Molecular analysis of post-harvest physiological deterioration of cassava

Han, Yuanhuai

Award date:
2000

Awarding institution:
University of Bath

Molecular Analysis of Post-harvest Physiological Deterioration of Cassava

Submitted by

Yuanhuai Han

For the degree of Ph.D.

of the University of Bath

2000

ProQuest LLC
789 East Eisenhower Parkway
P.O. Box 1346
Ann Arbor, MI 48106-1346

Abstract

Post-harvest physiological deterioration (PPD) of cassava storage roots is associated with mechanical wounding caused during harvest. It is similar to the wound response in other plant systems in some aspects but lacks adequate wound healing processes, leading to an unlocalised wound response throughout detached roots. To understand the mechanism of PPD, a molecular approach was adopted. A cDNA library was constructed from cassava storage roots undergoing physiological deterioration. Using heterologous probes, cDNA clones encoding phenylalanine ammonia-lyase, hydroxyproline-rich glycoprotein, β-1,3-glucanase and ACC oxidase were isolated from the PPD-related cDNA library and characterized. Non-preconceived target genes including those encoding putative ubiquitin-activating enzyme, serine/threonine kinase, glycinamide ribonucleotide synthetase (GARS) and phosphoserine aminotransferase (PSAT) were also identified from the cDNA library.

Two PAL cDNA clones, cMePAL1 and cMePAL3, were isolated from the library. They were 2253bp and 1126bp respectively, and share 79% and 93% similarity over 1126bp and 316 amino acid sequences respectively. Cassava PAL genes shared high identity in deduced amino acid sequences with PAL genes in other plant species, especially dicotyledonous tree plants such as poplar and lemon. Southern hybridisation indicated that there might be at least four genes in the PAL gene family of cassava. cMePAL1 in cultivar Mcol 22 was expressed in young leaves, stems and vegetative roots. It was not expressed in unwounded storage roots but was induced within 8hr after harvest; its expression stopped for at least 20 hrs and was induced again 40 hr post-harvest.

Nine positive clones with inserts ranging from 100bp to 1900bp were isolated from the cDNA library using a probe made from parsley HRGP cDNA. One cDNA, cMeHRGP1, was fully sequenced and comprised 1649bp. Its deduced amino acid sequence consisted of repetitive motifs such as $Ser(Pro)_4$ and $(Tyr)_3$, which are present in HRGPs of other plants. It also had large and tandem repetitive motifs such as $Ser(Pro)_4$-$(Tyr)_3$-His-$Ser(Pro)_4$-Val-Lys and $Ser(Pro)_4$-$(Tyr)_3$-His-$Ser(Pro)_4$-Ser-Pro, which was a combination of motifs present in dicot and monocot HRGPs. Partial sequences of other positive clones showed that three of the clones had high similarity to cMeHRGP1 in nucleotide or deduced amino acid sequences, and two of them showed high similarity to the sequences of the probe used for their isolation.

Eight positive clones were isolated with heterologous β-1,3-glucanase probes and two of them, cMeGLUC1 & 2, were characterised. They were 1015bp and 1694bp respectively, and cMeGLUC1 corresponded to 3' part of cMeGLUC2 with only one base difference between them in this region. The deduced amino acid sequences of cMeGLUC2 shared higher identity with acidic tobacco β-1,3-glucanases than basic ones; notably it shared more residues with PR-Q', a possible elicitor-releasing glucanase in tobacco. Northern analysis showed that β-1,3-glucanase genes were expressed 72 hr post-harvest.

During PPD, ACC oxidase gene was expressed, which was in line with previous reports on the production of ethylene. Ethylene and / or other post-harvest signals may lead to the expression of wound response-related and / or defence response-related genes. At least three different PAL genes (two of them were reported here) were expressed during PPD, which probably contributed to the large amount of biosynthesis of phenolic compounds and vascular streaking, the visible symptom of PPD. The expression of PR protein β-1,3-glucanases was also detected during PPD. PPD was considered a wound response but lacking healing response, but the isolation of putative HRGP cDNA clones from the cDNA library indicated that there might be some component(s) of healing, at least at a transcriptional level. It is not known if there is a link between the expression of these genes and ethylene production during PPD.

The possibility of genetic manipulation of PPD using these isolated genes, and possible approaches to identify genes for genetic manipulation and to understand PPD were discussed.

Acknowledgements

I would like to thank my supervisor Dr. John Beeching and Dr. Richard Cooper for providing me with the opportunity, guidance and encouragement throughout the lab work and write-up. I would like to acknowledge University of Bath for the studentship, without which this study would not have been possible.

I would like to give my thanks to Dr. Andy Bailey for his technical support and Dr. John Clarkson for discussions and suggestions. My thanks also to Felicity, Alison and Anne in teaching lab for their helps and the wonderful times during my practical demonstration with them. I would also like to say thanks to Paul, Janet, Natasha and Kim for their kindness and parties full of funs. I would also like to acknowledge the cassava group – Rob, Rocio, Kim, Maria and Holger for the times together. I would like to give my thanks to the numerous workers in lab 1.52 for help and support.

Finally I would like to thank my parents for their constant support. Last, I would like to say a big thank you to my wife Hongying for her love, support and encouragement for all these years.

Abbreviation

aa	amino acid
ABA	abscisic acid
ACC	1-amino-cyclopropane-1-carboxylic acid
ACMD	African cassava mosaic disease
ACMV	African cassava mosaic virus
ACO	ACC oxidase
AFLP	amplified fragment length polymorphism
AIRC	aminoimidasole ribonucleotide carboxylase
AIRS	aminoimidasole ribonucleotide synthetase
AMV RT	Avian myeloblastosis virus reverse transcriptase
AVG	aminoethoxyvinylglycine
BAC	bacterial artificial chromosome
bp	base pair
4CL	4 coumarate:CoA ligase
C4H	cinnamic acid-4-hydroxylase
CBB	cassava bacteria blight
CBN	Cassava Biotechnology Network
cDNA	complementary deoxyribonucleic acid
CHI	chalcone isomerase
CHS	chalcone synthase
CIAT	Centro International de Agricultura Tropical, in Colombia
cMeGLUC	cDNA clone of β-1,3-glucanase in cassava
cMeHRGP	cDNA clone of hydroxyproline-rich glycoprotein in cassava
cMePAL	cDNA clone of PAL in cassava
CsCMV	cassava common mosaic virus
DNA	deoxyribonucleic acid
dNTP	2′-deoxyribonucleoside triphosphate
E1	ubiquitin activating enzyme
EDTA	diamino ethanetetra-acetic acid, disodium salt
FW	fresh weight
GARS	glycinamide ribonucleotide synthetase

GART	glycinamide ribonucleotide formyltransferase
GCG	Genetics Computer Group
GDH	glutamate dehydrogenase
GRPs	glycine-rich proteins
GUS	β-glucorunidase
HAL	histidine ammonia-lyase
HNL	hydroxynitrile lyase
hr	hour
HRGP	hydroxyproline-rich glycoprotein
IAA	indole acetic acid
IITA	International Institute of Tropical Agriculture, Ibadan, Nigeria
IPTG	isopropyl β-D-thiogalactopyranoside
JA	jasmonate acid
Kb	kilobase
L-AOPP	L-α-aminooxy-β-phenylpropionic acid
LB	Luria-Bertani media
LRR	leucine-rich repeats
min	minute
mRNA	messenger ribonucleic acid
NCBI	National Center for Biotechnology Information
OD	optical density
PAL	phenylalanine ammonia-lyase
PCR	polymerase chain reaction
PEG	polyethylene glycol
Pin2	proteinase inhibitor II
PPD	post-harvest physiological deterioration
PR	pathogenesis-related
PRPs	phosphoserine aminotransferase
PVC	polyvinyl chloride
PVP	polyvinyl pyrollidone
RACE	rapid amplification of cDNA ends
RAPD	random amplified polymorphic DNA
RFLP	ramdom fragment length polymorphism

RH	relative humidity
RLK	receptor-like protein kinase
RNase	ribonuclease
ROS	reactive oxygen species
rpm	revolution per minute
RT-PCR	reverse transcriptase PCR
SA	salicylic acid
SAICARS	5-aminoimidasole-4-N-succinocarboxyamide ribonucleotide synthetase
SAR	systemic acquired resistance
SDS	sodium dodecyl sulphate
TMV	tobacco mosaic virus
Tris	Tris (hydroxymethyl) methylamine
UTR	untranslated region
UV	ultra-violet
VsP	vegetative storage protein
X-Gal	5-bromo-4-chloro-3-indolyl-β-D-galactopyranoside
Y_3	tyrosine-tyrosine-tyrosine

Contents

Chapter Five: Hydroxyproline-rich Glycoprotein (HRGP) Genes were expressed during PPD

Chapter Six: Isolation and Characterisation of β-1,3-glucanase cDNAs expressed during PPD

Chapter Seven: Identification of cDNA Clones of non-preconceived Target Genes and an ACC Oxidase Gene

Chapter Eight: General Discussion

Chapter one

Introduction

1.1 Cassava, the Plant and the Crop

Cassava (*Manihot esculenta* Crantz) is a dicotyledonous perennial plant of the *Euphorbiaceae* family, which also includes other important crop plants such as rubber (*Hevea braziliensis*), and castor bean (*Ricinus communis*). Some other names of cassava are manioc, mandioca, aipim, yuca, tapioca and sagu. The plant is grown as a crop for its swollen starchy roots (Fig. 1.1). The cassava contains laticifers from which latex is produced. The plants grow to a height of 1 to 3 metres or more depending on cultivars and conditions of cultivation. Cassava has erect, glabrous stems and varying degrees of branching. The mature stem is tough and woody due to the development of lenticels and cork on the surface and the formation of extensive xylem. The leaves, which are arranged spirally and cluster to the tops of the stems, are usually dark-green and have three to several obovate-lanceolate, acute lobes upto 20cm long. Flowering is frequent in some cultivars but it is rare or non-existent in others (Onwueme, 1978).

Cassava is propagated vegetatively from cuttings of lignified stems. The cuttings from the mature stem generally give a better yield than those from younger stems. The cuttings are usually planted upright (or at an angle) with two-thirds of their length in the soil, or are planted horizontally. In Thailand, cuttings are planted in the inverted position. This leads to the formation of compact storage root clusters close to the soil surface and branches near the base of the stem, which are easy for manual and mechanical harvesting (personal communication with Sriroth K, Kasetsart University, Thailand). Storage root formation commences as early as eight weeks as a result of secondary thickening in adventitious roots and there is no such secondary thickening change later than six to nine months after planting. A well-grown plant produces five to ten storage roots close to the base of the stem. The size of the storage root is about 15-100cm in length and 0.5-2.0kg in weight (Onwueme, 1978). The storage roots are ready for harvest about 8 to 18 months after planting depending on cultivars. The roots can be left in the ground for long periods up to two years and thus can be used as a famine reserve (Cock, 1985). Cassava is harvested just before consumption or processing as the roots deteriorate rapidly once harvested. It cannot be stored in good condition for more than one or two days after harvesting. The nature of this deteriorative process is the theme of this thesis.

Fig. 1.1 Cassava plants. Top: cassava plants in field. Bottom left and right are cassava leaves and cassava roots.

The best climate for cassava growth is warm (25-29°C), humid with an annual rainfall of 100-150cm. The crop is able to tolerate drought a few weeks after planting. During the drought period, the plant grows slowly or ceases growth and when the drought is over the growth resumes quickly. It can grow and yield reasonably well on soils of low fertility, where it would be uneconomical if other crops were cultivated. Compared to other crops cassava is affected less by pests and diseases. Its ability to withstand locust attack was thought to be an important factor for its spread in Africa (Coursey and Booth, 1977).

The composition of cassava roots has been well documented. The edible fleshy portion makes up 80-90% of the root and is composed of about 62% water, 35% carbohydrate, 1-2% protein, 0.3% fat, 1-2% fibre and 1% mineral matter. Most of the carbohydrate is starch, which makes up 20-25% of the root flesh. The protein in cassava roots is not only low in quantity but also in quality, with low methionine, lysine, tryptophan, phenylalanine and tyrosine (Oyenuga, 1968). However, cassava leaves contain high level of protein, about 29.3-32.4% dry weight (Awoyinka *et al.*, 1995). All cassava tissues except the seeds contain cyanogenic glycosides linamarin (more than 90% of total cyanogen) and lotaustralin (less than 10% of total cyanogen) (McMahon *et al.*, 1995). Leaves contain 5.0g linamarin/kg fresh weight, the highest cyanogenic glycoside level, whereas the level of total root linamarin ranges from 100 to 500mg/kg fresh weight depending on the cultivars. Based on the taste of the roots, cassava cultivars are divided into two groups, bitter types and non-bitter types (or sweet types or cool types). In the bitter types cyanogenic glycosides are present throughout the storage root and at high levels while in the other type only in the cortex and they are at low levels. It was discovered recently that cyanogenic glycosides were the sole contributor of bitterness in the parenchyma but in two out of six cultivars studied isopropyl-β-D-glucopyranoside contributed more to the bitterness of cortex (King and Bradbury, 1995).

Fresh cassava roots of most cultivars have to be processed soon after harvest and before consumption because of two factors. One is that the roots would deteriorate rapidly after harvest. The other one is that hydrocyanic acid would be produced in unprocessed roots from the relatively high level of cyanogenic glycosides when the roots are consumed.

Only roots of cultivars with very low level of cyanogenic glycosides can be consumed with little or no processing.

Cassava roots can be processed into four main products including meals, flour, chips (or pellets) and starch. Chips and starch, to be used for industry, are forms for cassava international trade. In most of processing methods, generally, the roots are peeled, grated or milled and baked or dried, during which the hydrolysis of cyanogenic glycosides to hydrocyanic acid is accelerated and the hydrocyanic acid evaporates. The most popular form of cassava consumption is *gari*, a dried form of cassava, which can be stored for long periods at room temperature. During the production of *gari*, peeling, milling, pressing, sieving and toasting are involved. Retting is also another way to prepare cassava meals with low hydrocyanic acid content; this is done by steeping the roots in water for several days until they are soft and then further processed by sieving and squeezing (Onwueme, 1978).

1.2 Origin, Distribution and World Production

Since cassava does not exist in the wild state and its wild ancestors are not known, its origin has been disputed. Cassava was thought to be native to South America and/or Central America, and was introduced to Africa in the late sixteenth century and to Asia during the late seventeenth century (Rogers and Appan, 1973; Allem, 1987). The hypotheses that cassava originated from South America or Brazil was supported with more evidence recently from the fertile relationship analysis and phylogenetic analysis based on microsatellite-primed PCR markers of the crop and South American wild *Manihot* species. These analyses revealed that South American wild cassava subspecies *M. flabellifolia* and *M. peruviana* were the closest wild relatives of cassava (Roa *et al.*, 1997; Allem *et al.*, 1998; Cabral *et al.*, 1998). Phylogeographic analysis of cassava and its wild relatives revealed that the haplotypes shared between cassava and wild *M. esculenta* subspecies occurred along the southern border of the Amazon basin but not in the eastern border, suggesting the southern Amazon border region as the geographical origin of cassava domestication (Olsen and Schaal, 1999).

Its tolerance of low fertility soil and drought, its ability to resist or recover from pest or disease damage and its particular function as a famine reserve, makes the cultivation widespread over tropical regions in the world. It is cultivated in more than 80 countries

between 30° south and 30° north of the equator. Cassava is the most important root crop in the tropics and it is among the world's ten most important food crops (Taylor *et al.*, unpublished). World annual cassava production from 1995 to 1997 was about 165 million tons, with Nigeria, Brazil and Thailand as the three largest producers. The yields range from 1.8 ton/ha (Sudan) to 27.3ton/ha (Barbados) (FAO, 1997). Serving as a staple food for 500 million people, it is in fourth place after rice, sugar cane and maize in terms of calories for human consumption in the tropics (CIAT, 1992). In addition, the roots are also used as animal feed and deployment as a raw material for industries (for production of starch, sugar, alcohol and acetone) has been increasing during the recent decades (Cooke and Cock, 1989; CIAT, 1992). During 1993 to 1995, the distribution of cassava among food, feed and industries was 59%, 24% and 17% respectively. With increasing population pressures and resulting demands for food and resources it is anticipated that cassava cultivation, production and its distribution in industry will increase world-wide.

1.3 Constraints to Cassava Production

Cassava production is affected by diseases, pests, cyanogenic toxicity and rapid post-harvest deterioration.

In spite of its relative tolerance, diseases and pests can cause considerable losses in cassava production. It was estimated that various diseases and pests cause 20-50% yield losses world wide, and sometimes result in total losses in some areas (Puonti-Kaerlas, 1998). The cassava diseases include cassava mosaic virus, bacteria blight, bacterial stem rot, brown leaf spots, white thread. Two major virus diseases are cassava common mosaic virus (CsCMV) and African cassava mosaic disease (ACMD). Symptoms show mosaic, leaf curl and significant reduction of yield. CsCMV caused yield losses up to 30% in some countries of South America (Calvert *et al.*, 1996). African cassava mosaic virus causes the most serious disease of cassava, resulting in severe losses of cassava production in some countries in Africa recent years (Harrison *et al.*, 1997). Cassava bacteria blight (CBB) is a very serious cassava disease, which is caused by *Xanthomonas axonopodis pv manihotis (Xam),* with the symptoms of angular spots on the leaves, blight and wilting of the leaves, vascular necrosis in the stem and root, and the die-back of the shoots. CBB is usually spread through infected cuttings and the loss can be up to 100%. Several pests including mealybugs (*Phanacoccus manihoti*), whiteflies (*Bemisia*),

stem borer (*Chilomina clarkei*), mites (including green spider / *Mononychellus tanajoa*, red spider / *Tetranychus telarius* and web spider / *Oligonychus spp.*) and root-knot nematodes (*Meloidogyne incognita*) affect yield, sometimes causing severe damage.

The toxicity of cassava due to the presence of cyanogenic glycosides has been a problem for the consumption of cassava as food or feed (Jones, 1998). Consumption of unprocessed or not fully-processed bitter cassava can cause food (or feed)-poisoning or even death. Various disorders such as hyperthyroidism and tropical ataxic neuropathy are associated with the consumption of cassava (White *et al.*, 1998). The high content of cyanogenic glycosides also presents a limitation to the world cassava trade, for which a low level of cyanogenic glycosides is required. In order to detoxify the roots for daily consumption, various methods of post-harvest processing and food preparation are used by farmers and processors in different regions as mentioned above.

Rapid post-harvest deterioration is another major constraint to cassava production. The rapid deterioration renders the roots unacceptable for human, animal and industrial use. With the increasing demand in industry and the increase in cassava yield there is a subsequent increase in the demand for marketing and large-scale processing. Storage of roots has become one of the major problems in cassava production, which severely constrains the marketability of the roots and the scale of processing; this problem therefore limits the scale of the crop cultivation.

1.4 Recent Advances in Cassava Research

Though of vital importance to millions, cassava has not been intensively studied, rather it has been neglected relative to all other major world crops. Cassava ranks the bottom of the list of important food crops in terms of research expenditure as a percentage of crop value (Taylor *et al.*, unpublished).

Early research on cassava improvement began in 1920s, with focus on starch extraction and breeding programmes on disease resistance. One of the problems in breeding programmes has been the narrow adaptation of most cassava cultivars, reflecting the long history of local selection. With the establishment of the international institutes IITA (International Institute of tropical Agriculture, Ibadan, Nigeria) and CIAT in 1960s, fully integrated cassava research programmes were made possible. IITA and CIAT have been

playing leading roles in breeding and improving production systems. Significant progress has been made in cassava breeding in increasing both yield and resistance to pests and diseases. It was estimated that yield increased from 7.7tons/ha in early 1960s to 9.5tons/ha in the early 1980s, which represents about 0.9% increase per year. Lines with resistance to ACMV and CBB have been developed at IITA (Taylor *et al.*, unpublished).

The importance of the crop has been recognised since the late 1980s. The cassava Biotechnology Network (CBN), which links the cassava research groups over the world, was established in 1988 with an aim to apply biotechnological approaches to the problems in cassava production and processing. CBN has been promoting the communication and collaboration among cassava researchers, and proposing research priorities, which have accelerated the cassava research to a great extent. As a result, significant progresses have been made in different aspects of cassava research in recent years.

Conservation of genetic resources, has been one of the research focuses as it is realised that the loss of plant populations and species happens relentlessly in the tropics. A core collection of cassava was established in CIAT, which included 651 genotypes (Hershey *et al.*, 1992). A Brazilian core collection of cassava was established, which consisted of 4100 accessions conserved in five regional Active Germplasm Banks (Cordereiro *et al.*, 1998). These collections provide a wider range of materials for breeding programmes. Part of the collections have been conserved *in vitro* (Mendes *et al.*, 1998), and cassava cryopreservation of shoot tips through an encapsulation-dehydration technique is being developed to make cassava germplasm conservation simple, and of low cost and high quality (Escobar *et al.*, 1998). The genetic diversity of cassava was assessed using random amplified polymorphic DNA (RAPD) markers or microsatellite-primed PCR markers. The valuation and relationship analysis of *Manihot* genetic resources suggested that cassava descends from two original wild subspecies of cassava as mentioned above, and that there were about 75 species for the entire genus (Allem *et al.*, 1998), less than was reported before which was 98 species (Rogers & Appan, 1973)

The advent of genetic engineering techniques has opened the possibility of genetic modification of cassava for higher yields and better quality. It was first reported in 1996 that transgenic cassava plants were produced using either particle bombardment or

Agrobacterium tumefaciens (Schopke *et al.*, 1996; Li *et al.*, 1996). Then much effort was spent on improving and optimizing transformation systems by transferring selectable markers and visible marker genes into cassava. Recently, the focus has moved onto transformation with genes of agronomic interest. Various gene constructs were transferred into cassava to improve certain aspects of cassava production. Transgenic cassava plants containing the coat protein gene of cassava common mosaic virus, the AC1 (replication-associated protein) gene and defective interfering (DI) genes of African cassava mosaic virus, viral antisense RNAs, truncated viral gene products and defective interfering DNA were produced (Schopke *et al.*, 1998; Taylor *et al.*, 1998; Scharer-Hernandez *et al.*, 1998) and the effects of these transgenes on the resistance of cassava to mosaic virus diseases are being assessed. Bacteria blight resistant gene *Xa*21 from rice was also introduced into cassava to study its effect on the resistance to cassava bacterial blight (Taylor *et al.*, 1998). Five BAC clones from cassava have been identified containing *Xa*21 homologue and these BAC DNAs are to be bombarded into cassava to investigate their effect on the defence of the cassava to CBB (Taylor *et al.*, unpublished). Engineering insect resistance in cassava is also being attempted, in which *cry*1A(b) gene from *Bacillus thuringiensis* was used to produce cassava resistant to stem borer (Legris *et al.*, 1998; Chavarriaga *et al.*, 1998). Transgenic plants containing hydroxynitrile lyase cDNA were produced in an attempt to reduce the cyanide toxicity (Arias-Garzon and Sayre, 1998). Cytokinin biosynthesis gene (from *Agrobacterium tumefaciens*) driven by a senescence-specific promoter (*sag* from *Arabidopsis*) was introduced into cassava to prolong the life-time of leaves (Li *et al.*, 1998). Transgenic plants with ADPG-pyrophosphorylase antisense gene were produced to decrease the content of starch and increase that of sugars (Munyikwa *et al.*, 1998). Tissue-specific promoters are of particular potential value for genetic engineering. Various research groups are searching for root-specific promoters. Using a differential screen or cDNA-AFLP, many cDNA clones showing root specific expression have been isolated and their corresponding genomic DNA sequences will be analysed to isolate root-specific promoters (Bohl-Zenger *et al.*, 1998; Huang *et al.*, 1998). Promoters isolated from cassava phenylalanine ammonia-lyase (PAL) gene and protein synthesis elongation factor gene are also being investigated for their potentials in genetic engineering (Li *et al.*, 1998; Suhandono *et al.*, 1998).

On the toxicity of cassava, the pathway from linamarin to cyanide is well characterized. Cyanogenic glycosides linamarin and lotaustralin in cassava are synthesized in leaf tissue from valine and isoleucine and transported to roots. Linamarin is the main form of cyanogenic glycoside and it is present in all tissues of cassava where it is stored in the vacuole. The cyanogenic enzymes, linamarase and hydroxynitrile lyase were characterised and their genes cloned (Hughes *et al.*, 1992; Pancoro & Hughes, 1992; Hughes *et al.*, 1994). Mechanical damage of the cassava tissue such as during food preparation leads to the deglycosylation of linamarin by linamarase to acetone cyanohydrin, which can break down to acetone and HCN spontaneously or by hydroxynitrile lyase. It was suggested that the toxicity of cassava is due in part to the absence of hydroxynitrile lyase (HNL) in the root. Transgenic cassava plants containing HNL cDNA were produced, which showed higher HNL activity and the effect of the transgene on the toxicity of the storage root is to be analysed (Arias-Garzon and Sayre, 1998).

Research on diseases in cassava has been enhanced by establishment of core collections and the development of molecular tools. Deployment of resistant varieties is one of the main approaches to control diseases. Ninety-three varieties of *M. esculenta* Crantz from the cassava core collection CIAT were screened for resistance to cassava bacteria blight and it was found that the resistance was broadly distributed in cassava germplasm (Sánchez *et al.*, 1998). A molecular genetic map, with markers linked to resistance genes to cassava mosaic disease (ACMV) and cassava bacterial diseases has been developed at CIAT (Fregene *et al.*, 1998; Akano and Dixon, 1998). Conventional breeding programmes have been carried out to produce disease resistance cultivars. Genetic engineering approaches are being used to produce disease-resistant or insect-resistant cassava plants as mentioned above. A suspension cultured cell system has been used to study the biochemical components of disease resistance of cassava. These cells showed a series of resistance-related responses such as a rapid (within 3min of elicitation) and intensive, extracellular oxidative burst, and induction of PAL activity, to microbial elicitors such as glucans from yeast and to certain endogenous elicitors such as oligogalacturonides, jasmonic acid and glutathione (Gomez-Vasquez *et al.*, 1998).

1.5 Post-harvest Physiological Deterioration

Cassava roots are not propagatable, unlike the tubers of other tropical crops such as yam, potato and sweet potato which are natural perennating organs and can be stored for certain time during physiological dormancy (Onwueme, 1978); instead cassava cuttings from lignified stems and sometimes seeds are used to propagate the crop. The cassava roots cannot be kept in a commercially satisfactory condition for more than a few days at ambient conditions after harvest due to rapid post-harvest deterioration. The rapid deterioration, which can develop within 24 hours after harvesting, leads to the decline in the quality of starch and palatability.

1.5.1 Post-harvest Storage

Because of rapid post-harvest deterioration, cassava roots are usually 'stored' in the living condition in the soil and harvested on a small scale only when they are needed, and immediately consumed, processed or marketed. But roots become more lignified and fibrous if they are left under the ground too long after their optimal development, and become low in the starch content and palatability (Onwueme, 1978; Ravi *et al.*, 1996). This also results in the occupation of large areas of land with already mature cassava. A number of underground storage methods have been studied such as storing in clamps, trenchs and pits of soil and sand (reviewed by Ravi *et al.*, 1996).

A number of chemicals were screened for effectively reducing PPD and some of them could reduce the deterioration to some extent in small scale experiments such as ethyl bromide, ethylene dibromide, formaldehyde, lactic acid, benzoic acid, sodium o-phenyl phenate, calcium and sodium hypochlorite, ethyl alcohol and benomyl (Booth, 1976). Treatment of cassava roots with benomyl and chlorine solutions reduced the level of fungal growth and incidence of bacteria respectively but PPD was not affected (Thompson and Marinia, 1977). It was therefore pointed out that chemicals for the control of pathogens were of no value unless PPD was also controlled (Plumbley and Rickard, 1991).

Storage of the roots in a polyethylene bag or dipping roots in wax to seal the roots can effectively prevent PPD and enable the roots to be stored for more than two weeks, especially if the roots are pre-treated with fungicide (thiabendazole-based) to control microbial deterioration (Wheatley *et al.*, 1989). Roots with minimal mechanical damage

and minimal exposure to the sun, if treated with fungicide (thiabendazole) and packed within 3 hours after harvest, can be stored for 15 days in polythene bags at about 30°C (Wheatley *et al.*, 1989). Roots dipped in paraffin wax at 90 to 95°C for short periods such as 45 seconds could be stored for up to 1 to 2 months and this technique is commercially used (Ravi *et al.*, 1996). Fresh harvested roots can also be stored for 2 to 4 weeks in an acceptable condition for marketing in boxes packed with moist sawdust (Booth, 1977). Roots dipped in 1% benomyl solution and stored in a polyethylene-lined cardboard box with moist sawdust or sand or their mixture could last for 13 to 14 weeks (Wickham and Wilson, 1988).

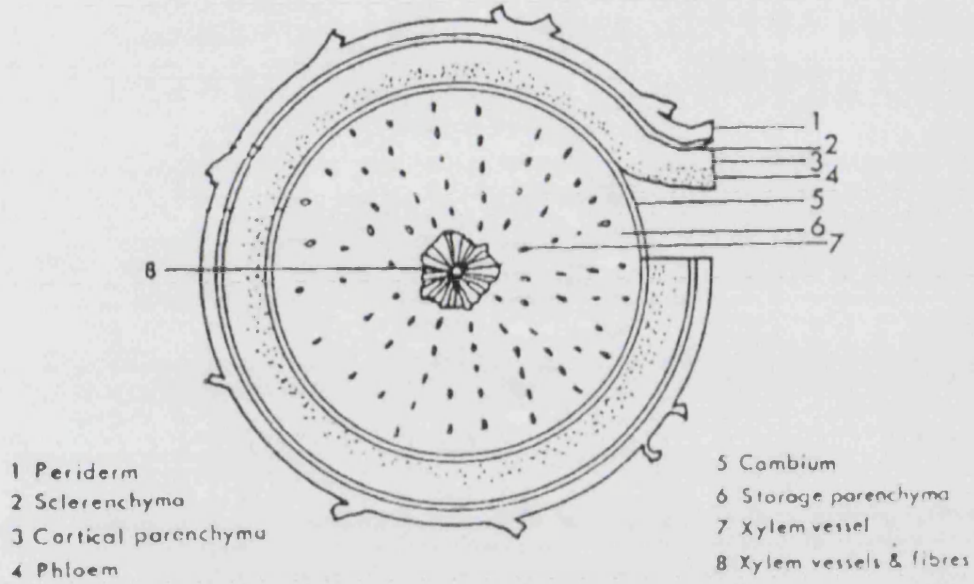
Cassava roots can be stored for several weeks at low temperatures of 0-5°C (Montaldo, 1973; Booth, 1976). However, PPD developed within 1 or 2 days after the roots were transferred to 24°C (Montaldo, 1973). Dipping in water of 60°C for 45min was effective in inactivating PPD or inhibiting the occurrence of vascular streaking during the 5-day observation after the treatment (Averre, 1967).

Curing, treating cassava at high temperature (35°C) and high humidity (80-85% relative humidity) for a period (two to four weeks) after harvest, can extend the shelf-life of the roots up to four more weeks. However, mechanical damage to the cured roots could lead to their rapid deterioration in a similar way to freshly-harvested roots (Booth, 1976).

These solutions are neither convenient, nor economical, nor suitable for large scale marketing and processing. In Thailand, the major exporting country of cassava (5.1 million tons in 1997, contributing about 81% of the world export of cassava), the storage problem is avoided by establishing large processing plants with excessive capacity to guarantee that all the roots collected could be processed within a few hours after harvesting (Sriroth K, personal communication). The processing plants have a break of a few months every year during which no cassava is harvested. The processing capacity is far from being efficiently exploited.

1.5.2 Post-harvest Deterioration

Cassava roots deteriorate rapidly within a few days after harvest. The symptoms of the deterioration are discoloration and decay or rotting. Initial research on the deterioration was on the identification of micro-organisms involved in the process. A number of micro-organisms were isolated from deteriorating roots, such as *Rhizopus sp.*, *Bacillus spp.*, *Lasiodiplodia theobromae*, *Trichoderma harzianum*, *Cylindrocarpon candidum*, *Aspergillus niger* and *flavus*, *Rigidoporus lignosus*, *Phytophthora drechsleri*, *Botryodiplodia theobromae*, *Pythium butleri* (Booth, 1976; Taniguchi *et al.*, 1984). Microbial infection leads to softened and disordered tissues, blue-black pigmentation of xylem vessels in tissues adjacent to the infected regions (Hirose, 1986) and other different symptoms depending on the pathogens (for review see Plumbley and Rickard, 1991). No microbes were isolated from the leading edge of discolored tissue (Averre, 1967; Noon and Booth, 1977). This was one of the bases on which Averre (1967) proposed that the rapid vascular streaking was a physiological disorder and the process of post-harvest deterioration consisted of at least two disorder stages, discoloration of the vascular tissues and microbial rotting. Other supportive evidence is that PPD could be delayed by exclusion of oxygen or keeping the roots under water, and PPD could be inhibited by hot water (53°C for 45 min) treatment (Averre, 1967). The two stages of post-harvest deterioration were classified as primary or physiological deterioration and secondary or microbial deterioration (Booth, 1976; Onwume, 1978). The visual evidence for physiological deterioration is vascular streaking, a blue-brownish coloration of the xylem vessels in the storage parenchymatous tissue (Fig. 1.2). A diagram of transverse section of cassava storage root (Hunt *et al*, 1977) is presented to show the anatomy of the root. During preparation of cassava root for human consumption, Part 1 to 4 are peeled and the storage parenchyma is processed.



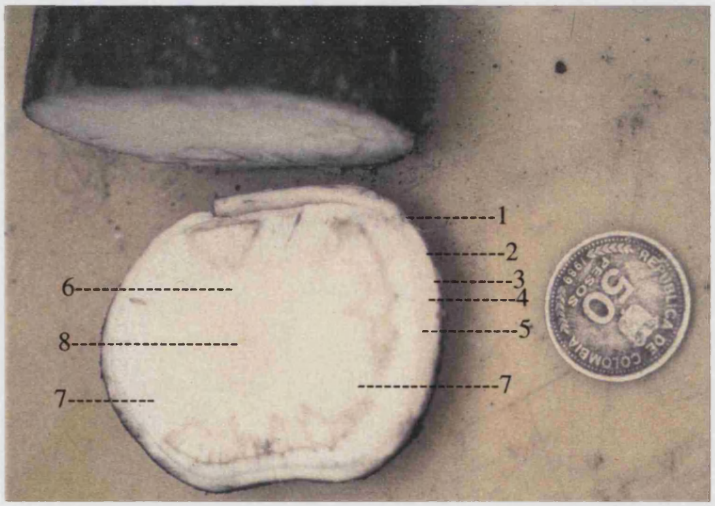


Fig.1.2 Diagram of transverse section of cassava storage root (Hunt *et al*, 1977) and transverse section of a deteriorating root. During preparation of cassava root for human consumption, Part 1 to 4 are peeled and the storage parenchyma is processed. During deterioration, xylem vessel are discolored, as indicated (7) in the bottom picture.

1.5.3 Factors Affecting PPD

The PPD responses vary with cultivars, stresses and storage conditions. Roots from different cultivars respond with different rates in the development of PPD. Montaldo (1973) investigated the response of 65 cultivars to PPD at 24°C and 80% relative humidity. Among these cultivars most of them (44 cultivars) showed streaking within 1 to 3 days after harvest and were classified as susceptible. Eleven cultivars did not develop streaking until seven days after harvest and two did not show streaking as long as 11 days after harvest, and these cultivars were grouped as PPD-resistant. Cultivars with low PPD response usually have a very low dry matter content, which makes these cultivars of less value (Wheatley, 1982). Defoliation due to the biotic or abiotic stresses of pests or drought led to reduced susceptibility to PPD. It was observed that non-locally adapted cultivars were less susceptible to PPD than the local ones (Wheatley, 1980). Pruning, cutting off the top part of the plant from 20-30cm above the base of the stem two to three weeks prior to harvesting, also reduced susceptibility to PPD (Lozano *et al.*, 1978; Tanaka *et al.*, 1984). However, the appearance of new shoots counteracts this effect. Removal of regrowth was necessary to achieve maximum reduction of PPD (Lozano *et al.*, 1978). The similar effects of these different stresses on the susceptibility of roots to PPD suggest that pre-stressing may prepare the roots for further stresses such as PPD, somewhat analogous perhaps to systemic acquired resistance and infection. Curing and storing in polyethylene bags affect PPD in a different way from pruning. The former two can prolong the shelf-life of the roots but still retain sensitivity to mechanical damage while pruning reduces the sensitivity. Storage roots of different developmental stages were investigated for their response to PPD and it was revealed that the development of PPD was not correlated with root age (Hirose and Data, 1984).

1.5.4 Physiological Deterioration

Cassava roots are inevitably wounded and usually suffer abrasions during harvesting. Handling and transport can further damage the root mechanically. It is from these wound sites that PPD is initiated and spreads through the root, and the more severe the mechanical damage of the roots is, the higher degree of primary deterioration (Booth, 1976). The development of physiological deterioration could be stimulated by water loss (Marriott *et al.*, 1978). When transverse cuts of roots were stored with the proximal end uncovered and the other end covered with PVC film, vascular discoloration occurred under the uncovered surface and extended along the pieces (Marriott *et al.*, 1978). Root

pieces with the cut covered by PVC film significantly reduced the water loss compared to uncovered ones, and vascular discoloration in covered pieces was of much less significance than in uncovered. It was observed that severe mechanical damage to the roots caused more fresh weight loss and higher degree of PPD (Booth, 1976). Wounded cassava roots responded differently to low and high humidity (Marriott *et al.*, 1978; Rickard, 1985). Stored at low humidity (45-55%), the wound sites formed a dry white layer within 72hr, beneath which a thin dark brown or greenish-brown band developed and vascular streaking was observed along the root. However, vascular streaking was not observed in the roots incubated at high humidity (80-90) (Rickard, 1985). The effect of the curing process which reduces water loss from the roots and promotes the suberization of the wounding sites and the formation of wound periderms (Coursey and Booth, 1977), also suggests that wounding and water loss is closely associated with PPD.

Microscopic observation revealed that the initial response to wounding in cassava roots was occlusion of the xylem vessels and production of phenolic compounds in the storage parenchyma. The main components of the occlusions were carbohydrates, lipids and lignin-like material (Rickard *et al.*, 1979; Rickard, 1983). The lignin-like materials were thought to be condensed tannins formed by polymerization of the leucoanthocyanidins and catechins.

A range of physiological and biochemical changes have been associated with the development of PPD. Development of PPD could be delayed in the absence of oxygen (Marriott *et al.*, 1978). When roots were stored in pure CO_2, or propane gas or oxygen-depleted air, PPD was significantly reduced (Marriott *et al.*, 1978). Respiration rate of the storage roots increased 2 fold within 1 day after harvest (Data *et al.*, 1984). Hirose (1986) observed that roots injured by removing partial cortex or periderm showed higher respiration rates than intact roots during storage. Two respiration peaks were also observed during storage, which occurred 1 day and 6 days after wounding. The first respiration peak was thought to be caused by wounding and the second was suggested to be due to biochemical changes induced by the development of PPD.

Intense fluorescence developed in the storage tissue before the appearance of general tissue discoloration, which is mainly due to the increase in coumarin components,

scopoletin (6-methoxy-7-hydroxycoumarin), scopolin (6-methoxy-7-hydroxycoumaroyl-7-β-D-glucoside), esculin (6,7-dihydroxycoumaroyl-6-β-D-glucoside) and two conjugates containing scopoletin and esculin respectively (Rickard, 1981; Tanaka, 1983). The fluorescence increased in both intensity and distance from the wound surfaces with time in roots incubated at low humidity, whereas the increased fluorescence remained localised below the wound sites at high humidity (Rickard, 1985). Scopoletin was reported to be produced first and peaked along with scopolin at 24 hours; esculin peaked at 40 hours after wounding (Wheatley and Schwabe, 1985). Total phenol content as well as leucoanthocyanidins and flavanols increased during the development of physiological deterioration (Rickard, 1985). Other phenolic compounds such as (+)-catechin, (+)-gallocatechin and proanthocyanidins were also identified in the wounded root during storage (Tanaka *et al.*, 1983; Uritani *et al.*, 1983).

Scopoletin was proved to be involved in PPD by applying exogenous application to cassava roots. A range of phenolic compounds including cinnamic acid, *p*-coumaric acid, caffeic acid, ferulic acid, coumarin, umbelliferone, esculetin, arbutin, catechol and scopoletin were applied at 1000mg/dm^3 to freshly-harvested roots to test their effect on PPD (Wheatley and Schwabe, 1985). Only scopoletin caused a significant deterioration after 18hr incubation and the visual symptom of the deterioration was like those of the natural occurring deterioration. Treatment of cassava roots with different concentrations of scopoletin showed that the roots responded to 500mg/dm^3 of scopoletin significantly and the reactions to 1000mg/dm^3 and 5000mg/dm^3 were intense in that all the vessels were discolored. Roots from plants with tops pruned two to three weeks before harvest responded to applied scopoletin as vigorously as roots from unpruned plants. However, if roots were stored in the absence of oxygen and then treated with scopoletin the roots did not respond. Analysis of the scopoletin content in pruned samples and unpruned samples 3 days after harvest showed that the former ones had 10% scopoletin of the latter ones (Wheatley and Schwabe, 1985). It was therefore thought that the effects of pruning and low or no oxygen storage on PPD was based on a different mechanism (Wheatley and Schwabe, 1985), possibly involving inactivation of different stages of the metabolism of scopoletin. It was postulated by Wheatley and Schwabe that pruning might be effective due to internally-reduced scopoletin supply or absence of some factors involved in the primary oxidation; while the absence of oxygen might lead to the loss of a scopoletin precursor or inactivation of related enzymes.

Phenylalanine ammonia-lyase (PAL), which is the key entry enzyme to phenylpropanoid biosynthesis, increased in activity and peaked at about 40 hours after cassava roots were wounded, which coincided with the increase of phenolic compounds during PPD (Tanaka, 1983; Uritani, 1983). Rickard (1985) investigated the activities of peroxidase and polyphenol oxidase during cassava deterioration. The appearance of coloured deposits in the xylem vessels and storage parenchyma was accompanied by increased activity of polyphenol oxidase and peroxidase, and a decrease in free phenols (Rickard, 1985).

Differences between roots from pruned and unpruned plants in physiological and biochemical changes during storage were observed. The activities of PAL, peroxidase and acid invertase, and the content of coumarins and phenols were significantly lower in roots from pruned plants than from unpruned plants 1 to 2 days after harvest (Tanaka *et al.*, 1984). Similar results were shown by Data *et al.* (1984) and it was also shown that roots from unpruned plants exhibited higher respiration rate, starch and HCN. Pruning had no effect on polyphenol oxidase acitivity. Pruning significantly increased the sugar content in the roots, which was reflected in the sweeter taste of the roots.

Ethylene is the only intercellular signal so far studied in cassava. Ethylene production in the root has been found to increase after a lag of 6 to 16 hours after wounding (Plumbley *et al.*, 1981; Hirose *et al.*, 1984). During PPD two-fold more ethylene was produced in roots susceptible to PPD than in roots from less susceptible cultivars (Hirose, 1986). During the incubation of separate parts of the root slices, most ethylene was found to be produced in the cortical parenchyma while the storage parenchyma where PPD occurs did not produce significantly more ethylene. Exogenous application of ethylene to roots showed no obvious effect on the development of PPD (Hirose *et al.*, 1984). It was suggested that ethylene may not be directly involved in the development of PPD (Wenham, 1995). However, it cannot be excluded that endogenous ethylene may be involved in the induction of the rapid PPD, since exogenous ethylene may not have the effect of endogenous ethylene. For example, exogenous ethylene did not induce Pin2 gene expression in tomato whereas endogenous did (Bowles, 1997).

During the development of PPD, the content of β-carotene in yellow roots decreased and the degree of PPD was in negative correlation with the content of β-carotene in the roots (Gloria and Uritani, 1984). It is not known whether cultivars with high content of the anti-oxidant β-carotene or yellow roots are more resistant to PPD than white roots.

1.6 Wound Responses in Other Plant Systems

The physiological and biochemical changes during the development of PPD show similar responses to those induced by wounding or to pathogen attack in other plant systems, such as the increase in respiration rate and induction of ethylene upon wounding or post-harvest. It may therefore be a rational way to analyse the special response of PPD by referring to the wounding responses in other well-studied plant systems.

1.6.1 Wound Responses

Plants respond to mechanical wounding by defense responses characterized by the activation of defense-related genes or expression of a number of proteins, functioning for wound healing and prevention of pathogen invasion.

During the wound response phenylpropanoid metabolism is activated. For example, expression of the genes coding for enzymes in the phenylpropanoid pathway such as PAL, C4H (cinnamic acid-4-hydroxylase), 4CL (4 coumarate:CoA ligase), chalcone synthase (CHS), chalcone isomerase (CHI) are induced or enhanced, which leads to the increased production of a range of phenolic compounds (Cottle and Kolattukudy, 1982; Lamb, 1977; Ryder *et al.*, 1987; Mizutani *et al.*, 1997; Faktor *et al.*, 1997; Mehdy and Lamb, 1987; Lawton and Lamb, 1987). Wound-induced chlorogenic acid, alkyl ferulate esters, and cell wall-bound phenolic esters may act directly as defense compounds or serve as precursors for the synthesis of lignin, suberin, and other wound-induced polyphenolic barriers (Halbrock and Scheel, 1989; Bernards and Lewis, 1992). Within 3 to 7 days of wound treatment of potato tubers, ferulates began to accumulate, and were temporally and spatially correlated with suberin formation, and restricted to the wound periderm (Bernards and Lewis, 1992).

Structural proteins such as hydroxyproline-rich glycoproteins (HRGPs) and glycine-rich proteins (GRPs), suberin, lignin, wall-bound phenolics, and callose, are expressed or formed to reinforce or repair the extracellular matrix upon wounding (Kolattukudy,

1980; Cassab and Varner, 1988; Bowles, 1990; Dixon and Lamb, 1990; Boudet *et al.*, 1995). But in the harvested cassava storage root, the healing process seems to be poorly expressed and the wound responses are not localized to the vicinity of the wounding site (Booth, 1976; Rickard and Coursey, 1981). This may be related to the fact that cassava roots are not functional in propagation as are those of other crop tubers. However, there are cases when the PPD could be suppressed as described in 1.5.3. High humidity and temperature (85%RH, 35°C), known as curing, can delay the development of PPD by up to 4 weeks in ambient conditions (Booth, 1976; Rickard, 1985). It was also shown that under such conditions there was a more typical wound response with localized production of phenols and formation of periderm (Richard, 1985).

In response to wounding, reactive oxygen species (ROS) are generated in plants. There are several mechanisms generating ROS in plants such as NADPH oxidase, pH-dependent cell wall peroxidase, exocellular germin-like oxidase, amine oxidase and protoplastic ROS-generating systems (for review see Bolwell and Wojtaszek, 1997). It was recently shown that the oxidative burst induced by elicitors in rose and French bean cells were generated by two distinct mechanisms. H_2O_2 was produced by a plasma membrane NAD(P)H oxidase or superoxide synthase in rose cells, but by a cell wall peroxidase in French bean cells (Bolwell *et al.*, 1998). Orozco-Cardenas and Ryan (1999) investigated the generation of hydrogen peroxide in leaves upon wounding using 3,3-diaminobenzidine (DAB) as a substrate for visual detection of hydrogen peroxide in 18 plant species such as tomato (*Lycopersicon esculentum*), cucumber (*Cucumis sativus*), maize (Zea *mays*) and pea (*Pisum sativum*), *Arabidopsis* and cotton (*Gossypium hirsutum*). It was revealed that 14 species showed the generation of H_2O_2 upon wounding. In some species such as cucumber the response was located primarily at wound sites, whereas in maize, pea, cotton and potato the response was strongly systemic. Hydrogen peroxide can be detected at wound sites and in distal leaf veins within 1 hr after wounding. Wound-induced H_2O_2 is involved in oxidative cross-linking of cell wall proteins. Resuspension of a cell wall fraction from soybean suspension-cultured cells in H_2O_2 led to rapid cross-linking of some proline-rich cell wall proteins, resulting in complete insolubilization of the proteins within 2 minutes (Bradley *et al.*, 1992). The rapid insolubilization was also induced by wounding in plants. Wounding of a soybean hypocotyl resulted in complete insolubilization in the 5mm section of tissue immmediately surrounding the incision. It was suggested that a rapid burst of H_2O_2

rather than peroxidase activity was the rate-determining step controlling the cross-linking of the proline-rich cell wall proteins, as elicitor treatment did not stimulate further *in vivo* peroxidase activity above a high basal level during insolubilization (Brisson *et al.*, 1994). The rapid insolubilization of pre-existing cell wall proteins is earlier than elicitor-induced defense gene transcription which reached maximum rate of transcription after 1 to 3 hrs (Lamb *et al.*, 1989), which means that a rapid defense response is switched on before the expression of defense mechanisms depending on transcriptional activation (Bradley *et al.*, 1992). ROS are also involved in lignification and suberization. Recently the anionic peroxidase associated with the suberization during potato wound healing was purified (Bernards *et al.*, 1999). It was localised to suberizing tissues in the immediate vicinity of the wound site and is more active on o-methoxyphenol-substituted hydroxycinnamates (accumulated in tubers during wound healing and incorporated into the suberized cell wall) than other phenolic compounds.

During wound response a number of defense-related genes are activated or induced such as proteinase inhibitors (as described in the following section), and hydrolases such as chitinases (Hedrick *et al.*, 1988) and β-1,3-glucanases (as described in section 1.9).

1.6.2 Systemic Wound Responses

During the interaction between plants and pathogens, systemic acquired resistance may develop in plants through long distance signalling. Similarly, wounding also stimulates systemic responses.

In response to wounding, a variety of genes are activated in plants. Some of them are not only expressed in the vicinity of the wound site, but also systemically activated in the non-damaged parts of the injured plants. These genes include proteinase inhibitor II (Peña-Cortês *et al.*, 1988; Farmer and Ryan, 1990), vegetative storage proteins (Mason and Mullet, 1990), prosystemin (McGurl *et al.*, 1992), polyphenol oxidase (Constabel *et al.*, 1995), sulfhydryl proteinase inhibitor, cathepsin D inhibitor (Hildmann *et al.*, 1992), proteolytic enzymes (such as carboxypeptidase, leucine aminopeptidase and aspartic proteinase) and threonine deaminase (Hildmann *et al.*, 1992; Walker-Simmons and Ryan, 1977; Pautot *et al.*, 1993). Among these systemically inducible genes, the proteinase inhibitor II (Pin2) gene family of potato and tomato (Peña-Cortês *et al.*, 1988; Farmer

and Ryan, 1990) and two vegetative storage protein genes (VsPA and VspB) of soybean (Mason and Mullet, 1990) are well-characterised.

Pin2 gene family is constitutively expressed in potato tubers and in the early stages of floral development. In tomato, Pin 2 mRNA was expressed in every organ of the adult flowers. Upon mechanical wounding or herbivore feeding, Pin2 mRNA was induced in the foliage in both tomato and potato plants. After a short delay, Pin2 mRNA started to accumulate in the non-damaged leaves of the wounded plants. Pin2 mRNA was at lower level in the systemically induced leaves than in locally wounded leaves (Peña-Cortês *et al.*, 1995).

Using the well established systemically inducible genes, especially Pin in tomato and potato, involvement of jasmonic acid (JA) and plant hormones in wound signal transduction pathway has been established.

Jasmonic acid, its volatile methylester (MeJA) and other conjugated forms, are octadecanoid compounds, derived from linolenic acid in a lipoxygenase-dependent pathway (Farmer, 1994; Creelman and Muller, 1997). A number of lines of evidence suggested that JA was involved in the wound response. Wounding triggers an increase in the endogenous levels of JA (Albrecht *et al.*, 1993; Laudert *et al.*, 1996), which is required for gene activation upon wounding (Peña-Cortês *et al.*, 1993). Exogenous application of JA or its methyl ester induced wound responses such as the expression of wound-responsive genes including Pin2 and Vsp (Mason and Mullet, 1990; Farmer *et al.*, 1992). Mutants with defective JA biosynthesis or in transformed plants with the gene for a key enzyme in JA biosynthesis suppressed, wound-induced gene expression is inhibited. A comparative study of wound responsive genes in wild-type and JA-insensitive mutant *Arabidopsis* plants demonstrated that wound responsive genes which are non-JA responsive accumulated in JA-insensitive mutant upon wounding (Titarenko *et al.*, 1997). They suggested that there were at least two different pathways of wounding signals. One is a JA-independent signalling pathway which is preferentially involved in gene activation in the vicinity of the wound site, whereas the other is a JA-dependent pathway activating gene expression in the aerial part of the plant.

