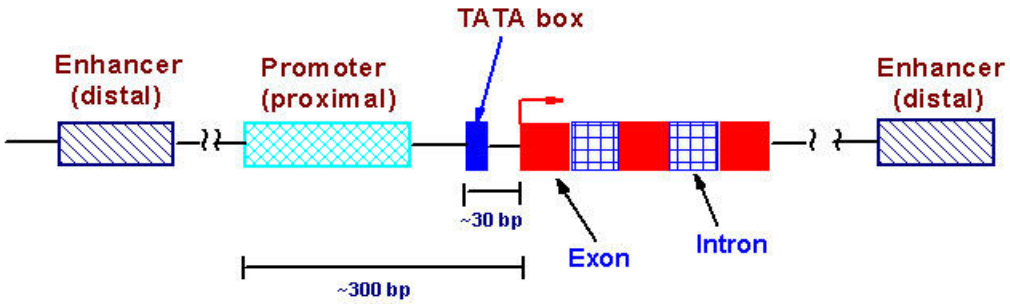


Figure 1 A common structure of eukaryote gene and transcription control regions. The promoter region consists of modular DNA sequences.

Source: Adapted from Klug and Cummings (1997)

The regulation of gene expression in eukaryote is a complicated process and extremely diverse. There are many regulatory sequences (*cis*-acting elements) located 5' upstream of the transcription start site, recognized by proteins known as transcription factors. The TATA box recognized by RNA polymerase II and transcription factors TFIIA, TFIIB and TFIID. For example, the regulation of cell-type specific expression requires at the least seven different transcription factors

binding with regulatory regions to direct the level of transcription of a given gene as shown in Figure 2 (Klug and Cummings,1997; King, 2003).

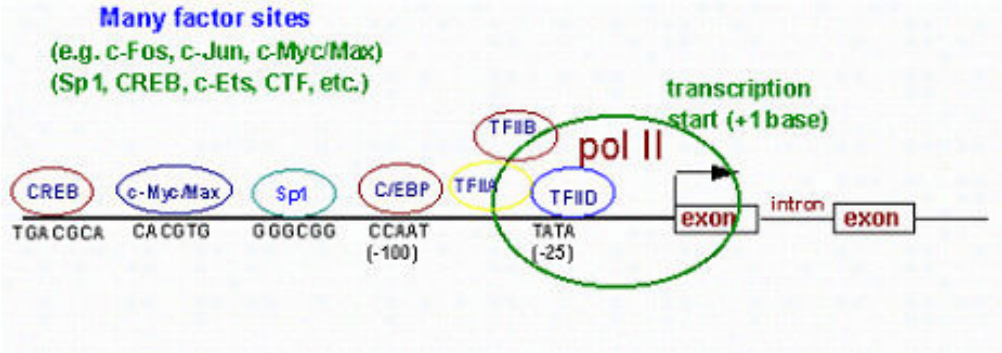


Figure 2 Diagram showed transcription binding sites and their transcription factors.

Source: King (2003)

2. Promoter Classification

The promoters can be generally divided into 4 groups according to the intended type of gene expression control; constitutive promoters, inducible promoters, tissue-specific promoters and synthetic promoters.

2.1 Constitutive Promoters

These promoters direct expression in virtually all tissues and are largely, if not entirely, independent of environmental and developmental factors. As their expression is normally not conditioned by endogenous factors, constitutive promoters are usually active across species and even across kingdoms. The CaMV 35S promoter is among widely known strong constitutive promoters driving high level of gene expression in dicot plants. Since its discovery in 1985 and application in driven chimeric gene, CaMV 35S has been used in the process of producing transgenic plants for many application including disease resistance, herbicide resistance and increase nutritional value. Recenly, Lin *et al*. (2007) reported *AtDREB1A*, a salt tolerance

gene, expression driven by *35S* promoter in tobacco. Similarly, *ZmOPR1* over expression driven by 35S promoter also enhanced the osmotic and salt tolerance in *Arabidopsis* during seed germination (Dan *et al*., 2007).

2.2 Inducible Promoters

The functions of inducible promoters are triggered by specific chemicals or physical factors. The chemically inducible promoters are triggered by either endogenous inducers such as hydrogen peroxide or exogenous inducers such as ethanol and tetracycline. The physically inducible promoters regulated by abiotic and external factors such as light, heat and mechanical injury (Rodriquez, 2003). Sang-Hoon *et al*. (2007) reported the oxidative stress-inducible promoter driven copper–zinc SOD (*CuZnSOD*) and *APX* in response to abiotic stresses. The heat shock inducible promoter of small heat shock genes (*sHSPs*) in tomato regulated strong expression of *GUS* reporter gene in roots, leaves, flowers, fruits and germinated seeds after heat treatment (Shu-Ying *et al*., 2006).

2.3 Synthetic Promoters

The synthetic promoter is a set of minimum consensus elements of the promoter regions found in nature. It is possible to synthesize consensus sequences that may work across different organisms and is not necessarily derived from a particular organism. Synthetic promoter was used to manipulation of transcriptional activity that will regulate single or multiple plant transgenes in direct response to specific environmental, physiological and chemical cues. The CaMV 35S core-promoter is ideal for transcription initiation and has been used in several plant promoter engineering strategies. The two most frequently used strategies are, combinatorial engineering of cis-motifs upstream of the core-promoter or combined with bidirectionalization of a unidirectional promoter (Figure 3) (Venter, 2006). Simran *et al.* (2003) revealed Mod2A1T, subdomain A1 of 35S promoter regulating the expression of reporter gene in seeding of the T_1 generation at the same strength to that of 35S promoter. In many plant biotechnological applications, recently, Moore *et al.*

(2006) reviewed using synthetic promoter that induced by chemical reagents. This system has proven highly flexible and could be used to either repress or activate plant transgene expression.

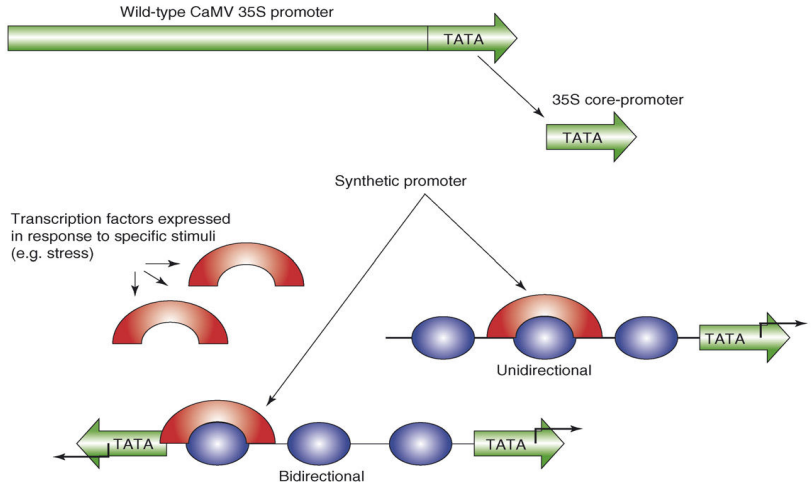


Figure 3 Strategic of a plant synthetic promoter construction. Core-region containing TATA-box of a wild type constitutive promoter, such as CaMV 35S, was used to driven transcription.

Source: Venter (2006)

2.4 Tissue-Specific Promoters

Tissue-specific promoters control direct the expression of genes in specific tissues or particular tissues and at certain developmental stages of plants. Tissue-specific promoters may be induced by endogenous and exogenous factors, so they may be also classified as inducible. The phosphate transporter gene promoter of *Arabidopsis thaliana* (*PHT1*) governed root-specific expression of transgene in rice (Takayashi *et al*., 2005). Tomato *E8, 2A11* and polygalacturonase promoter control the expression of genes in mature ovary tissue of a fruit and in the receptacle tissue of accessory fruits (Van Haaren and Houck, 1993; Speirs *et al*., 1998). A seed specific promoter, Glutelin B1 (*GluB1*) promoter, generate transgenic rice seed-based edible vaccine against house dust mite allergy (Lijun *et al*., 2007).

3. Function of Promoter in Eukaryotes

3.1 Introduction

In eukaryotes, the control of gene expression is complicated. Because multicellular eukaryotes need generate large numbers and types of cells, gene expression control depends on an array of interaction regulatory elements that turn genes on and off in the right places at the right times. The most complex controls observed in eukaryotic genes are those that regulate the expression of RNA pol II-transcribed genes, the mRNA genes. Almost all eukaryotic mRNA genes contain a basic structure consisting of coding exons and non-coding introns and basal promoters of two types and any number of different transcriptional regulatory domains. The basal promoter elements are termed CCAAT-boxes and TATA-boxes because of their sequence motifs (Hartwell *et al*., 2004). The cells of multicellular organism must also regulate their gene expression on a more long term basic. During development of multicellular organism, its cells undergo a process of cell differntiation, in form and function, resulting in several or many differentiated cell types (Mallery, 2001).

3.2 Level of Gene Regulation

The levels of gene expression in eukaryotes include transcriptional control, processing control, transport control, mRNA degradation control, translational control and protein degradation control (Figure 4) (Campbell and Reece, 2004; Mallery, 2001).

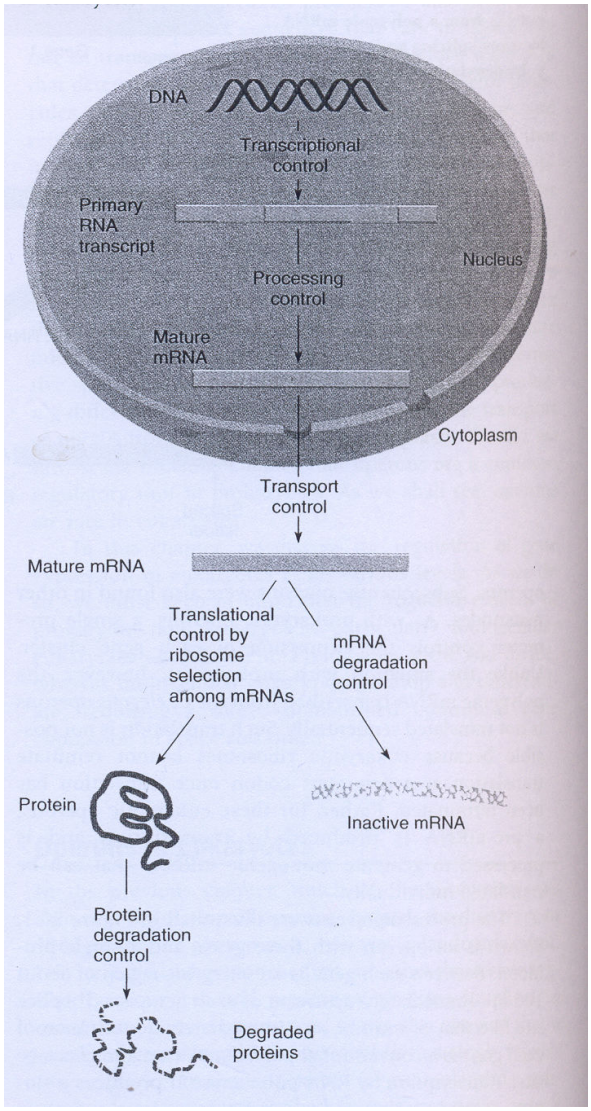


Figure 4 The diagram showing levels at which gene expression can be controlled in eukaryotes.

Source: Russell (2006)

3.2.1 Transcriptional Control

The regulation of genes expression in eukaryotes is mostly at the level of transcription initiation. Iinitiation of the transcription of genes is under the control of the promoter immediately upstream of the gene and of enhancers that are distant from the gene (Lewin, 2006). In general transcription machinery which assembles on the core promoter alone is capable of only a basal level of transcription. Regulation of the transcription up to its maximal level is possible via regulatory proteins binding to promoter proximal elements and to enhancer elements. In addition, chromatin remodeling or DNA methylation and acetylation also cause the inhibition of transcription. Specific factors that control transcription include the strength of promoter elements, the presence or absence of enhancer sequences and the interaction between these sequences and regulatory proteins (Russell, 2006).

3.2.2 RNA Processing Control

After transcription, RNA processing control regulates production of mature RNA molecules from precursor-RNA molecules. There are many cases used to produce different functional mRNAs such as alternative polyadenylation or alternative splicing (Mallery, 2001). Which products are generated depend on regulatory signals. This products are proteins that are encoded by the same gene, but that differ structurally and functionally (King, 2003).

3.2.3 RNA Transport Control

The mRNA transport control is the regulation of the number of transcripts that exit the nucleus to the cytoplasm. The mRNAs must be caped for them to exit through the nuclear pore complexes. In addition, the export process involves proteins that bind to mRNA molecules and interact with proteins at the nuclear pore complexes to direct the movement of the mRNAs to the cytoplasm (Campbell and Reece, 2004). Generally, transportation of mRNAs is a mechanism to achieve an appropriate protein distribution within the cell.

3.2.4 Translational Control

Messenger RNA molecules are subject to translational control by ribosome selection among mRNAs. Differential translation can greatly affect gene expression. The mRNAs are modified at their 3' end by the addition of a poly A-tail that can vary in length from fewer than 20 adenines to more than 200 adenines. Cellular enzymes in the cytoplasm slowly shorten the poly A-tails of all mRNAs after the leave the nucleus. Once the A-tail has disappeared, the core mRNA molecule is quickly degrade (Lewin, 2006).

3.2.5 mRNA Degradation Control

All RNA species are subjected to degradation control, in which the rate of RNA breakdown is regulated. Usually, both rRNA and tRNA are highly stable species whereas mRNA molecules exhibit a diverse range of stability. The half-life of individual mRNAs within a given eukaryotic cell may vary by several orders of magnitude, from a few minutes to many hours. The stability of particular mRNA molecules may change in response to regulatory signals (Campbell and Reece, 2004).

3.2.6 Protein Degradation Control

There are many mechanisms display after protein has been produced. Some proteins are not active when they are first formed. They will undergo modification such as glycosylation, acetylation, acylation, disulfide bond formation (King, 2003). Many of these modifications occur extreme rapidly, active protein will be transported to their site of action. Many proteins are rapidly degraded, whereas others are highly stable. Specific amino acid sequences in some proteins have been shown to bring about rapid degradation (King, 2003).

4. Isolation of Plant Promoter

Promoter isolation is an essential step to understand of regulation of gene expression in plants. Using of the Polymerase Chain Reaction (PCR) to isolate unknown 5'flanking region of known DNA sequence is becoming popular since the method is efficient, fast and there is no need to construct and screen libraries. Different PCR methods, such as inverse PCR (Ochmanet al., 1988), thermal asymmetric interlaced (TAIL)-PCR (Liu and Whittier,1995), single primer amplification (Hermann *et al*., 2000), the biotinylated gene-specific primers and streptavidin-coated magnetic beads (Rosenthal and Jones, 1990) and the ligation-mediated PCR (Rosenthal and Jones, 1990; Doven *et al*., 1995; Siebert *et al*., 1995; Balavoine, 1996; Zhang and Chiang, 1996; Cottage *et al.*, 2001) have been used to isolate the unknown flanking regions.

4.1 Inverse PCR

In this method, genomic DNA sequences were digested with restriction endonuclease. The digested DNA fragments were self-ligated to form circularised chromosomal DNA fragments for the amplification reaction. Two primers complement in each end of the known fragment and pointing away from each other (Ochman *et al*., 1988; Arendse *et al*., 1999). This method eliminates the laborious steps of shotgun cloning, colony screening and culturing of cells. However, as information on restriction sites is usually not available, the success is limited. This method was used to isolation ACC oxidase 5'flanking region from tomato (Blume and Grierson, 1997).

4.2 The Thermal Asymmetric Interlaced (TAIL)-PCR

This method has been successfully used to isolate insert-end segments of P1 and YAC clones (Liu and Whittier, 1995). Furthermore, 5'flanking regions of *Pal* and *Pgi* genes of yams were isolated with modified this technique (Terauchi and Kahl, 2000). This method is simple, but nevertheless efficient for genomic walking which

does not require any restriction or ligation steps. PCR is carried out with long sequence-specific primers in combination with short degenerate primers of arbitrary sequence.

4.3 Single Primer Amplification

This method was used to isolate the 5' regions of genes encoding banana sucrose phosphate synthase (*SPS*), sugarcane actin (*ACT1*), and *Dunaliella tertiolecta* S15 ribosomal protein (*S15*) (Hermann *et al*., 2000). This strategy is a consequence of both specific template capture, purification and the use of a single, nested primer in subsequent amplifications.

4.4 Biotinylated Gene-Specific Primers and Streptavidin-Coated Magnetic Beads

This method was used to separate PCR product of the target sequence from non-target sequences by using a specific primer which had been biotinylated at its 5' end. Isolation of biotin-labelled product was performed using streptavidin coated beads. (Rosenthal and Jones 1990; Hermann *et al*., 2000). This method is simple reliable for promoter isolation. This method was used to isolate 5'flanking regions of gene encoding SPS ACT1 and S15 in banana (Hermann *et al*., 2000).

4.5 Ligation-Mediated PCR

This method was comprised of 3 steps. First, genomic DNA was digested with restriction endonuclease. Secondly, digested DNA fragments were ligated to a DNA cassette or oligonucleotides. Thirdly, the unknown sequence was ampliflied by two rounds of nested PCR with specific primers. Siebert *et al*. (1995) improved the efficiency upon the adaptor ligation method by combining vectorette PCR with a newly developed method suppression PCR. The suppression PCR technology requires an adapter primer which is shorter in length than the adaptor and is capable of hybridizing to the outer primer binding site. Zhang and Chiang (1996) had devised

a single- stranded DNA by T4 RNA ligase method for cloning a 5' non-coding region of phenylalanine ammonia lyase (*Pal*) gene from lobloly pine. Using this technique, less than 2.0 kb of *Pal* gene 5' non-coding region was cloned. Cottage *et al.* (2001) presented a method for identification of DNA sequences flanking T-DNA insertions by digesting genomic DNA with restriction enzymes yielding blunt-ended fragments that are ligated to asymmetric adaptors. Using PCR primers specific to the adaptor sequences combined with PCR primers designed for the T-DNA, it is feasible to amplify unknown genomic regions flanking the T-DNA insertion site.

5. Plant Promoter Analysis by Bioinformatics Approach

Bioinformatics approach is commonly used to identify *cis*-acting elements in plant promoters. Currently, there are a number of databases with information on *cis*-acting elements that control the transcription initiation by binding corresponding nuclear factors. TRANSFAC (Wingerder *et al*., 2001), PLACE (Higo *et al*.,1999), PlantCARE (Lescot *et al*., 2002) and PlantProm DB (Shahmuradov *et al*., 2003) databases are wildly used for searching and *cis*-acting elements analysing in plant promoter.

5.1 TRANSFAC Database

The database comprises of extensive information on transcription factors, their structures, functions and expression patterns. The broad compilation of binding sites allows the derivation of positional weight matrices. The programs, match and patch via matrices and sequence location provide for performing matrix-or pattern-based search of transcription factor binding sites in regulatory DNA sequences. Thus, it is possible to make predictions for most gene promoters, which have not been studied in detail yet. TRANSFAC also includes a tool to automatically visualize gene-regulatory networks being based on interlinked factor and gene entries in the database. The TRANSFAC database is accessible via the http://www.gene-regulation.com/cgi-bin/pub/databases/transfac/search.cgi.

5.2 PLACE Datadase

PLACE is based on an in-house database of plant DNA *cis*-acting elements. These *cis*-acting elements have been extracted from review articles as well as original literatures on genes in vascular plants. In addition to motifs originally reported, their variations in other genes or in other plant species in later reports are also compiled. Documents for each motif in the PLACE database contain, in addition to a motif sequence, a brief definition and description of each motif in EMBL format, and relevant literature with PubMed ID numbers and GenBank accession numbers where available. Users can search their query sequences for *cis*-acting elements using the Signal Scan program at its web site. The results will be reported in one of the three forms including 'grouped by signal', mapped to sequence' and by sequence order. Drs. Ugawa and Higo created this database system in 1999. The database server is maintained in MAFF DNA Bank (Ministry of Agriculture, Forestry and Fisheries) of Japan in the framework of the Research Project for New Recombinant DNA Techniques at National Institute of Agrobiological Sciences in collaboration with Dr. Y. Nagamura and his colleagues. The PLACE database is accessible via the http://www.dna.affrc.go.jp/PLACE/

5.3 PlantCare Database

PlantCARE is a database of plant *cis*-acting regulatory elements, enhancers and repressors. Regulatory elements are represented by positional matrices, consensus sequences and individual sites on particular promoter sequences. These queries result in listing of entries with links to other information within the database or beyond through accession numbers from other databases, such as EMBL, GenBank, TRANSFAC and MEDLINE are provided when available. Data about the transcription sites are extracted mainly from the literature, supplemented with increasing number of in silico predicted data. Apart from a general description for specific transcription factor sites, levels of confidence for the experimental evidence, functional information and the position on the promoter are given as well. Furthermore, database is now provided to a new clustering and motif search method

to investigate clusters of co-expressed genes. The database can be queried on names of transcription factor (TF) sites, motif sequence, function, species, cell type, gene and literature references. The PlantCARE relational database is available via the http://bioinformatics.psb.ugent.be/webtools/plantcare/html/

5.4 Plant Prom DB

Plant Prom DB comprises of a proximal promoter sequences for RNA polymerase II and transcriptional factors with experimentally determined transcription start sites from various plant species. Users can search *cis*-acting elements using the NSITE-PL and TSSP programs at its web site. Currently, this database contains 1434 regulatory element from monocot, dicot and other plants respectively. It is more descriptive information than PLACE, including several fields on expression and some others. This software was developed by Softberry in collaboration with Department of Computer Science, University of London. The Plant Prom DB is accessible via http://www.softberry.com/berry.phtml?topic=nsitep&group=programs&subgroup=promoter

6. The use of Promoters for Plant Improvement

The alteration and/ or improvement of plant phenotypic characteristics for productivity or quality using gene manipulation requires the expression of introduced gene in plant tissues. Therefore, the availability of promoter to drive and control transgene expression is required. Promoters from bacteria, viruses, fungi and plants have been used to control gene expression in transgenic plants (Grierson *et al*., 2001). The length of promoter sequence can be varied from 100 bp-2 kb (Grierson *et al*., 2001). The most well known promoter for driving and control transgene expression is Cauliflower mosaic virus (CaMV) 35S. CaMV is a plant virus infecting several crop plants including turnip, cabbages and cauliflowers. When this virus infects a plant, the genome will move to plant nucleus and viral RNAs, 35S and 19S RNA, were synthesized. These promoter, namely 35S and 19S, functioned in several plant species (Matthews, 1991). These promoters contained TATA box and a core sequence

of promoter. Over the years, more elements such as CCAAT box were found and several versions of these promoters were constructed. The 35S promoter was found to be 10-15 stronger than 19S promoter.

The common version of a 350 bp 35S promoter contains all signals needed for activity including a 42 bp core promoter responsible for basal activity and a variety of upstream elements conferring particular tissue specificities and quantitative determinant. The roles of these elements make 35S promoter efficiently expresses in most plant tissues. Modification of 35S promoter by enhancer duplication results in 2 to 10 fold increase in transcription level. This promoter functions in both dicots and monocots although the activity in monocots is relatively lower (Potrykus *et al.*, 1998).

Recently, many plant promoters have been identified such as Polygacturonase promoter (Montgomery *et al*., 1993), tobacco RD2 gene promoter, beta-amylase gene, barley hordein gene promoters, banana TRX promoter, melon actin promoter ,tomato pz7 and pz130 gene promoters (Rodriquez, 2003), and ACC oxidase promoter from tomato, apple, banana, peach and papaya (Blume and Grierson.,1997; Atkinson *et al*., 1998; Moon and Callahan,2004; Chen *et al*., 2003 and Rodolfo *et al*., 2004).

ACC Oxidase Gene

1. Introduction

ACC oxidase is a key enzyme that converts 1-aminocyclopropane-1-carboxylate (ACC) to ethylene in the final step of ethylene biosynthesis pathway (Figure 5). The level of *ACO* expression involves with including wounding, flower development, seed germination, ethylene regulated events, leaf and flower senescence and fruit ripening (Kumdee *et al*., 2003; Chen *et al*., 2002).

2. *ACO* Gene Molecular Techniques Study

ACO enzyme is encoded by a mutigene family in several plant species (Alexander and Grierson, 2001; Gray *et al*., 1992.). The first *ACO* gene was identified through antisense expression of a clone pTOM13 in tomato. Five members of the tomato *ACO* gene family were characterized (Barry *et al*., 1996; Holdsworth *et al*., 1988; Nakatsuka *et al*., 1998; Sell and Hehl, 2005). Later, *ACO* genes were isolated from petunia, mung bean, broccoli, and sunflower (Tang *et al*., 1993; Kim *et al*., 1994; Pogson *et al*., 1995; and Liu *et al.,* 1997). In papaya, *ACO* genes have been reported in Taiwan (Chen *et al*., 2003), Mexico, Hawaiian (Neupane *et al*., 1998). In Thailand, two *ACO* genes designated *CP-ACO1* and *CP-ACO2* were identified from fruit of *Carica papaya* Khaek Nuan variety (Kumdee *et al*., 2003). The differential expression suggested that *CP-ACO1* is ripening associated whereas *CP-ACO2* might play the role in seed germination process. The level of *CP-ACO1* expression was increased during fruit ripening. This result indicated that *CP-ACO1* might play an important role in ripen process of papaya fruit. Within a fruit, *CP-ACO1* expression was highest in endocarp (Kumdee *et al*., 2003).

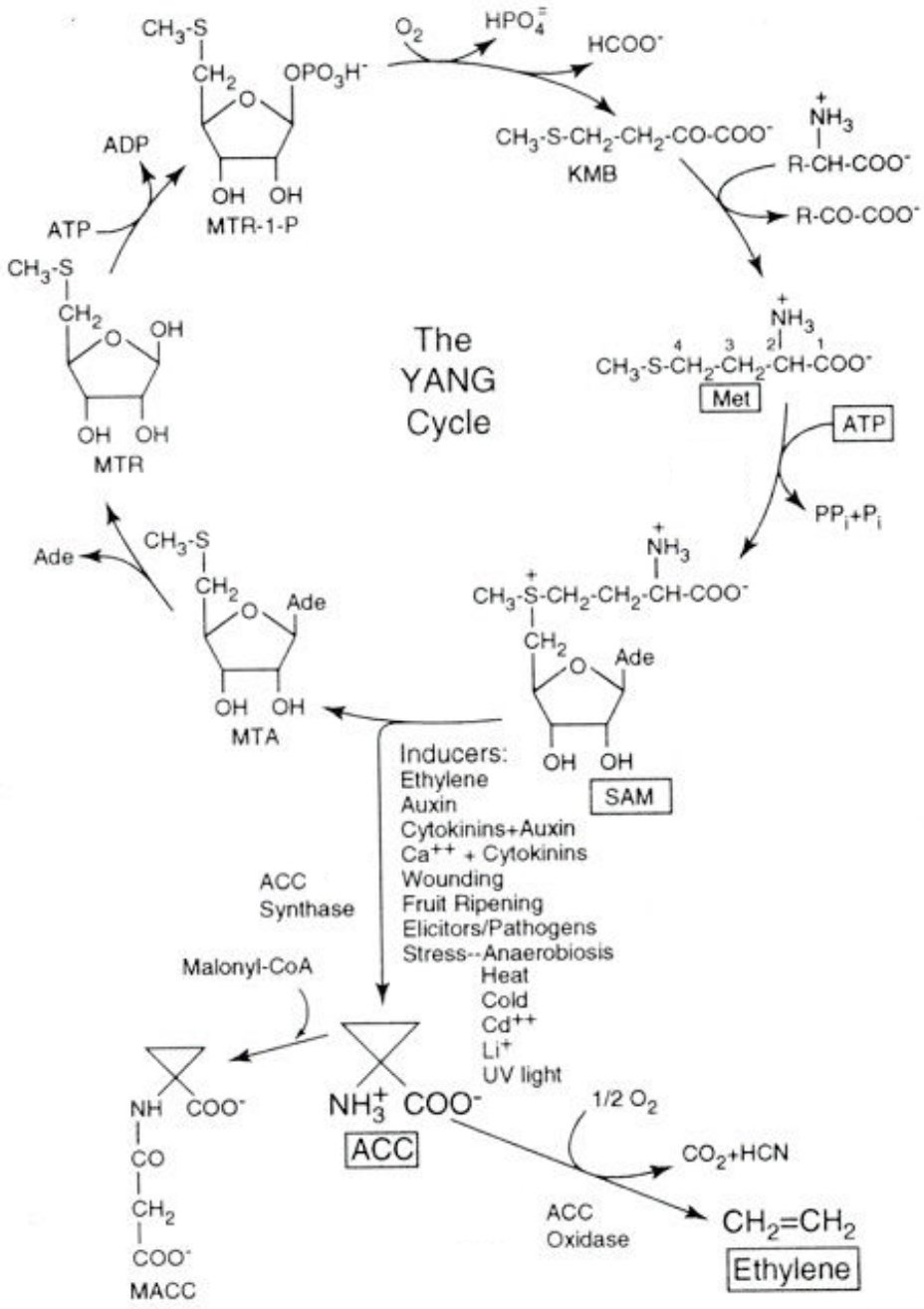


Figure 5 The ethylene biosynthetic pathway of higher plants

Source: Yang and Hoffman (1984)

3. Regulation of *ACO* Gene Expression

ACO gene is regulated by developmental signals and *cis*-acting elements. Blume and Grierson (1997) shown that *ACO* gene regulation occurred at the transcriptional level and developmental and environmental cues.

Blume *et al*. (1997) reported the isolation and sequencing of approximately 2 kb of 5'-flanking sequence of three tomato *ACO* genes *(LEACO1, LEACO2, LEACO3)* and the occurrence of class I and class II mobile element-like insertions in promoter and intron regions of two of them. The nucleotide sequence of all tomato *ACO* promoters (*LEACO1, LEACO2* and *LEACO3*) indicated that the ATG start codon is located between nucleotide 97-99. The putative TATA-box is between nucleotide -29 and -24 (Grierson *et al*., 2001). The *LEACO1* upstream region contains a 420-bp direct repeat which is present in multiple copies in the tomato genome and is very similar to promoter sequences of the tomato *E4* and *2A11* genes. The region covering the repeats resembles the remnant of a retrotransposon. Two copies of a small transposable element, belonging to the *Stowaway* inverted repeat element family, have been found in the 5'flanking sequence and the third intron of *LEACO3* (Blume *et al*., 1997). The *LEACO1* promoter is the strongest *ACO* promoter during ripening and in response to wounding. The *LEACO2* promoter could have utility in specific circumstances or cell types (Grierson *et al*., 2001).

Moreover, promoter sequence of climacteric fruits such as apple, banana, peach and papaya *ACO* genes were reported (Atkinson *et al*., 1998; Moon and Callahan, 2004; Chen *et al*., 2003; Rodolfo *et al*., 2004). In papaya, there are two *ACO* promoters reported in papaya (Chen *et al*., 2003; Rodolfo *et al*., 2004). Chen *et al*. (2003) reported the isolation and sequencing of 600 bp of 5'flanking region of *ACO2* from papaya Tainong2 variety. Several *cis*-acting elements were identified including a putative wound-inducible element (WUN) and three putative ethylene responsive elements (ERES). These *cis*-acting elements suggested that *CP-ACO2* promoter from papaya Tainong variety might be under the regulation of ethylene and wounding. *ACO1* promoter 1330 bp in length from Mexico papaya was identified by

Rodolfo *et al*. (2004). The putative GCC box (TAAAGAGCC) shown to be an ethylene responsive motif that is both necessary and sufficient for the regulation for transcription by ethylene. Furthermore, the sequences TATTTAAT which belongs to the *cis*-regions that have been identified as controlling the spatial or developmental specificity of some gene expression were detected.

4. The Utilization of *ACO* Promoter

The *ACO* promoter can be applied to drive gene expression in transgenic plants in the tissue-specific manner for improving plant quality and quantity. For examples, generating transgenic plants that are delayed ripening, high nutrition and physiological changes such as color, texture, and size.

Active variants of the *ACO* promoter can also be generated. It can be possible to alter the level or type of activity of the *ACO* promoters by manipulating their sequences such as altering the nucleotide sequence in key regulatory regions, truncating the sequence or deleting parts within the sequence. Segments of the oxidase promoter sequences of between 100 and 2000 bases in length may be useful as plant-operative promoters (Grierson *et al*., 2001). In practice, the *ACO* promoter can be inserted as a promoter sequence in a recombinant gene construct destined for use in a plant. This construct can be used in plant transformation. Any plant species may be transformed with the construct, and any suitable transformation method may be employed (Grierson *et al*., 2001).

MATERIALS AND METHODS

This study was comprised of 3 parts including 1) Isolation of *CP-ACO II* 5'flanking region 2) Identification of *cis*-acting element by database–assisted bioinformatics approache 3) Determination of promoter activity of *CP-ACO I* and *II* fragment DNA via *Agrobacterium*-mediated infiltration transient expression assay.

1. Isolation of *CP- ACO II* 5'Flanking Region

The *CP-ACO II* 5'flanking region was isolated by ligation-mediated PCR method. This sequence amplified with gene specific primer and adapter primer. Flow diagram of the strategy used for isolation was shown in Figure 6.

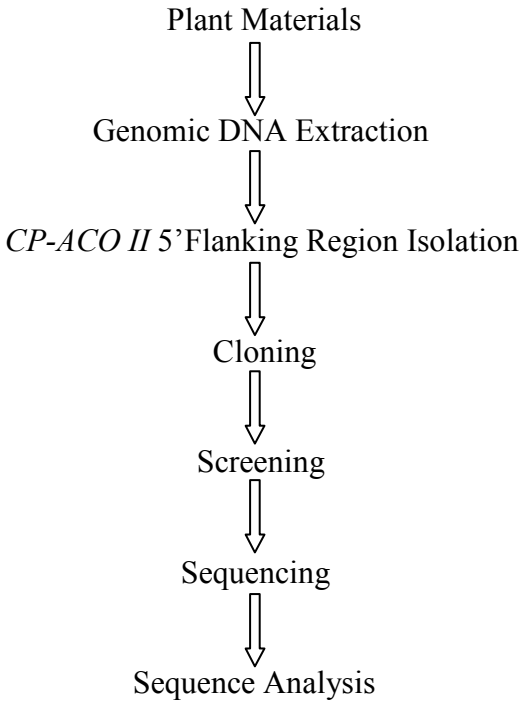


Figure 6 Flow diagram of the strategy for the isolation of *CP-ACO II* 5'flanking region

1.1 Plant Materials

Papaya (*Carica papaya*) Khaek Nuan variety was grown on the plot of The Central Laboratory and Greenhouse Complex, Kasetsart University Kamphaeng Sean Campus, Nakhon Pathom. Young leaves were collected, washed with distilled water, 70% ethanol and allowed to air dry. Then, the leaves were frozen with liquid nitrogen and kept at -80 °C until use.

1.2 Genomic DNA Extraction

Papaya genomic DNA extraction was modified from CTAB method (Warner, 1996). Papaya young leaves were ground into fine powder after freezing in liquid nitrogen. Four hundred microliters of CTAB buffer (2% CTAB, 1.4 M NaCl, 100 mM Tris-HCl, 200 mM EDTA, pH 8.0) were added. The suspension was mixed by vortexing and incubated for 30 min at 65 °C. The content was separated by centrifugation (Eppendorf, Germany) at 12,000 rpm for 10 min at 4 °C. The pellet was discarded and the supernatant was extracted with an equal volume of phenol: chloroform: isoamyl alcohol (25: 24: 1) and mixed by inverting for 2 min. The two phases were separated by centrifugation at 12,000 rpm for 10 min at 4 °C and upper phase was transferred to a new microcentrifuge tube. An equal volume of absolute ethanol and 0.1 volume of 3 M sodium acetate (pH 5.4) was added and mixed by inverting. The content was incubated for 15 min at -20 °C. The DNA was precipitated by centrifugation at 12,000 rpm for 20 min at 4 °C. The pellet was washed with 70% ethanol, dried under room temperature (RT) and resuspended in 30 µl of dH_2O containing 20µg/µl of RNase A. The DNA was incubated for 15 min at 55 °C to digest the RNA and stored at -20 °C for long term or 4 °C for short term use. The concentration and purity of DNA were determined using spectrophotometer (Amersham Biosciences, UK) and agarose gel electrophoresis (Bio-Rad, U.S.A.).

1.3 *CP-ACO II* 5'Flanking Region Isolation

1.3.1 Primer Design

The primers used for amplification were shown in Table 1. The location of primer were shown in Appendix Figure 2. TAKIIN was designed using sequence information from *CP-ACO II* gene papaya Kheak Nuan (Chuaboonmee, 2004). Adapter primer (API) and nested adapter primer (NAPI) were adapted from Cottage, *et al.* (2001). Reverse primers comprised StartUTR, MLP, RTATA and MLPN. Forward primers comprised FSPI and FSPNI.

Table 1 List of primers using in PCR for *CP-ACO II* 5'flanking region isolation

Primer names	**Sequence (5'-3')**
Reverse primers	
TAKIIN	ggt gag acc ctc cat gtt g
StartUTR	tgc aga aag aat ttc tcg gg
MLP	atg aat gca gac tcg aga g
RTATA	atg aat gca gac tcg aga ga
MLPN	gat cca gaa tga aga cag
Forward primers	
API	gga tcc taa tac gac tca cta tag ggc
NAPI	tat agg gct cga gcg gc
FSPI	gat tac tac att ttt aat ttg c
FSPNI	agt gtc tgc gtc ctt cat

1.3.2 Strategy for Ligation-Mediated PCR

This strategy was comprised of 3 steps 1) restriction digestion, 2) adaptor annealing and ligation and 3) PCR amplification were shown in Figure 7.

a. Restriction Digestion

Ten micrograms of genomic DNA was restriction digested overnight at 37 °C, with 50 units of each enzyme including *Ssp*I, *Dra*I, *Hpa*I, *BspLI*, *KspAI*, *PdmI*, *RsaI* and *PvuII* (Fermentas, Canana) in a final volume of 100 µl. The enzymes were heat inactivated for 10 min at 65 °C. The digested DNA was precipitated by adding of 0.1 volume of 3 M sodium acetate (pH 5.4) and 2 volumes of absolute ethanol. The mixture was vortexed, and centrifuged at 12,000 rpm for 15 min. The pellet was washed with 70% ethanol, centrifuged at 12,000 rpm for 10 min, air dried and resuspended in 20 µl of dH_2O.

b. Adapter Annealing and Ligation

Oligonucleotides ADAPL (5’ cta ata cga ctc act ata ggg ctc gag cgg ccg ccc ggg cag gt 3’) and ADSPS (5’ P-acc tgc cc-H_2N 3’) were used as adaptor following Cottage *et al*. (2001). The nucleotides were resuspended in dH_2O at a concentration of 100 pmol/ µl. A volume of 20 µl of each adaptor was pipetted into a 0.5 ml microcentrifuge tube and overlaid with mineral oil. The adaptors were heated for 2 min at 100 °C and the solution was allowed to cool for 1 h at RT. The annealed adaptors were decanted from under oil and stored at -20 °C. Then, 10 µl of the digested genomic restriction was ligated to 1 µl of the annealed adaptors with 2 µl of T4 DNA ligase (3 units/ml; Fermentas). The ligation was incubated overnight at 16 °C and heat inactivated for 10 min at 65 °C, this ligation mix was called the adaptor library.

c. PCR Amplification

CP-ACO II 5'flanking region was isolated from the adapter library (*Ssp*I, *Dra*I, *Hpa*I, *BspLI*, *KspAI*, *PdmI*, *RsaI* and *PvuII*) by ligation-mediated amplification. The first PCR reaction mixture consisted of 5 µl adaptor library, 5 µl 10X PCR buffer (10 mM Tris, pH 8.8, 50 mM KCl and 0.1% triton X-100), 2 mM $MgCl_2$, 2 mM of dNTPs and 1U of *Taq* polymerase (Fermentas), 0.6 pmol/µl of the reverse *CP-ACO II* gene specific primer TAKIIN and the adaptor primer AP1, was adjusted to a final volume of 50 µl with dH_2O. The cycle of the amplification was as followed; 94 °C 3 min 1 cycle and 94 °C 30 sec, 55 °C 40 sec and 72 °C 3 min for 35 cycles and followed by 1 cycle of 72 °C 10 min. The first undiluted PCR products and the serially diluted product to 1:100 and 1:1,000 µl with dH_2O were used as templates for the nested PCR. The second PCR reaction mixture consisted of 10 µl diluted first PCR products, 5 µl 10X PCR buffer (10 mM Tris, pH 8.8, 50 mM KCl and 0.1% triton X-100), 2 mM $MgCl_2$, 2 mM of dNTPs and 1U of *Taq* polymerase (Fermentas), 0.6 pmol/µl of the reverse *CP-ACO II* gene specific primer TAKIIN and the nested adaptor primer NAP1, was adjusted to a final volume of 50 µl with dH_2O. The cycle of the amplification was as followed; 94 °C 3 min 1 cycle and 94 °C 30 sec, 52 °C 40 sec and 72 °C 3 min for 35 cycles and followed by 1 cycle of 72 °C 10 min. The second PCR products were determined by agarose gel electrophoresis.

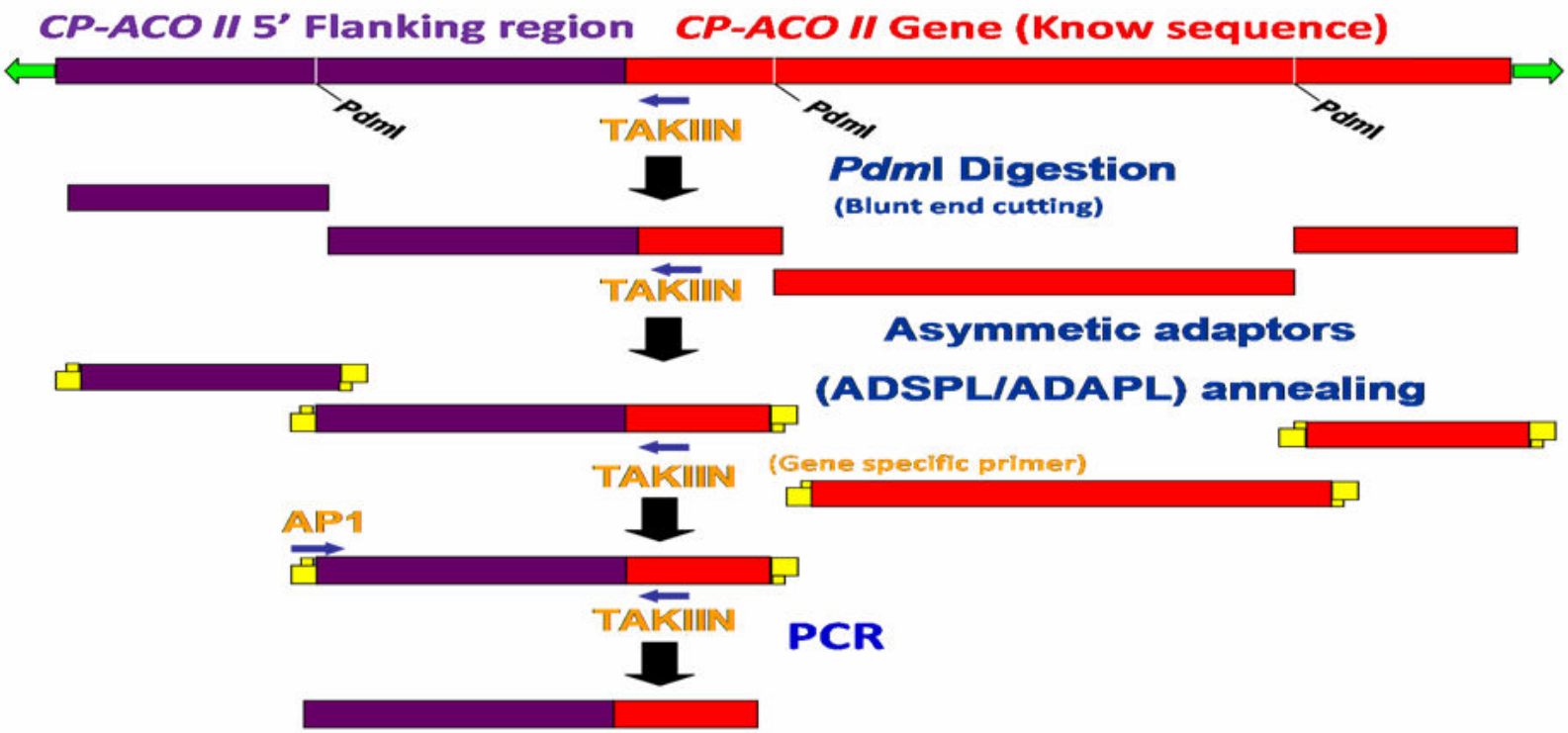


Figure 7 Strategy of ligation-mediated PCR for *CP-ACO II* 5'flanking region isolation. Plant genomic DNA was digested with blunt-ended cutting enzymes (*Pdm*I). ADSPL/ADAPL adaptors were annealed to the blunt-ended fragments. An adaptor primer AP1 and *CP-ACO II* 5'flanking reverse specific primer TAKIIN were used to amplify *CP-ACO II* 5'flanking region

1.4 Cloning

The second PCR products were adjusted to a final volume of 100 µl with dH_2O. An equal volume of absolute ethanol and 0.1 volume of 3 M sodium acetate (pH 5.4) was added and the solution was mixed by inverting. The content was incubated for 15 min at -20 °C. The second PCR products were precipitated by centrifugation at 12,000 rpm for 20 min at 4 °C. The pellet was washed with 70% ethanol, dried under RT and resuspended in 10 µl of dH_2O. The PCR products were determined using spectrophotometer and agarose gel electrophoresis. Expected *CP-ACO II* 5'flanking region DNA was ligated into pDrive vector (Qiagen) essentially as recommended by the manufacturer. The reaction mix contained 5µl of rapid ligation buffer (2X), 1 µl of pDrive vector (50 ng/µl), 1-4 µl PCR product. The reaction mix was made up to 10 µl with dH_2O, spun down and incubated at 16 ^{o}C overnight.

Two µl of the ligation mix was added to 100 µl aliquots of *Escherichia coli* strain DH5α competent cells and the mixture was incubated on ice for 30 min. The competent cells were then heat shocked for 90 sec at 42 °C and placed on ice for 2 min. The competent cells were resuspended in 800 µl of 2xYT medium (16 g/l bacto- tryptone, 10 g/l bacto-yeast extract and 5g/l NaCl) and incubated in a shaking incubator for 60 min at 37 °C. The dilutions were plated out on 2xYT agar (16 g/l bacto-tryptone, 10 g/l bacto-yeast extract, 5g/l NaCl and 1.5% bacto agar) containing IPTG (29 µg/ml), X-gal (0.06 % in dimethyl formamide) and amplicilin (50 µg/ml) and incubated overnight at 37 °C. The recombinant clones were identified by blue/white colony selection and used to incubate overnight on 2xYT medium at 37 °C, containing ampicillin (50 µg/ml).

1.5 Screening

The recombinant clones were grown overnight at 37 °C with vigorous shaking in 3 ml of 2xYT containing 50 μg/ml of ampicillin. Recombinant plasmids were extracted by a modification method of Sambrook *et al*., (1989). A 1.5 ml aliquot of cell suspension was centrifuged at 12,000 rpm for 3 min and the supernatant was discarded. The pellet was resuspended in 100 μl of chilled solution I (50 mM glucose, 25 mM Tris-HCl, pH 8.0, 10 mM EDTA, pH 8.0) mixed by vortexing and incubated on ice for 5 min. Two hundred microliters of freshly prepared solution II (200 mN NaOH, 1% SDS) was added, and the content mixed by inverting rapidly for 10-15 times. The solution was then neutralized by adding 150 μl of ice-cold solution III (3 M potassium acetate, 5 M glacial acetic acid) mixed by vortexing for 1 min and placed on ice for 5 min. The content was centrifuged at 12,000 rpm for 10 min and the supernatant was transferred into a new tube and extracted with an equal volume of chloroform. The recombinant plasmids DNA was precipitated with 2 volume of chilled absolute ethanol and centrifuged at 12,000 rpm for 10 min. The pellet was washed with 70% ethanol and centrifuged at 12,000 rpm for 5 min. The pellet was air-dried, resuspended in 20μl with dH_2O containing 10 μg/ml RNaseA and incubated for 30 min at 55 °C and stored at -20 °C.

