การโคลนนิ่งและการแสดงออกของยีนเบต้ากาแลคโตซิเดสจากข้าว

นางสาวมัลลิกา จันทรังษี

วิทยานิพนธ์นี้เป็นส่วนหนึ่งของการศึกษาตามหลักสูตรปริญญาวิทยาศาสตรมหาบัณฑิต
สาขาวิชาเทคโนโลยีชีวภาพ
มหาวิทยาลัยเทคโนโลยีสุรนารี
ปีการศึกษา 2544

ISBN 974-533-009-4

Cloning and Expression of β-galactosidase from rice (*Oryza sativa* L.)

Miss Mallika Chantarangsee

A Thesis Submitted in Partial Fulfillment of the Requirements

For the Degree in Master of Science in Biotechnology

Suranaree University of Technology

Academic Year 2001

ISBN 974-533-009-4

THESIS TITLE

Cloning and Expression of β-galactosidase from rice (*Oryza sativa* L.)

Accepted by University Council, Suranaree University of Technology in Partial Fulfillment of the Requirements for the Master's Degree.

THESIS COMMITTEE

.....Chairman
(Asst. Prof. Sunthorn Kanchanatawee, Ph.D.)

.....Advisor
(Asst. Prof. James Ketudat-Cairns, Ph.D.)

.....Committee
(Asst. Prof. Mariena Ketudat-Cairns, Ph.D.)

.....Committee
(Asst. Prof. Neung Teaumroong, Dr. rer. nat.)

.....
(Assoc. Prof. Kasem Prabriputaloong, Ph.D.)
Acting, Vice Rector for Academic Affairs

.....
(Assoc. Prof. Kanok Phalaraksh, Ph.D.)
Dean, Institute of Agricultural Technology

มัลลิกา จันทรังษี : การโคลนนิ่งและการแสดงออกของยีนเบต้ากาแลคโตซิเดสจากข้าว
(CLONING AND EXPRESSION OF β-GALACTOSIDASE FROM RICE (*Oryza sativa* L.))
อาจารย์ที่ปรึกษา : ผศ. ดร. เจมส์ เกตุทัต-คาร์นส์, 61 หน้า. ISBN 974-533-009-4

เบต้ากาแลคโตซิเดสเป็นเอนไซม์ที่พบได้ในพืชหลายชนิด เอนไซม์นี้มีส่วนเกี่ยวข้องในกระบวนการสลายตัวของผนังเซลล์ของพืช ตัวอย่างเช่นในกระบวนการสุกของผลไม้และการแก่ จากการศึกษาเกี่ยวกับการทำงานของเอนไซม์ในเมล็ดข้าวช่วงที่กำลังงอกโดยใช้สาร X-Gal ทำหน้าที่เป็นสับสเตรทพบว่ายีนดังกล่าวมีการแสดงออกในเมล็ดข้าวช่วงที่กำลังงอกตรงบริเวณ aleurone layer และบริเวณเนื้อเยื่อที่อยู่รอบๆ เพื่อศึกษาลักษณะรหัสพันธุกรรม (cDNA) ของเอนไซม์ดังกล่าวในข้าวจึงได้ทำการทดลองโดยใช้ PCR product ที่มีความจำเพาะซึ่งได้ออกแบบมาจากลำดับเปปไทด์ของโปรตีนเบต้ากาแลคโตซิเดสมาใช้เป็นตัวติดตามในการตรวจหายีนดังกล่าวจากห้องสมุดดีเอ็นเอ จากนั้นชิ้นส่วนของยีนดังกล่าวได้ถูกนำมาออกแบบ primer เพื่อใช้ในการหาชิ้นส่วนสารพันธุกรรมทั้งหมดของเอนไซม์ ผลจากการวิเคราะห์ด้วยวิธี Southern blot hybridization แสดงให้เห็นว่าทั้งข้าวสายพันธุ์ indica และ japonica มียีนที่มีความสัมพันธ์กันประมาณ 3 ยีน ส่วนการวิเคราะห์การแสดงออกของยีนโดยวิธี Northern blot hybridization นั้นพบว่ายีนนี้ตรวจพบในเมล็ดข้าวที่กำลังงอกในระดับที่ต่ำ นอกจากนี้ยังได้ศึกษาชิ้นส่วนของยีนที่มีความสัมพันธ์กันนี้โดยได้รับชิ้นส่วนของรหัสพันธุกรรมนี้มาจากโครงการวิจัยข้าวที่ประเทศญี่ปุ่น (โคลน EST) จากการวิเคราะห์ชิ้นส่วนของยีนตามด้วยการตัดด้วยเอนไซม์ตัดจำเพาะที่เหมาะสมแล้วนำไปเชื่อมต่อในพลาสมิดและถูกนำไปหาลำดับนิวคลีโอไทด์ทำให้ทราบว่าชิ้นส่วนสารพันธุกรรมทั้งหมดของเอนไซม์ประกอบด้วยกรดอะมิโน 717 ตัว โดยมีเปปไทด์สัญญาณที่ปลายด้านอะมิโนของโปรตีน 20 ตัว ส่วนรหัสพันธุกรรมของยีนนี้ที่ได้จากโคลนห้องสมุดดีเอ็นเอพบว่าประกอบด้วยกรดอะมิโนจำนวน 841 ตัว โดยมีสัญญาณเปปไทด์ 25 ตัว จากการวิเคราะห์ด้วยโปรแกรม PSORT พบว่าโปรตีนจากโคลนห้องสมุดดีเอ็นเอและโคลน EST จะเคลื่อนที่ไปยังเอนโดพลาสมิคเรติคูลัมและที่ด้านนอกของเซลล์ตามลำดับ

สาขาวิชาเทคโนโลยีชีวภาพ
ปีการศึกษา 2544

ลายมือชื่อนักศึกษา……..
ลายมือชื่ออาจารย์ที่ปรึกษา……..
ลายมือชื่ออาจารย์ที่ปรึกษาร่วม……
ลายมือชื่ออาจารย์ที่ปรึกษาร่วม……....

MALLIKA CHANTARANGSEE : CLONING AND EXPRESSION OF β-GALACTOSIDASE FROM RICE (*Oryza sativa* L.)
THESIS ADVISOR : ASSISTANT PROFESSOR JAMES R. KETUDAT-CAIRNS, Ph.D., 61 PP.
ISBN 974-533-009-4

β-galactosidase has been found in many species of plants, and its activity may be involved in the degradation and remodeling of cell walls, such as during growth, fruit ripening, and senescence. Activity staining for β-galactosidase in germinating rice by X-Gal assay indicated that β-galactosidase is expressed in germinating rice and it showed strong localization to aleurone layer and surrounding tissues. In order to characterize the primary structure from its cDNA encoding β-galactosidase from rice and its expression during rice germination, a specific PCR product was amplified with degenerate primers designed from conserved β-galactosidase peptide sequences. The fragment was used as a probe to screen a rice immature seed cDNA library, and for Southern and Northern hybridization. Southern blot analysis revealed approximately 3 highly related β-galactosidase genes in both indica and japonica rice. Northern blot analysis revealed that the β-galactosidase mRNA was detectable at low level in rice shoot on the fifth days of germination. The full length cDNA sequence of the cDNA library cloning encoded 841 amino acid, including a single signal peptide of about 25 residues . Another related clone (E1405) from panicle at flowering stage was identified from its EST, requested from the MAFF clone bank, was found to be a full-length cDNA which encoded 717 amino acid residues, including a single signal peptide of 20 residues. The prediction, by PSORT indicated that the signal sequences of the library clone and EST clone may target these proteins to the endoplasmic reticulum and outside the cell respectively.

สาขาวิชาเทคโนโลยีชีวภาพ ลายมือชื่อนักศึกษา……..…..
ปีการศึกษา 2544 ลายมือชื่ออาจารย์ที่ปรึกษา……..…..
ลายมือชื่ออาจารย์ที่ปรึกษาร่วม……
ลายมือชื่ออาจารย์ที่ปรึกษาร่วม……....

ACKNOWLEDGEMENT

I would like to express my deepest and sincere gratitude to my advisor, Asst. Prof. Dr. James Ketudat-Cairns for his kindly providing me a good opportunity to study in this field, supervision, the laboratory facilities, encouragement, valuable suggestions and supporting scholarship throughout my study.

Deep appreciation and admiration is extended to Asst. Prof. Dr. Mariena Ketudat-Cairns and Asst. Prof. Dr. Neung Teaumroong for their valuable advises, kindness and encouraging guidance.

I also wish to thank Prof. Dr. Tatsuhito Fujimura and Dr. Kouichi Mizuno for their guidance, use for their laboratory facilities and suggestions. I would also like to thank all of my friends in Prof. Dr. Fujimura's laboratory for their help and warmth toward me throughout. They were always nice and friendly. Special thanks are extended to At chan for his help with practice in cDNA library screening, Dai chan and Oki san for their kindness and helps in the automated DNA sequencing preparation and analyzing.

Finally, I would like to express my deepest appreciation and sincere gratitude to my dear parents for their love, devoting, understanding and encouragement of my entire study.

Mallika Chantarangsee

TABLE OF CONTENTS

Page

TABLE OF CONTENTS (CONTINUED)

Page

TABLE OF CONTENTS (CONTINUED)

Page

LIST OF TABLE

Table **Page**

LIST OF FIGURE

Figure **Page**

LIST OF FIGURE (CONTINUED)

Figure **Page**

LIST OF ABBREVIATIONS

Abs	Absorbance
Amp	Ampicillin
ATP	Adenosine triphosphate
bp	Base pairs
°C	Degree celsius
cDNA	Complementary deoxynucleic acid
con't	Continue
DEPC	Diethyl pyrocarbonate
DNA	Deoxyribonucleic acid
dNTPs	dATP, dCTP, dGTP and dTTP
EDTA	Ethylene diamine tetraacetic ccid
g	Gravitational acceleration
(m, n) g	(Milli, Nano) gram
h	Hour
min	Minute
(m, μ, η) M	(Milli, Micro, Nano) Molar
(m, μ) l	(Milli, Micro) Litter
mRNA	Messenger ribonucleic acid
OD	Optical density
PAGE	Polyacrylamide gel electrophoresis
PCR	Polymerase chain reaction
pmol	Picomole
RNase	Ribonuclease
rpm	Revolutions per minute
sec	Second
SDS	Sodium dodecyl sulfate
SSC	Saline sodium citrate
SSPE	Saline sodium hydrogen phosphate EDTA

TEMED	N,N,N’,N’-tetramethylenediamine
Tris-HCl	Tris-(hydroxymethyl)-aminoethane
UV	ultraviolet

CHAPTER 1
INTRODUCTION

1. Overview of β-galactosidase

β-galactosidase (EC 3.2.1.23) is an enzyme that catalyzes the hydrolysis of β-glycosidic bonds between galactose and hydroxyl groups, such as cleavage of lactose to glucose and galactose. This enzyme occurs widely among microorganisms, animals and plants. It is commercially important enzyme which has potential applications in biotechnology, medicine, agriculture and the food industry. β-galactosidase has been one of the most carefully studied glycosidases because of its ready inducibility in microorganisms. Many microorganisms which have β-galactosidase activity have been shown to possess the capability to produce oligosaccharide sugars, which are widely used as functional sweeteners in the food industries. (Pluemsab *et al*., 1998). The enzyme catalyzes the hydrolysis of β-D-galactosides and α-L-arabinosides. Among its naturally occurring substrates, its action on lactose has been most studied.

In mammals, acid β-galactosidase is a lysosomal enzyme which hydrolyzes the terminal β-D-Galactoside residue from glycolipids and glycoproteins during the normal catabolic process of these tissue constituents. (Nanba and Suzuki, 1990). Its deficiency has been shown to cause a number of diseases. A complete or partial deficiency of β-galactosidase has been observed in diseases of various molecular pathology: G_{M1}- gangliosidosis and Morquio B disease, caused by gene defects in the enzyme; mucolipidosis II and III, caused by defect in transport of precursor enzyme and galactosialidosis, caused by abnormal degradation of the enzyme molecule due to a defect in a protective protein. (Oshima *et al*., 1988).

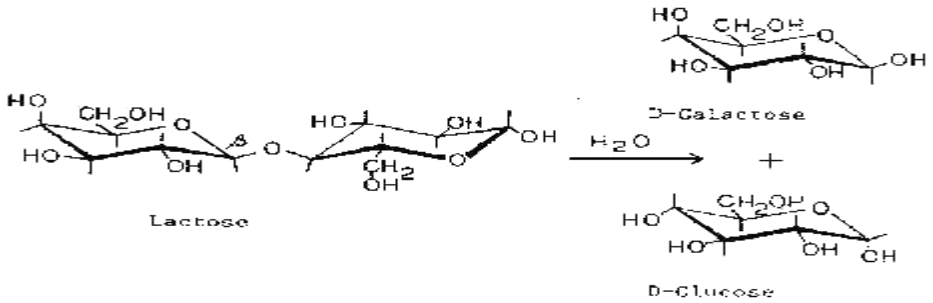


Figure 1. Hydrolysis of β-glycosidic bonds between galactose and hydroxyl group (Whitaker, 1994).

Several methods are available for following the activity of β-galactosidases. One of the best methods is to use a chromogenic substrate. A chromogenic substrate is one that gives a colored product on hydrolysis. As shown in Table.1, either *o*- or *p*- nitrophenyl galactoside may be used since the products, *o*- or *p*- nitrophenol, have an absorbance maximum near 400 nm. Another common substrate is X-Gal (5-bromo-4-chloro-3-indolyl-β-D-galactoside) which is also used as a chromogenic substrate such as identification of bacterial colonies that contain recombinant plasmids. The plasmids contain the *lacZ* gene, encoding *E. coli* β-galactosidase.

Table. 1 Substrate Specificity of β-galactosidase of *E. coli*, ML 309[a] (Whitaker, 1994)

Substrate	K_m (mol/L)	V_{max} (μmol/min/mg enzyme)	$V_{max}/_{Km}$ (X 10^{-4})
***o*-Nitrophenyl-β-D-galactoside**	**1.61×10^{-4}**	**178**	**110**
***p*-Nitrophenyl-β-D-galactoside**	**5.31×10^{-5}**	**22.4**	**44**
Methylsalicylate-β-D-galactoside	**2.50×10^{-3}**	**4.2**	**0.17**
Phenyl-β-D-galactoside	**1.47×10^{-3}**	**10.4**	**0.71**
α-Lactose	**1.90×10^{-3}**	**6.55**	**0.35**
allo-Lactose	**8.32×10^{-3}**	**12.8**	**0.15**
Thio- (*o*-Nitrophenyl)-β-D-galactoside	**1.20×10^{-4}**	**0.00**	**0**

[a]The reaction were performed at pH 7.6, 20° C, 0.05 M Tris-HCl buffer, and 0.05 M NaCl

2. Application of β-galactosidase

This enzyme is of interest to the food scientist for several reasons. Hydrolysis of lactose leads to glucose and galactose, which are much sweeter than lactose. Lactose also has a limited solubility and tends to precipitate out in frozen dairy products, which gives a grainy texture to the product (Whitaker, 1994).

Many of the vectors for current use carry a short segment of *E.coli* DNA that contains the regulatory sequences and the coding information for the first 146 amino acid of β-galactosidase gene (*lacZ*). Embedded in this coding region is a polycloning site that does not disrupt the reading frame but results in the harmless interpolation of a small number of amino acids into the amino-terminal fragment of β-galactosidase. However, cloning of a DNA fragment into one of the cloning sites will disrupt the protein coding sequence, resulting in a lack of β-galactosidase activity. Thus, cell containing recombinant clones appear white when grown on media containing X-Gal, while those with nonrecombinant plasmid are blue due to the hydrolysis of X-Gal by the β-galactosidase (Sambrook *et al*., 1989). Vectors of this type are used in host cells that code for the carboxy-terminal portion of β-galactosidase. Although neither the host encoded nor the plasmid-encoded fragments are themselves active, they can associate to form an enzymatically active protein. This type of complementation, in which deletion mutants of the operator-proximal segment of *lacZ* gene are complemented by β-galactosidase-negative mutants that have the operator-proximal region intact, is called α–complementation

3. The role of β-Galactosidase in plants

In plants, β-galactosidase has been found in the seeds of many species including cucumber, pumpkin, Thai jute, bitter melon, soybean, mungbean, cowpea, cotton and blackwood (Surarit *et al*., 1995), as well as in ripening fruit and senescencing tissues. Its activity may also be related to the degradation of fruit cell walls such as in ripening tomatoes and apples. β-Galactosidase, sometimes referred to as exo-galactanase, has been shown to possess *in vitro* activity against extracted cell wall material resulting in the release of galactose from wall polymers containing β (1→4)-D-galactan. (Smith *et al*., 1995). Numerous studies have shown that β-galactosidase catalyzes the hydrolysis of terminal galactosyl residues from carbohydrates, glycoproteins and galactolipids. β-Galactosidase action has been proposed to release stored galactose as a source of

energy for rapid growth, release free Gal during normal metabolic recycling of galactolipids, glycoproteins, and cell wall components, and degrade cell wall components during senescence (Smith and Gross, 2000). The study of the role of β-galactosidase in tomato fruit has resulted from physiological and biochemical data showing that Gal is the most dynamic sugar residue of the cell wall during tomato fruit development (Kim *et. al*, 1991). In particular, these physiological studies showed that there was a significant net loss of galactosyl residues from the wall throughout fruit development and the rate of galactosyl residue loss increasing during ripening. The presence of at least 7 isozymes of β-galactosidase in tomato fruit, suggests they may indeed play several roles (Smith and Gross, 2000).

3.1 Cell wall structure

An early model of the plant primary cell wall was that cellulose fibrils which were coated with hemicellulose, are found in a matrix composed of pectin and protein. Cellulose consists of linear chains of β(1-4)-linked glucose residues which aggregrate together by hydrogen bonds to form fibrils. Hemicellulose are composed of a variety of polymers, the major ones in dicots being xyloglycans, glucomannans, and galactoglucomannans. Pectin are also composed of a variety polymers, the neutral pectins being arabinans, galactans or arabinogalactans. It is thought that these polymers are linked together in the three dimentional cell wall by a variety of covalent and noncovalent bonds. Cellulose fibrils are held together by hydrogen bonds and similar bonds account for the interaction of cellulose with hemicellulose. However it has been postulated that in addition to simply coating the cellulose fibril, hemicellulose molecules may cross-link with the cellulose fibrils. They are certainly of sufficient length for this, and if such linkages do occur then cleavage of these bridges could be a major cause of cell wall loosening and hence softening (Seymour *et al*., 1993).