An 18-amino-acid polypeptide, systemin, extracted from tomato leaves, induces systemic expression of Pin genes. Transgenic tomato with antisense prosystemin cDNA did not express Pin genes systemically upon wounding (McGurl *et al.*, 1992). In contrast, over-expression of the gene in transgenic tomato constitutively activated defense genes and caused the plants to act as if they were constantly wounded (McGurl *et al.*, 1994)

ABA was proposed to be the primary signal in the systemic wounding-signalling cascade. The evidence for the proposal was that ABA-deficient mutants of potato and tomato did not respond to wounding with increased accumulation of Pin2, levels of ABA in tomato leaves increased dramatically 6hr after wounding and detached tomato leaves accumulated Pin2 when treated with ABA (Peña-Cortês *et al.*, 1996). However, recent evidence showed that ABA is not a component of the wound-inducible signal transduction pathway (Birkenmeier and Ryan, 1998). When ABA was supplied to excised tomato plants *via* cut stems at concentration up to 100μM only low levels of Pin2 were induced in comparison to the levels induced by systemin, and other wound-response genes such as Pin1, Cys proteinase inhibitor, polyphenol oxidase and prosystemin were weakly induced (Birkenmeier and Ryan, 1998). Accumulation of ABA upon wounding was significant in the tissue near the wounding site but not in distal leaves. Based on these observations, the failure of the ABA-deficient tomato mutants to activate defensive genes upon wounding (Peña-Cortês *et al.*, 1996) might suggest that a certain level of ABA was required for the wounding response (Birkenmeier and Ryan, 1998).

Ethylene is a signal mediating wound responses. Exogenous application of ethylene to tomato plants showed that it did not induce wound-responsive Pin genes (Kernan and Thornburg, 1989). However, ethylene was produced rapidly within 30min to 2 hrs after wounding or application of oligogalacturonide fragments of pectic polysaccharides (OGAs), systemin or JA (O'Donnell *et al.*, 1996). It was also shown that silver thiosulphate blocked the induction of Pin gene expression by wounding (O'Donnell *et al.*, 1996). Transgenic plants expressing antisense gene of ACC oxidase, the last enzyme in ethylene biosynthesis, did not express Pin gene upon wounding (Bowles, 1997). Therefore, endogenous ethylene is involved in the induction of wound-response gene.

Ethylene seemed to be placed downstream of JA in the wound transduction pathway, based on the evidence that application of JA induced ethylene production and ethylene action inhibitors could block the induction of Pin gene expression by JA (O'Donnell *et al.*, 1996). However, it was also shown that JA was significantly reduced by modifying ethylene action or synthesis and that pre-treatment of plants with aspirin blocked completely the induction of JA during wounding response whereas specific inhibition of ethylene synthesis or action reduced the JA level to 20-30% of the wild type. It was then suggested that JA and ethylene were both required for wound induction of Pin gene expression; at least one site of ethylene action in the wound response was the regulation of JA and two processes were related to the wound-induced increase in JA, one of which was ethylene-dependent (O'Donnell *et al.*, 1996).

1.7 Experiment Strategy

Molecular biological studies into the deterioration by referring to the expression of well-studied genes induced in response to wounding or pathogen attack in other plants could not only help to elucidate the mechanism, but also possibly provide approaches to the genetic manipulation of physiological deterioration of cassava roots.

The experimental strategy would be to use clones of important wound-response-related genes from other plants as probes to screen a cassava cDNA library made from physiologically deteriorating roots. The isolated clones can then be used to determine the expression of the corresponding genes during PPD. The proposed genes for the study include genes for PAL, hydroxyproline-rich glycoproteins (HRGPs) and β-1,3-glucanase; therefore these will be described below, mainly in the context of wound responses.

1.8 Phenylpropanoid Metabolism and Phenylalanine Ammonia-lyase

1.8.1 Phenylpropanoid Metabolism and Its Functions

Phenylpropanoid metabolism consists of a core pathway and many branch pathways, leading to the biosynthesis of phenolic compounds. The core pathway refers to the reactions catalysed by PAL, C4H and 4CL. The branch pathways lead to the biosynthesis of lignin, suberin, antimicrobial compounds, salicylic compounds and phenolic

compounds (Fig. 1.3). These phenylpropanoid metabolites, though many are classified as secondary metabolites, play important roles in the development of plant and in the response to biotic and abiotic stresses.

Lignin, a complex phenylpropanoid polymer, is the second most abundant polymer after cellulose. Basic lignin molecule is derived from three monomeric units, the monolignols *p*-coumaryl alcohol, coniferyl alcohol and sinapyl alcohol (Boudet, 1998). The deposition of lignin in the cell walls of supporting and conducting tissues such as fibres and tracheary elements strengthens these tissues due to the mechanical rigidity of lignin. Lignification in response to infection or wounding forms a physical barrier to block the penetration of pathogen or seal the wound (Boudet *et al.*, 1995).

Suberin is a matrix of polyphenolic and polyaliphatic domains in cell walls, and the phenolic domain comprises primarily (poly)hydroxycinnamates (Bernards and Lewis, 1998). Suberin strengthens the cell walls, and suberization during wound healing forms a physical barrier which reduces water loss and enhances defense against pathogens (Kolattukudy, 1980). PAL was so important to suberization that its inhibition by S-carvone or L-α-aminooxy-β-phenylpropionic acid (L-AOPP) could prevent the polyphenolic deposition (Oosterhaven *et al.*, 1995; Street *et al.*, 1986). PAL activity was enhanced 10-15 fold within 12hrs during the wound healing of potato tubers and remained at high levels for several days after wounding (Oosterhaven *et al.*, 1995), beyond the completion of suberization. During the suberization of potato tuber after wounding, the demand for phenylpropanoids was so high that even the shikimic acid pathway leading to the biosynthesis of phenylalanine and tyrosine was activated (Bernards and Lewis, 1998). It was shown recently that separate depositional patterns of two major domains of suberin including phenolic and aliphatic domains had separate roles in the development of resistance to bacterial and fungal infection during wound-healing in potato tuber (Lulai and Corsini, 1998). Total resistance to infection by *Erwinia carotovora* subsp. *carotovora* (causing bacterial soft rot) was observed after the completion of phenolic deposition on the outer tangential wall of the first layer of cells 2-3 days after wounding. However, no resistance to fungal infection by *Fusarium sambucinum* (fungal dry rot) was developed even after phenolic deposition was completed on adjoining radial and inner tangential cell walls. Resistance was developed only after deposition of the suberin aliphatic domain was initiated.

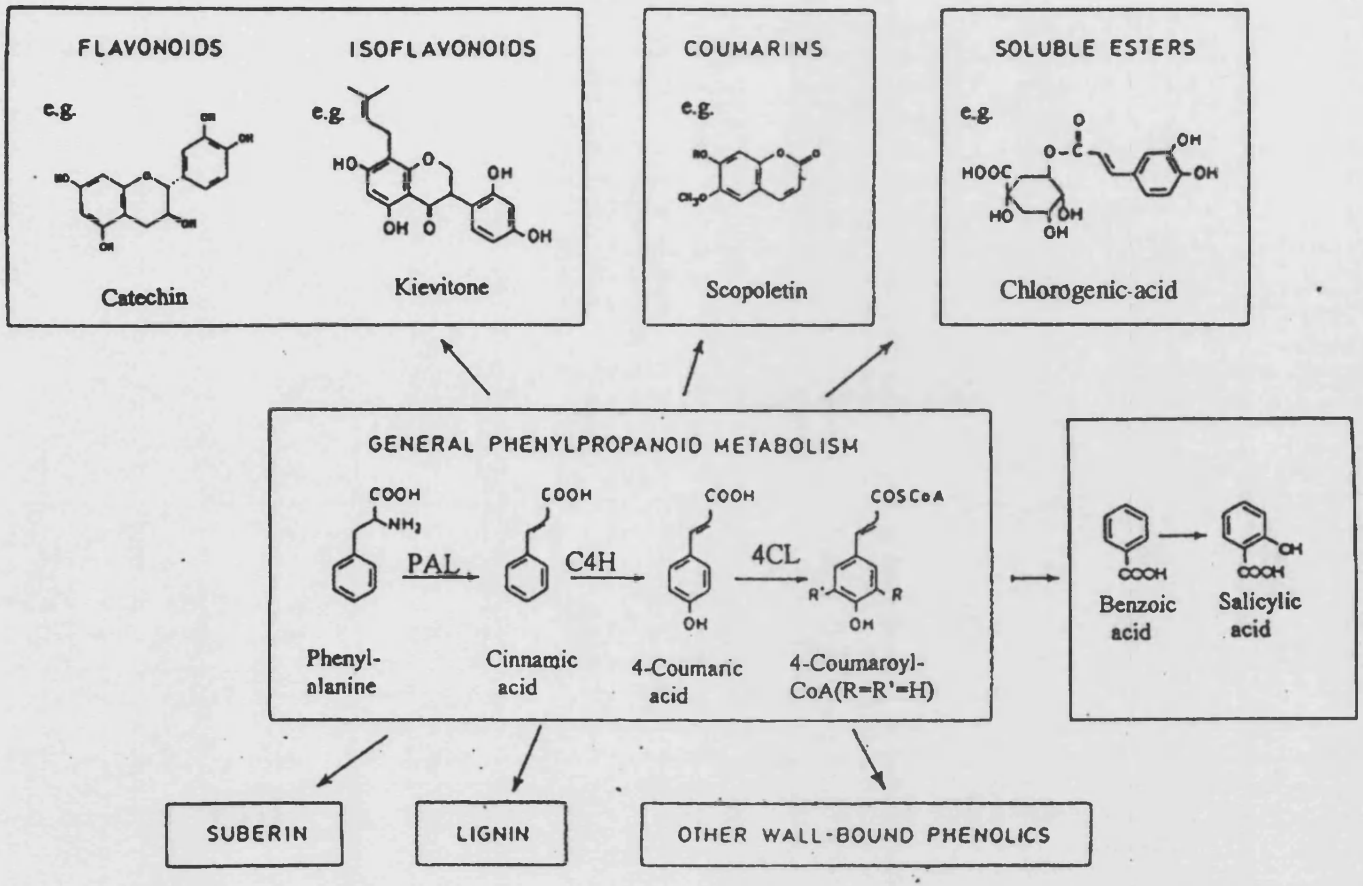


Fig1.3 Schematic illustration of the flux of phenylalanine-derived intermediates from the core reactions of general phenylpropanoid metabolism to major branch pathways (Modified from Hahlbrock & Scheel, 1989). PAL, phenylalanine ammonia lyase; C4H, cinnamic acid-4-hydroxylase; 4CL, 4-coumarate: CoA ligase.

Salicylic acid, synthesised *via* another branch pathway from general phenylpropanoid metabolism, is an important signalling component in plant defence responses. Transgenic tobacco and *Arabidopsis* expressing the *nah*G gene, which encodes salicylate hydroxylase from *Pseudomonas putida*, accumulated little SA and as a consequence showed reduced or no PR gene expression, and failed to establish SAR (Mur *et al.*, 1997). SA can induce the expression of many defense-related genes, potentiate the production of H_2O_2, the induction of cell death and the activation of several genes induced by fungal elicitors and wounding (Thomma *et al.*, 1998; Shirasu *et al.*, 1997).

Many phenylpropanoids are antimicrobial compounds synthesized and accumulated to a high level around the infection site in response to pathogen attack and are classified as phytoalexins. These compounds include pterocarpans (eg. glyceollin), isoflavans, prenylated isoflavonoids (eg. kievitone), stilbenes, psoralens, coumarins, 3-deoxyanthocyanidins, flavonols (eg. quercetin, kaempferol) and aurones (Bennett and Wallsgrove, 1994; Dixon *et al.*, 1995).

Wounding also induces the accumulation of phenylpropanoids. In potato tubers, chlorogenic acid accumulated within 12 days during wound healing (Cottle and Kulattukudy, 1982). After 3-5 days of storage at 5 or 10°C in wounded lettuce midrib sections the phenolic compounds detected were 5-caffeoylquinic acid (chlorogenic acid), 3,5-dicaffeoylquinic acid (isochlorogenic acid), caffeoyltartaric acid and dicaffeoyltartaric acid. Among these compounds chlorogenic acid accumulated to the highest level, with a 5-fold increase at 5°C and 10-fold increase at 10°C after 72hr storage (TomasBarberan *et al.*, 1997). Carrot root slices, stored for 4 days at 20°C, also reacted with a strong accumulation of total phenolics, especially chlorogenic acid (Leja *et al.*, 1997)

1.8.2 Involvement of Scopoletin in Wound and Defence responses

Scopoletin, a phenylpropanoid, accelerated PPD when it was applied to harvested root and different responses to scopoletin was observed between cured roots and roots from pruned plants (Wheatly and Schwabe, 1985). These observation indicated that scopoletin might play a role in the development of PPD.

Scopoletin can accumulate in plants upon wounding or insect or pathogen attack. For example, mechanical damage led to high level of accumulation of scopoletin within 11

days in sunflower (Olson and Roseland, 1991). Also in sunflower leaves, scopoletin concentration increased significantly 22 days after attack by beetles (*Zygogramma exclamationis*) or 11 days after damage by flower thrips (*Frankliniella tritici*). Methyl jasmonate, which can simulate wounding, induced the synthesis of the coumarin, scopoletin and scopolin in tobacco cell culture (Sharan *et al.*, 1998).

Like many other phenolic compounds, scopoletin has antimicrobial properties. Chlorogenic acid, esculin and scopoletin were produced in potato tuber tissue after pink eye (associated with *Verticillium* and *Rhizoctonia*, and the bacterium *Pseudomonas fluorescens*) and corky patch symptoms developed on the periderm of affected tubers. It appears that these compounds were produced in response to pathogen invasion and also as part of wound healing process (Nolte *et al.*, 1993). In the tobacco tissue surrounding the dead cells of hypersensitive reactions elicited by a *Phytophthora megasperma* glycoprotein elicitor, high amounts of scopoletin were synthesized, before pathogenesis-related (PR) transcripts and the corresponding PR proteins accumulated in high amounts in the same tissue (Dorey *et al.*, 1997).

The level of scopoletin accumulation is a balance between its synthesis and its degradation. When scopoletin at 0.005mM was fed to sunflower leaf discs the coumarin was initially slowly metabolized by glycosylation after a delay of 24hr (Edwards *et al.*, 1997). Scopoletin then disappeared rapidly, which was associated with the increased activity of a peroxidase, which metabolized the coumarin to a coloured insoluble metabolite. The scopoletin-peroxidase was purified; its activity increased locally around wound sites and its activity could be elevated by $CuCl_2$ or SA (Edwards *et al.*, 1997). It was also revealed that isoperoxidases from radish can use scopoletin as a substrate (Lee and Kim, 1998).

Sunflower IAA oxidase was inhibited by ferulic acid and scopoletin. The enzyme was not detected in healthy plants but in downy-mildew infected sunflower, indicating the induced oxidase may lead to retardation of the growth of the infected plants. Treatment with exogenous phytohormones revealed a stimulation of elongation growth in downy mildew-infected sunflower with gibberellic acid, but not with auxin (Benz and Spring, 1995). Scopoletin produced during PPD may cause competitive inhibition of auxin oxidase, resulting in a decrease in auxin degradation and hence might lead to an increase

in auxin concentration which could affect root metabolism (Wheatley and Schwabe, 1985).

1.8.3 Phenylalanine Ammonia-lyase

Phenylalanine ammonia-lyase (PAL, E.C. 4.1.3.5.) catalyses the deamination of L-phenylalanine to *trans*-cinnamate and NH_4^+. PAL is the first and key enzyme of the phenylpropanoid pathway for the biosynthesis of phenolic compounds. The subcellular location of PAL is mainly cytoplasmic, may also be loosely associated with the membrane-bound cinnamate hydroxylase or benzoate synthase systems (Hanson and Havir, 1981; Amrhein and Zenk, 1971, as cited by Jones, 1984). PAL genes have been extensively studied in many plants. PAL is encoded by a multigene family with 2 to 6 members in most studied plant systems (Hahlbrock, 1989; Cramer *et al.*, 1989; Lee, 1992; Lois *et al.*, 1992; Subramaniam *et al.*, 1993; Tanaka *et al.*, 1989; Wanner *et al.*, 1995) except that there are over 40 PAL genes in potato (Joos and Hahlbrock, 1992) and only one in loblolly pine (Whetten and Sederoff, 1992). PAL genes have been shown to be regulated developmentally, induced by wounding and pathogen attack (Jones, 1984; Lois, 1992; Collinge and Slusarenko,1987; Joos and Hahlbrock, 1992) and other various stresses (Joos and Hahlbrock, 1992; Edward *et al.*, 1985).

The *de novo* synthesis of PAL mRNA was rapidly induced by wounding in potato tubers (Ishizuka *et al.*, 1991). A sharp and high peak of PAL mRNA was detected at the early stage (0-6hr) of the wound response. Sharp expression peaks of three PAL genes were also detected 2h after wounding in potato leaves, which returned to original levels about 6h later (Joos and Hahlbrock, 1992). Specific inhibition of PAL with 2-aminoindan-2-phosphonic acid (AIP) caused depression of lignin accumulation associated with secondary wall thickening in isolated *Zinnia* mesophyll cells (Nakashima *et al.*, 1997). S-carvone inhibited PAL activity and wound healing in potato tubers, and PAL activity preceded the appearance of the first suberized cell layers after S-carvone was removed (Oosterhaven, 1995). In tomato plants resistant to a wilt fungus *Verticillium albo-atrum*, one of the earliest responses to the infection is the rapid suberin coating in the xylem, which is highly correlated with a rapid increase in PAL; whereas there was a substantial suppression of PAL mRNA and activity and significantly less vascular coating in susceptible plants (Lee, *et al.*, 1992). However, since *trans*-cinnamate the product of PAL activity can lead to several branch pathways, the increase in PAL activity may affect

only some of the pathways but not others and may be specific to a certain defense or wound response. Therefore, increased PAL may not be associated with resistance to disease. For example, PAL activity increased regardless of the resistance or susceptibility of barley cultivars to powdery mildew pathogen *Erysiphe graminis f.sp. hordei* (Shiraishi *et al.*, 1995). Similarly, increased PAL activity may be suggested to be essential to wound healing but does not necessarily lead to it. During wound healing of potato tubers the development of suberin coating was related to PAL activity (Oosterhaven *et al.*, 1995). But in cassava there is little wound healing response though there is a rapid increase in PAL during the storage of wounded roots (Tanaka, 1983; Uritani, 1983).

1.9 Hydroxyproline-rich glycoprotein

Plant cell walls provide mechanical support and shape to the cells and to the whole plant. The cell walls also have important physiological roles such as transport, absorption, secretion, control of cell growth, protection against pathogen and predator attack and cell to cell communication (Brett and Waldron, 1996)

The complicated plant cell wall may be visualised as consisting of four distinct polymer networks each interacting with each another. Cellulose microfibrils coated with hemi-cellulose mainly xyloglucan constitute the first network, which is embeded in the second network, a matrix of pectin. The hydroxyproline-rich glycoproteins and lignin form the third and the last network respectively (Brett and Waldron, 1996).

Hydroxyproline-rich glycoproteins (HRGPs) are the most abundant among the plant cell wall proteins. The HRGPs in the cell walls of plants are also termed as extensins, which was originally coined to suggest their role in cell wall extension (Lamport, 1967). HRGPs and four classes of other cell wall abundant proteins, glycine-rich proteins (GRPs), proline-rich proteins (PRPs), solanaceous lectins and arabinogalactan proteins may be evolutionarily related to one another in that they are similar in nucleotide sequence or / and rich in hydroxyproline (Showalter, 1993).

HRGPs are particularly abundant in dicotyledonous plants and have very similar characteristics in these plants. They are rich in hydroxyproline which is generated by post-translation hydroxylation of proline, and also rich in serine and in some combination of Val, Tyr, Lys, and / or His. Most of the hydroxyproline residues are glycosylated with

oligo-arabinosides through O-glycosidic bonds (Varner and Lin, 1989) and serine residues with single galactose residues. The most distinctive aspect is that the amino acid sequences in almost all studied dicot HRGPs contain highly repetitive Ser-(Pro)4 pentapeptide blocks (Showalter, 1993), with a few exceptions such as two HRGPs characterised in sugar beet and soybean wich lack repetitive tetrahydroxyproline blocks(Li *et al.*, 1990; Ahn *et al.*, 1996). HRGPs also contain various other repetitive peptide motifs, which may play important roles as functional sites during the process of cell wall assembly, X-(Hyp)n including Ser-(Hyp)4 for molecular rigidity, Pro-Hyp-Val-Tyr-Lys and variants for putative intermolecular cross-links, adhesion, cohesion and possible beta-turns, Tyr-X-Tyr-Lys for intramolecular and possible intermolecular isodityrosine (IDT) cross-links increasing rigidity and hydrophobicity (Kieliszewski and Lamport, 1994)

HRGPs have been proposed to strengthen cell walls and control cell wall extension by insolubilization of the molecules through intermolecular cross-links such as isodityrosine cross-links. The deposition of extensins and glycine-rich proteins in cell walls are related to lignin deposition and secondary wall thickening (Cassab and Varner, 1988). The importance of HRGPs in cell wall structure indicates their physiological functions in plant development and defence against pathogen attack. So far, many HRGP genes have been isolated from many plant species.

The expression of HRGP genes usually takes place in roots and stems (Showalter, 1993; Tire *et al.*, 1994). HRGPs have been shown to be developmentally regulated in a tissue-specific way and these genes are expressed mainly in meristems or in tissues rich in dividing cells in soybean and maize (Ye and Varner, 1991; Menossi *et al.*, 1997). The expression of HRGP genes have also been detected in flowers, namely in ovaries (Hirsinger *et al.*, 1997), styles and transmitting tissues in *N. alata* (Chen *et al.*, 1992; Lind *et al.*, 1994), *Antirrhinum* (Baldwin *et al.*, 1992), and *N. tabacum* (Goldman *et al.*, 1992). Extensin mRNAs are also accumulated in cell suspension cultures in response to elicitor treatment (Showalter *et al.*, 1985; Lawton and Lamb, 1987; Kawalleck *et al.*, 1995). It was demonstrated that extensin genes in tobacco and maize were induced by ethylene (Memelink *et al.*, 1993; GarciaMuniz *et al.*, 1998).

HRGPs were found to be inducible by wounding in bean (Corbin *et al.*, 1987), maize (Ludevid *et al.*, 1990), tobacco (Memelink *et al.*, 1993; Hirsinger *et al.*, 1997), *Nicotiana plumbaginifolia* (Tire *et al.*, 1994), *Nicotiana sylvestris* (Parmentier *et al.*, 1995), soybean (Suzuki *et al.*, 1993), and *Brassica napus* (Shirsat *et al.*, 1996). It has been shown that proline-rich protein was rapidly insolubilised in the cell walls in wounding sites (Bradley *et al.*, 1992). It was found that the wound healing process could be impaired by gamma irradiation in potato through its influence on wound-induced signal transduction systems and subsequent Ca^{2+} dependent protein phosphorylation for the activation of the HRGP gene (Ussuf *et al.*, 1996). After gamma irradiation, the wound periderm formation of potato was suppressed, rendering the tubers susceptible to pathogens. HRGP gene expression upon wounding was 30% higher in gamma irradiated potatoes than the untreated during the initial stage of storage. Three to five weeks post-irradiation, HRGP gene expression in the treated samples upon wounding was significantly lower than the control. Treatment with 5mM Ca^{2+} retrieved the low expression of HRGP; whereas calcium channel blocker – trifluoperazine reduced the HRGP expression in the control. A significant higher *in vivo* protein phosphorylation upon wounding was observed in irradiated and Ca^{2+} treated potato after 5 week storage than the irradiated potatoes.

HRGPs also play their roles in plant protection against pathogens. It was observed that the induction or inhibition of HRGP deposition could lead to increased or decreased plant disease resistance respectively (Esquerré-Tugayé *et al.*, 1979). It was reported that HRGP increased earlier and to a greater extent in resistant than in susceptible cultivars in tomato; whereas the increased HRGPs in the compatible interaction appeared to be from cell wall damage, HRGPs in the incompatible interaction accumulated in the walls of uninvaded cells (Benhamou *et al.*, 1991). Using immunoelectron microscopy, threonine-hydroxyproline-rich glycoprotein (THRGP) was found to be deposited in the extrahaustorial matrix in both incompatible and compatible plant-pathogen interaction, suggesting the formation of a modified physiological barrier with localised deposition of HRGPs in response to infection (Hippesanwald *et al.*, 1994).

1.10 β-1,3-glucanase

Endo-β-1,3-glucanases (BGlu; EC 3.2.1.39) are hydrolases and many β-1,3-glucanases have been purified and characterised (Boller, 1988). They are abundant proteins widely

distributed in seed-plant species (reviewed by Meins et al., 1992). They are usually monomers with molecular mass from 25-35 kDa (Stintzi, 1993). The substrate for the enzyme is (1,3)- and (1,3, 1,6)-β-glucans, a major cell wall constituents of some plants (especially *Gramineae*) (McNeil, 1984) and of common fungal pathogens (Wessels and Sietsma, 1982). β-1,3-glucanases are widely distributed in higher plants and considered as important representatives of the pathogenesis-related (PR) proteins (Dixon and Lamb, 1990). Another class of abundant β-glucan endohydrolases is 1,3-1,4-β-glucanases which hydrolyze internal 1,4-β-glucosyl linkages only in 1,3-1,4-β-glucans. These glucanases are essential for the deploymerization of plant cell wall β-glucans in germinating seeds and in young vegetative tissues (Slakeski *et al.*, 1990).

β-1,3-glucanases are known to accumulate developmentally during cell division (Waterkeyn, 1967, as cited by Beffa and Meins, 1996), microsporogenesis (Worrall, 1992), pollen germination (Roggen and Stanley, 1969), fertilization (Ori *et al.*, 1990) and seed germination (Woodward and Fincher, 1982). Some are induced under tissue culture conditions. In tobacco, β-1,3-glucanases accumulated under the control of auxin and cytokinin (Eichholz *et al.*, 1983; Felix and Mains, 1985). A β-glucanase was shown to be regulated hormonally in *N. plumbaginifolia* cell suspensions, accumulating with auxin only in the medium and decreasing after the addition of cytokinins (de Loose *et al.*, 1988).

β-1,3-glucanase gene expression can be induced by pathogen infection and wounding (Brederode *et al.*, 1991; Simons *et al.*, 1992). In TMV-infected tobacco leaves, different isoforms of β-1,3-glucanase were differentially regulated by tobacco mosaic virus infection at mRNA levels (Ward *et al.*, 1991). The mRNAs of β-1,3-glucanase PR-2 and PR-Q' were detectable by Northern analysis within 2d after inoculation, peaked 4 to 6 days and then decreased to basal level after 14 d. These two genes were coordinately expressed during the infection. But another β-1,3-glucanase was expressed with a different pattern, which was detectable in uninfected leaves and increased slightly during the infection. Yoshikawa *et al.* (1981) found that a soluble elicitor of phytoalexin glyceollin was released from insoluble mycelial walls of *Phytophthora megasperma* var. *sojae* 2 min after incubation with soybean cotyledon. The factor releasing elicitor from mycelial cell walls was identified as β-1,3-glucanase in soybean (Ham *et al.*, 1991). It

was proposed that β-1,3-glucanase induced by pathogen infection or chemical stress, functions in defense through breaking down fungal cell walls and releasing a phytoalexin elicitor from the mycelial walls of a pathogenic fungus. The cDNA of this β-1,3-glucanase was cloned and the translatable β-1,3-glucanase mRNA (detected by in vitro translation and immunoprecipitation) increased 50 – 100 fold in soybean cotyledons upon ethylene treatment, suggesting its strong regulation by ethylene (Takeuchi *et al.*, 1990).

Glucanase genes either in their sense or anti-sense orientations or in combination with chitinase genes have been transferred into different plant systems in order to study their functions. Transformation with anti-sense basic vacuolar β-1,3-glucanase in *Nicotiana sylvestris* effectively blocked the constitutive and induced expression of the corresponding gene but not of other isoforms, without increasing the susceptibility to *C. nicotianae* infection (Neuhaus *et al.*, 1992). But it was observed in transgenic tobacco that constitutive expression of a β-1,3-glucanase cDNA from tobacco increased the resistance to glucan-containing fungi *Peronospora tabacina and Phytophthora parasitica* var. *nicotiana* while the introduction of anti-sense glucanase cDNA resulted in decreased resistance (Lusso and Kuc, 1996). Constitutive over-expression of an alfalfa acidic glucanase in transgenic alfalfa did not show significant protection against leaf pathogens *Phoma medicaginis* or *Stemphylium alfalfae* or the leaf and stem anthracnose fungus *Colletotrichum trifolii*, but exhibited significant increase in tolerance to infection with the oomycete *Phytophthora megasperma f sp. medicaginis* (Masoud *et al.*, 1996). Transgenic tobacco plants expressing a β-1,3-endoglucanase from soybean showed increased resistance to an oomycete fungus, *Phytophthora parasitica var.nicotianae* (Yoshikawa *et al.*, 1993).

Constitutive co-expression of chitinase and glucanase in transgenic plants can increase the protection against fungal pathogens. Transgenic tobacco with genes encoding basic chitinase from rice and acidic glucanase from alfalfa (Zhu *et al.*, 1994) or with genes encoding chitinase, glucanase and plant secretion peptide ribosome-inactivating protein (spRIP) from barley (Jach *et al.*, 1995), or transgenic tomato with tobacco chitinase and glucanase genes showed substantially greater protection against the fungal pathogen *Cercospora nicotianae*, and soil-borne fungal pathogens *Rhizoctonia solani* and *Fusarium oxysporum f.sp. lycopersici*, respectively, than either transgene alone.

However, it was also reported that constitutive expression of the combination of basic chitinase and basic glucanase did not affect the colonization of the mycorrhizal fungus *Glomus mosseae* on tobacco (Vierheilig *et al.*, 1995).

1.11 Aims of the project

The physiology and biochemistry of PPD has been studied but the knowledge of the molecular processes of PPD is almost blank. The first aim of the project is to construct a cassava physiological-deterioration-related cDNA library. By screening the cDNA library, clones corresponding to genes for the above enzymes or proteins would be isolated and characterised. To elucidate the mechanism of physiological deterioration, the expression patterns of these clones would be studied during the development of physiological deterioration of cassava roots. The final aim of the project is, through the above studies, to identify the genes that are of importance to the development of PPD; eventually these genes may be used for the genetic manipulation of physiological deterioration of cassava root. In addition, the characterised clones could also be used for the genomic mapping of cassava in CIAT.

Chapter Two

Materials and Methods

2.1 Plant Materials and Methods

Cassava (*Manihot esculenta* Crantz) cultivar M NGA 1 cuttings was grown in pots (30cm diam) in a glasshouse, 25°C day and 15°C night, with a photoperiod of at least 16 hours a day. After about 6 to 9 months, storage roots were harvested and rinsed with tap water, washed with 1% sodium hypochlorite solution for 1 minute and rinsed briefly with autoclaved distilled water. The roots were then cut into 2cm transverse slices with a sterile blade. The root slices were then incubated in a sterile container at 27+3°C, 55-65% relative humidity (RH) and in the dark. Roots (cultivar Mcol 22) from CIAT were sealed in wax soon after harvest and air-delivered to UK within about a week.

2.2 RNA Extraction

Extracting RNA from cassava storage root was proved to be problematic and different protocols were therefore tried and used during the various stages of this research.

All the glassware used for RNA work was baked at 180°C overnight. All the solutions except Tris was treated with 0.1% DEPC for 1hour and autoclaved.

2.2.1 RNA for constructing cDNA library

Method A: Root slices were grated into liquid nitrogen, ground into fine powder and dispersed into a 1:1 mixture (2ml/gFW) of RNA extraction buffer (100 mM Tris-HCl pH9, 100mM LiCl, 10mM EDTA and 1%SDS) and phenol, which was pre-warmed to 90°C. The mixture was vortexed for 5min and centrifuged at 10,000g for 30min. The aqueous phase was extracted with chloroform twice and the aqueous phase then transferred to a centrifuge tube, to which 8M LiCl was added to final concentration 2M. RNA was precipitated by incubating at 4°C and centrifuging at 12,000g for 20min at 4°C. The pellet was washed with 70% ethanol and then resuspended in 300µl 0.3M sodium acetate pH5.2. After chloroform extraction, the RNA was then precipitated by adding 2.5vol ethanol and incubating at -20°C for at least 30min. The pellet was washed with 70% ethanol, vacuum-dried, dissolved in water and kept at - 80°C (Pawlowski *et al.*, 1994).

Method B: Root slices were grated into liquid nitrogen, ground into fine powder and dispersed into a mixture of RNA extraction buffer (1.5ml/g FW of 100 mM Tris-HCl pH9, 100mM NaCl, 2mM EDTA, 1%SDS), a 1:1 mixture of phenol / chloroform pH8 (1.5ml/g FW) and 2-mercaptoethanol (0.33ml/g FW). The mixture was incubated at room temperature for 60 minute and centrifuged at 12,000xg for 30 minutes. The supernatant was extracted twice with phenol/chloroform and once with chloroform. Then 8M LiCl was added to a final concentration of 2M and the RNA was precipitated overnight at 4°C and collected by centrifuging at 12,000xg for 30 minute. The resultant RNA pellet was washed twice with 2M LiCl and twice with 70% ethanol, air-dried and dissolved in water.

2.2.2 Total RNA of leaf, stem, sap and vegetative root

Total RNA from these tissues was extracted using RNeasy Plant Mini Kit (Qiagen) according the instructions from the manufacturer.

2.2.3 Total RNA of storage roots for expression studies using RT-PCR

Method C: Fine powder of cassava obtained as above was first extracted with 1:1 extraction buffer (100mM Tris-HCl pH8.5, 100mM NaCl, 20mM EDTA, 1% w/v SDS, 10% 2-mercaptoethanol, 2% w/v PVP) and phenol without incubation, followed with similar procedures as RNA preparation for cDNA library construction as described above except that an isopropanol precipitation step was added before LiCl precipitation. The total RNA was cleaned with RNeasy Plant Mini Kit (Qiagen) before being used for expression study.

2.2.4 Total RNA used for Northern blots

Method D: The first extraction was performed with 25:24:1 extraction buffer (100mM Tris-HCl pH7.5, 2M NaCl, 25mM EDTA, 1% w/v sodium lauroylsarcosine or SDS, 2% w/v PVP and 2% 2-mercaptoethanol just before use) / chloroform / isoamylalcohol at 8ml/gFW. The extraction mixture was centrifuged at 6300g to 10,000g for 20min at room temperature. The aqueous phase was then extracted twice with chloroform. The aqueous phase was transferred to another centrifuge tube, to which 0.33vol of 8MLiCl was added to final concentration 2M and incubated at 4°C overnight. The RNA was precipitated by centrifuging at 7,800g to 10,000g for 20min at 4°C. The pellet was then

dissolved in 500µl water and transferred to 1.5ml Eppendof centrifuge tubes. Chloroform extraction and phenol/chloroform extraction was carried out once respectively. The aqueous phase was transferred to another centrifuge tube, to which 0.25vol 5M NaCl and 2vol ethanol were added to precipitate RNA. After 30min incubation at -80°C, the RNA was precipitated by centrifuging at 10,000g for 20min at 4°C. The pellet was washed with 70% ethanol, air-dried, dissolved in 50 µl water and kept at -80°C.

Of all the methods tried, method D produced highest yield of total RNA (150µg/g FW leaf or 50 µg/g FW root) of good quality. The hot mixture of phenol and extraction buffer in Method A resulted in swollen starch, which greatly reduced the efficiency of the extraction. The significant increase in extraction efficiency in method D may result from two factors. One was the high concentration of sodium chloride in the extraction buffer, which helped to precipitate the carbohydrate effectively. Another was the chloroform extraction, which was much more powerful or effective than phenol extraction or phenol/chloroform extraction during the RNA extraction from cassava tissue (observed by the author). It was also observed that using chloroform in the first extraction mixture allowed more starch fraction to be mixed into the solution than using phenol or phenol/chloroform.

2.3 cDNA Library Construction

2.3.1 Messenger RNA isolation

Messenger RNA was isolated by using Oligotex™ suspension from Qiagen. The RNA solution was diluted to a 500µl final volume with MilliQ water and mixed with 500µl 2 x binding buffer (20mM Tris-HCl, pH 7.5, 1000mM NaCl, 2mM EDTA, 0.2% SDS) and 30µl Oligotex suspension. The bottom of the tube was gently flicked and the tube incubated for 3 minute at 65°C in a water bath. Then it was incubated for 10 minute at room temperature and centrifuged for 2 minutes at top speed of tabletop centrifuge and the supernatant was removed by aspiration. The Oligotex pellet was washed twice with 1ml of wash buffer (10mMTris-HCl, pH 7.5, 150mMNaCl, 1mM EDTA). Each time the mixture was vortexed to resuspend the pellet and spun at top speed for 2 minute. The mRNA was eluted twice with 20µl of preheated (70°C) elution buffer (5mM Tris-HCl, pH 7.5). Each time the mixture was vortexed to resuspend the pellet, spun down at top

speed for 2 minutes and the supernatant of mRNA was removed to a fresh Eppendof tube.

2.3.2 cDNA synthesis from messenger RNA

(Amersham kits were used)

2.3.2.1 First strand cDNA synthesis

The first strand cDNA synthesis reaction was set up in a microcentrifuge tube on ice.

First strand synthesis reaction buffer	4µl
(250mM Tris-HCl pH 8.3, 250mM KCl, 50mM $MgCl_2$)	
Sodium pyrophosphate solution (80mM)	1µl
Human placental ribonuclease inhibitor (HPRI)(20units/µl)	1µl
dNTP mix(10mM each dATP, dGTP, dTTP, 5mM dCTP)	2µl
Random hexanucleotide primers (70µM)	2µl
mRNA (1µg)	9µl

The components were mixed gently and spun for a few seconds in a microcentrifuge. Then 1µl of AMV reverse transcriptase (20 units/µl) was added. The mixture was incubated at 42°C for 60 minutes and then the first strand synthesis reaction was placed on ice.

2.3.2.2 Second strand cDNA synthesis

To the first strand cDNA synthesis reaction on ice, the following components were added in order.

Standard reaction mix for 1µg original mRNA	
First strand cDNA reaction	20µl
Second strand synthesis buffer	40µl
(80mM Tris-HCl pH7.5, 200mM KCl ,10mM MCl_2, 0.13mg/ml BSA)	
[α-^{32}P] dCTP (20µ Ci)	2µl
Ribonuclease H (0.8 units)	1µl
DNA polymerase I (23 units)	6.6µl
Water to a final volume of 100µl	30.4µl

The components were mixed gently and then incubated sequentially at 12°C for 60 minutes, then 22°C for 60 minutes, and finally at 70°C for 10 minutes. Then the mixture was spun for a few seconds in a microcentrifuge and placed on ice. Two units T4 DNA polymerase per microgram of original mRNA template was added, mixed gently and incubated at 37°C for 10 minutes. The reaction was stopped by adding 4µl of 0.25M EDTA pH8.0 per 100µl of final reaction mix.

2.3.2.3 Purification of double-stranded cDNA

To a completed standard second strand synthesis reaction one volume of phenol/chloroform (1:1) was added and vortexed briefly to form an emulsion. The mixture was spun for 1 minute in a microcentifuge to separate the phases and the aqueous (upper) phase was transferred into a fresh tube. The phenol/ chloroform extraction was repeated. One volume of chloroform was then added to the aqueous phase and vortexed briefly. It was spun for a few seconds and the aqueous phase transferred into a fresh tube, to which one volume of 4M ammonium acetate pH5.8 and twice the combined volume of 100% ethanol was added. The DNA was precipitated at -20°C for 30minute and then warmed to room temperature with gentle shaking (to dissolve precipitated, unreacted nucleotides). It was then spun for 10 minutes and the supernatant removed. The pellet was washed twice by adding 50µl 2M ammonium acetate and 100µl 100% ethanol (-20°C). To the washed pellet 200µl 100% ethanol (-20°C) was added. It was spun for a few minutes and the supernatant removed. The pellet was dried in a vacuum desiccator for 2-3 minutes. The cDNA was resuspended in TE buffer (10mM Tris-HCl, 1mMEDTA, pH8.0) and stored at -15°C to -30°C.

2.3.3 cDNA cloning from messenger RNA

(Amersham kits were used)

2.3.3.1 Ligation of adaptors to double stranded cDNA

To set up the adaptor ligation reaction, the components below were added in the order to a sterile 1.5ml microcentrifuge tube.

cDNA (1µg)	8.5µl
Ligase/kinase buffer (L/K-buffer)	2µl
Enzyme enhancer	5µ
*Eco*RI adaptors (100pmoles/µl)	2.5µl

The components were mixed gently and spun for a few seconds in a microcentrifuge. T4 DNA ligase (2.5 units/μl) 2μl was added, mixed gently and the reaction was performed at 16°C for 30 minutes, after which it was stopped by adding 2μl of 0.25M EDTA. STE buffer (TE buffer containing 0.1MNaCl) was added to give a final volume of 100μl.

2.3.3.2 Column purification/size fractionation of 'adapted' cDNA

The column was removed from storage at -20°C and allowed to thaw completely at room temperature for approximately 20 minutes. The column was inverted several times when thawed, to resuspend the matrix prior to use. It was then placed upright in a rack and allowed to drain fully, but not dried out. STE buffer 2ml was added to the column, which was then inverted several times to resuspend the gel. The caps were removed and the column allowed to drain. The STE equilibration was repeated twice more, using an additional 2ml of STE buffer for each wash. When the meniscus just reached the top of the gel bed, the column was placed in a 15ml conical centrifuged tube, centrifuged for 2 minutes at 400x g in a swing-out bucket rotor. Ligated sample 100μl was slowly and very carefully applied to the centre of the flat surface at the top of the compacted bed. The column was then placed in a clean 15ml conical centrifuge tube and spun again for 2 minutes at 400xg. The collected column elute containing 'adapted' cDNA was transferred to a clean microcentrifuge tube.

2.3.3.3 Phosphorylation of 'adapted' cDNA

The following components were mixed together in a sterile 1.5ml microcentrifuge tube:

Collected column sample	100μl
L/K buffer	20μl
T4 polynucleotide kinase (32 units)	4μl
Water to a final volume of	200μl

The mixture was incubated at 37°C for 30 minutes. The mixture was then extracted twice with an equal volume of phenol: chloroform 1:1, followed by chloroform: isoamyl alcohol 24:1 (twice). The DNA was ethanol precipitated by adding 0.1 volumes of 3M sodium acetate and 2.5 volumes of ice cold 100% ethanol, at -20°C for at least 2 hours. The precipitated DNA was spun down for 20 minutes at top speed in a microcentrifuge. Supernatant was removed and the pellet was washed by adding 1ml ice cold 70%

ethanol. After the supernatant was removed, the cDNA pellet was dried briefly and resuspended in 20µl TE buffer.

2.3.3.4 Ligation into λgt10 vector arms

Ligation reactions were set up on ice as listed in Table 2.1 omitting T4 DNA ligase. After the components were mixed gently, 1µl of T4 DNA ligase was added, mixed gently and incubated at 16°C for 30 minutes in a water bath. The tubes containing λgt10 arms + insert were warmed at 37°C for 2 minutes prior to the addition of ligase.

Table 2.1 Ligation reactions

	insert DNA	Whole λgt10	λgt10 arms	L/K-buffer	Water	T4 DNA ligase
Tube 1	-	5µl(0.5µg	-	1µl	3µl	1µl(2.5 units)
Tube 2	-	-	2µl(1µg)	1µl	6µl	1µl(2.5 units)
Tube 3	100ng EcoRI-ended control DNA	-	2µl(1µg)	1µl	-	1µl (2.5 units)
Tube 4	15-50ng 'adapted' cDNA	-	2µl(1µg)	1µl	-	1µl(2.5 units)

2.3.4 *In vitro* Packaging of Ligation Mixtures

For each packaging reaction to be performed, one blue tube (extract A) and one yellow tube (extract B) were removed from the packaging module (stored at -70°C) and placed on ice immediately. As soon as the extracts were thawed, 10µl from a blue tube was added to the ligation reaction. Immediately 15µl of the contents of a yellow tube was transferred to the reaction tube and mixed gently using the tip of the pipette. The mixture was incubated in a 20°C water bath for 2 hours, after which pre-cooled (4°C) 470µl SM buffer (5.8g NaCl, 2g $MgSO_4.7H_2O$, 6.05g Tris base and 5ml gelatin in 1000ml, pH7.5) was added to the packaging mixture. Finally 10µl chloroform was added and mixed gently. The cDNA library was stored at 2-8°C. Amplification of the PPD-related cDNA library was carried out by incubating the mixture of 30µl of the phage and 270µl of the plating cells at 37°C for 15min, plating the mixture together with warm 0.8% agarose on to LB plate, incubating overnight, and harvest the phage with SM buffer.

2.3.4.1 Preparation of phage plating cells

The host *E.coli* L87 and *E.coli* NM514 were streaked on to separate LB-agar plates and incubated overnight at 37°C. A single colony from each LB-agar plate of both strains was inoculated into 10ml of M-broth and incubated with shaking at 37°C overnight. The overnight culture 1ml was added to 50ml of pre-warmed growth medium LB or M-broth (20% maltose was added to LB to a final concentration of 0.4%) (if the cells were for liquid lysate) and incubated at 37°C with vigorous shaking until the cells had grown to an OD_{600} of 0.5 (2.5×10^8 cells/ml). The culture was then cooled on ice and centrifuged at 3000rpm (Sorvall SS34 rotor) for 10 minutes 4°C. The pellet was then resuspended in 15ml of ice cold sterile 10mM $MgSO_4$, mixed thoroughly and these plating cells were stored on ice or at 2-8°C.

2.3.4.2 Titration of λgt10 recombinants

The final packaged phage mixture 30µl was added to 270µl of SM buffer (10^2 dilution), mixed well on ice with a fresh pipette tip. Then 30µl of the 10^2 dilution was added to 270µl SM buffer (10^3 dilution). This was repeated using the previous dilution until the 10^7 dilution was reached. A duplicate dilution series was prepared. A series of tubes for each of the two cell types (L87 and NM514) was prepared by aliquoting 100µl of each type of plating cells. To the appropriate tubes for both L87 and NM514 series of cells 100µl of each phage dilution (10^2 to 10^7) was added to form a plating mix. The mixture was incubated at 37°C in a water bath for 15 minutes. To each plating mix 4ml of warm (47°C) liquid M-top agar was added, mixed quickly and poured on to a pre-warmed L-agar plate. The plates were incubated (inverted) at 37°C overnight. The phage titre/ml was calculated by multiplying the total number of plaques by the dilution number. Cloning efficiency was expressed as pfu (plaque forming uints)/µg insert DNA.

2.3.5 Plaque stock and plate lysate

Single plaques (either from screening or for checking inserts) or positive areas from the first screening were cored out and each was put into a microcentrifuge tube with 0.5ml pre-cooled SM buffer. The tubes were then slightly shaken in cold room (4 to 8°C) for 2 hours. Chloroform 10µl was added to the stock and it was stored at 4°C.

Using the selected stock, approximately 10^5 pfu was adsorbed on fresh NM514 plating cells and plated out, used L-top agarose. After incubation at 37°C for 8 hours to overnight, 4ml of SM buffer was added to the best of the plate, which was gently shaken on a tray at 2-8°C for 2-3 hours. The SM plate lysate was then removed into microcentrifuge tubes. A few drops of chloroform was added to the pooled SM butter and mixed briefly. The lysate was then spun at 4000 rpm for 10 minutes at 4°C and the supernatant was transferred to fresh tubes to be stored at 2-8°C.

2.3.6 Mini-plate lysate and Mini-liquid lysate

These techniques were developed to facilitate the screening process for positive clones from cDNA library and solve the problem of PCR amplification of the inserts in lambda clones. To isolate a single positive plaque, a secondary screening usually was performed. But sometimes the positive plaques were not well separated from other plaques; in this case a third screening had to be carried out. Mini-plate lysate was designed to eliminate the third screening. The plaque stock from a single positive plaque (as described in 2.3.5) was not concentrated enough to be used as templates for PCR amplification of the inserts in the lambda clones, so a plate lysate had to be carried out to get more concentrated phage solutions. However, it was time-consuming and the plate lysate would be of no use afterwards if this particular plaque proved to be false positive. Mini-liquid lysate was designed to prepare efficiently concentrated lysate in a small volume, which is especially very convenient when there are several positives in the first screening.

Mini-plate lysate: Mixture of 100µl plating cells and 3ml top agarose (45-47°C) was poured onto a LB plate. After the top agarose was set, the reverse side of the plate was marked into many square panels (one panel for one positive clone) depending on how many positives were to be used. Phage particles were inoculated from a stock plate (of secondary or third screening) to the LB plate using sterile a yellow tip or p10 tip to touch the positive plaque and then the top agarose of the LB plate. This step was repeated 3 or 4 times for each positive clone on the same panel to form a cluster of plaques of the same clone. The plate was then incubated at 37°C overnight. The plaques were then transferred into an Eppendof tube containing 500µl of SM buffer by coring out the top agarose of the plaque area. The phage particles were eluted into the buffer by gently shaking the tube in cold room. The lysate was concentrated enough to be used as template for PCR amplification.

Mini-liquid lysate: Mixture of 10µl of plating cells and 10µl of the initial elute of a plaque (from the secondary or third screening) was prepared and incubated at 37°C for 15min. Then 80µl of LB was added. The mixture was incubated overnight at 37°C and centrifuged at 4000g for 5min. The supernatant was ready to be used in PCR to check the cDNA insert of the clone.

2.3.7 PCR amplification of the inserts in lambda clones

Specific λgt10 forward and reverse primers 5'd(AGCAAGTTCAGCCTG GTTAAG) 3' and 5'd(CTTATGAGTATTTCTTCCAGGGTA) 3', were used to amplify insert sequences present in isolated λgt10 clones. These primers, anneal to sites directly flanking the *Eco*R I cloning site in λgt10 enabling amplification of inserts up to 2kb in length.

The PCR reaction mixture was prepared as follows: For 12 reactions, 72µl of distilled water, 12µl 10x Taq buffer, 12µl of 2mM dNTP mix, 12µl each of 1µM forward and 1µM reverse λgt10 primers, 3.6µl 50mM $MgCl_2$, 15 units Taq polymerase (Bioline). For each reaction 1µl eluted phage stock was added to 9µl of the PCR mixture and mixed gently. The concentration of phage in the phage stock from a single plaque was usually found to be too low, so mini-plate lysate and mini-liquid lysate were developed to get concentrated phage (see 2.3.6). The mixture was then covered with a drop of mineral oil. PCR reaction was carried out in a thermal cycler with the following programme: 94°C, 3min; 35 cycles of 94°C for 1 minute, 50-52°C for 1 minute, 72°C for 3 minutes; completed by 1 cycle of 72°C, 5 minutes. Ten microlitres of PCR product was analysed on a 1% agarose gel containing ethidium bromide.

2.4 DNA Extraction

2.4.1 Plasmid DNA extraction and purification

Plasmid DNA was extracted using mini-preparation as described by Sambrook *et al.* (1989). Plasmid DNA used for sequencing was extracted with Wizard Plus SV Minipreps DNA purification system (Promega) according to the instructions from the manufacturer.

2.4.2 Lambda DNA extraction

Method A: One milliliter of overnight culture of LE392 in NZYCM (5g NaCl, 5g bacto-yeast extract, 2g MgSO4, 11g Casein hydrolysate enzymatic in 1 litre) was pipetted into 150ml NZYCM of 250ml flasks and incubated at 37°C, 120rpm until OD_{600}=0.5-0.6. Then 1ml plate lysate was added to the *E.coli* culture and shaken vigorously at 37°C. After 4-5 hours, when lysis happened, the lysate was centrifuged at 10,000g (Sorvall GSA) for 15 minutes at 4°C. The supernatant was transferred to a flask. Fifty micro-litres of RNase (10mg/ml) and 50μl DNase (10mg/ml) were added to a final concentration of 1μg/ml. The digestion was carried out at 37°C for 1 hour. NaCl 9g was dissolved into each sample (final concentration 1M) and the flasks were kept on ice 1 hour. To each sample 15g PEG 8,000 (final concentration 10%w/v) was added and dissolved slowly using a stirrer. The flasks were kept on ice or at 4°C overnight to precipitate phage particles. After centrifugation at 12,000g for 15minutes at 4°C, the supernatant was discarded and any drops of fluid in the tube were removed. Then 4ml SM buffer was added to the pellet, which was resuspended slowly and thoroughly, and then 4ml chloroform was added to it and vortexed gently to form an emulsion. After centrifugation at 4,000g (GSA Sorvall) for 15 minutes at 4°C, the aqueous phase was transferred to another tube. EDTA (1.5M pH=8.5) 160μl (to final conc. 20mM), proteinase K (final conc. 50μg/ml) and 10% SDS 200μl (final conc. 0.5%) were added to each sample. After mixing by inversion and incubation for 1 hour at 56°C, the mixtures were extracted with 1 volume phenol (equilibrated with 0.5M Tris-HCl to pH 8) once, phenol:chloroform: isoamyl-alcohol (25:24:1) twice and chloroform: isoamyl-alcohol once. Then 0.1 volume sodium acetate 3M pH7.0 and 0.7vol of isopropanol were added and kept at -20°C for 30 minutes to precipitate the λ DNA. After centrifugation at 15,000g for 30 minutes at 4°C, the pellet was washed with 10ml 70% ethanol three times. The air-dried λ DNA pellet was resuspended in 500μl TE buffer and stored at -20°C.