1.6 Sequencing

The recombinant plasmids were prepared for automated sequencing by mixing with 40 μl of 13% PEG (v/v) and 8 μl of 4 M NaCl. The mixture was vortexed and incubated on ice for 20 min. DNA was pelleted by centrifugation for 15 min and then washed with 500 μl of 70% ethanol at RT for 10 min. The pellet was air-dried, resuspended in 20 μl of dH_2O and stored at -20 °C. The prepared recombinant plasmid DNA (100-500 ng) was sequenced using an ABI Prism® BigDye™ Terminator Cycle Sequencing Ready Reaction Kits (PE Biosystems) as described by the manufacturer. The sequencing reactions was contained 4 μl of terminator mix, 3.2 pmol of primer, 0.8 μl of 10X PCR buffer (Qiagen) and 100-500 ng of recombinant

plasmid DNA. The reaction mix was made up to 20 µl with dH_2O. Perkin Elmer Cetus of Hybaid thermal cyclers was used for PCR. The cycle of PCR reaction was as followed; 96 °C 30 sec, 50 °C 15 sec and 60 °C 4 min for 35 cycles. The unicorporated dye terminators from sequencing reactions were removed by using DyeEx™ Kits (Qiagen) as described by the manufacturer. The spin columns were prepared before used by vortexing to resuspend the resin, loosened the cap, snaped off the bottom closure and placed in a 2 ml collection tube before being centrifugation at 3,000 rpm for 3 min. The spin columns were transferred to a 1.5 ml microcentrifuge tubes and the sequencing reaction was applied to the gel bed before centrifugation at 3000 rpm for 3 min. The eluted DNA (20 µl) was mixed with 60 µl of isopropanol, 1 µl of glycogen (20 ng/ml) and 19 µl of dH_2O. The mixture was vortexed and incubated at RT for 15 min before centrifugation at 14,000 rpm for 20 min. The pellet DNA was washed by adding 250 µl of 70% ethanol and centrifuged at 14,000 rpm for 5 min. The pellet DNA was allowed to air-dry, resuspended in dH_2O and used for the determination of nucleotide sequences by using the DNA automated sequencer.

1.7 Sequence Analysis

Seqman program (Lasergene) was used to assemble *CP-ACO II* 5'flanking region sequences. The sequences were compared to their in GenBank database using BLASTN program (NCBI). Comparison of all nucleotide sequences was done using the MegAlign program (Lasergene).

2. Identification of *Cis*-Acting Element by Database-Assisted Bioinformatics Approach

Cis-acting elements of *CP-ACO I* and *II* 5'flanking regions were analyzed via 3 databases including PLACE, PlantCARE and Plant Prom DB. Flow diagram of this strategy was shown in Figure 8.

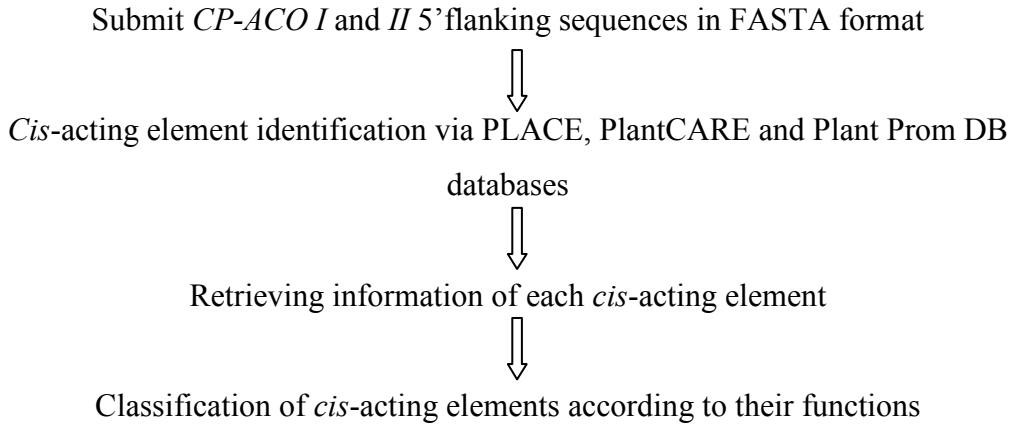


Figure 8 Flow diagram showing the strategy for the analysis of *cis*-acting elements by database-assisted bioinformatics approach

3. Determination of Promoter Activity of *CP-ACO I* and *II* Fragment DNA via Agrobacterium-Mediated Infiltration Transient Expression Assay.

The ability and efficiency of *CP-ACO I* and *II* 5'flanking DNA to regulate gene expression was determined using Agrobacterium-mediated transient expression assay. The strategy of this assay was summarized in Figure 9.

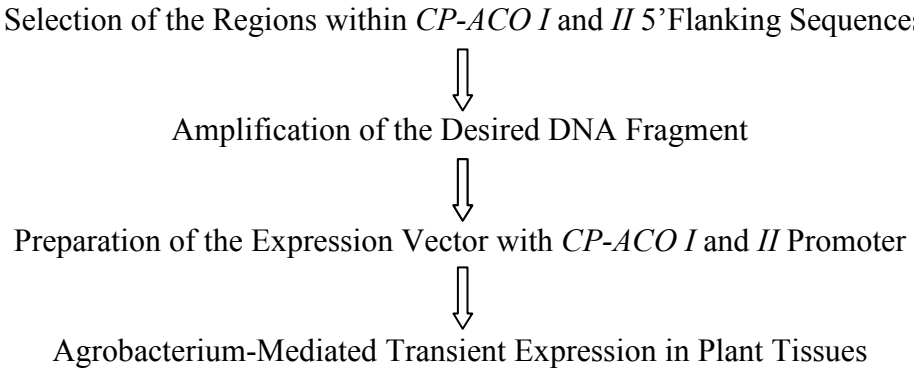


Figure 9 Flow diagram of Agrobacterium-mediated transient expression

3.1 Selection of the Regions within *CP-ACO I* and *II* 5'flanking Sequences

Identified *cis*-acting elements from database-assisted bioinformatics approach were mapped into *CP-ACO I* and *II* 5'flanking sequences. Locations of populated and absent specific elements were marked and used to design DNA regions for promoter study via Agrobacterium-mediated transient expression assay.

3.2 Amplification of the Desired DNA Fragment

3.2.1 *CP-ACO I* 5'Flanking Region Preparation

Three fragments of *CP-ACO I* 5'flanking regions were generated for expression vector construction including, *Carica papaya ACC oxidase I* short promoter (CPACOI-SPP) size 644 bp, *C. papaya ACC oxidase I* long promoter (CPACOI-LPP) with the size 1044 bp and *C. papaya ACC oxidase I* promoter that

consist of repeated DOFCOREZM *cis*-acting elements (CPACOI-DOFCORN) ith the size 579 bp. These fragments were generated by PCR method with *CP-ACO I* 5'flanking region specific primers (Chuaboonmee, 2004). The reaction mix was composed of 1 µl of genomic DNA (1µg/µl), 5 µl of 10X PCR buffer (10 mM Tris, pH8.8, 50 mM KCl and 0.1 % triton X-100), 4 µl of 25 mM $MgCl_2$, 1 µl of 10 mM dNTPs, 1 µl of 30 pmol *CP-ACO I* 5'flanking region specific primers (Appendix Figure 1) 1U of *Taq* polymerase (Fermentas) and adjusted to a final volume of 50 µl with dH_2O. The cycle of PCR reaction was as followed; 94 ^oC 3 min 1 cycle and 94 ^oC 30 sec, 52 ^oC 30 sec and 72 ^oC 1:30 min for 35 cycles and followed by 1 cycle of 72 ^oC 10 min. PCR products were checked by agarose gel electrophoresis. PCR products were purified and resuspened 10 µl with dH_2O.

3.2.2 *CP-ACO II* 5'Flanking Region Preparation

Four fragments of *CP-ACO II* 5'flanking regions were generated for expression vector construction including, *C. papaya ACC oxidase II* short promoter I (CPACOII-SP1) size 594 bp, *C. papaya ACC oxidase II* short promoter lacking seed and endosperm *cis*-acting elements (CPACOI-NO-SEboxN) size 198 bp, *C. papaya ACC oxidase II* short promoter III (CPACOII-SP3) size 283 bp and *C. papaya ACC oxidase II* short promoter IV (CPACOII-SP4) size 285 bp. These fragments were generated by PCR method with *CP-ACO II* 5'flanking region specific primers. The reaction mix was identical to that of *CP-ACO I* 5'flanking fragment amplification except the use of *CP-ACO II* 5'flanking region specific primers (Appendix Figure 2). The cycle of PCR reaction was as followed; 94 ^oC 3 min for 1 cycle and 94 ^oC 30 sec, 42 ^oC 30 sec and 72 ^oC 30 sec for 35 cycles and followed by 1 cycle of 72 ^oC 10 min. For CPACOII-SP1 and CPACOII-SP4, annealing and extension steps were changed to 51 ^oC 30 sec, 72 ^oC 45 sec and 65 ^oC 30 sec, 72 ^oC 10 min, respectively.

3.2.3 Ligation of *CP-ACOs* Promoter into pGEM-T Easy Vector

Seven fragments of *CP-ACO I* and *II* 5'flanking region were ligated into vector pGEM-T Easy (Promega) (Appendix Figure 3) essentially as recommended by the manufacturer. The reaction mix contained 5µl of 2 X rapid ligation buffer, 0.5 µl of 50 ng/µl pGEM-T Easy vector, 1-4 µl of DNA fragments and 1 µl of T4 DNA ligase (3U/µl). The reaction mix was made up to 10 µl with dH_2O, spined down and incubated at 16 ^oC overnight. The recombinant plasmid was transformed into *E coli.* DH5α competent cells. The recombinant clones were identified by blue/white colony and recombinant plasmids were confirmed by PCR and double digestion. The PCR product and digestion products were determined for the correct size by agarose gel electrophoresis.

Plasmids containing *CP-ACO I* and *II* 5'flanking regions were digested with *Nco*I. The reaction contained 10 µg of each recombinant plasmid, 10 µl of 10 X buffer Tango, 2 µl of *Nco*I (Fermentas) (10U/µl) and was adjusted to a final volume of 100 µl with dH_2O. The reaction was incubated at 37 ^oC overnight and heated at 65 ^oC for 10 min to inactivate the reaction. Digested recombinant plasmids were purified by adding an equal volume of absolute ethanol and 0.1 volume of 3 M sodium acetate (pH 5.4) then mixed by inverting. The content was incubated for 15 min at -20 °C. Digested recombinant plasmids were precipitated by centrifugation at 12,000 rpm for 20 min at 4 °C. The pellet was washed with 70% ethanol, dried under RT and resuspended in 20 µl of dH_2O. The digestion mix was incubated at RT for 10 min. Then, the *Nco*I digested recombinant plasmids were used for secondary digestion. The reaction contained 20 µl of *NcoI* digestion mix, 10 µl of 10X buffer O, 2 µl of *Pst*I (Fermentas) (10U/µl) and was adjusted to final volume of 100 µl with dH_2O. The reaction was incubated at 37 ^oC overnight and heated at 65 ^oC for 10 min to inactivate the reaction. The double digestion mix was precipitated and resuspended in 10 µl of dH_2O.

3.3 Preparation of the Expression Vector with *CP-ACO I* and *II* Promoter

3.3.1 Ligation of *CP-ACO I* and *II* 5'Flanking Fragment into pCAMBIA Expression Vector

pCAMBIA1304 (CAMBIA, Australia) expression vector containing 35S driving *gfp* and *gus* reporter genes was used to construct a new expression vector (Appendix Figure 8). CaMV 35S promoter was removed through double restriction digestion. The reaction contained 10 μg of pCAMBIA1304, 10 μl of 10X buffer Tango , 2 μl of *Nco*I (Fermentas) (10 U/μl) and adjusted final volume to 100 μl with dH_2O. The reaction was incubated at 37 ^oC overnight and heated at 65 ^oC for 10 min to inactivate the reaction. The digested vector pCAMBIA1304 was precipitated and resuspended in 20 μl of dH_2O. The digestion mix was incubated at RT for 10 min. Then, the *Nco*I digested vector pCAMBIA1304 mix was used for secondary digestion. The reaction contained 20 μl of *NcoI* digestion mix, 10 μl of 10X buffer 0.2 μl of *Pst*I (Fermentas) (10 U/μl) and was adjusted to a final volume of 100 μl with dH_2O. The reaction was incubated at 37 ^oC overnight and heated at 65 ^oC for 10 min to inactivate the reaction. The double digestion mix was precipitated and resuspended in 10 μl of dH_2O.

Seven fragments of *CP-ACO I* and *II* 5'flanking regions were ligated into 35S removed vector pCAMBIA1304. The ligation mix contained 0.5 μl of 35S removed pCAMBIA1304 (10 ng/μl), 2 μl of 10X reaction buffer, 1-4 μl of digested recombinant plasmids, and 0.5 μl of T4 DNA ligase (4 U/μl) (Fermentas) and was adjusted to a final volume of 20 μl with dH_2O. The ligation mix were incubated at 16 ^oC overnight and heated at 65 ^oC for 10 min to inactivate the reaction. The ligation mix was used directly for transformation into *E. coli* competent cells of DH5α (Appendix B). The positive clones were confirmed by PCR.

3.3.2 Agrobacterium Transformation

One microliter of recombinant plasmid pCAMBIA1304 containing *CP-ACO I* and *II* 5'flanking fagment fused to *gfp* and *gus* reporter genes were added to 100 µl aliquots of *A. tumefaciens* strain AGL1 competent cells. The mixture was added into a chilled cuvette and incubated on ice for 30 min. Competent cells were transformed essentially as described by the manufacturer (Bio-Rad) under condition; cuvette gap 0.1 cm, voltage 2.5 kV, capacitor 25 µF, resistor 400 Ω (pulse controller) and time constant 8-9 msec. After pulsing, the cells were resuspended in 1 ml of 2xYT medium and incubated in a shaking incubator at 28 ^{o}C for 60 min. The dilutions were plated out on 2xYT agar containing kanamycin (25 µg/ml) and incubated 48 h at 28 ^{o}C. Recombinant clones were identified by the ability to grow on 2xYT medium containing kanamycin (25 µg/ml). The recombinant plasmids were purified and confirmed by PCR method.

3.3.3 Agrobacterium Suspension Preparation

Agrobacterium suspension preparation was modified from Jing *et al*. (2006). Culture of *A. tumefaciens* containing *CP-ACO I* and *CP-ACO II* driven construct was grown in 10 ml 2xYT medium (16 g/l bacto-tryptone, 10 g/l bacto-yeast extract and 5g/l NaCl) supplemented with 100 mg/l kanamycin and 200 µmol/l acetosyringone at 28 °C for 12 h for an ultimate OD_{600} of 0.8. It was centrifuged for 3 min at 8000 rpm and then the pellet was resuspended in a new 10 ml standard medium MS (MS salts, 10 mmol/l, MES pH 5.6, 20 g/l sucrose) consisting of 200 µmol/l acetosyringone and kept at 25 °C for 1 h and used for infiltration. In case of transient expression in flower, Agrobarcterium suspension was resuspend with 1M glucose (Yongjin *et al*., 2007).

3.4 Agrobacterium-Mediated Transient Expression

3.4.1 Roots

Mung bean (*Vigna radiata*) seeds were surface-steriled with 20**%** Chlorox for 15 min and rinsed by sterilized water. Fifteen mung bean seeds were sown in a bottle (7 cm X 10 cm) containing a sterilized tissue paper and sterilized water and kept in a dark room at 28 ^{o}C for four days until primary roots were approximately 3-4 cm in length. The Agrobacterium suspension was poured into the bottles until mung bean roots were submerged. The plants were incubated at 25 °C for 2 days (light for 16 h/day).

3.4.2 Leaves

Arabidopsis thaliana (ecotype Columbia) were grown under natural light in air condition glasshouse, at Kasetsart University Kamphaeng Sean Campus, Nakhon Pathom. Their expanding leaves with the midribs 2.5-3.0 cm in length from 3-to 4-week-old seeding were collected, washed with distilled water, 70% ethanol and were transformed immediately. The leaves were prepared by removing the petiole and making three partially transverse and equidistant cut through the midrib from the petiole end to the leaf tip. The three-cut leaves were submerged in Agrobacterium suspension and apply 1 mbar vacuum for 30 min. The vacuum pump was turned off. The pressure released rapidly and the leaves were rinsed in sterile water three times and put facing up on a wet Whatman paper in a Petridish and incubated at 25 ^{o}C under 16 h photoperiod of artificial light for 2 days.

3.4.3 Flowers

White petunia plants (*Petunia hybrida*, White Wave) were maintained at Screen house, Plant Research Group, Kasetsart University Kamphaeng Sean Campus, Nakhon Pathom in room temperature under natural light. Flowers were cut into half horizontally by surgical knife. The flowers were incubated in

Agrobacterium suspension and apply 1 mbar vacuum for 10 min dried with Whatman paper and kept in Petridishes containing moistened Whatman paper. The agroinfiltrated detached flowers were cultured at 25 °C under artificial lights (16 h photoperiod) for 2 to 2.5 days before examining GUS staining.

3.4.4 Fruits

Tomato (*Lycopersicon esculentum*) Seedathip 4 variety was kindly provided by Tropical Vegetable Research Center (TVRC), Kasetsart University Kamphaeng Sean Campus, Nakhon Pathom. Four developmental stages; immature, mature, color break and ripen were used for this experiments. The experiment was done in triplicate. Fruits were washed with distilled water, 70% ethanol, air dried and kept at 4 ^{o}C until use. Tomato fruits were infiltrated using a 1-ml syringe. *Agrobacterium* containing expression vector was gently injected into the fruit tissue through the stylar apex. Total volume of solution was 100 µl. Excess bacteria was removed and the injected point was covered with sterile tape. Only completely infiltrated fruits were used in the experiments. The fruits were kept at 25 ^{o}C for 2 days with 16 h protoperiod.

4. Determination of Promoter Activity of *CP-ACO I* and *II* Fragment DNA via *GUS* Histochemical assay.

4.1 Roots

The submerged mung bean roots were washed and immediately sectioned using a razor blade. Three roots were immersed in GUS staining solution [5-bromo-4-chloro-indolyl-b-D-glucuronide (0.05%) in 100 mM sodium phosphate buffer (NaH_2PO_4 and Na_2HPO_4) pH7.0, 0.1% Triton-X, 0.5mM $K_4Fe(CN)_6H_2O$ and 0.5 mM $K_3Fe(CN)_6H_2O$] and allowed to stain for 16 h in the dark at room temperature. The reaction was stopped by removal of GUS staining buffer and the addition of 70% ethanol. (Jefferson *et al*., 1987).

4.2 Leaves

The transformed leaves were placed in GUS staining solution and applied 1 mbar vacuum for 30 min for substrate distribution and followed by incubation in the dark at 37 ^{o}C overnight. Chlorophyll was bleached out from the leaves by ethanol rinses.

4.3 Flowers

Flower samples were incubated in the dark for 12 h at 37 °C in GUS staining solution and then placed in 70% (v/v) ethanol to remove the chlorophyll and preserve the sample.

4.4 Fruits

The injected tomato fruits were sliced horizontally. Only the middle sections were used for GUS staining. The *GUS* histochemical assay was performed by immersing tissues in the GUS staining solution for 2 h at 37 ^{o}C in the dark. Following staining, the samples were fixed in 70% (v/v) ethanol.

5. Determination of Location of *GUS* Expression in Transformed Mung Bean Roots via Agarose-Embedding Technique

One centimetre of *GUS* stainted root samples were submerged in molten 5 % agarose in 15 mm X 15 mm metal base molds (MICROM, Denmark). After the agarose solidified, each embedding ring was trimmed to give a trapezoid shape. The sections were cut with a microtome to 100 µM thick (Chen-Yi *et al*., 1998). The root section was immediately studied using light microscope at 60X magnification. The outcome was reported by photographs using digital camera (Olympus, Japan).

6. Determination of Promoter Activity of *CP-ACO I* and *II* Fragment DNA via *GUS* Enzymatic Assay

6.1 Plant and Sample Preparation

Transformed mung bean roots were washed with dH_2O. Three roots were ground in 5 volume (w/v) of cold *GUS* extraction buffer [50 mM $NaPO_4$, pH 7.0, 1 mM Na_2EDTA, 0.1% Triton X-100, 10 mM dithiothreitol (DDT) and 0.1% Sodium Lauryl Sarcosine] and mixed by vortexing for 5 min. The content was separated by centrifugation at 12,000 rpm at 4 ^oC for 10 min and the supernatant was transfered into a new microcentrifuge tube and kept at 4 ^oC until use. The total protein was determined using polyacrylamide gel electrophoresis. Quantitative *GUS* assays were performed as described by Jefferson *et al*. (1987).

6.2 Protein Standard Curve

One microlitre of total protein was diluted by using 199 µl of dH_2O. The standard curve preparation using a serial dilution series (0.1-1.0 mg/ml) of Bovine Serum Albumin (BSA) (Table 2). One microlitre of Coomassie Brilliant Blue solution [0.25 g of Coomassie Brilliant Blue R250, 90 ml of methanol:H_2O(1:1 v/v) and 10 ml of glacial acetic acid] was added into each tube of the BSA standard and samples, mixed well and incubated for 5 min at RT before reading the absorbance at 595 nm. The standard curve was prepared by plotting the absorbance against the concentration of standard BSA (µg). Total protein concentration for each sample was determined by comparing with BSA standard curve.

Table 2 Protein standard curve preparation using dilution of Bovine Serum Albumin (BSA)

Protein standard (µg)	**BSA solution stock (0.1 mg/ml)**	**Distilled water (µl)**
0	0	200.00
1.25	12.50	187.50
2.50	25.00	175.00
5.00	50.00	150.00
7.50	75.00	125.00
10.00	100.00	100.00

6.3 SDS-PAGE

Total protein in each sample was adjusted to 50 µg in total volume 25 µl and mixed with an equal volume of 2X sample buffer [0.125 M TrispH6.8, 4% SDS, 0.02% Bromophenal blue, 16% glycerol and 5% mercapto ethanol (2X)]. The samples were boiled for 5 min and centrifuged for 2 min. Proteins were electrophoresed in discontinuous gels consisting of 12% separating gel with a 5% stacking (Appendix Table 2 and 3) using Tris-glycine buffer. A miniprotein apparatus (Bio-Rad) with the comb thickness of 0.75 mm was used. The gel was run at 50V until the dye was migrated out of stacking gel. The voltage was increased to 100 V. The run was terminated when the dye front reached the bottom of the gel.

6.4 Silver Staining

Silver staining was determined with a PlusOne^{TM} silver strain kit, supplied by Amersham Biosciences. The SDS-PAGE gel from 6.3 was submerged in 250 ml of fixing solution (100 ml of ethanol, 25 ml of glacial acetic acid and 125 ml of water) for 30 min. It was agitated by rocker. Then, the gel was nutralized by 250 µl of sensitizing solution [75 ml of ethanol, 1.25 ml of glutardiadehyde (25%w/v), 10 ml

of sodium thiosulphate (5%w/v) 1 packet of sodium acetate (17 g)] for 30 min. It was washed by dH_2O three times for 5 min each time. The gel was stained by 250 of silver solution [25 ml of silver nitrate solution (2.5%w/v), 0.1 ml of formaldehyde (37%w/v)] for 20 min. It was washed by dH_2O two times for 1 min each time. Gel was submerged in 250 ml of developing solution [1 packet of sodium carbonate (6.25 g), 0.05 ml of formaldehyde (37%w/v)] for 2 min. The reaction was stopped by 250 µl of stopping solution [1 packet of EDTA-$Na_2.2H_2O$ (3.65 g)] for 10 min. The gel was washed by dH_2O three times for 5 min and kept in sealed bags.

6.5 *GUS* Enzymatic Assay

GUS activity assay was determined with a Fluor AceTM β-glucuronidase reporter assay kit supplied by Bio-Rad Laboratories. The assay was initiated by adding 150 µg of total protein to 500 µl of warm assay buffer [(1mM 4-methy-lumbelliferyl-β-D-glucuronide (MUG) in *GUS* extraction buffer)] and mixed thoroughly with vortex. The reaction mix was incubated in 37 ^{o}C water bath for 30 min. The reaction was terminated by adding 1 µl of 1X Stop Buffer (0.2 M Na_2CO_3) at room temperature. Fluorescence of 4-Methylumbelliferone (4-MU) was measured with the excitation at 365 nm and emission at 455 nm with spectofluorometer (JASCO, Japan). Reading was calibrated using 4-Methylumbelliferone. A standard curve was made from a series of dilutions (0-10,000 nM) of 4-Methylumbelliferone dissolved in 1X Stop Buffer as shown in Table 3. Due to the interference, 4-MU solution must be kept away from light. The minimum and maximum relative fluorescence values were set by 0 and 10,000 nM of 4-MU. *GUS* activity of transformed root was calculated as the formula below.

$$\frac{\text{GUS Activity}}{\text{(pmol 4-MU/min/µg protein)}} = \frac{\text{Intensity of Sample/min}}{\text{Intensity of Sample /pmol 4-MU}} \times \frac{\text{Reaction Vol (µl)}}{\text{Sample Vol (µl)}} \times \frac{1}{\text{Vol.per test (µl)}} \times \frac{1}{\text{Extract conc. (µg protein/µl)}}$$

Table 3 Strategic dilution of 4-Methylumbelliferone in 1X Stop Buffer for the standard curve

Cuvette	4-MU	1X Stop Buffer (ml)	4MU (nM)
1	20 µl of mM 4MU stock	1.98	10,000
2	200 µl of cuvettle 1	1.80	1,000
3	200 µl of cuvettle 2	1.80	100
4	200 µl of cuvettle 3	1.80	10
5	0	2.00	0

RESULTS

1. Isolation of *CP- ACO II* 5'Flanking Region

1.1 Papaya Genomic DNA Extraction

Papaya genomic DNA was extracted from papaya young leaves. The quality and concentration of papaya genomic DNA was determined using 0.7 % agarose gel electrophoresis and spectrophotometer. Agarose gel electrophoresis showed genomic DNA size approximately 23 kb (Figure 10).

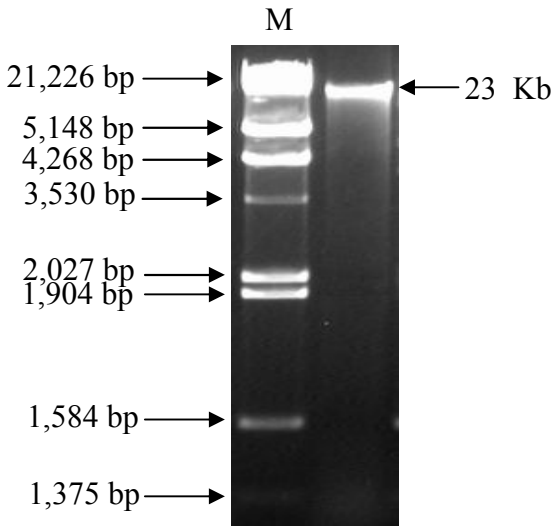


Figure 10 Agarose gel electrophoresis of papaya genomic DNA at 500 ng. The electrophoresis was performed at 90 V for 1 h and agarose was stained for 15 min with ethidium bromide. Lambda DNA/*Eco*RI + *Hind*III markers (M) was shown on the left.

1.2 First Ligation-Mediated PCR

Papaya genomic DNA was digested with restriction enzymes (*Ssp*I, *Dra*I and *Hpa*I). Digested genomic DNA was shown in Figure 11a. Only *Hpa*I yielded complete digestion after overnight reaction. *Hpa*I digested genomic DNA was ligated

to adaptor ADAPL/ADSPS and the ligation mix was called adaptor ligated library. The *Hpa*I adaptor ligated library (10, 5 and 1 µl) was amplified with adaptor forward primer AP1 and *CP-ACO II* 5'flanking reverse specific primer TAKII. The results revealed smear bands (Figure 11b). PCR products from lane 2 were serial diluted at 1:10, 1:100 and 1:1000 and were used as template for nested PCR with nested adaptor forward primer NAP1 and *CP-ACO II* 5'flanking nested reverse specific primer, TAKIIN. The reaction produced two PCR products at 240 and 375 bp in length (Figure 11c).

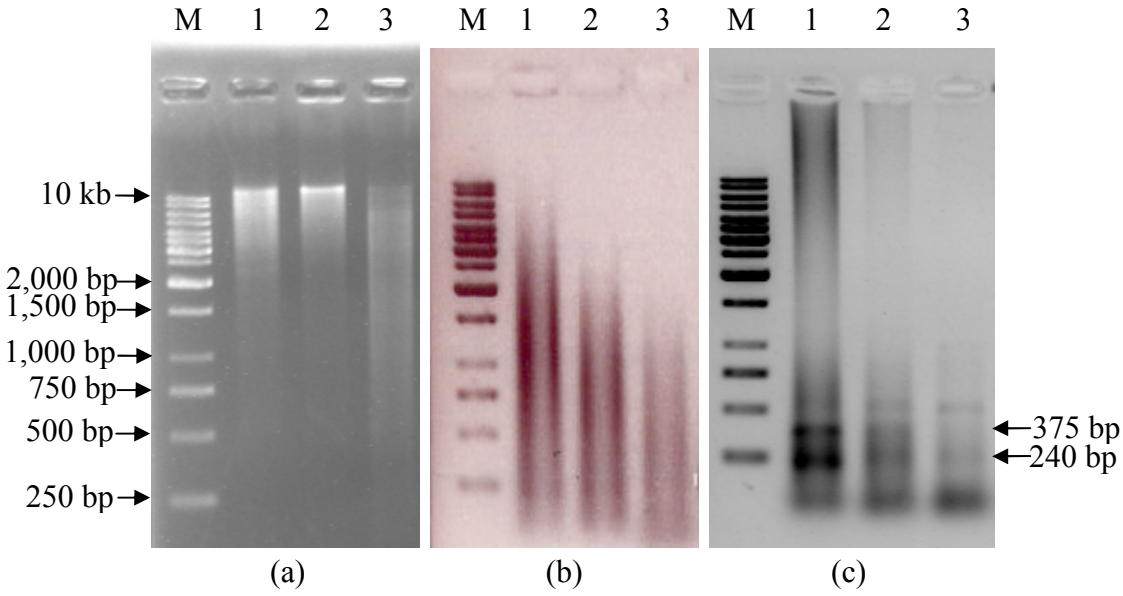


Figure 11 The first ligation-mediated PCR. (a) digested genomic DNA with blunt-ended restriction enzyme *Ssp*I (lane1), *Dra*I (lane 2) and *Hpa*I (lane 3); (b) PCR products of *Hpa*I adaptor ligated library 10 µl (lane 1), 5 µl (lane 2) and 1 µl (lane 3); (c) PCR products of *Hpa*I adaptor ligated library dilution 1:10 (lane 1), 1:100 (lane 2) and 1:1000 (lane 3) amplified wish nested adaptor forward primer NAP1 and *CP-ACO II* specific reverse primer TAKIIN. One kb ladder molecular weight markers (M) was shown on the left.

PCR products from nested PCR were cloned into pDrive cloning vector. The size of recombinant plasmid was confirmed by double digestion with *Mlu*I and *Xho*I. The results indicated that insert fragments from clones TAKIIN#50 (lane 1), TAKIIN#51 (lane 2) and TAKIIN#52 (lane 3) were approximately 240 bp in size (Figure 12). Clone TAKIIN#52 was sequenced. Sequence comparison indicated that clone TAKIIN#52 (240 bp) is a part of *CP-ACO II* 5'flanking region.

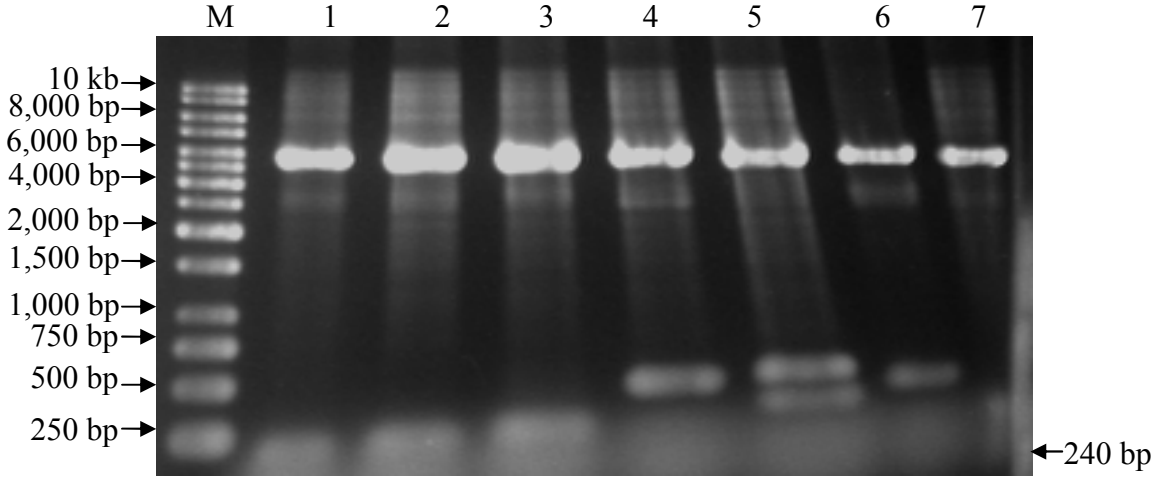


Figure 12 Agarose gel electrophoresis showed insert fragment of clones TAKIIN#50 to TAKIIN#58 (lane 1-7 respectively) were digested with *Mlu*I and *Xho*I. One kb ladder molecular weight markers (M) was shown on the left.

1.3 Second Ligation-Mediated PCR

Papaya genomic DNA was digested with blunt-ended restriction enzymes consisted of *BspL*I, *KspA*I, *Pdm*I, *Rsa*I and *Pvu*II. Digested genomic DNA were shown in Figure 13a. All five digested genomic DNA pools were ligated to adapter ADAPL/ADSPL to create adapter ligated library. Five microlitre of adaptor ligated library *BspL*I, *KspA*I, *Pdm*I, *Rsa*I and *Pvu*II (lane 1-5 respectively) were amplified with adaptor forward primer AP1 and *CP-ACO II* 5'flanking reverse specific primer TAKII. The results revealed smear bands (Figure 13b). PCR products (1:100) of *Pdm*I adaptor ligated library was used as template for nested PCR and were amplified with nested adaptor forward primer NAP1 and *CP-ACO II* 5'flanking nested reverse

specific primer MLPN. The result showed a major PCR product at approximately 500 bp in size (Figure 13c).

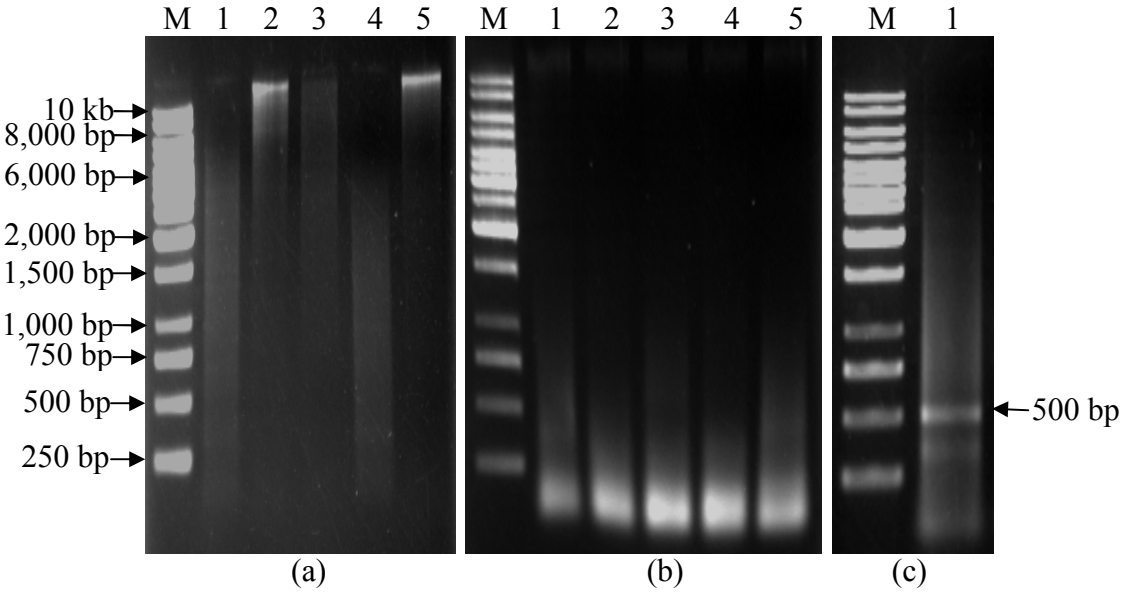


Figure 13 The ligation-mediated PCR (a) digested genomic DNA with blunt-ended restriction enzyme *BspL*I, *KspA*I, *Pdm*I, *Rsa*I and *Pvu*II (lane 1-5 respectively); (b) PCR products adaptor ligated library *BspL*I, *KspA*I, *Pdm*I, *Rsa*I and *Pvu*II (lane 1-5 respectively) using adaptor forward primer AP1 and *CP-ACO II* 5'flanking reverse specific primer TAKII; (c) PCR products of *Pdm*I adaptor ligated library (1:100) amplified with nested adaptor forward primer NAP1 and *CP-ACO II* specific reverse primer MLPN. One kb ladder molecular weight markers (M) was shown on the left.

PCR products were cloned into pDrive cloning vector. The size of recombinant plasmid was confirmed by PCR with universal primer M13F and M13R. The result revealed that insert fragment from clones MLPN#1, #4, #6, #7 and #10 (lane 1,4,6,7 and 10 respectively) were approximately 750 bp in size (Figure 14). Clone MLPN#7 was sequenced. The results indicated that the insert fragments from clone MLPN#7 is a part of TAKIIN#52 sequence. The sequences of MLPN#7 and TAKIIN#52 were assembled to the sequence of *CP-ACO II* gene. The result showed

CP-ACOII 5'flanking region was 591 bp in length from translation start site (Figure 15).

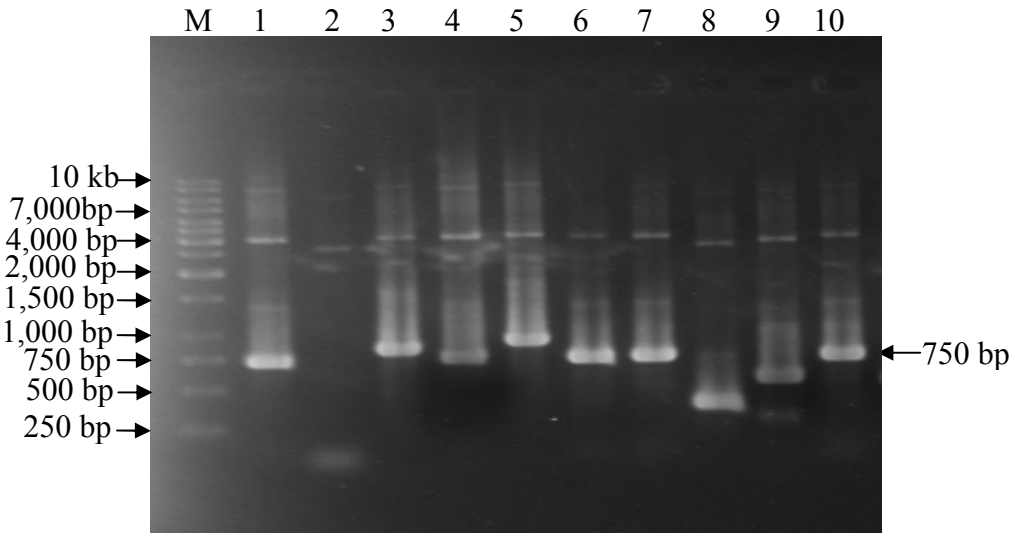


Figure 14 Agarose gel electrophoresis showed PCR products of clones MLPN that amplified with universal forward primer M13F and universal reverse primer M13R. One kb ladder molecular weight markers (M) was shown on the left. Clones MLPN#1-10 were shown on lane 1-10, respectively.

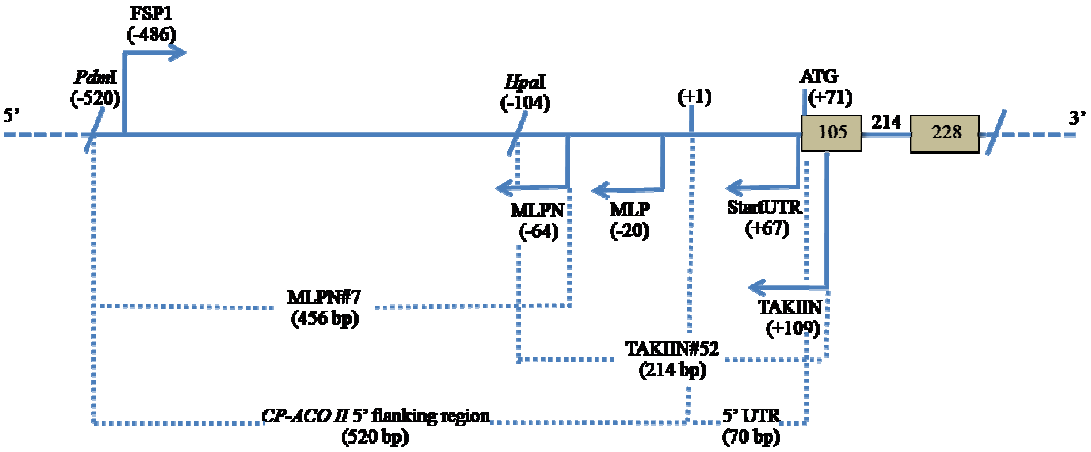


Figure 15 Schematic diagram of the assembled *CP-ACO II* 5'flanking region. Total sequence of *CP-ACO II* 5'flanking region was 591 bp in length from the translation start site.

CP-ACO II 5'flanking sequences and 5'untranslated region from *Carica papaya* Khaeknuan variety was compared to the promoter sequences of other plant *ACO* genes including *C. paapaya* Tainong (*ACO II*), *C. papaya* KhaekNuan (*ACO I*), *C. paapaya* Mexico [*ACO I* (AF 379855)], *Cucumis melo* [*CM-ACO I* (Q04644) and *CM-ACO III* (X9554)], *Lycopersicon esculentum* [*Le-ACO I* (X58273), *Le-ACO II* (Y00478) and *Le-ACO III* (Z54199)], *Malus domestica* (AF030859), *Musa acuminata* (X95599) and *Prunus persica* [*ACO I* (AF 532976)]. The 5'flanking region of *C. papaya* Khaeknuan *ACO II* was 44.5% identical to that of *C. papaya* Khaeknuan *ACO I* (Table 4).

Table 4 Percent identity of *CP-ACO I* and *II* 5'flanking sequences and 5'untranslated region from *Carica papaya* Khaeknuan variety and the promoter sequences of other plant *ACO* genes. Cluster analysis was done using the MegAlign program (DNASTAR).

Percent Identity

	1	2	3	4	5	6	7	8	9	10	11	12		
1		89.7	44.3	35.5	31.9	42.2	36.9	33.4	34.8	31.4	88.7	38.7	1	Carica papaya Tainong ACO2.SEQ
2	11.1		45.3	36.3	33.3	40.1	38.9	34.5	35.1	30.7	98.1	38.5	2	Carica papaya Mexico ACO AF 379855.SEQ
3	104.4	99.6		37.0	32.2	41.1	35.2	36.3	35.7	31.1	44.5	35.8	3	CP-ACOII flanking region 591 bp.seq
4	148.1	142.4	139.7		35.5	36.0	41.3	30.8	34.7	28.7	35.7	35.9	4	Cucumis melo ACO3 X95553.seq
5	190.4	169.1	176.1	152.3		35.4	38.5	32.1	37.3	29.5	32.9	33.8	5	Cucumis melo ACO1 X95551.seq
6	111.7	121.2	118.0	144.5	148.4		36.8	32.5	31.5	29.9	38.6	32.3	6	Malus domestica ACO AF030859.seq
7	140.3	126.8	153.6	114.7	131.7	140.2		32.6	33.3	33.8	38.2	34.9	7	Lycopersicon esculentum ACO2 Y00478.se
8	166.9	156.9	148.8	192.8	178.1	172.5	171.7		35.5	27.2	32.5	28.5	8	Lycopersicon esculentum ACO1 X58273.seq
9	162.0	150.6	148.6	153.7	137.2	183.6	165.0	148.9		30.2	35.7	34.2	9	Lycopersicon esculentum ACO3 Z54199.seq
10	208.8	209.4	195.7	228.4	283.1	217.4	163.1	350.0	218.4		30.8	28.7	10	Musa acuminata ACO AF221107.seq
11	12.2	1.9	103.8	146.9	172.5	128.1	130.7	174.4	146.1	208.4		37.2	11	promoter ACO1 complete.seq
12	130.1	129.6	152.3	145.4	161.2	175.0	152.0	235.0	160.1	254.6	137.0		12	Prunus persica ACO1 AF532976.seq
	1	2	3	4	5	6	7	8	9	10	11	12		

The sequence of 591 bp of *CP-ACO II* flanking region and 5' untranslated region was compared to Genbank database using BLASTN at 3 levels of comparison [highly similar (megablast), more dissimilar (discontiguous megablast) and somewhat similar (blastn)]. The results indicated that only the area between 1- 34 bp upstream from translation start site showed highly similarity to 5'UTR of *C. papaya ACO* accession number L76283 (Figure 16).

```
>   gb|L76283.1|CPA1A1C  Carica papaya 1-aminocyclopropane-1-carboxylate
oxidase mRNA, complete cds
Length=1199

 Score = 63.9 bits (34), Expect = 7e-07
 Identities = 34/34 (100%), Gaps = 0/34 (0%)
 Strand=Plus/Plus

Query  521  AAAAACTTTATTTACAAACAAACCCAGAAAAAGT  554
            ||||||||||||||||||||||||||||||||||
Sbjct  1    AAAAACTTTATTTACAAACAAACCCAGAAAAAGT  34
```

Figure 16 Pairwise alignment of the sequence 591 bp of *CP-ACO II* 5'flanking region and 5'untranslated region to *Carica papaya ACO* mRNA sequence

2. Identification of *Cis*-Acting Element by Database-Assisted Bioinformatics Approach

The analysis of *CP-ACO I* and *II* 5'flanking region using PLACE, Plantcare and PlantProm databases were shown in Table 5-10. The location of each element presented in *CP-ACO I* and *II* 5'flanking region were found in both plus and minus strands. These elements were also reportedly found in both monocot and dicot plants.

These *cis*–acting elements were divided according to their functions into 4 groups including biotic and abiotic responsive, hormone responsive, cell and tissue specific expression and others. Regardless to the databases used, the commonly found elements in biotic and abiotic responsive groups in both *CP-ACO I* and *CP-ACO II* 5'flanking regions were light responsive and dehydration responsive elements respectively. Despite approximately 3 times shorter, *CP-ACO II* 5'flanking region contained more elicitor responsive elements (3) than in *CP-ACO I* 5'flanking region (1) (Table 5 and 8). Interestingly, ARE1, an antioxidant element (RGTGACNNNGC) in this group was only found in *CP-ACO II* 5'flanking region (Table 8) (Rushmore *et al*., 1991). A 20 bp element, BOX A-1 (ATATATATATATTATATAAT) was found twice in *CP-ACO I* 5'flanking region. It was reported to be a sucrose responsive element (Grierson *et al*., 1994).

Seed and endosperm specific elements were the most commonly found in cell and tissue specific expression group both in *CP-ACO I* and *CP-ACO II* 5'flanking regions followed by root and pollen specific elements. Among seed and endosperm specific elements, DOFCOREZM was found 23 times in *CP-ACO I* 5'flanking region where most of them located within 591 bp upstream from translation start site. ROOTMOTIFTAPOX1 and NODCON2GM, root specific elements were found 14 and 8 times respectively in *CP-ACO I* 5'flanking region while only ROOTMOTIFTAPOX1 was found once in *CP-ACO II* 5'flanking region (Elmayan and Tepfer, 1995; Sandal *et al*., 2007).

Hormonal responsive elements to plant hormones such as auxin, cytokinin and abscisic acid were found in both *CP-ACO I* and *CP-ACO II* 5'flanking regions. Interestingly, we also found ethylene responsive element (ERE, ATTTTAAA) only with manual searching in *CP-ACO I* 5'flanking region (Itzhaki *et al*., 1994).