3.2 Cell wall hydrolase

A wide range of cell wall hydrolases can be identified in fruit tissue. Many of these are either constitutive throughout development and ripening. The enzyme most studied are pectinase (pectinesterase) (EC 3.1.1.15), polygalacturonase (EC 3.2.1.15), β-(1-4)glucanase or cellulase (EC 3.2.1.4) and β-galactosidase (EC3.2.1.23). Pectinesterase (PE) acts to remove the methyl

group from the C-6 position of a galacturonic acid. Polygalacturonase (PG) hydrolyses the α(1-4) link between adjacent glucose residues. The natural substrate for this enzyme is unknown, but this is likely to be a hemicellulose polymer rather than cellulose itself. β-galactosidase also often measured using artificial substrates. In many cases these enzymes are incapable of degrading native galactans and their natural substrates remain unknown. However in several cases β-galactosidases from ripening fruit have been shown to attack native galactan polymer. These enzymes can be considered as true β-galactosidases and could well be implicated in wall degradation.

4. **Rice plant**

Cultivated rice (*Oryza sativa* L.) is one of the world's most important crops and a primary food for more than a third of the world's population (Gurdev and Gary, 1991). Production and consumption are concentrated in Asia, where more than 90% of all rice is produced and consumed. Knowledge of rice molecular genetics and physiology are essential for application of molecular crop improvement techniques. Rice has numerous features that make it a model species for the grasses (Yuan *et al*., 2001). Its small genome size is compatible with current genomic technologies and sequencing efforts are underway to determine the complete sequence of the rice genome. Methods have been developed to produce transgenic rice to study the effects of protein over expression. Tools and resources are being developed to maximally interpret the rice genome sequences, and identify genes responsible for traits of interest for molecular breeding purpose.

4.1 Origin of cultivated rice.

There are two cultivated species of rice. *Oryza sativa*, the Asian rice, is grown worldwide. (Chang, 1976). *Oryza glaberrima,* the African rice, is grown on a limited scale in West Africa. Like wheat, corn, rye, oats and barley, rice belongs to Gramineae or grass family. Chang (1976) theorized that the genus *oryza* to which cultivated rice belongs, probably originated at least 130 million years ago and spread as a wild grass in Gondawanaland the super continent that eventually broke up and drifted apart to become Asia, Africa, the Americas, Australia and Antarctica. However recent analysis of the genetics of grasses have indicated this genus originated no more than 70 million years ago (Kellogg, 2001). Varieties of *Oryza sativa* are classified into six groups on the basis of genetic affinity. Widely known indica rices correspond

to group I and japonica to group VI. The so called javanica rices also belong to group VI and are designated as tropical japonicas in contrast to temperate japonicas grown in temperate climates (Khush, 1997). Indica rices are primarily grown in the tropics. The grain is medium to long with higher amylose content resulting in dry and fluffy cooked rice that shows little disintregration. Mostly indica varieties are grown in the Indian subcontinent, Southeast Asia, Southern China, and South America. A large proportion of the rice grown in the United States is also of the indica type. Japonica varieties most likely originated in China and are grown in regions of the world having temperate climates. Japonica varieties have a lower tillering capability than indicas and they are also more resistant to lodging. They are more responsive to fertilizer but exhibit lower resistances to insects and diseases. The amylose content of the grain is lower so cooking results in stickier and glossier grains that tend to disintregrate if boiled too long. Japonica varieties are primarily grown in Japan, Southern Europe and Central and Northern China (Christou, 1994).

4.2 Morphology and growth of rice

Cultivated rice is generally considered a semiaquatic annual grass. At maturity the rice plant has a main stem and a number of tillers. Plant height varies with variety and environmental conditions, ranging from approximately 0.4 to over 5 m in some floating rice. The morphology of rice is divided into the vegetative phases and reproductive phases. The rice seed, commonly called a grain , consists of the true fruit or brown rice (caryopsis) and the hull, which encloses the brown rice(Fig. 2). A single grain weights about 10-45 mg at 0% moisture content. Grain length, width, and thickness vary widely among varieties. Hull weight averages about 20% of the total grain weight. Brown rice consists mainly of embryo and endosperm. The surface contains several thin layers of differentiated tissues that enclose the embryo and endosperm. The outermost layer of endosperm is the aleurone layer, which differs both in morphology and function from the starchy endosperm. The aleurone layer may be one to seven cells thick and is thicker on the dorsal (back) than along the lateral (front) surface. Germination and seedling development start when seed dormancy has been broken, the seed absorbs adequate water and is exposed to a temperature ranging from about 10-40◦ C. The physiological definition of germination is usually the time when the radical or coleoptile (embryo shoot) emerges from the ruptured seed coat. The rice grain is similar to other cereal grains in terms of its mechanism of starch, protein and lipid biosynthesis

during grain ripening and in their breakdown during germination, when many metabolic and structural changes have been characterized. Many enzymes, including phosphatase, protease, esterase, lipase, peroxidase, catalase, β-glucosidase, α- and β-galactosidase increase by the fourth day of germination and have peak activities by the fifth to seventh day (Palmiano and Juliano, 1973). The site of production of many of these hydrolases during seed germination is the aleurone layer.

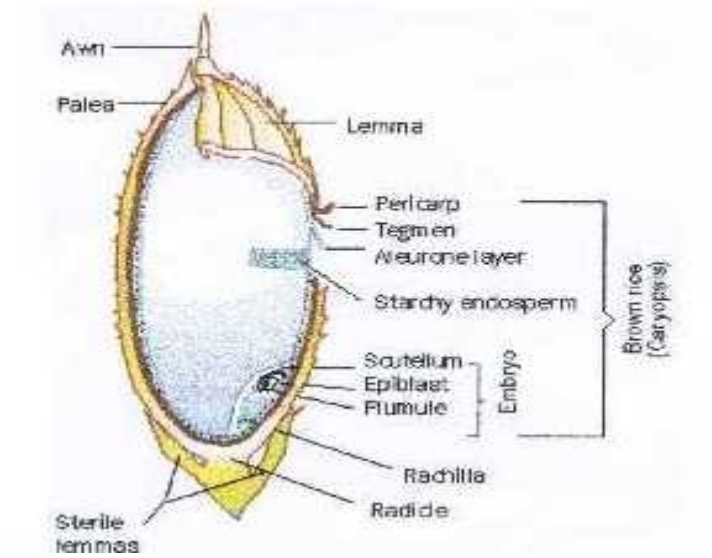


Figure 2. Structure of rice grain (adapted from Juliano and Adama, 1937)

5. How to obtain full-length of rice's cDNA

Most attempts to identify and isolate a novel cDNA result in the acquisition of clones that represent only a part of the mRNA's complete sequence (Fig. 3). These approaches include cDNA library screening and PCR amplification. Additional approaches may be used to attain the full length cDNA sequence and clone.

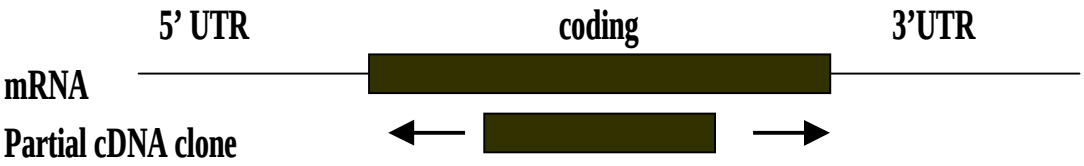


Figure 3. Schematic representation of an mRNA for which a cDNA representing only an internal portion of the transcript has been obtain.

Such circumstances often arise, for example, when open reading frame fragments are obtained from expression library, or Genbank Expressed Sequence Tag search.

The approaches described here are screening of cDNA library and screening of Expressed Sequence Tag (EST).

5.1 Screening of cDNA library

Each organism has a unique population of messenger RNA (mRNA) molecules. These mRNA populations are difficult to maintain, sequence, and amplify; therefore, they must be converted to more stable complimentary DNA molecules (cDNA). Successful cDNA synthesis should yield full-length copies of the original population of mRNA molecules. After synthesis, the cDNA is inserted into an *E. coli*–based vector (plasmid or λ phage), and the library is screened for clones of interest. Lambda has been the vector system of choice for cDNA cloning (Tangpradubkul, 1990). The fundamental reasons are that *in vitro* packaging of λ generally has a higher efficiency than plasmid transformation, and λ libraries are easier to handle (amplify, plate, screen, and store) than plasmid libraries. By using phage vector such as λgt10 and λgt11, libraries consisting 10^5-10^7 clones may be obtained (Lucotte and Baneyx, 1993). This high number of recombinant clones is the result of the efficiency and reproducibility of the *in vitro* encapsidation process of λ phage DNA in *E.coli*.

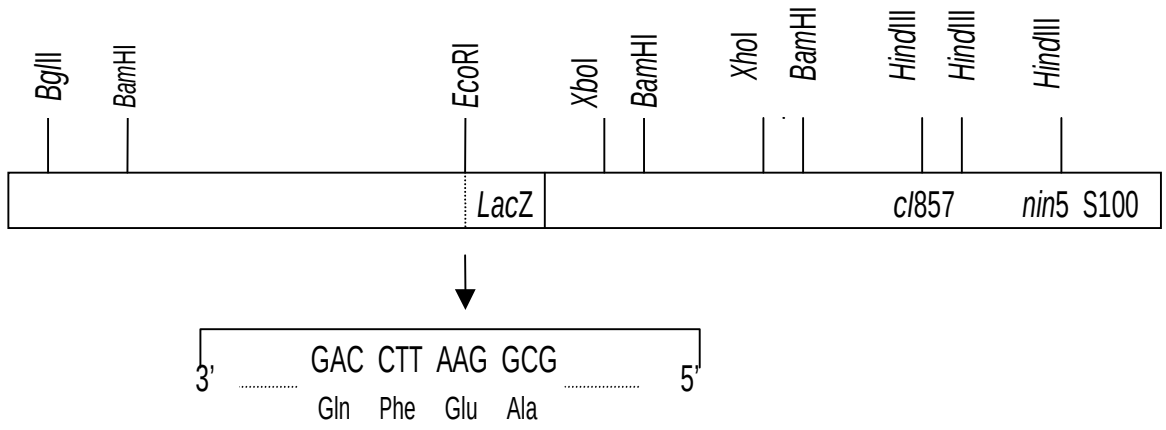


Figure 4. Map of λgt11. The position of the *lac*Z fragment is shown.

Lucotte and Baneyx (1993) stated that Bacteriophage λgt11 (*lac*5 *cI*857 *nin*5 S100) is an expression vector constructed by Young and Davis (1983). Cloning is performed at the unique *EcoR* I site located 53 basepair upstream of the translational termination codon of the lacZ gene (Fig.4). Approximately 7.2 kbp of DNA can be inserted at this position. λgt11 encodes a temperature-sensitive repressor (*cI*857) that is inactivated at 42◦ C and an amber mutation (*S*100) that prevent the phage from lysing host cells that do not carry a suppressor for this mutation. Foreign cDNA sequences cloned into λgt11 may be expressed as fusion proteins to β-galactosidase (the *lacZ* gene product). Figure 4. illustrates the expression vector for cDNA clones. For instance, the cDNA of interest can be fitted with *Sal* I and *EcoR* I adaptors and following digestion of both the modified cDNA and the vector by these enzymes, the cDNA can be ligated into the vector in the desired orientation. Probably the most commonly used method to screen a cDNA library is hybridization to a labeled DNA probe (Ian and Austin, 1997). This probe may be a single strand oligonucleotide or a double strand cDNA or PCR product. The DNA may be either radioactively or nonradioactively labeled. The sequence of an oligonucleotide probe may be derived from a number of sources, for example, a degenerate probe may be obtained by back translating a peptide sequence of an unknown protein, or may be a short region of conserved sequence within a cDNA from another species. Double-stranded DNA probes maybe a partial cDNA obtained from previous screening of a another library or a PCR product, or a cDNA from another member of the same gene family.

5.2 Screening of Expressed Sequence Tag (EST clone)

The rice genome project has provided another potential source of rice enzyme cDNA sequences in the form of expressed sequence tags (ESTs). Information about the number and expression of active genes is needed to completely utilize information in the rice genome. It is time consuming to make full-length cDNAs and sequence them completely, so partial sequencing of incomplete cDNA is used. These sequence data are used mainly for cataloguing expressed rice genes. The partial sequence data of such cDNAs are now available in the public database dbEST, held by the National Center for Biotechnology Information (NCBI), at http://www.ncbi.nlm.nih.gov/dbEST/index.html. About 80% of these EST data are from the Rice Genome Project (RPG). RPG has partially sequenced about 40,000 cDNA clones, mainly from their 5' ends (Shimamoto, 1999). The sources of these libraries were callus grown under several different conditions, young roots, green shoots, and panicles at different growing stages. So, they provide some expression information, as well as nucleic acid sequence. Clones for the sequence of interest can also be acquired from the EST source. Therefore, ESTs provide an important resource for cloning in genome project organisms, such as rice. The EST catalogue is used to identify the putative function of gene products by similarity search against public protein database with programs such as BLAST. About 30% of the translated amino acid sequences show significant similarity to known sequences from other species, including bacteria, plants, and animals (Sasaki *et al*., 1996). The huge amount of data from the large scale sequencing of many species have been incorporated into the protein database, and some match the rice ESTs, suggesting they may code for proteins with similar functions.

6. Objectives

To characterize β-galactosidase from germinating rice, including its primary structure from its cDNA sequence, and to characterize its expression during rice germination.

CHAPTER 2
MATERIALS AND METHODS

1. Materials

1.1 Plant Materials

Indica rice CV. jasmine (*Oryza sativa* L.) seeds were obtained from the Department of Agriculture, Bangkok. Japonica rice CV. Nipponbare (*Oryza sativa* L.) seeds were obtained from Prof. Dr. Tatsuhito Fujimura, Institute of Agricultural and Forest Engineering, University of Tsukuba, Ibaraki, Japan.

1.2 cDNA library

The immature rice seed cDNA library was kindly provided by Dr. Kouichi Mizuno, Institute of Agricultural and Forest Engineering, University of Tsukuba, Ibaraki, Japan. The cDNA library were constructed using rice plants *Oryza sativa* L. Japonica which were field grown in 1991 at Life Science Laboratory of Mitsui Toatsu Chemicals, Mobara (Mizuno *et al*., 1992).

1.3 Chemical and molecular reagents

Ampicillin, chloroform, chloroform/isoamyl alcohol (24:1), diethyl pyrocarbonate, and ethidium bromide (EtBr) were purchased from Sigma Chemical Company, St. Louis, MO, U.S.A. Agarose, deoxyribonucleotide (dATP, dCTP, dGTP,dTTP) and X-Gal were products from Promega, Madison, WI, U.S.A.

Superscript Reverse Transcriptase II and RnaseH were purchased from Gibco-BRL. Taq polymerase, T4 DNA ligase and restriction endonuclease, *Pst* I, *Sac* II, *Eco*R I and *Hin*c II were purchased from Promega (Madison, WI). ABI PRISM Dye terminator cycle sequencing kit was producted from Applied Biosystems Inc., Foster city, CA, U.S.A. Geneclean II kit was producted from BIO 101, Inc., La Jolla, U.S.A. All chemicals used in this study were of molecular biology or reagent grades. Other chemicals and solvents used but not listed here were purchased from variety of suppliers (Merck, Germany and Takara, Tokyo, Japan).

1.4 PCR primer

Degenerate oligonucleotides for DNA amplification were designed from conserved sequences between plant β-galactosidase proteins as shown in Table 2. were ordered from GIBCO-BRL (Life Technologies, Grand Island, NY).

Some DNA sequences were used to design primer for sequencing as described in the following methods. These primers including BgalM, rbRACE1f, rbRACE1r, rbRACE2f, rbRACE2r were synthesized from ESPEC OLIGO SERVICE CORP. (Ibaraki, Japan) except that the MOSM198f1, ME1405f1, and ME1405r1 were obtained from Geneset Oligos Ltd. (Singapore).

2. Methods

2.1 Crude enzyme extraction

The japonica rice seeds were germinated in the dark and prepared for analysis as previously described by Palmiano and Juliano (1973). Twenty grains were homogenized with a mortar and a pestle in ice cold 0.05 M sodium acetate buffer, pH 5.0 (2 ml per 1 g weight seed) The homogenate was centrifuged at 13,000 rpm for 30 min at 4◦C and the supernatant (crude enzyme) was used for assays.

2.2 Determination of β-Galactosidase activity

The crude extracts of germinating seed at 0-7 days were assayed by measuring the hydrolysis of p-nitrophenyl-β-D-galactoside. The reaction mixture consisted of 0.05 ml crude extract, 0.10 ml of 1 M sodium acetate buffer (pH 5.0), and 0.85 ml of 0.01 M of p-nitrophenyl substrate in 1 M sodium acetate buffer (pH 5.0) at 30◦ C for 10 min before being stopped by addition of 2M Na_2CO_3 (Suginta, 1995). The liberated *p*-nitrophenol was quantitated spectrophotometrically at 400 nm. One unit (U) of enzyme is the amount of enzyme that will release one μmole of *p*-nitrophenol per minute in this assay.

2.3 X-Gal staining

The japonica rice seeds at 0-7 days of germination were cut in half through the long axis, and the rice seeds were incubated in the staining solution (50 mM sodium phosphate buffer, 10 mM EDTA, 0.1% sarkosyl, 20% MeOH, 1 mM X-Gal in distilled water) for 30 min and destained

in MeOH (modified from Kosugi, 1996). After incubation the staining was photographed under a stereomicroscope (Nikon SMZ-U) at 10X magnificant.

2.4 Cloning and sequencing of β-Galactosidase fragment

2.4.1 Oligonucleotide primer design

Plant and related β-Galactosidase protein sequences were retrieved from the NCBI protein database and aligned using the Propack programs of Feng and Doolittle (1996). Conserved sequences were identified by inspection of multiple sequence alignments (Cairns *et al*., 1999) Protein sequences were back translated to yield all possible nucleotide sequences. Degeneracy was calculated and sequences with lowest degeneracy were used to make nested primers (inner and outer sets) for PCR. The derived sequences were used to order degenerate oligonucleotides.