Conventional λ DNA extraction protocols such as method A, chloroform extraction and phenol/chloroform extraction were performed to extract PEG or phenol and proteins in the λ preparation. Ethanol or isopropanol precipitation was also carried out in both conventional and some commercial extraction kits such as Qiagen lambda DNA

Extraction kit. These procedures were time-consuming and a certain amount of λ DNA was lost during each step. A common problem in routine λ DNA extraction was that the extracted λ DNA was contaminated with RNA (Fig.2.1A). Sometimes there was so much RNA in the λ DNA preparation that the released insert DNA fragment(s) from lambda DNA digestion could not be seen in gels after electrophoresis, as it was covered by the image of RNA. Lambda DNA preparation with no RNA contamination is especially critical when the size of the insert or fragment is small. Lambda DNA of high quality is also essential to direct sequencing from λ DNA. In order to overcome these problems and obtain λ DNA of high quality, a λ DNA extraction/purification system in combination with Hybaid Recovery™ Spin Mini (plasmid) Prep system or silicon bead suspension was developed as method B.

Method B: The initial steps were the same as in method A. The resulted pellet after PEG precipitation was resuspended in 400μl of SM buffer (dissolve NaCl 5.8g, $MgSO_4.7H_2O$ 2g, Tris base 6.05g and gelatin 5ml in 800ml distilled water, adjust to pH7.5 with HCl, add distilled water to final volume 1000ml, sterilize by autoclaving). Then 50μl 20% SDS was added and mixed. The mixture was then incubated at 70°C for 10min. After the mixture was cooled down to room temperature 225μl 3M KAc was added and spun down at top speed in a microcentrifuge at 4°C for 10min. The supernatant was transferred into a fresh 1.5ml microcentrifuge tube. To the supernatant 300μl of 6M guanidine thiocyanate and 300μl of binding buffer including the binding beads from the Hybaid Cosmid Mini Prep kit were added. Half of the mixture was then transferred to a Hybaid spin column and spun at top speed for 1min. The second half mixture then was added and spun down again. The collected solution in the tube was poured out and column was put back into the tube. To the bead pellet 350μl wash buffer (Hybaid kit) was added and the tube was spun at top speed for 2min. The column was then transferred into a fresh collection tube. To the column 100μl TE buffer was added and vortexed briefly (15seconds) on a vortexer. λ DNA was eluted by spinning the tube at top speed for 1min. The binding beads can be prepared from silica (Sigma) (S-5631) as described by Boyle and Lew (Boyle and Lew, 1996) with modifications. After 10g of silica was mixed with 100ml of PBS, the silica was allowed to settle for 1 hr. This procedure was repeated. After centrifugation (2000g for 2min) the silica pellet was resuspended in 6M guanidine thiocyanate.

The λ DNA extracted using this method was digested with restriction enzymes and run in 1.2% agarose gel (1xTBE) (Fig.2.1A). As it can be seen there is no contamination of RNA. DNA fragment or insert was released after restriction digestion. The insert or cDNA, encoding putative serine-threonine kinase in cassava, has been successfully subcloned into plasmid after being purified from the gel.

The system could also be used for cleaning λ DNA, for example, to get rid of RNA contamination. For 10μg of λ DNA, water was added to make up the volume to 250μl and then 250μl binding buffer and beads were added. The rest of the procedure was the same as described above. The performance of the system in eliminating RNA contamination is shown in Fig.2.1B.

This system needed neither phenol/chloroform extraction, nor protease K digestion, gravity elution nor ethanol/isopropanol precipitation. Compared to conventional λ DNA extraction methods and commercial λ DNA purification systems, this system was much simpler, efficient and time-saving. The quality of the purified λ DNA with this system was of automated sequencing standard as satisfactory results of direct sequencing from λ DNA purified using the system were achieved.

2.4.3 Cassava genomic DNA extraction

Young fully expanded leaf tissue (7-8g) was harvested into liquid nitrogen and ground in liquid nitrogen into a fine powder. Then 4g powder was put into each tube, to which 15ml pre-warmed (40-50°C) Dellaporta extraction buffer (100mM Tris-HCl pH8.0, 50mM EDTA and 500mMNaCl; 0.07% v/v mercaptoethanol and 1% PVP) was added. After adding 1ml 20% SDS to each tube and mixing, it was incubated in a shaking water bath at 65°C for 10 minutes, during which the tubes were inverted every 2 minutes. Five ml of 5M potassium acetate was then added and mixed vigorously. The mixture was centrifuged at 15,000g (Sorvall SS34) for 20 minutes at 4°C after which it was incubated on ice for 40-50 minutes. The supernatant was filtered through miracloth (or a double thickness of gauze) into a tube capable of centrifugation at 15,000g (Sorvall SS34), containing 10ml of isopropanol pre-cooled at-20°C. To get high molecular weight DNA the following steps were carried out carefully. The tubes were gently inverted a few

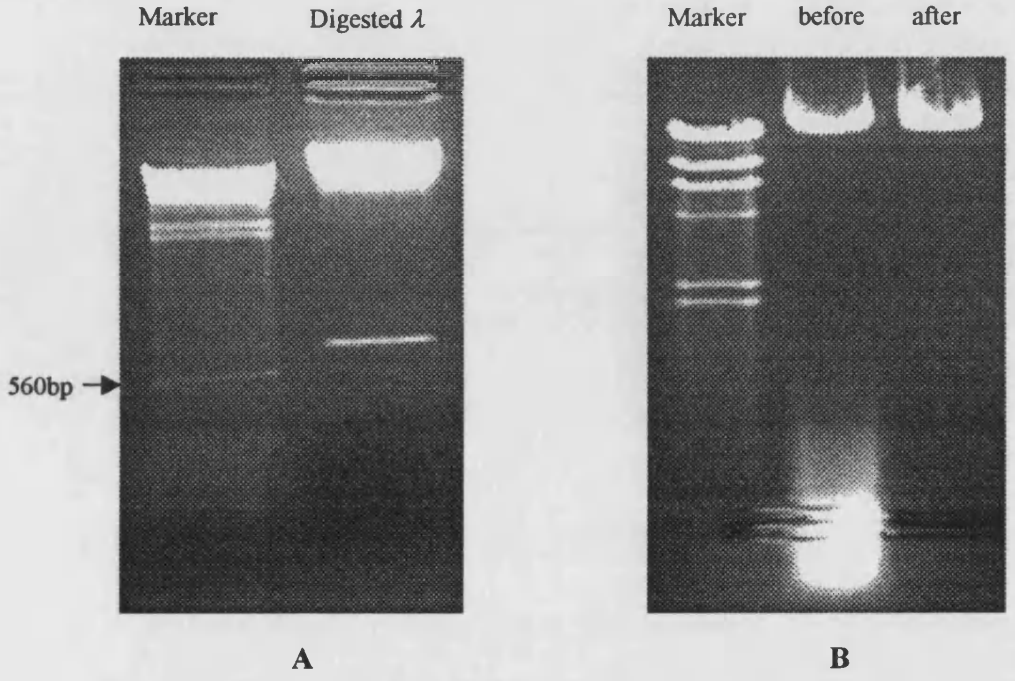


Fig.2.1 Improved method for λ DNA extraction. A: Purified λ DNA (using method B) was digested with *Eco*RI and a band of about 750bp DNA fragment was released (lane on right). **B:** λ DNA extracted using method A was contaminated with RNA (middle lane); the DNA was cleaned using method B (lane on right).

times to mix and then kept at -20°C overnight. The DNA was precipitated by centrifugation at 15,000g (Sorvall SS34) for 15 minutes at 4°C and the pellet was washed with 2ml 70% ethanol three times. After the pellet was air dried at room temperature, it was re-dissolved in 700µl TE buffer (sometimes incubation at 65°C for 5-10 minutes was carried out to improve the process). Then the mixture was transferred to a new Eppendoff tube and centrifuged at 12,000g for 10 minutes at room temperature in a microcentrifuge. The supernatant was transferred to a new Eppendoff tube and DNase free RNase was added at a final concentration of 10µg/ml. Followed incubation in a water bath at 37^{0}C for 30 minutes, phenol/chloroform extraction was applied once. To the aqueous phase 75µl 3M sodium acetate and 500µl isopropanol (pre-cooled) were added and then kept at -20^{0}C overnight. Then mixture was centrifuged at 15,000g for 15 minutes at 4^{0}C. The supernatant was discarded and the pellet was washed with 500µl 70% ethanol (pre-cooled at -20^{0}C) three times and air-dried. The DNA was dissolved in 200µl sterile water and stored at -20^{0}C.

2.5 Restriction Digestion of DNA

A certain amount of DNA was mixed with a certain volume of appropriate 10xbuffer (supplied by the restriction enzyme manufacturer) and water to create 1xbuffer condition suitable for the restriction enzyme for target digestion. Restriction enzyme was added to the mixture according to the quantity of DNA and the instruction about the restriction enzyme from the manufacturer of the enzyme. The digestion mixture was then incubated at 37°C for 2-3 hours for complete digestion. In the case of genomic DNA digestion, genomic DNA samples were digested with relevant restriction enzymes as described below: To 10µg genomic DNA, 10X buffer 5µl, 0.1M spermidine 1.1µl were added, then MilliQ water was added to a final volume 48µl and finally restriction enzyme 2µl (12 units/µl) was added. The digestion was incubated at 37°C overnight.

2.6 Gel Electrophoresis

2.6.1 DNA gel electrophoresis

Agarose was dissolved in 1xTBE buffer or 1xTAE by heating to make up a certain percent of agarose gel depending on the sizes of the DNA. When the gel solution was cooled down to about 50-60°C, ethidium bromide was added to a final concentration 0.5µg/ml. The mixture was then poured to a gel tray. After the gel was set, it was put

into a gel tank containing 1xTBE or 1xTAE buffer. DNA samples were mixed with 1/10 volume of loading buffer (0.25% bromophenol blue and 40% w/v sucrose in water) and loaded. Gels were run at appropriate voltage for appropriate period depending on the feature of the experiment.

2.6.2 Electrophoresis of RNA through formaldehyde gel

Formaldehyde gel was prepared and RNA samples were run as described by Sambrook *et al.* (1989).

2.7 DNA Band Purification

PCR products or the DNA fragments released by restriction digestion, were first run in 0.8% -1% agarose (1xTBE) gel. Then the DNA was purified from the gel-band with a Sephaglas™ Bandprep kit from Pharmacia Biotech. The DNA band was cut out and placed in a microcentrifuge tube and weighed. Gel Solubilizer 1μl/mg gel was added to the tube. Five μl of 50% glacial acetic acid was added to every 250μl Gel Solubilizer to enhance the DNA extraction efficency from the 1xTBE gel. After vigorous vortexing, the mixture was incubated at 60°C for 5-10 minute until the gel was dissolved completely. Sephaglas BP, 5μl for each μg of DNA, was added and vortexed gently. After incubation at room temperature for 5 minutes, and pulse centrifuged for 30 seconds, the supernatant was discarded and the pellet was washed with 70% ethanol 80μl for every 5μl of Sephaglas BP three times and then air-dried. TE buffer was added to the pellet and resuspended. The mixture was incubated at room temperature for 5 minutes with periodic agitation. After 30 seconds centrifugation, the supernatant or the purified DNA was transferred to a fresh tube, and stored at -20°C.

2.8 Probe Preparation

Purified DNA fragments were used to prepare probes using an Oligolabelling kit from Pharmacia Biotech. Template DNA 30-50ng was denatured at 100°C for 5 minutes and kept in ice-bath for 3 minutes. Then 10μl Reagent mix, 2-5μl (α $^{-32}$P) dCTP was added followed by ice-cold MilliQ water to a final total volume of 49μl. Klenow Fragment 1μl(5 units/μl) was added to the mixture. The labelling reaction mixture was incubated at 37°C for 30 minutes. The probe was either stored at -20°C or added directly to pre-hybridisation buffer after denaturation.

2.9 Hybridisation Screening of the λgt10 cDNA Library

Plating cells *E.coli* NM514 was infected with an appropriate dilution of the library and plated out to 90mm LB plates using LB-top agarose so that either high density screening (10^4 plaques/plate) for first screening or low density screening or secondary screening (200-500 plaques/plate) could be performed. The plates were inverted and incubated at 37°C until plaques appeared. These plates were then stored at 4°C. A Hybond™N$^+$ nylon disc was placed on to plaques on the agarose surface of the plate. The filter was left in contact with the plate for 1 minute, during which the filter and agar plate were punctured to mark the relative position for later location of positive areas or plaques on the plate. One filter replica was taken by leaving the filter on the plate for three minutes. The filters were placed with plaque side up for 7 minutes on a sheet of thick filter soaked in denaturing solution (1.5M NaCl, 0.5M NaOH). The filters were transferred to a second sheet of thick filter paper soaked in neutralizing solution (1.5M NaCl 1M Tris-HCl pH7) for 3 minutes. Neutralisation was repeated with a fresh filter paper soaked in the same solution. The filters were finally rinsed in 2x SSC and placed on filter paper with plaque side up to dry. The lambda DNA was fixed by vacuum-dried at 80°C for two hours.

The filters were hybridised and washed as described in 2.9. A film was exposed to the membrane in the dark at - 80°C overnight and then developed. The positive signals on the film indicated that there were positive plaques in the corresponding positions on the plate. As the first screening was at high density it was impossible to get separate positive plaques. The areas where showed a positive signal was cored out into SM buffer. These stocks were plated out at low density (about 200 plaques /plate of 90mm diameter) and a secondary screening was carried out to obtain single positive plaques. For example, 9 positive signals (Fig2.2A) appeared on the film when the PPD related cDNA library was screened for HRGPs clones. By matching the film with the membrane and the plate, nine areas (3-4mm in diameter) corresponding to the positions of the signals were cored out into nine microcentrifuge tubes containing SM buffer. Each was plated out at low density after infection with plating cells. The plaques on each of the plates were lifted to membranes and hybridized again with the same probe. After hybridization, each positive signal on the film (Fig.2.2B) represented one separate positive plaque on the plate. Only one of these positive plaques was eluted for further analysis as they were all the same (from the same positive area identified in the first screening).

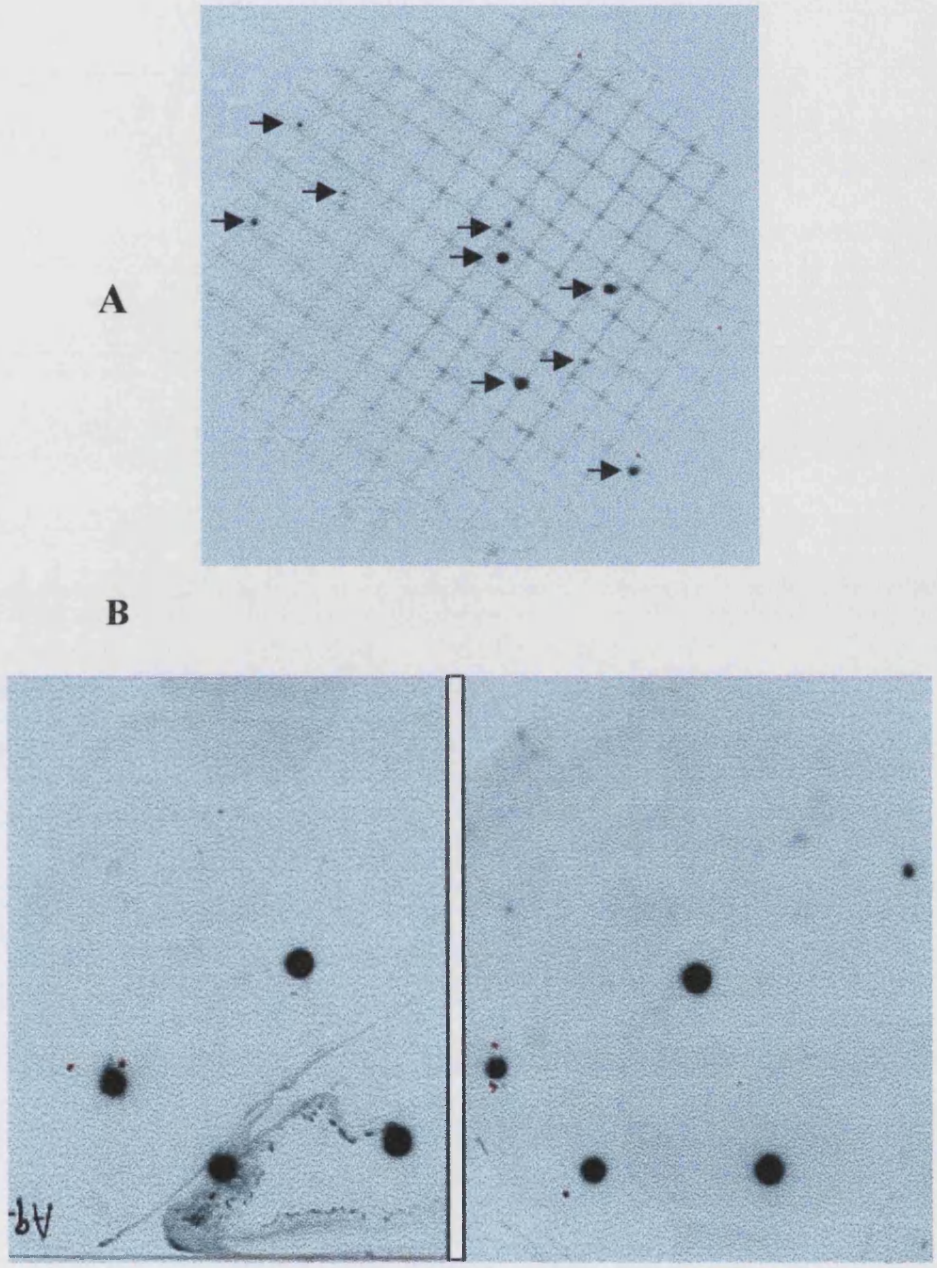


Fig. 2.2 Examples of the results of the first and the second screening of the PPD-related cDNA library. **A:** The positives obtained during the first screening for HRGP cDNA. **B:** the result of the secondary screening of clone 9 from the first screening, The left and right parts were the duplicates. Each positive signal was from a well-separated plaque. (Note: dupicate plaque lift was also screened for the first screening but the signal on the other film was weak and not shown here).

2.10 Southern Blot and Hybridisation

DNA samples, either PCR products or from restriction digestion, were mixed with 1/10 volume loading buffer (0.25% bromophenol blue and 40% w/v sucrose in water) and run in 0.8-1.5% agarose gel (1xTBE) with ethidium bromide (0.5mg/litre) at 3 volts/cm. In the case of genomic DNA, the samples were run in 1% (1xTBE) agarose gel without ethidium bromide at 0.5 volts per cm overnight and then stained in EtBr 100μg /1litre solution for 30 minutes. *Hind*III cut λ DNA or other markers were also run in the gel.

After electrophoresis was finished the gel was photographed against a scale. Afterwards the gel was soaked in 0.25M HCl for 15 minutes, denatured in 1.5M NaCl + 1.5N NaOH with gentle shaking for 30 minutes and neutralized in 1M Tris-HCl (pH 7.4) + 1.5M NaCl for 30 minutes. The gel was then rinsed with distilled water and trimmed. Finally the gel was blotted onto Hybond N^+ overnight as described (Sambrook *et al.*, 1989). The membrane was rinsed in 2xSSC for 15 minutes, air dried briefly and vacuum-fixed at 80°C for 2 hours.

Filters were placed in 50 -100ml preheated prehybridization buffer (Sambrook *et al.*, 1989) in a hybridisation sandwich box. Sufficient sonicated heterologous DNA (herring sperm DNA) (to final concentration 0.1mg/ml in prehybridisation buffer) was denatured first in boiling water bath for three minutes and then cooled down in ice bath before it was added to the prehybridisation solution to give a final concentration of 100μg/ml. The filter was prehybridised by incubating at 60°C with slight shaking for 1-2 hour in a sandwich box in a shaking incubator. The radio-labelled probe was added after denaturation at 95°C for 5 minutes. Hybridisation was carried out overnight at the relevant temperature. Following hybridisation, the filters were washed at relevant stringency from 2xSSC/0.1%SDS to 0.1xSSC/0,1%SDS. The filters were then rinsed with 2xSSC and drained briefly. Finally the filters were wrapped in cling film, and Kodak X-ray films were exposed to the filters in cassette in -80°C for one to several days depending on the strength of the expecting signals and developed.

2.11 Northern Blot and Hybridisation

After electrophoresis RNA gel was soaked in 0.005N NaOH for 20 minutes, rinsed in RNAase-free water and soaked in 20xSSC for 45 minutes. The RNA was transferred to

a nylon filter and hybridised against DNA probe as described as described (Sambrook *et al.*, 1989).

2.12 Competitive Reverse Transcription Polymerase Chain Reaction (RT-PCR)

Competitive RT-PCR using artificial internal RNA standards or competitor is the most accurate technique for quantitation of mRNA (Efrel *et al.*, 1997). It can therefore be used to compare relative expression levels of one gene in different treatments by using control RNA of the same concentration for different samples.

Construction of competitor RNA or DNA. Different methods could be used to make competitor constructs based on the features of the DNA sequence of a particular gene and the vector in which the gene is. The guideline is that the PCR product from the competitor should be different in size to that from the testing mRNA so that these two products could be easily separated on a gel but not so far apart.

RT-PCR reaction mixture was prepared according to the instructions of the manufacturer (Promega, Access RT-PCR System): nuclease-free water (to a final volume of 50 µl), 5µl AMV Tfl 5x reaction buffer, 1µl dNTP mix (10mM each dNTP), 50pmol of each forward and reverse primers, 2 µl of 25mM $MgSO_4$, 1µl Reverse Transcriptase (5u/µl), 1µl Tfl DNA polymerase(5u/µl), certain amount of testing RNA and internal RNA or DNA standard. The reverse transcription and PCR amplification was performed in the same tube. The conditions for the RT-PCR was based on the following programme, 1 cycle of reverse transcription at 48°C for 45min, 1 cycle of RNA/cDNA/primer denaturation and AMV RT inactivation at 94°C for 2min, 40 cycles of denaturation at 94°C for 30sec, annealing at certain temperature depending the primers used for 1min and extension at 68-72°C for 2min, 1 cycle of final extension at 68-72°C for 7min.

2.13 Sub-cloning

DNA was digested with appropriate restriction enzymes to generate compatible ends for cloning. The digestion released DNA or insert DNA which was then run in 1xTBE or 1xTAE agarose gel and purified as described in 2.5. The concentration of the purified insert DNA was estimated by agarose gel electrophoresis along with molecular weight

standard of known concentration. The cloning vector was also digested with corresponding restriction enzymes in multiple cloning region and then the restriction enzymes were inactivated by heating the reaction mixture at 70°C for 5 minutes.

For each ligation reaction, molar ratio of insert DNA : Vector DNA was 3:1 or 1:3 or 1:1 with the total DNA about 50ng, 1µl 10x ligation buffer (NBL) and the final volume was adjusted to 10µl with MilliQ water before 1µl T4 DNA ligase (NBL) was added. The ligation was performed at 12°C overnight.

Ligation mixture was then transformed into fresh competent cells of *E.coli* DH5α (ϕ80*dlacZ*ΔM15, *recA*1, *endA*1, *gyr A*95, *thi*-1, *hsd R*17 (r_K^-, m_K^+), *supE*44, *relA*1, *deoR*,Δ(*lac ZYA-argf*)U169) strain with calcium chloride (Sambrook *et al.*, 1989) and transformants were selected by white-blue screening.

2.14 DNA Deletion Using Exonuclease III and Mung Bean Nuclease

Deletion using exonuclease III and mung bean nuclease has been used routinely for sequencing DNA inserts in plasmids. Deletion was used here to make an artificial construct from pUC-MeP1E/H (containing *Eco*RI/*Hind*III cMePAL1 fragment) to be used in competitive RT-PCR.

About 5 to 6 µg of pUC-MeP1E/H was fully restriction digested with *Xho*I, gel-purified and resuspended the DNA in 100µl of 1x exonuclease buffer (Promega, M1811). About 0.7µl of β-mercaptoethanol(14M stock) was added to the DNA solution (to a final concentration about 10mM). About 60 to 100 units of exonuclease-III (4µl of 1/10 dilution of exonuclease, Promega) was added to the DNA, mixed immediately, incubated at 22°C. After 15sec, 25µl of the deletion mixture was transferred into a microcentrifuge tube and frozen quickly in liquid nitrogen. Aliquotes were also taken by 25sec, 35 sec and 45 sec. The exonuclease was then inactivated by incubating at 68°C for 15min. To make the DNA blunt ended, 3units of mung bean nuclease (NBL) was used every microgram DNA. The reaction was performed at 30°C for 30min and stopped with ethanol precipitation. After gel-purification, the open and deleted plasmid was ligated back to a functional one and transformed into *E. coli*. Plasmid DNA was extracted from

the transformants and further analysed by restriction digestion and sequencing to check the deletion.

2.15 DNA Sequencing and Sequence Analysis

DNA sequencing analysis was performed on an ABI 377 using dye primer automatic sequencing. Sequence search and analyses were performed using FASTA, MAP and BESTFIT routines of telnet gnome GCG package (Version 8, 1994), and National Centre for Biotechnology Informarion (NCBI, http://www.ncbi.nlm.nih.gov/).

Chapter Three

Construction of a Post-harvest Physiological Deterioration-related cDNA Library

A cDNA library represents the information encoded in the messenger RNA (mRNA) of a particular tissue or organism. As RNA molecules are exceptionally labile and difficult to amplify in their natural form, the information from RNA is converted first into a stable double strand DNA (cDNA) and then is inserted into a self-replicating lambda vector. The pool of lambda containing the cDNAs from original population of mRNA constitutes the cDNA library. Individual cDNA can be isolated and characterised with relative ease. The transcription and processing of mRNA can be analysed and interpreted to study the expression of the relevant gene. A cDNA library constructed from a population of mRNA of cassava storage root undergoing PPD would permit the isolation of PPD-related genes in order to understand the mechanism of PPD on the level of gene expression.

3.1 Construction of PPD-related cDNA Library

Studies on biochemical changes during PPD suggested that very active metabolism, including changes related to wounding responses, occur within 24 to 48 hours post-harvest (see 1.5). It was shown by *in vivo* labeling of proteins in cassava storage root discs that there was a massive synthesis of proteins, including novel ones, during PPD (Beeching *et al.*, 1995). The mRNA population expressed during this period probably plays critical roles in the development of PPD.

3.1.1 Total RNA Extraction and mRNA Isolation

Cassava root discs were kept in liquid nitrogen after incubation for 24 to 48 hours post-harvest for extracting RNA to construct PPD-related cDNA library. Obtaining total RNA of good quality has proved very difficult due to the presence of large amount of phenolic compounds and starch. About 440μg total RNA was used to isolate mRNA with Oligotex™ suspension (see 2.3.1) and 4 to 5μg of mRNAs was estimated to be isolated.

3.1.2 Synthesis of cDNA

The mRNA was *in vitro* reverse-transcripted into the first strand cDNA. Then double strand cDNA was synthesized, during which [α-^{32}P] dCTP was included in two reactions to predict the yield of cDNA. After double strand cDNA synthesis was completed, a sample was taken to measure the incorporated [α-^{32}P] dCTP into the double strand cDNA. The incorporation efficiency was about 1.6%, from which the yield of double strand cDNA was estimated as 840ng. The cDNA synthesis efficiency from mRNA was about 9.55%, which is about the suggested efficiency (>9%) from the kit manufacturer.

3.1.3 cDNA Cloning and *in vitro* Packaging

The purified double strand cDNA was ligated with *Eco*RI adaptors. The 'adapted' cDNA was then purified and size-fractionated using one column, and phosphorylated with kinase prior to ligation with the dephosphorylated λgt10 arms to allow covalent insertion. The ligation mixture was packaged and kept in SM buffer. The cDNA library was constructed.

3.2 The Titre and the Quality of the Library

The titre of the library was 4.5×10^5 pfu /ml. From the library 12 phage plaques were randomly picked and subjected to PCR using primers flanking the sequence of lambda arms to analyse the size of the cDNAs and insert ratio of the library. PCR amplification revealed that 7 out 12 tested clones or 58.33% had inserts ranging from 0.5 to 2.0 kb (Fig. 3.1). The actual range of the insert size was proved to be from 100bp to 2.3kb after a 100bp putative HRGP clone (see 5.2.1) and a 2.3kb PAL clone (see 4.2.1) were isolated from the library.

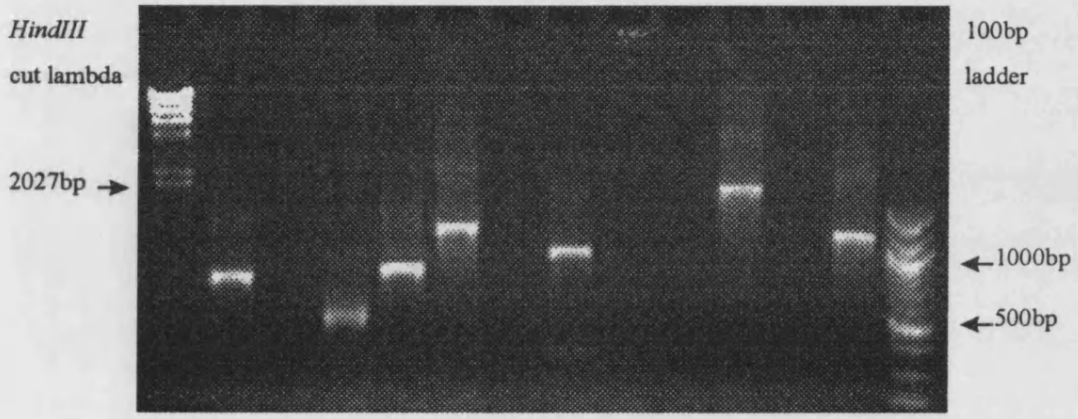


Fig.3.1 PCR amplification products, derived from 12 individual random plaques from the PPD-related cDNA library. Products were analysed by gel electrophoresis through a 1% (1xTBE) agarose containing ethidium bromide and photographed. The first lane from the left is *Hin*dIII cut lambda DNA markers and the last lane on the right is 100 base pair ladder, with particular markers indicated. The PCR was run in a thermal cycler with the following programme: 94°C, 3min; 35 cycles of 94°C for 1 minute, 50-52°C for 1 minute, 72°C for 3 minutes; completed by 1 cycle of 72°C, 5 minutes.

Chapter Four

The Phenylalanine Ammonia-lyase Gene Family in Cassava

4.1 Introduction

It has been demonstrated that PAL activity or/and gene expression is associated with wound (healing) responses in other plants (Street *et al.*, 1986; Liang *et al.*, 1989; Ishizuka *et al.*, 1991; Joos and Hahlbrock, 1992; Oosterhaven *et al.*, 1995). Though there is a remarkable increase in PAL activity and phenolic compounds during the storage of wounded cassava roots (Tanaka, 1983; Uritani, 1984), there is little wound healing response in these roots. The xylem vessels in storage parenchyma responded dramatically to wounding with discoloration, which was associated with the oxidation of phenolics (Rickard, 1985). Does this result from possible rapid activation of the expression of certain PAL genes in xylem parenchyma cells, and / or the changed redistribution of products of core phenylpropanoid metabolism to various branch pathways? Analysis of the PAL gene organisation in cassava and the expression of PAL genes during PPD may help to understand the possible role of PAL genes in the development of PPD. One PAL gene, MePAL, had been partially cloned using PCR with degenerate primers in another research group by the time the PAL cDNA clones described in this chapter were isolated. MePAL was expressed during PPD but was responsible for just a part of the PAL activity during PPD (Periera *et al.*, 1999). Therefore other PAL genes must be involved in PPD. This chapter addresses the possible roles of PAL genes through the isolation and characterisation of members of the cassava PAL gene family.

4.2 Results

4.2.1 Isolation and characterisation of PAL cDNA clones

Most of the PAL genes identified in plants are highly conserved. Many of them have been isolated, either by screening a cDNA or genomic DNA library with a probe made from the DNA fragment of an identified PAL gene from other plants (Wanner *et al.*, 1995), or by PCR amplification using degenerate primers designed from the conserved regions of PAL genes (Butland *et al.*, 1998). Using bean *Phaseolus vulgaris* PAL DNA probes, PAL genes have been isolated from other plants such as *Arabidopsis thaliana* and melon (Wanner *et al.*, 1995; Diallinas and Kanellis, 1994). Bean gPAL2 was shown to be activated by elicitor and wounding of hypocotyls (Edwards *et al.*, 1985; Cramer *et*

al., 1989). Therefore a probe prepared from 1.8kb *Hin*dIII/*Eco*RI fragment of bean gPAL2 was used to screen for PAL cDNA clones from the PPD-related cDNA library (Chapter 3). Two positives were obtained after first and secondary screening. The positive clones were confirmed by sequencing the PCR products amplified from the insert regions and comparing the sequences with other PAL genes in plants. The cDNA inserts were either subcloned into plasmids to be fully sequenced, or sequenced directly from the PCR product amplified from the λ DNA of the positive clones. The sequences of the cDNAs were further studied using sequence analysis programmes.

4.2.1.1 Isolation of PAL cDNA clones cMePAL1 and cMePAL3

About $5x10^4$ phage lambda particles from the PPD-related cDNA library were incubated with plating cells (*E. coli*, MN514) at 37°C for 15min, plated out with soft top agarose onto a LB plate at high density and incubated at 37°C overnight. The plaques were lifted onto Hybond N+ nylon membranes, which were then hybridised with the probe prepared from bean gPAL2. Pre-hybridisation and hybridisation were performed at 60°C for one to three hours and 50°C overnight respectively as described in 2.9. Following hybridisation, the filters were washed with 2x SSC, 0.1% (w/v) SDS at room temperature for 2x10 min; 2x SSC, 0.1% (w/v) SDS at 50°C for 2x10min, 1x SSC, 0.1% SDS at 50°C for 15 minutes. From the first screening two positives were obtained. The positives were then separately plated out at low density and screened under the same conditions as for the first screening, which led to the isolation of single positive plaques for two clones (cMePAL1 and cMePAL3).

In order to test the size and the fidelity of the cDNA inserts, these positive clones were subjected to PCR using a set of λgt10 primers, which would amplify the insert region. The size of the clones was about 2.3kb for P1 and 2.1kb for P3 (Fig.4.1).

Partial sequences of the PCR products were obtained by sequencing with λgt10 primers. The sequences were compared to sequences in DNA database using the computer program BLASTx (NCBI, 1996 or http://www.ncbi.nlm.nih.gov/). The search revealed that the deduced amino acid sequences of these two clones were of high identity (62-81%) to those of pea (*Pisum sativum*) PAL1 and PAL2, PAL of sweet potato (*Ipomoea batatas*), PAL of Aspen (*Populus kitakamiensis*) and avocado (*Persea americana*).

These indicated that the clones contained putative PAL cDNAs. The 2.3kb and 2.1kb cDNA inserts were designated as cMePAL1 and cMePAL3 respectively, as by the time cMePAL1 was fully analysed a genomic DNA clone had been isolated (H-Y Li *et al.*, unpublished results). The sequences of the genomic DNA clone were similar to but different from cMePAL1 in the coding region and this genomic clone was designated as gMePAL2.

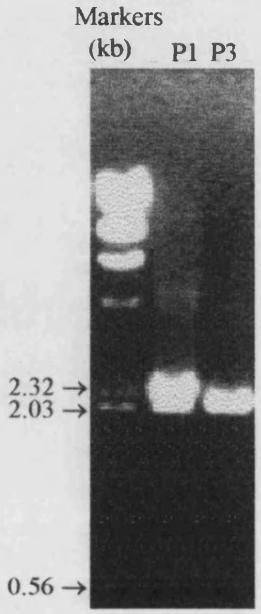


Fig. 4.1 PCR amplification of inserts from positive PAL plaques. PCR was performed as described in 2.3.7. The PCR products were run in 0.8% (1xTBE) agarose gel. *Hin*dIII cut λ DNA was used as markers. The size of the PCR products is 2.3kb for P1 (cMePAL1) and 2.1kb for P3 (cMePAL3).

4.2.1.2 Characterisation of cMePAL1

Plasmid vectors are generally easier to manipulate than phage λ in terms of amplification, DNA extraction and restriction digestion. It is a routine cloning procedure to subclone the insert from a λ clone into a plasmid vector if the cDNAs are initially cloned into vectors such as λgt10. The cDNA cMePAL1 was subcloned into pUC18 and sequenced as illustrated in Fig.4.2. First, the λ DNA was extracted from cMePAL1 and digested with *Eco*RI, which released two fragments, 1910bp and 387bp in size. This indicated that in addition to the *Eco*RI sites in the two adaptors on both ends of the cDNA, there was an internal *Eco*RI site in cMePAL1. The adaptors was attached to double strand cDNA to enable the cDNA to be ligated with the *Eco*RI site of λgt10 arms during the cDNA cloning process. Therefore, the released large and small fragments contained one adaptor each on one of their ends. The large and small cDNA fragments were subcloned into pUC18, forming pUC-MeP1Eb and pUC-MeP1Es respectively. Then a 798bp

*Eco*RI/*Hind*III fragment from pUC-MeP1Eb was subcloned into pUC18, resulting in pUC-MeP1E/H.

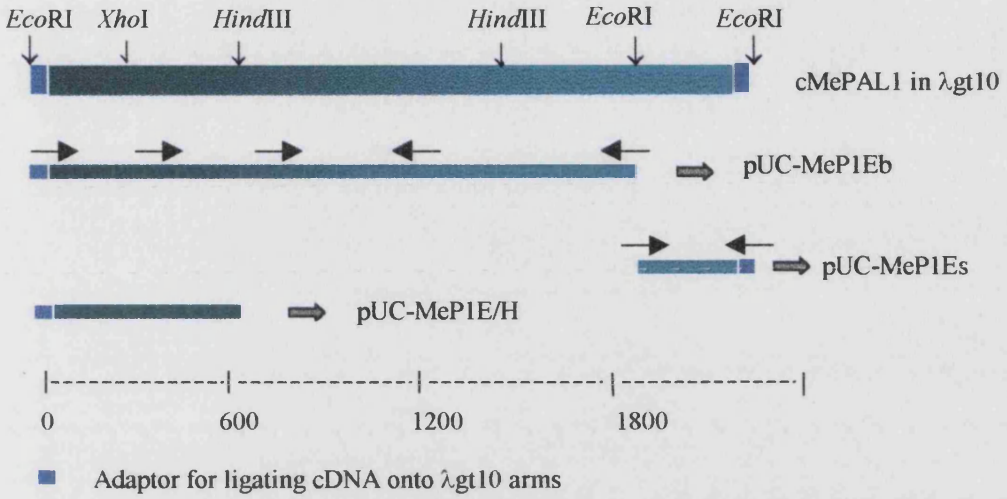


Fig. 4.2 MePAL1 cDNA subcloning and sequencing strategy. The MePAl1 cDNA was released from λgt10 with *Eco*RI digestion, which cut the cDNA into one long fragment (1910bp) and one small fragment (387bp). These two fragments, designated as MeP1Eb and MeP1Es respectively, were subcloned into pUC18 and named as pUC-MeP1Eb and pUC-MeP1Es. The *Eco*RI/*Hind*III fragment in pUC-MeP1Eb was subcloned into pUC18, forming the pUC-MeP1E/H. The sequence of the cDNA was obtained by using universal primers to sequence from both ends of the cDNA fragment in the plasmid vectors and by designing primers (as shown with arrows in bold) from obtained sequences to extend the sequence.

These clones were then checked by restriction digestion. The cDNA fragments of the expected sizes were released when pUC-MeP1Eb, pUC-MeP1Es and pUC-MeP1E/H were restriction digested with *Eco*RI, *Eco*RI and *Eco*RI/*Hind*III, respectively (Fig.4.3).

Sequence determination was carried out in both plasmid pUC-MeP1Eb and pUC-MeP1Es in parallel (Fig.4.2). The large cDNA fragment (1910bp) in pUC-MeP1Eb was first sequenced from both ends of the cDNA insert pUC-MeP1Eb using universal forward and reverse primers. Based on the obtained sequences, primers 375f, 775f and 1500r were designed to determine the rest of the sequence of the cDNA fragment. The small cDNA fragment (387bp) in pUC-MeP1Es was completely sequenced by using universal forward and reverse primers. The sequence segments obtained from each plasmid were assembled into separate consensus contigs using the Gel-assembly program in GCG. Then the two consensus contigs were joined together. Since there was a 25bp adaptor sequence at one end of the big fragment and the small fragment, while at the

other end there was an *Eco*RI site derived from the internal *Eco*RI site, the two consensus contigs were joined at the *Eco*RI site. The sequences obtained from PCR product from the lambda (4.2.1) were also used to check the combination of these two consensus contigs and the sequences near the both ends of cMePAL1. The complete sequence of cMePAL1 was obtained after the adaptor sequences were removed from both ends.

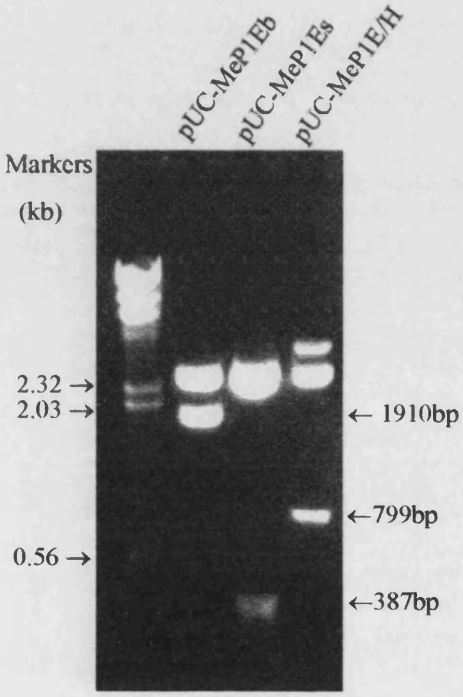


Fig.4.3 Restriction digestion of pUC-MeP1Eb, pUC-MeP1Es and pUC-MeP1E/H. The sizes of the released PAL cDNA fragments are indicated on the right.

The sequence of cMePAL1, after removal of the adaptors, was 2253bp, consisting of a 77bp 5'UTR, a 2130bp open reading frame beginning with an translation initiator ATG and ending with stop codon TAA, and a 46bp 3'-untranslated region (Fig.4.4). The open reading frame encoded a peptide containing 710 amino acid residues. The deduced amino acid sequence was compared with the protein patterns in PROSITE (http://www.expasy.ch/sprot/prosite.html), which led to the identification of the phenylalanine ammonia-lyase signature or active site, GTITA SGDLV PLSYIA (1246bp to 1290bp) in the polypeptide.

```
   1  CTCAAACTTCCTATACCTATATTTTTCTCCCACTTTCTCACCACTCCTCTACCTTTCCGG   60
 61  CAAAATAAATCAGCAAAATGGAGTTTTGTGAGGCTCACAATGTAACTGCTTCACCTGGAT  120
                     M  E  F  C  E  A  H  N  V  T  A  S  P  G  F
121  TCTCCTCTGCTGACCCATTGAACTGGGGCATGGCTGCAGAGTCACTGAAGGGCAGCCACC  180
       S  S  A  D  P  L  N  W  G  M  A  A  E  S  L  K  G  S  H  L
181  TTGATGAGGTGAAGCGCATGGTGGATGAGTACAGGAAGCCTGTGGTGAGGCTAGGGGGTG  240
       D  E  V  K  R  M  V  D  E  Y  R  K  P  V  V  R  L  G  G  E
241  AGACCCTGACTATAGCCCAAGTTACAGCAATTGCGAACCATGACTCAGGTGTCAAGGTTG  300
       T  L  T  I  A  Q  V  T  A  I  A  N  H  D  S  G  V  K  V  E
301  AGCTGTCTGAGGAGGCTCGAGCTGGGGTCAAGGCCAGTAGTGACTGGGTTCTTGATTCCA  360
       L  S  E  E  A  R  A  G  V  K  A  S  S  D  W  V  L  D  S  M
361  TGAATAAAGGAACCGATAGTTATGGTGTCACCACTGGCTTTGGTGCAACCTCCCATAGAA  420
       N  K  G  T  D  S  Y  G  V  T  T  G  F  G  A  T  S  H  R  R
421  GAACCAAGCAGGGGGGTGCCCTTCAGAGAGAACTCATAAGATTCTTGAATGCTGGGATCT  480
       T  K  Q  G  G  A  L  Q  R  E  L  I  R  F  L  N  A  G  I  F
481  TTGGAAATGGACAAGAATCTTGCCACACATTGTCTCACACTGCAACTAGAGCAGCAATGC  540
       G  N  G  Q  E  S  C  H  T  L  S  H  T  A  T  R  A  A  M  L
541  TGGTGAGAATCAACACCCTCCTGCAAGGTTATTCAGGCATAAGATTTGAAATCCTGGAAG  600
       V  R  I  N  T  L  L  Q  G  Y  S  G  I  R  F  E  I  L  E  A
601  CCATTACCAAGTTCATCAACAACAATGTTACTCCGCGTTTGCCCCTCAGAGGCACAATCA  660
       I  T  K  F  I  N  N  N  V  T  P  R  L  P  L  R  G  T  I  T
661  CAGCCTCTGGTGACCTGGTCCCACTGTCCTACATTGCCGGGCTTTTGACCGGCCGGCCCA  720
       A  S  G  D  L  V  P  L  S  Y  I  A  G  L  L  T  G  R  P  N
                                                    ↓HindIII
721  ACTCCAAGTCGTTAGGGCCCAATGGAGAATCCTTGGATGCAGCTGAAGCCTTTAAGCTTG  780
       S  K  S  L  G  P  N  G  E  S  L  D  A  A  E  A  F  K  L  A
781  CTGGGATCAATGGTGGATTTTTTGAATTGCAGCCCAAAGAGGGTCTAGCTTTAGTAAATG  840
       G  I  N  G  G  F  F  E  L  Q  P  K  E  G  L  A  L  V  N  G
841  GTACTGCAGTTGGTTCAGGCTTAGCTTCTATGGTTCTTTTTGAGGCCAATGTGTTGGCAG  900
       T  A  V  G  S  G  L  A  S  M  V  L  F  E  A  N  V  L  A  V
901  TCCTGTCAGAAGTTTTATCAGCAATTTTTGCAGAAGTTATGCTTGGAAAACCAGAGTTTA  960
       L  S  E  V  L  S  A  I  F  A  E  V  M  L  G  K  P  E  F  T
961  CAGATCATTTGACACATAAATTGAAGCATCATCCTGGACAAATTGAAGCTGCAGCAATTA 1020
       D  H  L  T  H  K  L  K  H  H  P  G  Q  I  E  A  A  A  I  M
1021 TGGAACATGTCTTGGATGGAAGTTCTTATATTAAAGCAGCTCAAAAGGTTCATGAAATTG 1080
       E  H  V  L  D  G  S  S  Y  I  K  A  A  Q  K  V  H  E  I  D
1081 ATCCATTGCAGAAGCCTAAACAAGACAGATATGCTCTTAGAACATCTCCTCAATGGCTTG 1140
       P  L  Q  K  P  K  Q  D  R  Y  A  L  R  T  S  P  Q  W  L  G
```

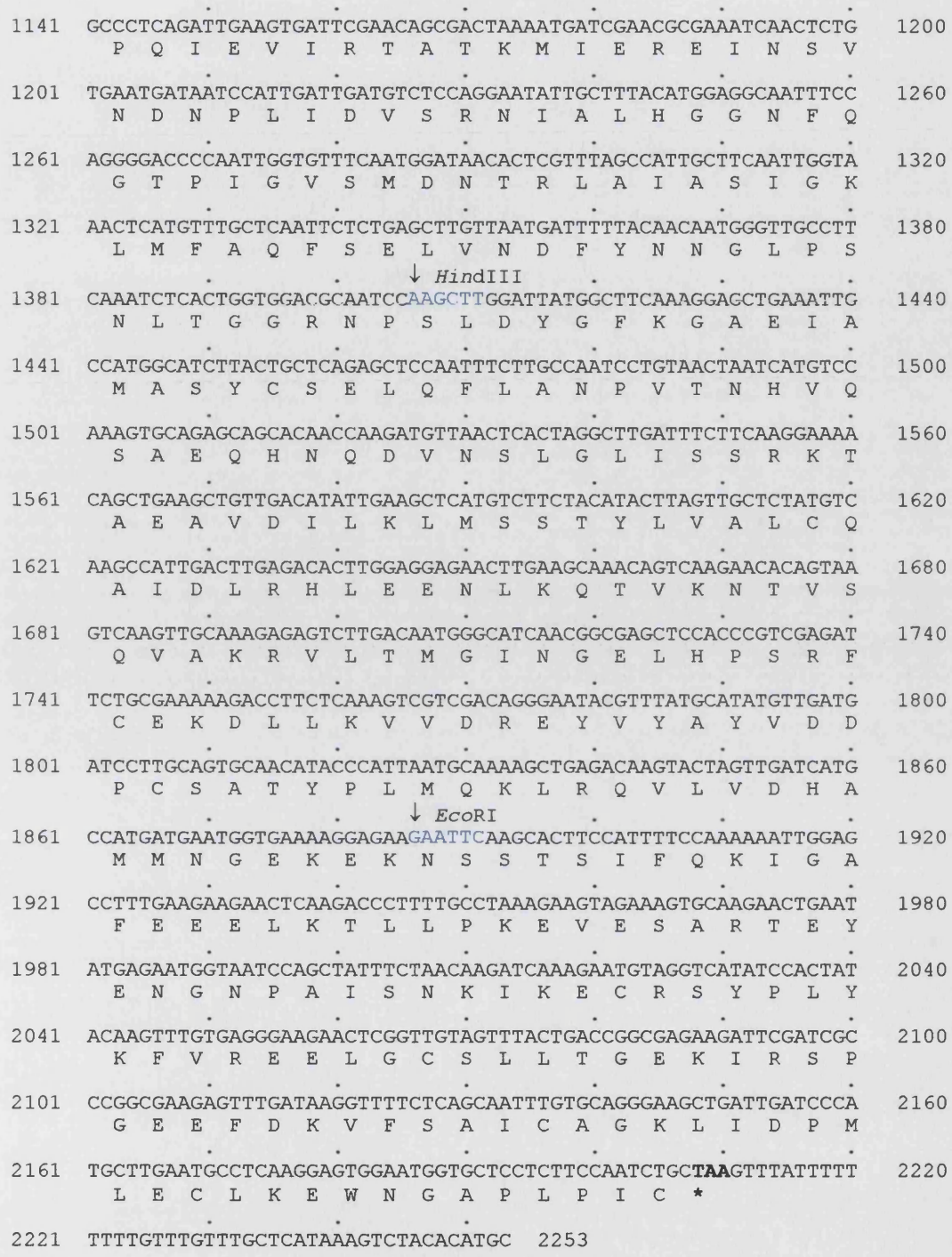


Fig.4.4 Nucleotide sequence of cMePAL1. Codons for translation initiation and termination are in bold at position 78 and 2210. The amino acid residues are placed under the first nucleotides in the codons. The PAL signature is indicated with double-underlined letters. Restriction sites of *Hin*dIII and *Eco*RI are in blue.

4.2.1.3 Characterisation of cMePAL3

The 2.1kb cDNA cMePAL3 was not subcloned into plasmid vector as there were no suitable restriction sites. The *Eco*RI sites in the adaptors might be destroyed during the

cDNA cloning process. Therefore, a different approach was adopted to characterise this clone. As was mentioned during the identification of this clone, the sequence of cMePAL3 was determined by sequencing the PCR product, first with λgt10 forward and reverse primers and then with primers designed from the determined sequences. The individual sequences were assembled into a contig encompassing the entire cDNA clone using the Gene Assembly program in GCG.