-10PEHVPSBD and CIACADIANLELHC elements involving in circadian expression were found in *CP-ACO I* and *CP-ACO II* 5'flanking regions, respectively (Table 5 and 8) (Thum *et al*., 2001; Piechulla *et al*., 1998).

Table 5 *Cis*-acting elements grouped by functions found in *CP-ACO I* 5'flanking region and 5'untranslated region (1044bp) using PLACE database with PLACE signal scan program revealed 73 *cis*-acting elements.

Groups	Functions	Factors or site names	Organisms	References
Biotic and abiotic responsive	Biotic and abiotic	ASF1MOTIFCAMV	*Arabidopsis thaliana*	Redman *et al*. (2002)
	Carbon dioxide	EECCRCAH1	*Chlamydomonas reinhardti*	Kucho *et al*. (2003)
	Dehydration	ABRELATERD1	*Arabidopsis thaliana*	Simpson *et al.* (2003)
		ACGTATERD1	*Arabidopsis thaliana*	Simpson *et al*. (2003)
		MYB1AT	*Arabidopsis thaliana*	Abe *et al*. (2003)
		MYB2CONSENSUSAT	*Arabidopsis thaliana*	Abe *et al*. (2003)
		MYBCORE	*Arabidopsis thaliana*	Solano *et al*. (1995)
		MYCATRD22	*Arabidopsis thaliana*	Abe *et al*. (1997)
		MYCATERD1	*Arabidopsis thaliana*	Simpson *et al.* (2003)
	Elicitor	BOXLCOREDCPAL	*Daucus carota*	Maeda *et al*. (2005)
	Light	GATABOX	*Petunia hybrida*	Lam *et al*. (1989)
		GT1CONSENSUS	*Nicotiana tabacum*	Zhou (1999)
		I BOX	*Lycopersicon esculentum*	Giuliano *et al.* (1988)
		IBOXCORE	*Lycopersicon esculentum*	Terzaghi *et al*. (1995)
		IBOXCORENT	*Nicotiana tabacum*	Martinez-Hernandez *et al*. (2002)
		INRNTPSADB	*Nicotiana tabacum*	Nakamura *et al*. (2002)
		PRECONSCRHSP70A	*Chlamydomonas reinhardti*	Von Gromoff *et al*. (2006)
		REALPHALGLHCB21	*Lemna gibba*	Degenhardt and Tobin. (1996)
		TBOXATGAPB	*Arabidopsis thaliana*	Chan *et al*. (2001)
		SORLIP1AT	*Arabidopsis thaliana*	Hudson and Quail. (2003)
		SORLIP3AT	*Arabidopsis thaliana*	Hudson and Quail. (2003)
		ZDNAFORMINGATCAB1	*Arabidopsis thaliana*	Ha and An (1988)
	Oxygen	CURECORECR	*Chlamydomonas reinhardtii*	Kropat *et al*. (2005)
	Pathgen	BIHD1OS	*Oryza sativa*	Luo *et al.* (2005)
		GT1GMSCAM4	*Glycine max*	Roger *et al.* (2001)
	Salt	GT1GMSCAM4	*Glycine max*	Roger *et al.* (2001)

Table 5 (Continued)

Groups	Functions	Factors or site names	Organisms	References
Biotic and abiotic responsive	Sugar	CGACGOSAMY3	*Oryza sativa*	Hwang *et al*. (1998)
		PYRIMIDINEBOXOSRAMY1A	*Oryza sativa*	Morita *et al*.(1998)
		WBOXHVISO1	*Hordeum vulgare*	Sun *et al*. (2003)
	Temperature	CCAATBOX1	*Glycine max*	Rieping *et al.* (1992)
		LTRECOREATCOR15	*Brassica napus*	Jaing *et al.* (1996)
		MYCCONSENSUSAT	*Arabidopsis thaliana*	Abe *et al.* (2003)
	Wounding	WBBOXPCWRKY1	*Ipomoea batatas*	Ishiguro and Nakamura (1994)
		WBOXATNPR1		
		WBOXNTERF3	*Arabidopsis thaliana*	Chen *et al*. (2002)
		WRKY71OS	*Nicotiana tabacum*	Nishiuchi *et al.* (2004)
			Oryza sativa	Zhang *et al*. (2004)
Cell and tissue specific expression	Cell cycle	E2FCONSENSUS	*Arabidopsis thaliana*	Vandepoele *et al*. (2005)
	Guard cell	TAAAGSTKST1	*Solanum tuberosum*	Plesch *et al*. (2001)
	Leave	RAV1AAT	*Arabidopsis thaliana*	Kagaya *et al*. (1999)
		ROOTMOTIFTAPOX1	*Agrobacterium rhizogenes*	Elmayan and Tepfer. (1995)
	Meristem	HEXAMERATH4	*Arabidopsis thaliana*	Chaubet *et al*. (1996)
	Mesophyll	CACTFTPPCA1	*Flaveria trinervia*	Gowik *et al*. (2004)
	Pigment synthesis	MYBPZM	*Zea mays*	Grotewold *et al*. (1994)
	Plastid	S1FBOXSORPS1L21	*Spinacia oleracea*	Zhou *et al*. (1992)
	Pollen	GTGANTG10	*Nicotiana tabacum*	Roger *et al*. (2001)
		POLLEN1LELAT52	*Lycopersicon esculentum*	Bate and Twell (1998)
	Root	NODCON1GM	*Glycine max*	Sandal *et al*. (2007)
		NODCON2GM	*Glycine max*	Sandal *et al.* (2007)
		OSE1ROOTNODULE	*Vicia faba*	Vieweg *et al*. (2004)
		OSE2ROOTNODULE	*Vicia faba*	Vieweg *et al*. (2004)
		RAV1AAT	*Arabidopsis thaliana*	Kagaya *et al.* (1999)
		ROOTMOTIFTAPOX1	*Agrobacterium rhizogenes*	Elmayan and Tepfer. (1995)
		SP8BFIBSP8BIB	*Ipomoea batatas*	Ishiguro and Nakamura (1994)

Table 5 (Continued)

Groups	Functions	Factors or site names	Organisms	References
Cell and tissue specific expression	Seed and endosperm	-300CORE	*Nicotiana tabacum*	Thomas *et al*. (1990)
		-300ELEMENT	*Triticum aestivum*	Thomas *et al*. (1990)
		2SSEEDPROTBANAPA	*Brassica napus*	Stalberg *et al*. (1996)
		CAATBOX1	*Pisum sativum*	Shirsat *et al*. (1989)
		CANBNNAPA	*Pisum sativum*	Ellerstrom *et al*. (1996)
		DOFCOREZM	*Zea mays*	Yanagisawa (2000)
		DPBFCOREDCDC3	*Daucus carota*	Kim *et al*. (1997)
		EBOXBNNAPA	*Brassica napus*	Stalberg *et al*. (1996)
		EMHVCHORD	*Hordeum vulgure*	Muller and Kundsen (1993)
		NAPINMOTIFBN	*Brassica napus*	Ericson *et al*. (1991)
		POLASIG1	*Pisum sativum*	Heidecker and Messing (1986)
		POLASIG2	*Oryza sativa*	O'Neill *et al.* (1990)
		PROXBBNNAPA	*Brassica napus*	Ezcurr *et al*. (1999)
		PYRIMIDINEBOXOSRAMY1A	*Oryza sativa*	Morita *et al*.(1998)
		SEF1MOTIF	*Glycine max*	Allen *et al*. (1989)
		SEF3MOTIFGM	*Glycine max*	Allen *et al*. (1989)
		SEF4MOTIFGM7S	*Glycine max*	Allen *et al*. (1989)
	Xylem development	XYLAT	*Arabidopsis thaliana*	Ko *et al*. (2006)
Homone responsive	ABA	MYCCONSENSUSAT	*Arabidopsis thaliana*	Abe *et al*. (2003)
	Auxin	NTBBF1ARROLB	*Agrobacterium rhizogenes*	Baumann *et al*. (1999)
	Cytokinin	ARR1AT	*Arabidopsis thaliana*	Sakai *et al*. (1996)
		CPBCSPOR	*Cucumis sativa*	Fusada *et al.* (2005)
	Gibberellin	PYRIMIDINEBOXOSRAMY1A	*Oryza sativa*	Morita *et al*. (1998)
		WRKY71OS	*Oryza sativa*	Zhang *et al*. (2004)
	Salicylic acid	WBOXATNPR1	*Arabidopsis thaliana*	Chen *et al*. (2002)
Others	Circadian expression	-10PEHVPSBD	*Hordeum vulgure*	Thum *et al*. (2001)
	Fermentative pathway	ANAERO1CONSENSUS	*Arabidopsis thaliana*	Mohanty *et al.* (2005)
	Poly A signal	POLASIG1	*Pisum sativa*	Joshi (1987)
		POLASIG2	*Oryza sativa*	O'Neill *et al.* (1990)
		POLASIG3	*Zea mays*	Heidecker and Messing (1986)

Table 5 (Continued)

Groups	Functions	Factors or site names	Organisms	References
Other	Promoter consensus	CAATBOX1	*Pisum sativa*	Shirsat *et al*. (1989)
		GATABOX	*Petunia hybrida*	Gidoni *et al*. (1989)
		TATABOX2	*Pisum sativa*	Shirsat *et al*. (1989)
		TATABOX3	*Pisum sativa*	Shirsat *et al*. (1989)
		TATABOX4	*Ipomoea batatas*	Shirsat *et al*. (1989)
		TATAPVTRNALEU	*Phaseolus vulgaris*	Yukawa *et al*. (2000)

Table 6 *Cis*-acting elements grouped by functions found in *CP-ACO I* 5'flanking region and 5'untranslated region (1044bp) using PlantCARE database with quality-based clustering and Gibbs Sampling indicated the presence of 23 *cis*-acting elements.

Groups	Functions	Elements	References
Biotic and abiotic stress responsive	Anaerobic	ARE	Manjunath and Sach (1996)
	Dehydration	MBS	Yamagushi-Shinozaki *et al.* (1994)
	Heat	HSE	Pastuglia *et al.* (1997)
	Light	ACE	No report
		BOX4	No report
		G-Box	Staiger *et al.* (1989)
		G-box	Giuliano *et al.* (1988)
		GAG-motif	Ki-Hong Jung *et al.* (2003)
		GATA-motif	Arguello-Astorga and Herrere-Estrella (1996)
		I-box	Arguello-Astorga and Herrere-Estrella (1996)
		TCCC-motif	No report
	Metabolism regulation	O2-site	Lohmer *et al.* (1991)
	Stress responsive	TC-rich repeats	Goldsbrough *et al.* (1993)
Cell and tissue specific responsive	Meristem	CAT box	No report
	Seed and endosperm	ATGCAAAT motif	Chuan-Yin *et al.* (2000)
		Skn-1_motif	Washida *et al.* (1999)
Hormone responsive	Abscisic acid	ABRE	Baker *et al.* (1994)
	Methy jasmonate	CGTCA-motif	Rouster *et al.* (1997)
		TGACG-motif	No report
Others	Enhancer	5UTR Py-rich stretch	Daraselia *et al.* (1996)
		TA-rich region	Karen *et al.* (1996)
	Promoter consensus	CAAT	Manjunath and Sachs (1996)
		TATA box	Pasquali *et al.* (1999)

Table 7 *Cis*-acting elements grouped by functions found in *CP-ACO I* 5'flanking region and 5'untranslated region (1044bp) using Plant Prom DB with NSITE-PL program indicated the presence of 33 *cis*-acting elements.

Groups	Functions	Factors or site names	Organisms	References
Biotic and abiotic responsive	Elicitor	Motif j/k	*Pisum sativum*	Seki *et al.* (1996)
	Light	AT-2a	*Pinus sylvestris*	Avila *et al*. (2002)
		box 3	*Nicotiana plumbaginifolia*	No report
		Gap box 3	*Arabidopsis thaliana*	Hawk-Bin *et al*. (1994)
		Z-DNA-motif	*Arabidopsis thaliana*	Ha and An (1988)
	Meristem	CCGTCG motif	*Arabidopsis thaliana*	Chaubet *et al.* (1996)
	Nitrate	NIT2 BS II	*Chlorella vulgaris*	Cannons and Shiflett (2001)
	Salt	Alfin1 BS3	*Medicago sativa*	Dhundy *et al*. (1998)
	Sucrose	BOX A-1	*Solanum tuberosum*	Grieson *et al*. (1994)
Cell and tissue specific expression	Ovule	GA-6	*Arabidopsis thaliana*	Kooiker *et al.* (2005)
	Plastid	CT-LB	*Spinacia oleracea*	Bolle *et al*. (1996)
	Seed and endosperm	ABRE	*Brassica napus*	Ezcurr *et al*. (1999)
		B2	*Zea mays*	Lohmer *et al*. (1991)
		B4	*Zea mays*	Lohmer *et al*. (1991)
		box b	*Zea mays*	Yanagisawa (2000)
		box e	*Zea mays*	Yanagisawa (2000)
		DLEC2, B	*Phaseolus vulgaris*	Chern *et al*. (1996)
		E-box	*Hordeum vulgare*	Muller and Kundsen (1993)
		E1-core	*Daucus carota*	Kim *et al*. (1997)
		EM1	*Triticum aestivum*	Colot *et al.* (1987)
		GC element	*Chlorella vulgaris*	Helmut *et al*. (1992)
		GCN4 motif	*Oryza sativa*	Yoshihara and Takaiwa. (1996)
		P-box 2	*Triticum aestivum*	Thomas *et al*. (1990)
		Zc2 A/T-2	*Zea mays*	Inmaculada *et al*. (1994)
Homone responsive	Methyl Jasmonate	TA-rich motif	*Nicotiana tabacum*	Bingfang *et al*. (2004)

Table 7 (Continued)

Groups	Functions	Factors or site names	Organisms	References
Others	Constitutuve promoter	C-rich motif	*Spinacia oleracea*	Bisanz-Seyer and Mache (1992)
	MADS Box	CArG box 1	*Lepidium africanum*	No report
		CArG box 2	*Brassica oleracea*	No report
	Promoter consensus	CLE-core	*Nocotiana Species*	Christopher *et al.* (2005)
	Unknown function	Box III	*Arabidopsis thaliana*	No report
		Box III	*Arabidopsis thaliana*	No report
		C-rich R	*Lycopersicon esculentum*	Dean *et al.* (1989)
		GA-2	*Arabidopsis thaliana*	No report

Table 8 *Cis*-acting elements grouped by functions found in *CP-ACO II* 5'flanking region and 5'untranslated region (554 bp) using PLACE database with PLACE signal scan program revealed 69 *cis*-acting elements.

Groups	Functions	Factors or site names	Organisms	References
Biotic and abiotic responsive	Antioxidant	ARE1	*Rattus norvegicus*	Rushmore *et al.* (1991)
	Biotic and abiotic	ASF1MOTIFCAMV	*Arabidopsis thaliana*	Redman *et al.* (2002)
	Carbon dioxide	EECCRCAH1	*Chlamydomonas reinhardti*	Kucho *et al.* (2003)
	Dehydration	ABRELATERD1	*Arabidopsis thaliana*	Simpson *et al* (2003)
		ACGTATERD1	*Arabidopsis thaliana*	Simpson *et al.* (2003)
		MYB2AT	*Arabidopsis thaliana*	Urao *et al.* (1993)
		MYB2CONSENSUSAT	*Arabidopsis thaliana*	Abe *et al.* (2003)
		MYBCORE	*Arabidopsis thaliana*	Solano *et al.* (1995)
		MYCATERD1	*Arabidopsis thaliana*	Simpson *et al.* (2003)
	Elicitor	BOXLCOREDCPAL	*Daucus carota*	Maeda *et al.* (2005)
		ELRECOREPCRP1	*Petroselinum crispum*	Rushton *et al.* (1996)
		PALBOXAPC	*Petroselinum crispum*	Logemann *et al.*(1995)
	Light	GATABOX	*Petunia hybrida*	Lam *et al.* (1989)
		GT1CONSENSUS	*Nicotiana tabacum*	Zhou (1999)
		I BOX	*Lycopersicon esculentum*	Giuliano *et al.* (1988)
		IBOXCORE	*Lycopersicon esculentum*	Terzaghi *et al.* (1995)
		INRNTPSADB	*Nicotiana tabacum*	Nagamura *et al.* (2002)
		PALBOXAPC	*Petroselinum crispum*	Logemann *et al.* (1995)
		PRECONSCRHSP70A	*Chlamydomonas reinhardti*	Von Gromoff *et al.* (2006)
		REALPHALGLHCB21	*Lemna gibba*	Degenhardt and Tobin (1996)
		SORLIP2AT	*Arabidopsis thaliana*	Hudson and Quail (2003)
		TATABOX5	*Pisum sativum*	Tjaden *et al.* (1995)
	Pathogen	ELRECOREPCRP1	*Petroselinum crispum*	Rushton *et al.* (1996)
		GT1GMSCAM4	*Glycine max*	Park *et al.* (2004)
		SEBFCONSSTPR10A	*Solanum tuberosum*	Boyle and Brison (2001)
	Sugar regulation	TATCCAOSAMY	*Oryza sativa*	Lu *et al.* (2002)

Table 8 (Continued)

Groups	Functions	Factors or site names	Organisms	References
Biotic and abiotic responsive	Sugar	CGACGOSAMY3	*Oryza sativa*	Hwang *et al*. (1998)
		CMSRE1IBSPOA	*Ipomoea batatas*	Morikami *et al*. (2005)
		WBOXHVISO1	*Hordeum vulgare*	Sun *et al*. (2003)
	Sulfer	SURECOREATSULTR11	*Arabidopsis thaliana*	Maruyama-Nakashita *et al*. (2005)
	Temperature	CAATBOX1	*Glycine max*	Rieping *et al*. (1992)
		LTRECOREATCOR15	*Brassica napus*	Jaing *et al*. (1996)
		MYCCONSENSUSAT	*Arabidopsis thaliana*	Abe *et al*. (2003)
	Wounding	ELRECOREPCRP1	*Petroselinum crispum*	Rushton *et al*. (1996)
		WBOXATNPR1	*Arabidopsis thaliana*	Chen *et al*. (2002)
		WBOXNTERF3	*Nicotiana tabacum*	Nishiuchi *et al*. (2004)
		WRKY71OS	*Oryza sativa*	Zhang *et al*. (2004)
Cell and tissue specific expression	Cell cycle	MYBCOREATCYCB1	*Arabidopsis thaliana*	Planchais *et al*. (2002)
	Guard cell	TAAAGSTKST1	*Solanum tuberosum*	Plesch *et al*. (2001)
	Leave	RAV1AAT	*Arabidopsis thaliana*	Kagaya *et al*. (1999)
		ROOTMOTIFTAPOX	*Agrobacterium rhizogenes*	Elmayan and Tepfer. (1995)
	Mesophyll	CACTFTPPCA1	*Flaveria trinervia*	Gowik *et al*. (2004)
	Plastid	BOXIINTPATAB	*Nicotiana tabacum*	Kapoor and Sugiura (1999)
	Pollen	GTGANTG10	*Nicotiana tabacum*	Rogers *et al*. (2001)
		POLLEN1LELAT52	*Lycopersicon esculentum*	Bate and Twell (1998)
	Root	NODCON1GM	*Glycine max*	Sandal *et al*. (2007)
		OSE1ROOTNODULE	*Vicia faba*	Vieweg *et al*. (2004)
		RAV1AAT	*Arabidopsis thaliana*	Kagaya *et al*. (1999)
		RHERPATEXPA7	*Arabidopsis thaliana*	Kim *et al*. (2006)
		ROOTMOTIFTAPOX	*Agrobacterium rhizogenes*	Elmayan and Tepfer. (1995)
	Seed and endosperm	-300ELEMENT	*Triticum estivum*	Thomas *et al*. (1990)
		2SSEEDPROTBANAPA	*Brassica napus*	Stalberg *et al*. (1996)
		AACACOREOSGLUB1	*Oryza sativa*	Wu *et al*. (2000)
		CANBNNAPA	*Pisum sativum*	Ellerstrom *et al*.(1996)

Table 8 (Continued)

Groups	Functions	Factors or site names	Organisms	References
Cell and tissue specific expression	Seed and endosperm	DOFCOREZM	*Zea mays*	Yanagisawa (2000)
		EBOXBNNAPA	*Brassica napus*	Stalberg *et al*. (1996)
		POLASIG1	*Pisum sativum*	Heidecker and Messing (1986)
		POLASIG2	*Oryza sativa*	O'Neill *et al.* (1990)
		RYREPEATBNNAPA	*Brassica napus*	Ezcurra *et al*. (1999)
		RYREPEATLEGUMINBOX	*Glycine max*	Fujiwara *et al.* (1994)
		SEF3MOTIFGM	*Glycine max*	Lessard *et al.* (1989)
Homone responsive	ABA	MYCCONSENSUSAT	*Arabidopsis thaliana*	Abe *et al.* (2003)
		RYREPEATBNNAPA	*Brassica napus*	Ezcurra *et al.* (1999)
	Auxin	NTBBF1ARROLE	*Agrobacterium rhizogenes*	Baumann *et al.* (1999)
		SURECOREATSULTR11	*Arabidopsis thaliana*	Maruyama-Nakashita *et al.* (2005)
	Cytokinin	ARR1AT	*Arabidopsis thaliana*	Sakai *et al*. (1996)
	Gibberellin	WRKY71OS	*Oryza sativa*	Zhang *et al*. (2004)
Others	Circadian expression	CIACADIANLELHC	*Lycopersicon esculentum*	Piechulla *et al*. (1998)
	Fermentative pathway	ANAERO1CONSENSUS	*Arabidopsis thaliana*	Mohanty *et al.* (2005)
	Lignin biosynthesis	MYBPLANT	*Antirrhinum majus*	Sablowski *et al.* (1994)
	MAD box	CARGCW8GAT	*Arabidopsis thaliana*	Tang *et al*. (2003)
	Poly A signal	POLASIG1	*Pisum sativa*	Joshi (1987)
		POLASIG2	*Oryza sativa*	O'Neill *et al*. (1990)
	Promoter consensus	GATABOX	*Petunia hybrida*	Gidoni *et al*. (1989)
		TATABOX2	*Pisum sativa*	Shirsat *et al.* (1989)
		TATABOX3	*Pisum sativa*	Shirsat *et al*. (1989)
		TATABOX5	*Pisum sativa*	Shirsat *et al*. (1989)
	Proto-oncogene	MYBST1	*Solanum tuberosum*	Baranowskij *et al*. (1994)
	Scaffold attachment region	MARABOX1	*Drodophila primaeva*	Gasser *et al.* (1989)
		MARTBOX	*Drodophila. primaeva*	Gasser *et al.* (1989)

Table 9 *Cis*-acting elements grouped by functions found in *CP-ACO II* 5'flanking region and 5'untranslated region (554 bp) using PlantCARE database with quality-based clustering and Gibbs Sampling indicated the presence of 22 *cis*-acting elements.

Groups	Functions	Elements	References
Biotic and abiotic stress responsive	Anaerobic	ARE	Manjunath and Sach (1996)
	Dehydration	MBS	Yamagushi-Shinozaki *et al.* (1994)
	Fungal elicitor	Box-W1	Rushton *et al.* (1996)
	Heat	HSE	Pastuglia *et al.* (1997)
	Light	BOX4	Lois *et al.* (1989)
		BOX I	Kuhlemeier *et al.* (1988)
		CATT-motif	No report
		G-Box	Staiger *et al.* (1989)
		G-box	Giuliano *et al.* (1988) and Gilmartin *et al.* (1990)
Cell and tissue specific responsive	Meristem	CCGTCC- box	No report
Hormone responsive	Abscisic acid	ABRE	Baker *et al.* (1994)
	Auxin	AuxRR-core	Sakai *et al.* (1996)
		TGA-element	Caroline *et al.* (2002)
	Methy jasmonate	CGTCA-motif	Rouster *et al.* (1997)
		TGACG-motif	No report
Others	Circadian control	circadian	Giuliano *et al.* (1988)
	Promoter consensus	CAAT	Manijunath and Sachs (1997)
		TATA box	Pasquali *et al.* (1999)
	Unknown function	A-box	Logemann *et al.* (1995)
		AAGAA-motif	Bruce *et al.* (1991)
		TCCACCT-motif	No report
		Unnamed_4	No report
		W box	Eulgem et *al.* (1999)

Table 10 *Cis*-acting elements grouped by functions found in *CP-ACO II* 5'flanking region and 5'untranslated region (554 bp) using Plant Prom DB with NSITE-PL program indicated the presence of 27 *cis*-acting elements.

Groups	Functions	Factors or site names	Organisms	References
Biotic and abiotic responsive	Anaerobic	9-mer box	*Zea mays*	Geffer *et al*. (2000)
	Elicitor	Box A	*Petroselinum crispum*	Logemann *et al*.(1995)
	Light	AT-1(2)	*Pisum sativum*	Arguello-Astorga and Herrere-Estrella (1996)
		Box 1	*Pisum sativum*	Janice *et al*. (1990)
		CUF-1	*Arabidopsis thaliana*	Shawn *et al*. (1994)
	Temperature	TATA/dyad motif	*Glycine max*	Czarnecka *et al*. (1990)
	Sucrose	Box A-1	*Solanum tuberosum*	Grieson *et al.* (1994)
Cell and tissue specific expression	Callus	C1-box	*Daucus carota*	Raghavan (1997)
	Cell cycle	MSA	*Catharanthus roseus*	Masaki *et al.* (1998)
		NON	*Nicotiana tabacum*	Shen and Gigol (1997)
		Site I	*Arabidopsis thaliana*	Tremousaygue *et al*. (2003)
	Nodule	Element 2	*Glycine max*	Jensen *et al*. (1988)
	Root	C2a-BS	*Oryza sativa*	He *et al*. (2002)
	Seed and endosperm	RY	*Glycine max*	Allen *et al*. (1989)
Homone responsive	ABA	-190 half G-box (core)	*Arabidopsis thaliana*	Lu *et al*. (1996)
		ABRE II	*Craterostigm plantagineum*	Andrea and Dorothea (2006)
		E2-core	*Daucus carota*	Chung *et al*. (2005)
		EM2	*Triticum aestivum*	Luo *et al*. (1993)
	Gibberellin	GARE/Box 2	*Hordeum vulgare*	Yamauchi *et al.* (2002)
Others	MADS Box	CArG box 1	*Lepidium africanum*	No report
	Promoter consensus	CT-B	*Spinacia oleracea*	Bolle *et al*. (1994)
	Retro transposon	Region B1	*Zea mays*	Sanmiguel *et al*. (1996)
	Unknown function	adjb	*Nicotiana tabacum*	No report
		CCGTCC motif	*Zea mays*	Rossitza *et al*. (1998)
		Dnase I footprint	*Arabidopsis thaliana*	No report
		EE 1	*Solanum melongena*	No report

Cis-acting elements found only in *CP-ACO I* or *CP-ACO II* 5'flanking and 5'untranslated regions (promoter region) from all 3 databases were further investigated for their roles in other plant species (Table 11 and 12). These specific elements were generally presented no more than 3 times. Element commonly found in both *CP-ACO I* and *II* 5'flanking regions were also identified (Table 13). The function analysis showed 19 important elements involving in root specific expression, mesophyll and plastid specific expression, ovule specific expression and seed and endosperm specific expression.

Table 11 C*is*-acting elements grouped by found only in *CP-ACO I* 5'flanking region and 5'untranslated region. * Location from +1 transcription start site, (+) = sense strand, (-) = antisense strand, ♠ = The element which found five times or more, their locations are not shown.

Factors or site names	Sequences	Frequency /location*	Description		
			Gene	Plant species	Function
-10PEHVPSBD (Thum *et al.,* 2001)	TATTCT	1/-584(-)	*psbD* gene	barley (*Hordeum vulgare*)	activated by blue, white or UV-A light
5UTR Py-rich stretch (Daraselia *et al.,* 1996)	TTTCTTCTCT	1/-535(-)	3-Hydroxy-3-Methylglutaryl Coenzyme A Reductase gene	tomato (*Lycopersicon esculentum*)	enhancer
BIHD1OS (Luo *et al.,* 2005)	TGTCA	2/+72(+), -47(-)	*OsBIHD1*gene	rice (*Oryza sativa*)	disease resistance responsiveness
CURECORECR (Kropat *et al.,* 2005)	GTAC	6♠	*Cyc6* and *Cpx1* genes	*Chlamydomonas reinhardtii*	involved in oxygen-response of these genes
GA-6 (Kooiker *et al.,* 2005)	GGAGAGAGA	1/-551(+)	SEEDSTICK gene	*Arabidopsis thaliana*	ovule- and septum-specific expression recognized by GA binding protein.
HEXAMERATH4 (Chaubet *et al.,*1996)	CCGTCG	1/-447(-)	histone (H4) gene	*Arabidopsis thaliana*	regulate its expression

Table 11 (Continued)

Factors or site names	Sequences	Frequency /location*	Description		
			Gene	Plant species	Function
IBOXCORENT (Martinez-Hernandez *et al.,* 2002)	GATAAGR	1/-546(-)	*RBCS* light-responsive gene	tobacco (*Nicotiana plumbaginifolia*)	activated by phytochrome, cryptochrome and plastid signals
MYBPZM (Grotewold *et al.,* 1994)	CCWACC	2/+94(+), -457(-)	*P* gene	maize (*Zea mays*)	myb homolog binding site of that specifies red pigmentation of kernel pericarp, cob and other floral organ
S1FBOXSORPS1L21 (Lagrange *et al.,* 1993)	ATGGTA	1/-480(+)	*RPS1* and *RPL21* genes	spinach (*Spinacia oleracea*)	encoding the plastid ribosomal protein S1 and L21, respectively
SP8BFIBSP8BIB (Ishiguro and Nakamura, 1994)	TACTATT	1/-755(+)	-330, -220, and -200 of *gSPO-B1* (sporamin) gene, -80 of *gB-Amy* (beta-amylase) gene	sweet potato (*Ipomoea batatas*)	recognized by SPB8F protein
TA-rich region (Karen *et al.,* 1996)	TATATATATATATATATATATA	1/-826(+)	peroxidase gene	tobacco (*Nicotiana tabacum*)	enhancer
TC-rich repeats (Goldsbrough *et al.,* 1993)	ATTTTCTTCA	1/-538(-)	β-1,3–glucanase gene	barley (*Hordeum vulgare*)	salicylic binding protein

Table 11 (Continued)

Factors or site names	Sequences	Frequency /location*	Description		
			Gene	Plant species	Function
WBBOXPCWRKY1 (Ishiguro and Nakamura, 1994)	TTTGACY	1/-259(-)	amylase gene	sweet potato (*Ipomoea batatas*)	activated by WRKY proteins
WRKY71OS (Zhang *et al.,* 2004)	TGAC	6♠	*Amy32b* gene	rice (*Oryza sativa*)	activated by WRKY71 which is a transcriptional repressor of the gibberellin signaling pathway
XYLAT (Ko *et al.*, 2006)	ACAAAGAA	1/+110(-)	core xylem gene set	*Arabidopsis thaliana*	involved regulating secondary xylem development

Table 12 *Cis*-acting elements grouped by found only in *CP-ACO II* 5'flanking region and 5'untranslated region. * Location from +1 transcription start site, (+) = sense strand, (-) = antisense strand, ♠ = The element which found five times or more, their locations are not shown.

Factors or site names	Sequences	Frequency /location*	Description		
			Gene	Plant species	Function
A-box (Logemann *et al.,* 1995)	CCGTCC	1/-403(-)	phenylalanine ammonia-lyase (*PAL*) gene	parsley (*Petroselinum crispum*)	associated with P- and L-box that involved in induced transcriptional activity
AACACOREOSGLUB1 (Wu *et al.,* 2000)	AACAAAC	1/+16(+)	glutelin gene	rice (*Oryza sativa*)	involved in controlling the endosperm-specific expression
ARE1 (Rushmore *et al.,* 1991)	RGTGACNNNGC	1/-355(-)	glutathione s-transferase and NAD(P)H quinone reductase genes	rat	antioxidant response element
CARGCW8GAT (Tang *et al.,* 2003)	CWWWWWWWWG	2/-58(+) and -58(-)	*AGL15* gene	*Arabidopsis thaliana*	activated by MADS protein
Circadian (Giuliano *et al.,* 1988)	CAANNNNATC	1/-11(+)	chlorophyll a/b-binding (cab) genes	*Lycopersicon sculentum*	involved in the expression of chloroplast-localized proteins and circadian expression
CT-B (Bolle *et al.,*1994)	ACCCACTTC	1/-119(+)	*PetH* gene	spinach (*Spinacia oleracea*).	element essential for the transcription

Table 12 (Continued)

Factors or site names	Sequences	Frequency /location*	Description		
			Gene	Plant species	Function
ELRECOREPCRP1 (Rushton *et al.,* 1996)	TTGACC	1/-240(-)	*PR1* gene	parsley (*Petroselinum crispum*)	elicitor responsiveness, recognized by WRKY protein
MYBPLANT (Sablowski *et al.,* 1994)	MACCWAMC	1/-13(+)	phenylpropanoid biosynthetic gene	*Antirrhinum majus*	involved phenylpropanoid and lignin biosynthesis
NON (Shen and Gigol, 1997)	CtATCCAAC	1/-407(-)	histone gene	*Nicotiana tabacum*.	related to phosphorylation and cell cycle alteration. It recognized by protein complexes.
PALBOXAPC (Logemann *et al.,* 1998)	CCGTCC	1/-402(-)	phenylalanine ammonia-lyase (*PAL*) gene	parsley (*Petroselinum crispum*)	involved elicitor and light responsive
Site I (Tremousaygue *et al.,* 2003)	CCAGtTGG	1/-418(+)	*eEF1A* gene	*Arabidopsis thaliana*	regulate several ribosomal protein genes in root meristems
TATABOX5 (Tjaden *et al.,* 1995)	TTATTT	7♠	glutamine synthetase gene	pea (*Pisum sativum*)	regulation of light-regulated gene

Table 13 *Cis*-acting elements commonly found in both *CP-ACO I* and *CP-ACO II* 5'flanking region and 5'untranslated region.
* Location from +1 transcription start site, (+) = sense strand, (-) = antisense strand, ♠ = The element which found five times or more, their locations are not shown.

Factors or site names	Sequences	Frequency /location*	Description		
			Gene	Plant species	Function
Box A-1 (Grierson, 1994)	ATATATATATATTATATAAT	2/-606 (+) and -362(+) in ACOI 1/-182(+) in ACOII	patatin gene	potato (*Solanum tuberosum*)	sucrose response
CACTFTPPCA1 (Gowik *et al.,* 2004)	YACT	13♠ in ACOI 9♠ in ACOII	phosphoenolpyruvate carboxylase (*ppcA1*) gene	*Flaveria trinervia*	mesophyll expression
DOFCOREZM (Yanagisawa, 2000)	AAAG	23♠ in ACOI 7♠ in ACOII	cytosolic orthophosphate kinase (CyPPDK) and a non-photosynthetic *PEPC*	maize (*Zea mays*)	binding of Dof proteins
GATABOX (Lam *et al.*, 1989)	GATA	7♠ in ACOI 5♠ in ACOII	chlorophyll a/b (*Cab22*) gene	*Petunia hybrid*	high level of light regulation and tissue specific expression
GT1CONSENSUS (Zhou, 1999)	GRWAAW	12♠ in ACOI 4/-158(+), +26(+), -218(-) and -206(-) in ACOII	light-regulated genes	pea (*Pisum sativum*)	consensus GT-1 binding site
GT1GMSCAM4 (Park *et al.,* 2004)	GAAAAA	6♠ in ACOI 2/-161(+) and +26(+) in ACOII	*SCaM-4* gene	soybean (*Glycine max*)	activated by pathogen and salt
GTGANTG10 (Rogers *et al.,* 2001)	GTGA	6♠ in ACOI 2/-349(-) and -274(-) in ACOII	late pollen gene (*g10*)	tobacco (*Nicotiana tabacum*)	homology to pectate lyase

Table 13 (Continued)

Factors or site names	Sequences	Frequency /location*	Description		
			Gene	Plant species	Function
NODCON1GM (Stougaard *et al.,*1990)	AAAGAT	1/-766(-) in ACOI 1/-195(-) in ACOII	*lbc* and *N23* genes	*Glycine max*	nodulin consensus sequences
NTBBF1ARROLB (Baumann *et al.*, 1999)	ACTTTA	2/-851(+),-686(-) in ACOI 1/+4(+) in ACOII	-341 to -306 of *rolB* gene	*Agrobacterium rhizogenes*	activated by (NtBBF1) Dof protein from tobacco
OSE1ROOTNODULE (Vieweg *et al.,* 2004)	AAAGAT	1/-387(+) in ACOI 1/-375(-) in ACOII	*VfLb29* gene	*Vicia faba*	nodule specific expression
POLLEN1LELAT52 (Bate and Twell., 1998)	AGAAA	8♠ in ACOI 7♠ in ACOII	-72 to -68 of *lat52* gene	tomato (*Lycopersicon esculentum*)	involved pollen specific activation
RAV1AAT (Kagaya *et al.,* 1999)	CAACA	2/+25(+),-287(-) in ACOI 2/-237(+) and -16(+) in ACOII	*RAV1* gene	*Arabidopsis thaliana*	recognized by transcription factor
ROOTMOTIFTAPOX1 (Elmayan and Tepfer. 1995)	ATATT	14♠ in ACOI 1/-174(+) in ACOII	*rolD* gene	*Agrobacterium rhizogenes*	root specific and strength expression
TAAAGSTKST1 (Plesch *et al.,* 2001)	TAAAG	7♠ in ACOI 2/-465(-) and +5(-) in ACOII	*KST1* gene	potato (*Solanum tuberosum*)	guard cell-specific gene expression

Table 13 (Continued)

Factors or site names	Sequences	Frequency /location*	Description		
			Gene	Plant species	Function
TATABOX2 (Shirsat *et al.,*1989)	TATAAAT	3/-28(+),-649(-) and -612(-) in ACOI 1/-57(+) in ACOII	*legA* gene	pea (*Pisum sativum*)	involved seed specific responsive
WBOXNTERF3 (Nishiuchi *et al.*,, 2004)	TGACY	3/-872(-),-259(-) and -219(-) in ACOI 3/-351(-),-276(-) and -240(-) in ACOII	*ERF3* gene	tobacco (*Nicotiana tabacum*)	involved in activation of *ERF3* gene by wounding
WRKY71OS (Zhang *et al.,* 2004)	TGAC	6♠ in ACOI 4/-372(-), -350(-), -275(-) and -239(-) in ACOII	*Amy32b*gene	rice (*Oryza sativa*)	transcriptional repressor of the gibberellin signaling pathway

3. Expression Vector Construction

3.1 Generation of *CP-ACO I* and *II* 5'Flanking Region Fragment

Three fragments of *CP-ACO I* 5'flanking region comprised *Carica papaya ACC oxidase I* short promoter (CPACOI-SPP) size 644 bp, *C. papaya ACC oxidase I* long promoter (CPACOI-LPP) size 1044 bp and *C. papaya ACC oxidase I* promoter that consisted of repeated DOFCOREZM *cis*-acting elements (CPACOI-DOFCORN) size 579 bp were amplified for expression vector construction. CPACOI-SPP, CPACOI-LPP and CPACOI-DOFCORN located between -478 to +147 bp, -878 to +147 bp and -579 to +1 bp from transcription start site, respectively (Figure 17). These fragments were generated by PCR method with *CP-ACO I* 5'flanking region specific primers (Chuaboonmee, 2004). PCR products of CPACOI-SPP, CPACOI-LPP fragments were shown in lane 1 and 2, respectively (Figure 19a). PCR product of CPACOI-DOFCORN fragment was shown in lane1 (Figure 19b).

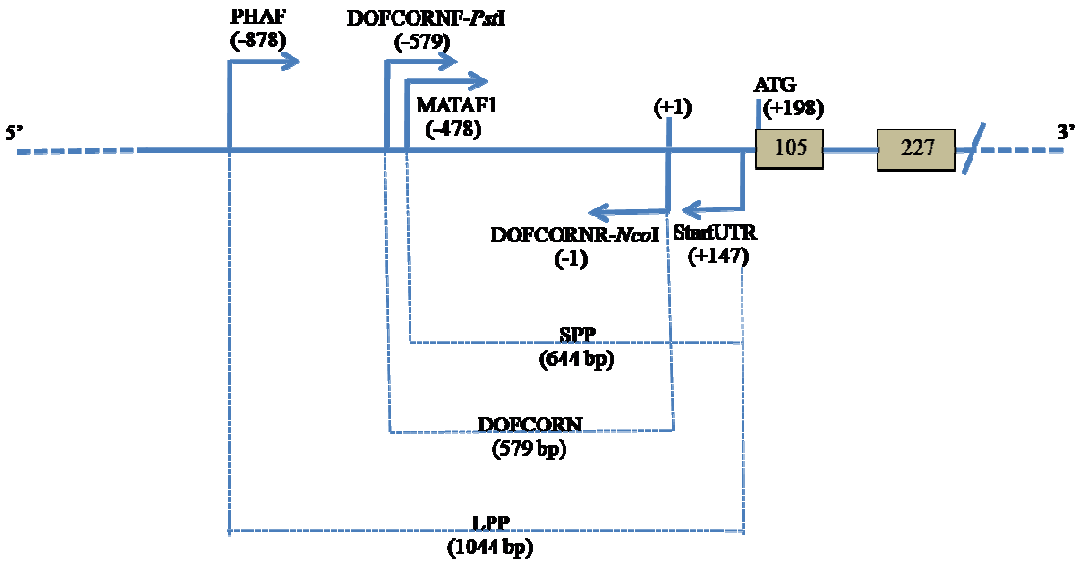


Figure 17 Schematic diagram showing locations of three fragments of *CP-ACO I* 5'flanking region

Four fragments of *CP-ACO II* 5'flanking region comprised *Carica papaya ACC oxidase II* short promoter I (CPACOII-SP1) size 594 bp, *C. papaya ACC oxidase II* short promoter that lack seed and endosperm *cis*-acting elements (CPACOI-NO-SEboxN) size 198 bp, *C. papaya ACC oxidase II* short promoter III (CPACOII-SP3) size 283 bp and *C. papaya ACC oxidase II* short promoter IV (CPACOII-SP4) size 285 bp were amplified for expression vector construction. CPACOII-SP1, CPACOI-NO-SEboxN, CPACOII-SP3 and CPACOII-SP4 located between -486 to +71 bp, -410 to -212 bp, -486 to -212 bp and -215 to +67 bp from transcription start site, respectively (Figure18). These fragments were generated by PCR method with *CP-ACO II* 5'flanking region specific primers. PCR products of CPACOII-SP1, CPACOI-NO-SEboxN, CPACOII-SP3 and CPACOII-SP4 were shown in Figure 20 a, b, c and d.

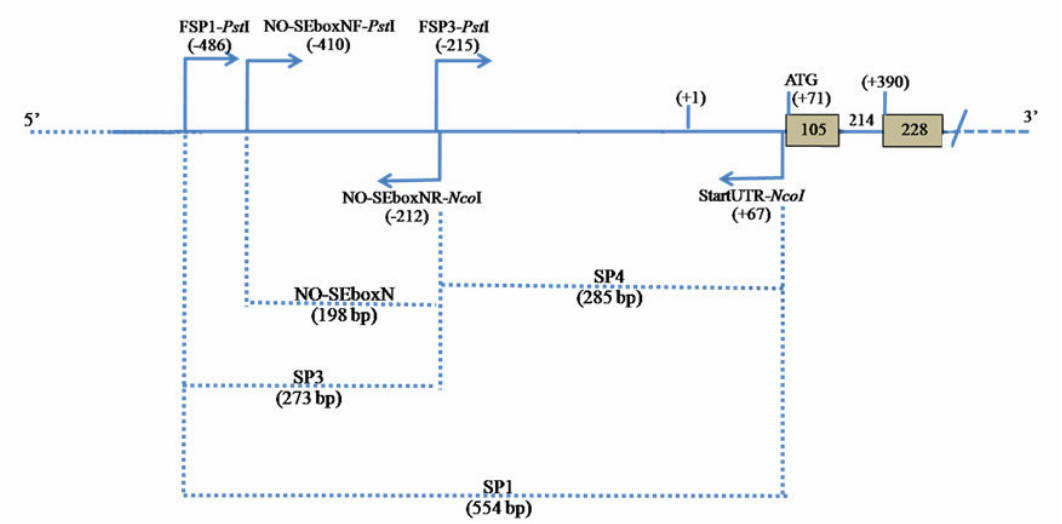


Figure 18 Schematic diagram showing locations of four fragments of *CP-ACO II* 5'flanking region

Seven fragments of *CP-ACO I* and *II* 5'flanking region from 3.1 were ligated into vector pGEM-T Easy (Promega). The recombinant plasmid was transformed into *E coli.* DH5α competent cells. The positive clones were identified by blue/white colony.

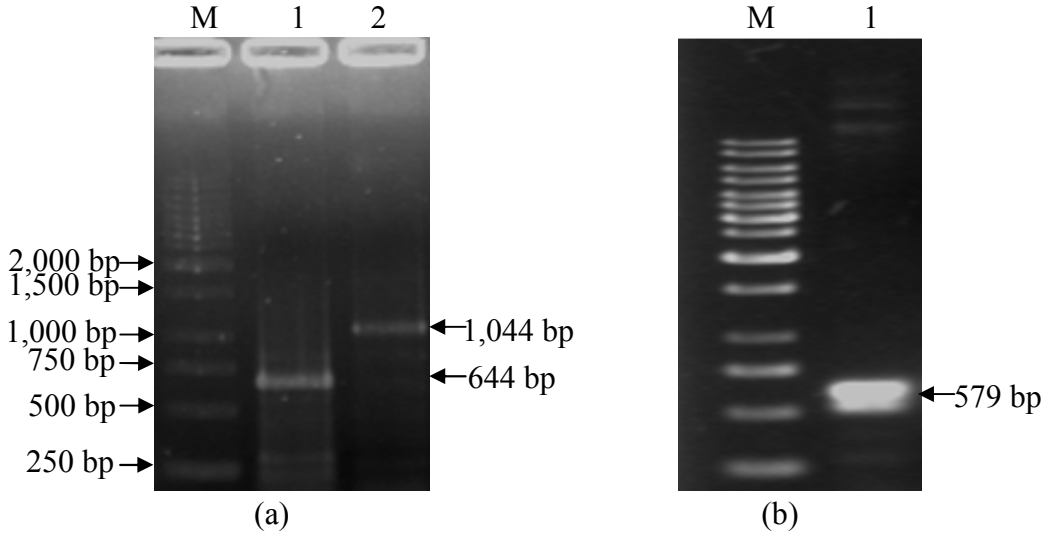


Figure 19 Agarose gel electrophoresis of PCR products of (a) CPACOI-SPP (lane 1) and CPACOI-LPP (lane 2) and (b) CPACOI-DOFCORN (lane1). One kb ladder molecular weight markers (M) was shown on the left.