Table 2. Degenerate oligonucleotides primers converting to nucleotide sequence coding for the consensus protein sequence using in amplification of β-galactosidase cDNA fragment

Primer name	Primer length	Sequence (5'to3')
bgalI173f (3f)	21	ATI SAR ACI TAY GTI TTY TGG
bgalG130f (3r)	18	GGI GGI TTY CCI GTI TGG
bgalW127f (6f)	26	GGA AYT WYG GIG GIT TYC CIG TNT GG
bgalD323r (6r)	20	TCI ARI GGI GCR TCR TAR TC
bgalF304r (7f)	17	AAR TTI GTI CCI CCR TG
bgalT259r (7r)	26	GTC CAI DYY TCI GTC CAC ATY TIN GG

I = inosine; other letters are according to IUPAC code for nucleotides and nucleotide combinations.

2.4.2 Total RNA Isolation

Isolation of total RNA from cells and tissues was done using TRIZOL Reagent (GIBCO-BRL, Grand Island, NY) developed by Chomczynski and Sacchi (1987). 100 mg of germinated rice seeds were homogenized by grinding with mortal and pestle in liquid N_2 and

transferred into a new sterile microcentrifuge tube containing 1 ml of Trizol reagent. The homogenized sample was incubated for 5 min at room temperature, followed by adding 0.2 ml of chloroform. The mixture was mixed vigorously by inversion for 15 sec, incubated at room temperature for 2-3 min, and then centrifuged at 12,000g for 1 min at 4° C. The colorless upper aqueous phase containing RNA was transferred into a fresh tube. The RNA from the aqueous phase was precipitated by mixing with 0.5 ml of isopropyl alcohol, left at room temperature for 10 min, and pelleted by centrifugation at 12,000g for 10 min at 4° C. The pelleted RNA was washed with 75% (V/V) ethanol (in DEPC-treated water) and dried briefly at 37° C. Finally the dried RNA pellet was resuspended in 20 μl of DEPC-treated water. The RNA solution was immediately used to synthesize first strand cDNA as described in next step or kept at -70° C. The concentration of RNA solution was determined by measuring the UV absorbance at 260 nm of a 1 μl aliquot, diluted to 1 ml in distilled water. One absorbance unit was assumed to be equivalent to 40 μg/ml RNA

2.4.3 First-strand cDNA synthesis

Total RNA extract was used as the template for the synthesis of first-strand cDNA using Superscript Reverse Transcriptase II (GIBCO-BRL). The reaction mixture was composed of 5 μg total RNA, 1 μl Oligo $(dT)_{17}$ (10 pmol/μl) and sterile distilled water up to 12 μl. The mixture was heated at 70°C for 10 min and quick chilled on ice. The contents of the tube were collected by brief centrifugation and 4 μl of 5X First Strand Buffer (250 mM Tris-HCl, pH 8.3, 375 mM KCl and 15 mM $MgCl_2$) 2 μl of 0.1 M DTT, 1μl of 10 mM dNTP mix (10 mM each dATP, dGTP, dCTP and dTTP at neutral pH) were added. The contents of the tube was mixed gently and incubated at 42° C for 2 min. Then 1 μl (200 units) of SUPERSCRIPT II was added and mixed by pipetting gently up and down. The mixture was incubated at 42° C for 50 min, and the reaction was stopped by heating at 70° C for 15 min. To remove RNA complementary to the cDNA, 1 μl (2 units) of *E. coli* RNase H was added and incubated at 37° C for 20 min. The cDNA was used as a template for amplification in PCR.

2.4.4 DNA amplification

In order to obtain the β-galactosidase cDNA fragment. Three sets of DNA amplification by polymerase chain reaction (PCR) were performed using 10% of the first-strand cDNA reaction as a template with the same component but difference primers (Table 3.). The

reaction mixture was composed of 10% first-strand cDNA, 50 pmol forward and reverse primers, 0.2 mM dNTP mix, 15 mM $MgCl_2$, 10X PCR buffer (500 mM KCl, 100 mM Tris-HCl, pH 9.0 and 1% Triton X-100), Taq Polymerase (5 unit/μl) and autoclaved distilled water to make the volume up to 25 μl. Three sets of PCR programs were used as shown in Table 3.

Table 3. PCR condition of rice cDNA with bgal primers

Template	Forward primer	Reverse primer	Annealing Temperature
Rice cDNA	bgalI73f	bgalD323r	40◦ C
Product from first set	bgalG130f	bgalF304r	50◦ C
Product from second set	bgalW127f	bgalT259r	55◦ C

2.4.5 DNA analysis

The amplified PCR products were analyzed by electrophoresis on 1% agarose gel in 1X TBE buffer. Aliquots of DNA sample were mixed with 6X loading dye (0.25% bromophenol blue, 0.25% xylene cyanol FF and 30% glycerol). The sample in loading dye and the DNA marker, were applied to the gel wells. Electrophoresis was performed at constant 80 V for 30 min using 1X TBE buffer. After electrophoresis, the gel was stained with 10 μg/ml of ethidium bromide solution for 10 min and destained with distilled water. The band in the gel was visualized by UV light transillumination.

2.4.6 Purification of DNA fragments from agarose gels

The DNA bands in agarose gels were extracted using the Gene Clean II Kit (BIO101, Vista, CA). The DNA bands on the agarose gel visualized under UV light, were cut out of the gel with a razor and transferred into 1.5 ml microtubes. The agarose was dissolved by adding 3 volumes of 6 M NaI and incubating at 50◦ C for 5 min with mixing every minute, then the solution was mixed with 5 μl of GLASSMILK® suspension and placed at room temperature for 5 min. The silica matrix with the bound DNA was pelleted by centrifugation at 12,000g for 5 sec. After the NaI solution was removed, the pellet was washed 3 times with 500 μl of NEW WASH solution. Finally, the bound DNA was eluted with 10 μl of sterile distilled water and then

centrifuged at 12,000g for 5 sec. The supernatant containing the eluted DNA was removed and transferred to a new tube. The concentration of DNA solution was determined by measuring the absorbance at 260 nm.

2.4.7 Ligation of DNA fragment into vector

Ligation of gel-purified PCR products into pGEM-T vector was done according to the plasmid supplier's recommendation. The reaction mixture (10 μl) was composed of 0.5 μl of pGEM-T vector (50 ng/μl), 7.5 μl of eluted DNA fragment (10-50 ng), 1 μl of 10X T4 DNA ligase Buffer, 1 μl of T_4 DNA ligase. The mixture solution was incubated at 14◦ C for overnight.

2.4.8 Preparation of competent cell

Competent *E. coli* DH5α cells were prepared using the $CaCl_2$ method (Sambrook et al., 1989). A single colony of *E. coli* was inoculated into 100 ml of LB broth and incubated the culture for 3 h at 37◦ C with vigorous shaking until the optical density (OD) at 600 nm was approximately 0.3-0.4. The cells were transferred into sterile disposable, ice cold 50 ml polypropylene tube, chilled on ice for 10 min and the cells were recovered by centrifugation at 4000 rpm for 10 min at 4◦ C. The media was decanted from the cell pellets, and each pellet was resuspended in 10 ml of ice cold 0.1 M $CaCl_2$ and stored on ice. The supernatant from the cell pellets was decanted, and the tube was placed in an inverted position for 1 min to allow the last traces of fluid to drain away. Each pellet was resuspended in 2 ml of ice cold 0.1 M $CaCl_2$ for each 50 ml of original culture. The suspension of competent cells was transferred in 100 μl aliquots to sterile microtubes. These competent cells were transformed immediately or kept at -70◦ C freezer.

2.4.9 Transformation and selection

An aliquot of frozen competent cells was thawed on ice. The ligation mixture was added to the thawed competent cells, mixed gently by swirling and store the tube on ice for 30 min. The cells were heat shocked at 42◦ C for 90 sec and the tube rapidly transferred to ice for 2 min. The transformed cells were grown by adding 800 μl of SOC medium and the cultures incubated at 37◦ C for 1 h. This culture was spread on an LB plate containing 100 μg/ml ampicillin along with 100 μl of 100 mM IPTG and 20 μl of 5 mg/ml X-Gal and incubated at 37◦ C overnight. The recombinant clones were identified as white colonies, while colonies containing nonrecombinant plasmid were blue.

Plasmids were prepared using QIAprep Spin Miniprep Kit protocol, and by restriction digest with *Pst I* and *Sac II* and 1% agarose gel electrophoresis to determine which clone has the recombinant plasmid, containing the expected fragment.

2.4.10 DNA sequencing and analysis

Purified plasmid was sequenced in cycle-sequencing using the ABI PRISM dye labeled terminator kit according to the recommended protocol. The reaction mixture composed of 4 μl of Terminator Ready Reaction Mix, 1 μl purified plasmid DNA (100 ng/μl), 1 μl primer (5 pmol/μl) and 4 μl of distilled water. Amplification was performed using Gene Amp PCR system. The program was conducted by rapidly increasing the temperature to 96◦ C, followed by 25 cycles with the program : 96◦ C 10 s, 50◦ C 5 s and 72◦ C 4 min. The amplified products were purified by ethanol precipitation as described in the kit manual and loaded onto an automated ABI 373 DNA automated sequencer. The peptide translation of the cDNA fragment sequence of rice β-Galactosidase determined was searched to find their location in protein sequence by using BLASTX at the Basic Blast facility of the Genbank Database (http://www.ncbi.nlm.nigh.gov). After the DNA sequence was interpreted and converted to single letter amino acid code, it was analyzed using the sequence analysis programs from the BCM Search Launcher Index.

2.5 Southern blot analysis

2.5.1 DNA isolation from rice by CTAB method

Three grams of japonica and indica rice shoot age 10 days were homogenized separately with mortar and pestle in liquid N_2. The homogenate was resuspended in 4.5 ml of 2X CTAB solution (modified from Doyle and Doyle, 1987) to which was added β-mercaptoethanol to a final concentration of 5% before used. The mixture was incubated at 56◦ C for 5 min, then 4.5 ml of 24:1chloroform/isoamyl alcohol was added, it was gently shaken and incubated at 56◦ C for 30 min. The mixture was pelleted by centrifugation at 5,000 g for 5 min at 4◦ C and the supernatant was transferred into a new tube. 1/10 volume of 10X CTAB solution was added, followed by 1 volume of chloroform/isoamyl alcohol and the mixture was incubated at 56◦C for 30 min. The solution was centrifuged at 5000 rpm for 5 min at 4◦ C and the supernatant was transferred into a new tube, 1 volume of precipitation solution (1% CTAB, 50 mM Tris-HCl pH 8.0, 10 mM EDTA) was added, gently shaken and incubated at –20◦ C for 10-20 min. The DNA

was pelleted by centrifugation at 5000 rpm for 5 min at 4◦ C and the pellet dissolved in 1.5 ml of TE. The DNA solution was transferred into two microcentrifuge tubes (about 750 μl/tube) and 750 μl of isopropanol was added to each. The mixture was incubated at –20◦ C for 5 min and centrifuged at 12,000 rpm for 5 min at 4◦ C. The supernatant was transferred into a new tube and 2 volumes of absolute ethanol was added. The DNA solution was incubated at –80◦C for 10 min and centrifuged at 12,000 rpm for 10 min at 4◦ C. The DNA pellet was rinsed with 500 μl of 70% ethanol and centrifuged at 1,200 rpm for 3 min at 4◦ C, followed by removal of the supernatant and drying of the DNA pellet by under vacuum. The DNA pellet was dissolved by adding 300 μl of distilled water. In order to get rid of the RNA, 30 μl of RNase 1 μg/μl was added into the DNA solution and incubated at 37◦ C for 10 min. Then 5/8 volume of polyethyleneglycol (PEG 800) solution was added and keep on ice for 10 min at 4◦ C. The DNA pellet was rinsed with 500 μl of 70% ethanol, centrifuged again at 12,000 rpm for 3 min at 4◦ C, the supernatant removed and the DNA pellet dried under vacuum. Finally, 300 μl of distilled water or TE buffer was added and kept at 4◦ C in the refrigerator to dissolve overnight. The concentration and purity of the DNA solution was determined by measuring the A_{260} and $A_{260/280}$ ratio

2.5.2 Transfer of DNA onto nitrocellulose membrane

Aliquots of the rice genomic DNA (15 μg) were digested with 30 U *Eco*R I , 30U *Bam* HI , and 30 U *Hin*d III at 37◦ C overnight. DNA digested with each enzyme was precipitated by adding 20 μl of 3M sodium acetate and 400 μl of absolute ethanol, chilling the DNA solution at –80◦ C for 5 min and centrifuged at 12,000 rpm for 10 min at 4◦ C. The DNA pellet was washed with 70% ethanol, centrifuged at 12,000 rpm for 5 min at 4◦ C, the supernatant removed and the pellet dissolved with 10 μl of TE. The DNA was separated by electrophoresis on a 0.8% agarose gel in TAE buffer. The DNA was denatured by soaking the DNA gel once in denaturation solution (1.5 M NaCl, 0.5 N NaOH) for 15 min, rinsed twice in neutralization solution (1 M Tris [pH 7.4], 1.5 M NaCl) for 15 min and DNA was transferred from gel onto nitrocellulose membrane for overnight by capillary transfer. The DNA was immobilized on the membrane by UV crosslinking for 3 min on UV transiluminator and then the membrane was soaked in 400 mM NaOH for 10 min followed by 2X SSC for 10 sec. The membrane was dried for 10 min at room temperature and kept at 4◦ C until hybridization.

2.5.3 Radio active labeling of cDNA fragment by the random primer method.

The blot was hybridized to the RT-PCR fragment of indica rice β-galactosidase labeled with random primer *Bca*BEST ™ Labeling Kit system. The reaction was carried out in a total volume of 25 μl containing 1.5 μl of β-Galactosidase RT-PCR DNA fragment (25 ng), 13.5 μl of distilled water, and 2 μl of random primer. The mixture was denatured at 100◦ C for 3 min and cooled on ice for 5 min followed by adding 2.5 μl of 10X buffer, 2.5 μl of dNTP mixture, 2 μl of α–^{32}P dCTP, and 1 μl of *Bca* BEST DNA polymerase were added. The mixture was incubated at 55◦ C for 15-30 min. To stop the reaction, 1 μl of 0.5 M EDTA was added and it was incubated at 70◦ C for 10 min. The mixture was adjusted to 100 μl with TES buffer (0.1 M NaCl, 10 mM Tris-Cl pH 8.0, 1 mM EDTA pH 8.0) and centrifuged through Sephadex G-50 spin column which was equilibrated with TES buffer to remove unincorporated nucleotides. The labeled probe was kept at –20◦ C. Prior to hybridization, the probe was heated at 100◦C for 5 min and immediately put on ice for approximately 10 min.

2.5.4 Hybridization and detection

The nitrocellulose membrane prepared in section 2.5.2 was put into a plastic hybridization bag, with 2 ml of prehybridization buffer (20X SSPE, 10% SDS, 50X Denhart solution). The membrane was prehybridized at 60◦ C for 2 h in a hybridization oven, then 36 μl of labeled denatured DNA probe (4 ng) mixed with 36 μl of salmon sperm DNA was added to the bag. Then, the hybridization reaction was performed at 60◦ C for 18 h. After hybridization, the nitrocellulose membrane was washed with 2X SSC for 10 min, 2X SSC containing 0.1% SDS for 10 min and 2X SSC for 10 min at room temperature and washed twice with 2X SSC containing 0.1% SDS at 65◦ C for 10 min, followed by a final wash with 2X SSC for 10 min at room temperature. The washing solution was aspirated and the membrane was dried out at room temperature for 1-2 min. The membrane was placed in the plastic bag and exposed to high sensitivity X-ray film (Hyper film ™-MP, Amersham). The membrane and film were placed inside a cassette and exposed at –80◦ C for 10 hr before developing.

2.6 Northern blot analysis

2.6.1 RNA isolation by CTAB method

RNA was isolated from shoot and root of indica and japonica rice at 1, 3, and 5 days of germination using CTAB isolation buffer modified from Clendennen and May (1997). Shoot and root tissues were ground to a powder in liquid nitrogen using a mortar and pestle. Three grams of frozen powder was added to 20 ml of 2X CTAB RNA isolation buffer [2% CTAB, 0.1 M Tris-HCl pH 9.5, 20mM EDTA, 1.4 M NaCl, 5% (v/v) β-mercaptoethanol]. The mixture was incubated at 65◦ C for 10 min. Samples were extracted twice in an equal volume of chloroform/isoamyl alcohol (24:1) and the phases were separated by centrifugation at 13,000 rpm for 10 min at room temperature and the aqueous phase transferred to a new tube. Following centrifugation, an equal volume of phenol: chloroform was added to the aqueous phase, mixed by shaking and centrifuged at 13,000 rpm for 10 min at room temperature. One fourth volume of lithium chloride was added to the supernatant, and the RNA was allowed to precipitate for 2 h at –20◦ C. RNA was pelleted by centrifugation at 13,000 rpm for 10 min at 4◦ C. The RNA pellet was dissolved with 500 μl of TE and the sample was centrifugated again at 13,000 rpm for 10 min at 4◦ C, 500 μl of TE saturated phenol pH 9.0 was added to the supernatant and it was mixed and centrifuged at 15,000 rpm for 10 min at room temperature. The aqueous phase was extracted by adding 500 μl of phenol-chloroform with vigorously shaking and the phases were separated by centrifugation at 13,000 rpm for 10 min at 4◦ C. Fivehundred microliters of chloroform: isoamyl alcohol (24:1) was added and centrifuged at 15,000 rpm for 10 min at room temperature. One fourth volume of 10 M lithium chloride was added, mixed and incubated at –20◦ C at for 2 h. The RNA was pellet by centrifugation at 15,000 rpm for 15 min at 4◦ C, washed with 70% ethanol and resuspended in 15 μl of 0.1% DEPC-treated water.