The sequence of the insert from this clone turn out to be chimeric (Fig.4.5). BLAST search with the cDNA sequence showed that the deduced amino acid sequence of the first part (938bp) was of high identity (up to 80%) to glycinamide ribonucleotide synthetase (GARS) identified in *Arabidopsis thaliana* and soybean (see 7.3), which is involved in purine biosynthesis. The second part (1126bp) was similar to PAL genes of other plants at the level of both DNA and deduced amino acid sequence. As it is shown in section 4.4 that this PAL cDNA fragment has 97% identity to the corresponding region of gMePAL2 at the DNA level and its deduced amino acid sequence is just one amino acid residue difference from that of gMePAL2. To avoid confusion, the PAL part of the clone is referred to cMePAL3 from now on. The cDNA fragment cMePAL3 is just part of the full sequence of the PAL gene, covering the coding region near the C-terminal of the peptide and 3' untranslated region (950bp-1126bp) (Fig4.6).

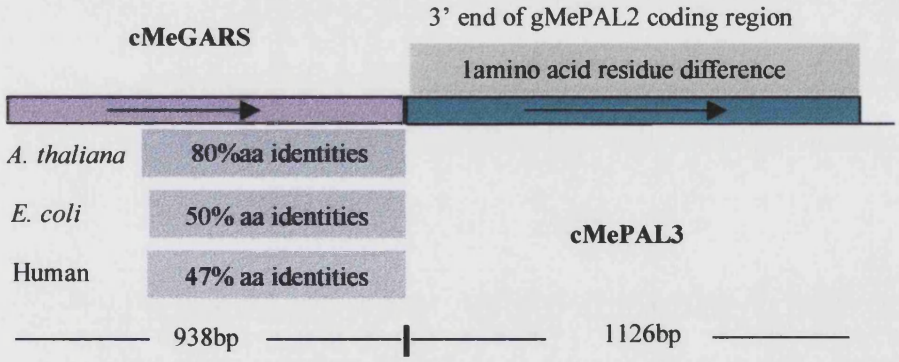


Fig.4.5 A chimeric positive clone consists of two fragments of cDNA showing high similarity to glycinamide ribonucleotide synthetase (GARS) or phosphoribosylamine-glycine ligase and PAL respectively in other organisms. The two cDNA coding regions are in the same orientation as indicated with solid arrows. The identities of the deduced amino acid (aa) sequences of the cDNAs to GARS in other organisms and gMePAL2 are shown. Sequence accession number, *A thaliana* GARS X74766, E coli GARS X51950 and human (*Homo sapiens*) tri-functional protein- GARS - aminoimidasole ribonucleotide synthetase (AIRS)-glycinamide ribonucleotide formyltransferase (GART) X54199.

```
   1  GGGGACCCCAATTGGAGTCTCAATGGATAATGCACGTTTGGCCATTGCATCAATAGGAAA   60
      G  T  P  I  G  V  S  M  D  N  A  R  L  A  I  A  S  I  G  K
  61  GCTCATGTTTGCTCAGTTCAGTGAGCTTGTAAATGATTTTTACAACAATGGGTTGCCATC  120
      L  M  F  A  Q  F  S  E  L  V  N  D  F  Y  N  N  G  L  P  S
 121  AAATCTCACAGCCAGCAGGAATCCAAGCTTGGATTACGGCTTCAAGGGAGCTGAAATTGC  180
      N  L  T  A  S  R  N  P  S  L  D  Y  G  F  K  G  A  E  I  A
 181  AATGGCTTCTTACTGTTCTGAGCTCCAATACCTTGCAAATCCAGTCACCAGCCATGTACA  240
      M  A  S  Y  C  S  E  L  Q  Y  L  A  N  P  V  T  S  H  V  Q
 241  AAGTGCAGAGCAGCACAATCAAGATGTAAACTCCTTGGGGCTAATTTCTTCAAGAAAGAC  300
      S  A  E  Q  H  N  Q  D  V  N  S  L  G  L  I  S  S  R  K  T
 301  AGAAGAAGCTGTAGACATCTTGAAGCTCATGTCCACAACTTTCTTAGTAGCACTTTGCCA  360
      E  E  A  V  D  I  L  K  L  M  S  T  T  F  L  V  A  L  C  Q
 361  AGCTATTGACTTGAGGCATTTGGAGGAGAACTTGAAGCACGCAGTCAAAAACACAGTAAG  420
      A  I  D  L  R  H  L  E  E  N  L  K  H  A  V  K  N  T  V  S
 421  CCAAGTAGCTAAGAGGATTCTAACTACAGGAGCTAGTGGAGAACTTCACCCATCAAGATT  480
      Q  V  A  K  R  I  L  T  T  G  A  S  G  E  L  H  P  S  R  F
 481  CTGCGAGAAGGACTTGCTCAAAGTGGTGGATCGCGAGCAAGTTTTCTCTTATGTCGATGA  540
      C  E  K  D  L  L  K  V  V  D  R  E  Q  V  F  S  Y  V  D  D
 541  CGCCTGCAGTGCTACCTATCCATTGATGCAAAAACTAAGGCAAGTTCTCGTGGACCATGC  600
      A  C  S  A  T  Y  P  L  M  Q  K  L  R  Q  V  L  V  D  H  A
 601  CTTGGCAAATGGCGAGAGTGAGAAGAATGCCAGCACTTCAATCTTCCAAAAGATCAGAGC  660
      L  A  N  G  E  S  E  K  N  A  S  T  S  I  F  Q  K  I  R  A
 661  TTTCGAGGAAGAATTGAAAGCCCTTTTGCCTAAAGAAGTTGAGAGTGCAAGAGAGGCATA  720
      F  E  E  E  L  K  A  L  L  P  K  E  V  E  S  A  R  E  A  Y
 721  CGAGAATGGGAATCCAGCAATTGCCAACAAGATCAAGGAATGCAGATCTTACCCATTGTA  780
      E  N  G  N  P  A  I  A  N  K  I  K  E  C  R  S  Y  P  L  Y
 781  TAAGTTTGTGAGAGAGGAAATAGGAACTGGGTTGCTGACCGGCGAAAAGGTCCGGTCACC  840
      K  F  V  R  E  E  I  G  T  G  L  L  T  G  E  K  V  R  S  P
 841  GGGAGAGGAATTTGATAAGGTTTTCACTGCCATGTGCCAAGGAAAGATCATTGATCCAAT  900
      G  E  E  F  D  K  V  F  T  A  M  C  Q  G  K  I  I  D  P  M
 901  GCTGGATTGTCTCAAAGAGTGGAATGGTGCCCCTCTTCCAATATGTTAAACTGTAACTTT  960
      L  D  C  L  K  E  W  N  G  A  P  L  P  I  C  *
 961  CTTGTTTTGTTTACACTTAAAGATTTGTTTTCCAATTGCTTTTTATGTACTTATAGTTTG 1020
1021  TGATGTAAAAAATCTGTAATGCATTTCTTTTAAATGTTCAATTGTTCTCTTCTCACTTTT 1080
1081  GTGCTTnAATTGAAGGCAGAAnGGCAATGGTAATTATTTAAAGATT 1126
```

Fig.4.6 The partial sequence of cMePAL3 and its deduced amino acid sequence. The translation stop codon is in bold.

4.2.2 Cloning of PAL Genomic DNA

While the cDNA mainly provides information about the coding region of a gene, the corresponding genomic DNA contains much more information about the gene including how the gene might be regulated (regulatory elements in the promoter) and where the splicing occurs (the boundaries between exons and introns). Promoter regions can be studied by fusing them to a reporter gene and transforming cassava in order to analyse the spacial and temporal expression of the promoter or the PAL gene. Introns can be used to make gene-specific probes to analyse not only the gene copy numbers in a genome but also the turnover of transcripts.

An intron was identified (by Hongying Li, unpublished results) by amplifying genomic DNA (from cultivar MNGA 1) using PCR with primers (forward primer from 375-395 in cMePAL1 and reverse from 1200 to 1180) designed (by the author) from cMePAL1. The intron was 582bp in length and located in between the second and third codon of arginine as it was in most PAL genes in plants (Fig.4.7). A specific PCR was obtained from the amplification as only one band appeared in gel analysis. It was later revealed, when gPAL2 was isolated, that the primers were specific to cMePAL1 compared to gPAL2 (see Appendix A4 & A5). The gMePAL1 or the genomic DNA corresponding to the known sequences of cMePAL1 was amplified from cassava genomic DNA by PCR with primers designed from cMePAL1 and sequenced (by Hongying Li). There was only one intron in the coding region of this gene.

Another PAL genomic DNA clone or gMePAL2 was isolated from cassava genomic DNA library using a probe made from cMeP1Eb fragment and characterised (by Hongying Li, unpublished results). For the convenience of comparison and discussion, the sequence of gMePAL2 is presented in Appendix A.1) (by the courtesy of Hongying Li).

```
   1  CTCAAACTTCCTATACCTATATTTTTCTCCCACTTTCTCACCACTCCTCTACCTTTCCGG   60
  61  CAAAATAAATCAGCAAAATGGAGTTTTGTGAGGCTCACAATGTAACTGCTTCACCTGGAT  120
                     M  E  F  C  E  A  H  N  V  T  A  S  P  G  F
 121  TCTCCTCTGCTGACCCATTGAACTGGGGCATGGCTGCAGAGTCACTGAAGGGCAGCCACC  180
        S  S  A  D  P  L  N  W  G  M  A  A  E  S  L  K  G  S  H  L
 181  TTGATGAGGTGAAGCGCATGGTGGATGAGTACAGGAAGCCTGTGGTGAGGCTAGGGGGTG  240
        D  E  V  K  R  M  V  D  E  Y  R  K  P  V  V  R  L  G  G  E
 241  AGACCCTGACTATAGCCCAAGTTACAGCAATTGCGAACCATGACTCAGGTGTCAAGGTTG  300
        T  L  T  I  A  Q  V  T  A  I  A  N  H  D  S  G  V  K  V  E
 301  AGCTGTCTGAGGAGGCTCGAGCTGGGGTCAAGGCCAGTAGTGACTGGGTTCTTGATTCCA  360
        L  S  E  E  A  R  A  G  V  K  A  S  S  D  W  V  L  D  S  M
                 375→
 361  TGAATAAAGGAACCGATAGTTATGGTGTCACCACTGGCTTTGGTGCAACCTCCCATAGAA  420
        N  K  G  T  D  S  Y  G  V  T  T  G  F  G  A  T  S  H  R  R
 421  GAACCAAGCAGGGGGGTGCCCTTCAGAGAGAACTCATAAG
        T  K  Q  G  G  A  L  Q  R  E  L  I  R
                                              gtaacaaatattaaaaatat  480
 481  ataaatatttataaactagttgattattttgagatcatccctataatagactagatccat  540
 541  aatggataaaatgatgaaagtacaagtttctttttctttttcttaaaagggtaatttaaaa  600
 601  aaatatttttatcagaaaataaataaataaagagtacacgaaacgaaactgccttctact  660
 661  aagtctaaatgcaaatcatctaccccaacagtactgttaattaataatctaaattatttt  720
 721  acagtattttcaccatcccattaaaatagatatattcatttttttatataatttaaaaaa  780
 781  tataattaaaaataatacatttatattgccttttttagtaatatatacattactcatatta  840
 841  ctcaattattaaatcatttttttaaattgataaatttttatttttttattatatattaaataa  900
```

```
901  gggtatattaaaaatattaaatgatagattcatataaatagctaattccaacaaatctga   960

961  aattaaaaatttaattgacttgagcattgtttcttgaacaaatggttagtaatttaaatg  1020

1021 tttttcttggaaatttttgttttattctccag

                                    ATTCTTGAATGCTGGGATCTTTGGAAAT  1080
                                      F  L  N  A  G  I  F  G  N

1081 GGACAAGAATCTTGCCACACATTGTCTCACACTGCAACTAGAGCAGCAATGCTGGTGAGA  1140
     G  Q  E  S  C  H  T  L  S  H  T  A  T  R  A  A  M  L  V  R

                                                          ←1200.
1141 ATCAACACCCTCCTGCAAGGTTATTCAGGCATAAGATTTGAAATCCTGGAAGCCATTACC  1200
     I  N  T  L  L  Q  G  Y  S  G  I  R  F  E  I  L  E  A  I  T

1201 AAGTTCATCAACAACAATGTTACTCCGCGTTTGCCCCTCAGAGGCACAATCACAGCCTCT  1260
     K  F  I  N  N  N  V  T  P  R  L  P  L  R  G  T  I  T  A  S

1261 GGTGACCTGGTCCCACTGTCCTACATTGCCGGGCTTTTGACCGGCCGGCCCAACTCCAAG  1320
     G  D  L  V  P  L  S  Y  I  A  G  L  L  T  G  R  P  N  S  K

1321 TCGTTAGGGCCCAATGGAGAATCCTTGGATGCAGCTGAAGCCTTTAAGCTTGCTGGGATC  1380
     S  L  G  P  N  G  E  S  L  D  A  A  E  A  F  K  L  A  G  I

1381 AATGGTGGATTTTTTGAATTGCAGCCCAAAGAGGGTCTAGCTTTAGTAAATGGTACTGCA  1440
     N  G  G  F  F  E  L  Q  P  K  E  G  L  A  L  V  N  G  T  A

1441 GTTGGTTCAGGCTTAGCTTCTATGGTTCTTTTTGAGGCCAATGTGTTGGCAGTCCTGTCA  1500
     V  G  S  G  L  A  S  M  V  L  F  E  A  N  V  L  A  V  L  S

1501 GAAGTTTTATCAGCAATTTTTGCAGAAGTTATGCTTGGAAAACCAGAGTTTACAGATCAT  1560
     E  V  L  S  A  I  F  A  E  V  M  L  G  K  P  E  F  T  D  H

1561 TTGACACATAAATTGAAGCATCATCCTGGACAAATTGAAGCTGCAGCAATTATGGAACAT  1620
     L  T  H  K  L  K  H  H  P  G  Q  I  E  A  A  A  I  M  E  H

1621 GTCTTGGATGGAAGTTCTTATATTAAAGCAGCTCAAAAGGTTCATGAAATTGATCCATTG  1680
     V  L  D  G  S  S  Y  I  K  A  A  Q  K  V  H  E  I  D  P  L

1681 CAGAAGCCTAAACAAGACAGATATGCTCTTAGAACATCTCCTCAATGGCTTGGCCCTCAG  1740
     Q  K  P  K  Q  D  R  Y  A  L  R  T  S  P  Q  W  L  G  P  Q

1741 ATTGAAGTGATTCGAACAGCGACTAAAATGATCGAACGCGAAATCAACTCTGTGAATGAT  1800
     I  E  V  I  R  T  A  T  K  M  I  E  R  E  I  N  S  V  N  D

1801 AATCCATTGATTGATGTCTCCAGGAATATTGCTTTACATGGAGGCAATTTCCAGGGGACC  1860
     N  P  L  I  D  V  S  R  N  I  A  L  H  G  G  N  F  Q  G  T

1861 CCAATTGGTGTTTCAATGGATAACACTCGTTTAGCCATTGCTTCAATTGGTAAACTCATG  1920
     P  I  G  V  S  M  D  N  T  R  L  A  I  A  S  I  G  K  L  M

1921 TTTGCTCAATTCTCTGAGCTTGTTAATGATTTTTACAACAATGGGTTGCCTTCAAATCTC  1980
     F  A  Q  F  S  E  L  V  N  D  F  Y  N  N  G  L  P  S  N  L

1981 ACTGGTGGACGCAATCCAAGCTTGGATTATGGCTTCAAAGGAGCTGAAATTGCCATGGCA  2040
     T  G  G  R  N  P  S  L  D  Y  G  F  K  G  A  E  I  A  M  A

2041 TCTTACTGCTCAGAGCTCCAATTTCTTGCCAATCCTGTAACTAATCATGTCCAAAGTGCA  2100
     S  Y  C  S  E  L  Q  F  L  A  N  P  V  T  N  H  V  Q  S  A

2101 GAGCAGCACAACCAAGATGTTAACTCACTAGGCTTGATTTCTTCAAGGAAAACAGCTGAA  2160
     E  Q  H  N  Q  D  V  N  S  L  G  L  I  S  S  R  K  T  A  E

2161 GCTGTTGACATATTGAAGCTCATGTCTTCTACATACTTAGTTGCTCTATGTCAAGCCATT  2220
     A  V  D  I  L  K  L  M  S  S  T  Y  L  V  A  L  C  Q  A  I

2221 GACTTGAGACACTTGGAGGAGAACTTGAAGCAAACAGTCAAGAACACAGTAAGTCAAGTT  2280
     D  L  R  H  L  E  E  N  L  K  Q  T  V  K  N  T  V  S  Q  V

2281 GCAAAGAGAGTCTTGACAATGGGCATCAACGGCGAGCTCCACCCGTCGAGATTCTGCGAA  2340
     A  K  R  V  L  T  M  G  I  N  G  E  L  H  P  S  R  F  C  E

2341 AAAGACCTTCTCAAAGTCGTCGACAGGGAATACGTTTATGCATATGTTGATGATCCTTGC  2400
     K  D  L  L  K  V  V  D  R  E  Y  V  Y  A  Y  V  D  D  P  C

2401 AGTGCAACATACCCATTAATGCAAAAGCTGAGACAAGTACTAGTTGATCATGCCATGATG  2460
     S  A  T  Y  P  L  M  Q  K  L  R  Q  V  L  V  D  H  A  M  M

2461 AATGGTGAAAAGGAGAAGAATTCAAGCACTTCCATTTTCCAAAAAATTGGAGCCTTTGAA  2520
     N  G  E  K  E  K  N  S  S  T  S  I  F  Q  K  I  G  A  F  E

2521 GAAGAACTCAAGACCCTTTTGCCTAAAGAAGTAGAAAGTGCAAGAACTGAATATGAGAAT  2580
     E  E  L  K  T  L  L  P  K  E  V  E  S  A  R  T  E  Y  E  N
```

```
2581  GGTAATCCAGCTATTTCTAACAAGATCAAAGAATGTAGGTCATATCCACTATACAAGTTT  2640
      G  N  P  A  I  S  N  K  I  K  E  C  R  S  Y  P  L  Y  K  F
2641  GTGAGGGAAGAACTCGGTTGTAGTTTACTGACCGGCGAGAAGATTCGATCGCCCGGCGAA  2700
      V  R  E  E  L  G  C  S  L  L  T  G  E  K  I  R  S  P  G  E
2701  GAGTTTGATAAGGTTTTCTCAGCAATTTGTGCAGGGAAGCTGATTGATCCCATGCTTGAA  2760
      E  F  D  K  V  F  S  A  I  C  A  G  K  L  I  D  P  M  L  E
2761  TGCCTCAAGGAGTGGAATGGTGCTCCTCTTCCAATCTGCTAAGTTTATTTTTTTTTGTTT  2820
      C  L  K  E  W  N  G  A  P  L  P  I  C  *
2821  GTTTGCTCATAAAGTCTACACATGC  2845
```

Fig.4.7 Nucleotide sequence of the MePAL1 gene. The sequence from MePAL1 cDNA is in upper case. Codons for translation initiation and termination are in bold at position 78 and 2802. The intron region (461 to 1052) is shown in lower case. The splicing sites are indicated with bold and underlined letters. The amino acid residues are placed under the first nucleotides in the codons. The PAL signature is indicated with bold and double-underlined letters. The primers used for PCR amplification of the intron of gMePAL1 were indicated with arrows.
Note: The intron region is identified and sequenced by Hongying Li.

4.2.3 PAL Gene Organisation

In order to estimate the number of PAL gene family members in cassava, Southern blot hybridisation of genomic DNA was performed. Cassava genomic DNA prepared from cultivar MNGA1 was digested with restriction enzymes *Kpn*I, *Eco*RV and *Xba*I which did not cut the three PAL genes in the known sequences. Two sets of digestions were carried out and the cleaved DNA samples of both sets were then run in one agarose gel in parallel. The DNA was then transferred to Hybond N^+ membrane and the two sets were separated for different hybridisation. Set A was hybridised with probes made from the 1.9kb *Eco*RI fragment or cMeP1Eb (Fig.4.8**A**) and set B with gMePAL2 *Eco*RII/*Hin*dIII fragment (2702-2959) or gMeP2E/H (Fig.4.8 **B**) at 60°C overnight. The membrane was washed at high stringency (60°C, 0.1xSSC, 0.1%SDS, 2x15min for **A** & **B**) and exposed to X-ray film. The probes on membranes A and B were stripped off and the resulted membranes were named E and D, were reprobed with cMeP1Eb and gMeP2E/H respectively. The membranes were washed at low stringency (60°C, 1xSSC, 0.1%SDS, 2x10min) (Fig.4.8 E & D). The membrane D was reprobed with gPAL2 intron part (*Hinc*II/*Eco*RII fragment), washed at high stringency and exposed to X-ray film (Fig.4.8C).

When the cassava genomic DNA blot was hybridised with the probe made from cMePAL1 1.9kb *Eco*RI fragment and washed at high stringency, specific signals to the MePAL1 probe appeared as there was only one band in each digestion (Fig.4.8A). If the blot was washed at low stringency, more bands appeared in each digestion (Fig.4.8E).

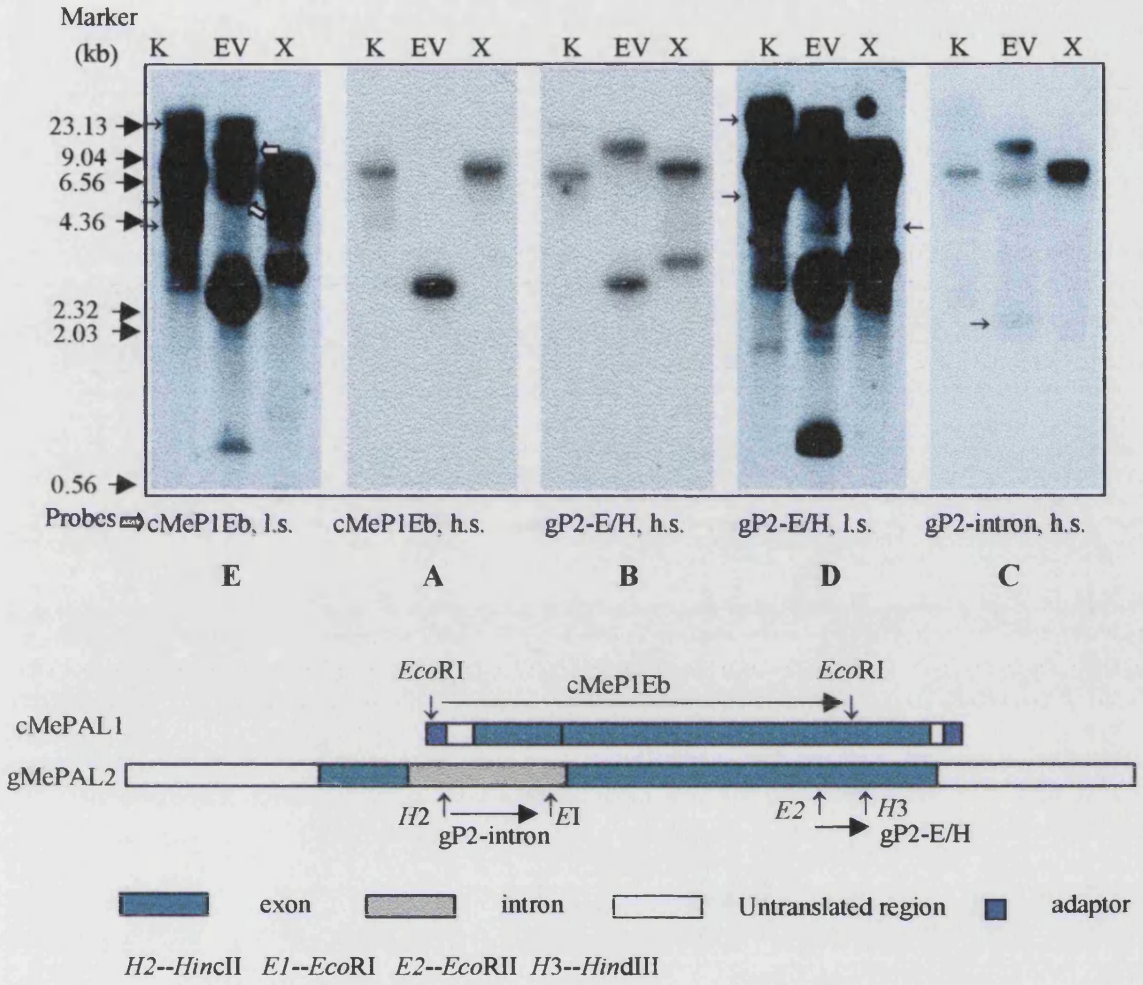


Fig.4.8 PAL gene organisation in cassava. About 10μg of genomic DNA was digested with different restriction enzymes *Kpn*I, *Eco*RV and *Xba*I. The cleaved DNA was then run in 1xTBE 1% agarose gel overnight at 1volt/cm. DNA was transferred to Hybond N+ by blotting overnight. After hybridization with probes made from the 1.9kb *Eco*RI fragment of cPAL1 (A & D) or gPAL2 *Eco*RII/*Hin*dIII fragment (B & E) at 60°C overnight, the membrane was washed at low or high stringency (l.s. or h.s.) The probe on the used membrane was stripped off with boiling solution of 0.1xSSC and 0.1%SDS, and the membrane was reprobed with gPAL2 intron part (*Hin*cII/*Eco*RI fragment) and washed at high stringency. **A, B & C:** high stringency 60°C, 0.1xSSC, 0.1%SDS, 2x15min. **D & E:** low stringency 60°C, 1xSSC, 0.1%SDS, 2x10min. ⇒, ⇐ or ← in D and E indicate the positions of obvious bands on the films, which are not so clear after the images were processed as presented here. DNA markers used were HindIII cut λ DNA.

Note: the similarity between cMeP1Eb and gMePAL2 is 79.68%; the similarity between gP2-E/H and cMePAL1 is 79.85%.

Similarly, more bands appeared when blot B (hybridised with MePAL2 probe from *Eco*RII/*Hin*dIII fragment or gP2-E/H) was washed at low stringency (Fig.4.8D) than at high stringency (Fig.4.8B). The signal patterns with the two probes (of MePAL1 & 2) at low stringency were similar.

Two bands appeared in lane EV and X of blot B (hybridised with the MePAL2 probe) at high stringency (Fig.4.8B). Were the small band in lane EV and big band in lane X from MePAL1 binding to gP2-E/H probe? Because the similarity between MePAL1 and MePAL2 is 79.68% and 79.85% respectively in the regions of the two probes cMePE1b and gP2-E/H, probe cMePeb1 would bind to MePAL2 at the same strength as probe gP2-E/H to MePAL1. Therefore, the different results in blot A and B indicated that the high stringency was adequate to discriminate non-specific binding and no signals in blot B were from MePAL1. Since there is no restriction site for *Eco*RV and *Xba*I in the known sequence of gMePAL2, only one of the two bands in either lane EV or X is specific to MePAL2. As the known sequence of gMePAL2 is 4791bp (see Appendix A1) and the smaller band in lane either EV or X is less than 4 kb, the bigger band in lane EV and X must be from gMePAL2. In order to confirm the specificity of these bands to gMePAL2, a *Hin*cII/*Eco*RI fragment from the intron of gMePAL2 was used to probe the B blot at high stringency. Indeed, the bigger band in EV or X showed strong signal (Fig.4.8C), which were of the same position as the bigger bands in EV or X in Fig.4.8B. Yet there were also weak signals. Both B and C suggested that there is another PAL gene which shares very high similarity with MePAL2, maybe even in the intron region. If there were no other PAL in cassava which share high identity with gMePAL2 in intron region, there would have been just one band on C, though it is very rare that the sequences of introns are of high similarity.

MePAL3 was 98% similar to MePAL2, suggesting that the signals on film B were from these two PAL genes. The appearance of two smaller bands in lane EV on C may suggest that there is an *Eco*RV site in the intron. It must be noted that there was only one band in lane K in both B and C, and in lane X of C, which indicates that the two PAL genes are not only high similar in sequence but may also be clustered together.

Considering that more bands appeared in low stringency in addition to the bands in A and B, it might be suggested that there might be more than four genes in the PAL gene family in cassava.

4.2.4 Isolation of Other PAL cDNA Clones

As the Southern blot hybridization of genomic DNA indicated there were more PAL genes in the cassava genome (4.2.3), the amplified library of original PPD-related cDNA library was screened for clones corresponding to these genes. It was also the aim to isolate the cDNA clones of gMePAL2 and MePAL (PCR fragment from genomic DNA) (Periera *et al.*, 1999), the other part of cMePAL3.

The amplified library was screened with cMePAL1 1.9kb *Eco*RI probe by hybridizing at 60°C overnight and washing finally with 1xSSC/0.1%SDS at 60°C for 2x10min. From the attempt no PAL clones were retrieved, but two interesting clones were identified which encoded putative ubiquitin-activating enzyme and receptor-like protein kinase or serine-threonine protein kinase. These two genes are briefly described in Chapter 7. The false positives might have resulted from the low stringency wash after the hybridization. There was 77% similarity between the sequences of cMePAL1 1.9kb fragment and kinase cDNA (gap weight penalty 2 and length weight penalty 0.3 during the comparison), 69% over 140bp between cMePAL1 1.9kb fragment and the cDNA clone encoding ubiquitin-activating enzyme (gap weight penalty 3 and length weight penalty 0.3) (Appendix A.2 & A.3).

4.2.5 Comparison of Nucleotide Sequences and Deduced Amino Acid Sequences of PAL Genes in Cassava

PAL genes including cMePAL1, gMePAL2, cMePAL3 and MePAL (Periera *et al.*, 1999) were compared at the level of nucleotide sequence and amino acid sequence (Table 4.1). Sequences of cMePAL3 and MePAL were not full length, corresponding to the second halves of cMePAL1 and gMePAL2 (Fig. 4.9). At the nucleotide sequence level, four genes were of high similarity and ranged from 79% to 84% between each other except that gMePAL2 and cMePAL3 shared very high similarity (98%). The deduced amino acid sequences of four genes were highly conserved, with similarity ranged from 93% to about 100%. The high similarity among four cassava genes in the level of amino acid sequence was also demonstrated by the multiple alignment (Fig.4.10)

Table 4.1 Comparison of sequences and amino acid sequences of the PAL genes in cassava

aa% \ bp%	cMePAL1	gMePAL2*	cMePAL3	MePAL
CMePAL1	**710aa \\ 2253bp**	79%	79%	83%
gMePAL2*	93%	**712aa \\ 2136bp**	98%	83%
cMePAL3	93%	~100%	**316aa \\ 1126bp**	84%
MePAL	95%	97%	97%	**173aa \\ 520bp**

Note: For the convenience of comparison only the coding region of gMePAL2 was compared with others. The sizes of different PAL genes and the length of polypeptides are in bold. The figures below the bold letters are the similarity percentage of amino acid residues; above the bold letters are similarity percentage of nucleotides. The comparison was made using the BESTFIT program in GCG.

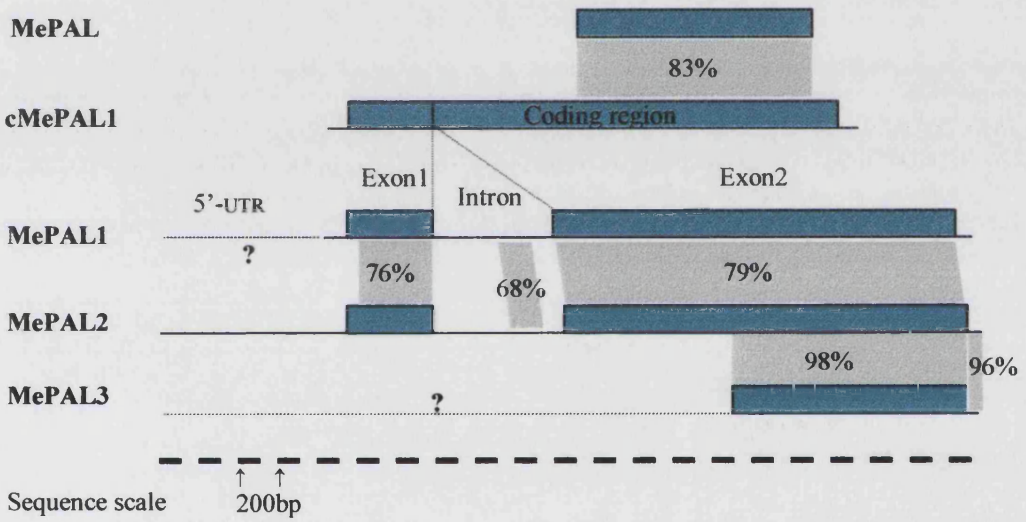
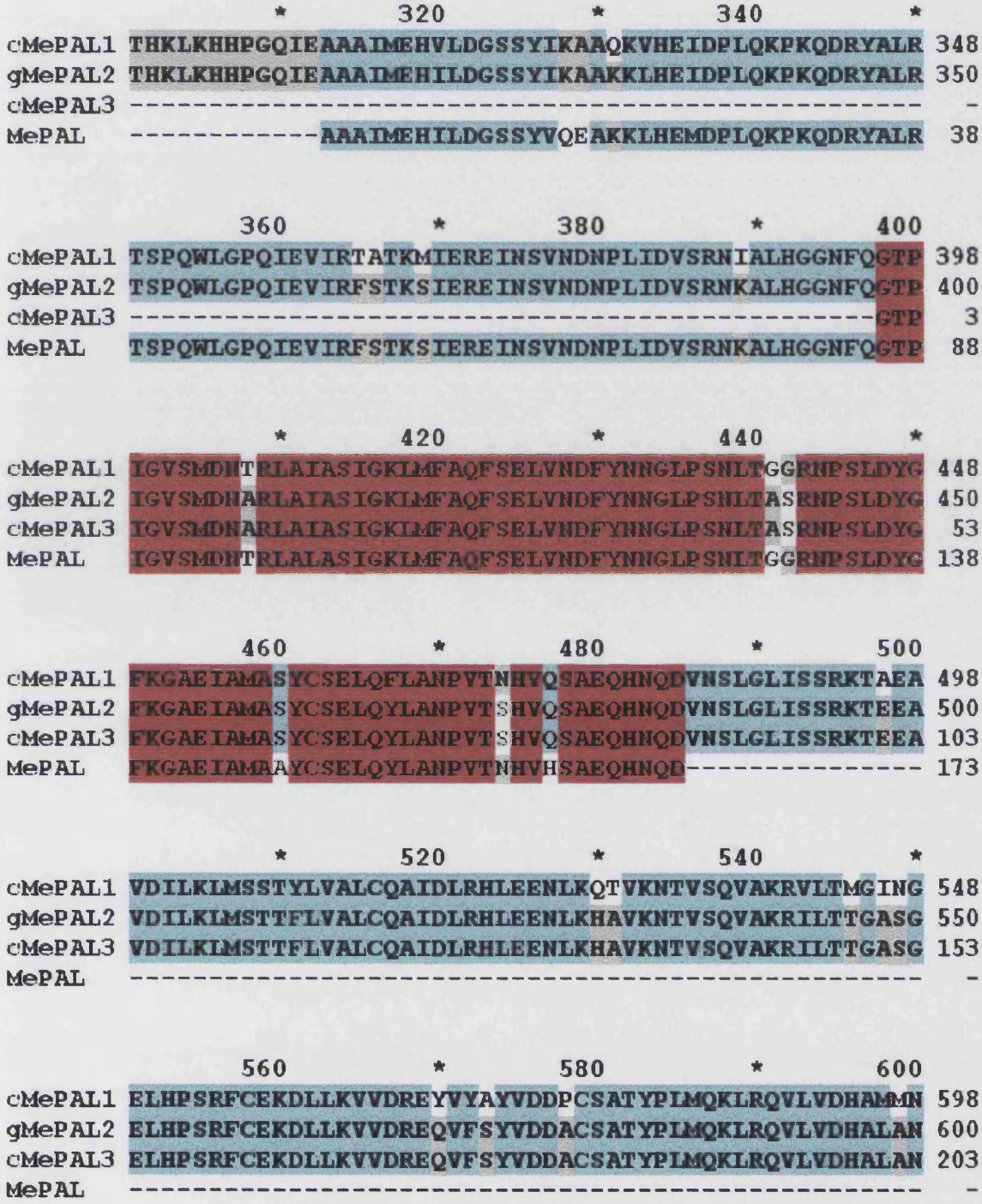


Fig. 4.9 Schematic diagram of the MePAL1, MePAL2 and MePAL3 sequences. The figures (in percentage) in the grey shades are the similarity in nucleotide sequence between the sequence blocks of different PAL genes. The dot lines indicate the undetermined sequences. 5'UTR refers to 5' untranslated region.

```
                  *         20          *         40          *
cMePAL1 --MEFCEAHNVTASPGFSSADPLNWGMAAESLKGSHLDEVKRMVDEYRKP  48
gMePAL2 MATISQNGHQNGSLDSLCTARDPLNWGLAAESMSGSHLDEVKKMVAEFRK  50
cMePAL3 ------------------------------------------------   -
MePAL   ------------------------------------------------   -

                60          *         80          *         100
cMePAL1 VVRLGGETLTIAQVTAIANHDSGVKVELSEEARAGVKASSDWVLDSMNKG  98
gMePAL2 PLVKLGGETLTVAQVAAIARESGLQVELAESARAGVKASSDWVMDSMSKG 100
cMePAL3 ------------------------------------------------   -
MePAL   ------------------------------------------------   -

                *         120         *         140         *
cMePAL1 TDSYGVTTGFGATSHRRTKQGGALQRELIRFLNAGIFGNGQESCHTLSHT 148
gMePAL2 TDSYGVTTGFGATSHRRTKQGGALQRELIRFLNAGIFGNKTESCHTLSHS 150
cMePAL3 ------------------------------------------------   -
MePAL   ------------------------------------------------   -

              160         *         180         *         200
cMePAL1 ATRAAMLVRINTLLQGYSGIRFEILEAITKFINNNVTPRLPLRGTITASG 198
gMePAL2 ATRAAMLVRINTLLQGYSGIRFEILEAITKLLNHNITPCLPLRGTITASG 200
cMePAL3 ------------------------------------------------   -
MePAL   ------------------------------------------------   -

                *         220         *         240         *
cMePAL1 DLVPLSYIAGLLTGRPNSKSLGPNGESLDAAEAFKLAGINGGFFELQPKE 248
gMePAL2 DLVPLSYIAGLLTGRPNSKAVGPNGESLDAQQAFHSAGIDSGFFELQPKE 250
cMePAL3 ------------------------------------------------   -
MePAL   ------------------------------------------------   -

              260         *         280         *         300
cMePAL1 GLALVNGTAVGSGLASMVLFEANVLAVLSEVLSAIFAEVMLGKPEFTDHL 298
gMePAL2 GLALVNGTAVGSGLASMVLFEANVLAVLSEVLSAIFAEVMNGKPEFTDHL 300
cMePAL3 ------------------------------------------------   -
MePAL   ------------------------------------------------   -
```


320
340
cMePAL1 THKLKHHPGQIEAAAIMEHVLDGSSYIKAAQKVHEIDPLQKPKQDRYALR 348
gMePAL2 THKLKHHPGQIEAAAIMEHILDGSSYIKAAKKLHEIDPLQKPKQDRYALR 350
cMePAL3 - 3
MePAL AAAIMEHILDGSSYVQEAKKLHEMDPLQKPKQDRYALR 38
360
380
400
cMePAL1 TSPQWLGPQIEVIRTATKMIEREINSVNDNPLIDVSRNIALHGGNFQGTP 398
gMePAL2 TSPQWLGPQIEVIRFSTKSIEREINSVNDNPLIDVSRNKALHGGNFQGTP 400
cMePAL3 GTP 3
MePAL TSPQWLGPQIEVIRFSTKSIEREINSVNDNPLIDVSRNKALHGGNFQGTP 88
420
440
cMePAL1 IGVSMDHTRLAIASIGKLMFAQFSELVNDFYNNGLPSNLTGGRNPSLDYG 448
gMePAL2 IGVSMDHARLAIASIGKLMFAQFSELVNDFYNNGLPSHLTASRNPSLDYG 450
cMePAL3 IGVSMDNARLAIASIGKLMFAQFSELVNDFYNNGLPSNLTASRNPSLDYG 53
MePAL IGVSMDHTRLALASIGKLMFAQFSELVNDFYNNGLPSNLTGGRNPSLDYG 138
460
480
500
cMePAL1 FKGAEIAMASYCSELQFLANPVTNHVQSAEQHNQDVNSLGLISSRKTAEA 498
gMePAL2 FKGAEIAMASYCSELQYLANPVTSHVQSAEQHNQDVNSLGLISSRKTEEA 500
cMePAL3 FKGAEIAMASYCSELQYLANPVTSHVQSAEQHNQDVNSLGLISSRKTEEA 103
MePAL FKGAEIAMAAYCSELQYLANPVTNHVHSAEQHNQD 173
520
540
cMePAL1 VDILKLMSSTYLVALCQAIDLRHLEENLKQTVKNTVSQVAKRVLTMGING 548
gMePAL2 VDILKLMSTTFLVALCQAIDLRHLEENLKHAVKNTVSQVAKRILTTGASG 550
cMePAL3 VDILKLMSTTFLVALCQAIDLRHLEENLKHAVKNTVSQVAKRILTTGASG 153
MePAL - 3
560
580
600
cMePAL1 ELHPSRFCEKDLLKVVDREYVYAYVDDPCSATYPLMQKLRQVLVDHAMMN 598
gMePAL2 ELHPSRFCEKDLLKVVDREQVFSYVDDACSATYPLMQKLRQVLVDHALAN 600
cMePAL3 ELHPSRFCEKDLLKVVDREQVFSYVDDACSATYPLMQKLRQVLVDHALAN 203
MePAL - 3

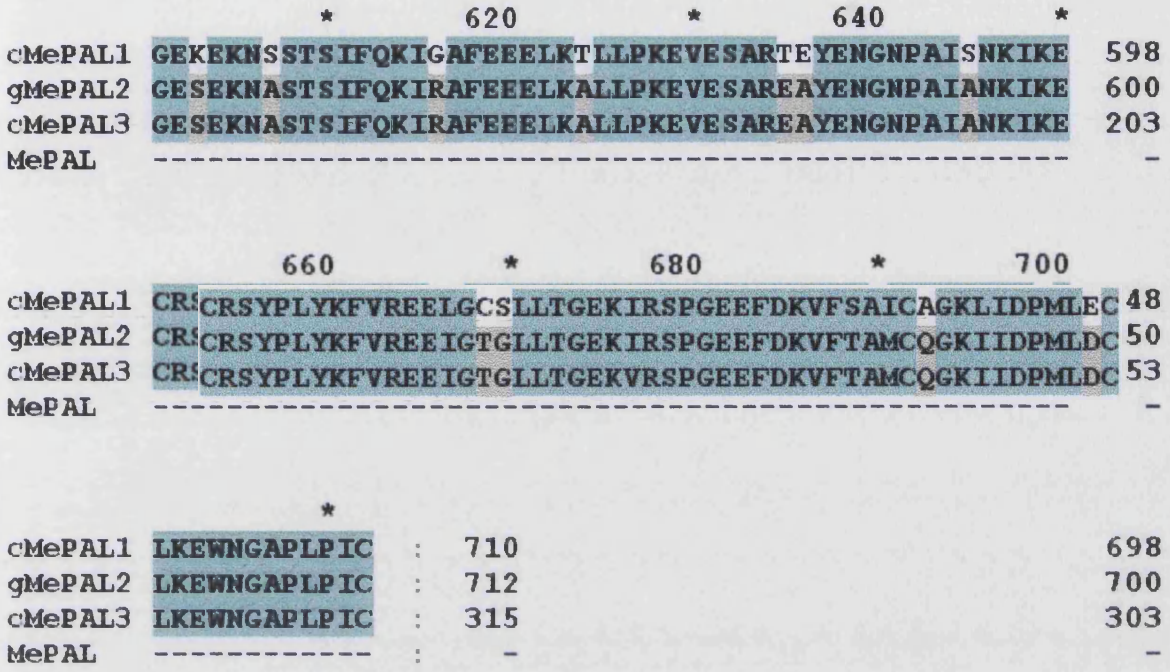


Fig.4.10 Multiple alignment of the deduced amino acid sequences of four cassava PAL genes. The residues in red shading are conserved among all the sequences; those in green shading conserved among three sequences (the fourth sequence maybe have not been determined in these regions); those in grey shading conserved between two sequences.

Between gMePAL1 and gMePAL2, the nucleotide sequences in the 5' untranslated region and the first 50bp of the exon1 were very divergent and there was no similarity. In the remaining region of exon1 and the region of exon2, the similarity between gMePAL1 and gMePAL2 in nucleotide sequences was 76% and 79% respectively (Appendix A.4 & 5). Usually the sequences of introns are very divergent among the genes in the same family. There was no similarity in most part of the introns of gMePAL1 and gMePAL2 but there was one 140bp region in which these two genes share 68% similarity (Fig.4.11). There was no similarity in the 3' untranslated region between these two genes. An overall nucleotide sequence alignment of gMePAL1 and gMePAL2 is shown in Appendix A.6.

```
MePAL1   854 TCATTTTTTAAATTGATAAATTTTATTTTTTA..TTATATATTAAATAAG 901
MePAL2  1851 TTATTATTAAATTAGTAAAATTTAATTATTTAGTTAATATATTAATCAAT 1900

         902 GGTATATTAAAAAT.ATTAAATGATA..............GATTCATA 934
        1901 TGAAAGAGAGATCTAACTAAATGATACTTTTGGCAGCCTTAGAATTCATA 1950

         935 ..TAAATAGCTAATTCCAACAAATCTGAAATTAAAAATTT 972
        1951 TTTTCCTAGAAAATTCATCCAAAACAGTAACTAAACATGT 1990
```

Fig.4.11 Comparison of the intron sequences of MePAL1 and MePAL2.

The comparison of the deduced amino acid sequences between cMePAL1 and gMePAL2 showed 93% similarity. The most divergent part was found in the N-terminal regions as has been noted in other PAL protein sequences (Fig.4.10). The higher identity and similarity in amino acid sequences than the similarity in nucleotide sequences meant that some of the differences in nucleotide sequences did not change the amino acid residue they encoded or changed the codon to a similar amino acid residue, as was also observed between gMePAL2 and cMePAL3 (described below).

MePAL had high similarity (83-84%) to both cMePAL1 and gMePAL2 or cMePAL3 DNA sequences, higher than the similarity (79%) between cMePAL1 and gMePAL2 or cMePAL3. At amino acid sequence level, MePAL showed higher similarity (95-97%) to cMePAL1 and gMePAL2 or cMePAL3 than the similarity (93%) between cMePAL1 and gMePAL2 or cMePAL3.

It was shown above that cMePAL3 and gMePAL2 were of extremely high similarity. Since gMePAL2 was isolated from a genomic library constructed from cassava cultivar MBRA 534, while cMePAL3 was isolated from the PPD-related cDNA library

constructed from cultivar MNGA1, it is possible that cMePAL3 might be the corresponding cDNA of gMePAL2. The difference in the sequences might be from the variance in the genome between the two cultivars. This hypothesis was supported with the fact that the deduced amino acid sequences of gMePAL2 and cMePAL3 were almost identical except for one residue I (isoleucine) or V (valine) which was coded by ATC at 3645 (gMePAL2) or GTC at 830 (cMePAL3). In this particular position of PAL amino acid sequences in other plants the amino acid was either I or V (position 687, Fig.4.12). Therefore, the differences in nucleotides between gMePAL2 and cMePAL3 might happen at the third nucleotide of the codon for amino acid residue. Indeed, ten out of total eleven nucleotide changes in the coding region of these two sequences happened in the third nucleotide and did not change the deduced amino acid sequence (Fig.4.12). The only difference between them was in the first nucleotide of the codon which changed the amino acid residue from I to V.

```
          A20rr→   .         .         .         .         .
gMePAL2 1782 tggactcctgttcctgtagctgtaggtagatttgcccattttgtttgact 1832
             ...       ...       ...       ...       ...       ...
                      .         .         .         .         .
gMePAL2 2815 GGGGACCCCAATTGGAGTCTCAATGGATAATGCACGTTTGGCCATTGCAT 2864
             ||||||||||||||||||||||||||||||||||||||||||||||||||||
cMePAL3 1    GGGGACCCCAATTGGAGTCTCAATGGATAATGCACGTTTGGCCATTGCAT 50
                      .         .         .         .         .
        2865 CAATAGGAAAGCTCATGTTTGCTCAGTTCAGTGAGCTTGTAAATGATTTT 2914
             ||||||||||||||||||||||||||||||||||||||||||||||||||||
        51   CAATAGGAAAGCTCATGTTTGCTCAGTTCAGTGAGCTTGTAAATGATTTT 100
                      .         .         .         .         .
        2915 TACAACAATGGGTTGCCATCAAATCTCACAGCCAGCAGGAATCCAAGCTT 2964
             ||||||||||||||||||||||||||||||||||||||||||||||||||||
        101  TACAACAATGGGTTGCCATCAAATCTCACAGCCAGCAGGAATCCAAGCTT 150
              .         .         .         .         .
        2965 GGATTATGGCTTCAAGGGAGCTGAAATTGCAATGGCTTCTTACTGTTCTG 3014
             ||||||||||||||||||||||||||||||||||||||||||||||||||||
        151  GGATTACGGCTTCAAGGGAGCTGAAATTGCAATGGCTTCTTACTGTTCTG 200
                  Y
                   .         .         .         .         .
        3015 AGCTCCAATACCTTGCAAATCCAGTCACCAGCCATGTACAAAGTGCAGAG 3064
             ||||||||||||||||||||||||||||||||||||||||||||||||||||
        201  AGCTCCAATACCTTGCAAATCCAGTCACCAGCCATGTACAAAGTGCAGAG 250
                      .         .         .         .         .
        3065 CAGCACAATCAAGATGTAAATTCATTGGGGCTAATTTCTTCAAGAAAGAC 3114
             ||||||||||||||||||| || ||||||||||||||||||||||||||||
        251  CAGCACAATCAAGATGTAAACTCCTTGGGGCTAATTTCTTCAAGAAAGAC 300
                        N  S
                      .         .         .         .         .
        3115 AGAAGAAGCTGTAGACATCTTGAAGCTCATGTCCACGACTTTCTTAGTAG 3164
             ||||||||||||||||||||||||||||||||||| ||||||||||||||
        301  AGAAGAAGCTGTAGACATCTTGAAGCTCATGTCCACAACTTTCTTAGTAG 350
                                        T
                      .         .         .         .         .
        3165 CACTTTGTCAAGCTATTGACTTGAGGCATTTGGAGGAGAACTTGAAGCAC 3214
             |||||||||||||||||||||||||||||||||||||||||||||||||||
        351  CACTTTGCCAAGCTATTGACTTGAGGCATTTGGAGGAGAACTTGAAGCAC 400
               C   .       ←cggttcatcgattctcctaag ra6      .
        3215 GCAGTCAAAAACACAGTAAGCCAAGTAGCTAAGAGGATTCTAACTACAGG 3264
             ||||||||||||||||||||||||||||||||||||||||||||||||||||
        401  GCAGTCAAAAACACAGTAAGCCAAGTAGCTAAGAGGATTCTAACTACAGG 450
```

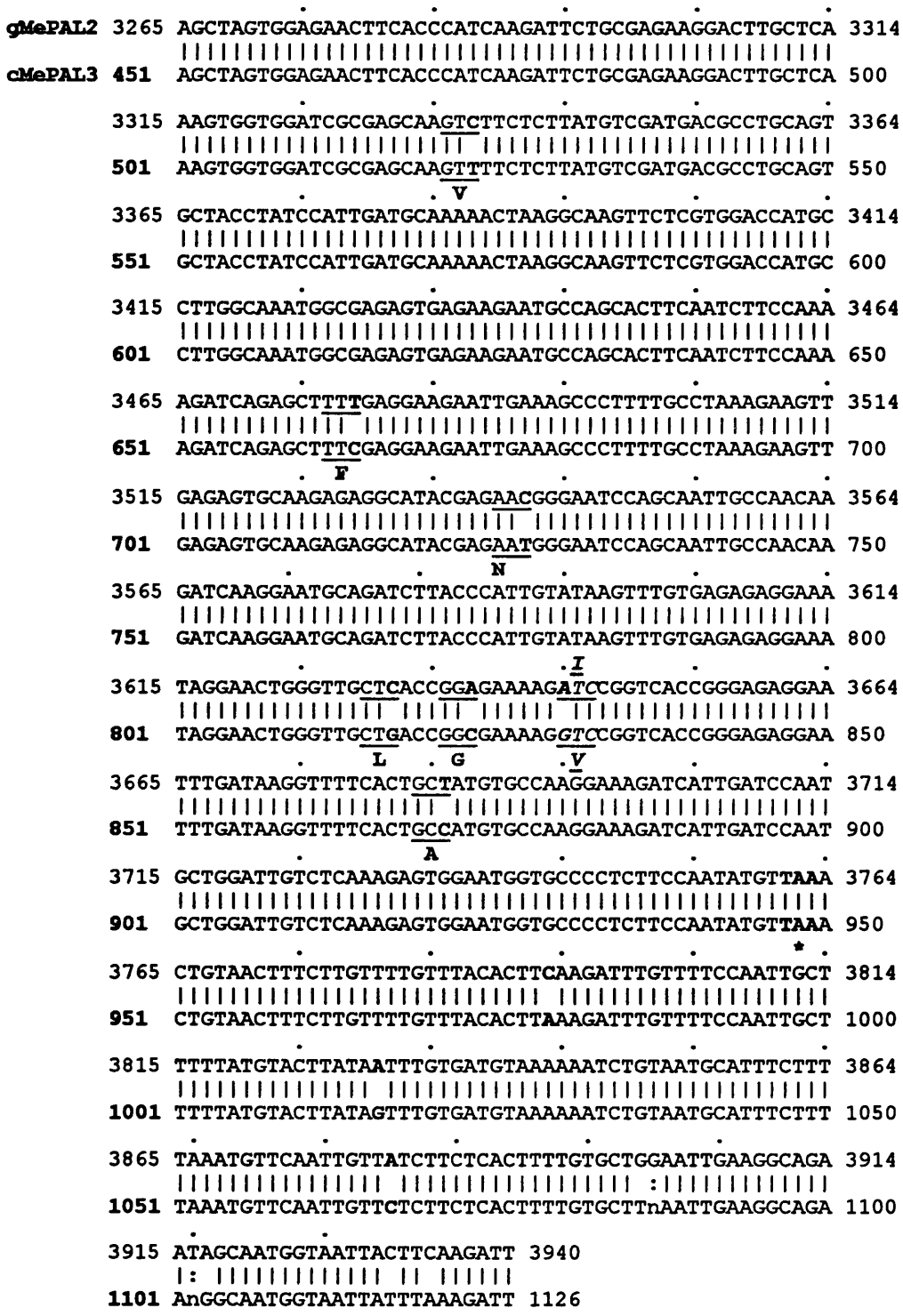


Fig.4.12 The comparison of nucleotide sequences between gMePAL2 and cMePAL3. The partial sequence of gMePAL2 intron shown is in lower case. The different nucleotides between these two sequences are in bold and underlined together with the other two nucleotides, which formed a code for an amino acid residue (the bold letter below the underlined code). The deduced amino acid sequences of gMePAL2 and cMePAL3 are almost identical except the ones coded by ATC at 3645 (gMePAL2) or GTC at 830 (cMePAL3). The translation stop codon was in bold and marked with *. The nucleotides in blue were the primers used for amplifying genomic DNA of gMePAL2 from cutivar MNGA1. The sequence from the PCR product with *ra6* covered the region from 2815 to 3184 of gMePAL2, indicated with blue bars.