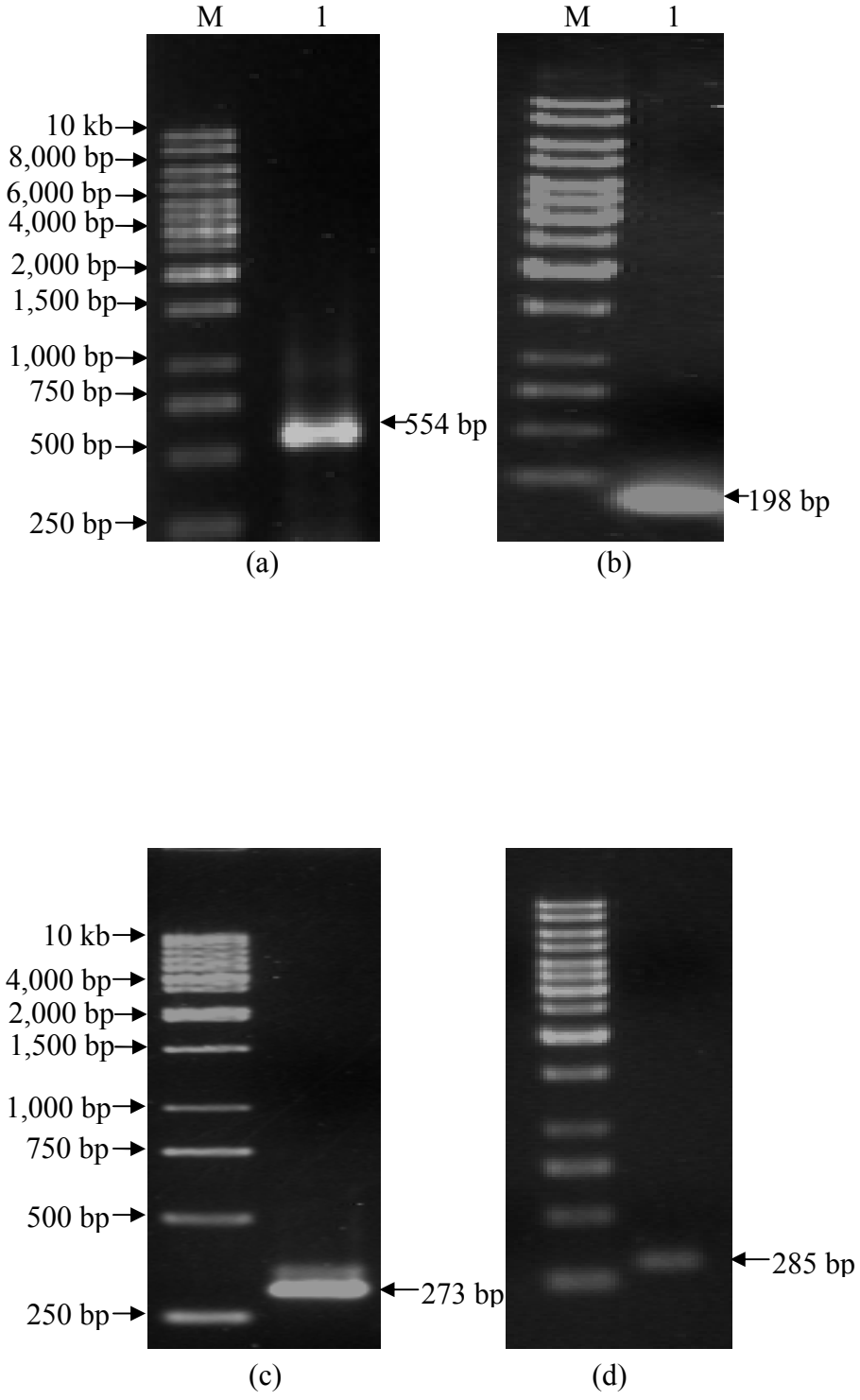


Figure 20 Agarose gel electrophoresis of PCR products of (a) CPACOII-SP1 (lane1), (b) CPACOII-NO-SEboxN (lane 1), (c) CPACOII-SP3 (lane 1) and (d) CPACOII-SP4 (lane1). One kb ladder molecular weight markers (M) was shown on the left.

3.2 Expression Vector Construction

pGEM-T Easy recombinant clones of seven fragments of *CP-ACO I* and *II* 5'flanking region were double digested with enzymes *Nco* I and *Pst* I at multiple cloning regions. The digested products were determined by agarose gel electrophoresis (Figure 21 and 22).

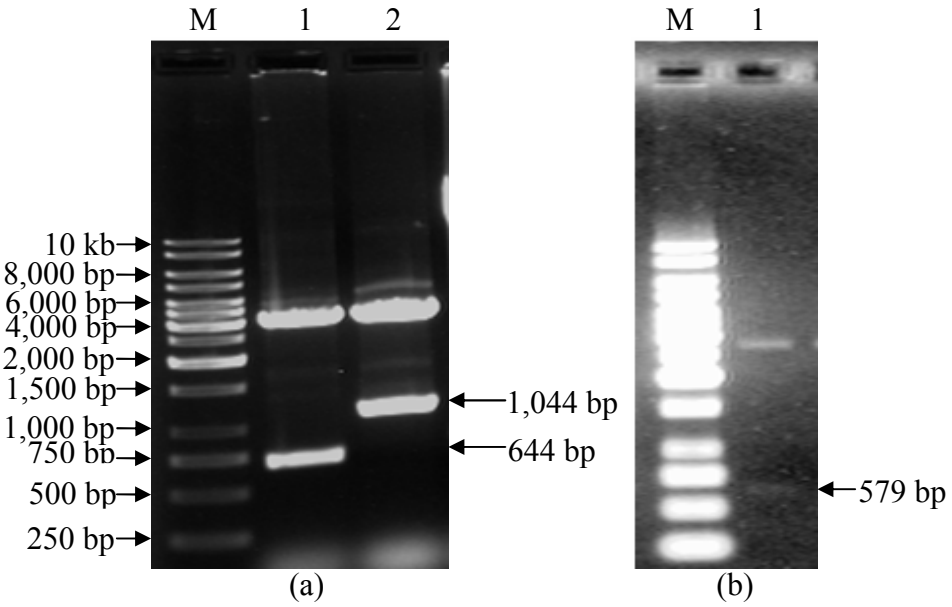


Figure 21 Agarose gel electrophoresis showed pGEM-T Easy recombinant clones of *CP-ACO I* 5'flanking fragments were double digested with *Nco* I and *Pst* I. (a) clone CPACOI-SPP and clone CPACOI-LPP (lane 1 and 2, respectively) and (b) clone CPACOI-DOFCORN (lane1). One kb ladder molecular weight markers (M) was shown on the left.

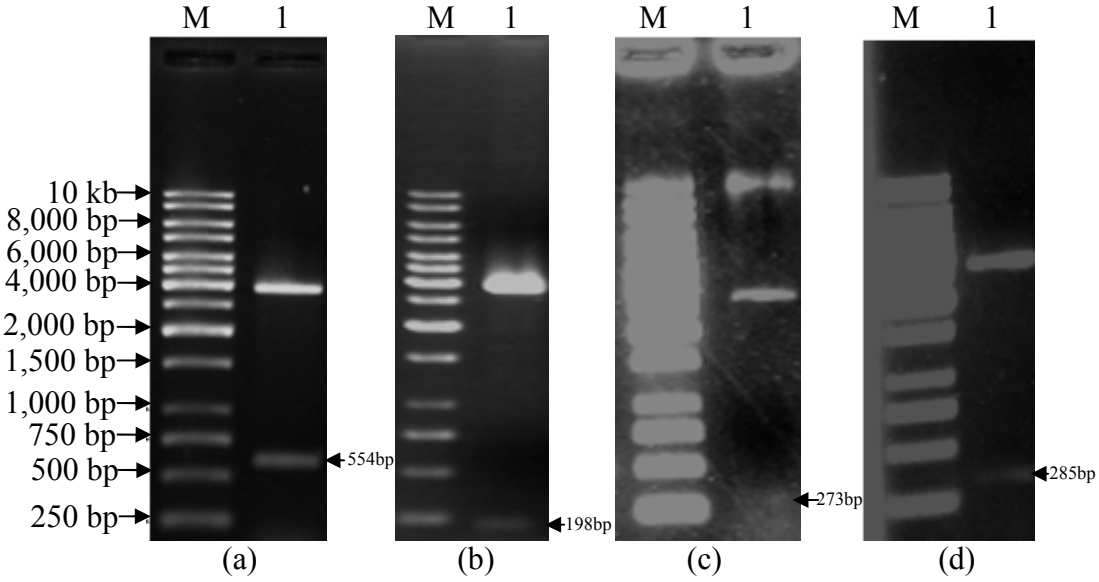


Figure 22 Agarose gel electrophoresis showed pGEM-T Easy recombinant clones of *CP-ACO II* 5'flanking fragment were digested with *Nco* I and *Pst* I. (a) clone CPACOII-SP1 (lane 1), (b) clone CPACOII-NO-SEboxN (lane 1), (c) clone CPACOII-SP3 (lane 1) and (d) clone CPACOII-SP4 (lane1). One kb ladder molecular weight markers (M) was shown on the left.

Binary vector pCAMBIA1304 containing 35S CaMV promoter for driving *mgfp* and *gus* reporter gene expression was double digested with enzyme *Nco* I and *Pst* I at position 11046-11842. *Lac Z* gene and 35S CaMV promoter for driving *mgfp* and *gus* reporter gene expression size 807 bp were deleted (Figure 23). The digested pCAMBIA1304 was purified by gel extraction kit (Fermentas) and used for expression vector construction.

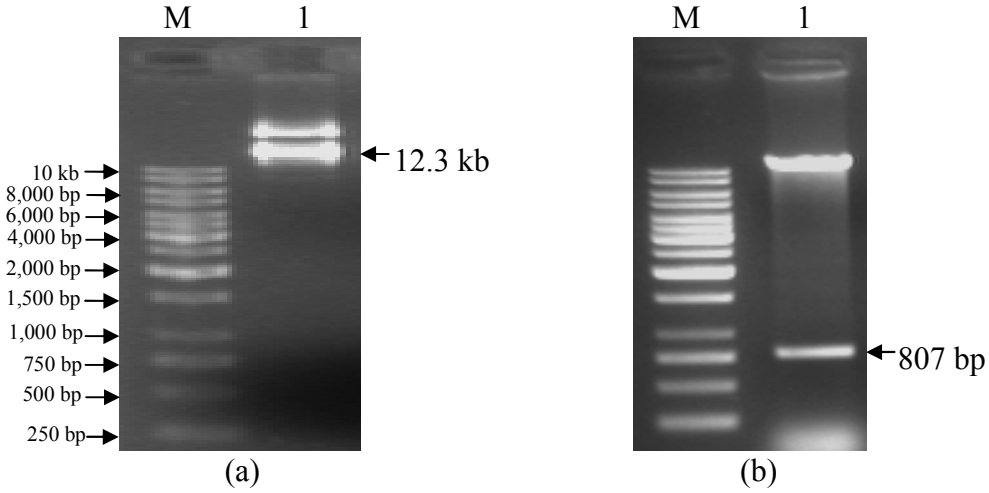


Figure 23 Agarose gel electrophoresis showed (a) pCAMBIA1304 (lane 1) and (b) pCAMBIA1304 were digested with *Nco* I and *Pst* I (lane 1). One kb ladder molecular weight markers (M) was shown on the left.

Three fragments of *CP-ACO I* and four fragments of *CP-ACO II* were ligated to digested pCAMBIA1304 producing CPACOI-SPP, CPACOI-LPP, CPACOI-DOFCORN, CPACOII-SP1, CPACOII-NO-SEboxN, CPACOII-SP3 and CPACOII-SP4, respectively. The position that *CPACO I* and *CPACO II* inserted in T-DNA of pCAMBIA1304 was shown in Figure 24.

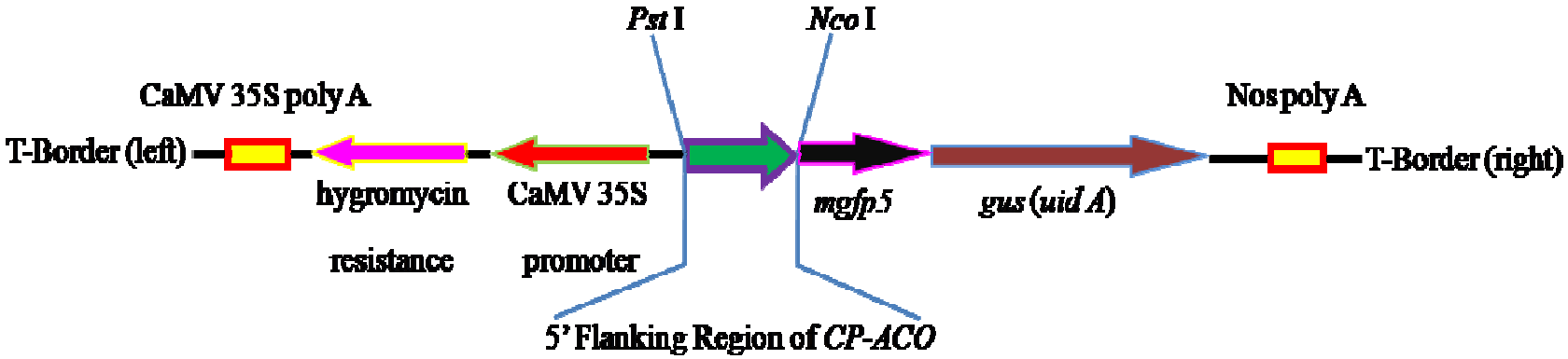


Figure 24 Schematic diagram of recombinant pCAMBIA1304 binary vector. Fragment of *CP-ACO I* and *II* 5'flanking regions were inserted in *Nco*I and *Pst*I recognition sites. Orientation of gene expression was displayed with arrows.

Seven constructs were transformed into *Agrobacterium tumefaciens* strain AGL-I competent cells by electroporation. Positive clones were checked by PCR with specific primers of each construct. The PCR products were determined by agarose gel electrophoresis (Figure 25 and 26).

For *CP-ACO I* 5'flanking region expression vector construction, the results indicated that clone CPACOI-SPP#1 and 2, CPACOI-LPP#1 in Figure 25a and CPACOI-DOFCORN#4, 5, 6 and 8 in Figure 25b were positive clones.

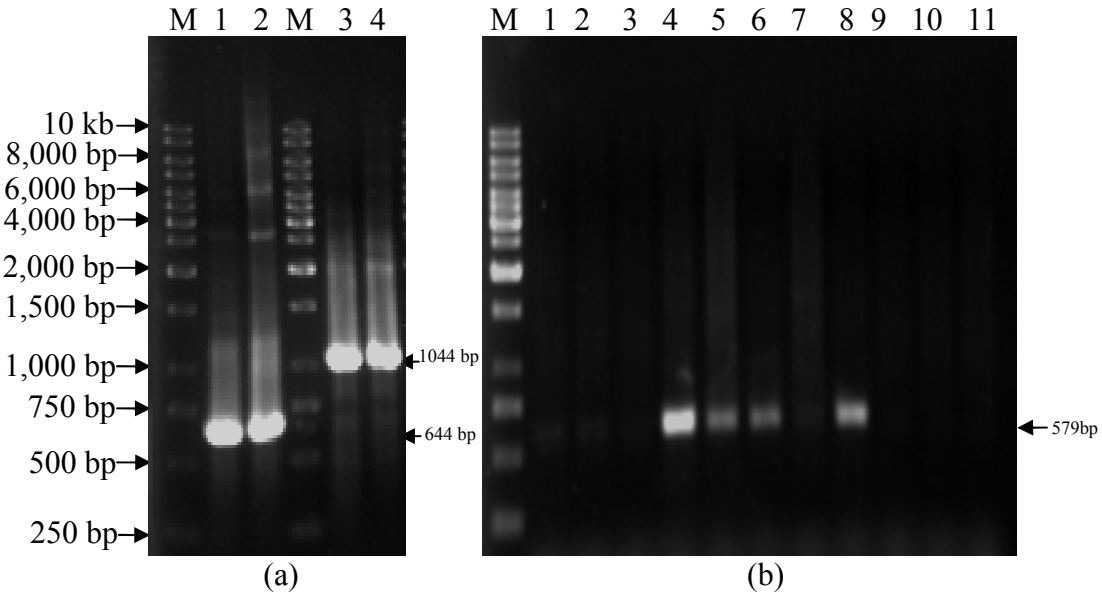


Figure 25 Agarose gel electrophoresis showed PCR products of clones (a) CPACOI-SPP#1 and 2 (lane 1 and 2) and CPACOI-LPP#1 and 2 (lane 3 and 4), (b) CPACOI-DOFCORN#1-11 (lane 1-11, respectively). One kb ladder molecular weight markers (M) was shown on the left.

For *CP-ACO II* 5'flanking region expression vector construction, the results indicated that clone CPACOII-SP1#1 and 2, CPACOII-NO-SEboxN#1 to 12, CPACOII-SP3#1 to 8 and CPACOII-SP4#1, 3 to 6 in Figure 26a-d were positive clones.

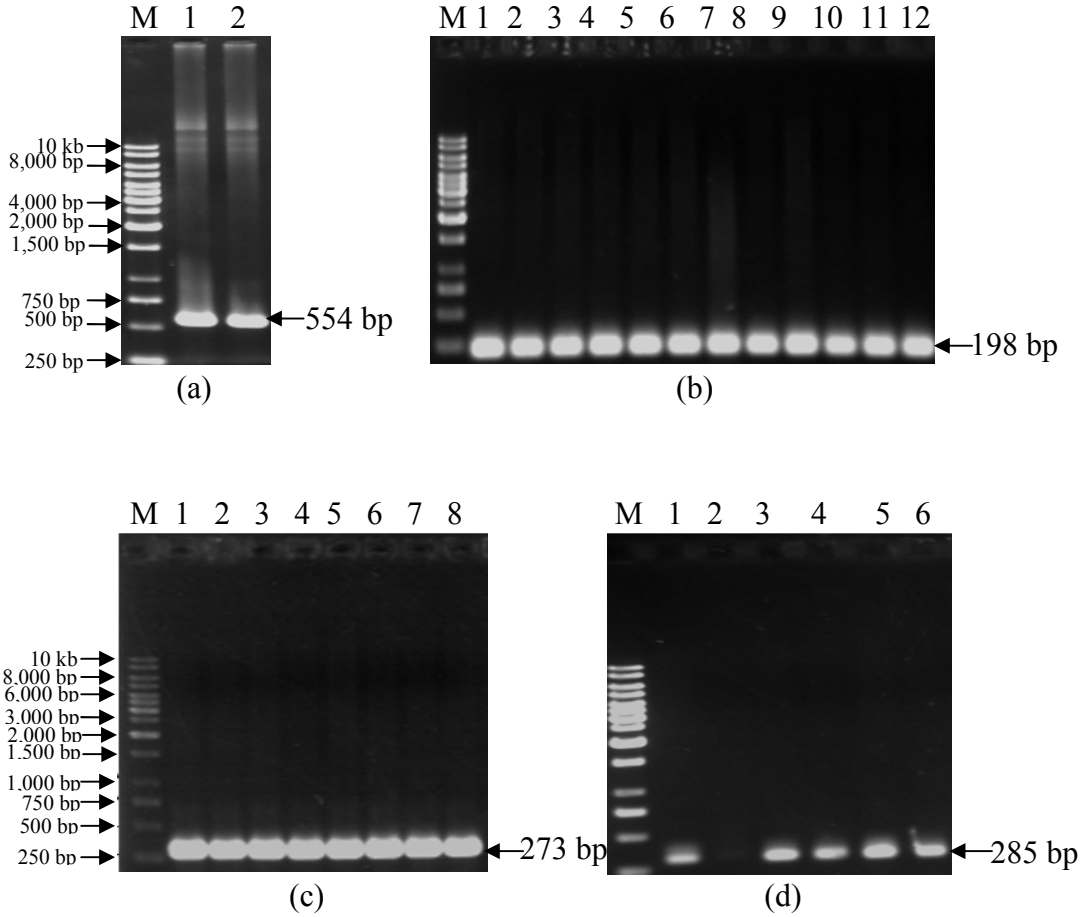


Figure 26 Agarose gel electrophoresis showed PCR products of clones (a)CPACOII-SP1#1 and 2 (lane 1 and 2), (b) CPACOII-NO-SEboxN#1 to 12 (lane 1 to 12), (c) CPACOII-SP3#1-8 (lane 1-8) and (d) CPACOII-SP4#1-6 (lane 1-6), respectively. One kb ladder molecular weight markers (M) was shown on the left.

4. *Agrobacterium*-Mediated Transient Expression in Roots

4.1 Determination of Promoter Activity of *CP-ACO I* and *II* Fragment DNA via *GUS* Histochemical Assay.

Five day old mung beans were incubated with *Agrobacterium* containing pCAMBIA1304, CPACOI-DOFCORN, CPACOII-NO-SEboxN and CPACOII-SP3. Mung bean roots were submerged in bacterial suspension at 25 ºC for 2 days under 16 h artificial light per day. After 2 days, there was no appearance of abnormality in plant growth. The mung bean plants were washed with distilled water and stained with X-Gluc solution. Only root tissues including primary root, lateral root and root hair were stained with X-Gluc solution. The presence and intensity of blue staining presented *GUS* activity in transiently transformed roots. Mung bean root transformed with *Agrobacterium* containing pCAMBIA1304, CPACOI-DOFCORN, CPACOII-NO-SEboxN and CPACOII-SP3 displayed strong blue staining indicating high *GUS* activity while roots transformed with *Agrobacterium* alone displayed clear to pale blue staining. The strength of *GUS* activity from 4 constructs was indistinguishable (Figure 27).

4.2 Determination of Location of *GUS* Expression in Transformed Mung Bean Roots via Agarose-Embedding Technique

Root samples from 4.1.1 were cut into 1 cm long from root cap and fixed in 5% agarose embedding block. After the gel hardening, the samples were sliced into 100 nm thick using microtome. The specimen was observed immediately under light microscope using 60X magnification. The cell types and their location were shown in Figure 28. The negative control, mung bean root transformed with *Agrobacterium* alone, displayed no *GUS* stanining in any part of the root. Roots transiently transformed with *Agrobacterium* containing pCAMBIA1304 displayed a strong GUS staining at root cap and elongation zone (Figure 29). Meristematic tissue, however, was only faintly stained. Vasculature, lateral root cap and columella root cap displayed deep blue color indicating very high *GUS* activity while cortex, endodermis

and epidermis were light blue with lower *GUS* activity. Transformed roots containing SP3 and DOFCORN displayed similar location of GUS staining to roots transiently transformed with *Agrobacterium* containing pCAMBIA1304. However, the intensity of blue staining was lower. Vasculature and columella root cap were stained with the noticeable absence of lateral root cap staining. Other cell types including cortex, endodermis and epidermis were also found to express *GUS*. Transformed roots with NO-SEboxN was found to express *GUS* strongly only at columella root cap. Other cell types, vasculature, were only faintly stained. Addition, transformed roots containing DOFCORN and SP3 showed low *GUS* activity in xylem whereas there was strong detectable *GUS* activity in roots transiently transformed with *Agrobacterium* containing pCAMBIA1304. However, transformed roots containing NO-SEboxN, *GUS* activity was considerable xylem specific expression.

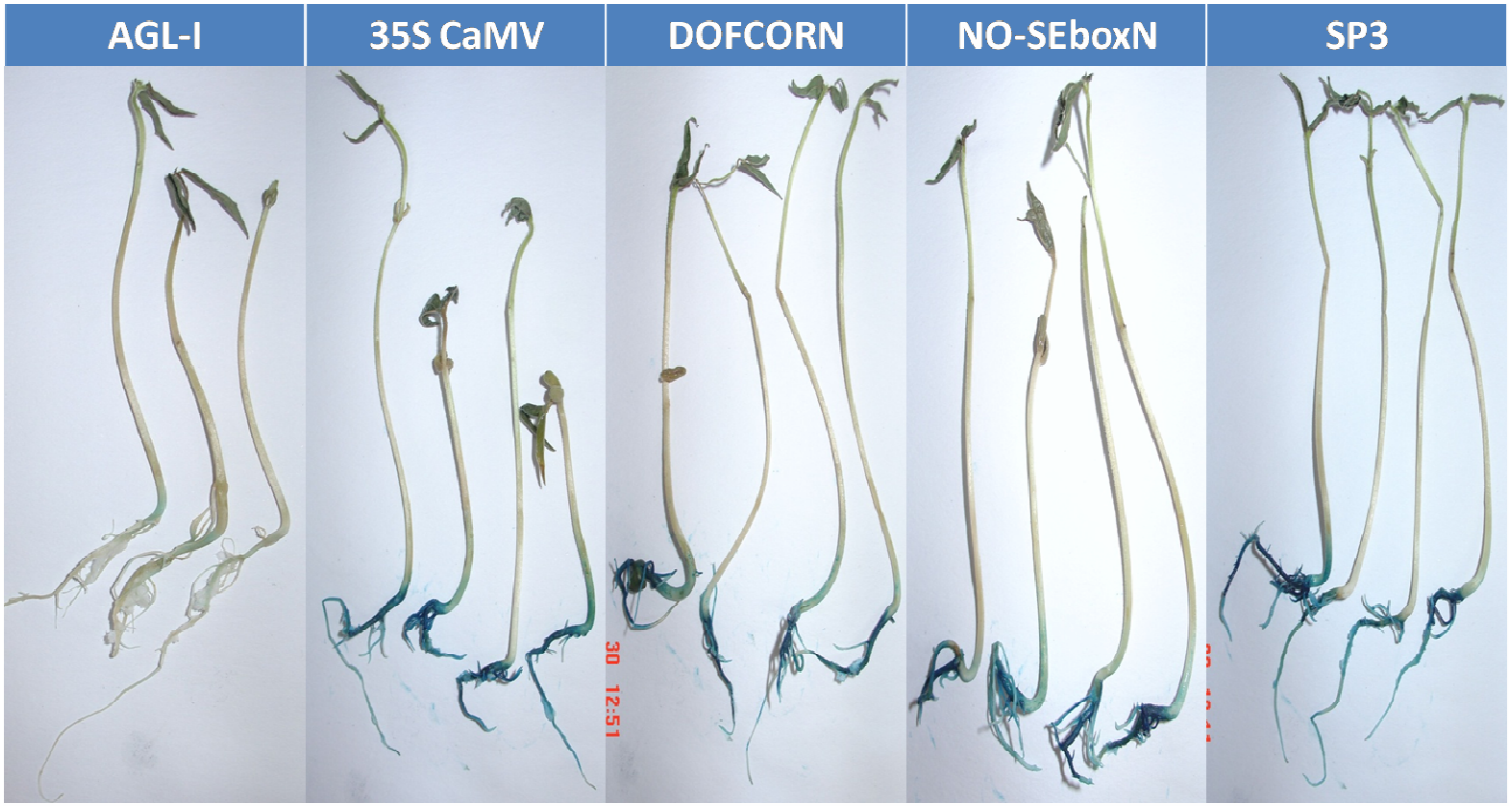
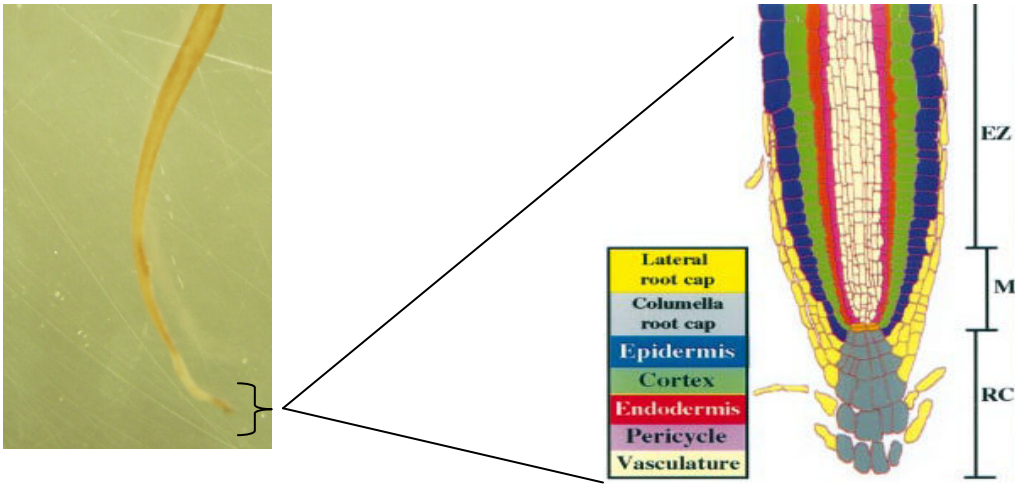
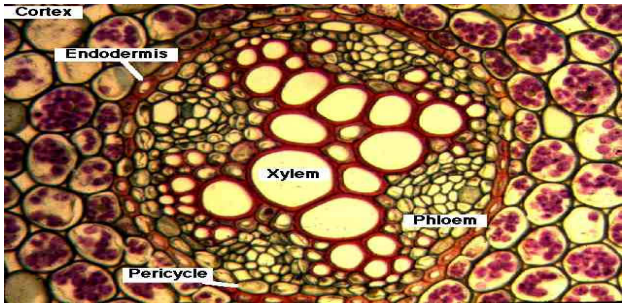


Figure 27 Transient expression of *GUS* in mung bean roots. Seedlings were submerged for 2 days (light for 16 h/day) with *Agrobacterium* without vector, *Agrobacterium* containing pCAMBIA1304, CPACOI-DOFCORN, CPACOII-NO-SEboxN and CPACOII-SP3.



(a) (b)



(c)

Figure 28 (a) Exterior view of developing root of mung bean. (b) Schematic illustration of root apical tissues delineating root cap (RC), meristematic (M) and elongation zone (EZ) and (c) Transverse section of dicot plant was shown composition of tissue

Source: Lemke (1999)

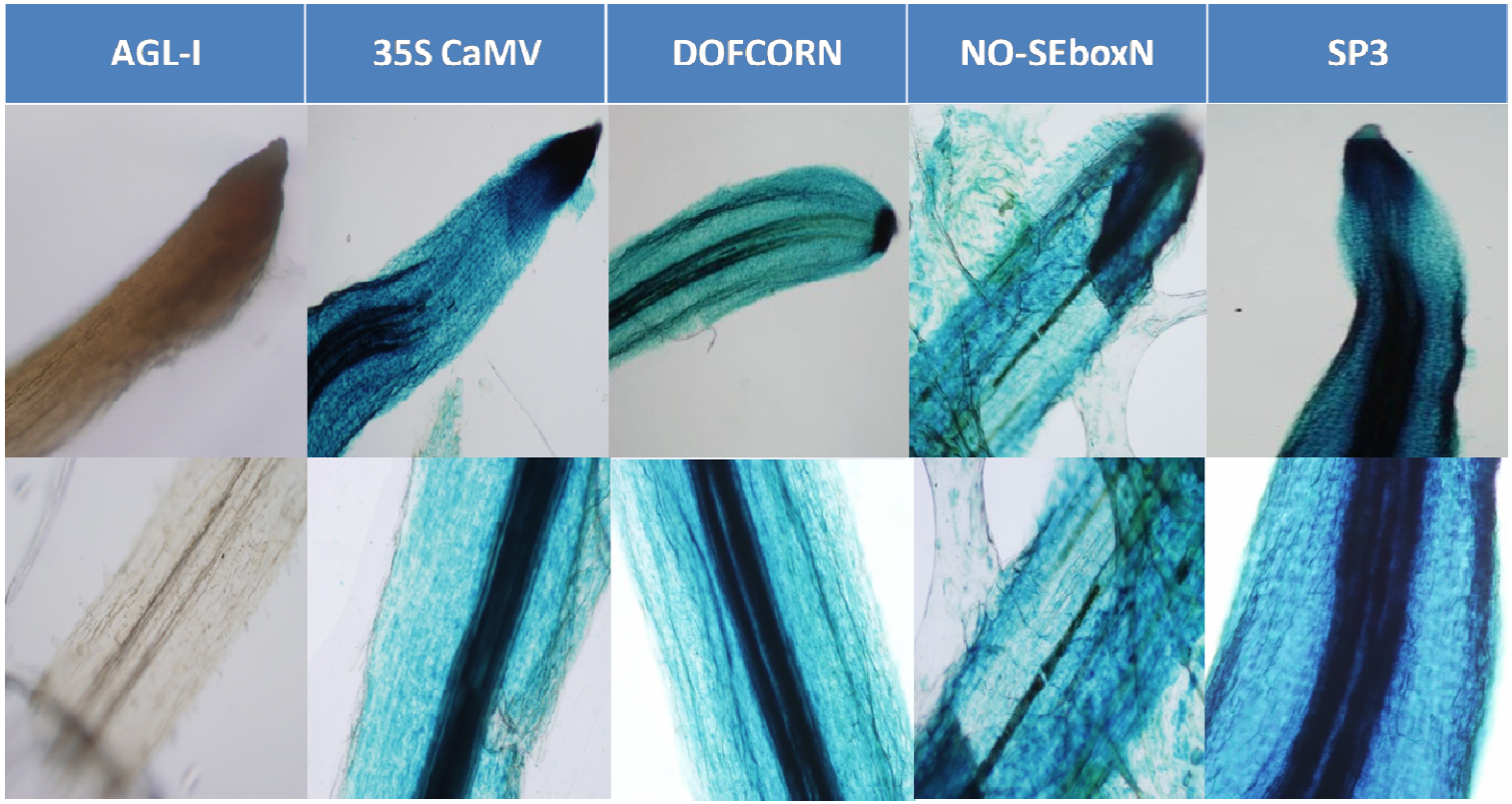


Figure 29 Transient expression of *GUS* in mung bean roots. Seedlings were submerged for 2 days (light for 16 h/day) with *Agrobacterium* without vector, *Agrobacterium* containing pCAMBIA1304, CPACOI-DOFCORN, CPACOII-NO-SEboxN and CPACOII-SP3. In circle were shown xylem tissue.

4.3 Determination of Promoter Activity of *CP-ACO I* and *II* Fragment DNA via *GUS* Enzymatic Assay

Transformed mung bean roots were washed with dH_2O. Three roots were extracted in chilled *GUS* extraction buffer. The total protein was determined using SDS-PAGE. *GUS* activity assay was determined with a Fluor AceTM β-glucuronidase reporter assay kit.

4.3.1 Protein Standard Curve

Bovine Serum Albumin (BSA) at 0-10 µg/µl was used as a standard protein (Appendix Table 10). The protein concentration was plotted against standard absorbance at O.D. $_{595}$ (Figure 30). The formula for protein concentration was done by Excel program. This formula was used for the determination of total protein concentration from root tissues.

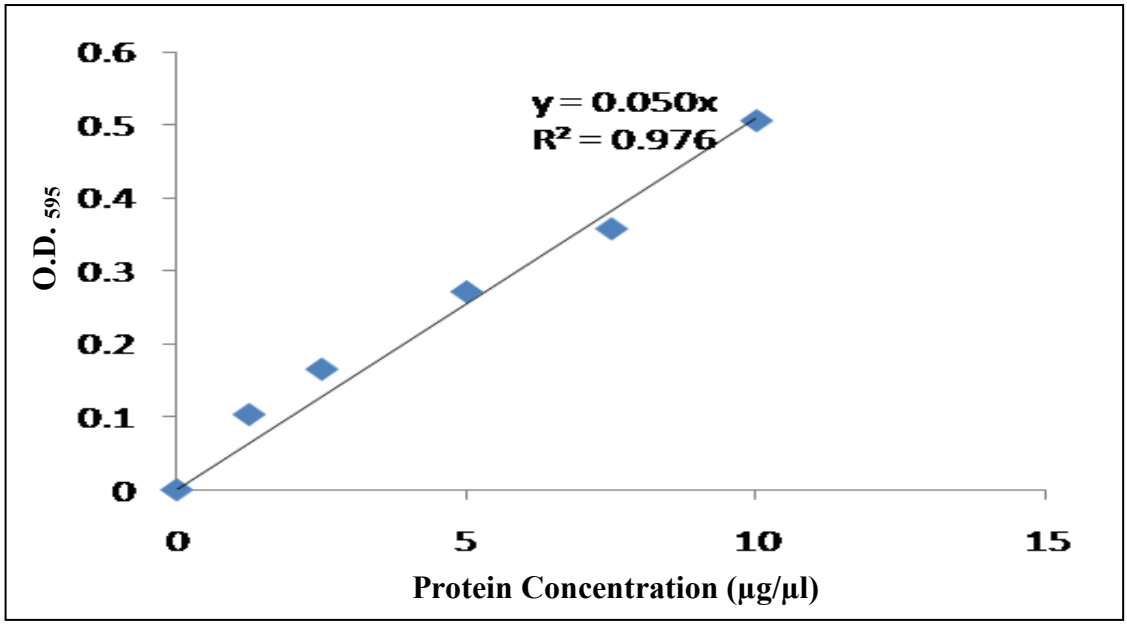


Figure 30 Bradford protein assay of Bovine Serum Albumin (BSA) standard protein. BSA concentration was between 0-10 µg/µl. X axis was shown concentration of protein (µg/µl). Y axis was shown optical density values at 595 wavelength of protein.

Table 14 Protein concentration (µg/µl) of transformed mung bean root extract calculated by standard equation.

Sample	O.D. $_{595}$	Protein Concentration (µg/µl)	50 µg of Total Protein
AGL-I	0.120	2.3576	12.71 µl
pCAMBIA1304	0.093	1.8271	16.41 µl
DOFCORN	0.097	1.9057	15.71 µl
NO-SEboxN	0.124	2.4361	12.31 µl
SP3	0.065	1.2770	23.49 µl

Fifty micrograms of total protein from root extract per each sample was loaded into SDS-Polyacrylamide gel (Table 14). After 1 h, the gel was stained with silver staining solution (Plusone silver staining kit). Sixty eight kDa protein, GUS protein molecular weight, was clearly seen in transformed root samples containing pCAMBIA1304, CP-ACOII-NO-SEboxN and CP-ACOII-SP3 (lane 2, 4 and 5 respectively). Such protein was absent in untransformed roots (lane 1) and transformed root containing CP-ACOII-DOFCORN (Figure 31).

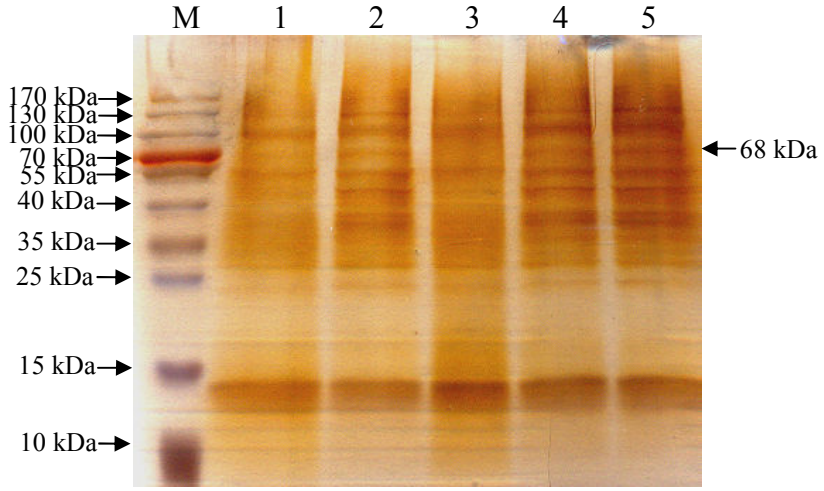


Figure 31 SDS-Polyacrylamide gel showed total protein extracted from infiltrated root samples; untransformed root, transformed roots with pCAMBIA 1304, DOFCORN, NO-SEboxN and SP3 (lane 1-8, respectively). One kb ladder molecular weight markers (M) was shown on the left.

4.3.2 *GUS* Enzymatic Assay

GUS activity assay was determined with the Fluor AceTM β-glucuronidase reporter assay kit. The fluorescence of 4-MU was measured with the excitation at 365 nm and emission at 455 nm with spectofluorometer (Appendix Table 11). A standard curve was made from a series of dilution (1:10, 1:100 and 1:1,000 pM) of 4-MU as shown in Figure 32. *GUS* activity expressed as fluorescence unit pmol MU per µg protein per min.

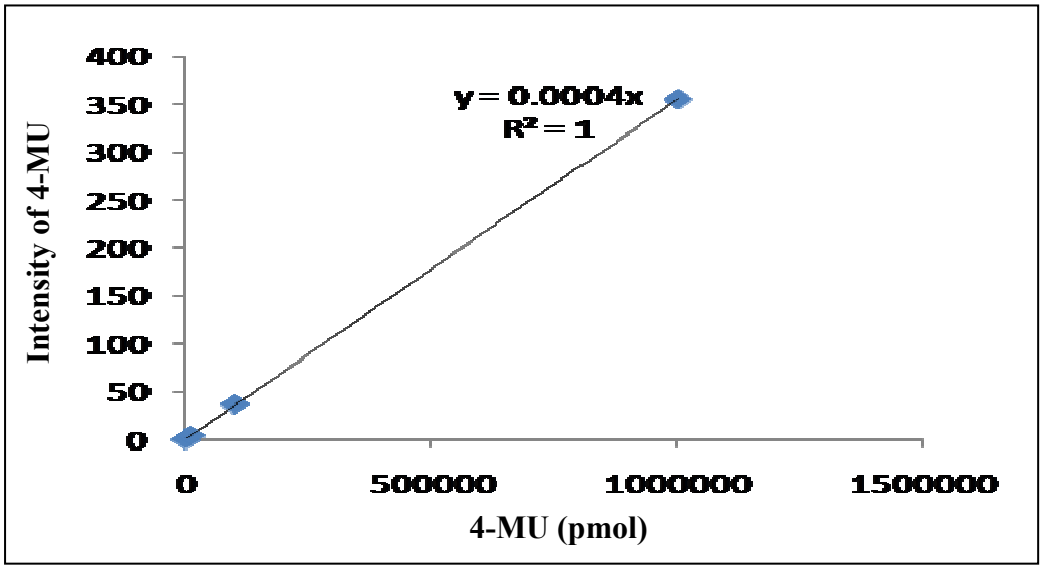


Figure 32 Standard curve of Relative Fluorescence Unit of 4-MU. 4-MU concentration for triple sample was determined by comparing with 4-MU standard curve. X axis represented concentration of 4-MU (pmol) while Y axis represented intensity value of 4-MU.

Quantitative *GUS* enzymatic activity was performed with 3 constructs (DOFCORN, NO-SEboxN and SP3) that induced *GUS* expression in transformed mung bean roots. The results indicated that *GUS* enzymatic activity was highest at in transformed roots with *Agrobacterium* containing pCAMBIA1304 (positive control) at 203.4 pmol MU/min/µg (Table 15). Among *CP-ACO* promoter driven constructs, DOFCORN had the highest *GUS* expression level at 10.8 pmol MU/min/µg followed by SP3 at 7.3 pmol MU/min/µg. The statistical analysis (T-test), however, indicated that there was no significant difference between these two experiments. This outcome was supported by the observation of similar GUS staining in various cell types in roots under light microscope of mung bean roots transformed with these 2 constructs. The roots transformed with NO-SEboxN, on the other hand, displayed lowest *GUS* enzymatic activity at 2.2 pmol MU/min/µg with very low GUS staining in most root cells (Figure 33).

Table 15 *GUS* enzymatic activity of mung bean roots transiently transformed with *Agrobacterium* without vector, *Agrobacterium* containing pCAMBIA 1304, CPACOI-DOFCORN, CPACOII-NO-SEboxN and CPACOII-SP3 constructs. Each sample was done in triplicate. Positive samples were provided by the kit. The T-test at p = 0.05 was used to compared *GUS* enzymatic activity between each sample.

Samples	**Intensity of Sample***	**4-MU (pmol)**	***GUS* Activity (pmolMU/min/µg)**	***GUS* Activity (Average)**
AGL-I	0	0	0	0
	0	0	0	
	0	0	0	
pCAMBIA1304 (35S CaMV)	209.16	522901.75	197.85	203.50[a]**
	212.26	530654.50	200.79	
	223.95	559874.25	211.84	
DOFCORN	12.14	30361.25	11.48	10.84[b]
	10.39	25970.75	9.82	
	11.86	29656.75	11.21	
NO-SEboxN	3.07	7666.75	2.89	2.25[c]
	0.61	1522.50	0.57	
	3.49	8720.50	3.28	
SP3	5.96	14912.00	5.69	7.28[b]
	8.46	21152.75	8.08	
	8.47	21167.75	8.08	

* Intensity of samples were calibrated with AGL-I (negative).

** *GUS* activity were grouped with statistical analysis (T-test) by Excel program.

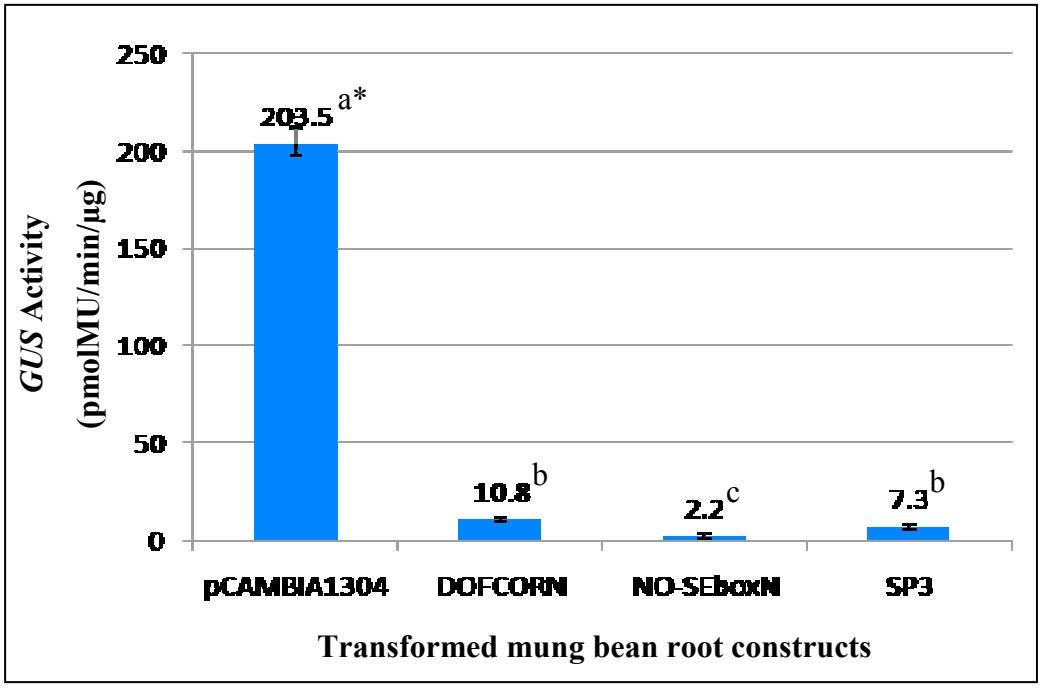


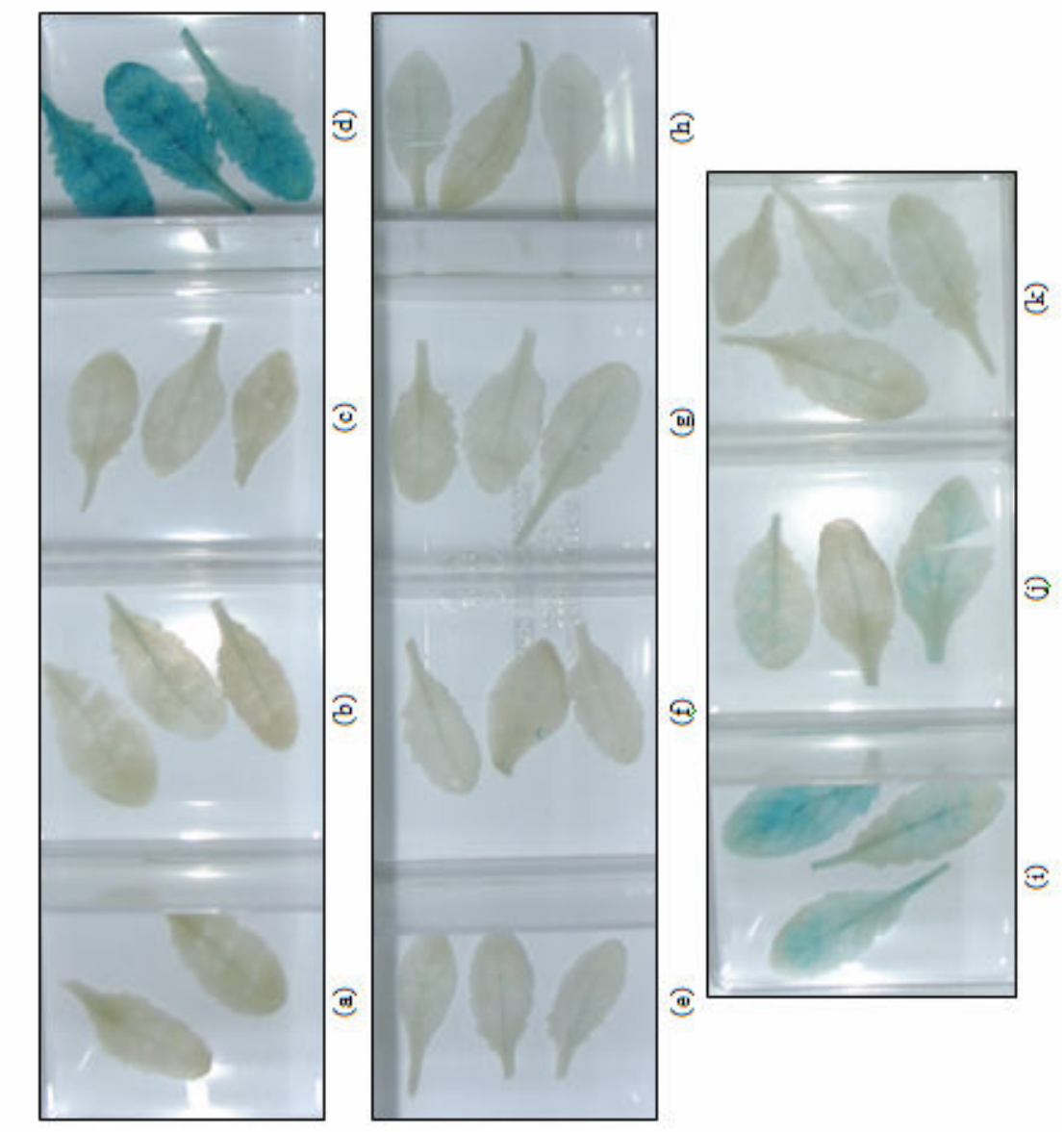
Figure 33 Comparison of *GUS* activity in transformed mung bean root with *Agrobacterium* containing pCAMBIA1304, CPACOI-DOFCORN, CPACOII-NO-SEboxN and CPACOII-SP3. An average of *GUS* activity from each sample was indicated on the top of each bar. * *GUS* activity were grouped with statistical analysis (T-test) by Excel program.