2.6.2 Transfer RNA onto nitrocellulose membrane

Ten micrograms of each of the RNA samples was denatured at 50◦C for 2 h and separated by electrophoresis through 1.2% agarose gel at 250 V for 2 h in 0.01M phosphate buffer and transferred to nitrocellulose membrane for overnight by capillary transfer. The denaturation reaction was carried out in 24 μl of 50% DMSO, 10 mM phosphate buffer and 1 M glyoxal. The membrane was treated with 50 mM NaOH for 5 min, washed in 2X SSC for 1 min, then dried at room temperature for 5 min and baked at 80◦ C for 2 h. The RNA blot was hybridized with a

cDNA probe prepared as previously described for Southern blot (section 2.5.3) and then hybridized at 60◦ C for 18 h. After hybridization the membrane was washed as previously described in Southern blot analysis. Detection was performed by autoradiography. The membrane and the X-ray film were placed inside an autoradiography cassette and exposed at –80◦ C for 4 days before developing.

2.7 Cloning and sequencing of the full-length cDNA encoding β-galactosidase enzyme from immature rice seed cDNA library

2.7.1 First round screening of the cDNA library and collection of plaques

A single colony of *E.coli* (Y1090r$^-$) was inoculated into 500 ml of LB medium containing 0.2% maltose, 10 mM $MgSO_4$. The suspension was incubated with shaking at 37◦ C for 4 h or until the OD_{600} was approximately 1.0, then 2 ml of this culture was mixed with 350 μl of 1000X dilution λ phage solution. The mixture was incubated with shaking at 37◦C for 15 min to allow phage particles to be adsorbed to the host cells. To propagate the λ phage, 45 ml of 0.8% top agarose containing 10 mM $MgSO_4$, 25 μg/ml ampicillin was added to the mixture and immediately poured onto a LB big plate (20X40 cm). After hardening of the agarose, the plate was inverted and incubated for 8 h at 42◦ C until cleared plaques appeared. Phage particles, the plaques on the LB plate were transferred to a 10X20 cm nitrocellulose membrane by applying the reactive side of the membrane on top of the agar. The phage were allowed to adsorb to the membrane for 5 min. Then the membrane was gently peeled off and dried for 5 min at room temperature. The phage particles were lysed by treating the membrane with 1X denaturing solution containing 0.2 M NaOH and 1.5 M NaCl on filter paper for 5 min. The membrane was dried for 5 min at room temperature, followed by treating the membrane twice with 1X neutralizing solution containing 1M Tris [pH 7.4], and 1.5 M NaCl. The DNA was immobilized by baking the treated membrane in a hot air oven at 80◦ C for 2 h. The nitrocellulose membranes prepared as described above were folded in a hybridization plastic bag. About 10 ml of hybridization solution was transferred into the plastic bag to soak the membrane. The membrane was prehybridized at 60◦ C for 6 h and then β-galactosidase RT-PCR fragment (labeled DNA probe), prepared as previously described (section 2.5.3) was added to the plastic bag. The

hybridization reaction was performed at 60◦ C for 18 h. After hybridization, the nitrocellulose membrane was washed twice with 2X SSC for 10 min at room temperature and twice with 2X SSC containing 1% SDS at 60◦ C for 10 min each and once with 2X SSC for 20 min at room temperature. The washing solution was aspirated and the membrane was dried at room temperature for 5 min. The membrane was covered with the syran wrap and exposed to high sensitivity X-ray film (Hyper film ^{TM}-MP, Amersham). The membrane and film were placed inside a metal cassette and incubated at –80◦C for 4 h before developing.

Plaques that gave very high hybridization signal were collected. Top agarose containing the positive plaques were picked from the plate by using the wide end of a pasteur pipette and transferred to 1 ml of sterile PDB buffer [100 mM NaCl, 8 mM $MgSO_4$, 50 mM Tris-HCl (pH 8.0), 0.01% (w/w) gelatin] in a microcentrifuge tube. The tube was then vigorously mixed by vortexing. About 3-4 drops of chloroform was added. The tube was kept for 1 h at room temperature, followed by centrifugation at 14,000 rpm for 5 min at 4◦C to separate the supernatant from the pellet before keeping the supernatant at 4◦ C.

2.7.2 Second and the third rounds screening of the cDNA library.

Two microliters of 1000X dilution λ phage from the first round screening was mixed with 200 μl of *E.coli* (Y1090r-) prepared as described in method 2.7.1. The mixture was incubated with shaking at 37◦C for 15 min to allow phage particles to be adsorbed to the host cell. To propagate the λ phage, 3 mL of 0.8% top agarose containing 10 mM $MgSO_4$, 25 μg/ml ampicillin was added to the mixture and immediately poured onto a 10 cm circular LB plate. After hardening of agarose, the plate was inverted and incubated at 42◦ C for 8-10 h or until clear plaques appeared. The plaques were transferred to 3X7 cm nitrocellulose membranes by applying the nitrocellulose membrane on top of the agar. Following these plaques lifts, hybridization, detection by autoradiography and collection of the positive plaque was done as previously described. A third round of screening was conducted in the same manner to allow isolation of an individual clone.

2.7.3 Isolation of plaque DNA

After detection of the third screening by autoradiography, plaques that gave high hybridizing signal were collected. Top agarose containing a single positive plaque was picked out from plate by using pasteur pipette and transferred to 200 μl of PDB buffer [50 mM Tris-Cl

(pH8.0), 100 mM NaCl, 8 mM $MgSO_4$, 0.01% (w/v) gelatin] in a sterile microtube. Chloroform was added to a final concentration of 10% (v/v). The tube was then vigorously vortex and centrifuged at 14,000 rpm for 10 min at 4◦ C to separate the supernatant before keeping at 4◦ C.

Six clones from the collection were randomly selected. Each of these selected plaques was transferred into host cells according to the method of section 2.7.2. After clear plaques appeared, 5 ml of PDB was added the plate and shaken for 1 h at room temperature, followed by completely collecting the supernatant into a 15 ml sterile tube using pasteur pipette. About 3-4 drops of chloroform was added. The tube was vigorously vortexed and spun down at 14,000 rpm for 10 min at 4◦ C. The supernatant containing phage particle was transferred into the new tube and kept at 4◦ C. To amplify the phage, 4 μl of the phage solution was mixed with 200 μl of the bacterial suspension (Y1090r-) and incubated at 37◦ C for 15 min. The mixture was then transferred to 20 ml of LB medium containing 25 μ/ml ampicillin and shaking continued at 37◦ C for 4-6 h until cell debris were appeared.

To complete cell lysis, 200 μl of chloroform was added and the tube were vortexed. Cell debris in the lysate was removed by centrifugation at 5,000 rpm for 5 min at 4◦ C and then the supernatant was transferred into a new tube. The solution was mixed with an equal volume of 20% (w/v) polyethyleneglycol (PEG800) and 10 ml of 2.5 M NaCl, and incubated on ice for 1 h. Phages particles were recovered by centrifugation of the mixture at 5,000 rpm for 10 min at 4◦C. The supernatant was discarded and the precipitate was allowed to dry at room temperature. The phage particles were resuspended in 1 ml of TE buffer pH 8.0. In order to digest the host's DNA and RNA, Dnase I and Rnase A were added to a final concentration of 1 mg/ml each, the mixture was incubated at 37◦ C for 1 h followed by addition of 200 μl of chloroform and centrifugation of the mixture at 12,000 rpm for 10 min at 4◦ C. The supernatant was transferred into new tube and 50 μl of 10% SDS, 40 μl of EDTA and 1 μl of proteinase K was added, followed by incubation at 65◦ C for 30 min. The recombinant DNA was extracted by the phenol-chlorofom method (see section 2.5.1) and the DNA solution was stored at 4◦ C until use.

2.7.4 Determination of the size of DNA insert

The size of the cDNA insert in the phage DNA from the previously step was determined by using the enzyme *EcoR* I to excise the cDNA insert. The restriction mixture was composed of 10 μl of DNA (50 ng), 3 μl of *EcoR* I (12 U/μl), 3 μl of 10XBuffer H and 14 μl of

distilled water. The mixture solution was incubated at 37◦ C overnight. DNA was separated on an 8% polyacrylamide gel. After staining with ethidium bromide, the restriction fragment band was excised from the gel and DNA eluted using Gilbert buffer (0.5M $AcONH_4$, 10mM $MgCl_2$, 1mM EDTA, 0.1% SDS) and ligated into vector pUC19 and transformed into competent DH5α cell followed by isolation of the plasmid. To determined DNA sequence, the recombinant plasmid was sequenced on a ABI 373 automated DNA sequencer using the ABI PRISM dye label terminator kit (section 2.4.10). In order to obtain the remaining DNA sequence, the recombinant plasmid insert was digested with *Hin*c II endonuclease, since there is a *Hin*c II site present in cDNA fragment. The reaction mixture contained 10 μl purified plasmid, 5 μl *Hin*c II (10U/μl), 10 μl 10X Buffer A and 75 μl distilled water. The mixture was incubated at 37◦ C overnight and 8% polyacrylamide gel electrophoresis was done to determine the DNA size. DNA fragments were purified from the gel and cloned into pUC19 at the *Sma* I site. Beside that, two sets of DNA sequences were designed to use as a sequencing primers to obtain the unknown region. (Fig. 5 and Table 4.).

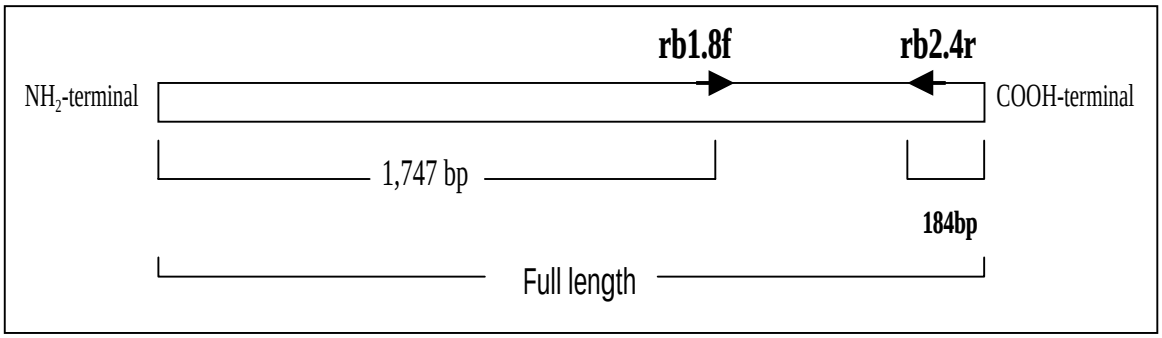


Figure 5. Estimated distance of primers obtained from DNA sequence of rice β-galactosidase for use in sequencing.

Table 4. Oligonucleotide primers using in sequencing of β-galactosidase cDNA fragment

Primer name	Primer Length	Sequence (5' to 3')
rb1.8f	20	GCA GAC TTG GTC CTA TCA GG
rb2.4r	20	AAT GTT CCG CAA GTT CCA AG

2.7.5 Amplification and sequencing of the cDNA encoding the 5' end of the β-galactosidase cDNA coding region.

To obtain the full length of cDNA, specific primers were designed from the corresponding gene in the Monsanto databank and use as a primer in reaction mixture. Total RNA from germinated rice seed was extracted by using TRIZOL Reagent as describe in method 2.4.2. First strand cDNA synthesis was performed with a rbRACE1r primer, derived from the internal cDNA β-galactosidase fragment. The mixture was prepared by using method 2.4.3 with 10 pmol rbRACE1r. After the reaction was finished, the reaction was treated with 1 μl of RNaseH to destroy the RNA template by incubating at 37◦ C for 20 min. For DNA amplification, 1 μl of the reverse transcribed cDNA was amplified with 10 pmol MOSM198f1 primer and rbRACE1r primer and amplified by using program described in Table 6. A second amplification was performed by using 1 μl of the first PCR with 10 pmol MOSM198f1 primer and rbRACE2r primer according to the program described in Table 7. After the PCR experiment was finished, PCR product was analyze by 1% agarose gel electrophoresis. The specific PCR product (expected size about 150-200 bp) on gel was eluted by using Gene clean II kit as described previously (method 2.4.6). The gel purified product was cloned into pGEM-T plasmid and sequenced. Then the peptide translation of the cDNA fragment was searched to find their location by using BLASTX at the NCBI Genbank database.

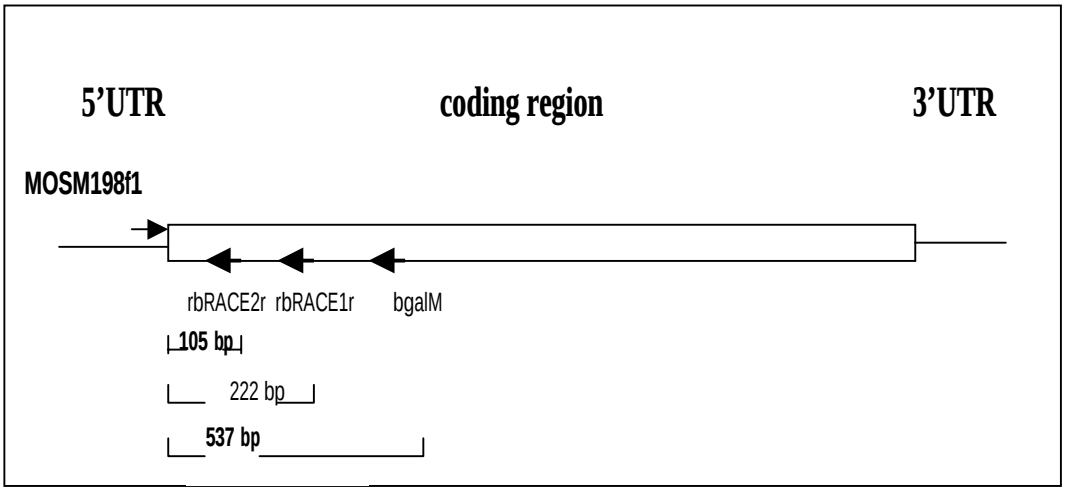


Figure 6. Estimated distance of oligonucleotide primer obtained from β-galactosidase gene using in amplification of 5'terminus cDNA fragment and sequencing primers obtained from DNA sequence of library clone.

Table 5: Oligonucleotide primers used in amplification of the cDNA encoding the N-terminus

Primer name	Primer Length	Sequence (5' to 3')
MOSM198f1	19	TGA GGA GAG TGG GTG AGT G
rbRACE1r	18	CTG GAT CAC ATC CAA GCC
rbRACE2r	21	GGG TGT GCT CCT CGG GTA ATG
bgalM	21	GGT CCT ATT ATC CTC TCT CAG

Table 6. First PCR program for amplification of the 5' terminus of β-galactosidase cDNA using MOSM198f1 and rbRACE1r primers

Cycle	Step	Temperature ◦C	Time (min)
Prerunning	denaturation	95	1.00
1-35	denaturation	94	1.00
	annealing	54	1.00
	extension	72	5.00
Postrunning	extension	72	5.00

Table 7. Second set of PCR program for amplification of the 5' terminus of β-galactosidase cDNA using MOSM198f1 and rbRACE2r primers

Cycle	Step	Temperature ◦C	Time (min)
Prerunning	denaturation	95	1.00
1-35	denaturation	94	1.00
	annealing	58	1.00
	extension	72	5.00
Postrunning	extension	72	5.00

2.8 Cloning and sequencing of the full length cDNA encoding β-galactosidase enzyme from panicle at flowering stage of rice (EST clone)

The closely related full length asparagus β-Galactosidase protein sequence was used to screen the EST database by protein sequence similarity (tBLASTN) and the rice β-Galactosidase from panicle at flowering stage of rice clone (E1405) in pBluescriptII SK +/- vector was requested from the Japanese Ministry of Agriculture, Forestry&Fisheries (MAFF) and analyzes by 8% polyacrylamide gel before and after digest with *Hin*c II. Since it appeared to be near full-length, it was then sequenced using T7 and T3 primers. To obtain the full length of cDNA

sequence, the restriction enzyme analysis was performed to cut cDNA into small fragment and cloned into pUC19 plasmid and sequenced by M13 forward and M13 reverse primer. The restriction reaction was composed of 10 μl of DNA (50ng), 5 μl of *Hinc* II (10 U/μl), 10 μl of Buffer A, and 75 μl of distilled water. The reaction was incubated at 37◦ C overnight. In order to obtain the remaining unknown region, four nucleotide sequences from 3' end was used to designed oligonucleotide primers for sequencing (Fig. 7 and Table 8).

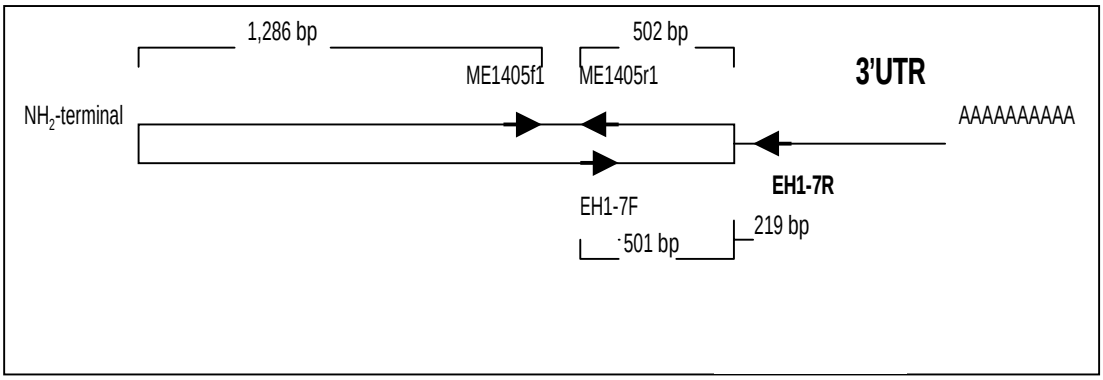


Figure 7. Estimated distance of sequencing primers obtained from DNA sequence of rice β-galactosidase EST clone

Table 8. Sequencing primers obtained from 3'end mRNA of rice β-galactosidase EST clone

Primer Name	Sequence 5' to 3'	Primer Length
ME1405f1	5' TCA CCT GGC AGT CGT ACG 3'	18
ME1405r1	5' AGC ACG CCG ATG TTC CAC 3'	18
EH1-7f	5' TGG AAC ATC GGC GTG CTC 3'	18
EH1-7r	5' GTA CGC CGG AGC GGC TG 3'	17

CHAPTER 3
RESULTS AND DISCUSSION

1. β-galactosidse activity during seed germination

Crude enzyme derived from twenty grains of japonica rice seed at 0-7 days germination was assayed by measuring the hydrolysis of p-nitrophenyl-β-D-galactoside by absorbance at 400 nm. β-galactosidase activity is present in dry and germinating rice seed. Enzyme activity, hydrolyzing pNP-galactose increased between 0 and 6 days germination. β-galactosidase activity, either in units per gram fresh weight or units per seed, reached a maximum after 6 days of germination and then decreased on the seventh day . (Table 9 and Fig. 8). Comparable results have been reported by Palmiano and Juliano (1973) during germination of IR8 rice seeds. In the IR8 seeds, many of the enzymes: phosphatases, proteases, esterases, peroxidase, catalase, β-glucosidase, α-and β-galactosidase increase along with soluble protein by the fourth day of germination and have peak activity by the fifth to seventh day.