These data raised the question as to whether gMePAL2 from MBRA 534 and cMePAL3 from MNGA1 represent different members of the PAL gene family or whether the minor nucleotide differences between them, which caused minimal amino acid residue differences, merely represents expected sequence variation between cultivars in the same gene. In order to address this question attempt was made to amplify and sequence the equivalent gene to gMePAL2 (from MBRA 534) in MNGA1.

Primers *a20rr* (located in the intron of gMePAL2 and specific to gMePAL2 by comparing with gMePAL1 intron) and *ra6* used for sequencing gMePAL2 were applied to amplify from genomic DNA prepared from cultivar MNGA1, the PAL gene fragment corresponding to the region from 1782 to 3254bp of gMePAL2 (Fig.4.12). The PCR product was of the expected size (1472bp) based on the region that the primers covered. The sequence of the PCR product was determined by using *ra6* primer. The sequence covered the region from 2815 to 3184 of gMePAL2. Comparison between these sequences revealed that the PCR sequence (from MNGA1 with gMePAL2 primers) was the same as the sequence of gMePAL2 (from MBRA 534) in the corresponding region, whereas in the region there were five nucleotide differences between cMePAL3 (from MNGA1) and gMePAL2. At least in the investigated region of gMePAL2 there was no difference between cultivar MBRA 534 and MNGA1. It is interesting to mention that gMePAL2 and cMePAL3 shared high similarity even in 3' untranslated region (Fig.4.12). It could be concluded that gMePAL2 and cMePAL3 were very similar at the level of both DNA and amino acid sequences but were not from the same gene. The comparison between the partial sequences of the same gene gMePAL2 from different cultivars MBRA 534 and MNGA1 also indicated that there was no divergence between cultivars in the analysed region.

4.2.6 Comparison between PAL Genes in Cassava and in Other Plants

Since gMePAL2 and cMePAL3 differ by just a few base pairs in nucleotide sequences and by one amino acid residue, and because cMePAL3 is not full length, cMePAL3 was not included in the comparison with PAL genes from other plants.

The similarity of cassava PALs to other PALs in other plants was demonstrated when the deduced amino acid sequences of cMePAL1 and gMePAL2 were compared to those of PAL genes from other plants (Tab.4.2). Both cassava PAL genes showed more than

60% identity and similarity to most PAL genes with which they were compared. Both cassava PALs showed higher identity and similarity to PALs from other dicotyledon plants than to those from rice, barley, pine and fungi. The gMePAL2 shared slightly higher identity or similarity to other dicots or other plants in amino acid sequences.

Tab. 4.2 The similarity and identity of deduced amino acid sequences between cassava PAL genes and plant PAL genes. (data from statistics of multiple alignment in Genedoc program)

Gene	PAL1	PAL2	Seq. length[b]	Seq. names	Accession NO.
Cassava PAL1	100/100[a]	85 / 91	710	cMePAL1	
Cassava PAL2	85 / 91	100/100	712	gMePAL2	
Poplar-ka[d]	83 / 90	88 / 94 √[c]	715		D30656
Poplar-kb	83 / 90	88 / 94 √	715	Pkpala	D30657
Poplar-g2b	86 / 93 √	82 / 89	710	Pkpalg2b	D43802
Lemon	85 / 92 √	83 / 90	722		U43338
Tea	84 / 91 √	82 / 90	714		D26596
Tobacco-1	83 / 91	84 / 92 √	715	Nttpal1a	M84466
Tobacco-3	83 / 91	84 / 92 √	712	ntpheal	X78269
Parsley-4	83 / 91 √	83 / 90	712	Pcpal4	L37357
Soybean-1	83 / 90	85 / 91 √	713	DMpal1	X52953
Tobacco-2	83 / 90	84 / 92 √	712	Ntpal1	D17467
Parsley-3	82 / 91	82 / 91	718	Pcpal3	X81159
Pea-2	82 / 90	83 / 91 √	724	Pspal2	X81158
S. humilis[d]	82 / 90	83 / 90 √	715	shpal	L36822
Clover	82 / 90	83 / 90 √	725	Tspal	M91192
Carrot	82 / 90	82 / 90	708		D85850
Alfalfa	82 / 88	82 / 90 √	725	Mspal	X58180
Tomato-5a	81 / 90	83 / 91 √	721	Lepal5a	M90692
Foxglove[d]	81 / 90	82 / 90 √	713	DLJ00222	AJ002221
A. thaliana-2	81 / 89	82 / 90 √	717		L33678
Pea-1	81 / 89	82 / 90 √	723	Pspal1	D10002
Parsley-2	81 / 89	81 / 90 √	716	Pcpal2	D10003
Tomato	80 / 89 √	79 / 88	703	Leapheaml	M83314
Sweet potato	80 / 88	81 / 88 √	707	Ibpal	M29232
Sweet potato	80 / 87	80 / 87	708	Ib640	D78640
Avocado	75 / 81 √	74 / 80	620	Pa16130	U16130
A. thaliana-3	72 / 82 √	71 / 81	695		L33679
Poplar-g4b	71 / 76 √	69 / 74	571	Pkpalg4b	D43803
Rice	67 / 79	67 / 79	701	Ospal	X16099
Pine	62 / 75 √	61 / 75	754	Pt39792	U39792
Kidney Bean	60 / 65	61 / 66 √	505	pvpal	M11939
Barley-7mr	59 / 65	59 / 65	549	Hvpal7mr	Z49147
V. vinifera	51 / 55	51 / 55	416	vvpal	X75967
R. toruloides	16 / 26	16 / 26	347	Rtpal (fungi)	M18261
Potato-1	12 / 14	13 / 15 √	138	Stpal1	X63103
Potato-2	12 / 14	13 / 14 √	141	Stpal2	X63104

[a] x / y means percentage of sequence identity or percentage of sequence similarity.

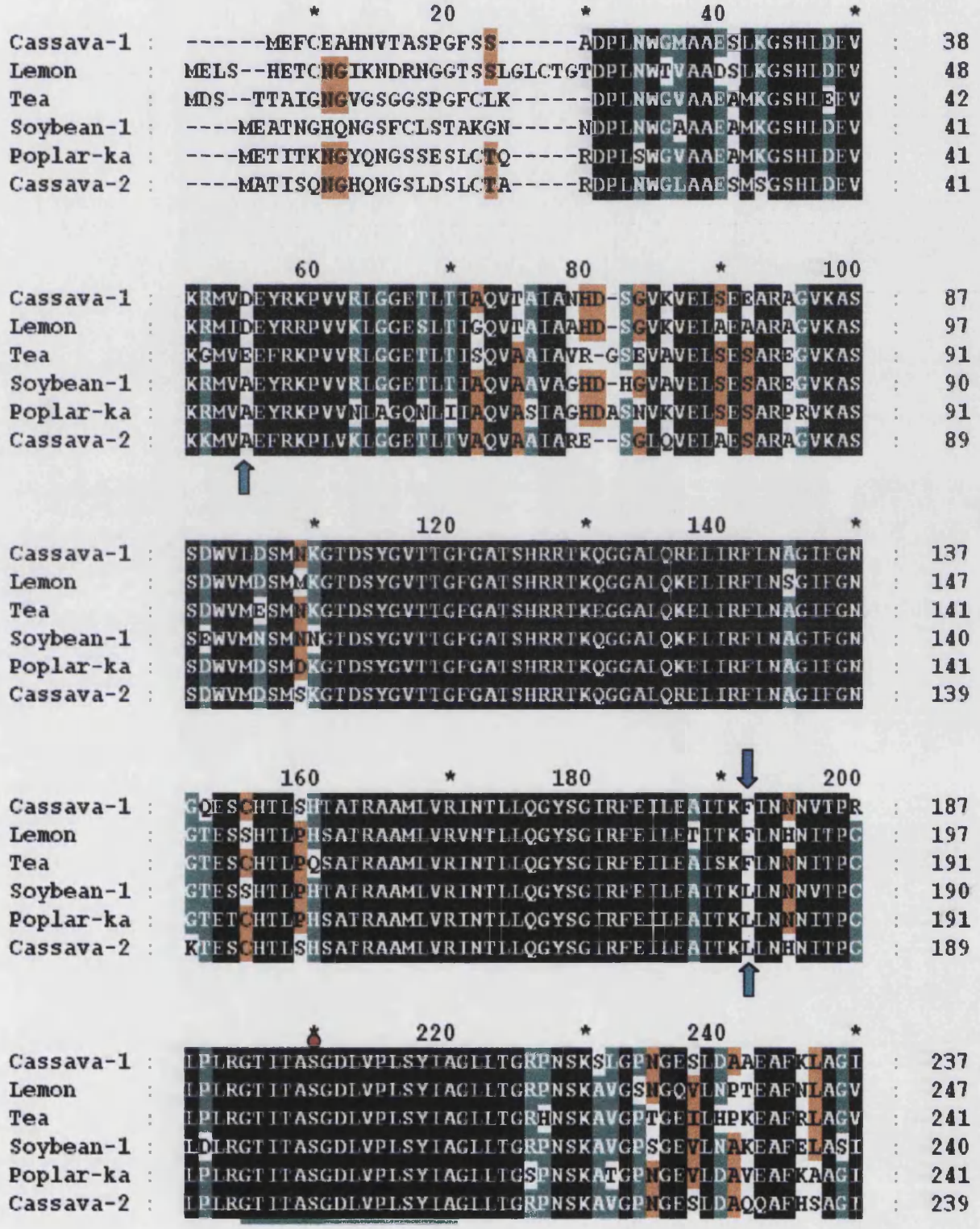
[b] Length of the deduced PAL amino acid sequence used in the comparison.

[c] √ means the this protein of cassava PAL has higher similarity or identity to the compared PAL than that of other cassava PAL.

[d] Poplar PAL was from poplar kitakamiensis. *S. humilis* for *Stylosanthes humilis*, *A. thaliana* for *Arabidopsis thaliana*, *V. vinifera* for *Vitis vinifera* and *R. toruloides* for *Rhodosporidium toruloides* (fungus). Foxglove is *Digitalis lanata*.

Multiple alignment was performed to illustrate further the conservation in amino acid sequences among PAL genes (Fig.4.13). PAL genes from *Limon citrus*, tea, soybean and poplar, showed high similarity to cassava cMePAL1 and gMePAL2, and therefore were selected to make the alignment. PAL genes in cassava, *Limon citrus*, tea, soybean and poplar had the same amino acid sequences in the phenylalanine ammonia-lyase active site (from residue 205 to 220). The PAL genes in these plants shared nearly the same deduced amino acid sequences except the short divergent N-terminal part and a few other residues. It is interesting to note that some of the unshaded residues or orange-shaded residues, which are not identical or not so similar among all six sequences in biochemical properties, are actually conserved among the sub-group cPAL1-lemon-tea or cPAL2-soybean-poplar such as residues at 192, 342, 452, 453 and 558.

The phylogenetic relationship of PAL genes in plants was analysed. A phylogenetic tree can be built from multiple alignment of (deduced) amino acid sequences or nucleotide sequences. If it is built on the aligned amino acid sequences alone, the relationships may not be fully expressed in the tree, as some changes in nucleotide sequence cannot be reflected through amino acid sequences. If the tree is built from aligned nucleotide sequences, it may not be a true reflection of the phylogenetic relationship as it is of high probability that there are some mis-alignments because there are only four different nucleotides. A much more reliable approach was applied to construct phylogenetic tree of PAL genes in plants, which avoided these drawbacks. The deduced amino acid sequences of PAL genes were aligned first. The nucleotide sequences of PAL genes were overlayed onto the aligned amino acid sequences and well aligned. In other words, the nucleotide sequences were aligned according to the multiple alignment of amino acid sequences. Then the phylogenetic tree was constructed from the aligned nucleotide sequences using Jukes and Canter (1969) distance calculation model and neighbour-joining algorhythm of the TREECON program (Van de Peer and De Wachter, 1993) with fungus *Rhodosporidium toruloides* as an out-group (Fig.4.14). The phylogenetic tree showed that the cassava PAL genes were very close to those of dicot tree plants such as poplar and *Limon citrus* in the evolution of PAL genes.


Cassava-1 : ------MEFCEAHNVTASPGFSS------ADPLNWGMAAESLKGSHLDEV : 38
Lemon : MELS--HETCNGIKNDRNGGTSSLGLCTGTDPLNWTVAADSLKGSHLDEV : 48
Tea : MDS--TTAIGNGVGSGGSPGFCLK------DPLNWGVAAEAMKGSHLEEV : 42
Soybean-1 : ----MEATNGHQNGSFCLSTAKGN-----NDPLNWGAAAEAMKGSHLDEV : 41
Poplar-ka : ----METITKNGYQNGSSESLCTQ-----RDPLSWGVAAEAMKGSHLDEV : 41
Cassava-2 : ----MATISQNGHQNGSLDSLCTA-----RDPLNWGLAAESMSGSHLDEV : 41
Cassava-1 : KRMVDEYRKPVVRLGGETLTIAQVTAIANHD-SGVKVELSEEARAGVKAS : 87
Lemon : KRMIDEYRRPVVKLGGESLTIGQVTAIAAHD-SGVKVELAEAARAGVKAS : 97
Tea : KGMVEEFRKPVVRLGGETLTISQVAAIAVR-GSEVAVELSESAREGVKAS : 91
Soybean-1 : KRMVAEYRKPVVRLGGETLTIAQVAAVAGHD-HGVAVELSESAREGVKAS : 90
Poplar-ka : KRMVAEYRKPVVNLAGQNLIIAQVASIAGHDASNVKVELSESARPRVKAS : 91
Cassava-2 : KKMVAEFRKPLVKLGGETLTVAQVAAIARE--SGLQVELAESARAGVKAS : 89
Cassava-1 : SDWVLDSMSKGTDSYGVTTGFGATSHRRTKQGGALQRELIRFLNAGIFGN : 137
Lemon : SDWVMDSMMKGTDSYGVTTGFGATSHRRTKQGGALQKELIRFLNSGIFGN : 147
Tea : SDWVMESMSKGTDSYGVTTGFGATSHRRTKEGGALQKELIRFLNAGIFGN : 141
Soybean-1 : SEWVMNSMNNGTDSYGVTTGFGATSHRRTKQGGALQKELIRFLNAGIFGN : 140
Poplar-ka : SDWVMDSMDKGTDSYGVTTGFGATSHRRTKQGGALQKELIRFLNAGIFGN : 141
Cassava-2 : SDWVMDSMSKGTDSYGVTTGFGATSHRRTKQGGALQRELIRFLNAGIFGN : 139
Cassava-1 : GQESGHTLSHTAIRAAMLVRINTLLQGYSGIRFEILEAITKFINNNVTPR : 187
Lemon : GTESSHTLPHSAIRAAMLVRVNTLLQGYSGIRFEILETITKFLNHNITPC : 197
Tea : GTESGHTLPQSAIRAAMLVRINTLLQGYSGIRFEILEAISKFLNNNITPC : 191
Soybean-1 : GTESSHTLPHTAIRAAMLVRINTLLQGYSGIRFEILEAITKLLNNNVTPC : 190
Poplar-ka : GTETGHTLPHSAIRAAMLVRINTLLQGYSGIRFEILEAITKLLNNNITPC : 191
Cassava-2 : KTESGHTLSHSAIRAAMLVRINTLLQGYSGIRFEILEAITKLLNHNITPC : 189
Cassava-1 : LPLRGTIIASGDLVPLSYIAGLLTGRPNSKSLGPNGESLDAAEAFKLAGI : 237
Lemon : LPLRGTIIASGDLVPLSYIAGLLTGRPNSKAVGSNGQVLNPTEAFNLAGV : 247
Tea : LPLRGTIIASGDLVPLSYIAGLLTGRHNSKAVGPTGEILHPKEAFRLAGV : 241
Soybean-1 : LDLRGTIIASGDLVPLSYIAGLLTGRPNSKAVGPSGEVLNAKEAFELASI : 240
Poplar-ka : LPLRGTIIASGDLVPLSYIAGLLTGSPNSKATGPNGEVLDAVEAFKAAGI : 241
Cassava-2 : LPLRGTIIASGDLVPLSYIAGLLTGRPNSKAVGPNGESLDAQQAFHSAGI : 239

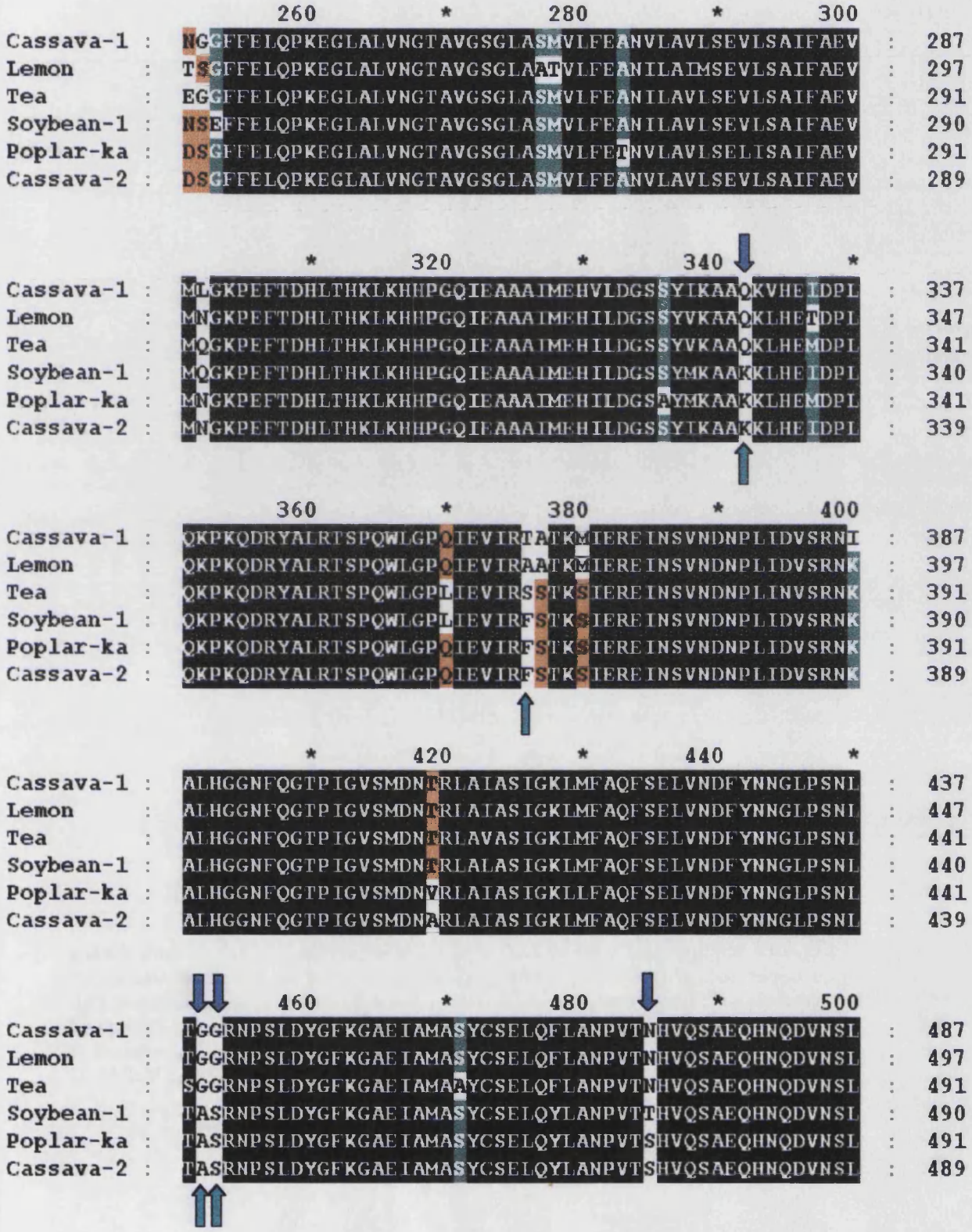
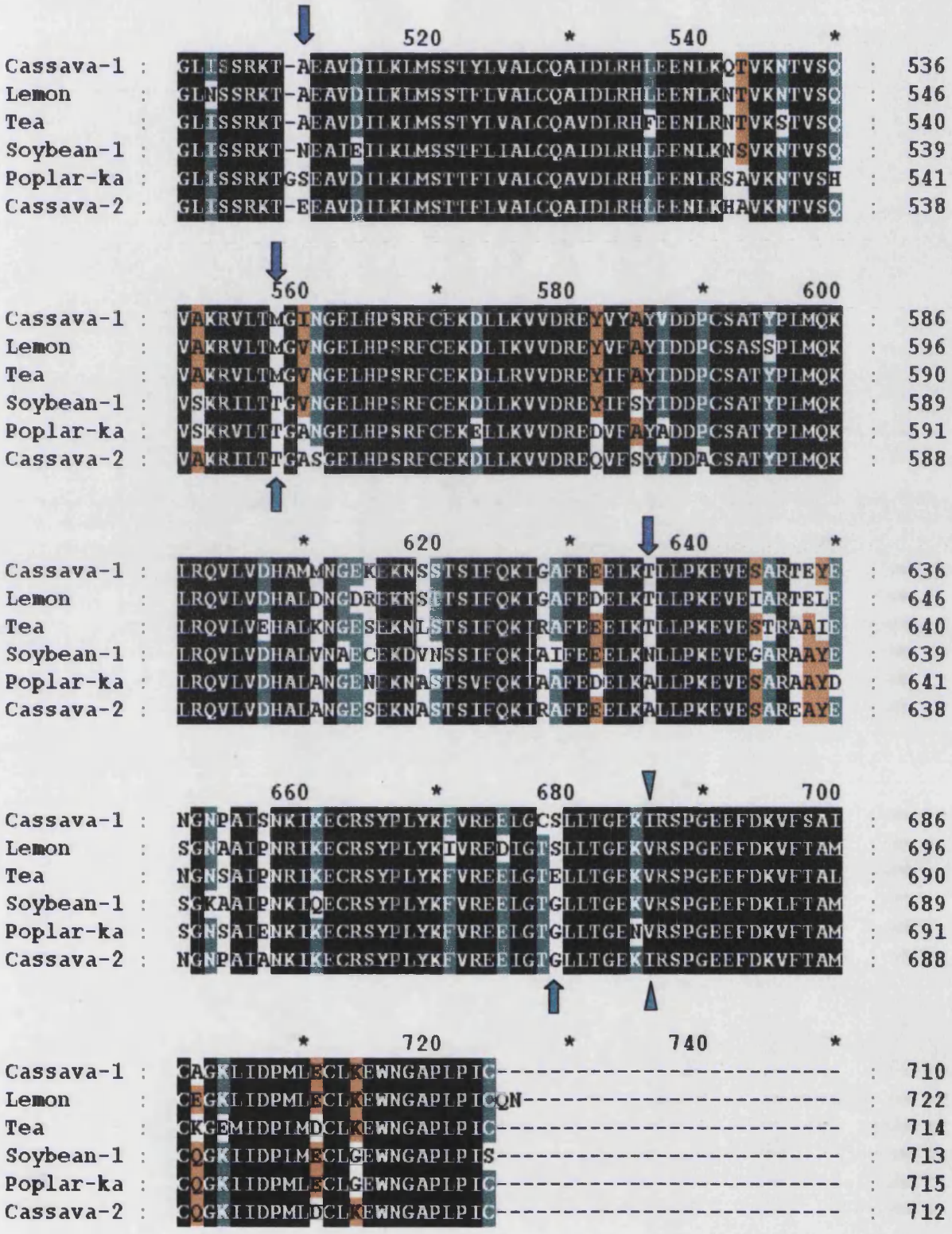

260 * 280 * 300
Cassava-1 : NGGFFELQPKEGLALVNGTAVGSGLASMVLFEANVLAVLSEVLSAIFAEV : 287
Lemon : TSGFFELQPKEGLALVNGTAVGSGLAATVLFEANILAIMSEVLSAIFAEV : 297
Tea : EGGFFELQPKEGLALVNGTAVGSGLASMVLFEANILAVLSEVLSAIFAEV : 291
Soybean-1 : NSEFFELQPKEGLALVNGTAVGSGLASMVLFEANILAVLSEVLSAIFAEV : 290
Poplar-ka : DSGFFELQPKEGLALVNGTAVGSGLASMVLFETNVLAVLSELISAIFAEV : 291
Cassava-2 : DSGFFELQPKEGLALVNGTAVGSGLASMVLFEANVLAVLSEVLSAIFAEV : 289
* 320 * 340 *
Cassava-1 : MLGKPEFTDHLTHKLKHHPGQIEAAAIMEHVLDGSSYIKAAQKVHEIDPL : 337
Lemon : MNGKPEFTDHLTHKLKHHPGQIEAAAIMEHILDGSSYVKAAQKLHETDPL : 347
Tea : MQGKPEFTDHLTHKLKHHPGQIEAAAIMEHILDGSSYVKAAQKLHEMDPL : 341
Soybean-1 : MQGKPEFTDHLTHKLKHHPGQIEAAAIMEHILDGSSYMKAAKKLHEIDPL : 340
Poplar-ka : MNGKPEFTDHLTHKLKHHPGQIEAAAIMEHILDGSAYMKAAKKLHEMDPL : 341
Cassava-2 : MNGKPEFTDHLTHKLKHHPGQIEAAAIMEHILDGSSYIKAAKKLHEIDPL : 339
360 * 380 * 400
Cassava-1 : QKPKQDRYALRTSPQWLGPQIEVIRTATKMIEREINSVNDNPLIDVSRNI : 387
Lemon : QKPKQDRYALRTSPQWLGPQIEVIRAATKMIEREINSVNDNPLIDVSRNK : 397
Tea : QKPKQDRYALRTSPQWLGPLIEVIRSSTKSIEREINSVNDNPLINVSRNK : 391
Soybean-1 : QKPKQDRYALRTSPQWLGPLIEVIRFSTKSIEREINSVNDNPLIDVSRNK : 390
Poplar-ka : QKPKQDRYALRTSPQWLGPQIEVIRFSTKSIEREINSVNDNPLIDVSRNK : 391
Cassava-2 : QKPKQDRYALRTSPQWLGPQIEVIRFSTKSIEREINSVNDNPLIDVSRNK : 389
* 420 * 440 *
Cassava-1 : ALHGGNFQGTPIGVSMDNTRLAIASIGKLMFAQFSELVNDFYNNGLPSNL : 437
Lemon : ALHGGNFQGTPIGVSMDNTRIAIASIGKLMFAQFSELVNDFYNNGLPSNI : 447
Tea : ALHGGNFQGTPIGVSMDNTRLAVASIGKLMFAQFSELVNDFYNNGLPSNL : 441
Soybean-1 : ALHGGNFQGTPIGVSMDNTRLALASIGKLMFAQFSELVNDFYNNGLPSNL : 440
Poplar-ka : ALHGGNFQGTPIGVSMDNVRLAIASIGKLLFAQFSELVNDFYNNGLPSNL : 441
Cassava-2 : ALHGGNFQGTPIGVSMDNARLAIASIGKLMFAQFSELVNDFYNNGLPSNL : 439
460 * 480 * 500
Cassava-1 : TGGRNPSLDYGFKGAEIAMASYCSELQFLANPVTNHVQSAEQHNQDVNSL : 487
Lemon : TGGRNPSLDYGFKGAEIAMASYCSELQFLANPVTNHVQSAEQHNQDVNSL : 497
Tea : SGGRNPSLDYGFKGAEIAMAAYCSELQFLANPVTNHVQSAEQHNQDVNSL : 491
Soybean-1 : TASRNPSLDYGFKGAEIAMASYCSELQYLANPVTTHVQSAEQHNQDVNSL : 490
Poplar-ka : TASRNPSLDYGFKGAEIAMASYCSELQYLANPVTSHVQSAEQHNQDVNSL : 491
Cassava-2 : TASRNPSLDYGFKGAEIAMASYCSELQYLANPVTSHVQSAEQHNQDVNSL : 489

Fig.4.13 Multiple alignment of PAL amino acid sequences. Residues in black shading are conserved among all the sequences; those in green-grey are conserved among five sequences; and those in orange are conserved among four sequences. The signature of PAL is underlined with bold green line and the active residue is marked with red dot. The upward or downward block arrows indicate the residues conserved among the first three sequences or the last three sequences. The sharp arrow indicates the different residue between MePAL1/2 and MePAL3.


520 * 540 *
Cassava-1 : GLISSRKT-AEAVDILKLMSSTYLVALCQAIDLRHLEENLKQTVKNTVSQ : 536
Lemon : GLNSSRKT-AEAVDILKLMSSTFLVALCQAIDLRHLEENLKNTVKNTVSQ : 546
Tea : GLISSRKT-AEAVDILKLMSSTYLVALCQAVDLRHFEENLRNTVKSTVSQ : 540
Soybean-1 : GLISSRKT-NEAIEILKLMSSTFLIALCQAIDLRHLEENLKNSVKNTVSQ : 539
Poplar-ka : GLISSRKTGSEAVDILKLMSTTFLVALCQAVDLRHLEENLRSAVKNTVSH : 541
Cassava-2 : GLISSRKT-EEAVDILKLMSTTFLVALCQAIDLRHLEENLKHAVKNTVSQ : 538
560 * 580 * 600
Cassava-1 : VAKRVLTMGINGELHPSRFCEKDLLKVVDREYVYAYVDDPCSATYPLMQK : 586
Lemon : VAKRVLTMGVNGELHPSRFCEKDLIKVVDREYVFAYIDDPCSASSPLMQK : 596
Tea : VAKRVLTMGVNGELHPSRFCEKDLLRVVDREYIFAYIDDPCSATYPLMQK : 590
Soybean-1 : VSKRILTTGVNGELHPSRFCEKDLLKVVDREYIFSYIDDPCSATYPLMQK : 589
Poplar-ka : VSKRVLTTGANGELHPSRFCEKELLKVVDREDVFAYADDPCSATYPLMQK : 591
Cassava-2 : VAKRILTTGASGELHPSRFCEKDLLKVVDREQVFSYVDDACSATYPLMQK : 588
* 620 * 640 *
Cassava-1 : LRQVLVDHAMMNGEKEKNSSTSIFQKIGAFEEELKTLLPKEVESARTEYE : 636
Lemon : LRQVLVDHALDNGDREKNSSTSIFQKIGAFEDELKTLLPKEVEIARTELE : 646
Tea : LRQVLVEHALKNGESEKNLSTSIFQKIRAFEEEIKTLLPKEVESTRAAIE : 640
Soybean-1 : LRQVLVDHALVNAECEKDVNSSIFQKIAIFEEELKNLLPKEVEGARAAYE : 639
Poplar-ka : LRQVLVDHALANGENEKNASTSVFQKIAAFEDELKALLPKEVESARAAYD : 641
Cassava-2 : LRQVLVDHALANGESEKNASTSIFQKIRAFEEELKALLPKEVESAREAYE : 638
660 * 680 * 700
Cassava-1 : NGNPAISNKIKECRSYPLYKFVREELGCSLLTGEKIRSPGEEFDKVFSAI : 686
Lemon : SGNAAIPNRIKECRSYPLYKIVREDIGTSLLTGEKVRSPGEEFDKVFTAM : 696
Tea : NGNSAIPNRIKECRSYPLYKFVREELGTELLTGEKVRSPGEEFDKVFTAL : 690
Soybean-1 : SGKAAIPNKIQECRSYPLYKFVREELGTGLLTGEKVRSPGEEFDKLFTAM : 689
Poplar-ka : SGNSAIENKIKECRSYPLYKFVREEIGTGLLTGENVRSPGEEFDKVFTAM : 691
Cassava-2 : NGNPAIANKIKECRSYPLYKFVREEIGTGLLTGEKIRSPGEEFDKVFTAM : 688
* 720 * 740 *
Cassava-1 : CAGKLIDPMLECLKEWNGAPIPIC---------- : 710
Lemon : CEGKLIDPMLECLKEWNGAPIPICQN---------- : 722
Tea : CKGEMIDPLMDCLKEWNGAPIPIC---------- : 714
Soybean-1 : CQGKIIDPLMECLGEWNGAPIPIS---------- : 713
Poplar-ka : CQGKIIDPMLECLGEWNGAPIPIC---------- : 715
Cassava-2 : CQGKIIDPMLDCLKEWNGAPIPIC---------- : 712

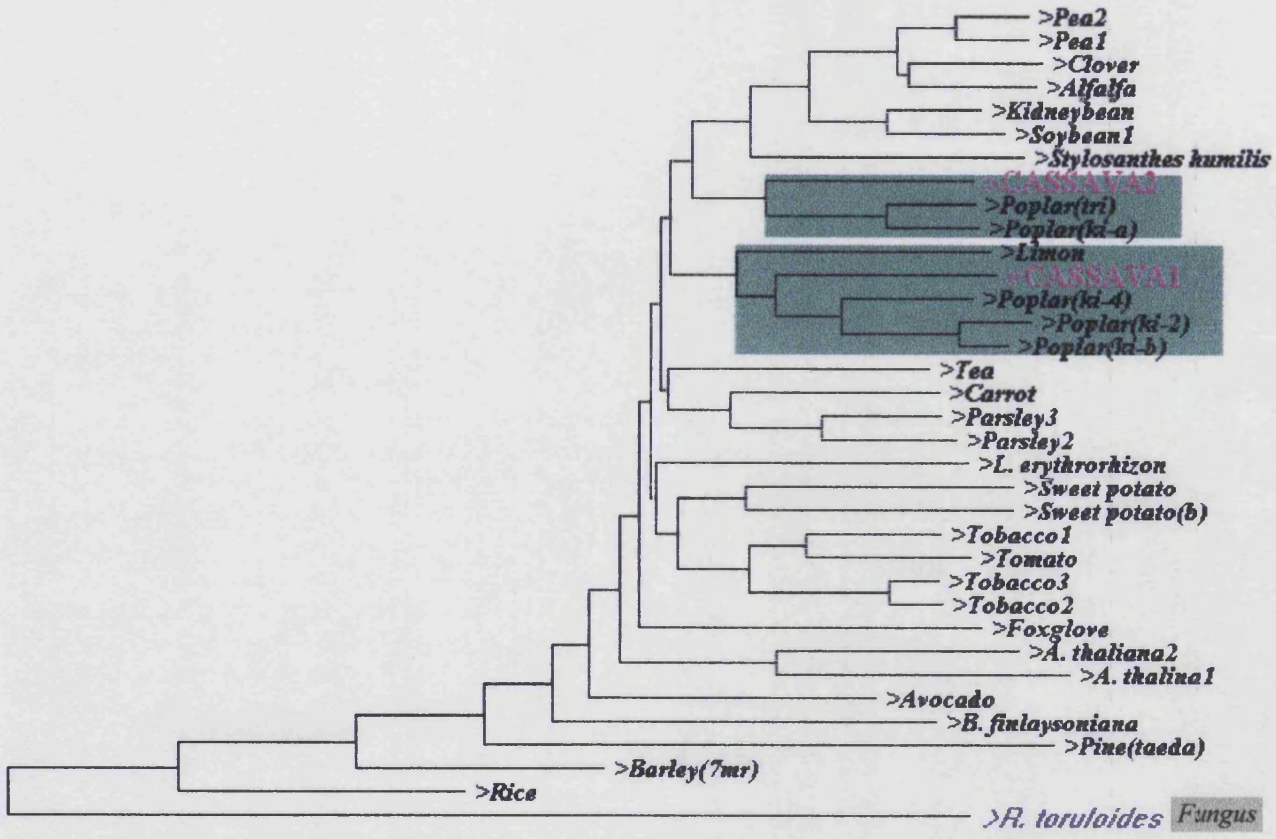


Fig.4.14 Phylogenetic analysis of PAL genes. The deduced amino acid sequences of PAL genes were aligned first. The nucleotide sequences of PAL genes were well aligned by overlaying them onto the aligned amino acid sequences. Then the aligned nucleotide sequences were used to construct the tree using neighbour-joining with fungus *Rhodosporidium toruloides* as an out-group. The groups to which cassava PAL1 and PAL2 belong are shaded. Note: the accession numbers for the PAL genes are illustrated in Table 4.1.

4.2.7 Expression of the PAL Genes

The isolation of PAL cDNAs cMePAL1 and cMePAL3 from the PPD-related cDNA library means that their mRNAs were present or these PAL genes were expressed in roots 48 hours after harvest. In order to understand the roles of these genes in the development and responses of cassava plant, especially in the development of PPD, the expression of MePAL1 in different organs and during PPD was investigated.

Many techniques have been developed to measure gene expression, including Northern hybridization, coupled reverse transcription and PCR amplification (RT-PCR), RNase protection assays, *in situ* hybridisation and dot blots. Northern hybridization is the conventional and most used method in gene expression studies. It is simple and straight-forward for analysing abundant mRNAs. But due to the low efficiency of the total RNA extraction from cassava storage roots and the limited availability of cassava roots, sufficient total RNA could not be prepared for Northern blot analysis at the stage when the expression experiments were to be carried out. RT-PCR is the most sensitive method among the techniques mentioned above, by which excellent amplification can be obtained using a total RNA template in the range of 10pg to 1μg, or mRNA template in the range of 1pg to 100ng. RT-PCR was chosen to study the gene expression with small amounts of total RNA because of its high sensitivity together with its requirement for minute amount of template RNA.

RT-PCR amplification is usually used to determine the presence or absence of a transcript or to clone cDNA products. In order to determine the abundance of a transcript, competitive RT-PCR has been developed. An RNA internal standard, which shares the same primer sequences as the target mRNA, is introduced into the RT-PCR reaction. The RT-PCR product from the internal standard can be distinguished from that of the target mRNA because of their difference in size. The use of an internal standard greatly improves the reproducibility of quantitation and competitive RT-PCR is the most accurate technique for quantitation of mRNA (Eferl *et al.*, 1997). As the focus of this research is on the expression patterns of PAL genes, competitive RT-PCR was not used to quantify the absolute mRNA abundance but to compare the levels of RT-PCR product from different treatments at the same concentration of internal standard.

4.2.7.1 The construction of competitor constructs for competitive RT-PCR and evaluation of the system

Before the construction, cMePAl1 specific primers were designed based on the sequence comparison between MePAL1 and MePAL2. The primers spanned the intron regions so that if there was any contamination of genomic DNA in the total RNA preparations there would be an extra RT-PCR product, which would be much larger in size. As the primers were located in the P1H/E fragment of pUC-P1H/E (subcloned from cMePAL1) and there was a unique *Xho*I site in the plasmid pUC-P1H/E in the middle of the region covered by the primers, the plasmid was used to make a deletion at the *Xho*I site in order to construct a competitor DNA of a smaller size (Fig.4.15). The plasmid was cut open with *Xho*I and exonuclease III was used to delete in both directions from the cut site. The deletion reactions were done in a time course of 15sec, 25sec, 35sec and 45sec, each of which was stopped by freezing the reaction tubes in liquid nitrogen to control tightly the deletion. Then the exonuclease was inactivated by heating at 75°C for 10min. Mung bean nuclease was used to blunt the overhang ends left after exonuclease deletion. The open plasmids were ligated back to generate functional plasmids and transformed into *E.coli* DH5α. Four transformants from each of the 15sec, 25sec and 35sec deletions were analysed by PCR amplification with the Spal (cMePAL1 specific, from 107 to 130) and Aspal (1128 to 1105) primers (see Appendix A.6 for location and specificity of these primers). Three transformants, d151 from 15sec deletion, d251 and d252 from 25sec deletion, showed slight deletions and were subject to plasmid extraction. Restriction digestion of these plasmids confirmed that deletion had occurred.

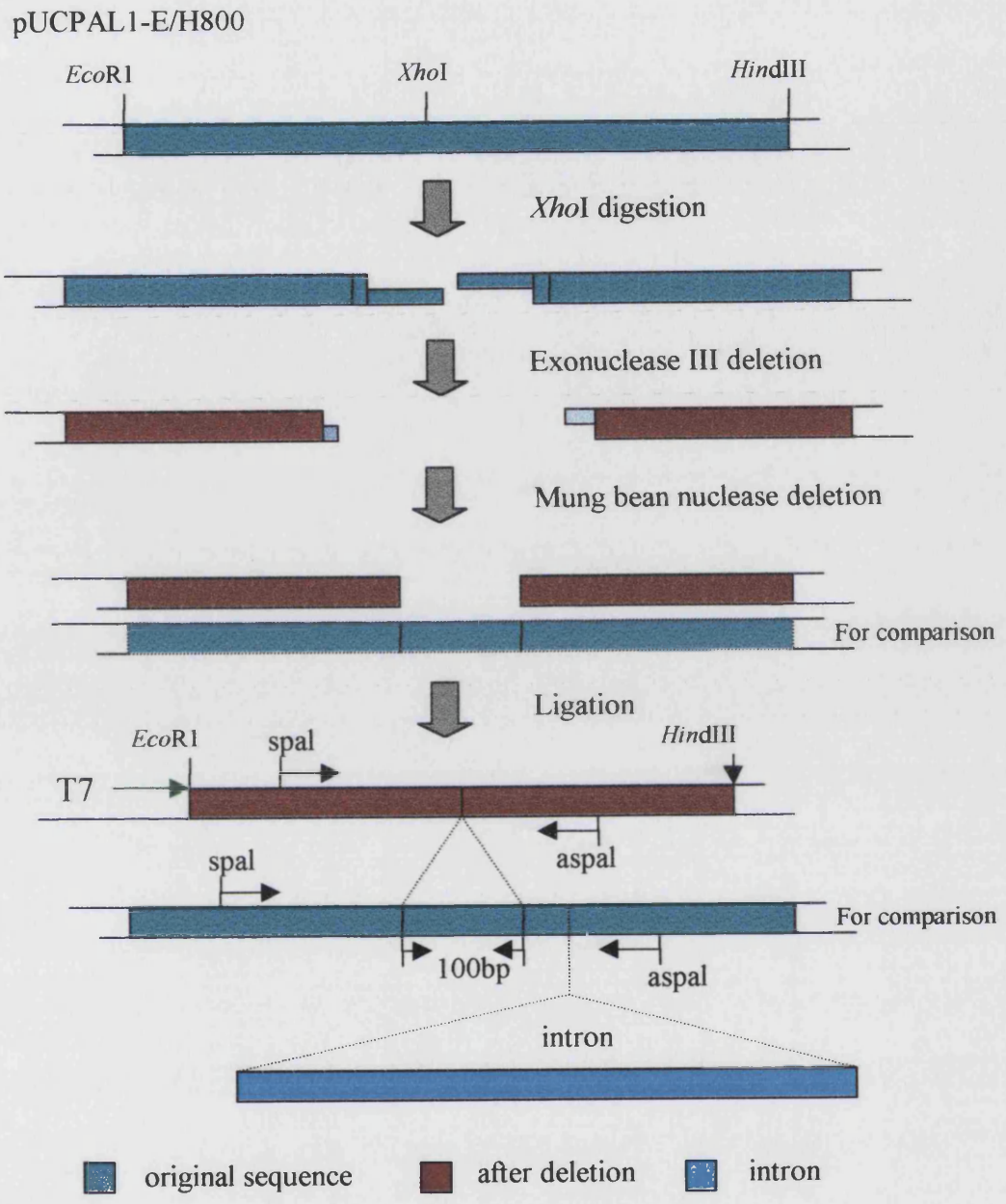


Fig.4.15 Construction of a control construct for competitive RT-PCR to analyse expression of cMePAL1 in cassava. spal and aspal are the primers to be used for the RT-PCR. Original DNA is aligned with the deleted ones to show the difference.

The deleted part of the plasmid d252 was sequenced using universal forward primer to check how many nucleotides were deleted. Comparison of the sequence from d252 with the cMePAL1 sequence shows that 100bp had been deleted (Fig4.16). The d252 deleted fragment was ideal for use as a competitor considering that the normal Spal-Aspal PCR

fragment was 431bp, because it is easy to separate the PCR product of the competitor from that of the testing template.

```
d252.rev 453 ACTATAGCCCAAGTTACAGCAATTGCGAACCATGACTCAGGTGTCAAGGT 404
             ||||||||||||||||||||||||||||||||||||||||||||||||||
cMePAL1  274 ACTATAGCCCAAGTTACAGCAATTGCGAACCATGACTCAGGTGTCAAGGT 323
                      .         .         .         .         .
         403 TGAGCTGTCT.......................................... 394
             ||||||||||        ↓ XhoI
cMePAL1  324 TGAGCTGTCTGAGGAGGCTCGAGCTGGGGTCAAGGCCAGTAGTGACTGGG 373

cMePAL1  374 TTCTTGATTCCATGAATAAAGGAACCGATAGTTATGGTGTCACCACTGGC 423
                      .         .         .         .         .
         393 ..........CCTCCCATAGAAGAACCAAGCAGGGGGGTGCCCTTCAGAG 354
                       ||||||||||||||||||||||||||||||||||||||||
cMePAL1  424 TTTGGTGCAACCTCCCATAGAAGAACCAAGCAGGGGGGTGCCCTTCAGAG 473
```

Fig.4.16 Comparison of the sequences between the competitor d252 and cMePAL1. Reverse sequence of d252 was used. The deleted region is in bold. The *Xho*I restriction site, from where the exonuclease III deleted, is in blue.

The deleted P1H/E fragment was subcloned into pBS KS(II)+, forming plasmid pBS-dP1H/E, in order to use the T7 or T3 promoter to prepare competitor RNA. To prepare competitor RNA, pBS-dP1H/E was cut with *Hind*III and *in vitro* transcription was performed using the T3 promoter to get sense dP1H/E RNA. The RNA was finally purified and a dilution series was prepared to be used as RNA competitor for RT-PCR.

RT-PCR conditions were optimised using total RNA extracted from cassava leaves and wounded leaves. Expected RT-PCR products were obtained as shown in Fig 4.17A. The competitor RNA was tested in normal (non-competitive) RT-PCR reaction. The RT-PCR products from leaf RNA and the RNA competitor showed the expected sizes and a clear difference between them in size (Fig.4.17B).

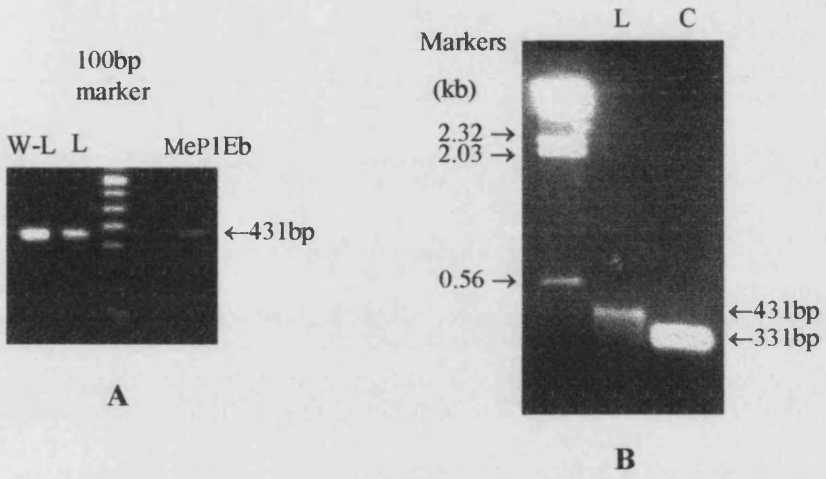


Fig.4.17 Tests of conditions for RT-PCR (A) and competitive RT-PCR (B, C & D).

A. Comparison of the RT-PCR products amplified from cassava leaf total RNA (L) and leaf wounded for 15min (W-L) with the PCR product from MeP1Eb using sense spal and antisense aspal primers. There is no sample in the first lane after the marker. The RT-PCR reaction mixture was set up as described in 2.11 with template RNA at 200ng per reaction. RT-PCR cycles: 1cycle of 48°C for 45min; 1 cycle of 94 °C for 2min; 40 cycles of 94 °C for 30sec, 70 °C for 1min, 72 °C for 1min and 30sec; 1 cycle of 70°C for 7min. The PCR reaction was set up as the following with a final volume of 10 μl: 3.5μl of MilliQ water, 1μl of 10xbuffer, 1 μl of 2mM dNTP mix, 1μl of 1μM each s/as primers, 0.3μl 50mM $MgCl_2$, 2μl of MeP1Eb fragment (1ng) and 3units Taq polymerase (Bioline). The thermal cycles were: 1 cycle of 94 °C for 2min; 30cycles of 94°C for 30sec, 65°C for 1min, 72°C for 1min and 30sec; 1 cycle of 72°C for 7min. The products of the RT-PCR and PCR were run on 1xTBE 1.2%(w/v) agarose gel with ethidium bromide (50μg/100ml).

B. Comparison between the RT-PCR products amplified from cassava leaf total RNA (L) and competitor RNA (C) with spal/aspal primers. The RT-PCR reaction mixtures were set up as described in 2.11 with 26ng leaf total RNA (L) or 10μl of 10^5x dilution of competitor RNA. The thermal cycles of the RT-PCR were the same as described in Fig.4.17 A.