5. *Agrobacterium*-Mediated Transient Expression in Leaves

Two weeks old Arabidopsis leaves were transiently transformed by *Agrobacterium* mediated infiltration. Seven constructs with *CP-ACO* promoter driven *GUS* constructs and pCAMBIA1304 were used. *Agrobacterium* without vector, distilled water and MS medium were used as negative controls. After 2 days, the transformed leaves were stained with X-Gluc solution and destained with 95% ethanol.

The results indicated that only Arabidopsis leaves transformed with pCAMBIA1304 and CPACOII-NO-SEboxN displayed blue colour of GUS staining (Figure 34). Upon 20X magnification of light microscope, the leaves transformed with pCAMBIA1304 displayed strong GUS staining at wounding areas, midrib, veins and trichomes while only wounding areas, midrib and veins were GUS stained in CPACOII-NO-SEbxN transformed leaves (Figure 35).

Figure 34 Transient expression of *GUS* in expanding Arabidopsis leaves. Two week old leaves were infiltrated for 2 days (light for 16 h/day) with (a) dH_2O, (b) MS medium, (c) *Agrobacterium* without vector, (d) *Agrobacterium* containing pCAMBIA1304, (e) CPACOI-SPP, (f) CPACOI -LPP, (g) CPACOI -DOFCORN, (h) CPACOII-SP1, (i) CPACOII-NO-SEboxN, (j) CPACOII-SP3 and (k) CPACOII -SP4.


(a)
(b)
(c)
(d)
(e)
(f)
(g)
(h)
(i)
(j)
(k)

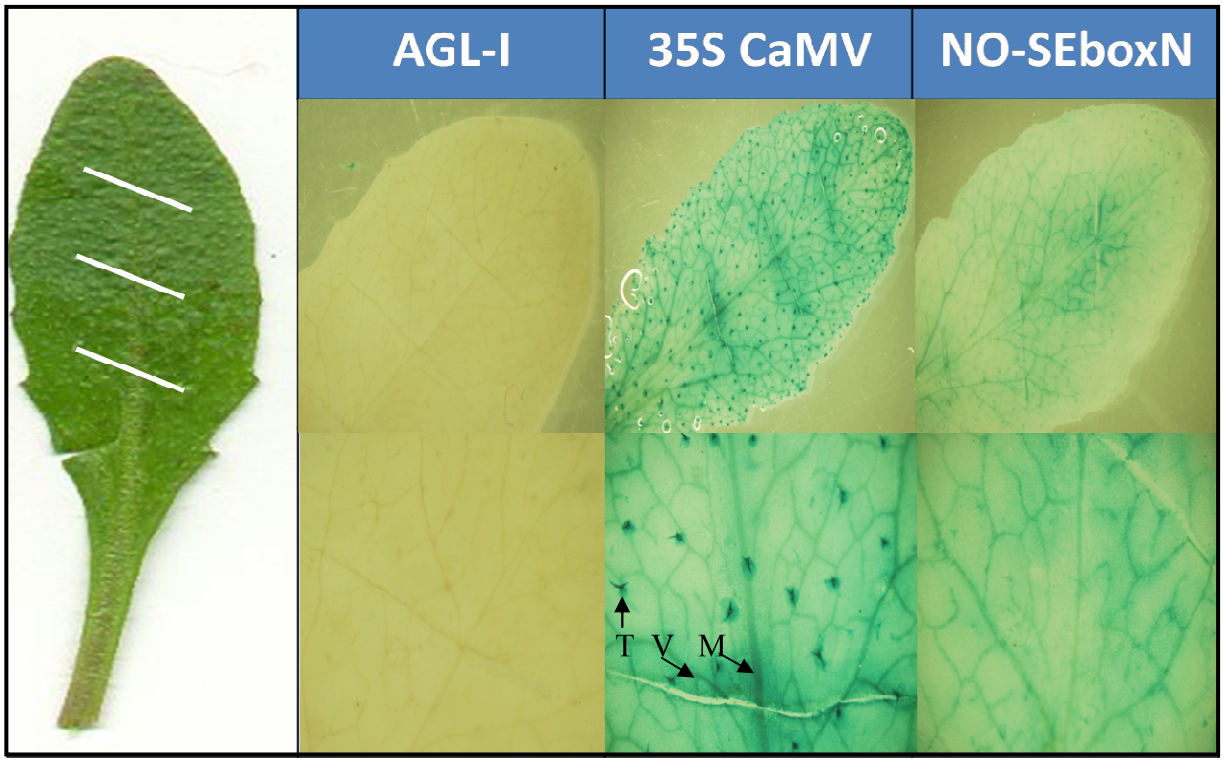


Figure 35 Transient expression of *GUS* in expanding Arabidopsis leaves. The left panel indicated leaves were made three partially transverse and equidistant cut through the midrib from the petiole end to the leaf tip. Two weeks old leaves were infiltrated for 2 days (light for 16 h/day) with *Agrobacterium* without vector, *Agrobacterium* containing pCAMBIA1304 and CPACOII-NO-SEboxN. Under 20X magnification of light microscope. Abbreviations, T: trichomes, V: vein and M: midrib.

6. *Agrobacterium*-Mediated Transient Expression in Flowers

White petunia hybrid flowers were transiently transformed by *Agrobacterium* mediated infiltration. After 2 days incubation, the flowers were stained with x-Gluc solution to determine *GUS* expression. Flowers transiently transformed with *Agrobacterium* alone, distilled water or glucose were clear indicating the absence of *GUS* expression (Figure 36). Flower transiently transformed with *Agrobacterium* containing pCAMBIA1304 displayed strong GUS staining. All seven constructs induced various degree of *GUS* expression in transformed flowers (Figure 36 and 37). Flowers transiently transformed with *Agrobacterium* containing CPACOI-SPP and CPACOII-NO-SEboxN, however, displayed strong *GUS* expression. Upon a closer look on petal, it was revealed that the reporter gene in pCAMBIA1304 *Agrobacterium* infiltration was expressed in all cell types. *GUS* signal was not observed in all negative control and other constructs. Flower transiently transformed with *Agrobacterium* containing CPACOII-NO-SEboxN displayed an identical *GUS* expression pattern to positive control (Figure 37). Although, SPP induced weaker *GUS* expression in flower, it exhibited strong vascular tissue specific (Figure 36).

Under 20X magnification of light microscope, pCAMBIA1304 transformed flowers showed strong GUS staining in ovary and receptacle. CPACOII-NO-SEboxN transformed flowers also displayed GUS staining in ovary and receptacle. *GUS* expression in receptacle was also found in CPACOII-SP4 transformed flowers (Figure 38).

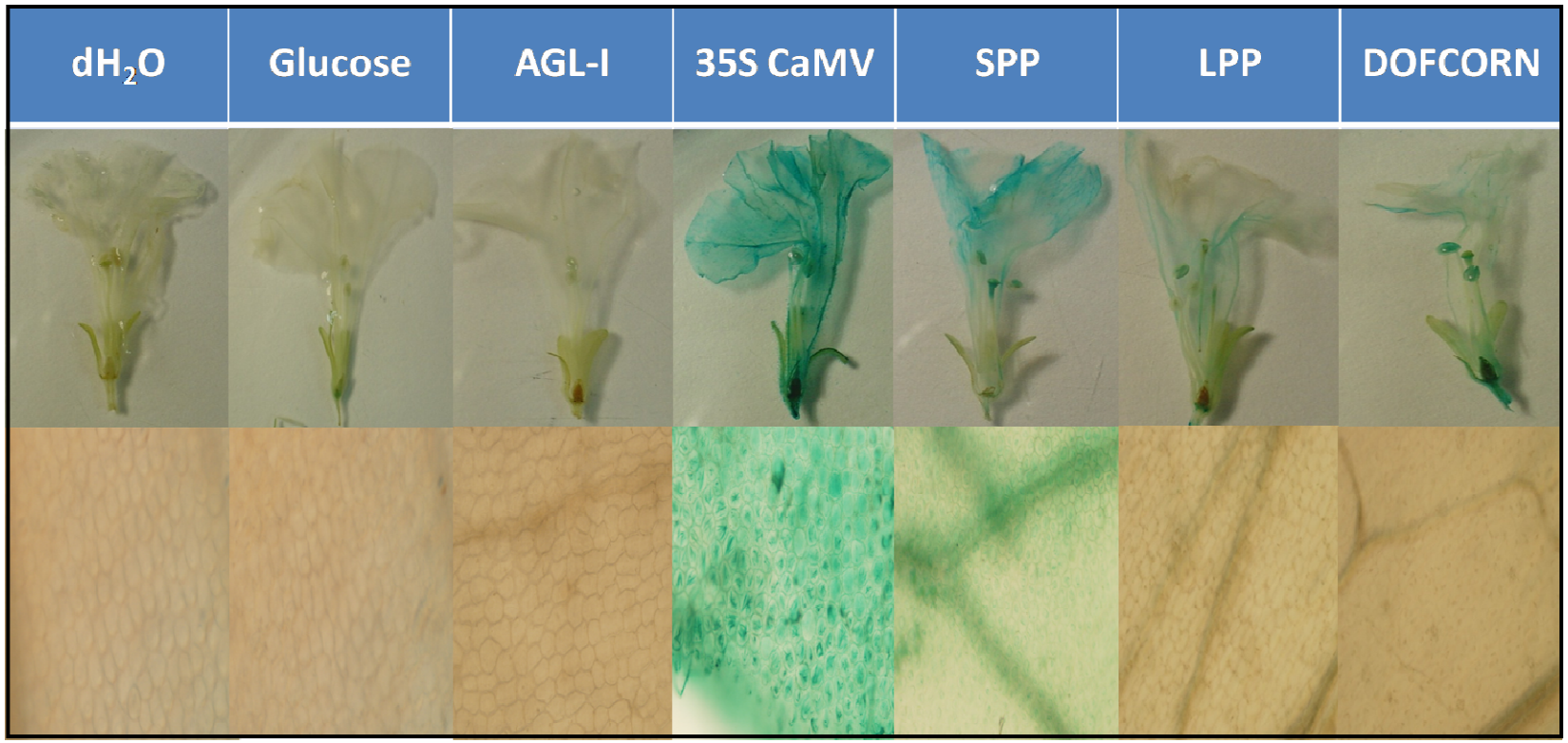


Figure 36 Transient *GUS* expression in white petunia hybrid flowers. The top panel indicated 4 constructs used in each experiment (*Agrobacterium* containing pCAMBIA1304, CPACOI-SPP, CPACOI-LPP and CPACOI-DOFCORN) and three negative control; dH_2O, glucose and *Agrobacterium* without vector. The middle panel was transiently transformed whole flowers. The bottom panel was flowers petals at 60X magnification.

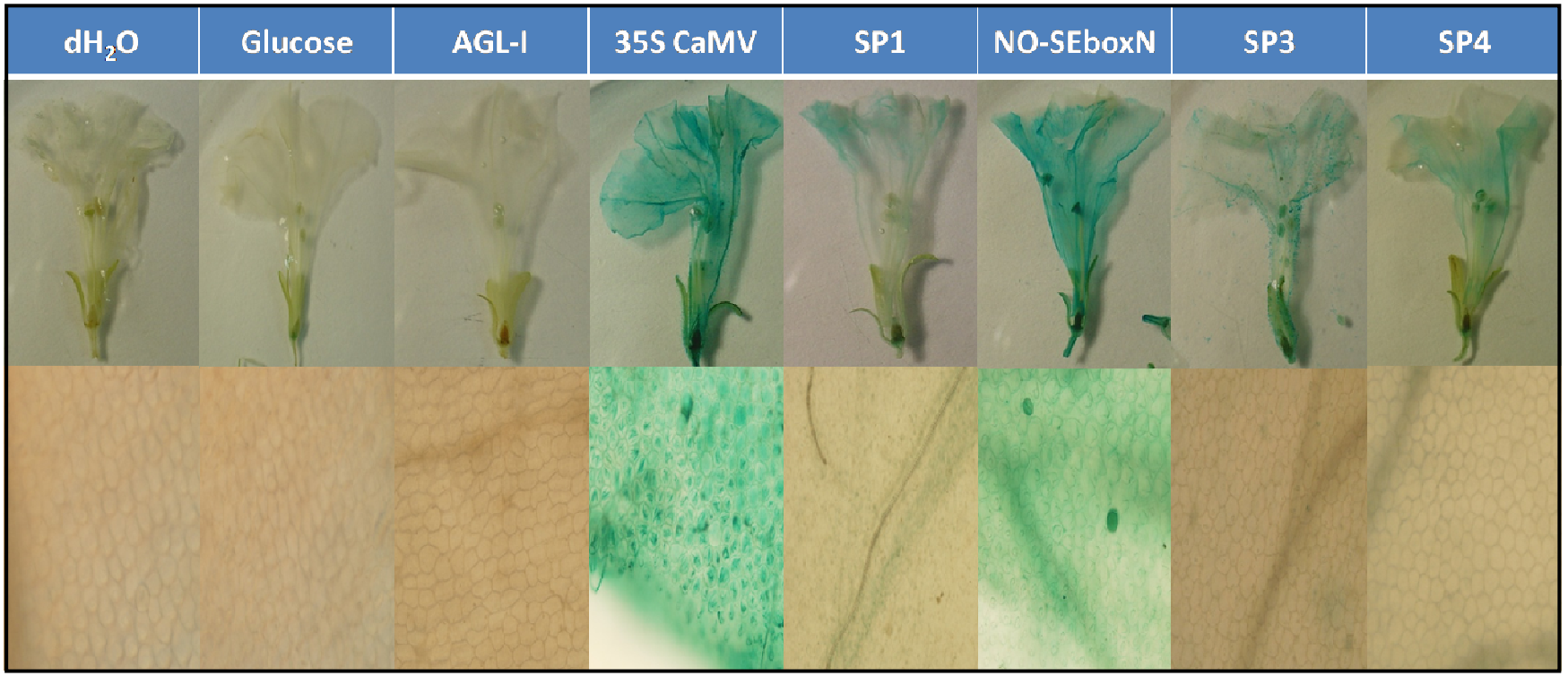


Figure 37 Transient *GUS* expression in white petunia hybrid flowers. The top panel indicated 5 constructs used in each experiment (*Agrobacterium* containing pCAMBIA1304, CPACOII-SP1, CPACOII- NO-SEboxN, CPACOII-SP3 and CPACOII-SP4) and three negative control; dH_2O, glucose and *Agrobacterium* without vector. The middle panel was transiently transformed whole flowers. The bottom panel was flowers petals at 60X magnification.

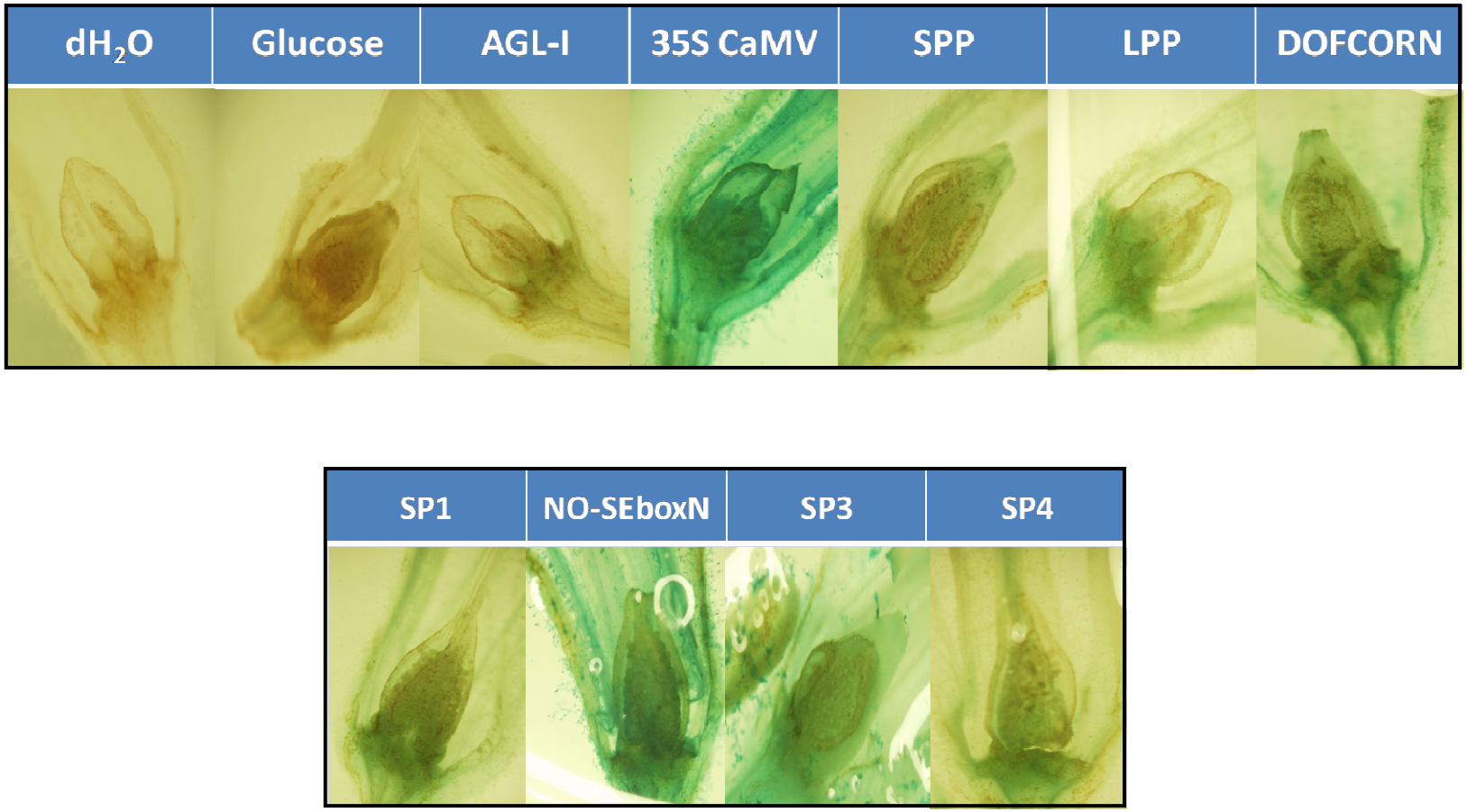


Figure 38 Transient expression of *GUS* in ovaries of petunia hybrid flowers. Ovaries were studied using stereo microscope at 20X magnification.

7. *Agrobacterium*-Mediated Transient Expression in Fruits

Tomato fruits at 4 developmental stages; immature, mature, colour break and ripen (Figure 39), were transiently transformed with *Agrobacterium* infiltration. After 2 days, the tomatoes were sliced and stained to determine *GUS* expression for 2 h at 37 ºC. The results showed that tomatoes transformed with distilled water, MS medium and *Agrobacterium* alone displayed no *GUS* staining in all four developmental stages (Figure 40 and 41). In positive controls, *GUS* expression was observed in all stages. Vascular bundle and columella were clearly blue stained.

In mature stage, all 7 constructs induced *GUS* expression. There was, however, variation in the intensity of *GUS* staining among the constructs. DORFCORN, SP3 and SP4 constructs induced faint blue staining while there was an absence in *GUS* expression in columella area with SPP, LPP and SP1 constructs. Tomato transformed with NO-SEboxN construct displayed distinct *GUS* expression in columella (Figure 40 and 41).

All seven constructs induced *GUS* expression in colour break stage. Columella, vascular bundle and placenta tissue were clearly blue stained (Figure 42). *GUS* expression was also detected in ripen tomato using all the constructs. However, *GUS* expression was absent in placental tissue using SP4 construct. None of the *ACO* promoter constructs induced any *GUS* activity in immature stage.

Figure 39 Developmental stages of tomato fruit; 1= immature, 2= mature, 3= color break and 4= ripen.

Figure 40 Transient expression of *GUS* in 4 developmental stages of tomato fruits; (a) immature, (b) mature, (c) color break and (d) ripen. Fruits were infiltrated with dH_2O, MS medium, *Agrobacterium* without vector, *Agrobacterium* containing pCAMBIA1304, CPACOI-SPP, CPACOI-LPP and CPACOI-DOFCORN.

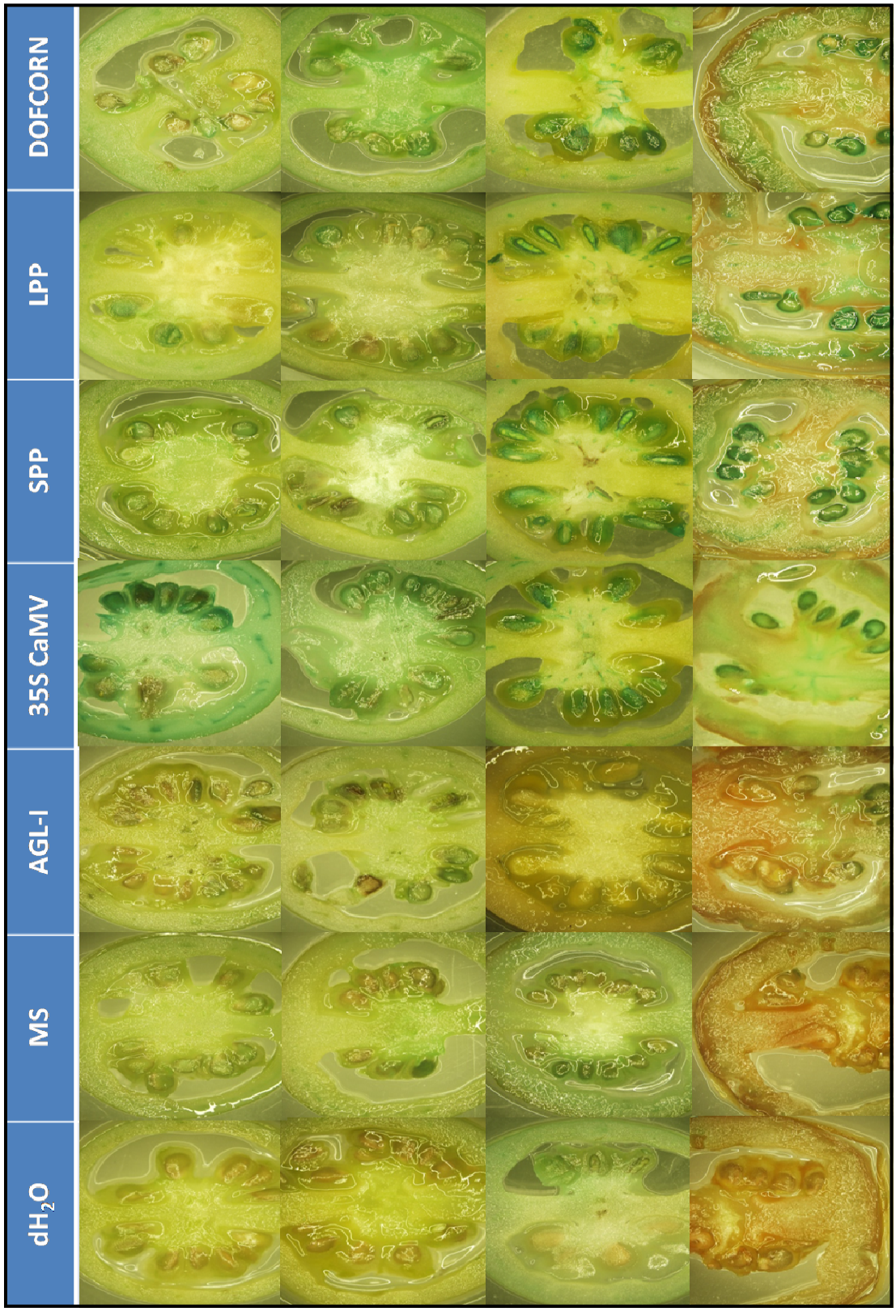
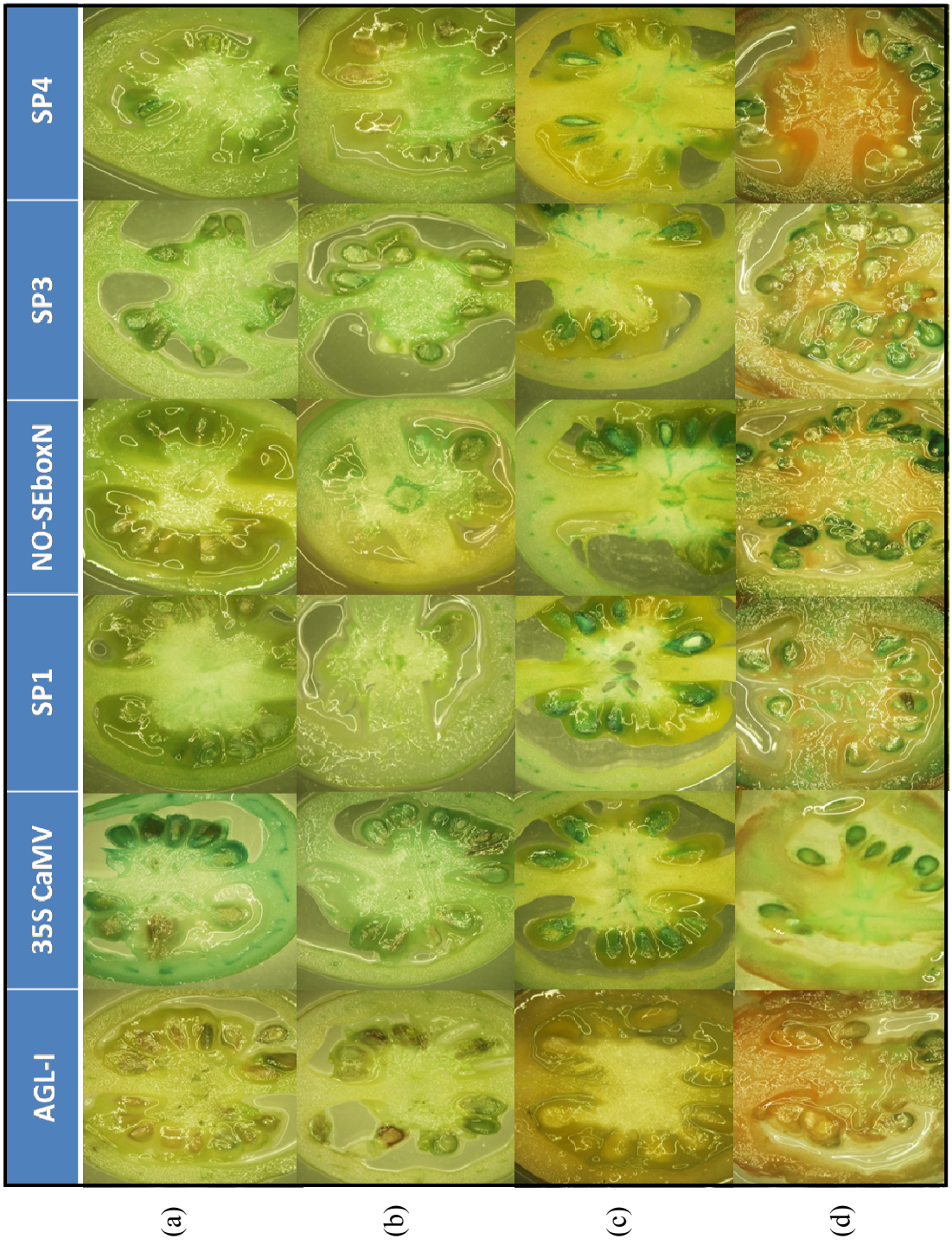

dH2O
MS
AGL-I
35S CaMV
SPP
LPP
DOFCORN
(a)
(b)
(c)
(d)

Figure 41 Transient expression of *GUS* in 4 developmental stages of tomato fruits; (a) immature, (b) mature, (c) color break and (d) ripen. Fruits were infiltrated with *Agrobacterium* without vector, *Agrobacterium* containing pCAMBIA1304, CPACOII-SP1, CPACOII-NO-SEboxN, CPACOII-SP3 and CPACOII-SP4.


AGL-I
35S CaMV
SP1
NO-SEboxN
SP3
SP4
(a)
(b)
(c)
(d)

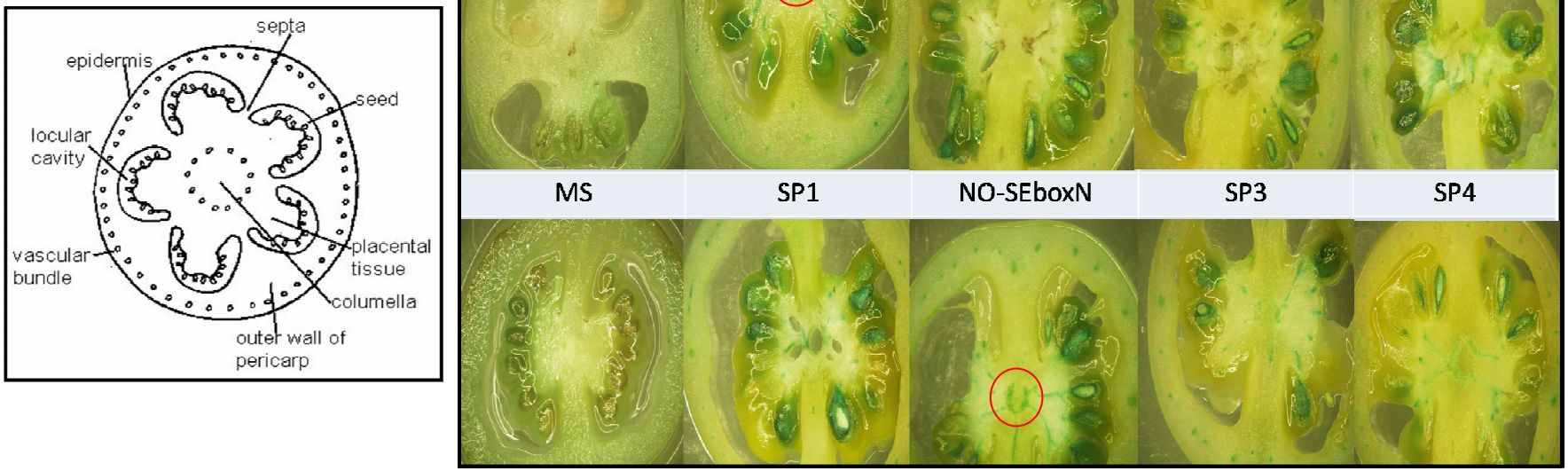


Figure 42 Transient expression of *GUS* in colour break stage of tomato fruits. Fruits were infiltrated with *Agrobacterium* containg pCAMBIA1304, CPACOI-SPP, CPACOI-LPP, CPACOI-DOFCORN, CPACOII-SP1, CPACOII-NO-SEboxN, CPACOII -SP3 and CPACOII -SP4 were shown strong *GUS* expression in all tissues. In circle were shown columella tissue.

Source: Rost (1996)

DISCUSSION

1. Identification and characterization of *CP-ACO I* and *CP-ACO II* 5'flanking region

In this study, identification of DNA sequences flanking T-DNA insertions was done by ligation-mediated PCR (Cottage *et al*., 2001). This method was applied by combining 'vectorette' (Lagerstrom, 1991) and 'suppression PCR' (Sibert *et al*., 1995). Concentration of first PCR products affected the quality of second PCR products. This could be solved by adjusting PCR product dilutions and the use of nested primers. The 591 bp *CP-ACO II* 5'flanking region (including 5'UTR) was cloned from *Pdm*I and *Hpa*I adaptor libraries. This region is likely covered most of promoter region of *CP-ACO II* gene because promoter regions are usually located within 300 bp upstream from translation start site (Klug and Cummings, 1997). Sequence comparison (BLASTN) indicated that *CP-ACO II* 5'flanking region had low similarity to sequences in Genbank Database. Percent identity of *CP-ACO II* 5'flanking region and other plant *ACO* promoter regions was between 31.1% (*Musa acuminata, ACO*) to 44.5% (*CP-ACO I*) using Megalign program (Lasergene). However, there was highly similarity region between 1-34 bp upstream from 5'UTR of *CP-ACO II* 5'flanking region and *ACO* mature RNA region (L76283).

2. Identification of *cis*-acting elements by database-assisted bioinformatics approach

Cis-acting elements presented in *CP-ACO I* and *CP-ACO II* 5'flanking region were predicted via 3 databases PLACE, PlantCARE and Plant Prom database. PLACE and PlantCARE databases were easily accessible and references were included. In addition, the result could be linked to other information within the database and other databases such as EMBL, GenBank, TRANSFAC and MEDLINE (Higo *et al*., 1999; Lescot *et al*., 2002). Although containing more 1400 regulatory elements, the accessibility of Plant Prom database was limited. The output provided regulatory element, organism and gene without references.

Analysis of *CP-ACO I* and *CP-ACO II* 5'flanking regions via three databases provided similar function groups of elements such as light responsive, hormonal responsive and seed and endosperm specific. However, PLACE detected more elements and their repeats in each functional group. The locations of elements were concentrated in specific areas. For example, DOFCOREZM (AAAG) element was found mainly between 1-579 bp upstream from the translation start site of *CP-ACO I* 5'flanking region. This element is a binding site for Dof protein family which activated the endosperm specific gene expression (Yanagisawa and Schmidt, 1999).

3. *Agrobacterium* transient expression in leaves, root, flower and fruit tissues

Agrobacterium transient expression was used as a tool to determine *CP-ACO I* and *II* promoter plant tissue specificity in four important plant organs including leaves, flowers, fruits and roots. Although similar methods were reported, our study had made some modification including the use of vacuum and a simple set up of transient system that will allow a more routine and repeatable protocol (Jing *et al*., 2006; Yongjin *et al*., 2007 and Orzaez, *et al*., 2006). In root transformation, the active absorption of *Agrobacterium* is needed therefore alive mung bean plants were used (Ray *et al*., 2005). The rest of the experiment was carried out using detached plant organs. The presence of blue staining resulting from *GUS* enzymatic activity allowed an easy way to detect *GUS* expression in several plant cell types including vasculature (xylem and phloem), midrib, vein and trichomes either through naked eyes or compound microscope.

pCAMBIA1304 binary vector was used as a positive control and for construction of new expression vector via replacing 35S promoter with *ACO* 5'flanking regions. It was proven to express well in every *Agrobacterium* transient expression system (root, leave, flower and fruit) tested. *GUS* expression was detected in every cell types while negative controls such as MS medium and glucose were shown to be negative (clear) in every experiments.

Seven expression constructs with *CP-ACO I* (SPP, LPP and DOFCORN constructs) and *CP-ACO II* (SP1, SP3, SP4 and NO-SEboxN constructs) promoters were generated. The location and size of each promoter regions were designed according to the presence of important *cis*-acting elements. LPP (-478 to +147 bp) and SPP (-878 to +147 bp) from *CP-ACO I* 5'flanking region were previously studied and reported to be fruit specific promoter (Chuaboonmee *et al*., 2004). In this study, SPP was also shown to regulate *GUS* expression in petal. DOFCORN region (-579 bp to +1 bp) contained mostly DOFCOREZM (AAAG) element. This element was commonly reported as seed and endosperm specific (Yanagisawa and Schmidt, 1999). In flower transient expression study, DOFCORN region was also induced expression in flower receptacle. Studies indicated that adjacent elements to DOFCOREZM could also play important roles in plant tissue specific expression (Yanakisawa *et al*., 1999).

SP1 was an entire region obtained from *CP-ACO II* 5'flanking region isolation. The SP1 region was divided into 2 fragments; SP3 and SP4. NO-SEboxN region was a subset of SP3 without seed and endosperm specific elements. NO-SEboxN region induced *GUS* expression in all plant tissues except in immature fruit, trichomes and vascular bundle in root. *GUS* expression in columella of fruit, was increased as ripening process started (a colour break stage). This result was related to *ACO* expression during fruit ripening (Kumdee *et al*., 2003). Two constructs with *CP-ACO II* promoter, SP3 and NO-SEboxN induced *GUS* expression in mung bean roots. Under light microscopy, cell specificity of *GUS* expression was, however, different in these two constructs. A root specific, RHERPATEXPA7, element was found twice in SP3 and once in NO-SEboxN region (Kim *et al*., 2006). This element could play an important role in strengthening *GUS* expression in mung bean roots transformed with SP3.

4. Determination of promoter activity of DOFCORN, NO-SEboxN and SP3 regions via *GUS* enzymatic assay

Mung bean roots were transiently transformed with seven constructs driven by *CP-ACO I* and *CP-ACO II* promoters, however only those transformed with DOFCORN, NO-SEboxN and SP3 constructs displayed blue staining of *GUS* expression. SDS-PAGE was used to qualify their total proteins. Sixty eight kDa GUS protein from transformed mung bean roots with pCAMBIA1304, NO-SEboxN and SP3 revealed clearly band, however transformed root containing DOFCORN was not shown band. Fluorogenic assay was used to determine *GUS* enzymatic activity of these three constructs as well as pCAMBIA1304 (positive control). The value of *GUS* enzymatic activity of SP3 (7.3 pmole MU/min/μg) and DOFCORN (10.8 pmole MU/min/μg) transformed roots were statistically similar at p=0.05. The *GUS* activity from NO-SEboxN transformed roots was the least at 2.2 pmol MU/min/μg. The results were correlated with the strength *GUS* expression in root cells in each sample. Hwang *et al*. (2006) suggested that longer *Agrobacterium* incubation also increased the *GUS* enzymatic activity in plant tissue.

CONCLUSION

In this study, *CP-ACO II* 5'flanking region was cloned, sequenced and characterized. *CP-ACO I* and *CP-ACO II* 5'flanking regions were compared and their promoter activity determined using *Agrobacterium* transient expression. We can conclude that:

1. *CP-ACO II* 5'flanking region 591 bp in length from translation start site was identified. Sequence comparison between *CP-ACO I* and *CP-ACO II* 5'flanking regions revealed similarity (44.5%). *CP-ACO II* 5'flanking sequence was also distinct from other *ACO* 5'flanking regions reported in Genbank database.

2. Three plant promoter databases were used and offered more efficient ways to analyze *cis*-acting elements.

3. Several important *cis*-acting elements were found to be unique in either *CP-ACO I* or *CP-ACO II* 5'flanking regions. BOX A-1 (ATATATATATATTATATAAT) was found twice and only in *CP-ACO I* 5'flanking region. It was reported to be a sucrose responsive element.

4. *Cis*-acting elements such as DOFCOREZM (AAAG) and NODCON1GM (AAAGAT) were found in both *CP-ACO I* and *CP-ACO II* 5'flanking regions suggesting common functions in both *CP-ACO* isoforms.

5. *Agrobacterium* transient expression can be used to screen promoter activity in various plant tissues. It has advantages over stable transformation with speed, low cost and high efficiency.

6. The region between -410 to -215 bp from *CP-ACO II* 5'flanking region (NO-SEboxN) was found to have constitutive promoter activity in all tested tissues

(flower, fruit, leaf and root). The sequence analysis revealed that there was ASF1MOTIFCAMV (TGACG) which was a core element of 35S CaMV.

7. Three constructs (DOFCORN, NO-SEboxN and SP3) driven *GUS* expression. *GUS* enzymatic activity of transformed root with DOFCORN was shown highest at 10.8 pmol MU/min/µg.

LITERATURE CITED

Abe, H., T. Urao, T. Ito, M. Seki, K. Shinozaki and K. Yamaguchi-Shinozaki. 2003. *Arabidopsis* AtMCY2 (dHLH) and AtMYB2 (MYB) function as transcriptional activators in abscisic acid signaling. **Plant Cell.** 15: 63-78.

Abe, H., K. Yamagushi-Shinozaki, T. Urao, T. Iwasaki, D. Hosokawa and K. Shinozaki. 1997. Role of *Arabidopsis MYC* and *MYB* homologs in drought- and abscisic acid-regulated gene expression. **Plant Cell.** 9: 1859-1868.

Alberts, B. 1998. **Molecular Biology of the Cell.** 3rd ed. Garland Science, New York.

Alexander, L. and D. Grierson. 2002. Ethylene biosynthesis and action in tomato: a activity. **JBC.** 266: 11632-11639.

Allen, R.D., F. Bernier, P.A. Lessard and R.N. Beachy. 1989. Nuclear factors interact with soybean Beta-conglycinin enhancer. **Plant Cell.** 1: 623-631.

Andrea, D. and B. Dorothea. 2006. Identification of a dehydration and ABA-responsive promoter regulon and isolation of corresponding DNA binding proteins for the group 4 *LEA* gene CpC2 from *C. plantagineum*. **Plant Mol. Biol.** 61: 643-663.

Arendse, M.S., L.A. Dubery and D.K. Berger. 1999. Isolation by PCR-based methods of a plant antifungal polygalactorunase inhibiting protein gene. **J. of Biotechnol.** 2: 152-159.

Arguello-Astorga G.R and L.R. Herrera-Estrella. 1996. Ancestral multipartite unit in light-responsive plant promoters have structural features correlating with specific phototransduction pathway. **Plant Physiol.** 112: 1151-66.

Atkinson, R.G., K.M. Bolitho, M.A. Wright, T.I. Bueno, S.J. Reid and G.S. Ross. 1998. Apple ACC oxidase and polygalacturonase: ripening specific gene expression and promoter analysis in transgenic tomato. **Plant Mol. Biol.** 38: 449-460.

Avila, C., R. Francisco, P.B. Cantóna, S. María-Fernanda, M. Pierre, R. Manuel, M.H. Jaime, O. Ricardo and M.C. Francisco. 2002. Structural and functional characterization of the 5'upstream region of a glutamine synthetase gene from Scots pine. **Ann. For Sci.** 59: 669-673.

Baker, S.S., S. Wilhelm and F. Thomashow. 1994. The 5'-region of *Arabidopsis thaliana corl5a* has *cis-acting* elements that confer cold-drought-and ABA-regulated gene expression. **Plant Mol. Biol.** 24: 701-713.

Balavoine, G. 1996. Identification of members of several homeobox genes in a planarian using a ligation-mediated polymerase chain reaction technique. **Nucleic Acids Res.** 24: 1547-1535.

Baranowskij, N., C. Frohber, S. Prat and L. Willmitzer. 1994. A novel DNA binding protein with homology to Myb oncoproteins containing only one repeat can function as a transcriptional activator. **EMBO J.** 13: 5383-5392.

Barry, C.S., B. Blume, M. Bouzayen, W. Copper, A.J. Hamilton and D. Grierson. 1996. Differential expression of the 1-Aminocyclopropane-1-Carboxylate (ACC) Oxidase gene family of tomato. **Plant J.** 9: 525-535.

Bate, N. and D. Twell. 1998. Functional architechture of a late pollen promoter: pollen specific transcriptional isdevelopmentally regulated by multistage-specific and co-depedent activatir elements. **Plant Mol. Biol.** 37: 859-869.

Baumann, K., A. De Paolis, P. Costantino and G. Gualberti. 1999. The DNA binding site of the dof protein NtBBF1 is essential for tissue-specific and auxin-regulated expression of the *rolB* oncogene in plants. **Plant Cell.** 11: 323-333.

Bingfang, X. and P.T. Michael. 2004. Methyl jasmonate induced expression of the tobacco putrescine *N*-methyltransferase genes requires both G-box and GCC-motif elements. **Plant Mol. Biol.** 55: 743-761.

Bisanz-Seyer, C and R. Mache. 1992. Organization and expression of the nuclear gene coding for the plastid-specific S22 ribosomal protein from spinach. **Plant Mol. Biol.** 18: 337-344.

Blume, B. and D. Grierson. 1997. Expression of ACC oxidase promoter GUS-fusions in tomato and *Nicotiana plumbaginifolia* regulated by development and environmental stimuli. **Plant J.** 12: 731-746.

_____, C.S., A.J. Hamilton, M. Bouzayen and D. Grierson. 1997. Identification of transposon-like elements in non coding regions of tomato ACC oxidase gene**.** **Mol. Gen. Genet.** 254: 297-303.

Bolle, C., G.H. Reinhold and O. Ralf. 1996. Different sequences for 5'-untranslated leaders of nuclear genes for plastid proteins affect the expression of the β-glucuronidase gene. **Plant Mol. Biol.** 32: 861-868.

Bolle, C., S.L. Sopory, G.H. Reinhold and O. Ralf. 1994. Segments encoding 5'-untranslated leaders of genes for thylakoid proteins contain cis-elements essential for transcription. **Plant J**. 6: 513-523.

Boyle, B. and N. Brisson. 2001. Repression of the defense gene PR-10a by the signal-stranded DNA binding protein SEBF. **Plant Cell.** 13: 2525-2537.

Bruce, W.B., X.W. Deng and P.H. Quail. 1991. A negatively acting DNA sequence element mediates phytochrom-directed repression of *phy A* gene transcription. **EMBO J.** 10: 3015-3024.

Campbell, N.A. and J.B. Reece. 2004. **Biology**. 7nd ed. Person Education, Inc., San Francisco.

Cannon, A.C. and S.D. Shiflett. 2001. Transcriptional regulation of the nitrate reductase gene in *Chlorella vulgaris*: identification of regulatory elements controlling expression. **Curr. Genet.** 40: 128-135.

Caroline, R., L. Magali, I. Dirk and D.V. Lieven. 2002. Effect of auxin, cytokinin, and sucrose on cell cycle gene expression in *Arabidopsis thaliana* cell suspension cultures. **Plant Cell, Tissue and Organ Culture.** 69: 167–176.

Chan, C.S., L. Guo and M.C. Shih. 2001. Promoter analysis of the nuclear gene encoding the chloroplast glyceraldehyde-3-phosphate dehydrogenase B subunit of *Arabidopsis thaliana*. **Plant Mol. Biol.** 46: 131-141.

Chaubet, N., M. Flenet, B. Clement, P. Brignon and C. Gigot. 1996. Identification of *cis*-elements regulating the expression of an *Arabidopsis* histone *H4* gene. **Plant J.** 10: 425-435.

Chen, W., N.J. Provart, J. Glazebrook, F. Katagiri, H.S. Chang, T. Eulgem, F. Mauch, S. Luan, G. Zou, S.A. Whitham, P.R. Budworth, Y.T. Xie. 2002. Expression profile matrix of *Arabidopsis* transcription factor genes suggests their putative functions in response to environmental stresses. **Plant Cell.** 14: 559-574.

Chen, Y.T., Y.R. Lee, Y.T. Wang, S.F. Yang and J.F. Shaw. 2003. A novel ACC oxidase *gene (CP-ACO2*) associated with the late stage fruit ripening and leaf senescence. **Plant Sci.** 164: 531-540.

Chen-Yi, H., Y. Lin, M. Zhang, S. Pollock, M.D. Marks and J. Schiefelbein. 1998. A common position-dependent mechanism controls cell-type patterning and GLABRA2 regulation in the root and hypocotyl epidermis of Arabidopsis. **Plant Physiol.** 117: 73–84.

Chern, M.S., A.S. Bobb and M.M. Bustos. 1996. The regulator of *MAT2* (ROM2) protein binds to early maturation promoters and represses PvALF-activated transcription. **Plant Cell.** 8: 305-321**.**

Christopher, I.C., J. Burke and J. Velten. 2005. Functional characterization of the geminiviral conserved late element (CLE) in uninfected tobacco. **Plant Mol. Biol.** 58: 465-481.