Table 9. β-galactosidase Enzyme activity (units) during 0-7 days of germination

Day	Enzyme activity (U)	Total activity	Units/ gram fresh weight	Units/seed
0	2.21×10^{-4}	5.30×10^{-3}	9.50×10^{-3}	2.70×10^{-4}
1	6.71×10^{-4}	1.90×10^{-2}	2.60×10^{-2}	8.70×10^{-4}
2	1.20×10^{-3}	3.40×10^{-2}	4.50×10^{-2}	1.60×10^{-3}
3	1.59×10^{-3}	4.95×10^{-2}	6.20×10^{-2}	2.20×10^{-3}
4	2.19×10^{-3}	8.09×10^{-2}	8.70×10^{-2}	3.50×10^{-3}
5	2.46×10^{-3}	9.18×10^{-3}	9.70×10^{-2}	3.90×10^{-3}
6	3.44×10^{-3}	1.24×10^{-1}	1.30×10^{-1}	5.50×10^{-3}
7	1.60×10^{-3}	6.41×10^{-2}	6.30×10^{-2}	2.70×10^{-3}

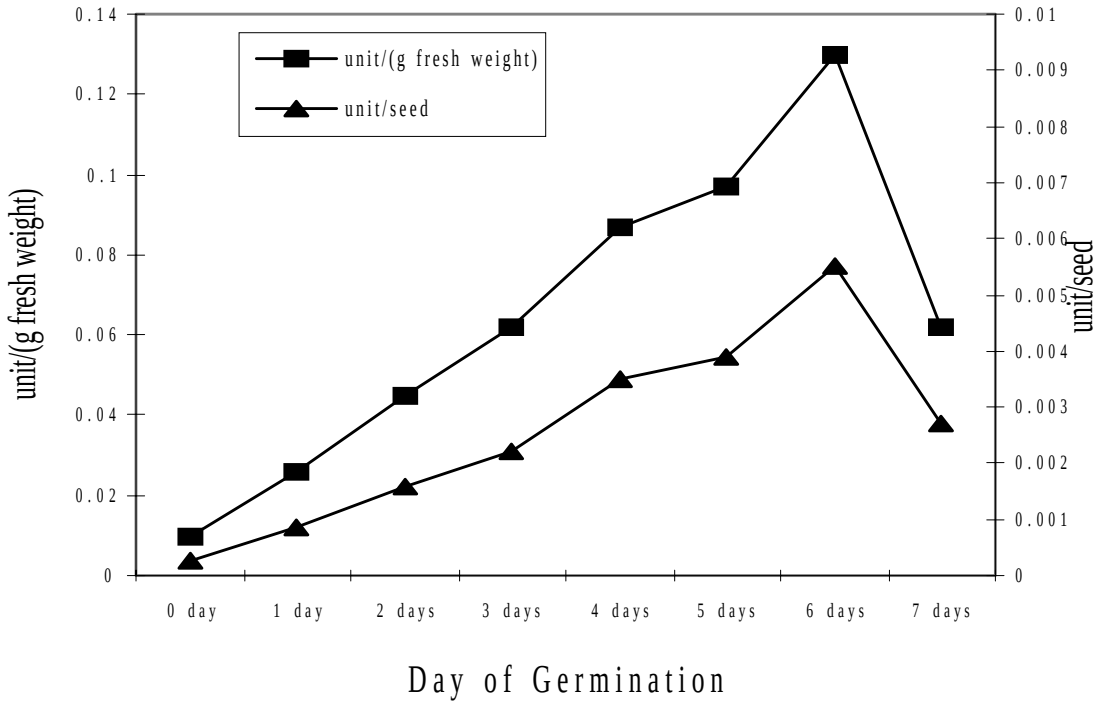


Figure 8. β-galactosidase activity in germinated rice seed. [Results are expressed as U/ (g fresh weight)and U/seed]

From the data it was apparent that β-galactosidase likely plays some role in the rice germination process. Support for this hypothesis is that the physiological definition of germination is usually the time when the radical or coleoptile (embryonic shoot) emerges from the seed coat. Loosening of cell walls to allow expansion of the cell, which is necessary for this process, is believed to require a cocktail of enzymes such as xyloglucan endotransglycosylase, β-galactosidases, polygalacturonases and expansins to undo polymers crosslinks before newly synthesized polymer are added (Thorpe et al., 1999). β-galactosidase is most likely to hydrolyze galactose residues from side chains of pectins or oligosaccharides produced from the solubilization of cell wall galactans. Increases in β-galactosidase activity have been reported in a wide range of ripening fruit, which also involves cell wall softening, putatively contributing to pectin solubilization and to flesh softening. Golden (1993) reported that β-galactosidase activity in coffee berries showed a progressive increase of more than four fold as the fruit developed from immature to ripe stage, with slight decrease in fully ripe fruit. This enzyme could hydrolyze galactan and therefore may be involved in the loss of galactose from cell wall. However, the

situation in rice appears to be less clear, since Caripta (1996) stated that a key difference between members of the Poaceae and other plants is the composition of the primary cell wall. The walls have very low amount of pectin and xyloglucan polymer, which in other genera are major cell wall polysaccharides, with the glucans playing an important role in cell wall turnover during cell elongation. So, it is unclear what are the primary substrates and functions of rice β-galactosidase during germination. To investigate this, further localization of the enzyme activity by X-gal staining was done.

2. Localization of β-galactosidase during seed germination

Japonica rice seeds at 0-7 days of germination were stained for β-galactosidase activity using X-Gal (5-bromo-4-chloro-3-indolyl-β-D-galactoside) as a substrate. β-galactosidase hydrolyses the colourless compound X-Gal into an insoluble blue product, which accumulates at the regions of the seed which have β-galactosidase activity. Figure 9 (a-h) represents the accumulate β-galactosidase activity in various tissues of the rice seed. Cross sections of rice seeds from 0-7 days germination indicated β-galactosidase is expressed and it showed strong localization to embryo and aleurone layer in the dried seed and during germination. β-galactosidase is also clearly detected in shoot tip and root tip at 5 –7 days germination (Fig. 9e-h). Similar trends were found by Muslim (1995) when β-galactosidase is detected in the rice seed staining by X-Gal, it showed strong localization to aleurone layer and surrounding tissues. These tissues may have high β-galactosidase activity because aleurone layer is the site of production of many hydrolase enzymes during seed germination.

Germination of rice seed is thought to occur primarily as rice seed structure is modified during germination, leading to a loss of tissue integrity allowing the embryo to emerge from the seed coat. Perhaps the most dramatic cell wall compositional change during germination is the loss of galactosyl residues. This loss may involve hydrolysis of galactosyl-containing polymers by β-galactosidase (Kang *et al*., 1994). In addition to the potential involvement of β-galactosidase in fruit softening and seed germination, this enzyme may be relevant in other ripening processes. Gross (1983) has been reported that Gal, the product of β-galactosidase action, increase 4-6 fold during tomato ripening.

Sekimata, *et. al* (1989) has reported the localization of β-galactosidase in radish radical tissues appears to be variable, its activity was located in chloroplasts, vacuoles, protein bodies, cell walls and intercellular fluid in which several isozyme of β-galactosidase was detected. It was thought to be involved in the degradation of such galactose containing cell wall polysaccharides as galactan-pectin polymers and xyloglucan in relation to cell growth and seed germination.

Figure 9a-h. The rice seeds at 0-7 days of germination after incubation in the staining solution. β-gal is expressed at aleurone layer (al), embryo (em), the growing shoot (sh) and root (rt) (a=0 day, b=1 day, c=2 days, d=3 days, e=4 days, f=5 days, g=6 days, h= 7 day)

3. Cloning and sequencing of β-galactosidase fragment

Degenerate primers were designed based on the most highly conserved regions of shared amino acid identity among plant β-galactosidases (Cairns *et al*., 1999). These primers were used in RT-PCR to produce a single product band of approximately 400 bp on agarose gel electrophoresis (Fig.10). The PCR product was cloned into the plasmid pGEM-T and then was transformed into competent DH5α.

The amplified fragment cloned in the recombinant plasmid was confirmed by restriction digest with *Pst* I and *Sac* II, which gave the expected 400 bp band on an agarose gel electrophoresis (Fig. 11). The sequenced is shown in Fig. 12. The peptide translation of the rice β-galactosidase cDNA fragment sequence was used to search the GenBank data base. Sequence comparison of the deduced amino acid sequence indicated it was homologous to a number of plant β-galactosidases with the closest match being asparagus β-galactosidase with 79% identity (Fig. 13). Sequencing of the PCR product indicated that it is composed of 396 bp of DNA sequence that translated to 132 amino acids. The amino acid sequence region contained many amino acid residues that were identical to other plant β-galactosidase sequences previously determined to be conserved by multiple sequence alignment. This indicated that this cDNA fragment is derived from a β-galactosidase mRNA as expected.

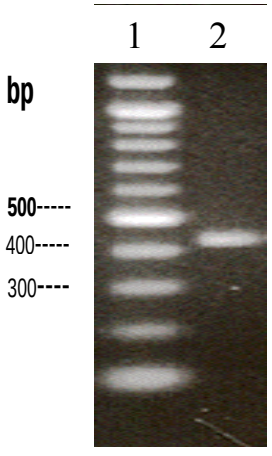


Figure 10. 1% agarose gel electrophoresis of the third PCR amplification using β-galactosidase primers

Lane 1 : 100 bp marker

Lane 2 : Product of the third PCR

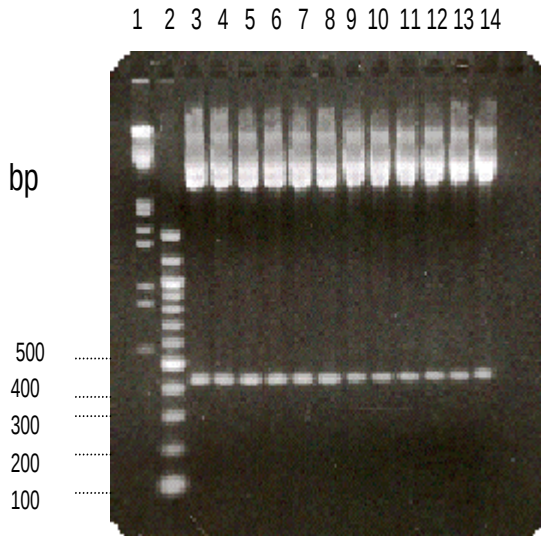


Figure 11. 1% agarose gel electrophoresis of restriction analysis of recombinant plasmids containing cDNA fragment of β-galactosidase mRNA with the *Pst* I and *Sac* II endonuclease

Lane 1 : Lambda/*Eco*R I+*Hin*d III marker

Lane 2 : 100 bp marker

Lane 3-14 : *Pst* I and *Sac* II digested recombinant plasmids from white colonies No.1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11,12 from left to right respectively.

```
         10        20        30        40        50        60        70        80        90
GGAATTACGGGGGGTTCCCGGTCTGGTTGAAGTATGTACCAGGCATCAGCTTCAGGACGGACAATGAACCTTTCAAGAATGCAATGCAGG

   N  Y  G  G  F  P  V  W  L  K  Y  V  P  G  I  S  F  R  T  D  N  E  P  F  K  N  A  M  Q

         100       110       120       130       140       150       160       170       180
GGTTCACAGAGAAAATTGTGGGCATGATGAAGAGTGAAAACCTCTTTGCTTCACAGGGCGGTCCTATTATCCTCTCTCAGATTGAGAACG

G  F  T  E  K  I  V  G  M  M  K  S  E  N  L  F  A  S  Q  G  G  P  I  I  L  S  Q  I  E  N

         190       200       210       220       230       240       250       260       270
AGTATGGGCCAGAAGGTAAAGAGTTTGGGGCTGCCGGCAAGGCATATATCAACTGGGCGGCAAAGATGGCCGTGGGATTGGACACCGGTG

E  Y  G  P  E  G  K  E  F  G  A  A  G  K  A  Y  I  N  W  A  A  K  M  A  V  G  L  D  T  G

         280       290       300       310       320       330       340       350       360
TGCCGTGGGTGATGTGCAAGGAGGATGACGCACCTGACCCAGTGATCAATGCATGCAATGGTTTGTATTGTGACACATTTTCTCCTAACA

V  P  W  V  M  C  K  E  D  D  A  P  D  P  V  I  N  A  C  N  G  L  Y  C  D  T  F  S  P  N

         370       380       390       400
AGCCTTACAAGCCCAAAATGTGGACCGAAATCTGGAC

K  P  Y  K  P  K  M  W  T  E  I  W  T
```

Figure 12. DNA sequence and deduce amino acid sequence of β-galactosidase PCR product

```
rice   NYGGFPVWLKYVPGISFRTDNEPFKNAMQGFTEKIVGMMKSENLFASQGGPIILSQIE

Aspar  NFGGFPVWLKYVPGIHFRTDNGPFKAAMGKFTEKIVSMMKAEGLYETQGGPIILSQIE
Athal  NFGGFPVWLKYVPGISFRTDNEPFKRAMKGFTERIVELMKSENLFESQGGPIILSQIE
Mango  NFGGFPVWLKYVPGIEFRTDNEPFKAAMQKFTEKIVSMMKAEKLFETQGGPIILSQIE
Apple  NYGGFPVWLKYVPGISFRTDNEPFKNAMQGFTEKIVGMMKSENLFASQGGPIILSQIE


rice   NEYGPEGKEFGAAGKAYINWAAKMAVGLDTGVPWVMCKEDDAPDPVINACNGLYCDTF

Aspar  NEYGPVEYYDGAAGKSYTNWAAKMAVGLNTGVPWVMCKQDDAPDPVINTCNGFYCDYF
Athal  NEYGRQGQLLGAEGHNYMTWAAKMAIATETGVPWVMCKEDDAPDPVINTCNGFYCDSF
Mango  NEFGPVEWEIGAPGKAYTKWAAQMAVGLDTGVPWVMCKQDDAPDPVINTCNGFYCENF
Apple  NEFGPVEWEIGAPGKAYTKWAAQMAVGLDTGVPWIMCKQEDAPDPVIDTCNGFYCENF


rice   SPNKPYKPKMWTEIWT   132 A.A.

Aspar  SPNKDNKPKMWTEAW    79% ID
Athal  APNKPYKPLIWTEAWS   78% ID
Mango  VPNQKNKPKMWTENWT   78% ID
Apple  KPNKDYKPKMWTEVW    77% ID
```

Figure 13. Multiple Sequence Alignment of β-galactosidase PCR product and deduced amino acid sequences of asparagus, arabidopsis, mango and apple β-galactosidase

4. Characterization of the genomic organization of β-galactosidase gene

A 0.4 Kb fragment derived from the cloned indica rice cDNA PCR product was used as a probe to determine the gene's copy number. Approximately 3 bands of different intensities were observed for each restriction enzyme digest (Fig. 14), suggesting that at least 3 homologous genes corresponding to β-galactosidase are present. However, characterization of gene copy number by Southern blot method is not ideal because bands from some, but not all, homologous genes seem to appear even when the blot was washed at fairly high stingency conditions. The appearance of only one band with the highest intensity in each lane of both southern blot, suggest only one gene exactly matching the PCR product found in rice, but 2 homologous genes may cross-react with the probe.

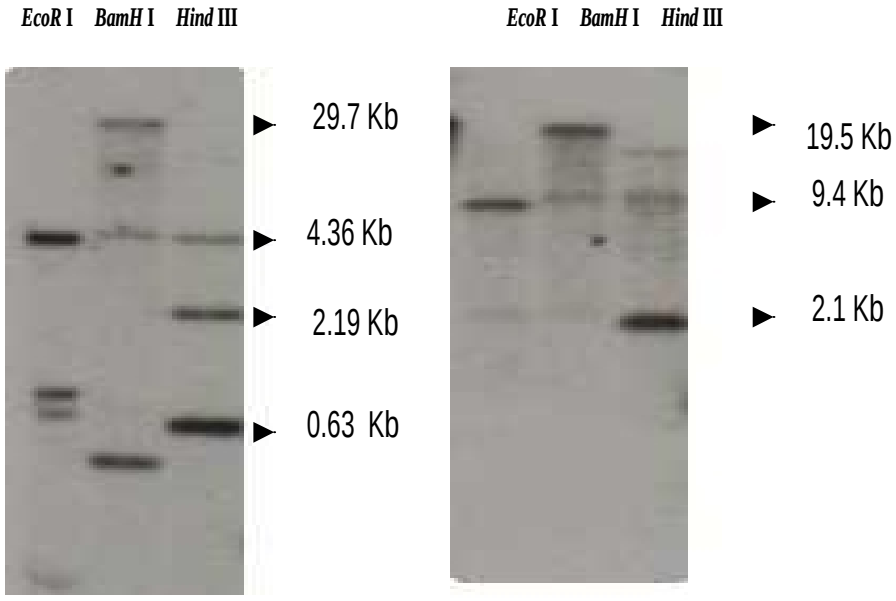


Figure 14. Autoradiograph of a gDNA gel blot. Fifteen micrograms of gDNA from a. Thai jasmine (indica) rice and b. Nippon Bare (japonica) rice digested with *Eco*R I, *Bam* HI, or *Hind* III and load in each lane.

5. Expression of β-galactosidase during seed germination

When the 0.4 Kb cDNA fragment was used as a probe in northern blot analysis of indica and japonica rice RNA from day 1, 3, and 5 of germination, a weak signal was detected in the RNA extract from five day of germination in rice shoot (Fig. 15). It is unclear why the transcript for β-galactosidase in germinating rice is only detected at 5 days germination using northern blot hybridization method. One of the most important factors is the quality of RNA that may be degraded in 1 and 3 day samples, which was difficult to obtain due to the large amounts of starchy endosperm. Another difficulty is that the probe may have hybridized to more than 1 gene. For the further work, the northern blot may be probed with a specific probe from 3' untranslated region of the cDNA, using good quality RNA on the blot.