When testing competitive RT-PCR, five reactions were prepared, each contained the same amount (260ng) of total RNA extracted from cassava leaves but in each reaction different amount of competitor RNA (10μl of $5x10^7$, $5x10^8$, $5x10^9$, $5x10^{10}$ dilution of original competitor RNA) was included. During the reaction the competitor RNA would compete for primers with target mRNA, therefore the less competitor there was, the more product there should be from the target mRNA. The analysis of the RT-PCR products showed the expected pattern (Fig.4.17C). While the product from competitor decreased when the amount added was reduced, the product from target mRNA increased. This reverse trend in amplification between two templates meant that the competing system was working. Since RNA is relatively unstable, which makes comparison between different experiments difficult, DNA competitor was tested in a similar way. In the tests, pUC-dP1E/H was used as competitive template. In the different reactions, each of them contained the same amount (300ng) of total RNA extracted from cassava roots 72 hours after harvest but in each reaction different amount of competitor DNA 1.5pg, 0.15pg, 0.015pg and 0.0015pg was included. The results showed similar trends as observed above (Fig.4.17D). These tests demonstrated that the method was reliable and DNA competitor had the same effect as RNA competitor.

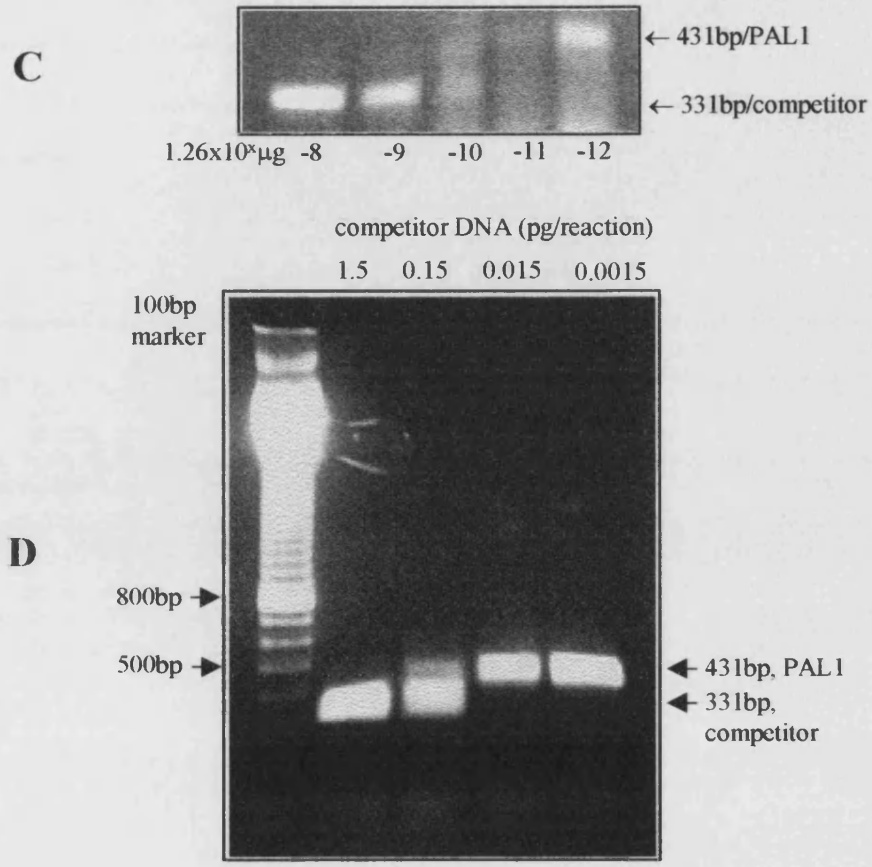


Fig.4.17C. A test of competitive RT-PCR system using competitor RNA. Five reactions were prepared as described in 2.11, each contained the same amount (260ng) of total RNA extracted from cassava leaves but in each reaction different amount of competitor RNA (10µl of $5x10^7$, $5x10^8$, $5x10^9$, $5x10^{10}$ dilution of competitor RNA) was included, 50pmol of sense spal and antisense aspal primers was added. RT-PCR cycles: 1cycle of 48°C for 45min; 1 cycle of 94 °C for 2min; 40 cycles of 94 °C for 30sec, 70 °C for 1min, 72 °C for 1min and 30sec; 1 cycle of 70°C for 7min. The products of the RT-PCR and PCR were run on 1xTBE 1.2%(w/v) agarose gel with ethidium bromide (50µg/100ml).

Fig.4.17D. A test of competitive RT-PCR system using competitor DNA. In the different reactions, each of them contained the same amount (300ng) of total RNA extracted from cassava roots 72hours after harvest but in each reaction different amount of competitor DNA pUC-dP1E/H1.5pg, 0.15pg, 0.015pg and 0.0015pg was included. Other conditions are as described in Fig.4.17C.

4.2.7.2 Expression of MePAL1

Total RNA samples extracted from young leaves, young leaves 3 hours after wounding, young stems, vegetative roots and storage root slices collected during a time-course after harvesting (cultivar, Mcol 22) were analysed using competitive RT-PCR with the *in vitro* transcribed RNA as competitor (Fig.4.18). Competitive RT-PCR reactions were set up with each reaction mixture containing the same amount of competitive template (10µl $5x10^{10}$dilution of the RNA competitor) and the same amount of total RNA (300ng) to be analysed.

The MePAL1 was expressed in young leaves, leaves wounded for three hours, young stems and vegetative roots. There was an increased expression in wounded leaves within three hours after wounding compared to normal leaves. The increased expression could be induced within 15min after wounding (Fig.4.17A). In storage roots just after harvest, there was no or very low expression of MePAL1. After incubation (27±3°C, 55-65% relative humidity) for 8 hours strong expression was evident but MePAL1 was not expressed from 16 to 36 hours after harvest. By 40 hours a weak signal appeared and then signals became stronger by 48 and 56 hours.

The expression of MePAL1 during the PPD was also studied using storage roots (Mcol 22) supplied by CIAT, Colombia. The roots were sealed in wax immediately after harvest and air-mailed to Bath. On arrival the wax and the cortex were peeled off the roots and the quality of the roots was checked for visual evidence of PPD. The roots (about 5cm in diameter & 25cm in length) that were of good quality, were cut into 2.5cm thick slices and incubated at 27±3°C, 55-65% relative humidity for 0, 24, 48 and 72 hours before they were used to extract total RNA. Normal RT-PCR (no competitor) was performed with 300ng of total RNA from the incubated roots and the RT-PCR products from the roots incubated 0, 24, 48 and 72 hours were compared (Fig.4.19A).

The RT-PCR product from MePAL1 mRNA increased from non-detectable in the control (0 hour) to very broad bands by 48 and 72 hours after harvest.

Competitive RT-PCR reactions were set up with each containing the same amount of competitive template pUC-dP1E/H (0.015pg) and the same amount of total RNA

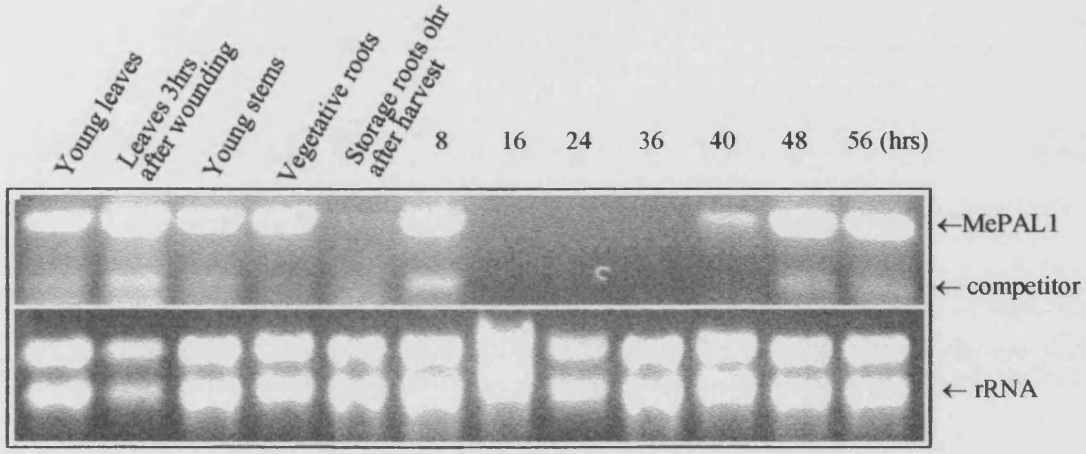


Fig.4.18 Expression analysis of MePAL1 using RT-PCR in cassava. The different organs of cassava (cultivar Mcol 22) analyzed were young leaves, young leaves 3hrs after wounding (by cutting into 0.25mm^2), young stems, vegetative roots, storage root slice just after harvest, root slices incubated for 8, 16, 24, 36, 40, 48 and 56 hours. Competitive RT-PCR reactions were set up as described in 2.11 with each reaction mixture containing the same amount of competitive template ($1.26x10^{-11}\mu g$) and the same amount of total RNA (300ng) to be analysed. The total RNA from different tissues and treatments was run in 1xTBE 1.2% (w/v) agarose gel to confirm the equal amount of templates to be used for RT-PCR among the samples by comparing the rRNA bands. Other conditions were the same as described in Fig.4.17C. The images of rRNA controls were enlarged from original gel picture.

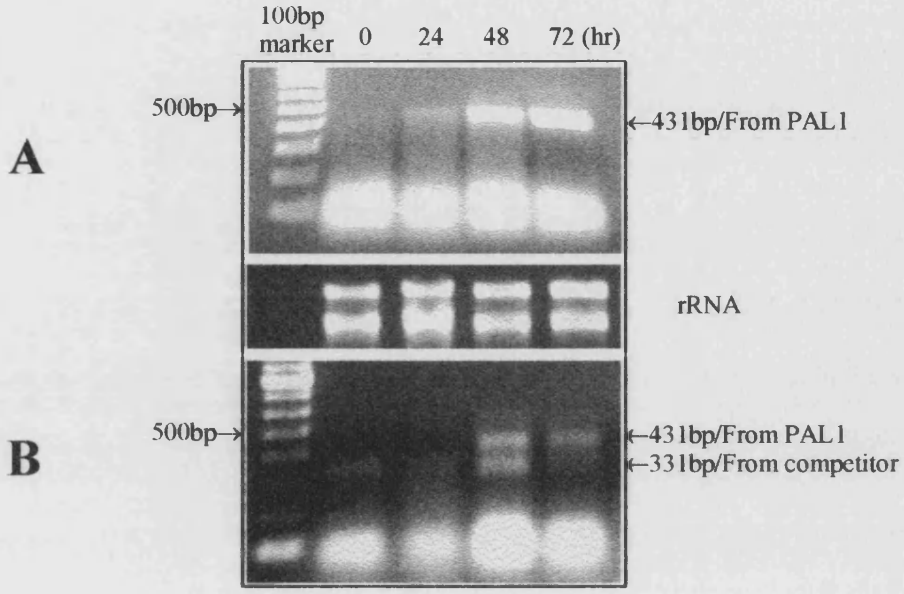


Fig.4.19 Expression analysis of MePAL1 in cassava using RT-PCR. Total RNA was extracted from 2.5cm thick slices which had been incubated at 27±3°C, 55-65% relative humidity for 0, 24, 48 and 72hrs. The cassava roots (cultivar Mcol 22) were wax-sealed after harvest and delivered to Bath from CIAT, Colombia.

A. RT-PCR analysis of MePAL1 expression. RT-PCR reactions were set up with each containing the same amount of total RNA (300ng) to be analysed.

B. Competitive RT-PCR reactions were set up with each containing the same amount of competitive template pUC-dP1E/H (0.015pg) and the same amount of total RNA (300ng) to be analysed. RT-PCR and competitive RT-PCR reactions were set up as described in 2.12 and other conditions were the same as described in Fig 4.17 C. The total RNA from different tissues and treatments was run in 1xTBE 1.2% (w/v) agarose gel to confirm the equal amount of templates to be used for RT-PCR among the samples by comparing the rRNA bands. The images of rRNA controls were enlarged from original gel image to match the bands in RT-PCR gel.

(300ng) to be analysed. The products of the amplification were run on 1xTBE 1.2%(w/v) agarose gel and visualised over UV light (Fig.4.19B). Under these competition conditions, there was no PCR product from MePAL1 mRNA by 0hr and 24 hr, a strong band by 48hr and weaker band by 72hr.

4.2.8 Summary of the Results

Two PAL cDNA clones cMePAL1 and cMePAL3 were isolated and characterised. The cMePAL1 consisted of a short 5' untranslated region, a coding region and 3'untranslated region. cMePAL3 was a partial clone of its corresponding mRNA, containing part of the coding region and the 3' untranslated region. The deduced amino acid sequences of these genes including gMePAL2 (which was isolated and characterized by Hongying Li, unpublished results) were of very high identity, especially between gMePAL2 and cMePAL3. The PAL genes shared high identity in deduced amino acid sequences with PAL genes from other plants, especially with dicot tree plants such as poplar and lemon. Southern blot hybridisation of cassava genomic DNA indicated that there were at least four genes in the cassava PAL gene family. MePAL1 expression was differentially regulated among organs and could be stimulated by wounding. MePAL1 was expressed in leaves, stems and vegetative roots at similar levels but not in unwounded storage roots. Mechanical wounding led to the increased expression of MePAL1 in leaves. MePAL1 was induced in the early stage of PPD, then declined and was induced again to high expression in later stages of PPD.

4.3 Discussion

PAL genes in plant systems are quite conserved as it has been demonstrated by the comparison of their deduced amino acid sequences. The sequences of the enzyme's active site in plants are 100% conserved, though there was a variation among plant PALs in amino acid sequence (14% to 40% identity between cassava PALs and other plant PALs). For example, PAL genes in rice and pine showed relatively low overall identity to cassava PAL genes at the amino acid level (67% and 62% respectively), but they share exactly the same putative active site sequences (Fig.4.20). The sequences at the active site were highly conserved not only among plants but also with fungi such as *Rhodotorula toruloides* and *Rhodotorula rubra*. What is more, the sequences were also shared by functionally similar genes in other kingdoms. Histidine ammonia-lyase (EC 4.3.1.3) or histidase, catalysing the deamination of histidine to produce urocanic acid, is

functionally and structurally related to PAL. Histidine ammonia-lyases in organisms ranging from micro-organisms (*Bacillus subtilis*, *Pseudomonas putida*, *Streptomyces griseus*) to mammals such as rat (liver) (*R. norvegicus*) and mouse (liver) (*M. musculus*) were of high identity and similarity in the sequences of their active sites to PALs from plants and fungi. The first conserved serine residue was identified as the active residue of PAL in parsley or HAL in rat liver by mutagenesis (Schuster and Rétey, 1994). It was proposed that the genes for PALs and HALs were diverged from a common ancestral gene and they were evolved to code for enzymes with similar active sites, but different substrate specificity (Taylor *et al.*, 1990). The substrate specificity may sometimes be not so strict, as it was shown that PAL from monocots utilise tyrosine as well as phenylalanine, whereas the enzyme from dicots can utilise only phenylalanine efficiently (Jangaard, 1974). This was further illustrated when a maize PAL expressed in *E. coli* showed tyrosine ammonia-lyase (TAL) activity at a very similar catalytic efficiency to PAL (Rosler *et al.*, 1997). The physiological implication of these two alternative pathways of coumaric acid production was not clear.

PAL1,2-cassava	GTITA**S**GDLVPLSYIA
PAL1-parsley	GTITA**S**GDLVPL*S*YIA*
PAL-rice	GTITA**S**GDLVPL**S**YIA
PAL-maize	GTITASGDLVPLSYIA
PAL-loblolly pine	GTITASGDLVPLSYIA
PAL-*R.toruliodes*	GTISASGDLSPLSYIA
PAL-*R.rubra*	GTISASGDLSPLSYIA
HAL-*P.putida*	GSVGA**S**GDLAPLATMS*
HAL-*B.subtilis*	GSLGA**S**GDLAPLSHLA
HAL-*Str.griseus*	GSLGCSGDLAPLSHCA
HAL-rat liver	GTVGA**S**GDLAPLSHLA*

Fig.4.20 The sequence alignment of the putative active sites of six PAL and five HAL (histidase). The active amino acid residue serine is in bold. * indicates the PAL or HAL in which the underlined serine residue was identified as the active residue by mutagenesis. The underlined italic serine residue in parsley, which was also conserved among PALs and HALs, was also tested and it was not an active residue. Accession number of PAL genes, parsley P24481, rice X16099, pine U39792, *R.toruliodes* M18261. Reference: maize PAL (Rosler *et al.*, 1997), HALs (Taylor *et al.*, 1990)

Comparison among the sequences of *cis*-acting elements in PAL and 4CL promoters from a number of plants revealed that three *cis*-acting element boxes were highly conserved not only among PAL promoters including gMePAL2 promoter (Li *et al.*, unpublished results) but also between PAL and 4CL promoters (Logemann *et al.*, 1995). While the high conservation among PALs in amino acid sequences and *cis*-acting elements may indicate that PAL is so essential that it has not diverged during evolution,

the conserved *cis*-acting elements between PAL and 4CL reflected the degree of coordination in the regulation of functionally interdependent genes.

The phenomenon of very high identity between two PAL genes (gMePAL2 and cMePAL3) in cassava was also observed in parsley (Logemann *et al.*, 1995) and jack pine (*P. banksiana*) (Butland *et al.*, 1998). In parsley, PAL1 and PAL2 were almost identical even including the 5' and 3' untranslated regions except for a few base difference. PAL2 and PAL5 in jack pine were of 94.0% and 98.4% identity in nucleotide and amino acid sequences respectively. There is no explanation why plants need two genes encoding almost identical proteins.

The nucleotide sequence of PAL1 in jack pine (*P.banksiana*) was exactly the same as the PAL sequence in loblolly pine (*P.taeda*) (Butland *et al.*, 1998). Therefore, it is not surprising that the gMePAL2 genes in cultivar MBRA 534 and MNGA1 are the same sequence in the region of about 500bp analysed.

Based on both the comparison between cassava PALs and PALs from other plants and phylogenetic analysis of the PALs, it could be suggested that PAL genes of cassava and dicot tree plants such as poplar, *Citrus limon* and tea were quite close in terms of evolution. They were less distant to an ancestral PAL gene than PALs in legume plants. *Arabidopsis thaliana* PAL1 and PAL2 were more close to an ancestral PAL gene than PALs in other dicots, as suggested by similar analysis (Wanner *et al.*, 1995).

PAL mRNA in other plant species is abundant in stems or petioles, low in leaves, and typically abundant in roots (Wanner *et al.*, 1995; Joos and Hahlbrock, 1992; Liang *et al.*, 1989). The expression study of MePAL1 in cassava revealed that it was expressed abundantly and at similar levels in leaves, stems and vegetative roots.

The accumulation of PAL mRNA upon wounding has been observed in many plants. Three PAL mRNAs in potato (*Solanum tuberosum* L.) leaves peaked rapidly within 3hr after wounding (Joos and Hahlbrock, 1992). Also in bean (*Phaseolus vulgaris* L.) all three PAL genes were induced to a very high level (5 to 10 fold of the levels in other organs) by wounding in hypocotyls (Liang *et al.*, 1989). When 20-day old melon (*Cucumis melo*) was wounded, PAL mRNA could be detected using Northern analysis as

early as 90min and it reached maximum levels 24hr after the stimulus (Diallinas and Kanellis, 1994). Similarly, in rice PALZB8 transcripts increased within 1hr after wounding and maximum induction appeared between 12 and 24 hrs after wounding (Zhu *et al.*, 1995). Preliminary analysis of the expression of MePAL1 in cassava leaves showed a possible enhancement of its expression by mechanical wounding as early as 15min after wounding. Wounding also induced MePAL1 expression in storage roots.

It is interesting to note that MePAL1 mRNA was not expressed or was expressed at a very low level in storage roots, while it was abundant in vegetative roots. In the roots of bean plants, it was suggested that the abundance of PAL gene transcripts may possibly be due to the mechanical damage to cells surrounding the endodermis at the sites of lateral root initiation during the root development (Liang *et al.*, 1989). Low or no PAL enzyme activity has been detected previously in cassava storage roots immediately after harvest (Tanaka, 1983; Rickard, 1985; Pereira, 1999). Wounding induced the expression of MePAL1 in the storage root as revealed from the expression data either from the roots harvested from the glass house in Bath or the roots delivered from Colombia. In the detailed time-course study, the expression of MePAL1 had two peaks. However, only one peak of MePAL1 expression was observed in the roots delivered from Colombia during a 24 hour scaled time course, which corresponded to the second peak in the detailed time course study. The two-peak pattern was also observed in the expression of another PAL gene during PPD (Pereira *et al.*, 1999). The first peak may be the initial response to the wounding and the second may be caused by the deterioration or may be part of the deteriorating process. Similar two-pattern appeared also in the expression of PAL genes in wounded leaves and wounded roots in parsley (*Petroselium crispum*). While the expression of PcPAL1, PcPAL2 and PcPAL3 peaked at 8 and 24 hours after wounding in leaves, the peaks appeared earlier and closer in wounded roots, at 2 and 8 hours after wounding (Logemann *et al.*, 1995). The detailed time-course analysis of MePAL1 expression may be applied to the roots delivered from Colombia to test whether there is a similar expression pattern of MePAL1. The expression patterns in the roots with different backgrounds could be compared to understand the effect of storage on MePAL1 expression. Though varying between different experiments, PAL activity generally increased to a peak in the storage roots by 40 to 72 hours after harvest (Tanaka, 1983; Uritani, 1984; Rickard, 1985; Pereira *et al.*, 1999), to which the main peak of MePAL1 expression in this study may contribute.

PPD is a complicated process and at least three PAL genes, MePAL1, MePAL3 and MePAL, were expressed during the process. It is probable that they contribute differently to the general PAL activity. Therefore, the expression pattern of one single PAL gene might not reflect the overall pattern of PAL activity changes during PPD. This was proven to be the case when the pattern of MePAL gene expression was compared with that of PAL activity (Pereira, 1999). While the PAL activity increased steadily during PPD, two peaks and one trough of MePAL and MePAL1 gene expression were observed. Northern analysis of all the isolated PAL clones in cassava and comparison between their expression patterns and vascular streaking may reveal the PAL genes closely related to PPD.

The role of a particular PAL gene in the process is difficult to elucidate by analysing gene expression due to the complexity of the gene family and of phenylpropanoid metabolism. However, the possibility to explore the role of individual PAL genes is now open with the availability of cassava transformation systems. Promoters can be fused to reporter genes such as β-glucuronidase (GUS) and transferred to cassava to investigate the temporal and spatial patterns of the promoter activities at the cell and tissue level. The PAL promoter ZB8/GUS fusion in transgenic rice revealed that strong GUS activities were shown in both inner and outer vascular bundles of stems. GUS activities were localized in epidermal and vascular tissues and root tips (Zhu *et al.*, 1995). In transgenic potato with a PAL (bean)-GUS fusion gene, GUS activity was localized in cells directly adjacent to the wounded edge in potato tubers 6hrs after slicing and in the wound periderm 72hrs after slicing. The location of GUS expression corresponded to the sites of suberin deposition (Shufflebottom *et al.*, 1993). Similar information about the activity and localization of cassava PAL promoters in storage roots during PPD would provide insight into the roles of PAL genes in the process. In particular the PAL genes involved in the deterioration process or the formation of wound periderm in high humidity may be identified in this way. The promoter of gMePAL2 gene has been fused with a reporter gene GUS and is being transferred into cassava to analyse the activity of the promoter (Li *et al.*, 1998). The availability of cassava transformation systems also makes possible the application of the well-established antisense gene-knock-out techniques to introduce antisense PAL genes into cassava to investigate their effects on the development of PPD. The approach would be very powerful in combination with the fingerprint analysis of

secondary metabolites, especially phenolic compounds; in this way an individual PAL gene may be linked with a particular pathway branch from the general phenylpropanoid metabolism pathway.

PAL genes in plants were found to be strongly induced during the hypersensitive reaction between tobacco and tobacco mosaic virus, potato and *Phytophthora infestans*, parsley and *Phytophthora megasperma* f.sp. *glycinea* (Pellegrini *et al.*, 1994; Guypers *et al.*, 1988; Schmelzer *et al.*, 1989). MePAL1 was expressed at the border between the cells undergoing hypersensitive reaction and the healthy tissues during the interaction between cassava and *Pseudomonas syringae* (Day *et al.*, unpublished results). PAL genes may play a role in the cell wall thickening or strengthening to form a barrier to limit the pathogen's spread as part of the hypersensitive reaction. MePAL was expressed during the interaction between *Xanthomonas axonopodis pv Manihotis* (*Xam*) (Pereira *et al.*, 1999).

There is a small PAL gene family with 2 to 6 members in most of the plants studied such as bean, parsley (4) (Logemann *et al.*, 1995), *Arabidopsis* (3) (Wanner *et al.*, 1995), poplar, alfalfa, tobacco (4) (Pellegrini *et al.*, 1994), and tomato (≥5) (Lee *et al.*, 1992). Southern blots of cassava genomic DNA with probes made from MePAL1 and 2 suggested that there were at least four PAL genes in the gene family of cassava and four PAL genes have been identified in cassava. However, it was estimated that there were about 40-50 PAL genes per haploid genome in potato (Joos and Hahlbrock, 1992), about 8 to 10 PAL genes in jack pine (Butland *et al.*, 1998), while there was only one in loblolly pine (Whetten and Sederoff, 1992). New PAL genes have been usually isolated using heterologous probes made from characterised PAL genes, which was based on the similarity of the target genes to the probes. This may partly explain the relatively high similarity (over 70% in nucleotide sequence) among the members of a PAL gene family in plants. Estimation of the size of a gene family by Southern analysis is also based on the similarity among the genes. Screening with a known gene and estimation with Southern blot could fail to detect target genes of low similarity to the probes. This problem may be partly overcome using more than one set of (degenerate) primers to clone from genomic DNA. This approach was applied successfully in jack pine to isolate PAL genes with relatively high diversity, from which it was suggested that there might be a similar multigene family of PAL in loblolly pine instead of just one PAL gene as was reported

(Butland *et al.*, 1998). Therefore, it is difficult to know or predict the exact size of the PAL gene family.

Chapter Five

Hydroxyproline-rich Glycoprotein (HRGP) Genes were expressed during PPD

5.1 Introduction

Hydroxyproline-rich glycoproteins (HRGPs), are the most abundant among the plant cell wall proteins. HRGPs strengthen cell walls and control cell wall extension by insolubilization of the molecules through intermolecular cross-links. HRGPs are induced in wound responses and involved in the wound healing process (Hirsinger *et al.*, 1997; Bradley *et al.*, 1992; Ussuf *et al.*, 1996). No wound healing processes have been observed at the histochemical level during PPD except for the formation of periderm under 'curing' conditions (high RH and high temperature). It remains to establish whether there are any wound healing responses at a molecular level during PPD. It would be important to test whether HRGP genes are expressed in cassava tubers both *in vivo* and post-harvest in an attempt to understand the apparent absence of wound healing responses. HRGPs may also play a role in plant defense against pathogens by forming a modified physiological barrier in host plants with localised deposition of HRGPs in response to infection (Hippesanwald *et al.*, 1994) or/and by agglutinating pathogens directly (Leach *et al.*, 1982; Mellon and Helgeson, 1982). The roles of HRGPs in cassava defense against pathogens including those involved in microbial deterioration during the storage of tubers, could be explored with the isolated HRGP clones from cassava.

5.2 Results

5.2.1 Isolation of HRGP cDNA clones

The PPD related cDNA library was screened with parsley hydroxyproline-rich glycoprotein (HRGP) cDNA ELI9 as a probe (Kawalleck *et al.*, 1995). The pre-hybridization was as described in 2.9 and hybridization was at 50°C overnight, the membrane was washed to 0.1xSSC / 0.1%SDS at 50°C for 2x10min. Then secondary screening was performed to obtain single positive plaques. Nine positive clones were isolated.

These positive clones were subjected to PCR using lambda gt10 primers to check the size of the inserts. The size of the clones were from about 100bp to 1.9kb (Fig.5.1a). The PCR

products of these clones were Southern-blotted and hybridized against ELI9 cDNA probe, which showed positive signals (Fig.5.1b) at high stringency though some clones showed a strong signal such as C and H and some showed very weak signal such as B, D, G and I.

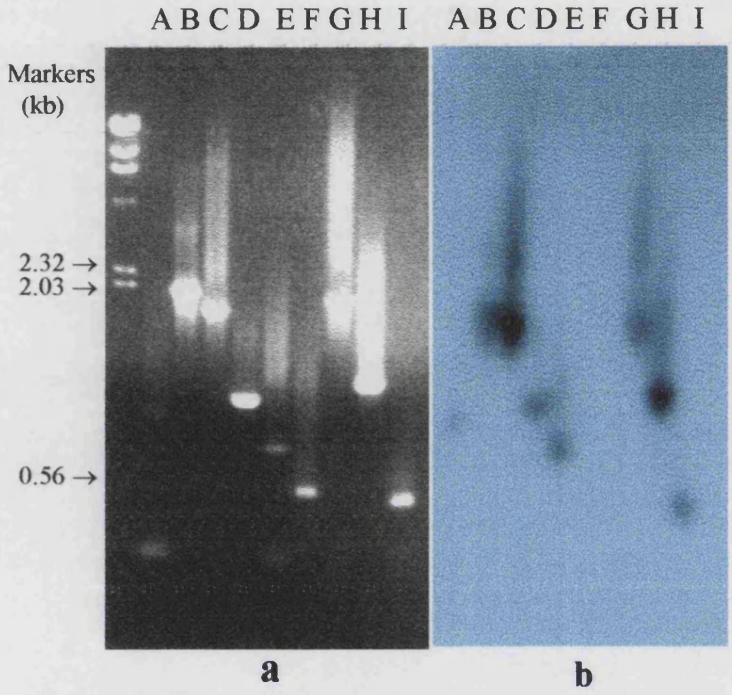


Fig. 5.1 Analysis of the sizes of the inserts in positive HRGP clones.

a. PCR amplification of inserts from positive plaques was performed as described in 2.2.7. The resulting PCR products were run in 1.0% (1xTBE) agarose gel containing ethidium bromide and photographed. Relevant *HindIII* cut lambda DNA marker was indicated. The size of the PCR products of A to I is 600bp, 1.9kb, 1.8kb, 600bp, 300bp,100bp, 1.8kb, 700bp and 100bp respectively.

b. Southern blot and hybridisation of the PCR products. The PCR products in the gel were blotted to Hybond N^+ nylon membrane and hybridised with the probe made from ELI9 of parsley at 50°C overnight. The membranes were washed subsequently with 2xSSC/0.1%SDS at room temperature for 2x10min., 2xSSC/0.1%SDS at 50°C for 2x10min., 1xSSC/0.1%SDS at 50°C for 2x10min., 0.1xSSC/0.1%SDS at 50°C for 2x10min. Films were exposed to the membranes at -80°C overnight and developed.

5.2.2 Characterisation of cMeHRGP1

Initial attempts were made to subclone clone B, C and G, but only clone G was successfully subcloned. The insert of clone G was released from the lambda vector with *Eco*RI digestion and subcloned into pUC18. The cDNA was sequenced using universal primers and

designed primers (Fig.5.2). The cDNA was 1649bp in size, with a 416bp 5'untranslated region, 669bp coding region, 494bp 3'untranslated region and a polyA tail (Fig.5.3). A putative polyadenylation site was located 28bp uptream of the polyA tail. The deduced amino acid sequences of the coding region contained repeat motifs such as SP_4, typical repeat units of HRGPs. The sequences showed 79% identity to the extensin genes in soybean (Genebank accession number L22031) and to bean (M18095) over 500bp. The cDNA was then designated as cMeHRGP1. The amino acid composition of the predicted polypeptide is 48.71% proline, 17.67% tyrosine, 13.36% serine, 6.47% lysine, 5.17% valine, 4.74% histidine and 3.88% other amino acids.

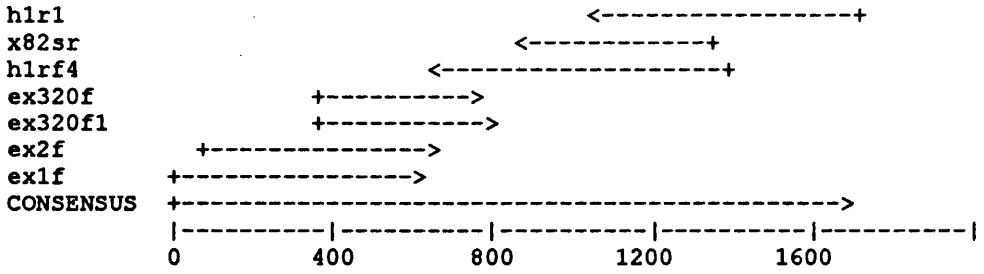


Fig.5.2 The sequence strategy of cMeHRGP1. Sequences h1r1 and ex1f were determined with universal primers. The other sequences were determined with primers designed from the known sequences.

```
   1  TCTCGTTAGCGCAAATATGCAAATATTTAGCATATGGGCTTTGCTCTAAACCATCAAAAG   60
  61  ATGGAATTTTTGTATAATTTTCAGATTTTCCGGACCAAATCACCTTGGGCTTGCAAAACT  120
 121  TTTGTGCCTCCTTGAAAGCCTTATCCCCCCACGACCCAGTGACCACGTAGTCAACCTGGT  180
 181  CATCGGGCTTGCACAGATTCAACGGGATCATGGCGAACTGCGTGGTTGCGCCGCCCTGGA  240
 241  GAAATAGTACGGCGAAATCTTCTGGGATGTTCAACAGAGCACGGAGATCAGCCTCGGCCT  300
 301  TCTGAATAATAGAGAGAAATTCTTTCCCTCTATGGCTCATCTCCATAACACTCATGCCAG  360
 361  ATCCGCGCCAGTTGTAGAGCTCCGCCTGGGCTTTCTTGAGGACGTTGGCAGGTAAGATGG  420
                                                        M  A
 421  CAGGACCGGCGGCGAAATTGAAGACATCACCATCACCTCCTCCTCCATACTACTATAAGT  480
        G  P  A  A  K  L  K  T  S  P  S  P  P  P  P  Y  Y  Y  K  S
 481  CCCCACCTCCACCATCTCCATCACCTCCACCTCCCTACTACTACAAATCTCCTCCACCAC  540
        P  P  P  P  S  P  S  P  P  P  P  Y  Y  Y  K  S  P  P  P  P
 541  CGTCTCCATCTCCTCCTCCCCCATACTACTACCACTCACCACCACCACCAGTGAAATCTC  600
        S  P  S  P  P  P  P  Y  Y  Y  H  S  P  P  P  P  V  K  S  P
```

```
601  CCCCTCCCCCATACTACTACCACTCGCCACCACCTCCTGTGAAATCACCTCCTCCTCCAT   660
       P  P  P  Y  Y  Y  H  S  P  P  P  P  V  K  S  P  P  P  P  Y

661  ACTACTACCACTCACCACCACCACCTGTGAAATCACCTCCTCCTCCATACTATTACCACT   720
       Y  Y  H  S  P  P  P  P  V  K  S  P  P  P  P  Y  Y  Y  H  S

721  CACCACCTCCTCCAGTAAAATCACCTCCTCCTCCATACTACTACCACTCACCACCACCAC   780
       P  P  P  P  V  K  S  P  P  P  P  Y  Y  Y  H  S  P  P  P  P

781  CCGTGAAATCACCTCCTCCCCCATACTACTACCACTCGCCTCCTCCTCCAGTGAAATCAC   840
       V  K  S  P  P  P  P  Y  Y  Y  H  S  P  P  P  P  V  K  S  P

841  CACCTCCTCCATACTATTACCACTCACCACCTCCTCCAGTGAAATCACCTCCTCCTCCAT   900
       P  P  P  Y  Y  Y  H  S  P  P  P  P  V  K  S  P  P  P  P  Y

901  ACTACTACCACTCACCACCACCACCCGTGAAATCACCTCCTCCCCCATACTACTACCACT   960
       Y  Y  H  S  P  P  P  P  V  K  S  P  P  P  P  Y  Y  Y  H  S

961  CGCCTCCTCCTCCAGTGAAATCACCACCTCCTCCATACTACTATCATTCACCACCACCAC   1020
       P  P  P  P  V  K  S  P  P  P  P  Y  Y  Y  H  S  P  P  P  P

1021 CAGTGAAATCACCTCCTCCTCCATACTACTACCACTCACCACCACCTCCTGTAAAATCAC   1080
       V  K  S  P  P  P  P  Y  Y  Y  H  S  P  P  P  P  V  K  S  P

1081 CTCCTCCCCCAGTTTACATTTACGCCTCACCATAGGCTCAGAAAGCTCAGTCACACACCA   1140
       P  P  P  V  Y  I  Y  A  S  P  *
1141 AAGTCGATCATATTTTAGTTTCAACAATGTAATAAAGGAAGGCTCCAAGAGGAAAAGATG   1200

1201 TGGACATCAAGCTTCTAATCCAAGTCCATTGAATAAGGAACTGAATTTTGCATCAATGAG   1260

1261 CTACAATTTGAATAATCCAAGCCAGGAATCTCCAATTTCAAAGCTCACCATTTTGCATCT   1320

1321 CATGCTCATGTTATGTTCCAGTAAAATTGGCTTTAAACATTTACCTCAACAAAGAAAAAC   1380

1381 TAGCAAGTGAAGTTAAAGATGAGGACTTGGATTCGAAGTGGGTGTTTATGTTTGGTTTTT   1440

1441 ATTTATTCATCCCCAATTTATTATTATATGTATAGCGAACATCTTTTATCGTTTTGATTG   1500

1501 GGCTTGTCTCTATTTATTGTGAGCGATTCTTTCTTGGACCTTCTTGTAAGTTATGCATTG   1560

1561 AGATTGTTTGCAATAAAATAAGAAATTTTCTATTACAAATTTACTTGTC(1609)(A)40
```

Fig.5.3 The nucleotide sequence and deduced amino acid sequence of cDNA cMeHRGP1. The translation start codon and stop codon are in bold. Putative polyadenylation signal is 28bp upstream to the poly(A) tail and underlined. The position where the antisense strand starts the codons for phosphoserine aminotransferase is marked by bold and underlined sequence.

5.2.3 Repetitive Motifs in Predicted Amino Acid Sequence of cMeHRGP1

There are two SP_4-Y_3K-SP_4-SP motifs, eleven SP_4-Y_3H-SP_4-VK repeat units and one SP_4 near the C- terminal in the deduced amino acid sequences (Fig.5.4). The deduced peptide contained 27 $S(P)_4$ repeats in total.

```
MAGPAAKLKT
                    SP
SPPPP YYYK SPPPP SP
SPPPP YYYK SPPPP SP
SPPPP YYYH SPPPP VK
SPPPP YYYH SPPPP VK
SPPPP YYYH SPPPP VK
SPPPP YYYH SPPPP VK
SPPPP YYYH SPPPP VK
SPPPP YYYH SPPPP VK
SPPPP YYYH SPPPP VK
SPPPP YYYH SPPPP VK
SPPPP YYYH SPPPP VK
SPPPP YYYH SPPPP VK
SPPPP YYYH SPPPP VK
SPPPP YYYH SPPPP VK
      SPPPP V
    YIY  ASP*
```

Fig. 5.4 Repeat motifs in the deduced amino acid sequences of cHRGP1 cDNA. Amino acid residues are arranged in order to display the various amino acid repeat units and their periodicities. S-serine, P-proline, Y-tyrosine, K-lysine, H-histidine, V-valine, I-isoleucine, T-threonine, A-alanine, L-leucine and G-glycine. * indicates the stop codon.

5.2.4 Is cMeHRGP1 a Chimeric cDNA Clone or Is It an Antisense Gene Pairing?

Searching the DNA database at NCBI using the Blastx program with the sequence of cMeHRGP1 revealed that the deduced amino acid of the reverse sequence of the first 445bp were of high similarity to phosphoserine aminotransferase in *Arabidopsis thaliana*, *Spinacia oleracea*, and some similarity to yeast and *E. coli*.. Phosphoserine aminotransferase catalyzes the conversion of phosphohydroxypyruvate to phosphoserine in the phosphorylation pathway of serine biosynthesis (see 7.4).

There were two aspects about this clone that needed to be resolved. Firstly, was this clone a chimeric clone consisting of two cDNA fragments in opposite orientations, cMePSAT for phosphoserine aminotransferase cDNA fragment and cMeHRGP1 for the rest of the cDNA clone (Fig. 5.5), or was it a genuine cDNA clone encoding two proteins with sense and antisense DNA strands or so called antisense gene pairing? Secondly, was the unusually long 3' untranslated region chimeric too?

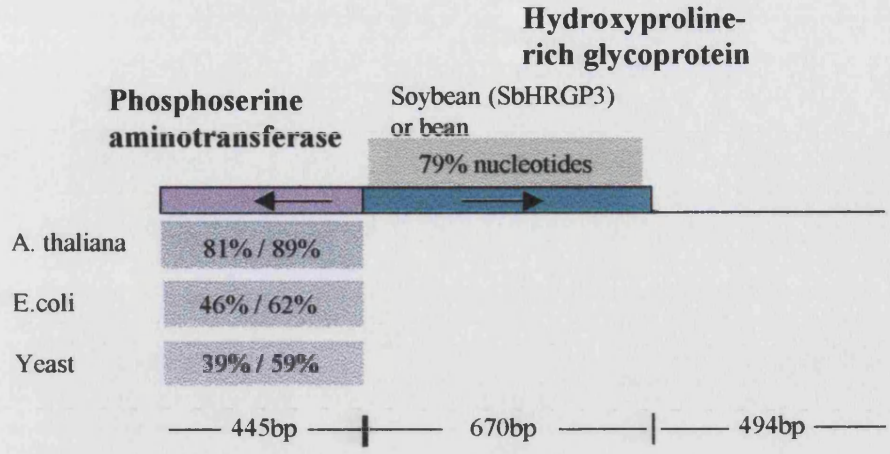


Fig.5.5 Each strand of cMeHRGP1 clone encoded a putative protein. Clone cMeHRGP1 consists of two fragments of cDNA shown with high similarity to phosphoserine aminotransferase and hydroxyproline-rich glycoproteins respectively in other organisms. The two sequences were in the opposite orientation as indicated with solid arrows. The identities and similarities of the deduced amino acid (aa) or nucleotide (for hydroxyproline-rich glycoprotein) sequences of the cDNAs to phosphoserine aminotransferase in other organisms and HRGP are shown in shadows. Sequence accession number, *A thaliana* D88541, *E. coli* P23721, yeast P33330, soybean U44838 and bean U18791.

In order to test whether cMeHRGP1 is chimeric or one genuine cDNA from one mRNA, attempts were made to investigate the possibility of amplifying the genomic DNA fragments corresponding to the cDNA by PCR with primers covering the junction of the two opposite coding regions on two strands. It was speculated that the clone would be chimeric if no PCR product could be amplified, or genuine if a product was produced.

Using primers h1rf (ACG GAG ATC AGC CTC GGC CTT CTG, from nucleotide 281 onwards) and xr1 (GTC CAA GAA AGA ATC GCT CAC, from nucleotide1540 backwards), PCR reactions under normal conditions were performed but no PCR products were obtained. Then hot start PCR reactions with different components added after the hot-start (3 to 4 min at 94°C) were performed and a fragment slightly bigger than the corresponding PCR product from cMeHRGP1 was amplified from genomic DNA (Fig.5.6 on the left). The fragment also appeared in a repeated experiment of hot start with *Taq* polymerase (Fig.5.6 on the right). A schematic diagram was drawn to illustrate the positions of the primers and the PCR products (Fig. 5.6)

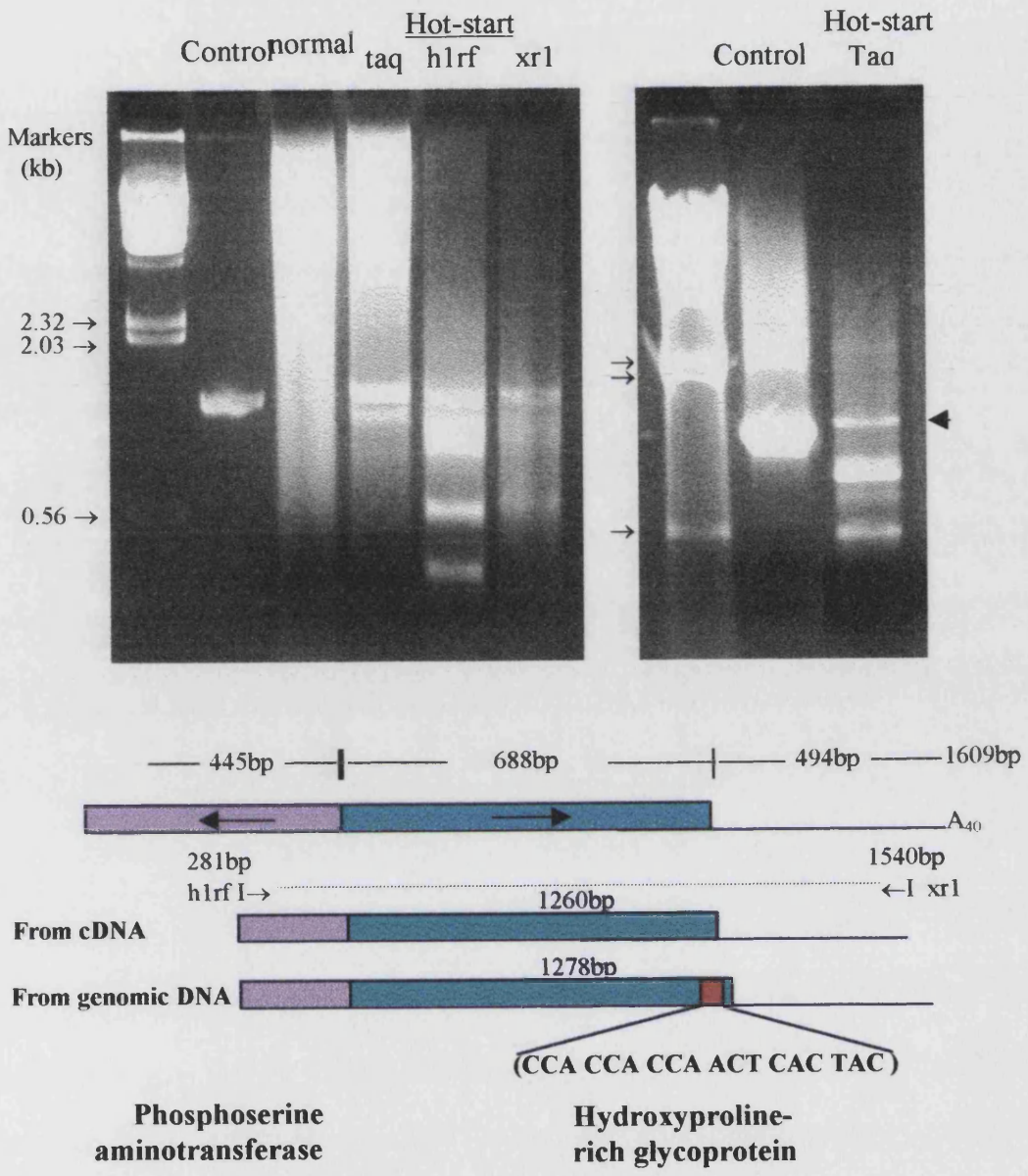


Fig.5.6 PCR amplification of the corresponding genomic DNA fragment to cMeHRGP1and confirmation of the fidelity of the cDNA clone. Normal PCR reaction was prepared by mixing water, 5 μ l 10x buffer, 5 μl of 2mM dNTP, 5 μl each of the 1mM h1rf and xr1 primers, 1.5 μl of 50mM $MgCl_2$, 12U of Bioline Taq polymerase, and 100ng cassava genomic DNA (from cultivar MNGA1) to a final volume 50 μl. PCR cycles were: 94 °C 5min for 1 cycle, 94 °C 1min / 62 °C 1min / 72 °C 2min for 35-38cycles, 72 °C 5min for 1 cycle. For hot-start PCR, a certain component (Taq polymerase or primers) was added only after the first denature cycle was completed to increase specificity. Normal PCR conditions were applied for amplification of the cDNA fragment from cMeHRGP1with the same set of primers as a control, the second lane in each gel. The primers were designed from cMeHRGP1 covered the junction of the coding regions of phosphoserine aminotransferase and hydroxyproline-rich glycoprotein. DNA markers used were *Hin*dIII cut λ DNA.
The schematic diagram showed the opposite positioned coding regions of two genes and the primers used to check the existence of the chimerical DNA in cassava genome. The sequence of the PCR product from genomic DNA had an extra 16 nucleotides at the end of coding region compared to cMeHRGP1.

The PCR fragment (xgcm3) larger than the control was then purified and sequenced partially from both ends with the same primers as was used for PCR amplification. The sequence derived from primer h1rf, xgcmf, was not of good quality but it showed very high identity (80%) to the sequence of cMeHRGP1(Fig.5.7). The difference between these two sequences was mainly from the ambiguous part of the sequence xgcmf. There were some nucleotides in xgcmf sequence which were clearly different from cMeHRGP1. This may indicate that the PCR product was amplified from a different HRGP gene rather than the HRGP gene corresponding to cMeHRGP1, though they were of high similarity.

```
cMeHRGP1  350 ACTCATGCCAGATCCGCGCC.AGTTGTAGAGCTCCGCCTGGGCTTTCTTG 398
              ||||||||||:|  ||:||: ||||||| |||||:|||||||||||||||
xgcmf      53 ACTCATGCCANAAACGNGCNAAGTTGTAAAGCTCNGCCTGGGCTTTCTTG 102

          399 AGGACGTTGGCAGGTAAGATGGCAGGACCGGCGGCGAAATTGAAGACATC 448
              ||||:|||::||:||||||||||||||||:| :||:||||||||||:||||
          103 AGGANGTTNNCANGTAAGATGGCAGGACCNGTNGCNAAATTGAAGNCATC 152

          449 ACCATCACCTCCTCCTCCATACTACTATAAGTCCCCACCTCCACCATCTC 498
              |: ||||||||:|||||  |||||||||||||||:|||||||:|::|||||:
          153 ANGATCACCTCNTCCTCTATACTACTATAAGTCNCCACCTCNANNATCTN 202

          499 CATCACCTCCACCTCCCTACTACTACAAATCTCCTCCACCACCGTCTCCA 548
              |||:||:|||||  ||:|||||||:|||||||:||:|||||::| |||:||||
          203 CATNACNTCCACTTCNCTACTANTACAAATNTCNTCCANNAGCGTNTCCA 252

          549 TCTCCTCCTCCCCCATACTACTACCACTCACCACCACCACCAGTGAAATC 598
              ||||||||||:|||||:||||||:||||:||||||||||||:|||||||||
          253 TCTCCTCCTCNCCCATNCTACTANCACTNACCACCACCACNAGTGAAATC 302

          599 TCCCCCTCCCCCATACTACTACCACTCGCCACCACCTCCTGTGAAATCAC 648
              |||:|||||||||||||||||||||||||:|:||||||||||||||||||:
          303 TCCNCCTCCCCCCATACTACTACCACTCNCNACCACCTCCTGTGAAATCAN 352

          649 CTCCTCCTCCATACTACTACCACTCACCACCACCACCTGTGAAATCACCT 698
              |||:|||||||||||||||:| |||||||:|: |||| ||||:|||||||||
          353 CTCNTCCTCCATACTACNATCACTCACNANGACCATCTGTNAAATCACCT 402

          699 CCTCCTCCATACTATTACCACTCACCACCTCCTCCAGTAAAATCACCTCC 748
              |||||||:|||||||||||:| |:| :| ||||||||||:|:|||||||||
          403 CCTCCTCNATACTATTACNAGTNATNATCTCCTCCAGTNANATCACCTCC 452

          749 TCCTCCATACTACTACCACTCACCACCACCACCCGTGAAATCACCTCCTC 798
              ||:|||||||||:||||||||:|| |||||:||||||||||: ||:|:
          453 TCNTCCATACTANTACCACTCANCATCACCANCCGTGAAATCANTTCNTN 502

          799 CCCCATACTACTAC 812
              :|||||| ||:|||
          503 NCCCATAGTANTAC 516
```

Fig.5.7 Comparison of the cMeHRGP1 sequence near the junction of two opposite positioned-coding regions with the sequence from the PCR product of genomic DNA with primers h1rf and xr1 designed from cMeHRGP1. There were some nucleotides in xgcmf.seq which were clearly different from cMeHRGP1, underlined. The junction of two coding regions is in bold. N refers to ambiguous sequence.