Chuaboonmee, S. 2004. **Identification and Characterization of a Tissue-Specific 1-Aminicyclopropane-1-Carboxylate (ACC) Oxidase Promoter in *Carica papaya*.** M.S. Thesis, Kasetsart University.

Chuan-Yin, W., W. Haruhiko, O. Yasuyuki, H. Kyuya and T. Fumio. 2000. Quantitative nature of the Prolamin-box, ACGT and AACA motifs in a rice glutelin gene promoter: minimal *cis*-element requirements for endosperm-specific gene expression. **Plant J.** 23: 415-421.

Chung, H.J., H.Y. Fu, L. Terry. 2005. Abscisic acid-inducible nuclear proteins bind to bipartite promoter elements required for ABA response and embryo-regulated expressionof the carrot *Dc3* gene. **Planta.** 220: 424-433.

Colot, V., L.S. Robert, T.A. Kavanagh, M.W. Bevan and R.D. Thompson. 1987. Localization of sequences in wheat endosperm protein genes which confer tissue-specific expression in tobacco. **EMBO J.** 6: 3559-3564.

Cottage, A., A. Yang, H. Maunder, R.C. Lacy and N.A. Ramsay. 2001. Identification of DNA sequences flanking T-DNA insertions by PCR-walking. **Plant Mol. Biol.** 19: 321-327.

Czarnecka, E., P.C. Fox and W.B. Gurley. 1990. In vitro interaction of nuclear proteins with the promoter of soybean heat shock gene *Gmhsp17.5E.* **Plant Physiol.** 94: 935-943.

Dan, G., L. Xihui, W. Maoyan, Z. Jun, H. Wei, W. Guoying and W. Jianhua. 2007. Overexpression of *ZmOPR1* in Arabidopsis enhanced the tolerance to osmotic and salt stress during seed germination. **Plant Sci.** 174: 124-130.

Daraselia, D.N., S. Tarchevskaya and J. O. Narita. 1996. The promoter for tomato 3-Hydroxy-3-Methylglutaryl coenzyme a reductase gene 2 has unusual regulatory elements that direct high-leve1 expression. **Plant Physiol.** 112: 727-733.

Dean, D.P., J.E. Lincoln, R.L. Fischer and A.B. Bennett. 1989. Transcriptional analysis of polygalacturonase and other ripening associated genes in rutgers, *rin, nor* and *Nr* tomato fruit. **Plant Physiol.** 90: 1372-1377.

Degenhardt, J. and E.M. Tobin. 1996. A DNA binding activity for one of two closely defined phytochrome regulatory elements in an *Lhcb* promoter is more abundant in etiolated than in green plants. **Plant Cell.** 8: 31-41.

Dhundy, R.B., V.P. Vaijaynti and W. Ilga**.** 1998. Alfin1, a novel zinc-finger protein in alfalfa roots that binds to promoter elements in the salt-inducible *MsPRP2* gene. **Plant Mol. Biol.** 38: 1123–1135.

Doven, R.S., D.J. Porteous and A.J. Brookes. 1995. Splinkerettes–improved vectorettes for greater efficiency in PCR walking. **Nucleic Acids Res.** 23: 1644-1645.

Ellerstrom, M., K. Stalberg, I. Ezcurra and L. Rash. 1996. Functional dissection of a napin gene promoter: identification of promoter elements required for embryo and endosperm- specific transcription. **Plant Mol. Biol.** 32: 1019-1027.

Elmayan, T. and M. Tepfer. 1995. Evaluation in tobacco of the organ specificity and strength of the *rol D* promoter, domain A of the *35S* promoter and the *35S 2* promoter. **Transgenic.** 4: 388-396.

Ericson, M.L., E. Muren, H.O, Gustavsson, L.G. Josefsson and L. Rask. 1991. Analysis of the promoter region of napin genes from *Brassica napus* demonstrates binding of nuclear protein in vitro to a conserved sequence motif. **Eur. J. Biochem.** 197: 741-746.

Eulgem, T., P.J. Rushton, E. Schmelzer, K. Hahlbrock and I.E. Somssich. 1999. Early nuclear events in plant defense signalling: rapid gene activation by KRKY transcription factors. **EMBO J.** 18: 4689-4699.

Ezcurra, I., M. Ellerström, P. Wycliffe, K. Stålberg and L. Rask. 1999. Interaction between composite elements in the *napA* promoter: both theB-box ABA-responsive complex and the RY/G complex are necessary forseed-specific expression. **Plant Mol. Biol.** 40: 699-709.

Fujiwara, T. and R.N. Beachy**.** 1994. Tissue-specific and temporal regulation of a β-conglycinin gene: roles of the RY repeat and other *cis-acting* elements. **Plant Mol. Biol.** 24: 261-272.

Fusada, N., T. Masuda, H. Kuroda, H. Shimada, H. Ohta and K. Takamiya. 2005. Identification of a novel cis-element exhibiting cytokinin-dependent protein binding in vitro in the 5'-region of NADPH-protochlorophyllide oxidoreductase gene in cucumber. **Plant Mol. Biol.** 59: 631-645.

Gasser, S.M., B.B. Amati, M.E. Cardenas and J.F.X. Hofmann. 1989. Studies on scaffold attachment sites and their relation to genome function. **Intnatl. Rev. Cyto.** 119: 57-96.

Geffers, R., R. Cerff and R. Hehl. 2000. Anaerobiosis-specific interaction of tobacco nuclear factors with *cis*-regulatory sequences in the maize *GapC4* promoter. **Plant Mol. Biol.** 43**:** 11–21.

Gidoni, D., P. Brosio, D. Bond-Nutter, J. Bedbrook and P. Dunsmuir. 1989. Novel *cis*- acting elements in Petunia *Cab* gene promoters. **Mol. Gen. Genet.** 215: 337-344.

Giuliano, G., E. Pichersky, V.S. Malik, M.P. Timko, P.A. Scolnik and A.R. Cashmore. 1988. An evolutionarity conserved protein binding sequence upstream of a plant light-regulated gene. **Proc. Natl. Acad. Sci.** 85: 7089-7093.

Goldsbrough, A.P., H. Albrecht and R. Stratfold. 1993. Salicylic acid–inducible binding of a tobacco nuclear protein to a 10 bp sequence which is higgly conserved amongst stress-inducible genes. **Plant J.** 3: 563-571.

Gowik, U., J. Burscheidt, M. Akyildiz, U. Schlue, M. Koczor, M. Streubel, and P. Westhoff. 2004. *Cis*-regulatory elements for mesophyll-specific gene expression in the C4 plant flaveria trinervia, the promoter of the C4 phosphoenolpyruvate carboxylase gene. **Plant Cell.** 16: 1077-1090.

Gray, J., S. Picton, J. Shabbeer, W. Schuch and D. Grierson. 1992. Molecular biology of fruit ripening and its manipulation with antisense gene. **Plant Mol. Biol.** 19: 69-87.

Grierson, C., D. Jian-Sheng, D.T.Z. Marta, B. Kyle, S. Caroline, H. Michssl and M. Bevan. 1994. Separate *cis* sequences and *trans* factors direct metabolic and

developmental regulation of a potato tuber storage protein gene. **Plant J.** 5: 815-826.

Grierson, D., B. Blume, A. Hamilton, M. Holdsworth and C. Barry. 2001. **DNA construcs and plants incorporating them.** U. S. Patent. 6, 204, 473.

Grotewold, E., B.J. Drummond, B. Bowen and T. Peterson. 1994. The myb-homologous *P* gene controls phlobaphene pigmentation in maize floral organs by directly activating a flavonoid biosynthetic gene subset. **Plant Cell.** 76: 543-553.

Ha, S. and G. An. 1988. Identification of upstream regulatory elements involved in the developmental expression of the *Arabidopsis thaliana cab1* gene. **Proc. Natl. Acad. Sci.** 85: 8017-8021.

Hartwell, L.H., L. Hood., M.L. Goldberg, A.E. Reynolds, L.M. Silver and R.C. Veres. 2004. **Genetis: from genes to genomes.** 2 nd ed. McGraw-Hill Companies, Inc., New York.

Hawk-Bin, K., P. Soo-Chul, P. Hsiao-Ping, M. Howard, J.D. Coodman and S. Ming-Che. 1994. Identification of a light-responsive region of the nuclear gene encoding the B Subunit of chloroplast glyceraldehydes 3-phosphate dehydrogenase from *Arabidopsis thaliana*. **Plant Physiol.** 105: 357-367.

He, X., J. Futterer and T. Horn. 2002. Contribution of downstream promoter elements to transcriptional regulation of the rice tungro bacilliform virus promoter. **Nucleic Acid Res.** 30: 497-506.

Heidecker, G. and J. Messing. 1986. Structural analysis of plant genes. Review**. Ann. Rev. Plant Physiol.** 37: 439-466.

Helmut, B., L. Nagyt, R. Villarroel, and U. Wobus. 1992. *Cis*-analysis of a seed protein gene promoter: the conservative RY repeat CATGCATG within the legumin gene. **Plant J.** 2: 233-239.

Hermann, S.R., J.A.C. Miller, S. O'neill, T.T. Tsao, R.M. Harding and J.L. Dale. 2000. Single primer amplification of flanking sequences. **Biotechniques.** 29: 1176-1180.

Higo, K., Y. Ugawa, M. Iwamoto and T. Korenaga. 1999. Plant *cis*-acting regulatory DNA elements (PLACE) database. **Nucleic Acids Res.** 27: 297-300.

Holdsworth, M.J., W. Schuch and D. Grierson. 1988. Organization and expression of a wound-ripening related small multigene family from tomato. **Plant Mol. Biol.** 11: 77-88.

.

Hudson, M.E. and P.H. Quail. 2003. Identification of promoter motif involved in the network of phytochrome A-regulated gene expression by combined analysis of genomic sequence and microarray data. **Plant Physiol.** 133: 1605-1616.

Hwang, H.-H., K.S. Myore and S.B. Gelvin**.** 2006. Transgenic *Arabidopsis* plants expressing *Agrobacterium tumefaciens* VirD2 protein are less susceptible to *Agrobacterium* transformation. **Molecular Plant Pathology**. 7**:** 473-484.

Hwang, Y.S., E.E. Karrer, R.B. Thomas, L. Chen and R.L. Rodriquez. 1998. Three *cis*-element required for rice alpha-amylase Amy3D expression during sugar stravation**. Plant Mol. Biol.** 36: 331-334.

Inmaculada, P., G. Pedro, M.D. Rosa, R. Manuel, A. Anna, E. Espel, F. Natale and J. Palau. 1994. Narrow A/T-rich zones present at the distal 5'-flanking sequences of the zein genes *Zcl* and *Zc2* bind a unique 30 kDa HMG-like protein. **Plant Mol. Biol.** 26: 1893-1906.

Ishiguro, S. and K. Nakamura. 1994. The nuclear factor SP8BF binds to the 5' upstream regions of three different genes coding for major proteins of sweet potato tuberous root. **Plant Mol. Biol.** 18: 97-108.

Itzhaki, H., J.M. Maxon and W.R. Woodson. 1994. An ethylene-responsive enhancer element is involved in the senescence- related expression of the carnation glutathione-S transferase (*GTS1*) gene. **Proc. Natl. Acad. Sci.** 91: 8925-8929.

Jaing, C., B. Iu and J. Singh. 1996. Requirement of a CCGAC *cis* -acting element for cold induction of the *BN115* gene from winter *Brassica napus.* **Plant Mol. Biol.** 30: 679-684.

Janice, W. E., L.W. Elsbeth and M.C. Gloria. 1990. Cell-specific expression in transgenic plants reveals nonoverlapping roles for chloroplast and cytosolic glutamine synthetase. **Proc. Natl. Acad. Sci.** 87: 3459-3463.

Jefferson, R.A., T.A. Kavanagh and M.W. Bevan. 1987. *Gus* fusions: ß-glucuronidase as a sensitive and versatile gene fusion marker in higher plants. **EMBO J.** 6: 3901-3909.

Jensen, E.O., K.A. Marcker, J. Schell and F.J. de Bruijn. 1988. Interaction of a nodule specific, trans-acting factor with distinct DNA elements in the soybean leghaemoglobin *Ibc3* 5' upstream region. **EMBO J.** 7: 1265-1271.

Joshi, C.P. 1987. Putative polyadenylation signals in nuclear genes of higher plants: a compilation and analysis. **Nucleic Acids Res.** 15: 9624-9640.

Kagaya, Y., K. Ohmiya and T. Hattori. 1999. RAV1, a novel DNA binding protein, binds to bipatite recognition sequence through two distinct DNA-binding domains uniquely found in higher plants. **Nucleic Acids Res.** 27: 470-478.

Kapoor, S. and M. Sugiura. 1999. Identification of two essential sequence elements in thenonconsensus type II *PatpB-290* plastid promoter by using plastid transcription extracts from cultured tobacco BY-2 cells. **Plant Cell.** 11: 1799-1810.

Karen, L.K. and L.M. Lagrimini. 1996. Phytohormone control of the tobacco anionic peroxidase promoter. **Plant Mol. Biol.** 31: 565-573.

Ki-Hong, J., J. Hur., R. Choong-Hwan, C. Youngju, C. Yong-Yoon, M. Akio, H. Hirohiko, and A. Gynheung. 2003. Characterization of a rice chlorophyll-deficient mutant using the T-DNA gene-trap system. **Plant Cell Physiol.** 44: 463-472.

Kim, D.W., S.H. Lee, S.B. Choi, S.K. Won, Y.K. Heo, M. Cho, Y.I. Park and H.T. Cho. 2006. Functional conservation of a root hair cell-specific *cis*-element in angiosperms with different root hair distribution patterns. **Plant Cell**. 18: 2958-2970.

Kim, S.Y., H.J. Chung and T.L. Thomas. 1997. Isolation novel class of bZIP transcription factor that interact with ABA responsive abg embryo specificelements in The *DC3* promoter using a modified yeast one-hybrid system**. Plant J.** 11: 1237-1251.

Kim, W.T. and S.F. Yang. 1994. Structure and expression of cDNAs encoding 1-Aminocyclopropane-1-Carboxylate (ACC) Oxidase homologs isolated from excised mung bean hypocotyls. **Planta.** 194: 223-229.

King, M.W. 2003. **Control of gene expression in eukaryote**. Available Source: http://web.indstate.edu/thcme/mwking/gene-regulation.html, November 24, 2007.

Klug, W.S. and M.R. Cummings. 1997. **Concepts of Genetics.** 5th ed., Prentic Hall, Inc., New Jersey.

Ko, J., E. Beers and K. Han. 2006. Global comparative transcriptome analysis identifies gene network regulating secondary xylem development in *Arabidopsis thaliana*. **Mol. Genet. Genomics.** 276: 517-531.

Kooiker, M., A.A. Airoldi, A. Losa, S.M. Priscilla, L. Finzi, M.M. Kater and L. Colombob. 2005. Basic Pentacysteine, a GA binding protein that induces conformational changes in the regulatory region of the homeotic *Arabidopsis* Gene *SEEDSTICK*. **Plant Cell.** 17: 722-729.

Kropat, J., S. Tottey, R. Birkenbihl, N. Depege, P. Huijser, R.S. Merchant. 2005. A regulator of nutritional copper signaling in Chlamydomonas is an SBP domain protein that recognizes the GTAC core of copper response element. **Proc. Natl. Acad. Sci.** 102: 18730-18735.

Kucho, K., S. Yoshioka, F. Taniguchi, K. Ohyama and H. Fukuzawa. 2003. Cis-acting elements and DNA-binding proteins involved in CO_2-responsive transcriptional activation of Cah1 encoding a periplasmic carbonic anhydrase in *Chlamydomonas reinhardtii*. **Plant Physiol.** 133: 783-793.

Kuhlemeierr, C., M. Cuozzo, J.G. Pamela, E. Goyvaerts, K. Ward and N.H. Chua. 1988. Localization and conditional redundancy of regulatory elements in *rbcS-3A*, a pea gene encoding the small subunit of ribulose-bisphosphate carboxylase. **Proc. Natl. Acad. Sci.** 85: 4662-4666.

Kumdee, O., S. Bandee, S. Chanprame and P. Burns. 2003. Comparative expression of 2 locis of ACC oxidase during the ripen of papaya cv. Khaek Nuan. **Agricultural Sci. J.** 34: 97-101.

Lagerstrom, M.J. 1991. PCR technique. **PCR Method Appl.** 1: 111-119.

Lam, E. and N. Chua. 1989. ASF-2: A factor that binds to the cauliflower mosaic virus 35S promoter and a conserved GATA motif in cab promoters. **Plant Cell**. 1: 1147-1156.

Lagrange, T., B. Franzetti, M. Axelos, R. Mache and S. Lerbs-Mache. 1993. Structure and expression of the nuclear gene encoding for the chloroplast ribosomal protein L21: Developmental regulation of a house-keeping gene by alternative promoters. **Mol. Cell Biol.** 13: 2614-2622.

Lemke, D.E. 1999. **Structure of Typical Dicot Root**. Herbaceous Dicot Roots, part 2. Available Source: http://www.bio.txstate.edu/~dlemke/botany/1410lab/lab_exercises/lab5/roots/dicot_root2.html, May 12, 2008.

Lescot, M., P. Déhais, Y. Moreau, B. De Moor, P. Rouzé and S. Rombauts. 2002. PlantCARE: a database of plant cis-acting regulatory elements and a portal to tools for in silico analysis of promoter sequences. **Nucleic Acids Res.** 30: 325-327.

Lewin, B. 2006. **Genes VIII.** Pearson Education, Inc., London.

Li, L. and J.C. Steffens. 2002. Overexpression of polyphenol oxidase in transgenic tomato plants results in enhanced bacterial disease resistance. **Planta.** 215: 239-247.

Lijun, Y., K. Hiroyuki, S. Kazuya, H. Sakiko, F. Kazuhito and T. Fumio. 2007. Generation of a transgenic rice seed-based edible vaccine against house dust mite allergy. **BBRC.** 365: 334–339.

Lin, C., C.Z. Hong, X.Z. Yu and Y.C. Tuan. 2007. *Arabidopsis DREB1A* confers high salinity tolerance and regulates the expression of GA dioxygenases in tobacco. **Plant Sci.** 174: 156-174.

Liu, J.H., S.H.L. Tamon and D.M. Reid. 1997. Differential and wound-inducible expression of two 1-Aminocyclopropane-1-Carboxylate (ACC) Oxidase gene in sunflower seedlings. **Plant Mol. Biol.** 34: 923-933.

Liu, Y.G. and R.F. Whittier. 1995. Thermal asymmetric interlaced PCR: automatable amplification and sequencing of insert end fragments from *P1* and YAC clones for chromosome walking. **Genomics.** 25: 674-618.

Logemann, E., M. Parniske and K. Hahlbrock. 1995. Modes of expression and common structural features of the complete phenylalanine ammonia-lyase gene family in parsley. **Proc. Natl. Acad. Sci.** 92: 5905-5909.

Lois, R., A. Dietrich, K. Hahlbrock and W. Schulz. 1989. A phenylalanine ammonia-lyase gene from parsley: structure, regulation and identification of elicitor and light responsive *cis*-acting elements. **EMBO J.** 8: 1641-1648.

Lohmer, S., M. Madaloni, M. Motto, N. Di Fonzo, H. Hartings, F. Salamini and R.D. Thompson. 1991. The maize regulation locus Opaque-2 encodes a DNA-binding protein which activates the transcription of the b-32 gene. **EMBO J.** 10: 17-24.

Lu, C.A., T.H.D. Ho, S.L. Ho, S.M. Yu. 2002. Three Novel MYB Proteins with one DNA binding repeat mediate sugar and hormone regulation of alpha-amylase gene expression. **Plant Cell.** 14: 1963-1980.

Luo, M.C., C. Yen and J.L. Yang. 1993. Crossability percentages of bread wheat collections from Tibet, China with rye. **Euphytica.** 70: 127-129.

Luo, H., F. Song, R. Goodman and Z. Zheng. 2005. Up-regulation of *OsBIHD1*, a rice gene encoding BELL homeodomain transcriptional factor, in disease resistance responses. **Plant Biol.** 7: 459-468.

Maeda, K., S. Kimura, T. Demura, J. Takeda and Y. Ozeki. 2005. *DcMYB1* acts as a transcriptional activator of the carrot phenylalanine ammonia-lyase gene (*DcPAL1*) in response to elicitor treatment, UV-B irradiation and the dilution effect. **Plant Mol. Biol.** 59: 739-752.

Mallery, C. 2001. **Gene regulation control in eukaryotes**. Available Source: http://cats.med.uvm.edu/cats_teachingmod/microbiology/courses/gene_regulation/euk., November 24, 2007.

Manjunath, S. and M.M. Sachs. 1996. Molecular characterization and promoter analysis of the maize cytosolic glyceraldehyde 3-phosphate dehydrogenase gene family and its expression during anoxia. **Plant Mol. Biol.** 33: 97-112.

Martinez-Hernandez, A., L. Lopez-Ochoa, G. Arguello-Astorga and L. Herrera-Estrella. 2002. Functional properties and regulatory complexity of a minimal RBCS light-responsive unit activated by phytochrome, cryptochrome, and plastid signals. **Plant Physiol.** 128: 1223-1233.

Maruyama-Nakashita1, A., Y. Nakamura, A. Watanabe-Takahashi1, E. Inoue and H. Takahashi. 2005. Identification of a novel *cis*-acting element conferring sulfur deficiency response in Arabidopsis roots. **Plant J.** 42: 305–314.

Masaki, I., M. Iwase, H. Kodama, P. Lavisse, K. Komamine, R. Nishihama, Y. Machida and A. Watanabe. 1998. A novel *cis*-acting element in promoters of plant B-type cyclin genes activates M phase–specific transcription. **Plant Cell.** 10: 331-341.

Matthews, R.E.F. 1991. **Plant Virology.** 3rd ed., Academic Press, California.

Mohanty, B., S. Krishnan, S. Swarup and V. Bajic. 2005. Detection and preliminary analysis of motifs in promoters of anaerobically induced genes of different plant species. **Ann. Bot.** 96: 669-681.

Montgomery, J., S. Goldman, J. Beiknan, L. Margossian and R.L. Fischer. 1993. Identification of an ethylene-responsive region in the promoter of a fruit ripening gene. **Proc. Nalt. Acad. Sci.** 90: 5939-5943.

Moon, H. and A.M. Callahan. 2004. Development regulation of peach ACC oxidase promoter-GUS fusions in transgenic tomato fruits. **J. Exp. Bot.** 55: 1515-1528.

Moore, I., M. Samalova and S. Kurup. 2006. Transactivated and chemically inducible gene expression in plants. **Plant J.** 45: 651-683.

Morikami, A., R. Matsunaga, Y. Tanaka, S. Suzuki, S. Mano and K. Nakamura. 2005. Two *cis*-acting regulatory elements are involved in the sucrose-inducible expression of the sporamin gene promoter from sweet potato in transgenic tobacco. **Mol. Genet. Genomics.** 272: 690-699.

Morita, A., T. Umemura, M. Kuroyanagi, Y. Futsuhara, P. Perata and J. Yamagushi. 1998. Functional dissection of a sugar repressed alpha-amylase gene. **FEBS. Lett.** 423: 81-85.

Muller, M. and S. Kundsen. 1993. The nitrogen respon of a barley C-hordein promoter is controlled by positive and negative regulation of the GCN4 and endosperm box**. Plant J.** 4: 343-355.

Nakamura, M., T. Tsunoda and J. Obokata. 2002. Photosynthesis nuclear genes generally lack TATA boxes: a tobacco photosystem I gene responds to light through an initiater. **Plant J.** 29: 1-10.

Nakatsuka, A., M. Shiho, O. Hironori, S. Shinjiro, N. Ryohei, K. Yasutaka and I. Akitsugu. 1998. Differential expression and internal feedback regulation

of 1-Aminocyclopropane-1-Carboxylate Synthase, 1-Aminocyclopropane-1-Carboxylate Oxidase, and ethylene receptor genes in tomato fruit during development and ripening. **Plant Physiol.** 118: 1295-1305.

Neupane, K.R., U.T. Mukatira, C. Kato and J.I. Stiles. 1998. Cloning and characterization of fruit-expressed ACC synthase and ACC oxidase from papaya (*Carica papaya* L.), pp. 329-337. *In* R.A. Draw. **ISHS Acta Horticulturae 461: International Symposium on Biotechnology of Tropical and Subtropical Species Part 2.**

Nishiuchi, T., H. Shinshi and K. Suzuki. 2004. Rapid and transient activation of transcription of the *ERF3* gene by wounding in tobacco leaves: Possible involvement of NtWRKYs and autorepression. **J. Biol. Chem.** 279: 55355-55361.

O'Neill, S.D., M.H. Kumagai, A. Majumdar, N. Huand, T.D. Sutliff and R.L. Rodriquez. 1990. The alpha- amylase genes in *Oryza sativa*: Characterization of cDNA clones and mRNA expression during seed germination. **Mol. Gen. Genet.** 221: 235-244.

Ochman, H., A.S. Gerber and D.L. Hartl. 1988. Genetic applications of an inverse polymerase chain recation. **Genetics.** 120: 621-629.

Orzaez, D., S. Mirabel, W.H. Wieland and A. Granell. 2006. Agroinjection of tomato fruits. A tool for rapid functional analysis of transgenes directly in fruit. **Plant Physiol**. 140: 3–11.

Park., H., M. Kim, Y. Kang, J. Jeon, J. Yoo, M. Kim, C. Park, J. Jeong, B. Moon, J. Lee, H. Yoon, S. Lee, W. Chung, C. Lim, S. Lee, J. Hong and M. Cho. 2004. Pathogen-and NaCl-induced expression of the *SCaM-4* promoter is mediated in part by a GT-1 box that interacts with a GT-1-like transcription factor. **Plant Physiol.** 135: 2150-2161.

Pasquali, G., A.S. Erven, P.B. Ouwerkerk, F.L. Menke and J. Memelimk. 1999. The promoter of the strictosidine synthase gene from periwinkle confers elicitor-inducible expression in transgenic tobaccoand binds nuclear factors GT-1 and GBF. **Plant Mol. Biol.** 39: 1299-1231**.**

Pastuglia, M., D. Roby, C. Dumas and J.M. Cock. 1997. Rapid induction by wounding and bacterial infection of an *S* gene family receptor like kinase gene in *Brassica oleracea*. **Plant Cell.** 9: 49-60.

Piechulla, B., N. Merforth and B. Rudolph. 1998. Identification of tomato *Lhc* promoter regions necessary for circadian expression. **Plant Mol. Biol.** 38: 655-662.

Pogson, B.J., C.G. Downs and K.M. Davies. 1995. Differential expression of two 1-Aminocyclopropane-1-Carboxylate (ACC) Oxidase gene in broccoli after harvest. **Plant Physiol.** 108: 651-657.

Planchais, S., C. Perennes, N. Glab, V. Mironov, D. Inze, R.A. Bergounioux. 2002. Characterization of cis-acting element involved in cell cycle phase-independent activation of Arath; CycB1;1 transcription and identification of putative regulatory proteins. **Plant Mol. Biol.** 50:111-127.

Plesch, G., T. Ehrhardt and B. Mueller-Roeber. 2001. Involvement of TAAAG elements suggests a role for Dof transcription factors in guard cell-specific gene expression. **Plant J.** 28: 455-464.

Potrykus, I., R. Bilang, J. Futterer, C. Sautter and M. Schrott. 1998. Chapter 6 Genetic Engineering of crop plant. **In: Agricultural Biotechnology.** Altman edited. Marcel Dekker Inc.

Raghavan, V. 1997. **Molecular embryology of flowering plants.** Cambridge University Press, New York.

Ray, C., B. Fuchs, N. Walter, W.K. Lutke and C.G. Taylor. 2005. Ex vitro composite plants: an inexpensive, rapid method for root biology. **Plant J**. 43: 449-457.

Redman, J., J. Whitcraft, C. Johnson and J. Arias. 2002. Abiotic and biotic stress differentially stimulate as 1-element activity in *Arabidopsis*. **Plant Cell Rep.** 21: 180-185.

Rieping, M., F. Schoffl. 1992. Synergistic effect of upstream sequences, CCAAT box elements, and HSE sequences for enhanced expression of chimaeric heat shock genes in transgenic tobacco. **Mol. Gen. Gene.** 231: 226-232.

Rodolfo, L.G., M.D. Franscisco, M.A. Omar and A.G.L. Miguel. 2004. Identification of a genomic clone to ACC oxidase from papaya (*Carica papaya* L.) and expression studies. **J. Agric. Food Chem.** 52: 794-800.

Rodriquez, C.R. 2003. **Promoter used to regulate gene expression**. Available Source: http://www.cambiaip.org/Whitepapers/Transgenic/Promoters/promoters.html., November 24, 2007.

Rogers, H., N. Bate, J. Combe, J. Sullivan, J. Sweetman, C. Swan, D. Lonsdale and D. Twell. 2001. Functional analysis of *cis*-regulatory elements within the promoter of the tobacco late pollen gene g10. **Plant Mol. Biol.** 45: 577-585.

Rombauts, S., K. Florquin, M. Lescot, K. Marchal, P. Rouze and Y.V. Peer. 2003. Computational approaches to identify promoters and *cis*-regulatory elements in plant genomes. **Plant Physiol.** 132: 1162-1176.

Rosenthal, A. and D.S.C. Jones. 1990. Genomic walking and sequencing by oligo-cassette mediated polymerase chain reaction. **Nucleic Acids Res.** 18: 3095-3096.

Rossitza, A., F. Martine, G. Claude and C. Nicole. 1998. Functional analysis of the promoter region of a maize (*Zea mays* L.) H3 histone gene in transgenic *Arabidopsis thaliana.* **Plant Mol. Biol.** 37: 275-285.

Rost, T.L. 1996. **Tomato anatomy**. Available Source: http://www plb.ucdavis.edu /labs/rost/Tomato/Reproductive/anat.html., November 30, 2007.

Rouster, J., R. Leah, J. Mundy and V. Cameron-Millis. 1997. Identification of a methyl jasmonate-responsive region in the promoter of a lypoxygenase1 gene expressed in barley gain. **Plant J.** 11: 513-523**.**

Rushmore, T., M. Morton and C. Pickett. 1991. The antioxidant responsive element. Activation by oxidative stress and identification of the DNA consensus sequence required for functional activity. **J. Biol. Chem.** 266: 11632-11639.

Rushton, P., J. Torres, M. Parniske, P. Wernert, K. Hahlbrock and I. Somssich. 1996. Interaction of elicitor-induced DNA-binding proteins with elicitor response elements in the promoters of parsley *PR1* genes. **EMBO J.** 15: 5690-5700.

Russell, P.J. 2006. **iGenetics: A Molecular Approach.** 2nd ed. Pearson Education, Inc., San Francisco.

Sablowski, R., E. Moyano, F. Culianez-Macia, W. Schuch, C. Martin and M. Bevan. 1994. A flower-specific Myb protein activates transcription of phenylpropanoid biosynthetic genes. **EMBO J.** 13: 128-137.

Sakai, T., Y. Takahashi and T. Nagata. 1996. Analysis of the promoter of the auxin-inducible gene, *parC*, of tobacco. **Plant Cell Physiol.** 37: 906-913.

Sambrook, J., T. Maniatis and E.F. Fritsch. 1989. **Molecular Cloning: A Laboratory Mannual**. Cold Spring Harbor Laboratory, Cold Spring Harbarbor.

Sandal, N., K. Bojsen and K. Marcker. 2007. A small family of nodule specific genes from soybean. **Nucleic Acids Res.** 5: 1507-1519.

Sang-Hoon, L., A. Nagib, L. Ki-Won, K. Do-Hyun, L. Dong-Gi, K. Sang-Soo, K. Suk-Yoon, K. Tae-Hwan and L. Byung-Hyun. 2007. Simultaneous overexpression of both CuZn superoxide dismutase and ascorbate peroxidase in transgenic tall fescue plants confers increased tolerance to a wide range of abiotic. **JPLPH.** 164: 1626-1638.

Sanmiguel, P., A. Tikhonov, Y.K. Jin, N. Motchoulskaia, D. Zakharov, A. Melake-Berhan, P.S. Springer, K.J. Edwards, M. Lee, Z. Avramova, J.L. Bennetzen. 1996. Nested retrotransposons in the intergenic regions of the maize genome. **Science.** 274: 737-738.

Seki, H., Y. Ichinose, H. Kato, T. Shiraishi and T. Yamada. 1996. Analysis of *cis-regulatory* elements involved in the activation of a member of chalcone synthase gene family *(PsChsl)* in pea. **Plant Mol. Biol.** 31**:** 479-491.

Sell, S. and Hehl, R. 2005. A fifth member of the tomato 1-Aminocyclopropane-1-Carboxylic Acid (ACC) oxidase gene family harbours a leucine zipper and is anaerobically induced. **DNA Sequence.** 16: 80-82.

Shahmuradov, I.A., A.J. Gammerman, J.M. Hancock, H.P. Bramley and Y.V. Solovyev. 2003. Plant Prom: a database of plant promoter sequences. **Nucleic Acids Res.** 31**:** 114-117.

Shawn, L.A., G.R. Teakle, S.J. Martino-Catt and S.A. Kay. 1994. Circadian clock- and phytochrome-regulated transcription is conferred by a 78 bp *cis*-acting domain of the *Arabidopsis CAB2* promoter. **Plant J.** 6: 457-470.

Shen, W.H. and C. Gigot. 1997. Protein complexes binding to cis elements of the plant histone gene promoters: multiplicity, phosphorylation and cell cycle alteration. **Plant Mol. Biol**. 33**:** 367-379.

Shirsat, A., N. Wilford, R. Croy and D. Boulter. 1989. Sequence responsible for the tissue specific promoter activity of a pea legumin gene in tobacco. **Mol. Gen. Genet.** 15: 326-331.

Shu-Ying, Y., S. Ai-Qing, S. Yan, Y. Jin-Ying, Z. Chun-Mei and L. Jian. 2006. Differential regulation of *Lehsp23.8* in tomato plants: Analysis of a multiple stress-inducible promoter. **Plant Science.** 171: 398-407.

Siebert, P.D., A. Chenchik, D.E. Kellogg, K.A. Lukyanov and S.A. Lukyanov. 1995. An improved PCR method for walking in uncloned genomic DNA. **Nucleic Acids Res.** 23: 1087-1088.

Simpson, S.D., K. Nakashima, Y. Narusaka, M. Seki, K. Shinozaki and K. Yamagushi-Shinozaki. 2003. Two different novel *cis*-acting elements of erd1, aclpA homologous *Arabidopsis* gene function in introduction by dehydration stress and dark-induced senescence. **Plant J.** 33: 259-270.

Simran, B., C. Suma, A. Sonia, D. Sudipta, P. Deepak and K.B. Pradeep. 2003. Strategies for development of functionally equivalent promoters with minimum sequence homology for transgene expression in Plants: *cis*-elements in a novel DNA context versus domain swapping. **Plant Physiol.** 132: 988-998.

Solano, R., C. Nieto, J. Avila, L. Canas, I. Diaz and J. Paz-Ares. 1995. Dual DNA binding specificity of a petal epidermis-specific MYB transcription factor (MYB.Ph3) from *Petunia hybrid*. **EMBO J.** 14: 1773-1784.

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

Spolaore, S., L. Trainotti and G. Casadoro. 2001. A simple protocol for transient gene expression in ripe fleshy fruit mediated by *Agrobacterium*. **J. Exp. Bot.** 52: 845-850.

Staiger, D., H. Kaulen and J. Schell. 1989. A CACGTG motif of the *Antirrhinum majus* chalcone synthase promoter is recognized by an evolutionarily conserved nuclear protein. **Proc. Natl. Acad. Sci.** 86: 6930-6934.

Stalberg, K., M. Ellerstrom, I. Ezcurra, S. Ablov and L. Rask. 1996. Disruption of an overlapping E box/ABRE motif abolished high transcription of the napA storage protein promoter in transgenic *Brassica napus* seed. **Planta.** 199: 515-519.

Stougaard, J., J. Jorgensen, T. Christensen, A. Kuhle and K. Marcker. 1990. Interdependence and nodule specificity of *cis*-acting regulatory elements in the soybean leghemoglobin l*bc3* and *N23* gene promoters. **Mol. Gen. Genet**. 220: 353-360.

Sun, C., S. Palmqvist, H. Olsson, M. Boren, S. Ahlandsberg and C. Jansson. 2003. A novel WRKY transcription factor, SUSIBA2, participates in sugar signalling in barley by binding to the sugar-responsive elements of the *iso1* promoter. **Plant Cell.** 15: 2076-2092.

Takayashi, K., O. Tosshiro, S. Masami, J. Tetsuro, M. Rie, T. Keiji, M. Norihiro, K. Tetsu, K. Tetsuya, O. Kunio and S. Kazuo. 2005. Promoter of *Arabidopsis thaliana* phosphate transporter gene drives root-specific expression of transgene in rice. **SBB.** 99: 38-42.

Tang, W. and S. Perry. 2003. Binding site selection for the plant MADS domain protein AGL15: an in vitro and in vivo study. **J. Biol. Chem.** 278: 28154-28159.

Tang, X., H. Wang, A.S. Brandt and W.R. Woodson. 1993. Organization and structure of the 1-Aminocyclopropane-1-Carboxylate (ACC) Oxidase gene family from *Petunia hybrida*. **Plant Mol. Biol.** 23: 1151-1164.

Terauchi, R. and G. Kahl. 2000. Rapid isolation of promoter sequences by TAIL-PCR: the 5'flanking regions of *Pal* and *Pgi* genes from yam (*Dioscorea*). **Mol. Gen. Genet.** 263: 554-560.

Terzaghi, W.B. and A.R. Cashmore. 1995. Light-regulated transctription. **Annu. Rev. Plant Physiol.** 46: 445-474.

Thomas, M.S. and R.B. Flavell. 1990. Identification of an enhancer element for the endosperm specific expression of high molecular weight glutenin. **Plant Cell.** 2: 1171-1180.

Thum, K.E., M. Kim, D.T. Morishige, C. Eibl, H.U. Koop and J.E. Mullet. 2001. Analysis of barley chloroplast *psbD* light responsive promoter elements in transcriptomic tobacco. **Plant Mol. Biol.** 47: 353-366.

Tjaden, G., J. Edwards and G. Coruzzi. 1995. *Cis* elements and trans-acting factors affecting regulation of a nonphotosynthetic light-regulated gene for chloroplast glutamine synthetase. **Plant Physiol.** 108: 1109-1117.

Tremousaygue, D., G. Lionel, C. Bardet, P. Dabos, C. Herve and B. Lescure. 2003. Internal telomeric repeats and 'TCP domain' protein-binding site co-operate to regulate gene expression in *Arabidopsis thaliana* cycling cells. **Plant J.** 33: 957-966.

Urao, T., K. Yamaguchi-Shinozaki, S. Urao and K. Shinozaki. 1993. An Arabidopsis myb homolog is induced by dehydration stress and its gene product binds to the conserved MYB recognition sequence. **Plant Cell.** 5: 1529-1539.

Van Haaren, M.J. and C.M. Houck. 1993. A functional map of the fruit-specific promoter of the tomato *2A11* gene. **Plant Mol. Biol.** 21: 625-640.

Vandepoele, K., K. Vlieghe, K. Florquin, L. Hennig, G. Beemster, W. Gruissem, Y. Van De Peer, D. Inze and L. De Veylder. 2005. Genome-wide identification of potential plant *E2F* target genes. **Plant Physiol.** 139: 316-328.

Venter, M. 2006. Synthetic promoters: genetic control through *cis* engineering. **TRENDS in Plant Science.** 12: 118-124.

Vieweg, M., M. Fruhling, H. Quandt, U. Heim, B. Helmut, A. Puhler, H. Kuster and M. Andreas. 2004. The promoter of the *Vicia faba L.* leghemoglobin gene *VfLb29* is specifically activated in the infected cells of root nodules and in the arbuscule-containing cells of mycorrhizal roots from different legume and nonlegume plants. **Mol. Plant Microbe Interact**. 17: 62-69.

Von Gromoff, E.D., M. Schroda, U. Oster, C.F. Beck. 2006. Identification of a plastid response element that acts as an enhancer within the Chlamydomonas HSP70A promoter. **Nucleic Acids Res.** 34: 4767-4779.

Warner, S.A.J. 1996. Genomic DNA isolation and lambda library construction. *In* G.D. Foster and D. Twell, eds. **Plant Gene Isolation Principles and Practice**. John Wiley & Sons Ltd., England.

Washida, H., C.Y. Wu, A. Suzuki, U. Yamanoushi, T. Akihama, K. Harada and F. Takaiwa. 1999. Identification of *cis*-regulatory elements required for endosperm expression of the rice storage protein glutelin gene *GluB-1*. **Plant Mol. Biol.** 40: 1-12.

Wingender, E., X. Chen, E. Fricke, R. Geffers, R. Hehl, I. Liebich, M. Krull, V. Matys, H. Michael, R. Ohnhäuser, M. Prüß, F. Schacherer, S. Thieleand and S. Urbach. 2001. The TRANSFAC system on gene expression regulation. **Nucleic Acids Res.** 29:281-283.

Wu, C.Y., H. Washida, Y. Onodera, K. Harada and F. Tahaiwa. 2000. Quantitative nature of the Prolamin-box, ACGT ang AACA motifs in a rice glutelin gene promoter: minimal *cis*-element requirements for endosperm-specific gene expression. **Plant J.** 23: 415-421.

Yamagushi-Shinozaki, K. and K. Shinozaki. 1994. A novel *cis*- acting element in an *Arabidopsis* gene is involved in responsiveness to drouht, low temperature, or high-salt stress. **Plant Cell.** 6: 251-264.

Yamauchi, D., Z. Rodolfo and D.H. Tuan-hua. 2002. Molecular analysis of the barley (*Hordeum vulgare L.*) gene encoding the protein kinase PKABA1 capable of suppressing gibberellin action in aleurone layers**. Planta.** 215: 319-326.

Yanagisawa, S and R.J. Schmidt. 1999. Diversity and similarity among recognition sequences of Dof transcription factors. **Plant J.** 17: 209-214.

_____. 2000. Dof1 and Dof 2 transcription factors aer associated with expression of multiple genes involved in carbon metabolism in maize**. Plant J.** 21: 281-288.

Yang, S.F. and N.E. Hoffman. 1984. Ethylene biosynthesis and its regulation in higher plants. **Annu. Rev. Plant Physiol.** 35: 155-189.

Yongjin, S., K.E. Schwinn, M.J. Bennett, D.A. Hunter, T.L. Waugh, N.N. Pathirana, D.A. Brummell, P.E. Jameson and K.M. Davies. 2007. Methods for transient assay of gene function in floral tissues. **Plant Methods.** 3:1-12.

Yoshihara, T. and F. Takaiwa. 1996. CM-regulatory elements responsible for quantitative regulation of the rice seed storage protein glutelin *GluA-3* gene. **Plant Cell Physio.** 37: 107-111.

Yukawa, Y., M. Sugita, N. Choisne, I. Small and M. Sugiura. 2000. The TATA, the CAA motif and the poly(T) transcription termination motif are all important for transcription re-initiation on plant tRNA genes. **Plant J.** 22: 439-447.

Zhang, X.H. and V.L. Chiang. 1996. Single-stranded DNA ligation by T4 RNA ligase for PCR cloning of 5'-noncoding fragments and coding sequence of a specific gene. **Nucleic Acids Res.** 24: 990-991.

Zhang, Z., Z. Xie, X. Zou, J. Casaretto, T. Ho and Q. Shen. 2004. A rice *WRKY* gene encodes a transcriptional repressor of the gibberellin signaling pathway in aleurone cells. **Plant Physiol.** 134: 1500-1513.

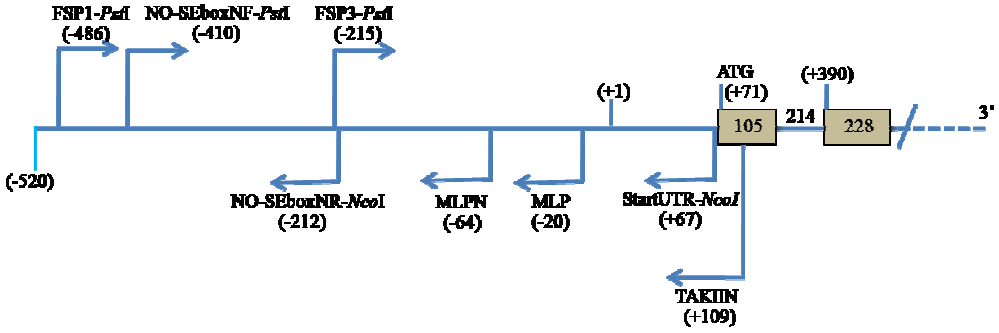
Zhou, D.X. 1999. Reguratory mechanism of plant gene transcription by GT-elements and GT-factors. **Trends in Plant Sci.** 4: 210-214.

Zhou, D.X., Y.F. Li, M. Rocipon and R. Mache. 1992. Sequence specific interaction between S1F, a spinach nuclear factor, and a negative *cis*-element conserved in plastid-related genes. **J. Biol. Chem.** 267: 23515-23519.

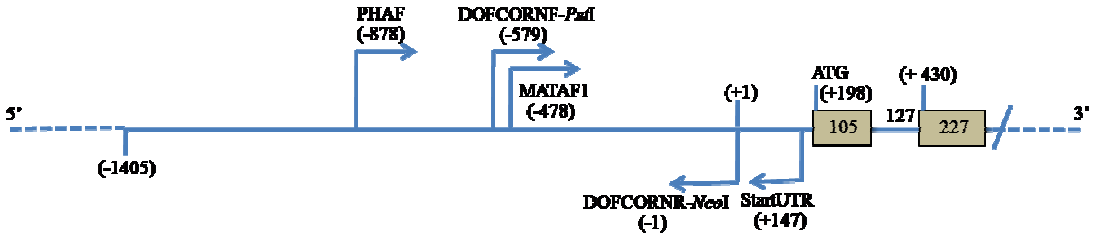
APPENDICES

Appendix A

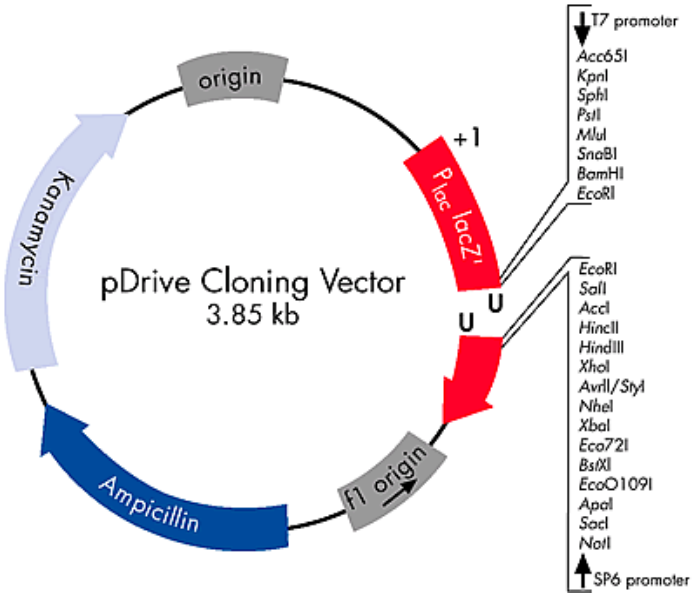
Isolation of *CP-ACO II* 5'flanking region and identification of *cis*-acting element



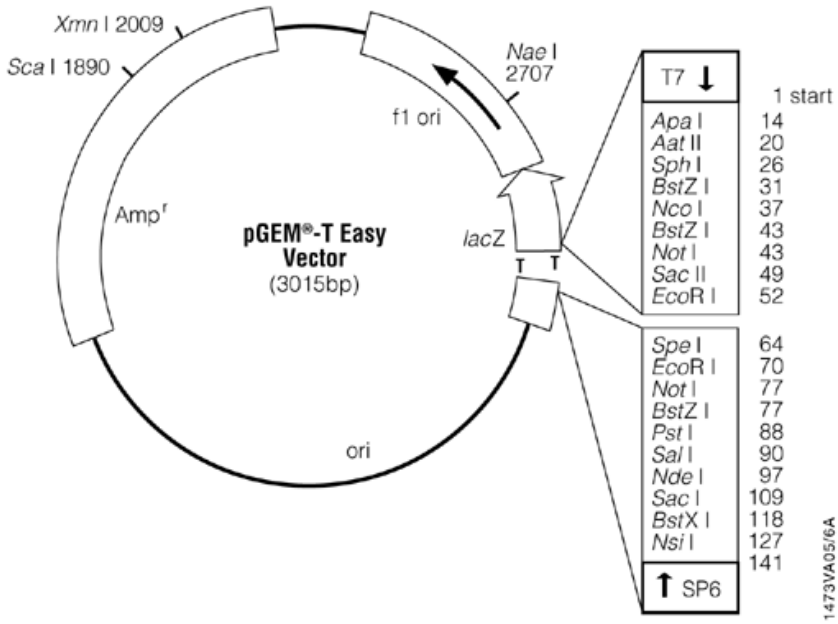
Appendix Figure 1 Location of *CP-ACO II* specific primers on 5'flanking region and 5'untranslated region of *CP-ACO II* gene. The position +1 referred to transcription start site, boxes refered to exons and linear line between boxes referred to introns. Primers were shown by arrows indicating the direction of extension.