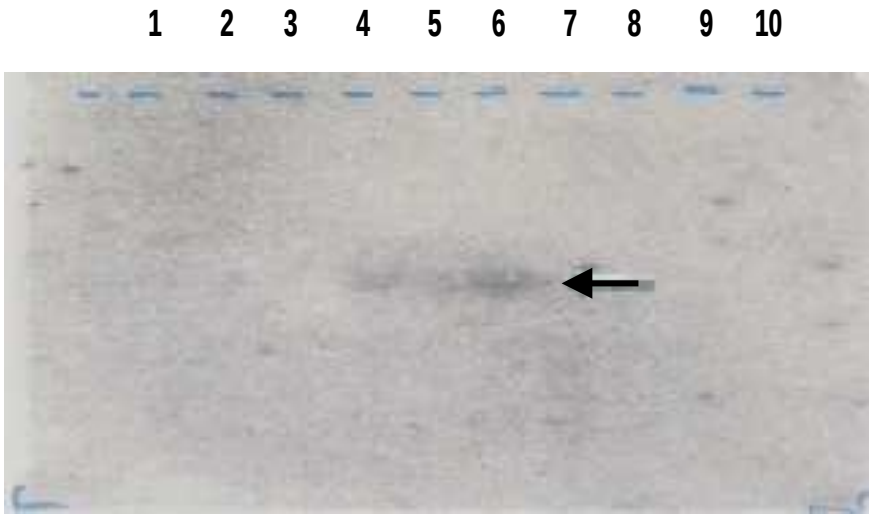


Figure 15. Detection of β-galactosidase RNA by gel-blot hybridization analysis. All lanes were prepared using 10 μg of total RNA. The blot was exposed to Kodak film at –80°C for 4 days

Lane 1: RNA isolation from indica rice seed 1 day germination
Lane 2: RNA isolation from japonica rice seed 1 day germination
Lane 3, 5: RNA isolation from indica rice shoot 3 and 5 day germination respectively
Lane 4, 6: RNA isolation from japonica rice shoot 3 and 5 day germination respectively
Lane 7, 9: RNA isolation from indica rice root 3 and 5 day germination respectively
Lane 8, 10: RNA isolation from japonica rice root 3 and 5 day germination respectively

6. Screening of the cDNA library and determination of DNA insert

Immature rice seed library was screened using the cloned PCR product as a probe. After the third screening, plaques that gave high hybridization signals were collected. Six clones from the collection were selected. Phage DNA insert from all six clones which was digested with enzyme *EcoR* I were composed of 2 restriction fragments of approximately 700 bp and 1800 bp (Fig. 16)

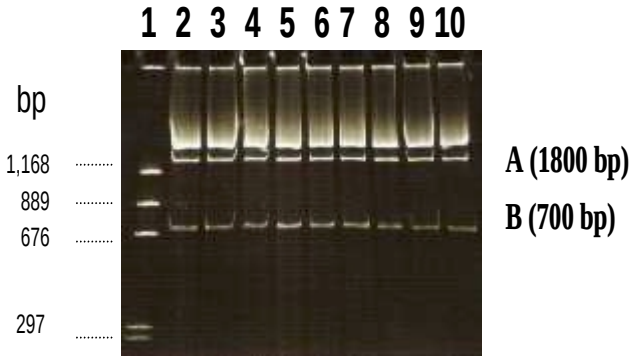


Figure 16. 8% polyacrylamide gel electrophoresis of restriction digest of recombinant phage λ gt11 DNA containing β-galactosidase cDNA with *Eco*R I.

Lane 1 : pGEMEX marker
Lane 2-4 : Phage DNA from clone I digested with the *EcoR* I
Lane 5-7 : Phage DNA from clone II digested with the *EcoR* I
Lane 8-10 : Phage DNA from clone III digested with the *EcoR* I

Large (1800 bp) and small (700 bp) DNA fragment from clone I were eluted from the gel and cloned into pUC19 vector to determined DNA sequence of β-galactosidase by using M13 forward and M13 reverse as a sequencing primers. The large fragment was also digested into smaller fragments with the enzyme *Hin*c II which composed of 3 restriction fragments of approximately 889 bp, 676 bp and 297 bp respectively (Fig. 17) but only the 676 bp DNA fragment was subcloned into pUC19 vector.

The small fragment (B) DNA sequence was also used to design the new sequencing primer, which was designated as bgalM (5'GGT CCT ATT ATC CTC TCT CAG3') in order to complete sequencing. Comparison of the sequence with the Monsanto rice genome database revealed the clone matched a gene in the OSM198 contig. The cDNA was apparently 27 nucleotide short of the full coding region at the 5' end, lacking the first 9 amino acids of a prepeptide signal sequence, as judged from comparison to related β-galactosidase sequence, the similarity of the derived peptide sequence to known prepeptide signal sequences, and the lack of any other ATG codon 5' of the first Methionine in this uninterupted reading frame (URF) of the gene.

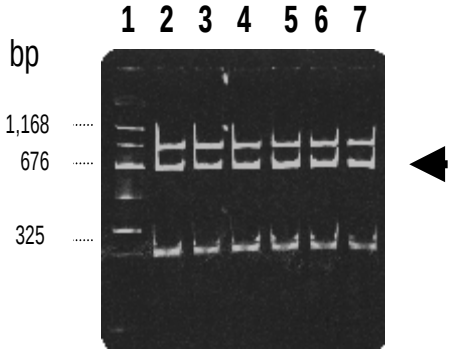


Figure 17. 8% polyacrylamide gel electrophoresis of restriction digest of recombinant plasmid containing β-galactosidase cDNA with *Hinc* II, the arrow indicate the DNA fragment which was subcloned into pUC 19

7. Amplification and sequencing of the cDNA encoding the 5' terminus of β-galactosidase cDNA library clone

Attempts at 5' RACE to generate the 5' region of the cDNA were unsuccessful, so the remaining coding sequence was amplified by RT-PCR with primers derived from the clone and the gene sequence to complete the cDNA coding region (CDS). To determine the 5' end of the mRNA β-galactosidase, the first strand cDNA was synthesized using reverse transcriptase with rbRACE 1r primer and total RNA extracted from germinating rice seeds. The first strand cDNA was amplified by PCR using Taq polymerase with the MOSM198f1 primer, derived from the apparent 5'untranslated region in the gene and the rbRACE1r primer (see Fig. 19). The PCR product was reamplified with the MOSM198f1 and rbRACE2r primers. In the second PCR reaction (Table 7), a major band about 200 bp (lane5, Fig. 18) and a large smear band were produced on 1% agarose gel analysis of the product. This specific band (200bp) was eluted and ligated with pGEM-T vector. The recombinant plasmid was isolated and sequenced.

Use of nucleotide sequences selected from the Monsanto rice genome database to synthesize a specific primer to determine the 5' end mRNA sequence using RT-PCR was found to be an effective way to construct the cDNA. Compared to 5'RACE, this method offered more specificity to the gene and took less time to perform. Using the specific primer, a PCR product of approximately 200 bp was produced but some contaminating bands were also amplified. However, the major PCR product separated quite well by using agarose gel electrophoresis, was easily eluted, and was successfully cloned and sequenced.

The sequence of the cloned PCR product was found to match the cDNA library clone. After combining the sequence of large DNA segment with the small DNA fragment and the 5' PCR product, the coding sequence of the cDNA was composed of 2,523 bp and translated to 841 amino acid residues (Fig.20). Starting from the methionine start (ATG) codon of 5' PCR product. The deduced amino acid sequence of the cDNA sequence shared significant identity with all published plant β-galactosidase amino acid sequences in the database (Fig.24). When the entire ORF of the β-galactosidase gene was compared with the best matches in the GenBank database, the shared amino acid sequence identity was 70% for

Lycopersicon esculentum (AF 154424), 69% for *Arabidopsis thaliana* (AJ270299) and 64% for *Asparagus officinalis* (X77319).

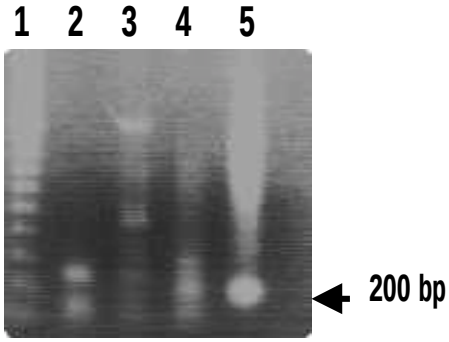


Figure 18. 1% agarose gel electrophoresis of PCR product from the 5' terminus amplified with MOSM198f1 primer and rbRACE2r primer

Lane 1 : 100 bp marker
Lane 3 : First PCR product from MOSM198f1 primer and rbRACE1r primer
Lane 5 : Second PCR product from MOSM198f1 primer and rbRACE2r primer

Interesting features of amino acid sequence from the library clone were analyzed using prediction programs available through the Expasy proteomics site (www.expasy.org). The predicted protein sequence was analyzed with several protein prediction programs indicating that the mature protein is composed of 816 amino acid residues with an expected, molecular weight = 89.75 KD and pI = 5.92 (Table 10). The SignalP program (Nielsen *et al*., 1997) predicted that the protein contained an N-terminal prepeptide signal sequence for transcript into the ER composed of 25 amino acid residues (MGRGCLAALLGGAVAVAVLVAVVHC). The prediction, by both PSORT and signalP, that the signal sequence of the library clone may target this protein outside the cell suggests that it may be involved in cell wall galactosyl modification.

Table 10. β-galactosidase cDNA nucleotide and deduced amino acid sequence data.

Clone	Length (bp)	Signal peptide	Mature protein	KD	pI	Target
Library	2,523	25	816	89.7	5.92	out
EST	2,151	20	697	76.57	7.65	ER

Figure 19. DNA sequence and deduced amino acid sequence of rice β-galactosidase fragment. Underline represent the DNA sequences of 5' RT-PCR primers

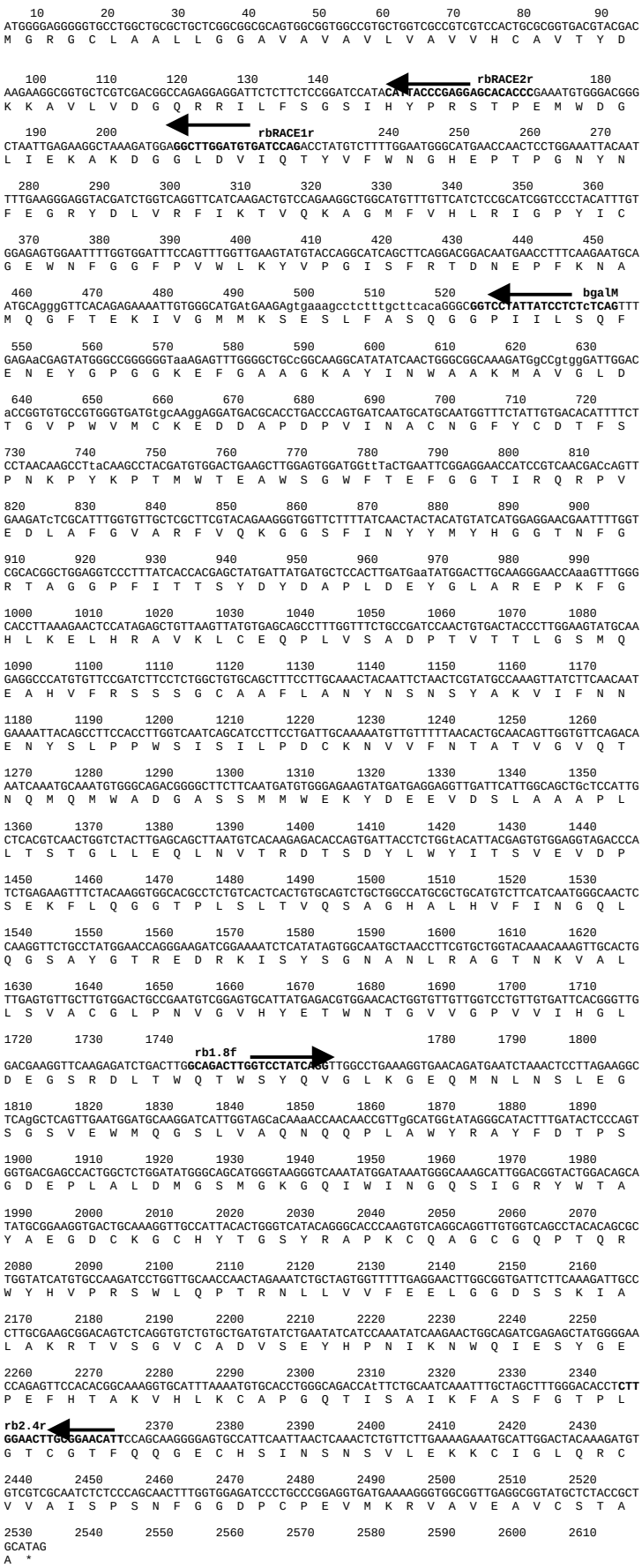


Figure 20. DNA sequence and amino acid sequence of full length β-galactosidase cDNA library clone

8. Sequencing of cDNA fragment encoding the 5'end and 3'end mRNA of β-Galactosidase enzyme from the EST clone.

By screening the dbEST database by protein sequence similarity (tBLASTN) using the closely related full length asparagus β-galactosidase several ESTs corresponding to putative β-galactosidase mRNA were identified. The rice clone named E1405 corresponding to a β-galactosidase homologue cDNA from panicle at flowering stage, was requested from the Japanese Ministry of Agriculture, Forestry and Fisheries (MAFF) and sequenced using an automated DNA sequencer with T7 and T3 primer complementary to the pBluescript II SK+/- plasmid containing the clone. The resulting DNA sequence was rearranged and translated to amino acids. The sequence appeared to be that of full length cDNA. The DNA sequence which was sequenced using the T7 sequencing primer was composed of 192 amino acid residue and contained the first methionine as judged by comparison with other plant β-galactosidases. On the other hand the DNA sequence which was sequenced using T3 as a sequencing primer was composed of a stop codon (TGA), 495 untranslated nucleotides and a poly A tails (15 A residues).

In order to determine the remaining DNA sequence of this clone, a restriction digest was performed using the enzyme *Hinc* II to cut the recombinant plasmid clone E1405 in pBluescript II SK+/- plasmid. The digested DNA was analyzed on an 8% polyacrylamide gel (Fig. 21). Four fragments with sizes of approximately 4130, 670, 460 and 250 bp, designated as EH1, EH2, EH3 and EH4, were eluted from the gel, cloned into pUC19 vector and sequenced using M13 forward and M13 reverse sequencing primers. After these DNA sequences were combined the DNA sequence was translated to amino acids. The full length coding sequence of the cDNA was composed of 2,151 base pair and was translated to 717 amino acid residue (Fig.23).

The SignalP program predicted that the protein contains a prepeptide signal sequence composed of 20 amino acid residues targetting it to the secretory pathway. The protein prediction programs at the Expasy site predicted that the mature protein is composed of 697 amino acid residues, molecular weight = 76.57 KD, pI = 7.65. Comparison of the deduced amino acid sequences of the EST clone and the cDNA library clone by protein sequence alignment showed they share 62% identity. The β-galactosidases which best matched this clone were *Asparagus officinalis* (Acession no. P45582) with 71.6% identity, *Lycopersicon esculentum* (Acession no.

AJ012796) with 66.6% identity, and *Fragaria ananassa* (Acession no. AJ278703) with 66.4% identity.

Although PSORT predicted that EST protein is targeted to the endoplasmic reticulum membrane, the SignalP predicted that EST protein are secreated. Therefore, it is conceivable that this enzyme is involved in cell wall galactosyl modification as well.

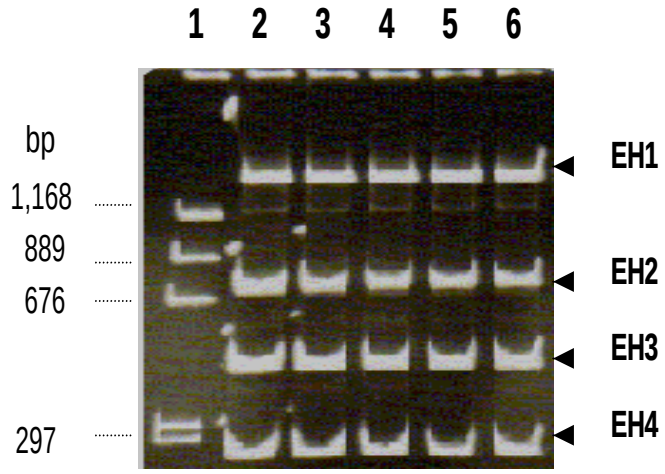


Figure 21. 8% polycrylamide gel electrophoresis of restriction analysis of recombinant plasmid containing β-Galactosidase cDNA with *Hinc* II enzyme

Lane 1 : Lambda/ *Eco*R I+*Hin*d III marker

Lane 2-6 : digested recombinant plasmid from EST clone

The open reading frames (ORFs) of two β-Galactosidase clones, which have a significant level of shared amino acid sequence identity to each other and other published plant β-galactosidases, were aligned (Fig. 24) Due to a number of amino acid deletions or insertions throughout their sequences, the deduced amino acid sequence of the cDNA library ORF is closet to previously described cDNA (accession no. AF 154424) from *Lycopersicon esculentum* (Smith and Gross, 2000). The ORF of EST most closely matches the cDNA clone (accession no P45582) isolated from *Asparagus officinalis* (King and Davies, 1995).

The large size of the protein in the cDNA library clone compared to the EST clone protein was due primarily to an addition of approximately 100 amino acids at its carboxyl terminus. The addition at the carboxyl terminus was remarkable due to the presence of a number of highly conserved residues in plant β-galactosidases, and in particular, containing this region eight conserved cysteines (Fig. 24, shown in bold). It is likely that covalent disulphide bonds between these cysteine residues of the cDNA library clone C-terminus give the protein greater stability. This high concentration of disulfide bonds is similar to many binding domains. In fact, BLAST searches of GenBank database, indicated the amino acid sequence of the library clone contained Galactose binding lectin domain at its C-terminus (Fig. 22), homologue to sea urchin galactose-binding lectin (Altschul, 1997). This suggests the C-terminal addition may allow efficient binding with Gal a galactose-containing substrates, which may be a cell wall component. Therefore, the difference in lengths of the cDNA library and EST clones may have an affect on binding to the cell wall.