The 3' partial sequence of the PCR product was identical to the 3' part of cMeHRGP1 except that there was 18 nucleotides more than cMeHRGP1 just before the translation stop codon (Fig.5.8). The additional 18 nucleotide encoded 6 amino acid residues, which added another SP_4 to the polypeptide (Fig.5.8).

```
cMeHRGP1  1042 CATACTACTACCACTCACCACCACCTCCTGTAAAATCACCTCCTCCCCCA 1091
               ||||||||||||||||||||||||||||||||||||||||||||||||||
PCR-xgcm3  353 CATACTACTACCACTCACCACCACCTCCTGTAAAATCACCTCCTCCCCCA 402

           1092 GTTTACATTTACGCCTCACCA..................TAGGCTCAGAA 1123
                ||||||||||||||||||||                  ||||||||||| 
            403 GTTTACATTTACGCCTCACCACCACCACCAACTCACTACTAGGCTCAGAA 452

           1124 AGCTCAGTCACACACCAAAGTCGATCATATTTTAGTTTCAACAATGTAAT 1173
                ||||||||||||||||||||||||||||||||||||||||||||||||||
            453 AGCTCAGTCACACACCAAAGTCGATCATATTTTAGTTTCAACAATGTAAT 502

           1174 AAAGGAAGGCTCCAAGAGGAAAAGATGTGGACATCAAGCTTCTAATCCAA 1223
                ||||||||||||||||||||||||||||||||||||||||||||||||||
            503 AAAGGAAGGCTCCAAGAGGAAAAGATGTGGACATCAAGCTTCTAATCCAA 552
```

```
xgcm3 332  CAGTGAAATCACCTCCTCCTCCATACTACTACCACTCACCACCACCTCCTGTAAAATCAC   391
             V  K  S  P  P  P  P  Y  Y  Y  H  S  P  P  P  P  V  K  S  P

      392  CTCCTCCCCCAGTTTACATTTACGCCTCACCACCACCACCAACTCACTACTAGGCTCAGA   451
             P  P  P  V  Y  I  Y  A  S  P  P  P  P  T  H  Y  *

      452  AAGCTCAGTCACACACCAAAGTCGATCATATTTTAGTTTCAACAATGTAATAAAGGAAGG   511

      512  CTCCAAGAGGAAAAGATGTGGACATCAAGCTTCTAATCCAAGTCCATTGAATAAGGAACT   571
```

Fig.5.8 Comparison between 3' part of cMeHRGP1 and the genomic PCR product xgcm3. Above the dashed line is the difference between cMeHRGP1 and the genomic PCR product. Below is the partial 3' sequence xgcm3 of the PCR product from genomic DNA with primers h1rf and xr1 covering the region from 281 to 1540 of the cDNA cMeHRGP1. The underlined nucleotides and amino acid residues are absent in the sequence of cMeHRGP1, all of the rest sequence is the same as cMeHRGP1.

The difference between the sequence of cMeHRGP1 and the partial sequences from both ends of the genomic PCR product indicated again that the PCR product was not the genomic counterpart of cDNA cMeHRGP1, but it shared high identity to cMeHRGP1. The success of PCR amplification of genomic DNA using primers from cMeHRGP1 and the high identity between genomic DNA and cDNA suggested that the cDNA cMeHRGP1 did have its genomic counterpart, though the PCR product may be from another HRGP gene of very high identity to the gene of cMeHRGP1. Then, a question remains as to why no PCR product corresponding to cMeHRGP1 was amplified from genomic DNA.

5.2.5 Investigation of the Unusually Long 3' Untranslated Region

When it was realized that cMeHRGP1 might be a chimeric clone, the unusually long untranslated region of the cDNA clone was also in doubt. The fidelity of this clone with respect to its long 3'UTR was tested by investigating the existence of its genomic counterpart using PCR.

Primers xf1 (TTA CGC CTC ACC ATA GGC TC, from 1100bp of cMeHRGP1) and xr1 were designed from the end of the coding region and the region near to the poly(A) tail (Fig.5.9b). A fragment was amplified from genomic DNA and it was bigger than the PCR product from the control (Fig.5.9a), which indicated that there might be an intron in the region covered by the primers.

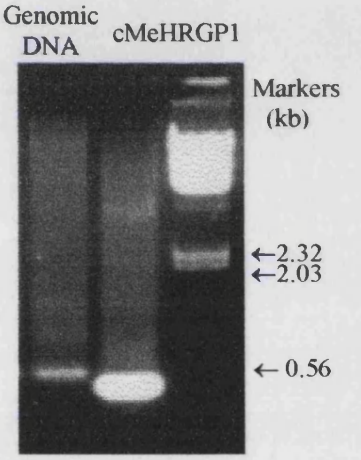


Fig.5.9a. PCR amplification of the genomic DNA fragment corresponding to the 3' untranslated region of cMeHRGP1. Primers xf1 and xr1 were indicated in the schematic diagram below. PCR condition and cycles were as described in 2.3.7 with template DNA 15ng for cMeHRGP1plasmid and 300ng for cassava genomic DNA, anealing temperature at 58°C.

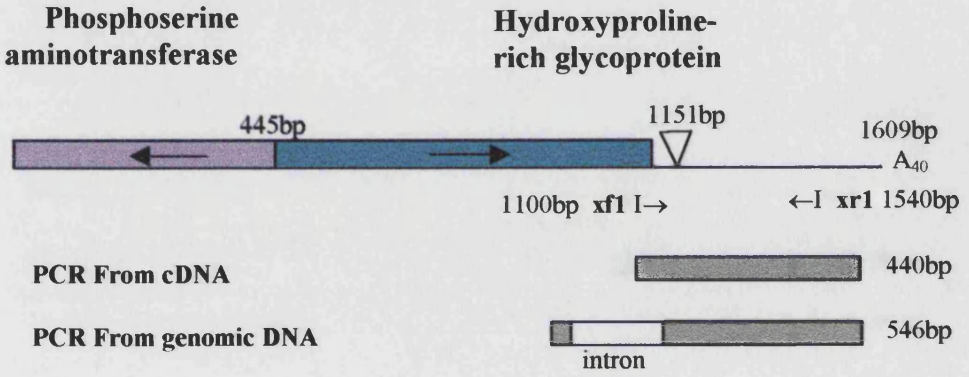


Fig.5.9b PCR amplification of the genomic DNA corresponding to the region of 3' end of cMeHRGP1 with primers xf1(from 1100) and xr1(from 1540 backwards). The intron was located in position 1151, indicated with an inverted open triangle. The intron was 106bp and with a spicing site GT AG conserved among plants.

Comparison between the sequences of the genomic PCR product and the cDNA revealed that these two sequences were identical except that there was an intron of 106bp just 36bp downstream of the translation stop codon in the genomic sequence (Fig.5.10). The splicing sites are GT AG, which are conserved in plant systems (Brown, 1986). Introns located in 3' untranslated regions also occurred in HRGP genes of other plants (Ahn *et al.*, 1996; Chen and Varner, 1985; Wycoff *et al.*, 1995)

The PCR amplification of the genomic counterpart of the cDNA demonstrated that the long 3'UTR was the true nature of this cDNA and the PCR product was from MeHRGP1 gene and not from counterpart of xgcm3 (amplified with h1rf and xr1).

```
  1  GCCTCACCATAGGCTCAGAAAGCTCAGTCACACACCAAAGTCGATCAT
     A  S  P  *
                                                  gtaagtttcttt   60
 61  aacacttctaacatcatttttcaagtttttttagtattagcaaaatgaatcattttgaagt   120
121  gtatgtctaattttctgccatttccatacaacag
                                                  ATTTTAGTTTCAACAATGTAATAAAG   180
181  GAAGGCTCCAAGAGGAAAAGATGTGGACATCAAGCTTCTAATCCAAGTCCATTGAATAAG   240
241  GAACTGAATTTTGCATCAATGAGCTACAATTTGAATAATCCAAGCCAGGAATCTCCAATT   300
301  TCAAAGCTCACCATTTTGCATCTCATGNTCATGTTATGTTCCAGTAAAATTGGCTTTAAA   360
361  CATTTACCTCAACAAAGAAAAACTAGCAAGTGAAGTTAAAGATGAGGACTTGGATTCGAA   420
421  GTGGGTGTTTATGTTTGGTTTTTATTTGTTCATCCCCAATTTATTATTATATGTATAGCG   480
481  TACATCTTTTATCGTATTGATTGGCTTTTCTATATTTATT  520
```

Fig. 5.10 Sequence of genomic PCR product corresponding to the 3' UTR of cMeHRGP1. The nucleotides in capital letters are shared sequences between cMeHRGP1 and genomic PCR product. The intron was in lower case and the splicing sites are in bold.

5.2.6 HRGP Gene Organization

Genomic DNA Southern blot analysis was performed using a probe made from the PCR product covering most of the coding region and 3' end of cMeHRGP1 to estimate the size of HRGP gene family. Primers used for the PCR amplification were xp1 (CCT CCT CCT CCA TAC TAC TAT, from 456bp of the cDNA) and xr1. As shown in Fig.5.11 (left), there are many positive signals with different density in lane *Hind*III and *Bgl*II when the

membrane was washed at low stringency. When the membrane was washed at higher stringency, a single prominent band of large size was present in lane *Hind*III and *Bgl*II but two bands (3.3kb and 3.5kb) were present in lane *Xba*I (Fig.5.11, right). This indicates that cassava HRGPs are probably encoded by a multigene family.

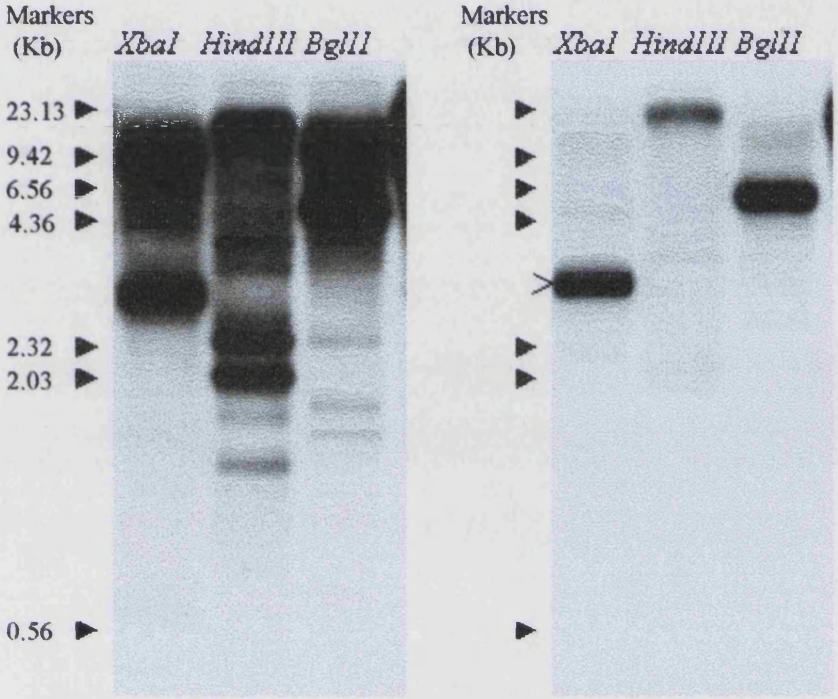


Fig.5.11 HRGP gene organisation in cassava. About 10μg of genomic DNA was digested with restriction enzymes. After hybridization with a probe made from the region corresponding to repetitive sequences at 60°C overnight, the membrane was washed at low or high stringency. **Left:** low stringency 60°C, 1xSSC, 0.1%SDS, 2x10min; **Right:** high stringency 60°C, 0.1xSSC, 0.1%SDS, 2x10min. The size marker on the right membrane was the same as on the left membrane.

5.2.7 Preliminary Analysis of Other HRGP cDNA Clones

Genomic Southern blotting indicated that HRGPs were encoded by a multigene family in cassava. PCR products of the inserts of the isolated positive cDNA clones appeared with hybridization signals of different intensity to the same probe (Fig.5.1a & b), suggesting that these clones were not the same at the nucleotide sequence level. Four of the eight un-analyzed positive clones, clone B, C, D, H, with relatively larger inserts, was partially sequenced and characterized from PCR-amplified cDNA inserts using λgt10 primers.

5.2.7.1 A cDNA clone with the same deduced amino acid sequence as cMeHRGP1

The partial sequence of the clone B cDNA showed 93% identity over 271bp to cMeHRGP (Fig. 5.12). All the identified nucleotide changes in the comparable region (271bp) occurred in the third nucleotide of the codons for amino acid residues. These did not change the encoded amino acid residues, as the cDNA of clone B had the same repetitive motif SP_4-Y_3-H-SP_4-VK. The partial sequence of the clone B cDNA was of high identity to cMeHRGP1 and coded for the same amino acid sequence or repetitive motifs.

```
cMeHRGP1 501 TCACCTCCACCTCCCTACTACTACAAATCTCCTCCACCACCGTCTCCATC 550
                |||||||| ||||| |||||||| || || || || ||||| ||||||||
Clone B   28 TCACCTCCTCCTCCATACTACTATAAGTCCCCACCTCCACCATCTCCATC 77
               S  P  P  P  P  Y  Y  Y  K  S  P  P  P  P  S  P  S

          551 TCCTCCTCCCCCATACTACTACCACTCACCACCACCACCAGTGAAATCTC 600
              ||||||||||||||||||||||||||||||||||||||||||||||||||
           78 TCCTCCTCCCCCATACTACTACCACTCACCACCACCACCAGTGAAATCTC 127
               P  P  P  P  Y  Y  Y  H  S  P  P  P  P  V  K  S  P

          601 CCCCTCCCCCATACTACTACCACTCGCCACCACCTCCTGTGAAATCACCT 650
              |||||||||||||||||||||||| |||||||||| ||||||||||||
          128 CCCCTCCCCCATACTACTACCACTCACCACCACCTCCCGTGAAATCACCT 177
                P  P  P  Y  Y  Y  H  S  P  P  P  P  V  K  S  P

          651 CCTCCTCCATACTACTACCACTCACCACCACCACCTGTGAAATCACCTCC 700
              ||||||||||||||||||||||||||||||||||||||||||||||||||
          178 CCTCCTCCATACTACTACCACTCACCACCACCACCTGTGAAATCACCTCC 227
              P  P  P  Y  Y  Y  H  S  P  P  P  P  V  K  S  P  P

          701 TCCTCCATACTACTATTACCACTCACCACCTCCTCCAGTAAAATCACCTCCTC 750
              ||||||||||||||||||||||||||||||||||||||||||||||||||||
          228 TCCTCCATACTACTATTACCACTCACCACCTCCTCCAGTAAAATCACCTCCTC 277
               P  P  Y  Y  Y  H  S  P  P  P  P  V  K  S  P  P  P

          751 CTCCATACTACTACCACTCAC 771
              |||||||||||||||||||||
          278 CTCCATACTACTACCACTCAC 298
                P  Y  Y  Y  H  S
```

Fig. 5.12 Comparison of the nucleotide sequences between cMeHRGP1 and clone B cDNA and the deduced amino acid residues. The different nucleotides between two sequences are not linked with vertical dashes and the related codons are underlined.

5.2.7.2 A cDNA encoding lysine-, tyrosine- and hydroxyproline-rich glycoprotein

Clone C hybridized strongly with the probe from parsley extensin Eli 09 (Fig.5.1b). The partial sequence of the clone C cDNA showed high similarity only with Eli09 when it was compared with the DNA data in NCBI using Blastn program. The cassava cDNA and parsley Eli09 shared 77% identity over 358bp. The cDNA was only 69% identical to cMeHRGP1 over 107bp (Fig. 5.13). The deduced polypeptide had SP_4 and YYY repeats, two AYGK repeats, two polyproline units WP_4 and HP_3, seven palindromes of at least four amino acid residues and many palindromes of three residues (Fig.5.14). The partial polypeptide is rich in proline (17mole%), lysine (K) (16%), tyrosine (Y) (9%), valine (8%), glycine (7%) and histidine (H) (5%).

```
cMeHRGP1 918 CCACCACCCGTGAAATCACCTCCTCCCCCATACTACTACCACTCGCCTCC 967
             |||||||||  ||||||||  | ||  ||   || ||  || |  ||||
Clone C  125 CCACCACCCCCAAAATCACCAGCACCAACACCTTATTATTACCCTTCTCC 174
                     .        .        .        .        .
         968 TCCTCCAGTGAAATCACCACCTCCTCCATACTACTATCATTCACCACCAC 1017
             || |||   ||  | |  ||||| || |||||||||  ||| |||||||
         175 ACCACCACCTAAGGCTCATCCTCCGCCGTACTACTATACTTCTCCACCAC 224

        1018 CACCAGT 1024
             |||| ||
         225 CACCTGT 231
```

Fig. 5.13 Comparison of nucleotide sequences between cMeHRGP1 and clone C.

```
121  ATGGCCACCACCCCCAAAATCACCAGCACCAACACCTTATTATTACCCTTCTCCACCACC  180
      W  P  P  P  P  K  S  P  A  P  T  P  Y  Y  Y  P  S  P  P  P

181  ACCTAAGGCTCATCCTCCGCCGTACTACTATACTTCTCCACCACCACCTGTTCCTTATCC  240
      P  K  A  H  P  P  P  Y  Y  Y  T  S  P  P  P  P  V  P  Y  P

241  TCACCCGTCTCCACATCCCCATCATTTCGTCGTAAAGGTGGTAGGAAAGGTCTATTGCTA  300
      H  P  S  P  H  P  H  H  F  V  V  K  V  V  G  K  V  Y  C  Y

301  CAGATGCTATGACTGGGATTATCCTGAAAAATCACATGACAAGAAGCATCTCAAAGGTGC  360
      R  C  Y  D  W  D  Y  P  E  K  S  H  D  K  K  H  L  K  G  A

361  CGTAGTGGAGGTAACATGCAAGACAGGTGAAAAGGAGATCAAGGCTTATGGTAAAACCAA  420
      V  V  E  V  T  C  K  T  G  E  K  E  I  K  A  Y  G  K  T  K

421  GATCAACGGTAAATACAGTATCACCGTTGAGGGCTTTGCGTATGGCAAATATGGAGCCGA  480
      I  N  G  K  Y  S  I  T  V  E  G  F  A  Y  G  K  Y  G  A  E

481  GGCTTGCAAGGCTAAGCTCCATAAGGCACCCAAAGGCTCACCATGCAACATACCAACTAA  540
      A  C  K  A  K  L  H  K  A  P  K  G  S  P  C  N  I  P  T  N

541  CCTCCACTGGGGCAAGAAGGGTGCCAAGCTCAAGGTGAAGTCCAAGACaaaagtatgaag  600
      L  H  W  G  K  K  G  A  K  L  K  V  K  S  K  T  k  v  *  s
```

Fig. 5.14 Deduced amino acid sequence from the partial nucleotide sequence of cDNA clone C. SP_4 and YYY are in bold, other XP_n and repeats are underlined. Palindrome regions of amino acid sequence are over-lined. Palindrome-like structures are indicated with double underlines. Palindrome overlapped with another one is dot-underlined. The ambiguous region of the sequence is in lower case.

5.2.7.3 A cDNA counterpart of genomic PCR product xgcm3

Partial sequence of the cDNA in clone D was identical to the 3'end of xgcm3, which was amplified from genomic DNA using primers designed from cMeHRGP1 to check the fidelity of cMeHRGP1 (see 5.2.4). Therefore the cDNA in clone D was highly identical to cMeHRGP1, at least in the determined sequence region, with just 3x6 extra nucleotides compared to cMeHRGP1 (Fig.5.15).

```
                     .        .        .        .        .
cMeHRGP1 985 CACCTCCTCCATACTACTATCATTCACCACCACCACCAGTGAAATCACCT 1034
             ||||||||||||||||||||||||||||||||||||||||||||||||||
Clone D  107 CACCTCCTCCATACTACTATCATTCACCACCACCACCAGTGAAATCACCT 156
```

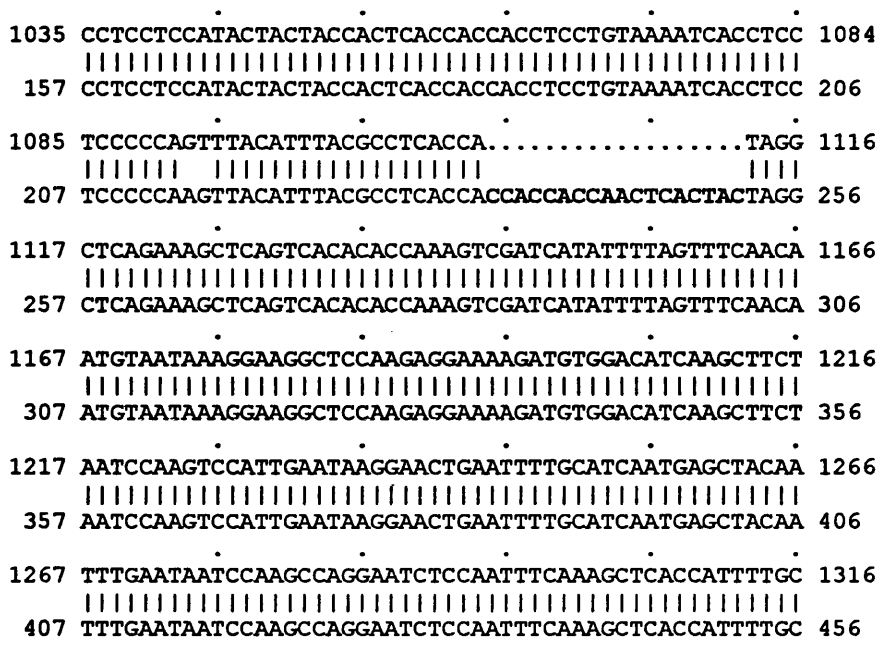


Fig. 5.15 Comparison between cMeHRGP1 and the partial sequence of clone D cDNA.

Clone F was identified to be identical to 3' coding region of cMeHRGP1. Very limited sequence information was obtained about clone H. The sequence determined from clone H cDNA was ambiguous but it showed 74% identity to cMeHRGP1 over 170bp and 69% identity over 170bp to the probe Eli09 used for library screening.

5.2.8 Summary

Nine positive cDNA clones were isolated from PPD-related library using parsley HRGP cDNA Eli09. One cDNA, cMeHRGP1, was fully characterised. The deduced amino acid sequence of cMeHRGP1 consisted of repetitive motifs which exist in HRGPs of other plants and large, combined repetitive motifs containing dicot repeat and monocot repeat units. Part of the antisense strand of cMeHRGP1 encoded phosphoserine aminotransferase, one of the enzymes in the phosphorylated pathway of serine biosynthesis. PCR amplification of genomic DNA using primers from cMeHRGP1 covering the junction of the coding region of HRGP and phosphoserine aminotransferase and sequence of the genomic PCR product excluded to a large extent the possibility that cMeHRGP1 is a chimeric clone. In other words, it is probable that MeHRGP1 and phosphoserine aminotransferase genes are an antisense gene pair.

Another five HRGP cDNA clones were preliminarily analysed. According to partial sequence, clone B was identified encoding the same polypeptide as cMeHRGP1, though its nucleotide sequence was slightly different from cMeHRGP1. Clone D had 18bp more just before the stop codon than cMeHRGP1. Clone F also showed high similarity to cMeHRGP1 in nucleotide sequence. Clones C and H were similar to parsley Eli09, the cDNA used as probe for screening the cDNA library. Clone C was identified encoding lysine tyrosine-rich HRGP with two SP_4 and some palindromes.

The isolation of at least three different HRGP cDNA clones and the result of genomic Southern blot hybridization suggested that HRGPs in cassava were encoded by a multigene family.

5.3 Discussion

5.3.1 HRGP cDNA Clones

One HRGP cDNA clone, cMeHRGP1 was fully characterized. Five of the remaining eight isolated positive HRGP cDNA clones were subjected to preliminary analysis. These clones fall into two groups according to their similarity. Three clones including clone B, D and F were of high similarity to cMeHRGP1. The other clones including clone C and H were not so similar to cMeHRGP1, but had higher similarity than all the other four analyzed clones to parsley extensin Eli09. This explained why only clone C and H showed strong signals when the PCR products (cDNA inserts) of the different clones were hybridized against the probe made from parsley Eli09 (Fig. 5.2b). The deduced amino acid sequences of all the clones had SPPPP repeats and YYY motifs, and were rich in proline and tyrosine. The deduced amino acid sequence of clone B did not comprise entirely repetitive motifs, with a large proportion of non-repetitive region.

5.3.2 Antisense Gene Pairing of cMeHRGP1

The antisense strand of the 5' part of the characterized cMeHRGP1 encoded putative phosphoserine aminotransferase. Investigation using PCR techniques concluded that it was most likely there was a genomic counterpart of this cDNA, which excluded to a large extent the possibility of its being a chimeric clone. The question remains as to whether the

genomic DNA of cMeHRGP1 encodes HRGP by the sense strand and phosphoserine aminotransferase by the antisense strand.

Long open reading frames on the antisense strand of open reading frames are more frequent than expected (Silke, 1997). The antisense of many genes may code for unidentified proteins (Knee and Murphy, 1997; Silke, 1997). Some deduced amino acid sequences of antisense strands in some yeast genes showed high homology to some sequences in Genebank database (Cebrat *et al.*, 1998). One gene encoding basic fibroblast growth factor (bFGF) in *Xenopus oocytes* was located on the antisense strand of another bFGF gene (Kimelman and Kirschner, 1989). The complementary strand of heat shock protein HSP70 gene in *Achlya klebsiana* codes for an NAD-specific glutamate dehydrogenase (NAD-GDH) gene inducible by L-glutamine. The antisense strand of the coding region of HSP70 corresponded to the last and the largest exon of NAD-GDH gene and the antisense strand of the 3'untranslated region in HSP70 corresponded to the the introns and exons of NAD-GDH gene (LeJohn *et al.*, 1994).

Although the antisense strand of the 5' untranslated region in the isolated HRGP clone coded for phosphoserine aminotransferase, the antisense strand of the coding region of the HRGP was unlikely to code for any proteins due to its repetitive sequences and the abundance of serine whose antisense codon is a stop codon. What is likely though, is that the major part of the coding region of cMeHRGP1 may be an intron of the antisense gene. Genes existing in the introns of other genes or overlapping genes were reported (Nonnelye and Laudet, 1994). The 3'end of cMeHRGP1 may serve as the 5'untranslated region and the initial coding region of the antisense gene (Fig.5.16).

By comparing the deduced amino acid sequences of the antisense strand and known phosphoserine aminotransferase genes in other organisms, it was found that the amino acid sequence of the antisense strand was the near N terminal part of the phosphoserine aminotransferase peptide, possibly only a few or about 50 amino acid residues away from the N terminal. Is the missing N terminal in the antisense strand of the 5' untranslated region of cMeHRGP1? One possible translation start codon was located next to the stop codon of cMeHRGP1, which led a possible short N terminal of phosphoserine

aminotransferase (Fig.5.17). Isolation of the cDNA and genomic clone(s) of phosphoserine aminotransferase needs to be carried out to confirm further its antisense gene pairing with cMeHRGP1, which could also provide more information for identifying the N terminal coding region and intron(s).

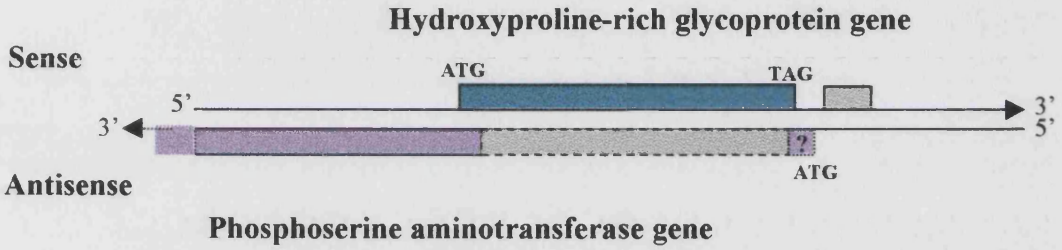


Fig.5.16 Schematic illustration of the organisation of MeHRGP1 and phosphoserine aminotransferase genes. The sense strand codes for MeHRGP1, with the green close box referrring to the coding region and grey box to the intron. The antisense strand codes for phosphoserine aminotransferase gene, with purple closed box for the known part of coding region, dot-line-closed purple box for possible leading sequence of the coding region, the purple shading by the 3'end for the unknown coding region, dot-line-closed grey box for the possible intron.

```
          S  P  P  P  P  Y  Y  Y  H  S  P  P  P  P  V  K  S  P  P  P
cMeHRGP1  TCACCTCCTCCTCCATACTACTACCACTCACCACCACCTCCTGTAAAATCACCTCCTCCC
    1029  -+---------+---------+---------+---------+---------+---------+--------- 1088
          AGTGGAGGAGGAGGTATGATGATGGTGAGTGGTGGTGGAGGACATTTTAGTGGAGGAGGG
             V  E  E  E  M  S  S  G  S  V  V  V  E  Q  L  I  V  E  E
           F  D  G  G  G  G  Y  *  *  W  E  G  G  G  G  T  F  D  G  G
            F  *  R  R  R  W  V  V  V  V  *  W  W  R  R  Y  F  *  R  R

          P  V  Y  I  Y
          CCAGTTTACATTTAC
          -+---------+---
          GGTCAAATGTAAATG
          G  L  K  C  K
           G  T  *  M  *
            W  N  V  N  V

          A  S  P  *
 PCR      GCCTCACCATAGGCTCAGAAAGCTCAGTCACACACCAAAGTCGATCATgtaagtttcttt
gxer1     ---------+---------+---------+---------+---------+---------+ 60
          CGGAGTGGTATCCGAGTCTTTCGAGTCAGTGTGTGGTTTCAGCTAGTAcattcaaagaaa
d   ←    R  R  V  M ?P  E  S  L  E  T  V  C  W  L  R  D  H  L  N  R   -
e         A  E  G  Y  A  *  F  A  *  D  C  V  L  T  S  *  T  L  K  K  -
f          G  *  W  L  S  L  F  S  L  *  V  G  F  D  I  M  Y  T  E  K -

          aacacttctaacatcatttttcaagtttttagtattagcaaaatgaatcattttgaagt
       61 ---------+---------+---------+---------+---------+---------+ 120
          ttgtgaagattgtagtaaaaagttcaaaaaatcataatcgttttacttagtaaaacttca
d         *  C  K  *  C  *  K  E  L  K  K  T  N  A  F  H  I  M  K  F   -
e          L  V  E  L  M  M  K  *  T  K  *  Y  *  C  F  S  D  N  Q  L  -
f           V  S  R  V  D  N  K  L  N  K  L  I  L  L  I  F  *  K  S  T -
```

```
        gtatgtctaattttctgccatttccatacaacagATTTTAGTTTCAACAATGTAATAAAG
    121 ---------+---------+---------+---------+---------+---------+ 180
        catacagattaaaagacggtaaaggtatgttgtcTAAAATCAAAGTTGTTACATTATTTC
d       H  I  D  L  K  R  G  N  G  Y  L  L  N  *  N  *  C  H  L  L  -
e        T  H  R  I  K  Q  W  K  W  V  V  S  K  L  K  L  L  T  I  F  -
f         Y  T  *  N  E  A  M  E  M  C  C  I  K  T  E  V  I  Y  Y  L -

        TACATCTTTTATCGTATTGATTGGCTTTTCTATATTTATT
    481 ---------+---------+---------+---------+ 520
        ATGTAGAAAATAGCATAACTAACCGAAAAGATATAAATAA
d       T  C  R  K  D  Y  Q  N  A  K  R  Y  K  N -
e        Y  M  K  *  R  I  S  Q  S  K  *  I  *   -
f         V  D  K  I  T  N  I  P  K  E  I  N  I  -
```

Fig.5.17 Sense and antisense translation of the 3' end of the coding region of cMeHRGP1 and PCR sequence of the genomic DNA corresponding to the 3' part of cMeHRGP1 with primers xf1 and xr1. The sequence of the identified intron is in bold and in lower case. The possible antisense translation region is underlined, with M marked as a question mark.

The biological impact of the transcript from the antisense strand or natural antisense RNA, may lie in the regulation of the expression of the sense gene. A growing number of eukaryotic genes are thought to be regulated at least in part by natural antisense RNA transcribed from the presumptive non-coding DNA strand (Knee and Murphy, 1997). Modification of A to G was detected in the sense transcripts of fibroblast growth factor gene, which was thought to be the result of the action of a modifying enzyme on the sense and antisense duplex. The modification may lead to an inactive polypeptide as the modifications occurred in the region that was suggested to be crucial for the binding to the bFGF receptor (Kimelman and Kirschner, 1989).

The possible roles of the antisense gene pair MeHRGP1 and phosphoserine aminotransferase gene in the gene regulation could be considered when sufficient expression data about these two genes has been obtained. However, it may not be a coincidence that phosphoserine aminotransferase (involved in the biosynthesis of serine) and HRGP1 (serine-rich) are encoded by complementary strand DNA. Considering that phosphorylated pathway of serine biosynthesis plays the primary role in the total cellular supply of serine in tissues of low photosynthetic activity and that fact that cassava storage root is poor in protein content and HRGPs may be one of the major proteins, also that serine is one of the few amino acids in HRGPs, it is tenable that serine biosynthesis may have a closer link with HRGP gene expression in roots than in leaves.

5.3.3 Repetitive Motifs and Peptide Palindromes in HRGPs

The deduced amino acid sequence showed typical repeat motifs as found in HRGPs of other plants (Table 5.1). The SP_4 motif, which leads to molecular inflexibility and wall self-assembly, is the most common repeat motif in HRGPs. There are 28 SP_4 repeats in HRGP1. There are two SP_4-Y_3K-SP_4-SP motifs and eleven SP_4-Y_3H-SP_4-VK repeat units. SP_4-Y_3K motifs were observed in HRGPs of bean cells, tomato and potato, SP_4-Y_3H repeats in rape and soybean, SP_4-VK-SP_4 in maize (Sommer-Knudsen, 1998), but the long and main repeat motif SP_4-Y_3H(or K)-SP_4-VK or SP_4-VK-SP_4-Y_3H(or K) has not been reported.

Table 5.1 Extensin repetitive motifs in cassava and other plant systems

Sequence	Source	Accession/References
SPPPPKH	Soybean	U44838
SPPPPKK	Rape	A18822
SPPPPKKPYYPP	Tobacco	L38908
SPPPPSP	Bean, Almond, tomato	Sauer *et al.*, 1990, X65718, X55685
SPPPPSPKYVYK	Tomato	X55683
SPPPPSPSPPPP	Bean cells, tomato	Chen *et al.*,1992, X55685
SPPPPYYYH	Soybean, bean	L22030, U18791
SPPPPYYYHSPPPPVK	**Cassava**	cMeHRGP1
SPPPPVK	Rape	Evans *et al.*, 1990
SPPPPVKSPPPP	Maize	Z34475
SPPPPVKSPPPPYYYH	**Cassava**	cMeHRGP1
SPPPPYYYK	tomato, soybean	X55685, L22030
SPPPPYYYKSPPPPSP	**Cassava**	cMeHRGP1
SPPPPTPVYK	Carrot, tobacco	X02873, U38908
SPPPPVYK	soybean, bean	U22029, U18791
SPPPPVYSPPPP	Bean	Sauer *et al.*, 1990
SPPPPVHSSPPPPVA	Tomato	Showalter *et al.*, 1985

The tyrosine in Tyr-X-Tyr-Lys (X refers to any residues), Y_3-H or Y_3-K provides the possibility of both inter- and intra-molecular cross-link through the formation of isodityrosine (IDT), increasing rigidity and hydrophobicity of the molecules (Kieliszewski and Lamport, 1994; Sommer-Knudsen, 1998), which may strengthen the cell wall and form a barrier to prevent water loss at wound sites or pathogen infection.

Palindromes PPPPSPPPP, YYY or YIY were present in the polypeptide of cMeHRGP1. The same structures also exist in other plants (Kieliszewski and Lamport, 1994). In the lysine and tyrosine-rich hydroxyproline-rich glycoprotein encoded by clone C, there were

different palindromes such as PYYYP, HPSPH, HHFHH, VVKVV, YDWDY and GKKG. The palindromic peptides could create centrosystemetric domains, which may act as self-assembly nucleation sites. In other words, an intermolecular interaction may establish a structure to initiate a succession of intermolecular reaction or 'growth' (Kieliszewski and Lamport, 1994; Lindsey, 1991).

5.3.4 Expression of HRGP Genes

Increases in the level of extensin transcripts upon wounding stimulation have been observed in many plant systems, which provided the evidence for the assumed function of HRGPs based on the characteristics of their amino acid sequences. Wounding of potato tubers caused a marked increase in extensin mRNA within 18 hr (Bown *et al.*, 1993). Two extensin mRNAs accumulated differentially after wounding in carrot (Tierney *et al.*, 1988); Excision of hypocotyls induced strong expression of an extensin gene within 1.5hr in bean (Corbin *et al.*, 1987). Two classes of extensins were accumulated locally after wounding in tomato (Showalter *et al.*, 1992). A root specific extensin gene was induced upon wounding in *Brassica napus* (Shirsat *et al.*, 1996). Induction of extensin mRNAs was also observed in tobacco hypocotyls, leaves and stems after wounding (Showalter *et al.*, 1985; Wycoff *et al.*, 1995; Parmentier *et al.*, 1995; Memelink *et al.*, 1993). Three HRGP genes in bean were differently regulated by wounding response. The expression of Hyp3.6 was increased after wounding, Hyp2.13 and Hyp4.1 were induced 4hr after wounding and Hyp4.1 showed the strongest responses (Corbin *et al.*, 1987). The effect of HRGP gene expression on wound healing was demonstrated by the impaired healing response in potatoes after gamma irradiation (Ussuf *et al.*, 1996).

Most of the wound inducible HRGP genes have also been found to be inducible upon pathogen infection. Wound-inducible HRGP genes in bean showed induction of transcripts upon infection with either compatible or incompatible pathogens (Corbin *et al.*, 1987). An extensin gene in *N. plumbaginifolia* was more highly expressed in roots than in the leaves, and the expression in leaves was enhanced by mechanical stimuli, salicylic acid and infection with the incompatible bacterium *Pseudomonas syringae* (Tire *et al.*, 1994).

The special feature of HRGP sequences and the increased expression of HRGP genes during wounding and pathogen infection suggested their functions in sealing-off the wound sites to prevent the desiccation or/and the penetration of pathogen. It has been shown that proline-rich cell wall protein was rapidly insolubilised at wound sites and it was demonstrated that H_2O_2 mediated the rapid insolubilization in soybean cells and bean cells incubated with a fungal elicitor (Bradley *et al.*, 1992). Tomato extensin insolubilization (measured with the elutability of extensin by $CaCl_2$) was detected immediately in tomato cells after being challenge with a yeast elicitor; this was earlier than the peak production of active oxygen species (Brownleader *et al.*, 1997). Brownleader *et al.* (1997) then suggested that there was a primary defence mechanism incorporating pre-existing HRGPs, peroxidase and peroxide-generating system. In bean, rapid wound activation of the HRGP4.1 promoter was localized primarily to the outer phloem with weaker expression in the inner phloem. It was suggested that the reinforcement of cell walls in the outer phloem by increased HRGP4.1 might be important in preventing entry of pathogens into the vascular system at wound sites (Wycoff *et al.*, 1995).

The isolation of HRGP cDNA clones from the PPD-related cDNA library indicated that HRGP genes were expressed during the wounding response. This suggested that at least one element of the healing process was present during the development of PPD. However, the expression of HRGPs genes might not be enough to seal off the wound sites and terminate PPD. The possible role of HRGPs in wound healing during PPD could be investigated by comparing the expression of HRGP genes during the wound healing process of the roots in natural conditions (attached to the plant) and the deteriorating process of the roots after harvest, and also by comparing the formation and insolubilization of the HRGPs during these different processes. Their role could also be tested by comparing the expression patterns of HRGP genes and rate of insolubilization in cultivars with different responses to PPD. It may be expected that cultivars showing resistance to PPD may express HRGP genes at higher level and HRGPs insolubilized at a higher rate so that the water loss from the wound sites could be reduced and the rate of PPD decreased.

Wound inducible extensin genes can be characterised by the presence of $(Y)_3$ tyrosine blocks in the protein. The phenomenon was first reported in tomato and in bean (Corbin *et*

al., 1987; Showalter *et al.*, 1991), and later in potato (Bown *et al.*, 1993) and tobacco (Parmentier *et al.*, 1995). Other wound inducible extensin genes such as tobacco pCNT1 and soybean SbHRGP3 were also found to be rich in tyrosine residues in their proteins (Memelink *et al.*, 1993; Ahn *et al.*, 1998). In tobacco, Ext 1.2 with Y3 repeat blocks was strongly induced by wounding (Parmentier *et al.*, 1995), while Ext 1.4 with much less tyrosine residues did not appear to be regulated by wounding (Hirsinger *et al.*, 1997). It was also reported that an extensin-like gene lacking tyrosine residues was not induced by mechanical wounding in *Nicotiana alata* (Chen *et al.*, 1992). Cassava MeHRGP1 was rich in tyrosine residues with many Y_3 blocks, which may indicate that MeHRGP1 was induced by wounding during PPD. The expression patterns of MeHRGP1 upon wounding or during PPD would provide answers to this assumption.

The high level of tyrosine organized in Y_3 blocks that are regularly distributed in the polypeptide, may facilitate erection of a structural barrier to infection by extensive intra- and inter-molecular HRGP oxidative cross-linking and provision of sites for the deposition of stress-induced lignin as a further structural barrier to infection (Corbin *et al.*, 1987; Whitmore, 1978). Immunolocalizations indicated an association of the extensin protein with lignin deposition in the xylem vessel cell walls in *N. plumbaginifolia* (Tire *et al.*, 1994). Therefore the extensins with Y3 repeat blocks may play a more important role than other extensins in sealing off the wounding sites.

5.3.5 HRGP Gene Family

In the graminaceous species studied so far (*Zea diploperennis teosinte*, maize, sorghum and rice), HRGP appeared to be encoded by a single gene showing high similarity between species (Menossi *et al.*, 1997). However, extensins are encoded by a multigene family in the dicot plant systems studied to date. Two extensin genes have been isolated from *Brassica napus* (Evans *et al.*, 1990) and carrot (Tierney *et al.*, 1988), three from soybean (Hong *et al.*, 1994) and bean (Corbin *et al.*, 1987), and at least seven from tomato (Showalter *et al.*, 1991). Southern blotting of genomic DNA indicated that there was a multigene family of extensin genes in cassava. The preliminary analysis of five cDNA clones besides cMeHRGP1 demonstrated that their partial sequences were different from each other. But it could not be excluded that some of them might be different parts of one

gene. What can be sure is that cMeHRGP1, clone C the lysine tyrosine-rich HRGP cDNA and clone D (with extra 18 bp compared to cMeHRGP1) are from three different genes. It was suggested by the result of high stringency hybridisation, that there might be two copies of genes corresponding to cMeHRGP1 or another gene of very high similarity to cMeHRGP1. The identification of a cDNA with almost identical sequence (18bp extra) to cMeHRGP1 by the 3'end supported the later suggestion.

Chapter six

Isolation and characterisation of β-1,3-glucanase cDNAs expressed during PPD

6.1 Introduction

β-1,3-glucanases accumulate developmentally (Ori *et al.*, 1990) and can be induced by pathogen infection and wounding (Brederode *et al.*, 1991; Simons *et al.*, 1992). Their role in plant defense against fungal pathogens has been the focus of much research. β-1,3-glucanase genes were extensively studied in tobacco and grouped into four classes. Class I β-1,3-glucanase genes encode proteins with basic isoelectric point, which were located in the vacuole of cultured cells and leaves (Van den Bulcke *et al.*, 1989; Keefe *et al.*, 1990). These genes were expressed in mature leaves and roots, and in response to pathogen invasion (Linlhorst *et al.*, 1990). Class II genes encode acidic (low pI) and extracellular β-1,3-glucanases. Genes in this group did not accumulate in healthy leaves but were induced by pathogen attack. Some of the originally designated PR proteins identified through protein analysis of the interaction between plant and pathogen were encoded by genes in this group, such as PR-2, PR-N, PR-O (Ward *et al.*, 1991). An acidic and pathogen-induced β-1,3-glucanase PR-Q' was typed as class III due to its highly diverged sequence from class I and II (Payne *et al.*, 1990). Class IV includes the β-1,3-glucanase genes specifically expressed in reproductive tissues, sp41a and sp41b in stylar matrix (Ori *et al.*, 1990) and *Tag*1 in anthers (Bucciaglia and Smith, 1994). Three out of four classes of the β-1,3-glucanases in tobacco were related to disease resistance. It was proposed that β-1,3-glucanase induced by pathogen infection, functions in defense through breaking down fungal cell walls and releasing an elicitor of phytoalexin biosynthesis from the mycelial walls of a pathogenic fungus (Ham *et al.*, 1991). But how they function in wound response is unknown. Studying the gene expression of β-1,3-glucanases may therefore lead to understanding part of the signal pathway during PPD and the interaction between cassava and the fungal pathogens involved in microbial deterioration.

6.2 Results

6.2.1 Isolation of β-1,3-glucanase cDNA Clones

About $5x10^4$ phage from the PPD-related cDNA library were plated onto agar plates (130mm in diameter) and screened for β-1,3-glucanase clones. The probe was prepared from mixed templates of β-1,3-glucanase cDNA pBEG (de Loose *et al.*, 1988) from *Nicotiana plumbaginifolia*, and cDNAs PR2 (acidic and extracellular PR protein) from tobacco (Ward *et al.*, 1991). Prehybridisation was carried as described in 2.9 and hybridization was performed at 50°C overnight, which was then followed by a low stringency wash with 2xSSC, at room temperature, 2x10minutes; 2xSSC, 0.1% SDS, 50°C, 2x10 minutes. From the first screening 8 positives were obtained. Each of the positives was then plated out at low density (about 200 plaques a plate 90mm in diameter) and screened under the same conditions as for the first screening. One single positive plaque was isolated from each of the 8 positives.

These positive clones were subjected to PCR using λgt10 primers to check the size of the cDNA inserts. Two clones E and F were about 2.2kb; the others were about 1.4 to 1.6kb (Fig.6.1a). The PCR products of these clones were Southern-blotted and hybridized against the mixed probe. All the clones showed positive signals (Fig.6.1b). The PCR products from D and F were gel-purified and partially sequenced using λgt10 primers. The determined sequence showed high similarity to those of glucanase genes from many plants such as *Prunus persica* β-1,3-glucanase gene (70% identity in 328bp overlap) and *Nicotiana tabacum* cDNA (63% in 418bp overlap).

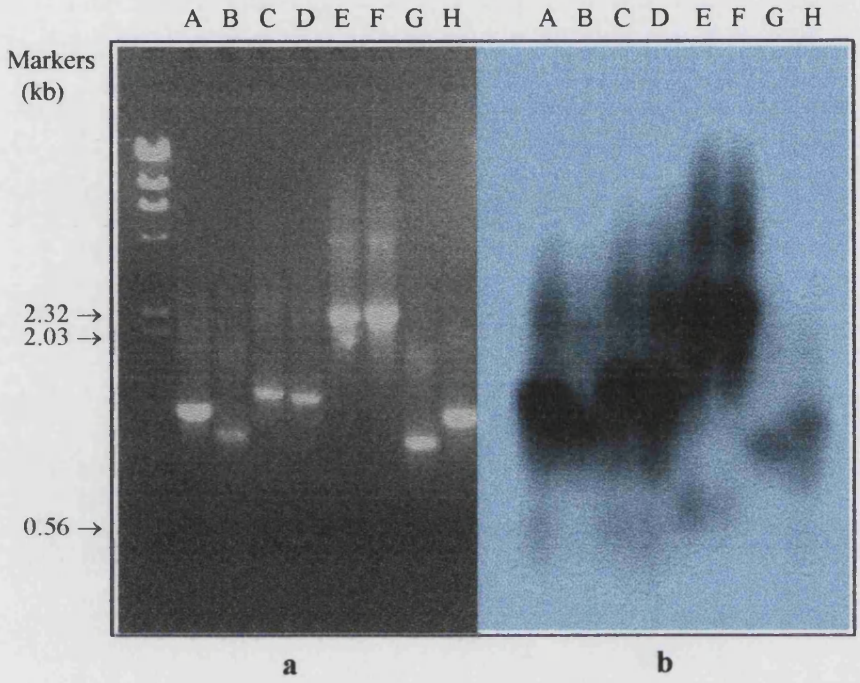


Fig.6.1 Analysis of positive β-1,3-glucanase clones with PCR and Southern hybridisation.

a: PCR amplification of inserts from positive plaques was performed as described in 2.2.7. The resulted PCR products were run in 1.0% (1xTBE) agarose gel containing ethidium bromide and photographed. The size of the PCR products for clone A to H is about 1.5kb, 1.4kb, 1.6kb, 1.6kb, 2.3kb, 2.3kb, 1.4kb and 1.5kb respectively.
b: Southern blot analysis of the PCR products. The PCR products in the gel were blotted to Hybond N^+ nylon membrane and hybridised with the mixed probes made from β-glucanase cDNA pBEG (de Loose *et al.*, 1988) from *Nicotiana plumbaginifolia*, glucanase cDNAs PR2 (Ward *et al.*, 1991) from tobacco at 50°C overnight. The membranes were then washed subsequently with 2xSSC/0.1%SDS at room temperature

6.2.2 Characterisation of the Clone cMeGluc1 & 2

The 1.3kb and 2.1kb cDNA inserts in clone D and F, cMeGluc1 and cMeGluc2 were subcloned into pUC18 and pBS KS(II+) respectively.

The sequences of cMeGluc1 and cMeGluc2 were 1015bp and 1694bp respectively (Fig.6.2). The cMeGluc1 consists of a partial coding region and a 117bp 3' end sequence

including the 3'untranslated region and a polyA tail. A putative polyadenylation signal AATAAA was located 86bp upstream of the polyA tail. The nucleotide sequence flanking the second methionine, ACATGGCT, is of very high similarity to the consensus sequence (A/C)aATGGCT flanking the translation start codon in dicots or plants (Joshi *et al.*, 1997). The sequence of cMeGluc1 was almost identical to the second half of cMeGluc2 except that there was just one base change, which lead the change of amino acid residue from lysine (K) in cMeGluc1 to threonine (T) in cMeGluc2 (Fig.6.2). A schematic diagram of these two cDNAs together with the sequence strategy was shown in Fig.6.3.