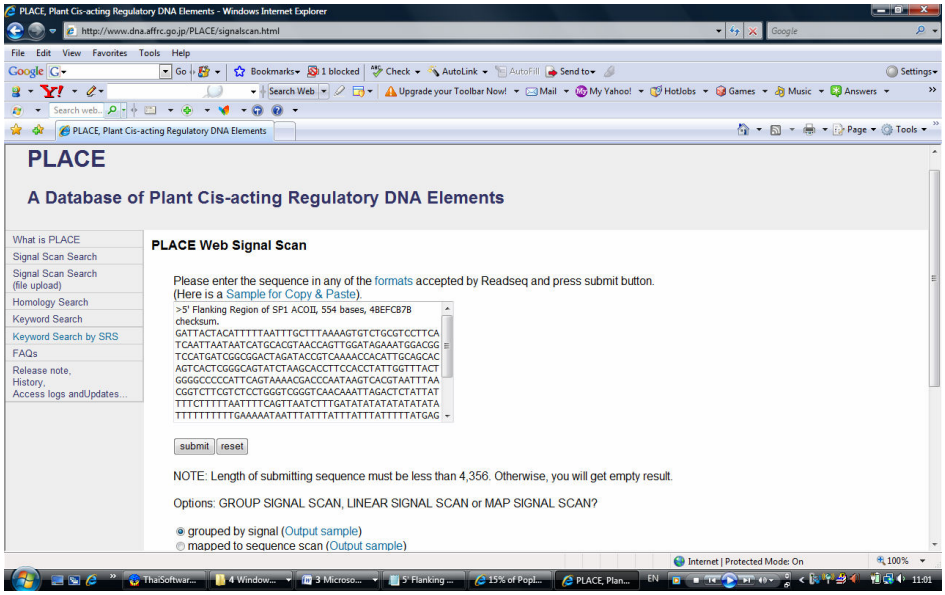
Appendix Figure 2 Location of *CP-ACO I* specific primers on 5'flanking region and 5'untranslated region of *CP-ACO I* gene. The position +1 referred to transcription start site, boxes refer to exons and linear line between boxes referred to introns. Primers were shown by arrows indicating the direction of extension.



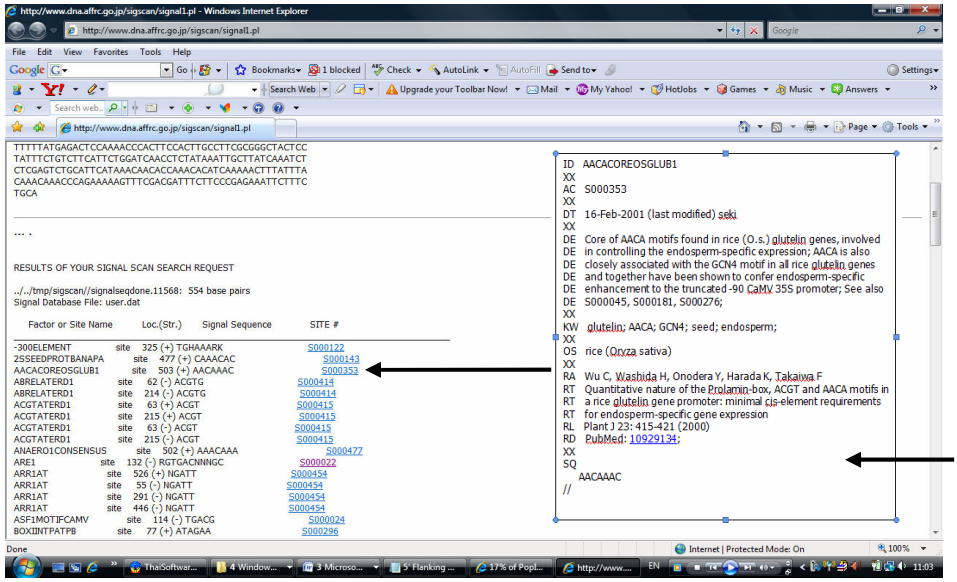
Appendix Figure 3 pDrive cloning vector (QIAGEN)



Appendix Figure 4 pGEM-T Easy vector (Promega)



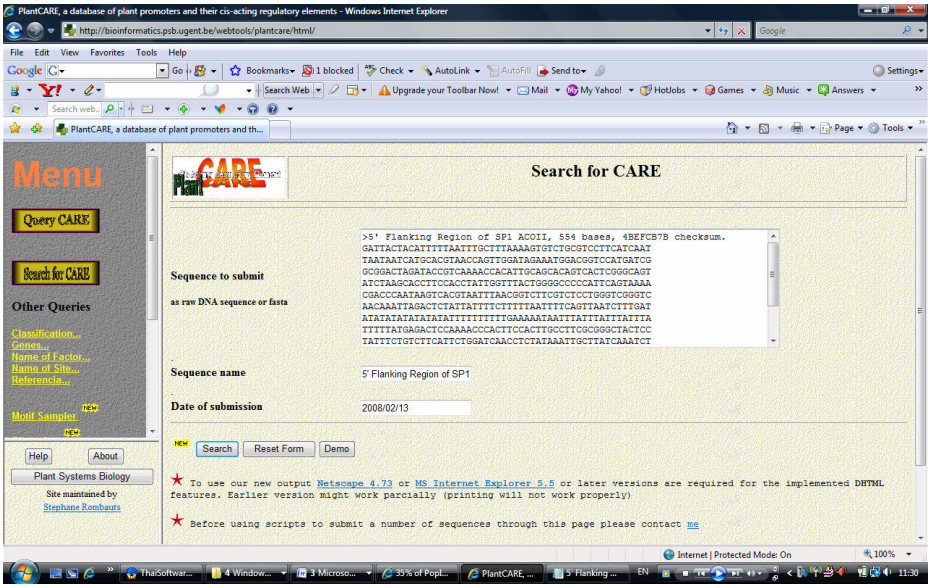
(a)



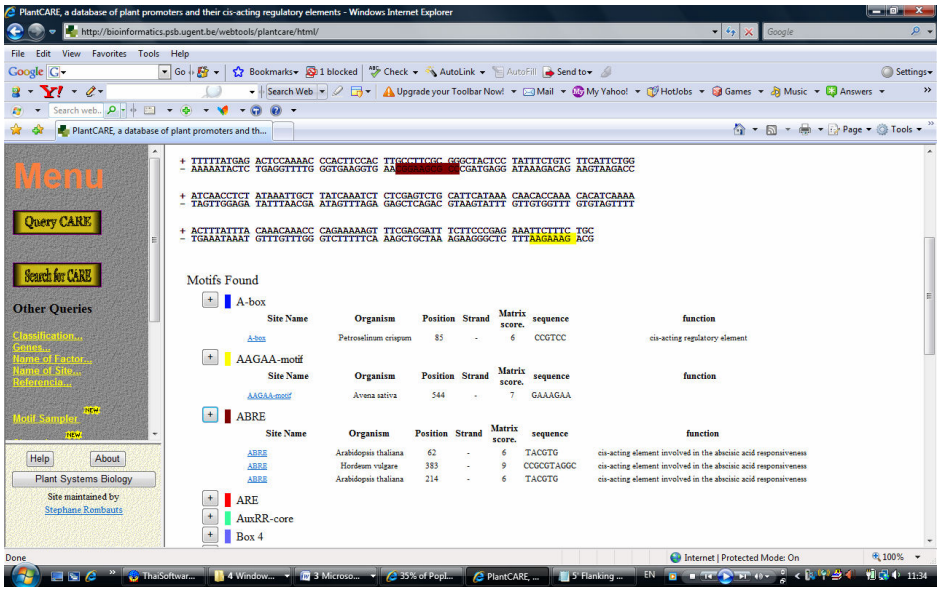
(c)

(b)

Appendix Figure 5 A demo of PLACE query using Signal Scan program. (a) A sample entry to the Signal Scan analysis page in the PLACE database. (b) A result reporting 'groupd-by-signal' format of Signal Scan (c) The document of the PLACE accession number S000353.

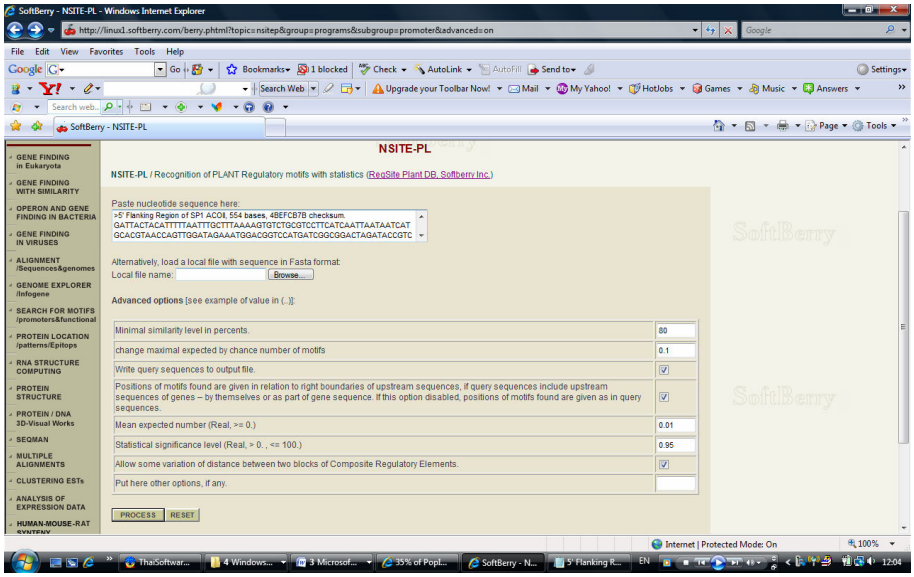


(a)

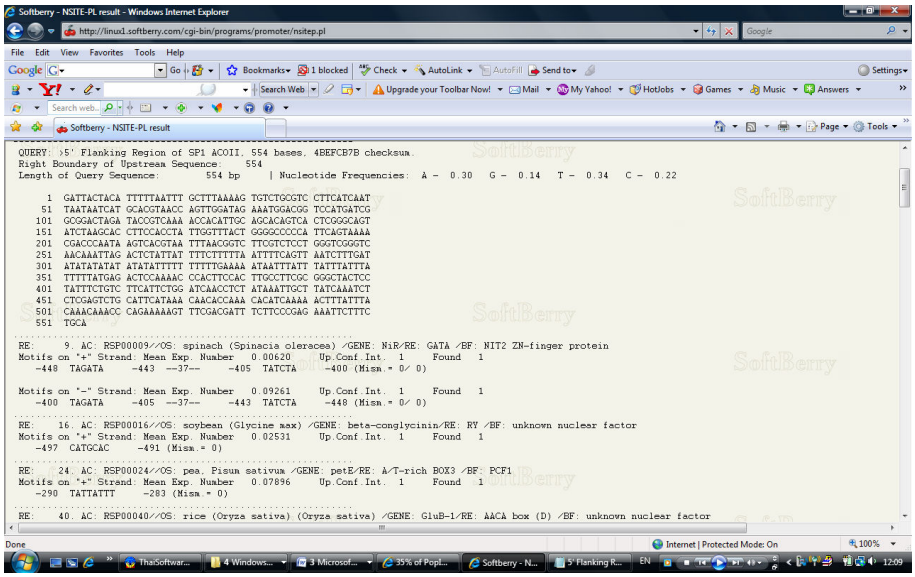


(b)

Appendix Figure 6 A demo of PlantCARE query. (a) A sample entry to the 'Search for CARE page' in the PlantCARE database. (b) A result reporting of 'Search for CARE page'

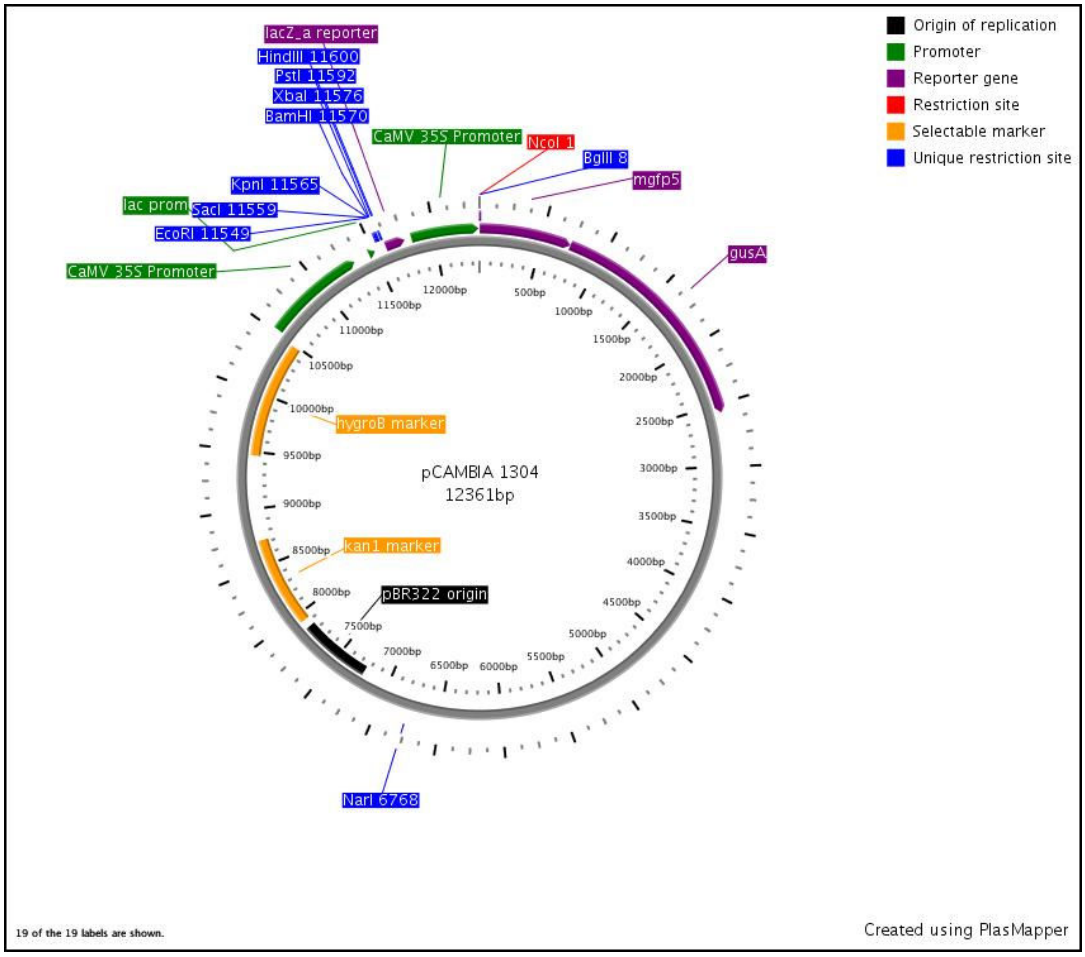


(a)



(b)

Appendix Figure 7 A demo of Plant Prom DB using NSITE-PL program. (a) A sample entry to the NSITE-PL page in the Plant Prom DB. (b) A result reporting 'groupd-by-signal' format of Signal Scan



Appendix Figure 8 Binary vector pCAMBIA1304

Appendix Table 1 List of primers used in PCR for *CP-ACO II* 5'flanking region isolation and expression vector construction.

Primer names	Sequences (5'-3')
Reverse primers	
DOFCORNR	aaa act gca gtt agt cgt aga aaa atg aat
StartUTR (*CP-ACO I*)	gtc tgc taa atc tct ccc tg
TAKIIN	ggt gag acc ctc cat gtt g
StartUTR (*CP-ACO II*)	tgc aga aag aat ttc tcg gg
MLP	atg aat gca gac tcg aga g
MLPN	gat cca gaa tga aga cag
StartUTR-*Nco*I	cat gca atg gtg cag aaa gaa ttt ctc ggg
NO-SEboxNR-*Nco*I	cat gca atg gaa aat aat aga gtc taa t
Forward primers	
DOFCORNF	aaa act gca gtt agt cgt aga aaa atg aat
PHAF	cag gga gtc acc tca gtt g
MataF1	gta aca ggc agg aca aaa gag
API	gga tcc taa tac gac tca cta tag ggc
NAPI	tat agg gct cga gcg gc
FSPI	gat tac tac att ttt aat ttg c
FSP1-PstI	aaa act gca gga tta cta cat ttt taa ttt gc
NO-SEboxNF-*Pst*I	aaa act gca gga tag aaa tgg acg
FSP3-*Pst*I	aaa act gca gtt ttc ttt tta att ttc agt t

Appendix Table 2 Solutions for preparing 12% separating gel for Tris-glycine SDS-Polyacrylamide gel electrophoresis

Solution components	**Component volumes (ml) (20 ml total volumes)**
H_2O	6.6
30% acrylamide mix	8.0
1.5M Tris (pH8.8)	5.0
10% SDS	0.2
10% ammonium persulfate	0.2
TEMED	0.008

Appendix Table 3 Solutions for preparing 5% stackingting gel for Tris-glycine SDS-Polyacrylamide gel electrophoresis

Solution components	**Component volumes (ml) (4 ml total volumes)**
H_2O	2.7
30% acrylamide mix	0.67
1.0M Tris (pH6.8)	0.5
10% SDS	0.04
10% ammonium persulfate	0.04
TEMED	0.004

Appendix Table 4 The presence of *cis*-acting elements in *CP-ACO I* 5'flanking region using PLACE Database

Factor or Site Name		Location	Strand	Signal Sequence	SITE
-10PEHVPSBD	site	294	(-)	TATTCT	S000392
-300CORE	site	28	(-)	TGTAAAG	S000001
-300ELEMENT	site	27	(-)	TGHAAARK	S000122
2SSEEDPROTBANAPA	site	883	(+)	CAAACAC	S000143
ABRELATERD1	site	47	(+)	ACGTG	S000414
ACGTATERD1	site	47	(+)	ACGT	S000415
ACGTATERD1	site	47	(-)	ACGT	S000415
ANAERO1CONSENSUS	site	991	(-)	AAACAAA	S000477
ANAERO3CONSENSUS	site	869	(+)	TCATCAC	S000479
ARR1AT	site	1032	(+)	NGATT	S000454
ARR1AT	site	460	(+)	NGATT	S000454
ARR1AT	site	997	(+)	NGATT	S000454
ARR1AT	site	727	(-)	NGATT	S000454
ASF1MOTIFCAMV	site	391	(+)	TGACG	S000024
BIHD1OS	site	950	(+)	TGTCA	S000498
BIHD1OS	site	831	(-)	TGTCA	S000498
BOXLCOREDCPAL	site	975	(+)	ACCWWCC	S000492
CAATBOX1	site	35	(+)	CAAT	S000028
CAATBOX1	site	637	(+)	CAAT	S000028
CAATBOX1	site	762	(+)	CAAT	S000028
CAATBOX1	site	958	(+)	CAAT	S000028
CAATBOX1	site	639	(-)	CAAT	S000028
CAATBOX1	site	781	(-)	CAAT	S000028
CAATBOX1	site	829	(-)	CAAT	S000028
CAATBOX1	site	841	(-)	CAAT	S000028
CAATBOX1	site	999	(-)	CAAT	S000028
CACTFTPPCA1	site	26	(+)	YACT	S000449
CACTFTPPCA1	site	465	(+)	YACT	S000449
CACTFTPPCA1	site	647	(+)	YACT	S000449
CACTFTPPCA1	site	906	(+)	YACT	S000449
CACTFTPPCA1	site	123	(+)	YACT	S000449
CACTFTPPCA1	site	717	(+)	YACT	S000449
CACTFTPPCA1	site	823	(+)	YACT	S000449
CACTFTPPCA1	site	167	(-)	YACT	S000449
CACTFTPPCA1	site	363	(-)	YACT	S000449
CACTFTPPCA1	site	589	(-)	YACT	S000449
CACTFTPPCA1	site	598	(-)	YACT	S000449
CACTFTPPCA1	site	770	(-)	YACT	S000449
CACTFTPPCA1	site	948	(-)	YACT	S000449
CANBNNAPA	site	883	(+)	CNAACAC	S000148
CANBNNAPA	site	590	(-)	CNAACAC	S000148
CCAATBOX1	site	636	(+)	CCAAT	S000030
CCAATBOX1	site	957	(+)	CCAAT	S000030
CCAATBOX1	site	781	(-)	CCAAT	S000030
CCAATBOX1	site	841	(-)	CCAAT	S000030
CGACGOSAMY3	site	431	(+)	CGACG	S000205
CPBCSPOR	site	298	(+)	TATTAG	S000491
CURECORECR	site	444	(+)	GTAC	S000493
CURECORECR	site	523	(+)	GTAC	S000493
CURECORECR	site	567	(+)	GTAC	S000493
CURECORECR	site	444	(-)	GTAC	S000493
CURECORECR	site	523	(-)	GTAC	S000493
CURECORECR	site	567	(-)	GTAC	S000493
DOFCOREZM	site	193	(+)	AAAG	S000265
DOFCOREZM	site	346	(+)	AAAG	S000265
DOFCOREZM	site	353	(+)	AAAG	S000265
DOFCOREZM	site	361	(+)	AAAG	S000265
DOFCOREZM	site	366	(+)	AAAG	S000265
DOFCOREZM	site	416	(+)	AAAG	S000265
DOFCOREZM	site	606	(+)	AAAG	S000265
DOFCOREZM	site	623	(+)	AAAG	S000265
DOFCOREZM	site	672	(+)	AAAG	S000265
DOFCOREZM	site	766	(+)	AAAG	S000265

Appendix Table 4 (Continued)

Factor or Site Name		Location	Strand	Signal Sequence	SITE
DOFCOREZM	site	929	(+)	AAAG	S000265
DOFCOREZM	site	936	(+)	AAAG	S000265
DOFCOREZM	site	966	(+)	AAAG	S000265
DOFCOREZM	site	1077	(+)	AAAG	S000265
DOFCOREZM	site	1014	(+)	AAAG	S000265
DOFCOREZM	site	28	(-)	AAAG	S000265
DOFCOREZM	site	114	(-)	AAAG	S000265
DOFCOREZM	site	470	(-)	AAAG	S000265
DOFCOREZM	site	477	(-)	AAAG	S000265
DOFCOREZM	site	615	(-)	AAAG	S000265
DOFCOREZM	site	804	(-)	AAAG	S000265
DOFCOREZM	site	814	(-)	AAAG	S000265
DOFCOREZM	site	990	(-)	AAAG	S000265
DPBFCOREDCDC3	site	537	(+)	ACACNNG	S000292
DPBFCOREDCDC3	site	48	(-)	ACACNNG	S000292
DPBFCOREDCDC3	site	434	(-)	ACACNNG	S000292
E2FCONSENSUS	site	784	(-)	WTTSSCSS	S000476
EBOXBNNAPA	site	14	(+)	CANNTG	S000144
EBOXBNNAPA	site	387	(+)	CANNTG	S000144
EBOXBNNAPA	site	538	(+)	CANNTG	S000144
EBOXBNNAPA	site	637	(+)	CANNTG	S000144
EBOXBNNAPA	site	14	(-)	CANNTG	S000144
EBOXBNNAPA	site	387	(-)	CANNTG	S000144
EBOXBNNAPA	site	538	(-)	CANNTG	S000144
EBOXBNNAPA	site	637	(-)	CANNTG	S000144
EECCRCAH1	site	788	(+)	GANTTNC	S000494
EECCRCAH1	site	787	(-)	GANTTNC	S000494
EMHVCHORD	site	27	(-)	TGTAAAGT	S000452
GATABOX	site	159	(+)	GATA	S000039
GATABOX	site	332	(+)	GATA	S000039
GATABOX	site	395	(+)	GATA	S000039
GATABOX	site	656	(+)	GATA	S000039
GATABOX	site	752	(+)	GATA	S000039
GATABOX	site	93	(-)	GATA	S000039
GATABOX	site	111	(-)	GATA	S000039
GT1CONSENSUS	site	309	(+)	GRWAAW	S000198
GT1CONSENSUS	site	349	(+)	GRWAAW	S000198
GT1CONSENSUS	site	357	(+)	GRWAAW	S000198
GT1CONSENSUS	site	358	(+)	GRWAAW	S000198
GT1CONSENSUS	site	926	(+)	GRWAAW	S000198
GT1CONSENSUS	site	933	(+)	GRWAAW	S000198
GT1CONSENSUS	site	1002	(+)	GRWAAW	S000198
GT1CONSENSUS	site	85	(-)	GRWAAW	S000198
GT1CONSENSUS	site	798	(-)	GRWAAW	S000198
GT1CONSENSUS	site	471	(-)	GRWAAW	S000198
GT1CONSENSUS	site	472	(-)	GRWAAW	S000198
GT1CONSENSUS	site	627	(-)	GRWAAW	S000198
GT1GMSCAM4	site	309	(+)	GAAAAA	S000453
GT1GMSCAM4	site	349	(+)	GAAAAA	S000453
GT1GMSCAM4	site	358	(+)	GAAAAA	S000453
GT1GMSCAM4	site	926	(+)	GAAAAA	S000453
GT1GMSCAM4	site	1002	(+)	GAAAAA	S000453
GT1GMSCAM4	site	471	(-)	GAAAAA	S000453
GTGANTG10	site	439	(+)	GTGA	S000378
GTGANTG10	site	586	(+)	GTGA	S000378
GTGANTG10	site	8	(-)	GTGA	S000378
GTGANTG10	site	380	(-)	GTGA	S000378
GTGANTG10	site	464	(-)	GTGA	S000378
GTGANTG10	site	872	(-)	GTGA	S000378
HEXAMERATH4	site	431	(-)	CCGTCG	S000146
IBOX	site	332	(+)	GATAAG	S000124
IBOXCORE	site	332	(+)	GATAA	S000199
IBOXCORENT	site	332	(+)	GATAAGR	S000424
INRNTPSADB	site	151	(-)	YTCANTYY	S000395
INRNTPSADB	site	311	(-)	YTCANTYY	S000395

Appendix Table 4 (Continued)

Factor or Site Name		Location	Strand	Signal Sequence	SITE
INRNTPSADB	site	336	(-)	YTCANTYY	S000395
LTRECOREATCOR15	site	430	(+)	CCGAC	S000153
MYB1AT	site	257	(+)	WAACCA	S000408
MYB1AT	site	650	(-)	WAACCA	S000408
MYB2CONSENSUSAT	site	14	(-)	YAACKG	S000409
MYBCORE	site	14	(+)	CNGTTR	S000176
MYBCORE	site	402	(-)	CNGTTR	S000176
MYBPZM	site	972	(+)	CCWACC	S000179
MYBPZM	site	421	(-)	CCWACC	S000179
MYCATERD1	site	538	(-)	CATGTG	S000413
MYCATRD22	site	538	(+)	CACATG	S000174
MYCCONSENSUSAT	site	14	(+)	CANNTG	S000407
MYCCONSENSUSAT	site	387	(+)	CANNTG	S000407
MYCCONSENSUSAT	site	538	(+)	CANNTG	S000407
MYCCONSENSUSAT	site	637	(+)	CANNTG	S000407
MYCCONSENSUSAT	site	14	(-)	CANNTG	S000407
MYCCONSENSUSAT	site	387	(-)	CANNTG	S000407
MYCCONSENSUSAT	site	538	(-)	CANNTG	S000407
MYCCONSENSUSAT	site	637	(-)	CANNTG	S000407
NAPINMOTIFBN	site	563	(-)	TACACAT	S000070
NODCON1GM	site	112	(-)	AAAGAT	S000461
NODCON2GM	site	497	(+)	CTCTT	S000462
NODCON2GM	site	812	(+)	CTCTT	S000462
NODCON2GM	site	417	(-)	CTCTT	S000462
NODCON2GM	site	607	(-)	CTCTT	S000462
NODCON2GM	site	767	(-)	CTCTT	S000462
NODCON2GM	site	942	(-)	CTCTT	S000462
NODCON2GM	site	967	(-)	CTCTT	S000462
NODCON2GM	site	1008	(-)	CTCTT	S000462
NTBBF1ARROLB	site	27	(+)	ACTTTA	S000273
NTBBF1ARROLB	site	192	(-)	ACTTTA	S000273
OSE1ROOTNODULE	site	112	(-)	AAAGAT	S000467
OSE2ROOTNODULE	site	497	(+)	CTCTT	S000468
OSE2ROOTNODULE	site	812	(+)	CTCTT	S000468
OSE2ROOTNODULE	site	417	(-)	CTCTT	S000468
OSE2ROOTNODULE	site	607	(-)	CTCTT	S000468
OSE2ROOTNODULE	site	767	(-)	CTCTT	S000468
OSE2ROOTNODULE	site	942	(-)	CTCTT	S000468
OSE2ROOTNODULE	site	967	(-)	CTCTT	S000468
OSE2ROOTNODULE	site	1008	(-)	CTCTT	S000468
POLASIG1	site	147	(+)	AATAAA	S000080
POLASIG1	site	248	(+)	AATAAA	S000080
POLASIG1	site	763	(+)	AATAAA	S000080
POLASIG2	site	211	(+)	AATTAAA	S000081
POLASIG2	site	235	(+)	AATTAAA	S000081
POLASIG2	site	721	(+)	AATTAAA	S000081
POLASIG2	site	115	(-)	AATTAAA	S000081
POLASIG3	site	702	(+)	AATAAT	S000088
POLASIG3	site	795	(-)	AATAAT	S000088
POLLEN1LELAT52	site	308	(+)	AGAAA	S000245
POLLEN1LELAT52	site	343	(+)	AGAAA	S000245
POLLEN1LELAT52	site	348	(+)	AGAAA	S000245
POLLEN1LELAT52	site	604	(+)	AGAAA	S000245
POLLEN1LELAT52	site	925	(+)	AGAAA	S000245
POLLEN1LELAT52	site	1011	(+)	AGAAA	S000245
POLLEN1LELAT52	site	494	(-)	AGAAA	S000245
POLLEN1LELAT52	site	911	(-)	AGAAA	S000245
PRECONSCRHSP70A	site	430	(+)	SCGAYNRNNNNNNNNNNNNNNHD	S000506
PROXBBNNAPA	site	883	(+)	CAAACACC	S000263
PYRIMIDINEBOXOSRAMY1A	site	476	(+)	CCTTTT	S000259
PYRIMIDINEBOXOSRAMY1A	site	352	(-)	CCTTTT	S000259
PYRIMIDINEBOXOSRAMY1A	site	928	(-)	CCTTTT	S000259
RAV1AAT	site	903	(+)	CAACA	S000314
RAV1AAT	site	591	(-)	CAACA	S000314
REALPHALGLHCB21	site	258	(+)	AACCAA	S000362

Appendix Table 4 (Continued)

Factor or Site Name		Location	Strand	Signal Sequence	SITE
ROOTMOTIFTAPOX1	site	160	(+)	ATATT	S000098
ROOTMOTIFTAPOX1	site	184	(+)	ATATT	S000098
ROOTMOTIFTAPOX1	site	264	(+)	ATATT	S000098
ROOTMOTIFTAPOX1	site	297	(+)	ATATT	S000098
ROOTMOTIFTAPOX1	site	491	(+)	ATATT	S000098
ROOTMOTIFTAPOX1	site	36	(-)	ATATT	S000098
ROOTMOTIFTAPOX1	site	181	(-)	ATATT	S000098
ROOTMOTIFTAPOX1	site	252	(-)	ATATT	S000098
ROOTMOTIFTAPOX1	site	263	(-)	ATATT	S000098
ROOTMOTIFTAPOX1	site	296	(-)	ATATT	S000098
ROOTMOTIFTAPOX1	site	504	(-)	ATATT	S000098
ROOTMOTIFTAPOX1	site	733	(-)	ATATT	S000098
ROOTMOTIFTAPOX1	site	738	(-)	ATATT	S000098
ROOTMOTIFTAPOX1	site	743	(-)	ATATT	S000098
S1FBOXSORPS1L21	site	398	(+)	ATGGTA	S000223
SEF1MOTIF	site	264	(+)	ATATTTAWW	S000006
SEF1MOTIF	site	248	(-)	ATATTTAWW	S000006
SEF3MOTIFGM	site	954	(+)	AACCCA	S000115
SEF4MOTIFGM7S	site	170	(+)	RTTTTTR	S000103
SEF4MOTIFGM7S	site	221	(+)	RTTTTTR	S000103
SEF4MOTIFGM7S	site	694	(+)	RTTTTTR	S000103
SEF4MOTIFGM7S	site	149	(-)	RTTTTTR	S000103
SORLIP1AT	site	599	(-)	GCCAC	S000482
SORLREP3AT	site	52	(+)	TGTATATAT	S000488
SORLREP3AT	site	37	(-)	TGTATATAT	S000488
SP8BFIBSP8BIB	site	123	(+)	TACTATT	S000184
TAAAGSTKST1	site	192	(+)	TAAAG	S000387
TAAAGSTKST1	site	365	(+)	TAAAG	S000387
TAAAGSTKST1	site	765	(+)	TAAAG	S000387
TAAAGSTKST1	site	965	(+)	TAAAG	S000387
TAAAGSTKST1	site	28	(-)	TAAAG	S000387
TAAAGSTKST1	site	114	(-)	TAAAG	S000387
TAAAGSTKST1	site	615	(-)	TAAAG	S000387
TATABOX2	site	850	(+)	TATAAAT	S000109
TATABOX2	site	229	(-)	TATAAAT	S000109
TATABOX2	site	266	(-)	TATAAAT	S000109
TATABOX3	site	144	(-)	TATTAAT	S000110
TATABOX3	site	245	(-)	TATTAAT	S000110
TATABOX4	site	268	(-)	TATATAA	S000111
TATABOXOSPAL	site	126	(+)	TATTTAA	S000400
TATAPVTRNALEU	site	267	(+)	TTTATATA	S000340
TBOXATGAPB	site	622	(-)	ACTTTG	S000383
WBBOXPCWRKY1	site	619	(-)	TTTGACY	S000310
WBOXATNPR1	site	830	(+)	TTGAC	S000390
WBOXATNPR1	site	620	(-)	TTGAC	S000390
WBOXATNPR1	site	951	(-)	TTGAC	S000390
WBOXHVISO1	site	6	(-)	TGACT	S000442
WBOXHVISO1	site	619	(-)	TGACT	S000442
WBOXHVISO1	site	659	(-)	TGACT	S000442
WBOXNTERF3	site	6	(-)	TGACY	S000457
WBOXNTERF3	site	619	(-)	TGACY	S000457
WBOXNTERF3	site	659	(-)	TGACY	S000457
WRKY71OS	site	391	(+)	TGAC	S000447
WRKY71OS	site	831	(+)	TGAC	S000447
WRKY71OS	site	7	(-)	TGAC	S000447
WRKY71OS	site	620	(-)	TGAC	S000447
WRKY71OS	site	660	(-)	TGAC	S000447
WRKY71OS	site	951	(-)	TGAC	S000447
XYLAT	site	988	(-)	ACAAAGAA	S000510
ZDNAFORMINGATCAB	site	45	(+)	ATACGTGT	S000321

Appendix Table 5 The presence of *cis*-acting elements in *CP-ACO II* 5'flanking region using PLACE Database

Factor or Site Name		Location	Strand	Signal Sequence	SITE
-300ELEMENT	site	325	(+)	TGHAAARK	S000122
2SSEEDPROTBANAPA	site	477	(+)	CAAACAC	S000143
AACACOREOSGLUB1	site	503	(+)	AACAAAC	S000353
ABRELATERD1	site	62	(-)	ACGTG	S000414
ABRELATERD1	site	214	(-)	ACGTG	S000414
ACGTATERD1	site	63	(+)	ACGT	S000415
ACGTATERD1	site	215	(+)	ACGT	S000415
ACGTATERD1	site	63	(-)	ACGT	S000415
ACGTATERD1	site	215	(-)	ACGT	S000415
ANAERO1CONSENSUS	site	502	(+)	AAACAAA	S000477
ARE1	site	132	(-)	RGTGACNNNGC	S000022
ARR1AT	site	526	(+)	NGATT	S000454
ARR1AT	site	55	(-)	NGATT	S000454
ARR1AT	site	291	(-)	NGATT	S000454
ARR1AT	site	446	(-)	NGATT	S000454
ASF1MOTIFCAMV	site	114	(-)	TGACG	S000024
BOXIINTPATPB	site	77	(+)	ATAGAA	S000296
BOXLCOREDCPAL	site	159	(+)	ACCWWCC	S000492
CAATBOX1	site	47	(+)	CAAT	S000028
CAATBOX1	site	206	(+)	CAAT	S000028
CAATBOX1	site	126	(-)	CAAT	S000028
CAATBOX1	site	170	(-)	CAAT	S000028
CAATBOX1	site	435	(-)	CAAT	S000028
CACTFTPPCA1	site	139	(+)	YACT	S000449
CACTFTPPCA1	site	372	(+)	YACT	S000449
CACTFTPPCA1	site	378	(+)	YACT	S000449
CACTFTPPCA1	site	4	(+)	YACT	S000449
CACTFTPPCA1	site	177	(+)	YACT	S000449
CACTFTPPCA1	site	395	(+)	YACT	S000449
CACTFTPPCA1	site	29	(-)	YACT	S000449
CACTFTPPCA1	site	148	(-)	YACT	S000449
CACTFTPPCA1	site	194	(-)	YACT	S000449
CANBNNAPA	site	477	(+)	CNAACAC	S000148
CARGCW8GAT	site	429	(+)	CWWWWWWWG	S000431
CARGCW8GAT	site	429	(-)	CWWWWWWWG	S000431
CCAATBOX1	site	205	(+)	CCAAT	S000030
CCAATBOX1	site	170	(-)	CCAAT	S000030
CGACGOSAMY3	site	523	(+)	CGACG	S000205
CIACADIANLELHC	site	477	(+)	CAANNNNATC	S000252
CMSRE1IBSPOA	site	84	(+)	TGGACGG	S000511
DOFCOREZM	site	27	(+)	AAAG	S000265
DOFCOREZM	site	516	(+)	AAAG	S000265
DOFCOREZM	site	22	(-)	AAAG	S000265
DOFCOREZM	site	274	(-)	AAAG	S000265
DOFCOREZM	site	294	(-)	AAAG	S000265
DOFCOREZM	site	492	(-)	AAAG	S000265
DOFCOREZM	site	546	(-)	AAAG	S000265
EBOXBNNAPA	site	70	(+)	CANNTG	S000144
EBOXBNNAPA	site	378	(+)	CANNTG	S000144
EBOXBNNAPA	site	70	(-)	CANNTG	S000144
EBOXBNNAPA	site	378	(-)	CANNTG	S000144
EECCRCAH1	site	540	(-)	GANTTNC	S000494
ELRECOREPCRP1	site	247	(-)	TTGACC	S000142
GATABOX	site	76	(+)	GATA	S000039
GATABOX	site	109	(+)	GATA	S000039
GATABOX	site	298	(+)	GATA	S000039
GATABOX	site	150	(-)	GATA	S000039
GATABOX	site	441	(-)	GATA	S000039
GT1CONSENSUS	site	326	(+)	GRWAAW	S000198
GT1CONSENSUS	site	513	(+)	GRWAAW	S000198
GT1CONSENSUS	site	269	(-)	GRWAAW	S000198
GT1CONSENSUS	site	281	(-)	GRWAAW	S000198
GT1GMSCAM4	site	326	(+)	GAAAAA	S000453

Appendix Table 5 (Continued)

Factor or Site Name		Location	Strand	Signal Sequence	SITE
GT1GMSCAM4	site	513	(+)	GAAAAA	S000453
GTGANTG10	site	138	(-)	GTGA	S000378
GTGANTG10	site	213	(-)	GTGA	S000378
IBOX	site	439	(-)	GATAAG	S000124
IBOXCORE	site	440	(-)	GATAA	S000199
INRNTPSADB	site	411	(+)	YTCANTYY	S000395
LTRECOREATCOR15	site	243	(-)	CCGAC	S000153
MARABOX1	site	335	(-)	AATAAAYAAA	S000063
MARABOX1	site	339	(-)	AATAAAYAAA	S000063
MARABOX1	site	343	(-)	AATAAAYAAA	S000063
MARTBOX	site	316	(+)	TTWTWTTWTT	S000067
MYB1AT	site	119	(+)	WAACCA	S000408
MYB1AT	site	66	(+)	WAACCA	S000408
MYB1AT	site	172	(-)	WAACCA	S000408
MYB2AT	site	286	(-)	TAACTG	S000177
MYB2CONSENSUSAT	site	223	(+)	YAACKG	S000409
MYB2CONSENSUSAT	site	70	(-)	YAACKG	S000409
MYB2CONSENSUSAT	site	286	(-)	YAACKG	S000409
MYBCORE	site	70	(+)	CNGTTR	S000176
MYBCORE	site	286	(+)	CNGTTR	S000176
MYBCORE	site	223	(-)	CNGTTR	S000176
MYBCOREATCYCB1	site	224	(+)	AACGG	S000502
MYBPLANT	site	474	(+)	MACCWAMC	S000167
MYBST1	site	75	(+)	GGATA	S000180
MYCCONSENSUSAT	site	70	(+)	CANNTG	S000407
MYCCONSENSUSAT	site	378	(+)	CANNTG	S000407
MYCCONSENSUSAT	site	70	(-)	CANNTG	S000407
MYCCONSENSUSAT	site	378	(-)	CANNTG	S000407
NODCON1GM	site	292	(-)	AAAGAT	S000461
NTBBF1ARROLB	site	491	(+)	ACTTTA	S000273
OSE1ROOTNODULE	site	292	(-)	AAAGAT	S000467
PALBOXAPC	site	85	(-)	CCGTCC	S000137
POLASIG1	site	335	(-)	AATAAA	S000080
POLASIG1	site	339	(-)	AATAAA	S000080
POLASIG1	site	343	(-)	AATAAA	S000080
POLASIG1	site	347	(-)	AATAAA	S000080
POLASIG1	site	493	(-)	AATAAA	S000080
POLASIG2	site	13	(-)	AATTAAA	S000081
POLASIG2	site	277	(-)	AATTAAA	S000081
POLASIG3	site	52	(+)	AATAAT	S000088
POLASIG3	site	330	(+)	AATAAT	S000088
POLASIG3	site	266	(-)	AATAAT	S000088
POLLEN1LELAT52	site	79	(+)	AGAAA	S000245
POLLEN1LELAT52	site	512	(+)	AGAAA	S000245
POLLEN1LELAT52	site	539	(+)	AGAAA	S000245
POLLEN1LELAT52	site	271	(-)	AGAAA	S000245
POLLEN1LELAT52	site	403	(-)	AGAAA	S000245
POLLEN1LELAT52	site	529	(-)	AGAAA	S000245
POLLEN1LELAT52	site	547	(-)	AGAAA	S000245
PRECONSCRHSP70A	site	78	(-)	SCGAYNRNNNNNNNNNNNNNNHD	S000506
RAV1AAT	site	250	(+)	CAACA	S000314
RAV1AAT	site	471	(+)	CAACA	S000314
REALPHALGLHCB21	site	171	(-)	AACCAA	S000362
RHERPATEXPA7	site	61	(+)	KCACGW	S000512
RHERPATEXPA7	site	213	(+)	KCACGW	S000512
ROOTMOTIFTAPOX1	site	313	(+)	ATATT	S000098
RYREPEATBNNAPA	site	58	(+)	CATGCA	S000264
RYREPEATLEGUMINBOX	site	58	(+)	CATGCAY	S000100
SEF3MOTIFGM	site	368	(+)	AACCCA	S000115
SEF3MOTIFGM	site	507	(+)	AACCCA	S000115
SEF4MOTIFGM7S	site	10	(+)	RTTTTTR	S000103
SEF4MOTIFGM7S	site	350	(+)	RTTTTTR	S000103
SEF4MOTIFGM7S	site	486	(-)	RTTTTTR	S000103
SORLIP2AT	site	182	(+)	GGGCC	S000483
SORLIP2AT	site	183	(-)	GGGCC	S000483

Appendix Table 5 (Continued)

Factor or Site Name		Location	Strand	Signal Sequence	SITE
SURECOREATSULTR11	site	358	(+)	GAGAC	S000499
SURECOREATSULTR11	site	234	(-)	GAGAC	S000499
TAAAGSTKST1	site	22	(-)	TAAAG	S000387
TAAAGSTKST1	site	492	(-)	TAAAG	S000387
TATABOX2	site	430	(+)	TATAAAT	S000109
TATABOX3	site	49	(-)	TATTAAT	S000110
TATABOX5	site	267	(+)	TTATTT	S000203
TATABOX5	site	336	(+)	TTATTT	S000203
TATABOX5	site	340	(+)	TTATTT	S000203
TATABOX5	site	344	(+)	TTATTT	S000203
TATABOX5	site	348	(+)	TTATTT	S000203
TATABOX5	site	494	(+)	TTATTT	S000203
TATABOX5	site	329	(-)	TTATTT	S000203
TATCCAOSAMY	site	74	(-)	TATCCA	S000403
WBOXATNPR1	site	115	(-)	TTGAC	S000390
WBOXATNPR1	site	248	(-)	TTGAC	S000390
WBOXHVISO1	site	136	(-)	TGACT	S000442
WBOXHVISO1	site	211	(-)	TGACT	S000442
WBOXNTERF3	site	136	(-)	TGACY	S000457
WBOXNTERF3	site	211	(-)	TGACY	S000457
WBOXNTERF3	site	247	(-)	TGACY	S000457
WRKY71OS	site	115	(-)	TGAC	S000447
WRKY71OS	site	137	(-)	TGAC	S000447
WRKY71OS	site	212	(-)	TGAC	S000447
WRKY71OS	site	248	(-)	TGAC	S000447

Appendix Table 6 The presence of *cis*-acting elements in *CP-ACO I* 5'flanking region using PlantCARE Database

5UTR Py-rich stretch

Site Name	Organism	Position	Strand	Matrix score.	sequence
5UTR Py-rich stretch	*Lycopersicon esculentum*	343	-	9	TTTCTTCTCT

Function : *cis*-acting element conferring high transcription levels

ABRE

Site Name	Organism	Position	Strand	Matrix score.	sequence
ABRE	*Arabidopsis thaliana*	46	+	6	TACGTG

Function : *cis*-acting element involved in the abscisic acid responsiveness

ACE

Site Name	Organism	Position	Strand	Matrix score.	sequence
ACE	*Petroselinum crispum*	44	-	9	GACACGTATG

Function : *cis*-acting element involved in light responsiveness

ARE

Site Name	Organism	Position	Strand	Matrix score.	sequence
ARE	*Zea mays*	257	-	6	TGGTTT
ARE	*Zea mays*	650	+	6	TGGTTT

Function : *cis*-acting regulatory element essential for the anaerobic induction

ATGCAAAT motif

Site Name	Organism	Position	Strand	Matrix score.	sequence
ATGCAAAT motif	*Oryza sativa*	176	+	8	ATACAAAT

Function : *cis*-acting regulatory element associated to the TGAGTCA motif

Site Name	Organism	Position	Strand	Matrix score.	sequence
Box 4	*Petroselinum crispum*	144	+	6	ATTAAT
Box 4	*Petroselinum crispum*	245	+	6	ATTAAT
Box 4	*Petroselinum crispum*	217	+	6	ATTAAT
Box 4	*Petroselinum crispum*	286	+	6	ATTAAT

Function : part of a conserved DNA module involved in light responsiveness

Appendix Table 6 (Continued)