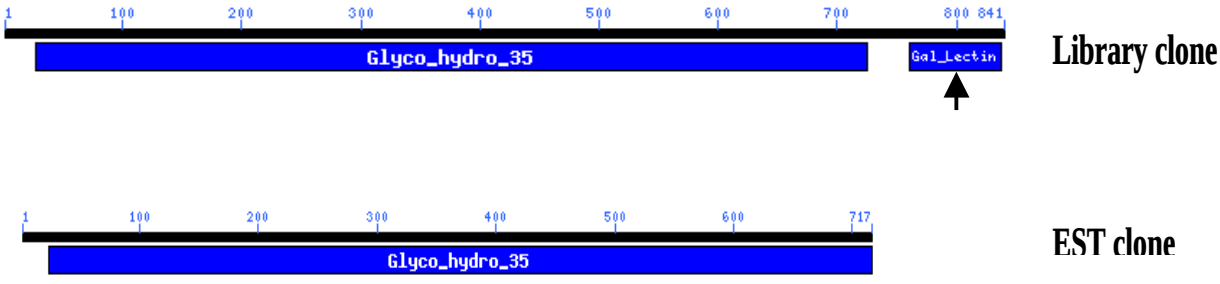


Figure 22. Schematic representation of Galactose binding lectin domain at C-terminus of β-Galactosidase . (the arrow indicated the lectin binding domain)

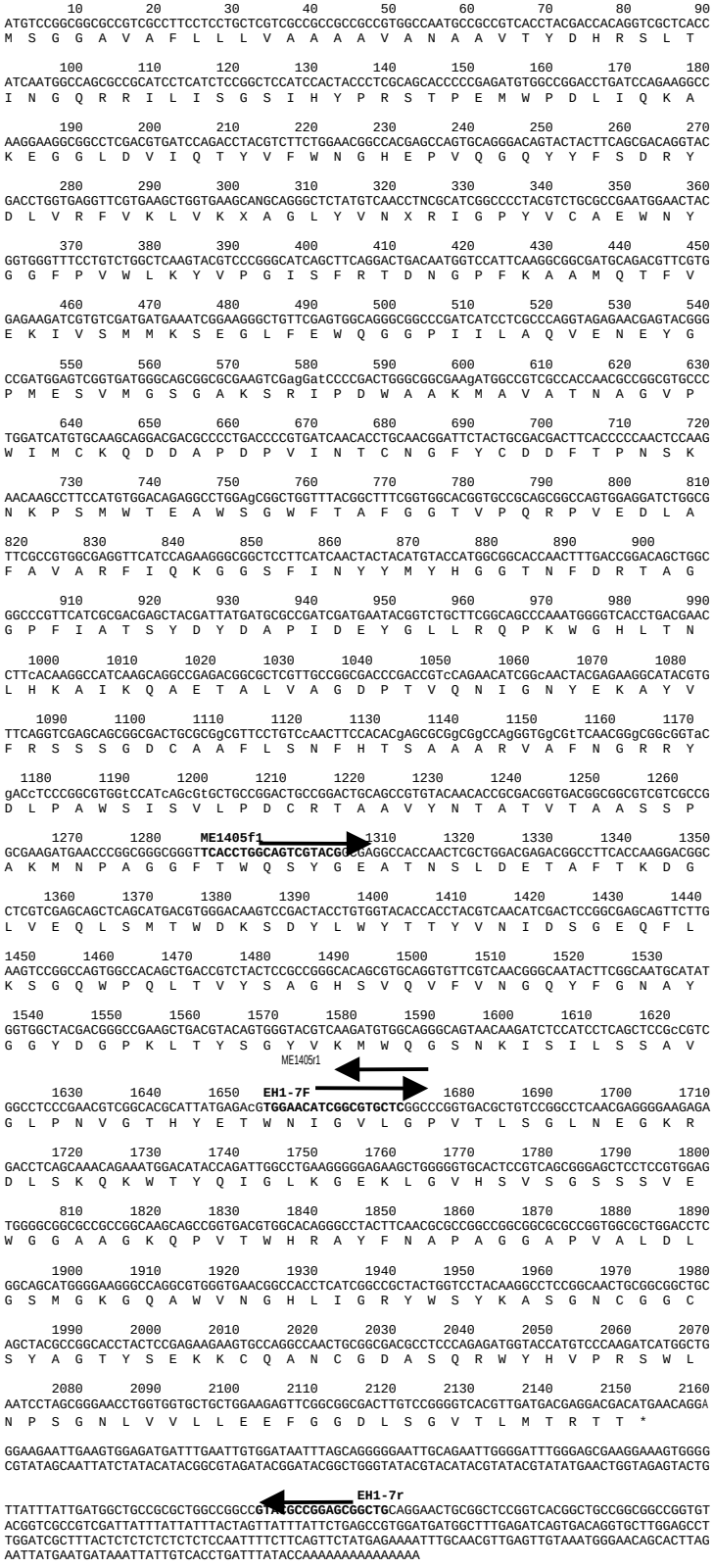


Figure 23. DNA sequence and deduced amino acid sequences of full length β-galactosidase from EST clone

```
lib             ---MGRGCLAALLGGAVAVAVLVA---VVHCAVTYDKKAVLVDGQRRILFSGSIHYPRST
ara             MREMGTGDSASRLILWFCLGFLILGVGFVQCGVTYDRKALLINGQRRILFSGSIHYPRST
tmt             ---MEVNSLQKWVLLWCIVLFISS--GLVHCDVTYDRKAIVINGQRRLLFSGSIHYPRST
cic             ---METNSVSKLLSLFFFLFFVCS--QLIHCSVTYDRKAIIINGQRRILISGSIHYPRST
asp             ---MALKLVLMLMVALLAAVWSPP---AVTASVTYDHKSVIINGQRRILISGSIHYPRST
est             ---MSGGAVAFLLLVAAAAVAN--------AAVTYDHRSLTINGQRRILISGSIHYPRST
                   *        :                . ****:::: ::****:*:**********

lib             PEMWDGLIEKAKDGGLDVIQTYVFWNGHEPTPGNYNFEGRYDLVRFIKTVQKAGMFVHLR
ara             PDMWEDLIQKAKDGGIDVIETYVFWNLHEPSPGKYDFEGRNDLVRFVKTIHKAGLYAHLR
tmt             PEMWEDLINKAKEGGLDVVETYVFWNVHEPSPGNYNFEGRYDLVRFVKTIQKAGLYAHLR
cic             PEMWEDLIQKAKVGGLDVIDTYVFWNVHEPSPSNYNFEGRYDLVRFIKTVQKVGLYVHLR
asp             PEMWPDLIQKAKDGGLDVIQTYVFWNGHEPSPGQYYFGGRYDLVRFLKLVKQAGLYAHLR
est             PEMWPDLIQKAKEGGLDVIQTYVFWNGHEPVQGQYYFSDRYDLVRFVKLVKXAGLYVNLR
                *:** .**.*** **:*.:.****** ***  . * * .* *****:* :: .*::.:**

lib             IGPYICGEWNFGGFPVWLKYVPGISFRTDNEPFKNAMQGFTEKIVGMMKSESLFASQGGP
ara             IGPYVCAEWNFGGFPVWLKYVPGISFRTDNEPFKRAMKGFTERIVELMKSENLFESQGGP
tmt             IGPYVCAEWNFGGFPVWLKYVPGISFRADNEPFKNAMKGYAEKIVNLMKSHNLFESQGGP
cic             IGPYVCAEWNFGGFPVWLKYVPGISFRTDNGPFKAAMQGFTQKIVQMMKNEKLFQSQGGP
asp             IGPYVCAEWNFGGFPVWLKYVPGIHFRTDNGPFKAAMGKFTEKIVSMMKAEGLYETQGGP
est             IGPYVCAEWNYGGFPVWLKYVPGISFRTDNGPFKAAMQTFVEKIVSMMKSEGLFEWQGGP
                ****:*.***:************* **:** *** **  :.::** :** . *:  ****

lib             IILSQFENEYGPGGKEFGAAGKA-YINWAAKMAVGLDTGVPWVMCKEDDAPDPVINACNG
ara             IILSQIENEYGRQGQLLGAEGHN-YMTWAAKMAIATETGVPWVMCKEDDAPDPVINTCNG
tmt             IILSQIENEYGPQAKVLGAPGHQ-YSTWAANMAVGLDTGVPWVMCKEEDAPDPVINTCNG
cic             IILSQIENEYGPQGRALGAVGHA-YSNWAAKMAVGLGTGVPWVMCKEDDAPDPVINSCNG
asp             IILSQIENEYGPVEYYDGAAGKS-YTNWAAKMAVGLNTGVPWVMCKQDDAPDPVINTCNG
est             IILAQVENEYGPMESVMGSGAKSRIPDWAAKMAVATNAGVPWIMCKQDDAPDPVINTCNG
                ***:*.*****      *. .      ***:**:.  :****:***::********:***

lib             FYCDTFSPNKPYKPTMWTEAWSGWFTEFGGTIRQRPVEDLAFGVARFVQKGGSFINYYMY
ara             FYCDSFAPNKPYKPLIWTEAWSGWFTEFGGPMHHRPVQDLAFGVARFIQKGGSFVNYYMY
tmt             FYCDNFFPNKPYKPAIWTEAWSGWFSEFGGPLHQRPVQDLAFAVAQFIQRGGSFVNYYMY
cic             FYCDDFSPNKPYKPKLWTESWSGWFSEFGGPVPQRPAQDLAFAVARFIQKGGSFFNYYMY
asp             FYCDYFSPNKDNKPKMWTEAWTGWFTGFGGAVPQRPAEDMAFAVARFIQKGGSFINYYMY
est             FYCDDFTPNSKNKPSMWTEAWSGWFTAFGGTVPQRPVEDLAFAVARFIQKGGSFINYYMY
                **** * **.  ** :***:*:***: ***.: :**.:*:**.**:*:*:***:.*****

lib             HGGTNFGRTAGGPFITTSYDYDAPLDEYGLAREPKFGHLKELHRAVKLCEQPLVSADPTV
ara             HGGTNFGRTAGGPFVTTSYDYDAPIDEYGLIRQPKYGHLKELHRAIKMCEKALVSADPVV
tmt             HGGTNFGRTAGGPFITTSYDYDAPIDEYGLIRQPKYGHLKELHRAVKMCEKSIVSADPAI
cic             HGGTNFGRSAGGPFITTSYDYDAPIDEYGLLREPKYGHLKDLHKAIKQCEHALVSSDPTV
asp             HGGTNFGRTAGGPFISTSYDYDAPIDEYGLLRQPKWGHLRDLHKAIKLCEPALVSGEPTI
est             HGGTNFDRTAGGPFIATSYDYDAPIDEYGLLRQPKWGHLTNLHKAIKQAETALVAGDPTV
                ******.*:*****::********:***** *:**:*** :**:*:* .* .:*:.:*.:

lib             TTLGSMQEAHVFRSSSG-CAAFLANYNSNSYAKVIFNNENYSLPPWSISILPDCKNVVFN
ara             TSIGNKQQAHVYSAESGDCSAFLANYDTESAARVLFNNVHYNLPPWSISILPDCRNAVFN
tmt             TSLGNLQQAYVYSSETGGCAAFLSNNDWKSAARVMFNNMHYNLPPWSISILPDCRNVVFN
cic             TSLGAYEQAHVFSSGTQTCAAFLANYHSNSAARVTFNNRHYDLPPWSISILPDCKTDVFN
asp             TSLGQNQESYVYRSKSS-CAAFLANFNSRYYATVTFNGMHYNLPPWSVSILPDCKTTVFN
est             QNIGNYEKAYVFRSSSGDCAAFLSNFHTSAAARVAFNGRRYDLPAWSISVLPDCRTAVYN
                 .:*  ::::*: : .  *:***:* .    * * **. .*.**.**:*:****:. *:*

lib             TATVGVQTNQMQMWADGASSMMWEKYDEEVDSLAAAPLLTSTGLLEQLNVTRDTSDYLWY
ara             TAKVGVQTSQMEMLPTDTKNFQWESYLEDLSSLDDSSTFTTHGLLEQINVTRDTSDYLWY
tmt             TAKVGVQTSKMEMLPTNSEMLSWETYSEDISALDDSSSIRSFGLLEQINVTRDTSDYLWY
cic             TARVRFQNSKIQMLPSNSKLLSWETYDEDVSSLAESSRITASGLLEQINATRDTSDYLWY
asp             TARVGAQTTTMKMQYLGG--FSWKAYTEDTDALNDN-TFTKDGLVEQLSTTWDRSDYLWY
est             TATVTAASSPAKMNPAGG--FTWQSYGEATNSLDET-AFTKDGLVEQLSMTWDKSDYLWY
                ** *   ..  :*   .   : *: * *  .:*     :   **:**:. * * ******
```

Figure 24. Multiple Sequence Alignment of β-Gal cDNA library deduced amino acid sequence (lib), the deduced amino acid sequence *of Arabidopsis thaliana* (ara), *Lycopersicon esculentum* (tmt), *Cicer arietum* (cic), and rice EST β-Gal sequences (est). Identical amino acid are marked by asterisks and conserved substitutions are represented by dots.

```
lib             ITSVE-VDPSEKFLQGGTPLSLTVQSAGHALHVFINGQLQGSAYGTREDRKISYSGNANL
ara             MTSVD-IGDSESFLHGGELPTLIIQSTGHAVHIFVNGQLSGSAFGTRQNRRFTYQGKINL
tmt             ITSVD-IGSTESFLHGGELPTLIVETTGHAMHVFINGQLSGSAFGTRKNRRFVFKGKVNL
cic             ITSVD-ISPSESFLRGGNKPSISVHSSGDAVHVFINGKFSGSAFGTREQRSCTFNGPINL
asp             TTYVD-IAKNEEFLKTGKYPYLTVMSAGHAVHVFINGQLSGTAYGSLDNPKLTYSGSAKL
est             TTYVGNIDSGEQFLKSGQWPALTVYSAGHSVQVEVNGQYFGNAYGGYDGPKLTYSGYVKM
                 * *  :   *.**: *    : : ::*.:::: :**:  *.*:*  .     :.*  ::

lib             RAGTNKVALLSVACGLPNVGVHYETWNTGVVGPVVIHGLDEGSRDLTWQTWSY--QVGLK
ara             HSGTNRIALLSVAVGLPNVGGHFESWNTGILGPVALHGLSQGKMDLSWQKWTY--QVGLK
tmt             RAGSNRIALLSVAVGLPNIGGHFETWSTGVLGPVAIQGLDHGKWDLSWAKWTY--QVGLK
cic             HAGTNKIALLSVAVGLPNGGIHFESWKTGITGPILLHGLDHGQKDLTWQKWSY--QVGLK
asp             WAGSNKISILSVSVGLPNVGNHFETWNTGVLGPVTLTGLNEGKRDLSLQKWTY--QIGLH
est             WQGSNKISILSSAVGLPNVGTHYETWNIGVLGPVTLSGLNEGKRDLSKQKWTY--QIGLK
                  *:*::::** : **** * *:*:*. *: * : : **..*. **:  .*:*  *:**:

lib             GEQMNLNSLEGSGSVEWMQGSLVAQNQQPLAWYRAYFDTPSGDEPLALDMGSMGKGQIWI
ara             GEAMNLAFPTNTPSIGWMDASLTVQKPQPLTWHKTYFDAPEGNEPLALDMEGMGKGQIWV
tmt             GEAMNLVSTNGISAVDWMQGSLIAQKQQPLTWHKAYFNTPEGDEPLALDMSSMGKGQVWI
cic             GEAMNLVSPNGVSSVDWVRESLASQNQPQLKWHKAYFNAPDGNEALALDMSGMGKGQVWI
asp             GETLSLHSLTGSSNVEWGEAS----QKQPLTWYKTFFNAPPGNEPLALDMNTMGKGQIWI
est             GEKLGVHSVSGSSSVEWGGAA----GKQPVTWHRAYFNAPAGGAPVALDLGSMGKGQAWV
                ** :.:    .   : *   :        : *::::*::* *  .:***:  ***** *:

lib             NGQSIGRYW-TAYAEGDCKGCHYTGSYRAPKCQAGCGQPTQRWYHVPRSWLQPTRNLLVV
ara             NGESIGRYW-TAFATGDCSHCSYTGTYKPNKCQTGCGQPTQRWYHVPRAWLKPSQNLLVI
tmt             NGQSIGRYW-TAYATGDCNGCQYSGVFRPPKCQLGCGEPTQKWYHVPRSWLKPTQNLLVL
cic             NGQSIGRYW-LVYAKGNCNSCNYAGTYRQAKCQLGCGQPTQRWYHVPRSWLKPTNNLMVV
asp             NGQSIGRYWPAYKASGSCGSCDYRGTYNEKKCLSNCGEASQRWYHVPRSWLIPTGNFLVV
est             NGHLIGRYW-SYKASGNCGGCSYAGTYSEKKCQANCGDASQRWYHVPRSWLNPSGNLVVL
                **. *****    * * *  * * * :   **  .**:.:*:******:** *: *::*:

lib             FEELGGDSSKIALAKRTVSGV<b>C</b>ADVSEYHPNIKNWQIESYGNPE-FHTAKVHLK<b>C</b>APGQT
ara             FEELGGNPSTVSLVKRSVSGV<b>C</b>AEVSEYHPNIKNWQIESYGKGQTFHRPKVHLK<b>C</b>SPGQA
tmt             FEELGGDPTRISLVKRSVTNV<b>C</b>SNVAEYHPNIKNWQIENYGKTEEFHLPKVRIH<b>C</b>APGQS
cic             FEELGGNPWKISLVKRTIH--TP--ASSEPNLRTNTTQV--------------------
asp             LEEWGGDPTGISMVKRSVASV<b>C</b>AEVEELQPTMDNWRTKAYGRPK------VHLS<b>C</b>DPGQK
est             LEEFGGDLSGVTLMTRTT-------------------------------------------
                :** **:   ::: .*

lib             ISAIKFASFGTPLGT<b>C</b>GTFQQGE<b>C</b>HSINSNSVLEKK-----<b>C</b>IGLQR<b>C</b>VVAISPSNFGGD
ara             IASIKFASFGTPLGT<b>C</b>GSYQQGE<b>C</b>HAATSYAILERK-----<b>C</b>VGKAR<b>C</b>AVTISNSNFGKD
tmt             ISSIKFASFGTPLGT<b>C</b>GSFKQGT<b>C</b>HAPDSHAVVEKK-----<b>C</b>LGRQT<b>C</b>AVTISNSNFGED
cic             ------------------------------------------------------------
asp             MSKIKFASFGTPQGT<b>C</b>GSFSEGS<b>C</b>HAHKSYDAFEQEGLMQN<b>C</b>VGQEF<b>C</b>SVNVAPEVFGGD
est             ------------------------------------------------------------


lib             P<b>C</b>PEVMKRVAVEAV<b>C</b>STAA-------
ara             P<b>C</b>PNVLKRLTVEAV<b>C</b>APETSVSTWRP
tmt             P<b>C</b>PNVLKRLSVEAH<b>C</b>TPTQN------
cic             -------------------------
asp             P<b>C</b>PGTMKKLAVEAI<b>C</b>E----------
est             -------------------------
```

Figure. 24 (con't)

CHAPTER 4
CONCLUSIONS

1. β-galactosidase enzyme activity increased between 0-6 days of rice germination. The activity reached a maximum at day 6 and then decreased on day 7.
2. β-galactosidase was found in rice seed and seedlings during 0-7 days germination by activity staining using X-Gal as a substrate. It showed strong localization to aleurone layer and embryo. After germination β-galactosidase was found throughout the growing shoot and root tip.
3. Characterization of the gene copy number in both of japonica and indica rice by Southern blot analysis revealed approximately 3 highly related β-galactosidase genes, which hybridized to the cDNA fragment are present, but only 1 is likely to encode the same mRNA.
4. A weak signal was detected in RNA extracted from five day germination in rice shoot by northern blot analysis.
5. The original clone full coding sequence was completed by cDNA library screening and PCR using sequences derived from the corresponding gene sequence from the Monsanto rice genome database. The full length coding sequence of β-galactosidase transcript is composed of 2,523 bases of DNA, which translated to a 841 amino acid sequence including a signal peptide of about 25 residues and a 816 amino acid mature protein. It also contained 617 bp 3'UTR.
6. The amino acid sequence of the clone β-galactosidase cDNA library contains a galactose binding lectin domain at C-terminus as indicated by the BLAST search, thus suggesting a function in binding to the cell wall
7. Another related β-galactosidase clone from panicle at flowering stage (EST clone) had a full length coding sequence composed of 2,151 base pairs DNA sequence, which translated to a 717 amino acid sequence including a signal peptide of about 20 residues and a 697 amino acid mature protein. The cDNA coding also contained 149 bp 5'UTR and 495 bp 3'UTR.
8. The data obtained from this study characterized the cDNA coding for 2 rice β-galactosidases. However, the precise role of the β-galactosidase in rice plant can only be speculated. To

evaluate the role of the gene in rice, expression of gene in *E.coli* or yeast expression system to produce enough enzyme to characterize the β-galactosidase activity will be necessary. Further study about the protein purified from rice should also be done further determine the complex structure eg. glycosylation, localization, etc.

REFERENCES

สุมาลี ตั้งประดับกุล. (2523). **คู่มือปฏิบัติการทางพันธุวิศวกรรม1 เรื่องการขยายยีนและการตัดต่อยีนจากโครโมโซม.** ภาควิชาเคมี คณะวิทยาศาสตร์และโครงการอณูพันธุศาสตร์-พันธุวิศวกรรมศาสตร์ มหาวิทยาลัยมหิดล.

Altschul, S. F., Boguski, M. S., Gish, W., and Wootton, J. C. (1994). Issues in searching molecular sequences databases. **Nature Genetics**. 6: 119-129.

Altschul, S. F., Thomas, L. M., Alejandro, A. S., Jinghui, Z., Zheng, Z., Webb, M., and David, J. L. (1997). Gapped BLAST and PSI-BLAST: a new generation of protein database search program. **Nucleic Acids Res.** 25: 3389-3402.

Cairns, J. K., Chantarangsee, M., Chaiwangrad, S., and Phawong, J. (1999). Primary Structure Based Screening for Glycosyl Hydrolases in Thai Plants.**Thai J. Biotechnol**. 1: 20-30.

Caripta, N. C. (1996). Structure and biogenesis of the cell walls of grasses. **Annu Rev Plant Physiol Plant Mol Bio.** 47: 445-476.

Chang, T. T. (1976). The original, evolution, cultivation, dissemination and diversification of Asian and African rices. **Euphytica.** 25: 435-444.

Chomczynski, P., and Sacchi, N. (1987). Single step method of RNA isolation by acid guanidinium thiocyanate-phenol-chloroform extraction. **Anal. Biochem**. 162: 150-159.

Christou, P. (1994). **Rice Biotechnology and Genetic Engineering**. Technomic Publishing Company . Inc.

Cowell, I., and Austin, C. A. (1997). **cDNA Library Protocol**. Humana Press Inc. Totowa, NewJersy. 269 P.

Clendennen, S. K., and May, G. D. (1997). Differential gene expression in ripening banana fruit. **Plant Physiol**. 115: 463-469.

Doyle, J. J., and Doyle, J. L. (1987). A rapid DNA isolation procedure for small quantities of fresh leaf tissue. **Phytochem Bull**. 19: 11-15.

Feng, D. F., and Doolittle, R. F. (1996). Progressive alignment of amino acid sequences and construction of phylogenetic trees from them. **Meth. Enzymol**. 226: 368.

Gurdev, S. K., and Gary, H. Y. (1991). **Rice Biotechnology**. International Rice Research Institute, UK. PP. 1-2.

Golden, K. D., John, M. A., and Kean, E. A. (1993). β-galactosidase from *Cofee arabica* and its role in fruit ripening. **Phytochemistry**. 34(2): 355-360.

Kang, I. K., Suh, S. G., Gross, K. C., and Byun, J. K. (1994). N-Terminal amino acid sequence of persimmon fruit β-galactosidase. **Plant Pysiol**. 105: 975-979.

Kellogg, E. A. (2001). Evolutionary History of Grasses. **Plant Physiol**. 125: 1198-1205.

Khush, G. S. (1997). Origin, dispersal, cultivation and variation of rice. **Plant Molecular Biology**. 35: 25-34.

Kim, J. G. and Solomos, T. (1991). Galactose metabolism and ethylene production during development and ripening of tomato fruit. **Postharv Biol Technol.** 1: 67-80.

King, G. A. and Davies, K. M. (1995). Cloning of a harvest-induced β-galactosidase from tips of harvest asparagus spear. **Plant Physiol.** 108: 419-420.

Kohorn, B. D. (2000). Plasma membrane-cell wall contacts. **Plant Physiol**. 124: 31-38.

Kosugi, S., Ohashi, Y., Nakajima, K., and arai, Y. (1996). An improved assay for β-glucuronidase in transformed cells: Methanol almost completely suppressed a putative endogenous β-glucuronidase activity. **Plant Science**. 70: 133-140.

Lucotte, G., and Baneyx, F. (1993). **Introduction to Molecular cloning Technique**. USA. New York. VCH Publishers Inc.

Lucotte, G., and Baneyx, F. (1993). **Introduction to Molecular cloning Technique**. USA. New York. VCH Publishers Inc. quoted in Young, R. A., and Davies, R. W. (1983). **Science**. 222: 778.

Mizuno, K., Kimura, K., Kawasaki, T., Shimada, H., and Baba, T. (1992). Starch branching enzymes from immature rice seeds. **J. Biochem**. 112: 643-651.

Muslim, C. (1995). **Isolation and Characterization of Rice (*Oryza sativa* L.) β-glucosidase**. Ph.D. Dissertation, Virginia Polytechnic Institute and State University, Blackburg, Virginia.

Nanba, E., and Suzuki, K. (1990). Molecular cloning of mouse acid β-galactosidase cDNA: sequence, expression of catalytic activity and comparison with the human enzyme. **Biochemical and Biophysical Research Communications.** 173: 141-148.

Nielsen, H., Engelbrecht, J., Brunak, S., and Heijne, G. (1997). Identification of prokaryotic signal peptide and prediction of their cleavage site. **Protein Engineering**. 10: 1-6.

Oshima, A., Tsuji, A., Nagao, Y., Sakuraba, H., and Suzuki, Y. (1988). Cloning, sequencing, and expression of cDNA from human β-galactosidase. **Biochemical and Biophysical Research Communications.** 157: 238-244.

Palmiano, P. E., and Juliano, O. B. (1973). Change in the activity of some hydrolases, peroxidase, and catalase in the rice seed during germination. **Plant Physiol**. 52: 274-277.

Pluemsab, et al. (1998). Isolation and screening of β-galactosidase producing microorganisms for oligosaccharide production. [Abstract]. **Biotechnology for a self-sufficient economy**. 103 P.

Ross, G. S., Wegrzynt, T., MacRae, E. A., and Redgwell, R. J. (1994). Apple β-galactosidase: Activity against Cell Wall Polysaccharides and Characterrization of related cDNA clone. **Plant Physiol**. 106: 521-528.

Sambrook, J., Fritsch, E. F., and Maniatis, T. (1989). **Molecular cloning: a laboratory manual**, 2 nd ed. Cold Spring Harbor Laboratory Press, New York.

Sasaki, T., Yano, M., Kurata, N., and Yamamoto, K. (1996). The Japanese rice genome research program. **Genome Res.** 6: 661-666.

Seymour, G. B., Tatlor, J. E., and Tucker, G. A. (1993). **Biochemistry of Fruit Ripening.** Chapman & Hall, London. PP. 418-422.

Shaikh, S. A., Khire, J. M., and Khan, M. I. (1997). Production of β-galactosidase from thermophillic fungus *Rhizomucor* sp. **Journal of Industrial Microbiology&Biotecnology.** 19: 239-245.

Shimamoto, K. (1999). **Molecular Biology of Rice**. PP. 3-8. Tokyo springer, Tokyo.

Smith, D. L., and Gross, K. C. (2000). A family of at least seven β-galactosidase gene is expressed during tomato fruit Development. **Plant Physiol**. 123: 1173-1183.

Smith, L. D., Starrett, A. D., and Gross, C. K. (1998). A gene coding for tomato fruit β-galactosidase II is expressed during fruit ripening. **Plant Physiol.** 117: 417-423.

Suginta, W., and Savasti, J. (1995). Purification and properties of β-galactosidase from *Herbicus sabdaiffa* L. var. *Altissima*. **J. Sci. Soc. Thailand**. 21: 183-196.

Surarit, R. et al. (1995). Screening of Glycosylhydrolase enzymes in Thai plant seeds for potential use in oligosaccharide synthesis. **J. Sci. Soc. Thailand**. 21: 293-303.

Thorpe, M. R., Mac Rae, E. A., Minchin, P. E. H., and Edwards, C. M. (1999). Galactose stimulation of carbon import into roots is confined to the Poaceae. **Journal of Experimental Botany**. 50(339): 161-1618.

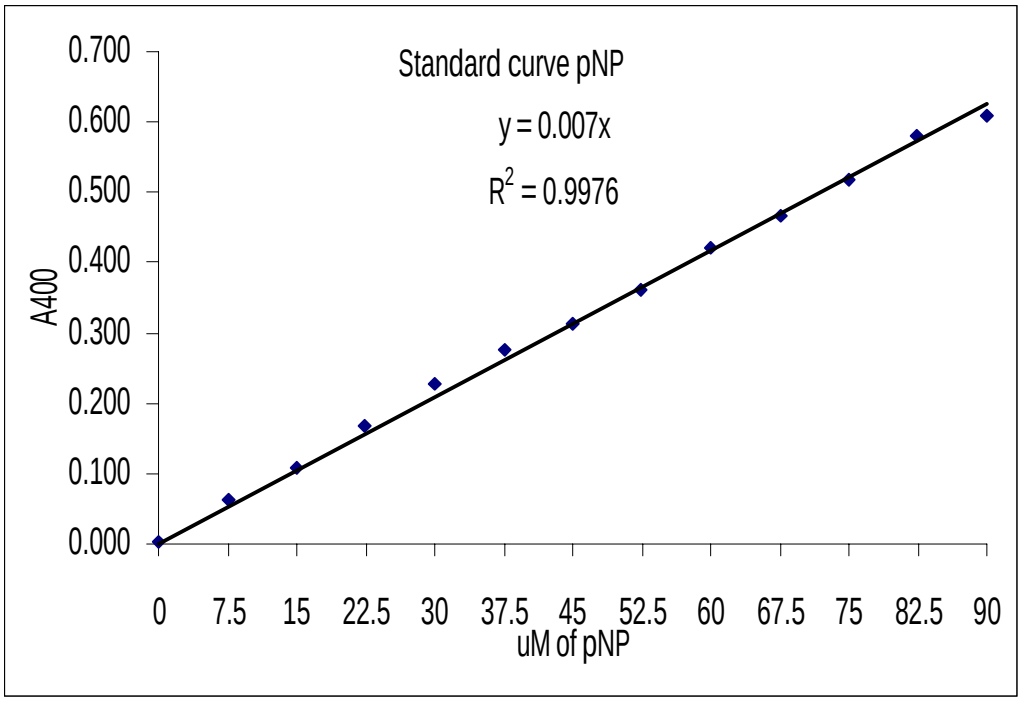
Whitaker, J. R. (1994). **Principle of Enzymology for the food Science**. 2 nd ed. New York: Marcell Dekker.

Yuan, Q. et al. (2001). Rice bioinformatics. Analysis of rice sequence data and leveraging the data to other plant species. **Plant Physiol**. 125: 1166-1174.

APPENDIX

1. Determination of β-galactosidase activity

1.1 Standard curve pNP



1.2 Weight/20 seed and total volume of extraction buffer using in crude enzyme extraction

Day	weight/20 seed	Total volume
0	0.56 g	1.2 ml
1	0.68 g	1.3 ml
2	0.69 g	1.3 ml
3	0.72 g	1.4 ml
4	0.81 g	1.6 ml
5	0.83 g	1.6 ml
6	0.84 g	1.6 ml
7	0.86 g	1.7 ml

1.3 Enzyme activity (U) = $\frac{\mu mol}{min}$

1.4 Units/ gram fresh weight = $\frac{(U)\ (Total\ volume)}{(volume)\ (\ weight/20\ seeds)}$

1.5 Units/ seed = $\frac{(U)\ (Total\ volume)}{(volume)\ (20)}$

1.6 Total activity = $\frac{(U)\ (Total\ volume)}{(volume)}$

2. Preparation for media and reagents

2.1 DEPC treated water

Add 0.1 ml of Diethylpyrocarbonate (DEPC) to 100 ml of distilled water to be treated and shake vigorously. Autoclave the solution at 121◦ C for 15 min to inactivate the remaining DEPC.

2.2 IPTG stock solution (0.1M)

Dissolve 1.2 g IPTG (Isopropyl thio-β-D-galactoside) to final volume of 50 ml distilled water and store at 4◦ C.

2.3 LB broth

Dissolve 25 g of Luria broth base (LB) to 1 liter of distilled water. Autoclave the solution at 121◦ C for 15 min.

2.4 LB plate with 100 ug/ml of ampicillin

Add 15 g selective agar to 1 liter of LB broth. Allow the medium to cool to 50◦ C before adding ampicillin to a final concentration 100 ug/ml. Pour medium into petri-dishes. Allow the agar to harden.

2.5 LB plate with100 ug/ml of ampicillin/IPTG/X-Gal

Make the LB plates with ampicillin as above. Then spread 100 ul of 100 mM IPTG and 20 ul of 50 mg/ml X-Gal over the surface of the plates and allow to absorb for 30 min at 37◦ C before use.

2.6 PDB

NaCl	5.8 g
MgSO4.7H2O	2.0 g
1M Tris-HCl (pH7.0)	50.0 ml
2% gelatin	5.0 ml

Adjust the volume to 1 liter with H_2O. Sterilize by autoclaving at 121◦ C for 20 min.

2.7 Top agarose

Dissolve 25 g of Luria broth base (LB) to 1 liter of distilled water. Autoclave the solution at 121◦ C for 15 min. Add 10 mM $MgSO_4$ and 0.8% agarose before use add ampicillin to a final concentration 100 ug/ml.

2.8 Medium for *E.coli* Y1090r-

Bacto Tryptone	10 g
Bacto Yeast Extract	5 g
NaCl	5 g

Adjust the volume to 1 liter with H_2O. Sterilize by autoclaving at 121◦ C for 20 min. Add 500 ul of20% Maltose and 200 ul of 25 mg-ml ampicillin.

2.9 20XSSC

Dissolve 175.3 g of NaCl and 88.2 g of sodium citrate in 800 ml of H_2O. Adjust the pH to 7.0 with a few drops of 10 N solution of NaOH. Adjust the volume to 1 liter with H_2O. Dispense into aliquots. Sterilize by autoclaving.

2.10 20XSSPE

Dissolve 175.3 g of NaCl, 27.6 g of NaH_2PO_4.H2O and 7.4 g of EDTA in 800 ml of H_2O. Adjust the pH to 7.4 with NaOH (~ 6.5 ml of 10 N solution of NaOH). Adjust the volume to 1 liter with H_2O. Dispense into aliquots. Sterilize by autoclaving.

2.11 50X TAE buffer

Tris-base	242 g
Glacial acetic acid	51.1 ml
EDTA	37.2 g

Adjust volume with distilled water to 1 liter.

2.12 5X TBE buffer

Tris-base	54.0 g
Boric acid	27.5 g
0.5 M EDTA, pH 8.0	20 ml

Adjust volume with distilled water to 1 liter.

BIOGRAPHY

Miss Mallika Chantarangsee was born on 7 August, 1975 in Nakhorn Ratchasima. She graduated with a Bachelor of Science degree in Biology from Khon Kaen University in 1997. She has been working in the position of research assistant to Dr. James Ketudat Cairns funded by the Thailand Research Fund at Suranaree University of Technology for 1 year. During her master degree enrollment in school of Biotechnology, Institute of Agricultural Technology, Suranaree University of Technology. She was support by the Association of International Exchange of Japan (AIEJ scholarship) to work at the University of Tsukuba, Ibaraki, Japan during 6 April-29 November, 2000.