CAAT-box

Site Name	Organism	Position	Strand	Matrix score.	sequence
CAAT-box	*Hordeum vulgare*	35	+	4	CAAT
CAAT-box	*Brassica rapa*	104	+	5	CAAAT
CAAT-box	*Brassica rapa*	162	-	5	CAAAT
CAAT-box	*Brassica rapa*	179	+	5	CAAAT
CAAT-box	*Brassica rapa*	261	+	5	CAAAT
CAAT-box	*Petunia hybrida*	592	-	7	TGCCAAC
CAAT-box	*Arabidopsis thaliana*	636	+	5	CCAAT
CAAT-box	*Glycine max*	637	+	5	CAATT
CAAT-box	*Glycine max*	638	-	5	CAATT
CAAT-box	*Hordeum vulgare*	639	-	4	CAAT
CAAT-box	*Hordeum vulgare*	762	+	4	CAAT
CAAT-box	*Arabidopsis thaliana*	781	-	5	CCAAT
CAAT-box	*Glycine max*	828	-	5	CAATT
CAAT-box	*Hordeum vulgare*	829	-	4	CAAT
CAAT-box	*Brassica rapa*	838	+	5	CAAAT
CAAT-box	*Arabidopsis thaliana*	839	-	8	CCCAATTT
CAAT-box	*Glycine max*	840	-	5	CAATT
CAAT-box	*Arabidopsis thaliana*	841	-	5	CCAAT
CAAT-box	*Arabidopsis thaliana*	957	+	5	CCAAT
CAAT-box	*Hordeum vulgare*	958	+	4	CAAT
CAAT-box	*Hordeum vulgare*	999	-	4	CAAT

Function : common *cis*-acting element in promoter and enhancer regions

CAT-box

Site Name	Organism	Position	Strand	Matrix score.	sequence
CAT-box	*Arabidopsis thaliana*	598	-	6	GCCACT

Function : *cis*-acting regulatory element related to meristem expression

CGTCA-motif

Site Name	Organism	Position	Strand	Matrix score.	sequence
CGTCA-motif	*Hordeum vulgare*	391	-	5	CGTCA

Function : *cis*-acting regulatory element involved in the MeJA-responsiveness

G-Box

Site Name	Organism	Position	Strand	Matrix score.	sequence
G-Box	*Antirrhinum majus*	46	-	6	CACGTA

Function : *cis*-acting regulatory element involved in light responsiveness

Appendix Table 6 (Continued)

G-box

Site Name	Organism	Position	Strand	Matrix score.	sequence
G-box	*Daucus carota*	46	+	6	TACGTG

Function : *cis*-acting regulatory element involved in light responsiveness

GAG-motif

Site Name	Organism	Position	Strand	Matrix score.	sequence
GAG-motif	*Arabidopsis thaliana*	608	+	7	AGAGAGT

Function : part of a light responsive element

GATA-motif

Site Name	Organism	Position	Strand	Matrix score.	sequence
GATA-motif	*Pisum sativum*	90	-	7	GATAGGG

Function : part of a light responsive element

HSE

Site Name	Organism	Position	Strand	Matrix score.	sequence
HSE	*Brassica oleracea*	131	-	9	AAAAAATTTC
HSE	*Brassica oleracea*	691	-	9	AAAAAATTTC

Function : *cis*-acting element involved in heat stress responsiveness

I-box

Site Name	Organism	Position	Strand	Matrix score.	sequence
I-box	*Zea mays*	90	-	7	GATAGGG
I-box	*Flaveria trinervia*	395	+	7	GATATGG

Function : part of a light responsive element

MBS

Site Name	Organism	Position	Strand	Matrix score.	sequence
MBS	*Arabidopsis thaliana*	14	-	6	CAACTG

Function : MYB binding site involved in drought-inducibility

Appendix Table 6 (Continued)

O2-site

Site Name	Organism	Position	Strand	Matrix score.	sequence
O2-site	*Zea mays*	392	+	9	GATGATATGG
O2-site	*Zea mays*	749	+	9	GATGATATGG

Function : *cis*-acting regulatory element involved in zein metabolism regulation

Skn-1_motif

Site Name	Organism	Position	Strand	Matrix score.	sequence
Skn-1 motif	*Oryza sativa*	390	-	5	GTCAT
Skn-1 motif	*Oryza sativa*	660	+	5	GTCAT

Function : *cis*-acting regulatory element required for endosperm expression

TA-rich region

Site Name	Organism	Position	Strand	Matrix score.	sequence
TA-rich region	*Nicotiana tabacum*	52	+	20	TATATATATATATATATATA
TA-rich region	*Nicotiana tabacum*	56	+	20	TATATATATATATATATATA
TA-rich region	*Nicotiana tabacum*	54	+	21	TATATATATATATATATATA

Function : enhancer

TATA-box

Site Name	Organism	Position	Strand	Matrix score.	sequence
TATA-box	*Arabidopsis thaliana*	36	-	9	tcTATATAtt
TATA-box	*Brassica napus*	37	+	6	ATATAT
TATA-box	*Arabidopsis thaliana*	38	+	4	TATA
TATA-box	*Arabidopsis thaliana*	40	+	4	TATA
TATA-box	*Helianthus annuus*	52	-	6	TATACA
TATA-box	*Arabidopsis thaliana*	54	+	8	TATATATA
TATA-box	*Brassica napus*	55	+	6	ATATAT
TATA-box	*Arabidopsis thaliana*	56	+	8	TATATATA
TATA-box	*Brassica napus*	57	+	6	ATATAT
TATA-box	*Arabidopsis thaliana*	58	+	8	TATATATA
TATA-box	*Brassica napus*	59	+	6	ATATAT
TATA-box	*Arabidopsis thaliana*	60	+	8	TATATATA
TATA-box	*Brassica napus*	61	+	6	ATATAT
TATA-box	*Arabidopsis thaliana*	62	+	8	TATATATA
TATA-box	*Brassica napus*	63	+	6	ATATAT
TATA-box	*Arabidopsis thaliana*	64	+	4	TATA
TATA-box	*Brassica napus*	65	+	6	ATATAT
TATA-box	*Arabidopsis thaliana*	66	+	4	TATA
TATA-box	*Helianthus annuus*	70	-	6	TATACA
TATA-box	*Arabidopsis thaliana*	72	+	4	TATA
TATA-box	*Brassica napus*	107	+	6	ATTATA

Appendix Table 6 (Continued)

TATA-box

Site Name	Organism	Position	Strand	Matrix score.	sequence
TATA-box	*Arabidopsis thaliana*	108	-	5	TATAA
TATA-box	*Arabidopsis thaliana*	109	+	4	TATA
TATA-box	*Lycopersicon esculentum*	120	+	5	TTTTA
TATA-box	*Arabidopsis thaliana*	126	+	8	TATTTAAA
TATA-box	*Lycopersicon esculentum*	130	-	5	TTTTA
TATA-box	*Zea mays*	137	-	8	TTTAAAAA
TATA-box	*Lycopersicon esculentum*	138	+	5	TTTTA
TATA-box	*Glycine max*	146	+	5	TAATA
TATA-box	*Lycopersicon esculentum*	149	-	5	TTTTA
TATA-box	*Arabidopsis thaliana*	172	-	7	TATAAAA
TATA-box	*Arabidopsis thaliana*	173	-	6	TATAAA
TATA-box	*Arabidopsis thaliana*	174	-	5	TATAA
TATA-box	*Arabidopsis thaliana*	175	+	4	TATA
TATA-box	*Brassica napus*	182	+	6	ATATAT
TATA-box	*Arabidopsis thaliana*	183	+	4	TATA
TATA-box	*Zea mays*	188	-	8	TTTAAAAA
TATA-box	*Lycopersicon esculentum*	189	+	5	TTTTA
TATA-box	*Zea mays*	222	-	8	TTTAAAAA
TATA-box	*Lycopersicon esculentum*	223	+	5	TTTTA
TATA-box	*Arabidopsis thaliana*	224	-	11	TATAAATATAAA
TATA-box	*Antirrhinum majus*	228	-	8	TATAAATT
TATA-box	*Ac*	229	-	7	TATAAAT
TATA-box	*Arabidopsis thaliana*	230	-	6	TATAAA
TATA-box	*Arabidopsis thaliana*	231	-	5	TATAA
TATA-box	*Arabidopsis thaliana*	232	+	4	TATA
TATA-box	*Lycopersicon esculentum*	238	-	5	TTTTA
TATA-box	*Glycine max*	247	+	5	TAATA
TATA-box	*Arabidopsis thaliana*	248	+	11	TATAAATATAAA
TATA-box	*Brassica oleracea*	253	+	6	ATATAA
TATA-box	*Arabidopsis thaliana*	254	+	6	TATAAA
TATA-box	*Arabidopsis thaliana*	265	-	9	taTATAAAtc
TATA-box	*Ac*	266	-	7	TATAAAT
TATA-box	*Arabidopsis thaliana*	267	-	11	TATAAATATAAA
TATA-box	*Arabidopsis thaliana*	268	-	7	TATATAA
TATA-box	*Arabidopsis thaliana*	269	+	8	TATATATA
TATA-box	*Brassica napus*	270	+	6	ATATAT
TATA-box	*Arabidopsis thaliana*	271	+	8	TATATATA
TATA-box	*Brassica napus*	272	+	6	ATATAT
TATA-box	*Arabidopsis thaliana*	273	+	8	TATATATA
TATA-box	*Brassica napus*	274	+	6	ATATAT
TATA-box	*Arabidopsis thaliana*	275	+	8	TATATATA
TATA-box	*Brassica napus*	276	+	6	ATATAT
TATA-box	*Arabidopsis thaliana*	277	+	4	TATA
TATA-box	*Brassica napus*	278	+	6	ATATAT
TATA-box	*Arabidopsis thaliana*	279	+	4	TATA
TATA-box	*Glycine max*	298	-	5	TAATA
TATA-box	*Lycopersicon esculentum*	479	+	5	TTTTA
TATA-box	*Arabidopsis thaliana*	489	-	5	TATAA
TATA-box	*Arabidopsis thaliana*	490	+	4	TATA

Appendix Table 6 (Continued)

TATA-box

Site Name	Organism	Position	Strand	Matrix score.	sequence
TATA-box	*Arabidopsis thaliana*	500	-	5	TATAA
TATA-box	*Arabidopsis thaliana*	501	+	4	TATA
TATA-box	*Glycine max*	503	+	5	TAATA
TATA-box	*Arabidopsis thaliana*	504	-	9	tcTATATAtt
TATA-box	*Brassica napus*	505	+	6	ATATAT
TATA-box	*Arabidopsis thaliana*	506	+	8	TATATATA
TATA-box	*Brassica napus*	507	+	6	ATATAT
TATA-box	*Arabidopsis thaliana*	508	+	8	TATATATA
TATA-box	*Brassica napus*	509	+	6	ATATAT
TATA-box	*Arabidopsis thaliana*	510	+	8	TATATATA
TATA-box	*Brassica napus*	511	+	6	ATATAT
TATA-box	*Arabidopsis thaliana*	512	+	4	TATA
TATA-box	*Brassica napus*	513	+	6	ATATAT
TATA-box	*Arabidopsis thaliana*	514	+	4	TATA
TATA-box	*Helianthus annuus*	542	-	6	TATACA
TATA-box	*Arabidopsis thaliana*	544	+	4	TATA
TATA-box	*Helianthus annuus*	548	-	6	TATACA
TATA-box	*Arabidopsis thaliana*	550	+	4	TATA
TATA-box	*Helianthus annuus*	554	-	6	TATACA
TATA-box	*Arabidopsis thaliana*	556	+	4	TATA
TATA-box	*Oryza sativa*	683	-	7	TACAAAA
TATA-box	*Arabidopsis thaliana*	715	+	4	TATA
TATA-box	*Lycopersicon esculentum*	724	-	5	TTTTA
TATA-box	*Glycine max*	732	+	5	TAATA
TATA-box	*Brassica oleracea*	734	+	7	ATATAAT
TATA-box	*Arabidopsis thaliana*	735	+	4	TATA
TATA-box	*Glycine max*	737	+	5	TAATA
TATA-box	*Brassica oleracea*	739	+	7	ATATAAT
TATA-box	*Arabidopsis thaliana*	740	+	4	TATA
TATA-box	*Glycine max*	742	+	5	TAATA
TATA-box	*Arabidopsis thaliana*	745	+	4	TATA
TATA-box	*Helianthus annuus*	756	-	6	TATACA
TATA-box	*Arabidopsis thaliana*	758	+	4	TATA
TATA-box	*Glycine max*	797	-	5	TAATA
TATA-box	*Lycopersicon esculentum*	820	+	5	TTTTA
TATA-box	*Daucus carota*	848	+	9	ccTATAAATT
TATA-box	*Ac*	850	+	7	TATAAAT

Function : core promoter element around -30 of transcription start

Appendix Table 6 (Continued)

TC-rich repeats

Site Name	Organism	Position	Strand	Matrix score.	sequence
TC-rich repeats	*Nicotiana tabacum*	340	-	9	ATTTTCTTCA

Function : *cis*-acting element involved in defense and stress responsiveness

TCCC-motif

Site Name	Organism	Position	Strand	Matrix score.	sequence
TCCC-motif	*Spinacia oleracea*	1026	-	7	TCTCCCT

Function : part of a light responsive element

TGACG-motif

Site Name	Organism	Position	Strand	Matrix score.	sequence
TGACG-motif	*Hordeum vulgare*	391	+	5	TGACG

Function : *cis*-acting regulatory element involved in the MeJA-responsiveness

Unnamed__1

Site Name	Organism	Position	Strand	Matrix score.	sequence
Unnamed__1	*Zea mays*	457	-	5	CGTGG

Function : no know function

Unnamed__2

Site Name	Organism	Position	Strand	Matrix score.	sequence
Unnamed__2	*Glycine max*	217	+	14	ATTAAATTTTAAATT

Function : no know function

Unnamed__3

Site Name	Organism	Position	Strand	Matrix score.	sequence
Unnamed__3	*Zea mays*	457	-	5	CGTGG

Function : no know function

Unnamed__4

Site Name	Organism	Position	Strand	Matrix score.	sequence
Unnamed__4	*Petroselinum hortense*	4	-	4	CTCC
Unnamed__4	*Petroselinum hortense*	774	-	4	CTCC
Unnamed__4	*Petroselinum hortense*	327	-	4	CTCC

Appendix Table 6 (Continued)

Unnamed 4

Site Name	Organism	Position	Strand	Matrix score.	sequence
Unnamed 4	*Petroselinum hortense*	981	+	4	CTCC
Unnamed 4	*Petroselinum hortense*	324	-	4	CTCC
Unnamed 4	*Petroselinum hortense*	919	+	4	CTCC
Unnamed 4	*Petroselinum hortense*	455	+	4	CTCC
Unnamed 4	*Petroselinum hortense*	1028	-	4	CTCC

Function : no know function

Unnamed 6

Site Name	Organism	Position	Strand	Matrix score.	sequence
Unnamed 6	*Zea mays*	262	-	10	taTAAATATct

Function : no know function

Appendix Table 7 The presence of *cis*-acting elements in *CP-ACO II* 5'flanking region using PlantCARE Database

A-box

Site Name	Organism	Position	Strand	Matrix score.	sequence
A-box	*Petroselinum crispum*	85	-	6	CCGTCC

Function : *cis*-acting regulatory element

AAGAA-motif

Site Name	Organism	Position	Strand	Matrix score.	sequence
AAGAA-motif	*Avena sativa*	544	-	7	GAAAGAA

Function : no know function

ABRE

Site Name	Organism	Position	Strand	Matrix score.	sequence
ABRE	*Arabidopsis thaliana*	62	-	6	TACGTG
ABRE	*Hordeum vulgare*	383	-	9	CCGCGTAGGC
ABRE	*Arabidopsis thaliana*	214	-	6	TACGTG

Function : *cis*-acting element involved in the abscisic acid responsiveness

ARE

Site Name	Organism	Position	Strand	Matrix score.	sequence
ARE	*Zea mays*	119	-	6	TGGTTT
ARE	*Zea mays*	172	+	6	TGGTTT

Function : *cis*-acting regulatory element essential for the anaerobic induction

AuxRR-core

Site Name	Organism	Position	Strand	Matrix score.	sequence
AuxRR-core	*Nicotiana tabacum*	89	+	7	GGTCCAT

Function : *cis*-acting regulatory element involved in auxin responsiveness

Box 4

Site Name	Organism	Position	Strand	Matrix score.	sequence
Box 4	*Petroselinum crispum*	49	+	6	ATTAAT

Function : part of a conserved DNA module involved in light responsiveness

Appendix Table 7 (Continued)

Box I

Site Name	Organism	Position	Strand	Matrix score.	sequence
Box I	*Pisum sativum*	323	-	7	TTTCAAA

Function : light responsive element

Box-W1

Site Name	Organism	Position	Strand	Matrix score.	sequence
Box-W1	*Petroselinum crispum*	247	-	6	TTGACC

Function : fungal elicitor responsive element

CAAT-box

Site Name	Organism	Position	Strand	Matrix score.	sequence
CAAT-box	*Brassica rapa*	17	-	5	CAAAT
CAAT-box	*Brassica rapa*	253	+	5	CAAAT
CAAT-box	*Arabidopsis thaliana*	205	+	5	CCAAT
CAAT-box	*Hordeum vulgare*	435	-	4	CAAT
CAAT-box	*Hordeum vulgare*	126	-	4	CAAT
CAAT-box	*Glycine max*	434	-	5	CAATT
CAAT-box	*Hordeum vulgare*	206	+	4	CAAT
CAAT-box	*Brassica rapa*	444	+	5	CAAAT
CAAT-box	*Glycine max*	47	+	5	CAATT
CAAT-box	*Arabidopsis thaliana*	170	-	5	CCAAT

Function : common *cis*-acting element in promoter and enhancer regions

CATT-motif

Site Name	Organism	Position	Strand	Matrix score.	sequence
CATT-motif	*Zea mays*	460	+	6	GCATTC

Function : part of a light responsive element

CCGTCC-box

Site Name	Organism	Position	Strand	Matrix score.	sequence
CCGTCC-box	*Arabidopsis thaliana*	85	-	6	CCGTCC

Function : *cis*-acting regulatory element related to meristem specific activation

CGTCA-motif

Site Name	Organism	Position	Strand	Matrix score.	sequence
CGTCA-motif	*Hordeum vulgare*	114	+	5	CGTCA

Appendix Table 7 (Continued)

Function : *cis*-acting regulatory element involved in the MeJA-responsiveness

G-Box

Site Name	Organism	Position	Strand	Matrix score.	sequence
G-Box	*Antirrhinum majus*	62	+	6	CACGTA
G-Box	*Antirrhinum majus*	214	+	6	CACGTA

Function : *cis*-acting regulatory element involved in light responsiveness

G-box

Site Name	Organism	Position	Strand	Matrix score.	sequence
G-box	*Daucus carota*	62	-	6	TACGTG
G-box	*Daucus carota*	214	-	6	TACGTG

Function : *cis*-acting regulatory element involved in light responsiveness

HSE

Site Name	Organism	Position	Strand	Matrix score.	sequence
HSE	*Brassica oleracea*	514	+	9	AAAAAATTTC
HSE	*Brassica oleracea*	540	-	9	AAAAAATTTC

Function : cis-acting element involved in heat stress responsiveness

MBS

Site Name	Organism	Position	Strand	Matrix score.	sequence
MBS	*Arabidopsis thaliana*	70	-	6	CAACTG
MBS	*Arabidopsis thaliana*	286	-	6	TAACTG

Function : MYB binding site involved in drought-inducibility

TATA-box

Site Name	Organism	Position	Strand	Matrix score.	sequence
TATA-box	*Lycopersicon esculentum*	12	+	5	TTTTA
TATA-box	*Lycopersicon esculentum*	25	-	5	TTTTA
TATA-box	*Glycine max*	51	+	5	TAATA
TATA-box	*Lycopersicon esculentum*	196	-	5	TTTTA
TATA-box	*Glycine max*	265	-	5	TAATA

Appendix Table 7 (Continued)

TATA-box

Site Name	Organism	Position	Strand	Matrix score.	sequence
TATA-box	*Lycopersicon esculentum*	276	+	5	TTTTA
TATA-box	*Arabidopsis thaliana*	298	-	9	taTATAAAtc
TATA-box	*Brassica napus*	299	+	6	ATATAT
TATA-box	*Arabidopsis thaliana*	300	+	8	TATATATA
TATA-box	*Brassica napus*	301	+	6	ATATAT
TATA-box	*Arabidopsis thaliana*	302	+	8	TATATATA
TATA-box	*Brassica napus*	303	+	6	ATATAT
TATA-box	*Arabidopsis thaliana*	304	+	8	TATATATA
TATA-box	*Brassica napus*	305	+	6	ATATAT
TATA-box	*Avena sativa*	306	+	12	TATATTTATATTT
TATA-box	*Brassica napus*	307	+	6	ATATAT
TATA-box	*Arabidopsis thaliana*	308	+	8	TATATATA
TATA-box	*Brassica napus*	309	+	6	ATATAT
TATA-box	*Arabidopsis thaliana*	310	+	4	TATA
TATA-box	*Brassica napus*	311	+	6	ATATAT
TATA-box	*Arabidopsis thaliana*	312	+	4	TATA
TATA-box	*Arabidopsis thaliana*	348	-	9	TAAAAATAA
TATA-box	*Lycopersicon esculentum*	352	+	5	TTTTA
TATA-box	*Arabidopsis thaliana*	428	+	9	tcTATATAtt
TATA-box	Ac	430	+	7	TATAAAT

Function : core promoter element around -30 of transcription start

TCCACCT-motif

Site Name	Organism	Position	Strand	Matrix score.	sequence
TCCACCT-motif	*Petroselinum hortense*	163	+	7	TCCACCT

Function : no know function

TGA-element

Site Name	Organism	Position	Strand	Matrix score.	sequence
TGA-element	*Brassica oleracea*	199	+	6	AACGAC

Function : auxin-responsive element

Appendix Table 7 (Continued)

TGACG-motif

Site Name	Organism	Position	Strand	Matrix score.	sequence
TGACG-motif	*Hordeum vulgare*	114	-	5	TGACG

Function : *cis*-acting regulatory element involved in the MeJA-responsiveness

Unnamed 4

Site Name	Organism	Position	Strand	Matrix score.	sequence
Unnamed 4	*Petroselinum hortense*	236	+	4	CTCC
Unnamed 4	*Petroselinum hortense*	397	+	4	CTCC
Unnamed 4	*Petroselinum hortense*	362	+	4	CTCC

Function : no know function

W box

Site Name	Organism	Position	Strand	Matrix score.	sequence
W box	*Arabidopsis thaliana*	247	-	6	TTGACC

Function : no know function

circadian

Site Name	Organism	Position	Strand	Matrix score.	sequence
circadian	*Lycopersicon esculentum*	477	+	6	CAANNNNATC

Function : *cis*-acting regulatory element involved in circadian control

Appendix Table 8 The presence of *cis*-acting elements in *CP-ACO I* 5'flanking region using Plant Prom DB. Abbreviations, Regulatory Element/Consensus: RE, Accession No of RE in a given DB: AC, Organism/Species: OS, Binding Factor or One of them: BF, Mismatches: Mism, Mean Expected Number: Mean. Exp. Number and Upper Confidence Interval: Up.Conf.Int.

```
.....................................................................................
 RE:    36. AC: RSP00036//OS: Brassica napus /GENE: napA/RE: ABRE Bp /BF: ABI3
 Motifs on "+" Strand: Mean Exp. Number   0.00902     Up.Conf.Int.  1     Found   1
     883  CAAACACC      890 (Mism.= 0)
.....................................................................................
 RE:    69. AC: RSP00069//OS: pea (Pisum sativum) /GENE: PSPAL2/RE: Motif j/k /BF:
epicotyl-specific nuclear factor
 Motifs on "-" Strand: Mean Exp. Number   0.02259     Up.Conf.Int.  1     Found   1
     832  CAATTTAGTAA      822 (Mism.= 2)
.....................................................................................
 RE:   210. AC: RSP00210//OS: French bean (Phaseolus vulgaris) /GENE: DLEC2/RE:
DLEC2,B /BF: MAT2 (ROM2)
 Motifs on "+" Strand: Mean Exp. Number   0.00230     Up.Conf.Int.  1     Found   1
       7  GtCACCTCAGtT       18 (Mism.= 2)
.....................................................................................
 RE:   445. AC: RSP00445//OS: maize /GENE: cyPPDK1/RE: box e /BF: DOF1
 Motifs on "+" Strand: Mean Exp. Number   0.01059     Up.Conf.Int.  1     Found   1
    1004  AAAAAAGAGA     1013 (Mism.= 0)
.....................................................................................
 RE:   447. AC: RSP00447//OS: maize /GENE: pepcZm2A/RE: box b /BF: DOF1
 Motifs on "+" Strand: Mean Exp. Number   0.03619     Up.Conf.Int.  1     Found   1
    1011  AGaAAAGCAA     1020 (Mism.= 1)
.....................................................................................
 RE:   473. AC: RSP00473//OS: alfalfa (Medicago sativa) /GENE: MSPRP2/RE: Alfin1 BS3
/BF: Alfin1
 Motifs on "-" Strand: Mean Exp. Number   0.00520     Up.Conf.Int.  1     Found   1
     543  CATGTGTGTGTgt      531 (Mism.= 2)
.....................................................................................
 RE:   492. AC: RSP00492//OS: maize (Zea mays) /GENE: Zc2/RE: Zc2 A/T-2 /BF: nuclear
factor of apparent molecular mass 30 kDa
 Motifs on "+" Strand: Mean Exp. Number   0.00000     Up.Conf.Int.  1     Found   1
      37  ATATATAcATAcgTgTgTATATATATATATATATgTATAcAc       78 (Mism.= 8)

 Motifs on "-" Strand: Mean Exp. Number   0.00000     Up.Conf.Int.  1     Found   1
      78  gTgTATAcATATATATATATATATAcAcAcgTATgTATATAT       37 (Mism.= 8)
.....................................................................................
 RE:   495. AC: RSP00495//OS: wheat (Triticum aestivum) /GENE: LMW-glutenin/RE: P-box
2 /BF: unknown nuclear factor
 Motifs on "-" Strand: Mean Exp. Number   0.03727     Up.Conf.Int.  1     Found   1
      34  TGTAAAGT       27 (Mism.= 0)
.....................................................................................
 RE:   521. AC: RSP00521//OS: carrot (Daucus carota) /GENE: Dc3/RE: E1-core /BF: DPBF-
1; DPBF-2;
 Motifs on "-" Strand: Mean Exp. Number   0.01024     Up.Conf.Int.  1     Found   1
     440  ACACCCG      434 (Mism.= 0)
.....................................................................................
 RE:   601. AC: RSP00601//OS: tomato (Lycopersicon esculentum), Lycopersicon
esculentum /GENE: rbcS2/RE: C-rich R /BF: unknown nuclear factor
 Motifs on "+" Strand: Mean Exp. Number   0.02082     Up.Conf.Int.  1     Found   1
     631  CCaCACCAA      639 (Mism.= 1)
.....................................................................................
 RE:   653. AC: RSP00653//OS: spinach (Spinacia oleracea) /GENE: petH/RE: CT-LB /BF:
unknown nuclear factor
 Motifs on "-" Strand: Mean Exp. Number   0.01158     Up.Conf.Int.  1     Found   1
     336  TTaTCTCTCCT      326 (Mism.= 1)
```

Appendix Table 8 (Continued)

```
.....................................................................................
 RE:   686. AC: RSP00686//OS: barley (Hordeum vulgare) /GENE: C-hordein/RE: E-box /BF:
unknown nuclear factor
 Motifs on "-" Strand: Mean Exp. Number   0.03727     Up.Conf.Int.  1     Found   1
      34  TGTAAAGT       27 (Mism.= 0)
.....................................................................................
 RE:   743. AC: RSP00743//OS: tobacco (Nicotiana plumbaginifolia) /GENE: Cab-E/RE: box
    3 /BF: GT-1
 Motifs on "-" Strand: Mean Exp. Number   0.04736     Up.Conf.Int.  1     Found   1
     636  GtGTGGTAAAcT      625 (Mism.= 2)
.....................................................................................
 RE:   773. AC: RSP00773//OS: Brassica oleracea /GENE: SLR1/RE: Box III /BF: unknown
transcription factor
 Motifs on "-" Strand: Mean Exp. Number   0.01397     Up.Conf.Int.  1     Found   1
     294  TGAaTTAATG      285 (Mism.= 1)
.....................................................................................
 RE:   778. AC: RSP00778//OS: arabidopsis (Arabidopsis thaliana) /GENE: AtS1/RE: Box
III /BF: unknown transcription factor
 Motifs on "-" Strand: Mean Exp. Number   0.03719     Up.Conf.Int.  1     Found   1
     294  TGAATtAATG      285 (Mism.= 1)
.....................................................................................
 RE:   802. AC: RSP00802//OS: potato (Solanum tuberosum) /GENE: patatin 21/RE: Box A-1
/BF: BABF
 Motifs on "+" Strand: Mean Exp. Number   0.04433     Up.Conf.Int.  1     Found   1
     272  ATATATATATATgcatTAAT      291 (Mism.= 4)

 Motifs on "-" Strand: Mean Exp. Number   0.04433     Up.Conf.Int.  1     Found   2
     518  ATATATATATATatTATAAg      499 (Mism.= 3)
     516  ATATATATATATTATAagAg      497 (Mism.= 3)
.....................................................................................
 RE:   838. AC: RSP00838//OS: arabidopsis (Arabidopsis thaliana) /GENE: H4A748/RE:
CCGTCG motif /BF: Unknown nuclear factor
 Motifs on "-" Strand: Mean Exp. Number   0.03401     Up.Conf.Int.  1     Found   1
     436  CCGTCG      431 (Mism.= 0)
.....................................................................................
 RE:   861. AC: RSP00861//OS: arabidopsis (Arabidopsis thaliana) /GENE: STK/RE: GA-2
/BF: BPC1
 Motifs on "+" Strand: Mean Exp. Number   0.01210     Up.Conf.Int.  1     Found   1
     604  AGAAAGAGA      612 (Mism.= 0)
.....................................................................................
 RE:   865. AC: RSP00865//OS: arabidopsis (Arabidopsis thaliana) /GENE: STK/RE: GA-6
/BF: BPC1
 Motifs on "+" Strand: Mean Exp. Number   0.04068     Up.Conf.Int.  1     Found   1
     327  GGAGAGAtA      335 (Mism.= 1)
.....................................................................................
 RE:   868. AC: RSP00868//OS: tobacco (Nicotiana tabacum) /GENE: NiPMT1a/RE: TA-rich
motif /BF: Unknown nuclear factor
 Motifs on "-" Strand: Mean Exp. Number   0.00359     Up.Conf.Int.  1     Found   1
     746  TATTATATTATA      735 (Mism.= 0)
.....................................................................................
 RE:   869. AC: RSP00869//OS: wheat (Triticum aestivum) /GENE: LMWG-1D1/RE: EM1 /BF:
Unknown nuclear factor
 Motifs on "-" Strand: Mean Exp. Number   0.00538     Up.Conf.Int.  1     Found   1
      34  TGTAAAGTG       26 (Mism.= 0)
.....................................................................................
 RE:   895. AC: RSP00895//OS: Geminiviruses (Nicotiana species); Pepper Huasteco Virus
(PHV) /GENE: Intergenic Regions; PHV coat protein (CP)/RE: CLE-core /BF: Unknown
nuclear factor
 Motifs on "-" Strand: Mean Exp. Number   0.01452     Up.Conf.Int.  1     Found   1
      28  GTGGACCC       21 (Mism.= 1)
.....................................................................................
 RE:   914. AC: RSP00914//OS: maize (Zea mays) /GENE: b-32/RE: B2 /BF: Opaque-2
 Motifs on "+" Strand: Mean Exp. Number   0.01647     Up.Conf.Int.  1     Found   1
     749  GATGATATGT      758 (Mism.= 1)
.....................................................................................
 RE:   916. AC: RSP00916//OS: maize (Zea mays) /GENE: b-32/RE: B4 /BF: Opaque-2
 Motifs on "-" Strand: Mean Exp. Number   0.04221     Up.Conf.Int.  1     Found   1
     873  GATGAAGGGA      864 (Mism.= 2)
```

Appendix Table 8 (Continued)

```
.....................................................................................
 RE:   933. AC: RSP00933//OS: pine (Pinus sylvestris) /GENE: GS1a/RE: AT-2a /BF:
nuclear proteins from cotyledons;
 Motifs on "+" Strand: Mean Exp. Number   0.01628     Up.Conf.Int.  1     Found   1
     128  TTTAaAATTTTTT      140 (Mism.= 1)
.....................................................................................
 RE:  1022. AC: RSP01015//OS: arabidopsis (Arabidopsis thaliana) /GENE: GapA/RE: Gap
box 3 /BF: GAPF
 Motifs on "-" Strand: Mean Exp. Number   0.02848     Up.Conf.Int.  1     Found   1
     823  aAAATGAAAG      814 (Mism.= 1)
.....................................................................................
 RE:  1068. AC: RSP01061//OS: rice (Oryza sativa) /GENE: GluA-3/RE: GCN4 motif /BF:
Hahb-4
 Motifs on "+" Strand: Mean Exp. Number   0.04672     Up.Conf.Int.  1     Found   1
       3  GgGAGTCAC       11 (Mism.= 1)
.....................................................................................
 RE:  1112. AC: RSP01105//OS: spinach (Spinacia oleracea) /GENE: rps22/RE: C-rich
motif /BF: unknown nuclear factor
 Motifs on "-" Strand: Mean Exp. Number   0.04210     Up.Conf.Int.  1     Found   1
     373  CCCCCT      368 (Mism.= 0)
.....................................................................................
 RE:  1175. AC: RSP01168//OS: Vicia faba /GENE: LeB4/RE: GC element /BF: Unknown
nuclear factor
 Motifs on "+" Strand: Mean Exp. Number   0.04281     Up.Conf.Int.  1     Found   1
     860  AAGtTCCCT      868 (Mism.= 1)
.....................................................................................
 RE:  1209. AC: RSP01201//OS: Chlorella vulgaris /GENE: NR/RE: NIT2 BS II /BF: NIT2
 Motifs on "+" Strand: Mean Exp. Number   0.00806     Up.Conf.Int.  1     Found   1
     807  TGAAGCTC      814 (Mism.= 0)
.....................................................................................
 RE:  1217. AC: RSP01209//OS: Lepidium africanum /GENE: LaCRC/RE: EM1 (CArG box 1)
/BF: MADS box proteins
 Motifs on "-" Strand: Mean Exp. Number   0.04025     Up.Conf.Int.  1     Found   1
     932  CTTTTTCTGG      923 (Mism.= 1)
.....................................................................................
 RE:  1223. AC: RSP01215//OS: Brassica oleracea /GENE: BoCRC/RE: EM2 (CArG box 2) /BF:
MADS box proteins
 Motifs on "-" Strand: Mean Exp. Number   0.02403     Up.Conf.Int.  1     Found   1
     933  CCTTTTTCTGG      923 (Mism.= 2)
.....................................................................................
 RE:  1296. AC: RSP01288//OS: Arabidopsis (Arabidopsis thaliana) /GENE: cab1/RE: Z-
DNA-motif /BF: unknown nuclear factor
 Motifs on "+" Strand: Mean Exp. Number   0.01777     Up.Conf.Int.  1     Found   1
      45  ATACGTGT       52 (Mism.= 0)
.....................................................................................
 Totally      36 motifs of    33 different REs have been found
_____________________________________________________________________________________
```

Appendix Table 9 The presence of *cis*-acting elements in *CP-ACO II* 5'flanking region using Plant Prom DB. Abbreviations, Regulatory Element/Consensus: RE, Accession No of RE in a given DB: AC, Organism/Species: OS, Binding Factor or One of them: BF, Mismatches: Mism, Mean Expected Number: Mean. Exp. Number and Upper Confidence Interval: Up.Conf.Int.

```
.....................................................................................
 RE:    16. AC: RSP00016//OS: soybean (Glycine max) /GENE: beta-conglycinin/RE: RY
/BF: unknown nuclear factor
 Motifs on "+" Strand: Mean Exp. Number   0.02531     Up.Conf.Int.  1     Found   1
      58  CATGCAC       64 (Mism.= 0)
.....................................................................................
 RE:    54. AC: RSP00054//OS: rice (Oryza sativa) tungro bacillform virus (RTBV; plant
pararetrovirus) /GENE: RTBV promoter/RE: C2a-BS /BF: root-specific nuclear factor(s)
 Motifs on "+" Strand: Mean Exp. Number   0.04465     Up.Conf.Int.  1     Found   1
     438  GCTTATC      444 (Mism.= 0)
.....................................................................................
 RE:    97. AC: RSP00097//OS: maize (Zea mays) /GENE: GapC4/RE: 9-mer box /BF: tobacco
nuclear factors
 Motifs on "+" Strand: Mean Exp. Number   0.00342     Up.Conf.Int.  1     Found   1
      87  ACGGTCCA       94 (Mism.= 0)
.....................................................................................
 RE:   140. AC: RSP00140//OS: pea,Pisum sativum /GENE: rbcS-3.6/RE: AT-1 (2) /BF: AT-1
 Motifs on "+" Strand: Mean Exp. Number   0.02158     Up.Conf.Int.  1     Found   1
     346  ATTTATTTTTATG      358 (Mism.= 2)
.....................................................................................
 RE:   170. AC: RSP00170//OS: soybean (Glycine max) /GENE: lbc3/RE: Element 2 /BF:
nodule specific factor
 Motifs on "+" Strand: Mean Exp. Number   0.03898     Up.Conf.Int.  1     Found   3
     330  aaTAAtTTATTTATTT      345 (Mism.= 3)
     334  aTTtAtTTATTTATTT      349 (Mism.= 3)
     338  aTTtAtTTATTTATTT      353 (Mism.= 3)
.....................................................................................
 RE:   249. AC: RSP00249//OS: Catharanthus roseus /GENE: CYM/RE: MSA /BF: unknown
nuclear factor
 Motifs on "-" Strand: Mean Exp. Number   0.03831     Up.Conf.Int.  1     Found   1
     231  AGACCGTTa      223 (Mism.= 1)
.....................................................................
 RE:   284. AC: RSP00284//OS: parsley (Petroselinum crispum) /GENE: PAL-1/RE: Box A
/BF: unknown nuclear factor
 Motifs on "-" Strand: Mean Exp. Number   0.01229     Up.Conf.Int.  1     Found   1
      90  CCGTCC       85 (Mism.= 0)
.....................................................................................
 RE:   405. AC: RSP00405//OS: tobacco (Nicotiana tabacum) /GENE: RNP2/RE: adjb /BF:
unknown nuclear factor
 Motifs on "-" Strand: Mean Exp. Number   0.01655     Up.Conf.Int.  1     Found   1
     368  TTTGGAGT      361 (Mism.= 0)
.....................................................................................
 RE:   522. AC: RSP00522//OS: carrot (Daucus carota) /GENE: Dc3/RE: E2-core /BF: DPBF-
1; DPBF-2;
 Motifs on "+" Strand: Mean Exp. Number   0.02914     Up.Conf.Int.  1     Found   1
     377  CCACTTG      383 (Mism.= 0)
.....................................................................................
 RE:   646. AC: RSP00646//OS: carrot (Daucus carota) /GENE: DC 59/RE: C1-box /BF:
unknown nuclear factor
 Motifs on "-" Strand: Mean Exp. Number   0.01628     Up.Conf.Int.  1     Found   1
      21  cAAATTAAAAATGt        8 (Mism.= 2)
.....................................................................................
 RE:   654. AC: RSP00654//OS: spinach (Spinacia oleracea) /GENE: PetH/RE: CT-B /BF:
unknown nuclear factor
 Motifs on "+" Strand: Mean Exp. Number   0.02170     Up.Conf.Int.  1     Found   1
     368  aACCCACTTC      377 (Mism.= 1)
```

Appendix Table 9 (Continued)

```
.....................................................................................
 RE:   680. AC: RSP00680//OS: barley (Hordeum vulgare) /GENE: Amy32b/RE: GARE/Box 2
/BF: aleurone layers nuclear protein ectracts
 Motifs on "-" Strand: Mean Exp. Number   0.02032     Up.Conf.Int.  1     Found   1
     269  TAAtAGAGTC      260 (Mism.= 1)
.....................................................................................
 RE:   683. AC: RSP00683//OS: arabidopsis (Arabidopsis thaliana) /GENE: Adh/RE: -190
half G-box (core) /BF: GBF3
 Motifs on "-" Strand: Mean Exp. Number   0.01077     Up.Conf.Int.  1     Found   1
     385  GgCAAGTGGA      376 (Mism.= 1)
.....................................................................................
 RE:   738. AC: RSP00738//OS: arabidopsis (Arabidopsis thaliana) /GENE: CAB2/RE: CUF-1
BS /BF: CUF-1
 Motifs on "+" Strand: Mean Exp. Number   0.03645     Up.Conf.Int.  1     Found   1
     212  GtCACGTAAt      221 (Mism.= 2)
.....................................................................................
 RE:   802. AC: RSP00802//OS: potato (Solanum tuberosum) /GENE: patatin 21/RE: Box A-1
/BF: BABF
 Motifs on "+" Strand: Mean Exp. Number   0.00745     Up.Conf.Int.  1     Found   1
     305  ATATATATATATTtTtTttT      324 (Mism.= 4)

 Motifs on "-" Strand: Mean Exp. Number   0.00745     Up.Conf.Int.  1     Found   1
     310  ATATATATATATcAaAgAtT      291 (Mism.= 4)
.....................................................
.......................
 RE:   849. AC: RSP00849//OS: maize (Zea mays) /GENE: H3C4/RE: CCGTCC motif /BF:
Unknown nuclear factor
 Motifs on "-" Strand: Mean Exp. Number   0.01229     Up.Conf.Int.  1     Found   1
      90  CCGTCC       85 (Mism.= 0)
.....................................................................................
 RE:   856. AC: RSP00856//OS: maize (Zea mays) /GENE: H4C7/RE: CCGTCC motif /BF:
Unknown nuclear factor
 Motifs on "-" Strand: Mean Exp. Number   0.01229     Up.Conf.Int.  1     Found   1
      90  CCGTCC       85 (Mism.= 0)
.....................................................................................
 RE:   870. AC: RSP00870//OS: wheat (Triticum aestivum) /GENE: LMWG-1D1/RE: EM2 /BF:
Unknown nuclear factor
 Motifs on "+" Strand: Mean Exp. Number   0.01517     Up.Conf.Int.  1     Found   1
      23  TtTAAAAGTG       32 (Mism.= 1)
.....................................................................................
 RE:   878. AC: RSP00878//OS: arabidopsis (Arabidopsis thaliana) /GENE: PCNA/RE: Site
I /BF: RPCF-1
 Motifs on "+" Strand: Mean Exp. Number   0.03550     Up.Conf.Int.  1     Found   1
      69  CCAGtTGG       76 (Mism.= 1)
.....................................................................................
 RE:   891. AC: RSP00891//OS: Solanum melongena /GENE: SmCP/RE: EE 1 /BF: Dof factors
 Motifs on "-" Strand: Mean Exp. Number   0.02347     Up.Conf.Int.  1     Found   1
     319  AAAATATAT      311 (Mism.= 0)
.....................................................................................
 RE:  1053. AC: RSP01046//OS: tobacco (Nicotiana tabacum) /GENE: histone genes
(oligo)/RE: NON /BF: unknown nuclear factor
 Motifs on "-" Strand: Mean Exp. Number   0.00703     Up.Conf.Int.  1     Found   1
      80  CtATCCAAC       72 (Mism.= 1)
.....................................................................................
 RE:  1082. AC: RSP01075//OS: Craterostigma plantagineum (Scrophulariaceae) /GENE:
CpC2/RE: ABRE II /BF: CpbZIP1; CpbZIP2 (short);
 Motifs on "-" Strand: Mean Exp. Number   0.02001     Up.Conf.Int.  1     Found   1
     146  CCCGaGTGaC      137 (Mism.= 2)
.....................................................................................
 RE:  1099. AC: RSP01092//OS: pea (Pisum sativum) /GENE: GS2/RE: box 1* /BF: unknown
nuclear factor
 Motifs on "-" Strand: Mean Exp. Number   0.02085     Up.Conf.Int.  1     Found   1
     267  ATAGAgTCTAA      257 (Mism.= 1)
.....................................................................................
 RE:  1147. AC: RSP01139//OS: maize (Zea mays) /GENE: Adh1/RE: Region B1 /BF: unknown
nuclear factor
 Motifs on "-" Strand: Mean Exp. Number   0.04190     Up.Conf.Int.  1     Found   1
     127  AtGTGGTTT      119 (Mism.= 1)
```

Appendix Table 9 (Continued)

```
.....................................................................................
RE:  1169. AC: RSP01162//OS: soybean (Glycine max) /GENE: Gmhspl7.5-E Heat Shock
Promoter/RE: TATA/dyad motif /BF: unknown nuclear factor
 Motifs on "-" Strand: Mean Exp. Number   0.03331     Up.Conf.Int.  1     Found   1
     552  cAgAAAGAATTTC      540 (Mism.= 2)
.....................................................................................
  RE:  1217. AC: RSP01209//OS: Lepidium africanum /GENE: LaCRC/RE: EM1 (CArG box 1)
/BF: MADS box proteins
 Motifs on "-" Strand: Mean Exp. Number   0.01164     Up.Conf.Int.  1     Found   1
     519  CTTTTTCTGG      510 (Mism.= 1)
.....................................................................................
 RE:  1359. AC: RSP01351//OS: Arabidopsis (Arabidopsis thaliana) /GENE: APX1/RE: DNase
I footprint /BF: CPRF1
 Motifs on "+" Strand: Mean Exp. Number   0.01803     Up.Conf.Int.  1     Found   1
     179  CTGGGgCCcCC      189 (Mism.= 2)
.....................................................................................
 Totally      30 motifs of    27 different REs have been found
_____________________________________________________________________________________
```

Appendix Table 10 Absorbance of BSA (0-10 μg/μl) at wave length at 595 nanometer

Concentration of BSA (μg/μl)	0	1.25	2.5	5	7.5	10
Absorbance of Standard	0	0.103	0.165	0.271	0.357	0.505

Appendix Table 11 Intensity of 4-MU (0-10^6 pmol) with excitation at 365 nm and emission at 455 nm

Concentration of 4-MU (pmol)	0	10^4	10^5	10^6
Intensity of Standard #1	-0.0400	3.7006	37.3679	349.8381
Intensity of Standard #2	0.1473	3.6222	35.9550	358.5081
Intensity of Standard #3	0.2390	3.8683	37.1748	356.3491
Intensity of Standard (Average)	0.12	3.7304	36.8326	354.8984

Appendix B

E. coli competent cell preparation

Protocol of preparation of *E. coli* competent cell

Competent cells made by this procedure preserved at -80 ºC, although there may be some deterioration in the efficiency of transformation during prolonged storage.

1. Pick a single colony (2-3 mm in diameter) from a plate freshly grown for 16-20 h at 37 ºC and transfer it into 100 ml of SOB medium in a 500 ml flask. Incubate the culture for 3 h at 37 ºC with vigorous shaking (250 cycles/min in a rotary shaker) until absorbance of cuture (0.4) at O.D. $_{600}$.

2. Aseptically transfer the cells to sterile , ice cold 250 ml tubes. Cool the cultures to 0 ºC by storing the tubes in ice for 15min.

3. Recover the cells by centrification at 3,500 rpm for15 min at 4 ºC.

4. Decant the media from the cell pellets. Stand the tubes in an inverted position for 1min to allow the last traces of media to drain away.

5. Resuspend each pellet in 33 ml of ice cold RF1 solution [100mM KCl, 50 mM $MnCl_2.4H_2O$, 30mM K-acetate, 10mM $CaCl_2$ and 15% glycerol: neutralized with acetic acid until to pH 5.8] and store on ice for 15 min.

6. Recover the cells by centrification at 3,500 rpm for15 min at 4 ºC.

7. Decant the fluid from the cell pellets and stand the tubes in an inverted position for 1min to allow the last traces of media to drain away.

8. Resuspend each pellet in 4 ml of ice cold RF2 solution [10 mM MOPs, 10 mM KCl, 75 mM $CaCl_2$ and 15% glycerol: neutralized with NaOH until to pH 6.8].

9. Using sterile pipette tip, transfer 100 µl of each suspension of competent cells to a sterile microcentrifuge tube. Store the tubes at -80 ºC.