By searching the protein pattern database in PROSITE (ExPASy web page, 1998, http://expasy.hcuge.ch/) with deduced amino acid sequence of cMeGluc2, the signature IVVSESGWPTAG for glycosyl hydrolases was located in the amino acid sequence (in the region of 1359bp to 1392bp) (Fig.6.2). The glutamate (E) in the signature was the putative catalytic nucleophile. The last glutamate E in the peptide was putative proton donor during hydrolysis. The marked E (glutamate) and K (lysine) were strictly conserved residues, which were thought to be important in influencing the protonation state of the catalytic glutamate (Varghese *et al.*, 1994).

```
   1  CGTCATCCTCGGGATTTGGGATACTTGGTCTTTTCCTGAGCGGATGGCGATCTAGAGTTG   60
        S  S  S  G  F  G  I  L  G  L  F  L  S  G  W  R  S  R  V  A
                 .        .        .        .        .        .
  61  CTGCTGATCCACAGTTCCCTTTCAAAGTTCTAATGGAGGAATTGGTTGGTGTCTCCGCTT  120
        A  D  P  Q  F  P  F  K  V  L  ?M  E  E  L  V  G  V  S  A  C
                 .        .        .        .        .        .
 121  GTGTTCTTGGTGACATGGCTTCGCGTCCTAACTTTGGACTCAATGAACTCGACTTCGTAT  180
        V  L  G  D  ?M  A  S  R  P  N  F  G  L  N  E  L  D  F  V  F
                 .        .        .        .        .        .
 181  TCTCGACTCTGGTAGTTGGTTCCATAATGAATTTCACCCTCATGTACCTCTTGGCACCTA  240
        S  T  L  V  V  G  S  I  ?M  N  F  T  L  M  Y  L  L  A  P  T
                 .        .        .        .        .        .
 241  CCGCGTCTGCTGCAAGTACCAGTCTCCCTGCCATCTTCGCAAGTTGCCCAACAAGCCACA  300
        A  S  A  A  S  T  S  L  P  A  I  F  A  S  C  P  T  S  H  M
                 .        .        .        .        .        .
 301  TGTTTGAGCCTGGTGCCTTTACCCTCATGAATCGACTGGGCACTTTTGTTTACAAAGGAA  360
        F  E  P  G  A  F  T  L  M  N  R  L  G  T  F  V  Y  K  G  T
                 .        .        .        .        .        .
 361  CCATCTTTGCAGCTGTTGGTTTCGGCGCTGGACTAGTAGGAACTGCAATCTCAAATGGGT  420
        I  F  A  A  V  G  F  G  A  G  L  V  G  T  A  I  S  N  G  L
                 .        .        .        .        .        .
 421  TGATTGCGATGAGGAAGAAGATGGATCCAACTTTTGAGACGCCAAACAAGCCACCTCCAA  480
        I  A  M  R  K  K  M  D  P  T  F  E  T  P  N  K  P  P  P  T
```

```
481  CAGTTCTAAATGCAGTGACATGGGCTCTTCACATGGGCATTAGCAGTAACTTGAGATACC  540
       V  L  N  A  V  T  W  A  L  H  M  G  I  S  S  N  L  R  Y  Q

541  AAACTCTGAACGGTGTAGAGTTTTTGCTGCAGAAAGGGCTTTCTCCTCTGGCTTTCAAGA  600
       T  L  N  G  V  E  F  L  L  Q  K  G  L  S  P  L  A  F  K  S

601  GCTCAGTAATTGTTCTTAGATGCTTGAACAACGTGCTGGGTGGAATGTCGTTTGTTATAT  660
       S  V  I  V  L  R  C  L  N  N  V  L  G  G  M  S  F  V  I  L

661  TGGCAAGGTTAACAGGATGCTATGGAATGCTTGGTAACCTGCCACCACCAGCAGAAGTCG  720
       A  R  L  T  G  C  Y  G  M  L  G  N  L  P  P  P  A  E  V  V

721  TAAGCCTCTACAACCAGAACGGCATCCGCAGAATGCGAATCTACGACCCAAATCCAGATG  780
       S  L  Y  N  Q  N  G  I  R  R  M  R  I  Y  D  P  N  P  D  A

781  CTCTCCGAGCCCTTGGAGGCTCTAATATA
       L  R  A  L  G  G  S  N  I

                (cMeGluc1 start)  GAGCTCATTCTTGGCCTTCCAAATGATAAAC  840
                                    E  L  I  L  G  L  P  N  D  K  L

841  TTCAGAGTATTGCTTCCAACCAAGCTGAAGCAGATTCATGGGTTCAAAACAACGTAAAAA  900
       Q  S  I  A  S  N  Q  A  E  A  D  S  W  V  Q  N  N  V  K  N

901  ACCATGGAAATGTCAAGTTTCGTTACATCGCAGTTGGAAATGAGGTAAAGCCCTCAGCTG  960
       H  G  N  V  K  F  R  Y  I  A  V  G  N  E  V  K  P  S  A  A

961  CAGAAGCAGGATCTCTGTTCCCTGCTATGAGAAATATTCGCAACGCACTCAATTCTGCTG  1020
       E  A  G  S  L  F  P  A  M  R  N  I  R  N  A  L  N  S  A  G

1021 GTCTTGGAGGTATCAAAGTTTCCACTGCTATTGATACTATAGGCCTTACTGCAGATTCCT  1080
       L  G  G  I  K  V  S  T  A  I  D  T  I  G  L  T  A  D  S  F

1081 TTCCTCCCTCTAGGGGCTCTTTCAAGCCAGAATATCGTCAACTTCTTGATCCTGTAATAC  1140
       P  P  S  R  G  S  F  K  P  E  Y  R  Q  L  L  D  P  V  I  Q

1141 AATTTCTAGTGAACAATCAATCTCCATTGCTGGTTAACTTGTATCCATACTTCAGTTACA  1200
       F  L  V  N  N  Q  S  P  L  L  V  N  L  Y  P  Y  F  S  Y  R

1201 GAGATAGTCAGGGAACTATCAATCTTGATTATGCTCTTTTCAGACCGGCGCCGCCAGTCC  1260
       D  S  Q  G  T  I  N  L  D  Y  A  L  F  R  P  A  P  P  V  Q

1261 AAGATCCCGACGTCGGACGTACTTACCAAAACCTTTTCGATGCCATACTTGATACTGTGT  1320
       D  P  D  V  G  R  T  Y  Q  N  L  F  D  A  I  L  D  T  V  Y

1321 ATGCTGCGGTGGAGAAGGCTGGCGGAGGAGCTTTGGAGATTGTTGTATCAGAAAGTGGTT  1380
       A  A  V  E  K  A  G  G  G  A  L  E  I  V  V  S  E  S  G  W
                                                C

1381 GGCCTACTGCTGGAGGATTTGGAACTTCAGTTGAAAATGCAAAAACTTATAATAACAACT  1440
       P  T  A  G  G  F  G  T  S  V  E  N  A  K  T  Y  N  N  N  L
                                                T

1441 TGATTCAAAAAGTGAAAAATGGGACTCCAAAGAAGCCTGGAAAGCCCATTGAAACTTACA  1500
       I  Q  K  V  K  N  G  T  P  K  K  P  G  K  P  I  E  T  Y  I

1501 TTTTTGCCATGTTTGATGAAAGCAACAAAGGAGGTGAAGAGCTGGAGAAACATTGGGGAC  1560
       F  A  M  F  D  E  S  N  K  G  G  E  E  L  E  K  H  W  G  L

1561 TCTTTTCTCCAAACAAGCAGCCTAAGTACCCAGTCAATTTCAATTGAAATATTCCCTATA  1620
       F  S  P  N  K  Q  P  K  Y  P  V  N  F  N  *

1621 AATTGCTTCTAAGGGCAATCTCATGTATAACTTAATAAGAGTTTTGTAACAACTCCCTAA  1680

1681 TTTAAAAGATCGGG (the end of cMeGluc2)
```

```
                                                                    .
                    GAAGCTACATATATGTATGCTAATAAAAAGTTAAAAGAGTAGTTCT   1740

                                                                    .
1741  TCAATTACTTTAGTAAAATTTTCTGAATTTGATAAGTGCACTCTTTGTATTAATGTGGAT   1800
                .
1801  ATTATAAGAAAAAAAAAAAAAAA  1824
```

Fig.6.2 The nucleotide sequences and deduced amino acid sequences of cMeGluc1 and cMeGluc2. The possible translation start codons are indicated with '?' and the methionines underlined. The stop codon TGA is in bold. Putative polyadenylation signal AATAA is underlined. In the over-lapped sequences between cMeGluc1 and cMeGluc2 there is only one nucleotide difference, which is A in cMeGluc1 instead of C in cMeGluc2 at 1323bp. The single nucleotide change leads to the change of amino acid residue from lysine (K) to threonine (T) in cMeGluc2. The block arrow-marked glutamate E in the signature of glycosyl hydrolases (underlined) is the putative catalytic nucleophile. The last glutamate E marked with block arrow is putative proton donor during hydrolysis. The E and K (lysine) marked with ◈ are strictly conserved residues which are thought to be important in influencing the protonation state of the catalytic glutamate.

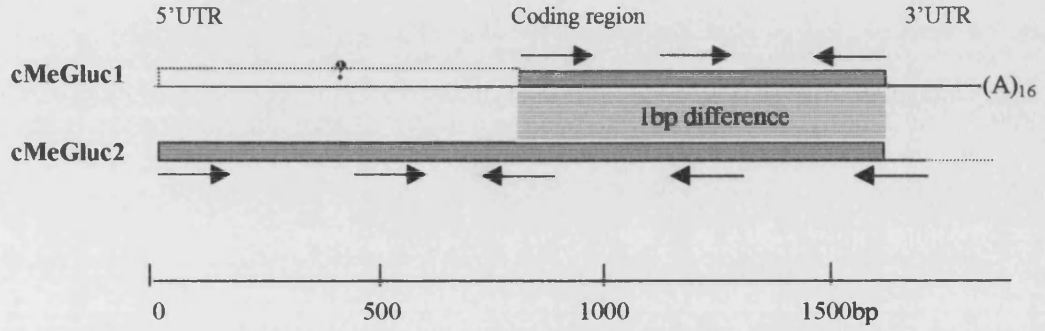


Fig.6.3 Schematic diagram of cMeGluc1 and cMeGluc2. The light shading between two cDNA clones indicates their identity in nucleotide sequences. The dotted line indicates the unknown sequences. The arrows indicate the primers used for sequencing.

6.2.3 Comparison of the Deduced Amino Acid Sequence of cMeGluc2 with β-1,3-glucanases in Other Plants

The deduced amino acid sequence upstream of 468bp in cMeGluc2 had no similarity to any protein sequences in the database of NCBI (BLASTp, 1998). The rest of the sequence showed moderate identity to most of the β-1,3-glucanases in other plants (Tab.6.1). The sequence had slightly higher identity and similarity to the basic β-1,3-glucanases in *Citrus sinensis*, *Prunus persica* and tomato, and to the acidic β-1,3-glucanases in tomato. Most interestingly cMeGluc2 had high similarity to the elicitor-releasing β-1,3-glucanase in soybean (Takeuchi *et al.*, 1990).

Table 6.1 Comparison of deduced amino acid sequences between cMeGluc2 and β-1,3-glucanases in other plants.

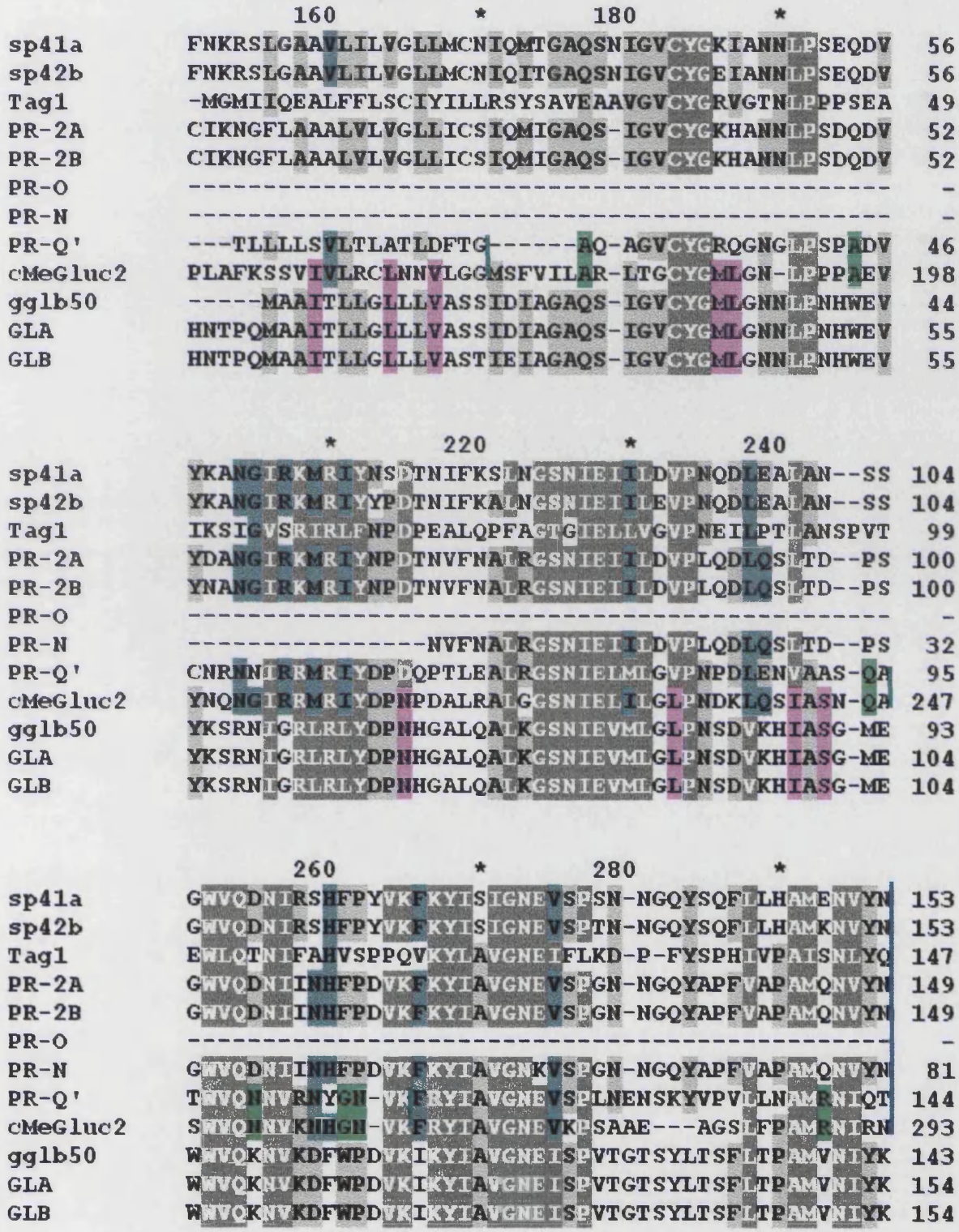
Sources	I% (S%)	No. of AA	Accession No.
Potato-basic	55 (71)	334	P52401
Tomato-basic	54 (71)	334	Q01413
Tomato-acidic	59 (76)	324	S44364
Soybean-1*	63 (79)	318	Q03773
Bean	53 (70)	318	P23535
Kidney bean	53 (69)	318	S13323
C. sinensis-b	67 (81)	314	AJ000081
Rubber	57 (73)	313	P52407
Garden pea	54 (70)	313	S28430
P. persica-b	65 (81)	309	P52408
Tomato-basic	62 (78)	309	S44365
Soybean-2	61 (79)	309	U41323
G.hirsutum	60 (74)	309	Z68154
Maize-PRm 6b	55 (70)	309	S82315
N.plumbaginif-olia	62 (76)	308	M63634
Potato-basic	58 (73)	308	P52400
Barley GII	52 (69)	308	P15737
Potato-basic 3	57 (72)	302	P52402
A. thaliana	63 (75)	284	M58464
A.thaliana-PR2	63 (75)	282	P33157
Sugar beet	54 (68)	344	A23400
Wheat	35 (57)	367	U30323
Yeast	15 (41)	485	X59259

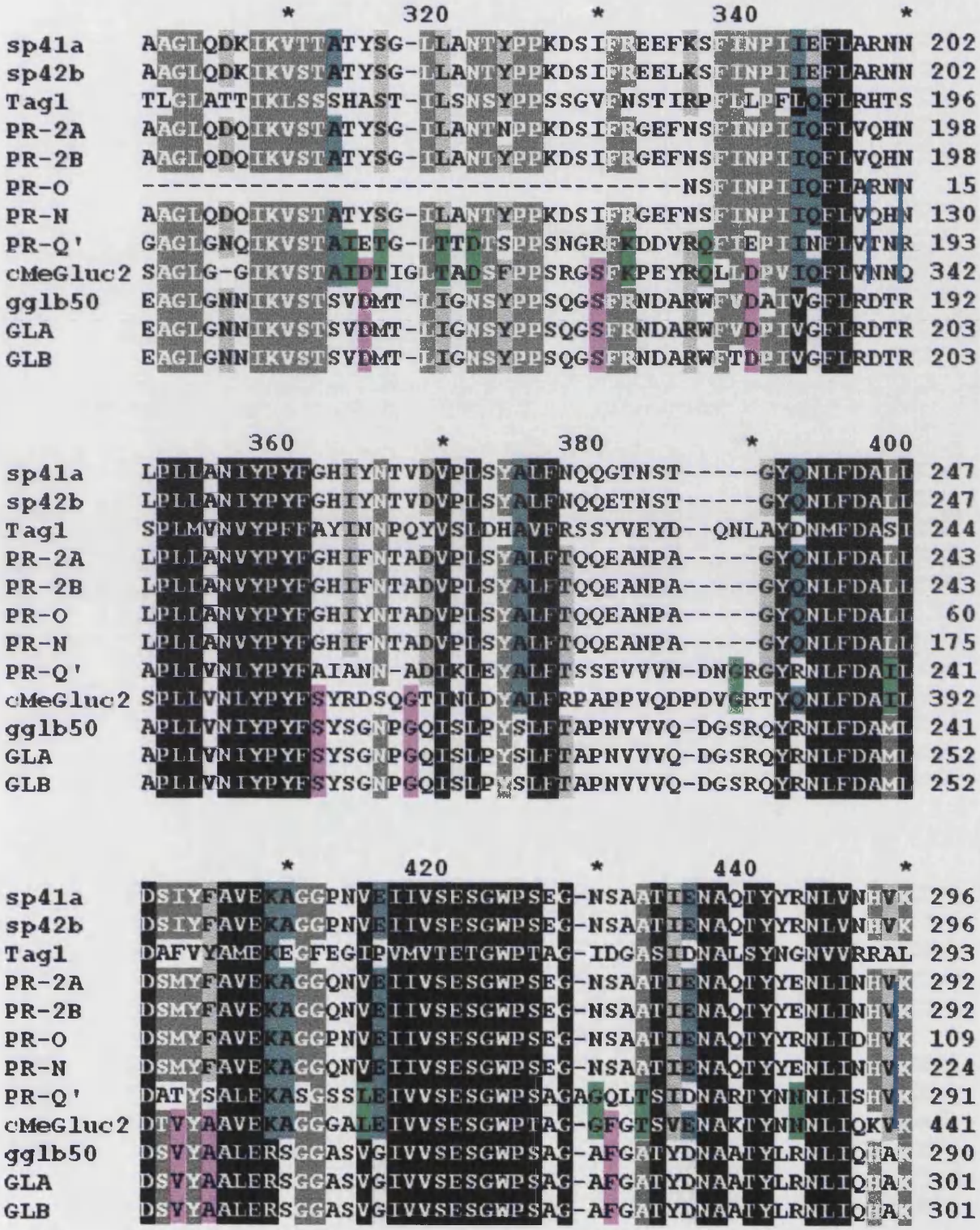
* elicitor-releasing factor

I refers to identical residue percentage

S refers to similar residue or conservative substitution percentage

Aligning the deduced amino acid sequence of cMeGluc2 with tobacco β-1,3-glucanases revealed that cMeGluc2 was of higher identity to acidic than to basic β-1,3-glucanases (Fig.6.4). Besides the residues conserved in all sequences, in cMeGluc2 peptide there were more residues conserved specifically between acidic (30 residues) than basic (18 residues) β-1,3-glucanases. In addition to the residues specifically conserved with acidic β-1,3-glucanases, the cMeGluc2 peptide had 18 residues specifically conserved between just PR-Q'. Comparing the deduced amino acid sequence of cMeGLuc2 with individual tobacco β-1,3-glucanase (Fig.6.4) showed that cMeGluc2 had 62% identity and 77% similarity with PR-Q', 58% identity and 73% similarity with all basic β-1,3-glucanases over 308 or 309 amino acid residues. Though cMeGluc2 had lower similarity (67 - 69%) to the other acidic β-1,3-glucanases, the similarity was over a longer region (328-335). Similar to tobacco acidic β-1,3-glucanases, the deduced amino acid sequence of cMeGluc2 lacks the C-terminal vacuolar sorting signal which is usually in basic β-1,3-glucanases and chitinases of tobacco (Fig.6.4) (Payne *et al.*, 1990).


160 * 180 *
sp41a FNKRSLGAAVLILVGLLMCNIQMTGAQSNIGVCYGKIANNLPSEQDV 56
sp42b FNKRSLGAAVLILVGLLMCNIQITGAQSNIGVCYGEIANNLPSEQDV 56
Tag1 -MGMIIQEALFFLSCIYILLRSYSAVEAAVGVCYGRVGTNLPPPSEA 49
PR-2A CIKNGFLAAALVLVGLLICSIQMIGAQS-IGVCYGKHANNLPSDQDV 52
PR-2B CIKNGFLAAALVLVGLLICSIQMIGAQS-IGVCYGKHANNLPSDQDV 52
PR-O ---------- -
PR-N ---------- -
PR-Q' ---TLLLLSVLTLATLDFTG------AQ-AGVCYGRQGNGLPSPADV 46
cMeGluc2 PLAFKSSVIVLRCLNNVLGGMSFVILAR-LTGCYGMLGN-LPPPAEV 198
gglb50 -----MAAITTLLGLLLVASSIDIAGAQS-IGVCYGMLGNNLPNHWEV 44
GLA HNTPQMAAITTLLGLLLVASSIDIAGAQS-IGVCYGMLGNNLPNHWEV 55
GLB HNTPQMAAITTLLGLLLVASTIEIAGAQS-IGVCYGMLGNNLPNHWEV 55
* 220 * 240
sp41a YKANGIRKMRIYNSDTNIFKSLNGSNIEIILDVPNQDLEALAN--SS 104
sp42b YKANGIRKMRIYYPDTNIFKALNGSNIEIILEVPNQDLEALAN--SS 104
Tag1 IKSIGVSRIRLFNPDPEALQPFAGTGIELLVGVPNEILPTLANSPVT 99
PR-2A YDANGIRKMRIYNPDTNVFNALRGSNIEIILDVPLQDLQSLTD--PS 100
PR-2B YNANGIRKMRIYNPDTNVFNALRGSNIEIILDVPLQDLQSLTD--PS 100
PR-O ---------- -
PR-N ----------NVFNALRGSNIEIILDVPLQDLQSLTD--PS 32
PR-Q' CNRNNIRRMRIYDPDQPTLEALRGSNIELMLGVPNPDLENVAAS-QA 95
cMeGluc2 YNQNGIRRMRIYDPNPDALRALGGSNIELILGLPNDKLQSIASN-QA 247
gglb50 YKSRNIGRLRLYDPNHGALQALKGSNIEVMLGLPNSDVKHIASG-ME 93
GLA YKSRNIGRLRLYDPNHGALQALKGSNIEVMLGLPNSDVKHIASG-ME 104
GLB YKSRNIGRLRLYDPNHGALQALKGSNIEVMLGLPNSDVKHIASG-ME 104
260 * 280 *
sp41a GWVQDNIRSHFPYVKFKYISIGNEVSPSN-NGQYSQFLLHAMENVYN 153
sp42b GWVQDNIRSHFPYVKFKYISIGNEVSPTN-NGQYSQFLLHAMKNVYN 153
Tag1 EWLQTNIFAHVSPPQVKYLAVGNEIFLKD-P-FYSPHIVPAISNLYQ 147
PR-2A GWVQDNIINHFPDVKFKYIAVGNEVSPGN-NGQYAPFVAPAMQNVYN 149
PR-2B GWVQDNIINHFPDVKFKYIAVGNEVSPGN-NGQYAPFVAPAMQNVYN 149
PR-O ---------- -
PR-N GWVQDNIINHFPDVKFKYIAVGNKVSPGN-NGQYAPFVAPAMQNVYN 81
PR-Q' TWVQNNVRNYGN-VKFRYIAVGNEVSPLNENSKYVPVLLNAMRNIQT 144
cMeGluc2 SWVQHNVKHHGH-VKFRYIAVGNEVKPSAAE---AGSLFPAMRNIRN 293
gglb50 WWVQKNVKDFWPDVKIKYIAVGNEISPVTGTSYLTSFLTPAMVNIYK 143
GLA WWVQKNVKDFWPDVKIKYIAVGNEISPVTGTSYLTSFLTPAMVNIYK 154
GLB WWVQKNVKDFWPDVKIKYIAVGNEISPVTGTSYLTSFLTPAMVNIYK 154


* 320 * 340 *
sp41a AAGLQDKIKVTTATYSG-LLANTYPPKDSIFREEFKSFINPIIEFLARNN 202
sp42b AAGLQDKIKVSTATYSG-LLANTYPPKDSIFREELKSFINPIIEFLARNN 202
Tag1 TLGLATTIKLSSSHAST-ILSNSYPPSSGVFNSTIRPFLLPFLQFLRHTS 196
PR-2A AAGLQDQIKVSTATYSG-ILANTNPPKDSIFRGEFNSFINPIIQFLVQHN 198
PR-2B AAGLQDQIKVSTATYSG-ILANTYPPKDSIFRGEFNSFINPIIQFLVQHN 198
PR-O ----------NSFINPIIQFLARNN 15
PR-N AAGLQDQIKVSTATYSG-ILANTYPPKDSIFRGEFNSFINPIIQFLVQHN 130
PR-Q' GAGLGNQIKVSTAIETG-LTTDTSPPSNGRFKDDVRQFIEPIINFLVTNR 193
cMeGluc2 SAGLG-GIKVSTAIDTIGLTADSFPPSRGSFKPEYRQLLDPVIQFLVNNQ 342
gglb50 EAGLGNNIKVSTSVDMT-LIGNSYPPSQGSFRNDARWFVDAIVGFLRDTR 192
GLA EAGLGNNIKVSTSVDMT-LIGNSYPPSQGSFRNDARWFVDPIVGFLRDTR 203
GLB EAGLGNNIKVSTSVDMT-LIGNSYPPSQGSFRNDARWFTDPIVGFLRDTR 203
360 * 380 * 400
sp41a LPLLANIYPYFGHIYNTVDVPLSYALFNQQGTNST-----GYQNLFDALL 247
sp42b LPLLANIYPYFGHIYNTVDVPLSYALFNQQETNST-----GYQNLFDALL 247
Tag1 SPLMVNVYPFFAYINNPQYVSLDHAVFRSSYVEYD--QNLAYDNMFDASI 244
PR-2A LPLLANVYPYFGHIFNTADVPLSYALFTQQEANPA-----GYQNLFDALL 243
PR-2B LPLLANVYPYFGHIFNTADVPLSYALFTQQEANPA-----GYQNLFDALL 243
PR-O LPLLANVYPYFGHIYNTADVPLSYALFTQQEANPA-----GYQNLFDALL 60
PR-N LPLLANVYPYFGHIFNTADVPLSYALFTQQEANPA-----GYQNLFDALL 175
PR-Q' APLIVNLYPYFAIANN-ADIKLEYALFTSSEVVVN-DNGRGYRNLFDAIL 241
cMeGluc2 SPLLVNLYPYFSYRDSQGTINLDYALFRPAPPVQDPDVGRTYQNLFDAIL 392
gglb50 APLLVNIYPYFSYSGNPGQISLPYSLFTAPNVVVQ-DGSRQYRNLFDAML 241
GLA APLLVNIYPYFSYSGNPGQISLPYSLFTAPNVVVQ-DGSRQYRNLFDAML 252
GLB APLIVNIYPYFSYSGNPGQISLPYSLFTAPNVVVQ-DGSRQYRNLFDAML 252
* 420 * 440 *
sp41a DSIYFAVEKAGGPNVELIVSESGWPSEG-NSAATIENAQTYYRNLVNHVK 296
sp42b DSIYFAVEKAGGPNVELIVSESGWPSEG-NSAATIENAQTYYRNLVNHVK 296
Tag1 DAFVYAMEKEGFEGIPVMVTETGWPTAG-IDGASIDNALSYNGNVVRRAL 293
PR-2A DSMYFAVEKAGGQNVELIVSESGWPSEG-NSAATIENAQTYYENLINHVK 292
PR-2B DSMYFAVEKAGGQNVELIVSESGWPSEG-NSAATIENAQTYYENLINHVK 292
PR-O DSMYFAVEKAGGPNVELIVSESGWPSEG-NSAATIENAQTYYRNLIDHVK 109
PR-N DSMYFAVEKAGGQNVELIVSESGWPSEG-NSAATIENAQTYYENLINHVK 224
PR-Q' DATYSALEKASGSSLEIVVSESGWPSAGAGQLTSIDNARTYNNNLISHVK 291
cMeGluc2 DTVYAAVEKAGGGALEIVVSESGWPTAG-GFGTSVENAKTYNNNLIQKVK 441
gglb50 DSVYAALERSGGASVGIVVSESGWPSAG-AFGATYDNAATYLRNLIQHAK 290
GLA DSVYAALERSGGASVGIVVSESGWPSAG-AFGATYDNAATYLRNLIQHAK 301
GLB DSVYAALERSGGASVGIVVSESGWPSAG-AFGATYDNAATYLRNLIQHAK 301

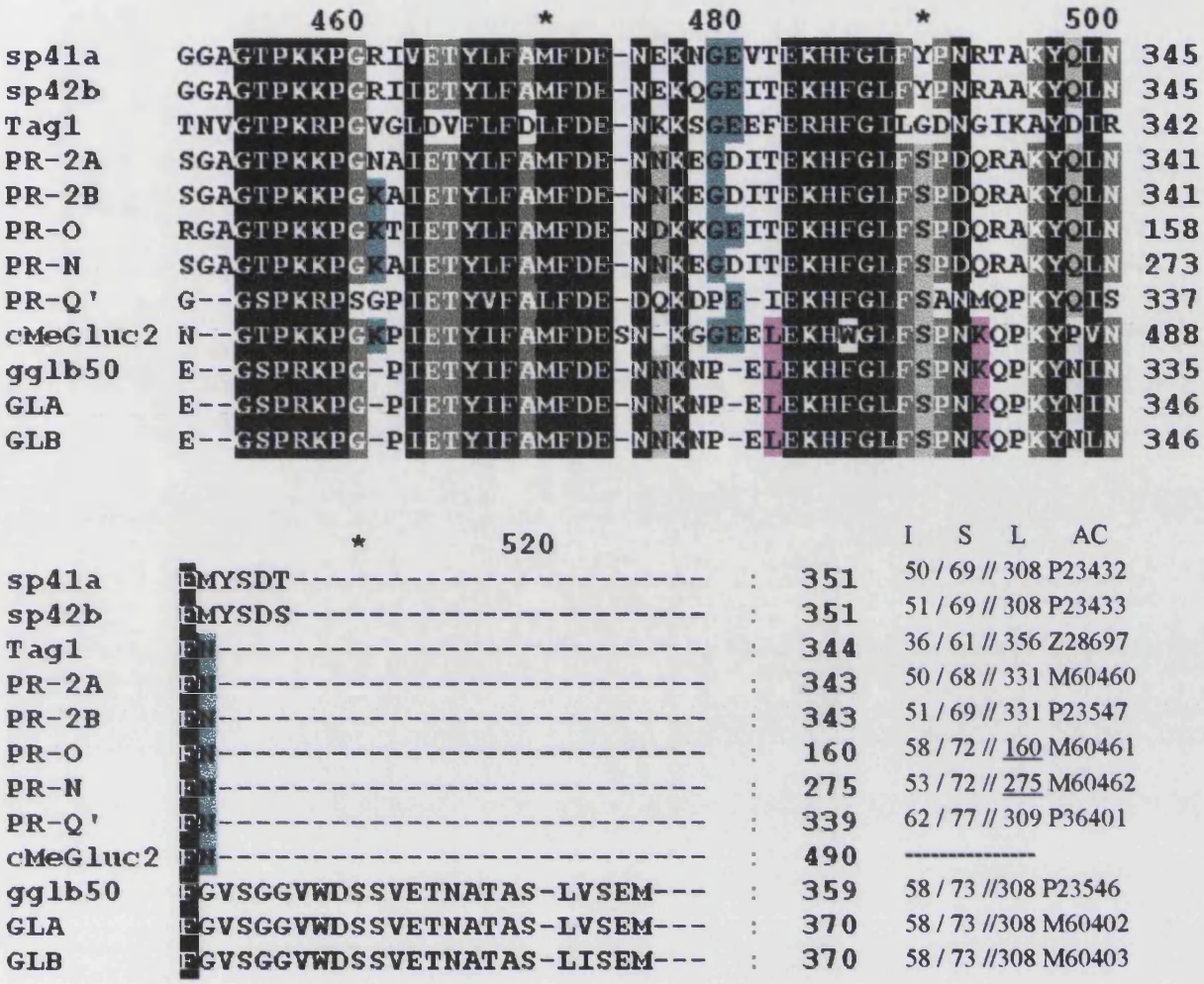


Fig.6.3 Multiple alignment of the deduced amino acid sequences of cMeGluc2 and tobacco β-1,3-glucanase genes. Residues conserved among all 12 sequences were in black shade; less conserved residues in dark-grey shade; residues conserved among less than 7 sequences were in light-grey shade. Residues that were similar in physiochemical properties were also considered conserved. About 127residues were conserved among all the sequences, most of which were located in the second half of the peptide. Specifically conserved residues between PR-Q' and cMeGluc2 were in green shade. Specifically conserved residues among basic glucanases and cMeGluc2 were in purple shade. Specifically conserved residues among most acidic glucanases and cMeGLuc2 were in green-blue shade.

The table by the end of the multiple alignment showed the identity (I)% and similarity (s)% between cMeGluc2 and tobacco β-1,3-glucanases in amino acid sequence, length of the polypeptide compared and the accession number of these genes.

6.2.4 β-1,3-glucanase gene organization in cassava.

Using the cMeGluc1 as a probe, cassava genomic DNA Southern blot analysis was performed to estimate the numbers of β-1,3-glucanase genes present in cassava. Restricted DNA 10μg was run in 1XTBE agarose (1%) gel and transferred to Hybond N+ nylon membrane by standard Southern blotting procedures. Prehybrization was performed for 2 hours at 60°C in pre-hybridization buffer (the same as hybridization buffer), followed with hybridization at 60°C in 6xSSC, 5xDenhardt's solution, 1%SDS and 0.1mg/ml denatured and sonicated herring sperm DNA for 20 to 24 hours.

After a low stringency wash, there were 3, 4, 2 and 6 bands in lanes of DNA cut with *Eco*RI, *Xba*I, *Hin*dIII and *Bgl*II respectively (Fig.6.5A). Whereas, after high stringency wash, there was only one prominent band in the first lane and the bands in the other lanes were weak, which may have been due to the extended wash at high stringency (Fig.6.5B).

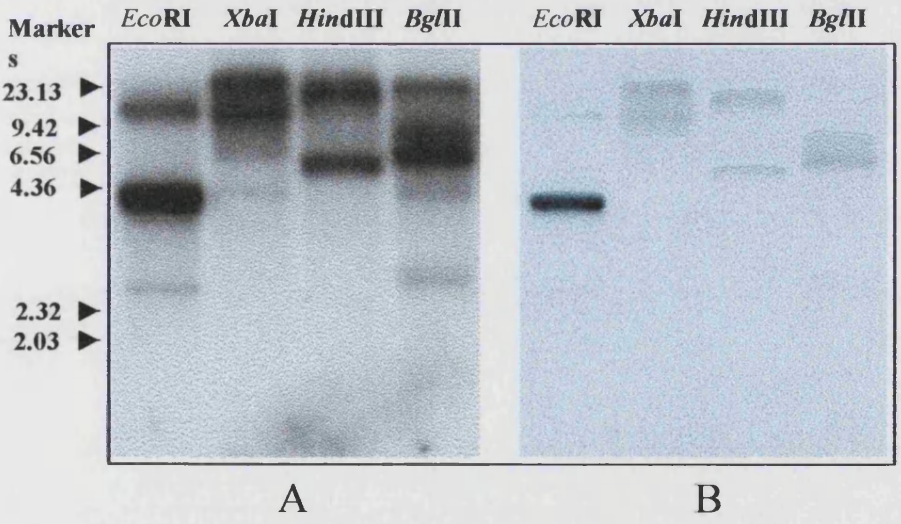


Fig.6.5 β-1,3-glucanase gene organisation in cassava. About 10μg of genomic DNA was digested with restriction enzymes. After hybridization with the probe made from cMeGluc1 at 60°C overnight, the membrane was washed at low or high stringency. **A:** low stringency 60°C, 1xSSC, 0.1%SDS, 2x10min; **B:** high stringency 60°C, 0.1xSSC, 0.1%SDS, 2x20min. The figures on the left side indicate the sizes of DNA markers in kb.

6.2.5 Expression of β-1,3-glucanases

The isolation of β-1,3-glucanase genes from PPD-related cDNA library indicated that these genes were expressed during PPD. In order to understand their roles in the wound response, the expression of cMeGluc1 and 2 in leaves after wounding and in storage roots during PPD was studied by Northern blot hybridization. The probes used was made from cMeGluc1, which would bind to mRNAs of cMeGluc1 & 2 as there was only one base difference between these two genes over the region covered by cMeGluc1.

No β-1,3-glucanase transcript was detected in normal young leaves or in leaves 2hrs after wounding. During the development of PPD, β-1,3-glucanase MeGluc1 & 2 were induced in the storage root slices but not before 72hrs post-harvest (Fig.6.6).

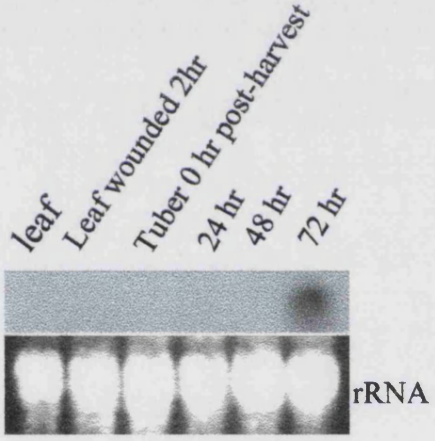


Fig.6.6 Northern blot analysis of β-1,3-glucanase in cassava. 25μg of total RNA for each sample and ribosome RNA was checked after Northern transfer for the equal loading and transfer of samples. Cassava leaves were wounded by cutting into 0.5x0.5cm pieces and incubating in a petri-dish with pre-wet filter paper in the green house. *Eco*RI fragment of cMeGLuc1was used as a probe. Hybridization at 42°C in standard SSPE/formamide buffer overnight and final wash with 0.1xSSC/0.1%SDS/2x10min at 68 °C.

6.3 Discussion

Eight putative beta-1,3-glucanase cDNA clones were isolated from a PPD-related cDNA library. Two of them, cMeGluc1 & 2 were further characterised. The PCR products of clone G and H had weak signals when hybridized with heterologous β-1,3-glucanase probes (Fig.6.1b), which indicated that they might be different from the other clones.

The gene family of β-1,3-glucanase in plants appears to be complicated. In potato, more than 10 bands with different densities were detected in genomic Southern hybridisation with a potato 1,3-glucanase probe, which suggested considerable complexity (Beerhues and Kombrink, 1994). In tomato, four β-1,3-glucanase cDNA clones have been isolated (Domingo *et al.*, 1994). At least ten β-1,3-glucanase genes have been isolated in tobacco as shown in 6.2.3. In cassava, the genomic DNA Southern blot indicated that there was a relatively small gene family of β-1,3-glucanases, which is similar to the possible size of the gene family of β-1,3-glucanase in rubber tree, in the same family (Euphorbiaceae) as cassava (Chye and Cheung, 1995). Under high stringency wash conditions, the band in *Eco*RI lane was very dense but the two bands in either *Xba*I or *Hin*dIII lane were faint. Based on the fact that cMeGluc1 & 2 were of extremely high identity in the probe region, it could be predicted that the strong band in *Eco*RI lane was due to the clustering of two genes in one *Eco*RI genomic fragment or to coincidental location in different genomic *Eco*RI fragments of the same size. If the genomic DNA was not digested with *Eco*RI but with *Xba*I or *Hin*dIII, the two similar genes were separated on different genomic fragments and gave signals of equal density.

The extremely high identity (one base difference) between cMeGluc1 and 2 in known sequences has not been described among glucanase genes in a gene family in other plant systems. These two clones could be from one mRNA and the difference may be due to the error caused during the cloning process. This could be tested by checking the sequence of the genomic DNA corresponding to the cDNAs. If it is true that these two clones are the same, there may be two copies of this gene in cassava genome, as indicated by the Southern blot hybridization followed high stringency wash.

Comparison of the deduced amino acid sequences of cMeGluc1 with tobacco and other plant β-1,3-glucanases revealed a high level of homology. It has 127 amino acid residues that were conserved among all classes of glucanases, and especially significant, it shared the same residues that are thought to be important in carbohydrate binding and glucan hydrolysis (Chen *et al.*, 1993). The (putative) catalytic sites, which were identified by amino acid sequence alignment and amino acid residue substitution, were not only highly conserved but also similarly positioned, and were surrounded by highly conserved amino acid residues. These indicate conservation in enzyme structure and activity, suggesting that the deduced protein is a β-1,3-glucanase.

The deduced peptide of cMeGluc2 had moderate identity to most of β-1,3-glucanases in other plants and higher similarity to both basic and acidic β-1,3-glucanases of some plants, including the elicitor-releasing glucanase (acidic) in soybean (Takeuchi *et al.*, 1990). The deduced amino acid sequences of cMeGluc2 had higher identity and similarity to acidic than to basic tobacco β-1,3-glucanases. Compared to other acidic β-1,3-glucanases in tobacco cMeGluc2 shared more residues with PR-Q', a possible tobacco analogue of the elicitor-releasing glucanase (acidic) from soybean based on sequence similarity and phylogenetic analysis (Payne *et al.*, 1990; Bucciaglia *et al.*, 1994). These analyses suggested that cMeGluc2 might fall in the category of PR-2 (β-1,3-glucanase) proteins.

The cMeGluc2 possesses an unusually long coding region with at least 156 amino acid residues at the N-terminal in addition when compared to other plant β-1,3-glucanases. It shared high similarity with the amino acid sequences of both basic and acidic β-1,3-glucanase of some other plants but its deduced peptide was much closer to acidic β-1,3-glucanase PR-Q' from tobacco glucanases.

The deduced amino acid sequence of cMeGluc2 showed high similarity to extracellular acidic tobacco β-1,3-glucanases (or class II and III) and lacks C-terminal vacuolar sorting signal sequence. Therefore, the cMeGluc2 glucanase may be extracellularly located. The tobacco class II such as PR-2 and class III β-1,3-glucanase PR-Q' were induced 2 or 3 days after TMV infection and the expression of these two β-1,3-glucanase genes was coordinately regulated during the interaction between TMV and tobacco (Ward *at al.*,

1991). The constitutive expression of PR-N cDNA for β-1,3-glucanase in tobacco was claimed to increase the resistance of the transgenic plants to glucan-containing fungi *P. parasitica var. nicotianae* and *P. tabacina*, though the resistance was not enhanced to viruses (Lusso and Kuc, 1996). The expression of antisense PR-N gene reduced the activity of β-1,3-glucanase and decreased the resistance to these pathogens (Lusso and Kuc, 1996). These results illustrate the importance of β-1,3-glucanases in the resistance of plants against fungal pathogens. The extracelliular location of β-1,3-glucanases may suggest their defence function early in the infection process as they might have a direct fungicidal action on the hyphae invading the extracellular space (Mauch and Staehelin, 1989). The similarity of amino acid sequences between PR-Q' and soybean elicitor-releasing β-1,3-glucanase might indicate their similar roles in plant defence against glucan-containing fungal pathogens (Payne *et al.*, 1990). The soybean β-1,3-glucanase from soybean cotyledons was shown to release elicitors from the cell walls of *Phytophthora megasperma f.sp. glycinea* which elicited synthesis of the phytoalexin glyceollin (Keen and Yoshikawa, 1983). The expression of the soybean glucanase gene in tobacco increased the resistance of the transgenic plants against *P. parasitica var. nicotianae* and *A. alternata*. The increased resistance was thought to involve the release of active elicitor molecules and not the direct action of the hydrolase on the fungus, as it was found that the accumulation of PAL mRNAs occurred earlier in the transgenic plants than in the control plants and the β-1,3-glucanase was not toxic to the fungi (Yoshikawa *et al.*, 1993). It is interesting to note that deduced amino acid sequences of cMeGluc2 had high similarity to the soybean β-1,3-glucanase, and especially that cMeGluc2 shares more amino acid residues with PR-Q' than other tobacco β-1,3-glucanases, which may indicate its possible function in cassava.

In the rubber tree, from the same family as cassava, a basic β-1,3-glucanase gene was highly expressed in laticifers. The glucanase gene was expressed much more abundantly in stems than in leaves reflecting the presence of more laticifers in the stems (Chye and Cheung, 1995). In cassava, laticifers are also abundant in stems, leaves and both vegetative and storage roots. It would be interesting to test whether cMeGluc1 & 2 are expressed in laticifers and also whether there is an analogue of the rubber glucanase in cassava.

The expression of β-1,3-glucanases in roots or upon mechanical wounding has not been well studied. A vacuolar basic β-1,3-glucanase gene was expressed at high levels in potato roots (Beerhues and Kombrink, 1994). The gene was expressed at a very low level in freshly harvested potato tubers and it was not detectable in stored tubers, though it was abundantly expressed in stems. In the storage root of cassava (cultivar M col 22) cMeGLuc1 & 2 were not detected until 72hr after wounding. The induction of these glucanase genes was extremely slow compared to 3 to 6 hrs needed for the induction of the basic β-1,3-glucanase by elicitor, ethylene, *Phytophthora infestans* and wounding in potato leaves (Beerhues and Kombrink, 1994). Was the late induction of β-1,3-glucanase transcripts linked with the high susceptibility of the cultivar to PPD? The isolation of eight positive β-1,3-glucanase clones (see 6.2.1) from the PPD-related cDNA library made from storage root (cultivar M NGA1) 24-48 hr after harvest indicated that these genes were expressed. The medium susceptibility of cultivar M NGA1 may be linked with its more rapid induction of glucanase genes, in contrast to the late induction of these genes in susceptible cultivar M col22. Expression patterns of β-1,3-glucanase genes in storage roots and wounded leaves from cultivars of different susceptibility to PPD needs to be investigated. The role of β-1,3-glucanases during PPD could also be investigated by over-expressing β-1,3-glucanase genes or down-regulating them through cassava transformation. If a link between resistance to PPD and early abundant expression of β-1,3-glucanase genes could be established, β-1,3-glucanase genes may be used as a marker for selecting or breeding PPD resistant cultivars.

Most of the β-1,3-glucanase genes in plants can be induced by exogenous ethylene. The expression of the elicitor-releasing β-1,3-glucanase in soybean could be enhanced 50 to 100 fold within 6 to 12hrs after application of exogenous ethylene (Takeuchi *et al.*, 1990). However, application of ethylene to cassava storage roots did not affect the development of PPD (Hirose *et al.*, 1984). It would be interesting to test how β-1,3-glucanase genes would respond to exogenous ethylene, which might provide some insight into the signal transduction during PPD.

Chapter Seven

Identification of cDNA Clones of Non-preconceived Target Genes and an ACC Oxidase Gene

During the isolation and cloning of target cDNAs, cDNAs of non-preconceived target genes were also identified. When the PPD-related cDNA library was screened for other possible PAL cDNA clones at low stringency, two false positive clones were isolated. The deduced amino acid sequences of one of these two clones were of high similarity to ubiquitin-activating enzyme in *Arabidopsis* and wheat; the other had high similarity to serine/threonine kinase protein or receptor-like kinase. Clone cMeGARs was identified during the cloning of cMePAL3, as the cDNAs formed a chimeric cDNA clone. The antisense sequence of the 5' end of cMeHRGP1 was found to be of high identity to phosphoserine aminotransferase gene in *Arabidopsis* and spinach.

An ACC oxidase cDNA was isolated from the PPD-related cDNA library using tomato ACC oxidase cDNA pTOM13 as a probe. The clone was preliminarily analysed here.

7.1 A cDNA Encoding a Putative Ubiquitin-activating Enzyme

Ubiquitin is a small and abundant protein of almost all eukaryote cells. Its functions are mediated via its linkage to target proteins. A number of essential processes in eucaryotes, such as elimination of abnormal proteins (Ciechanover *et al*, 1984; Driscoll & Finley, 1992) and the control of the turnover of important regulatory proteins, are regulated by selective protein breakdown via the ubiquitin-dependent proteolytic pathway (Vierstra, 1993; Ciechanover and Schwartz, 1994). The known targets include transcriptional regulators (Hochstrasser, 1991), plant phytochrome (Shanklin *et al*, 1987), tumour suppressor p53 (Scheffner, 1990) and cyclins (Glotzer *et al*, 1991). The selectivity is mediated through the ubiquitin-conjugation system. The attachment or conjugation of ubiquitin to target proteins is carried out in a chain of three enzymatic reactions. First, ubiquitin is activated through an ATP-mediated transfer of the C-teminus of ubiquitin to a thiol group of an internal cysteine residue of the ubiquitin-activating enzyme (E1) (Hass and Rose, 1982). Activated ubiquitin is then transferred to a specific cysteine residue of one

of several ubiquitin-conjugating enzymes (E2 or ubiquitin carrier proteins) by transthiolation. Finally, ubiquitin is covalently attached to a target protein by an isopeptide bond between the C-terminal glycine residue of ubiquitin and a specific internal lysine residue of the target protein or by ubiquitin ligase or E3. The final breakdown is done by proteasome. The plant proteasome has been partially characterised and cloned (Genschik *et al*, 1994). A plant 26S proteasome isolated from spinach showed very high ultrastructural similarity to mammalian 26S proteasome (Fujinami *et al*, 1994).

There are only very few substrates for ubiquitinylation known from plants. Plant regulatory photoreceptor phytochrome is one example (Shanklin *et al*, 1987). Red light was found to be the necessary signal for ubiquitin conjugation as only the P_{fr}-form was ubiquitin-dependently degraded.

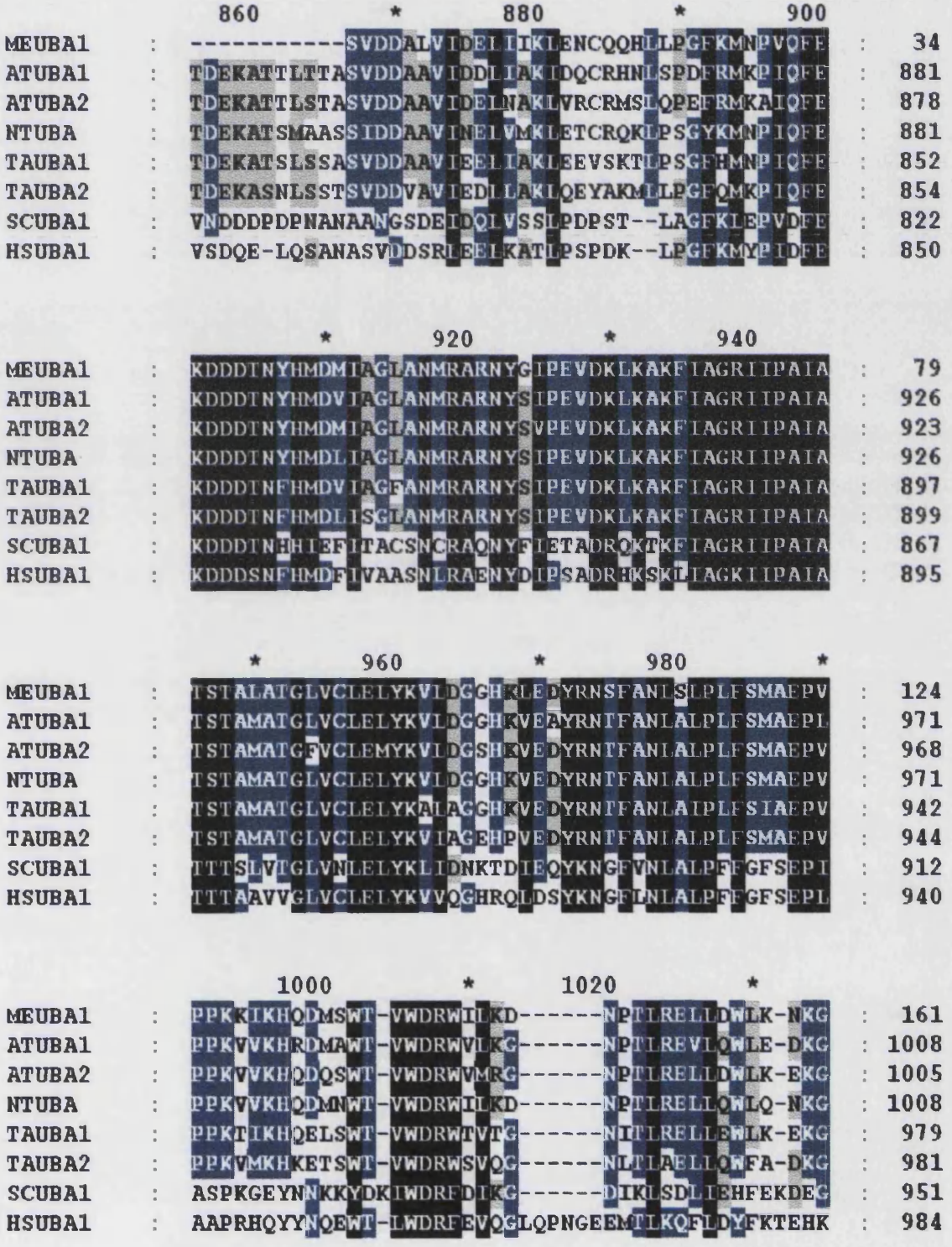
There seems to be an association between ubiqutinylation and stress response in plants and involvement of ubiquitin in plant and pathogen interaction. It was demonstrated in plants (Callis, 1995) that heat-denatured abnormal proteins triggered the activation of polyubiquitin genes, resulting in a burst of ubiquitin-dependent proteolysis to cope with the denatured proteins. Upon chilling stress, a remarkable increase in the amount of polyubiquitin mRNA was observed in *Chlamydomonas reinhardtii* (von Kampen and Wettern, 1995) but no changes were reported in the level of ubiquitin-encoding mRNAs in barley, tomato and rice (Cattivelli and Bartels, 1990). Wounding and heavy metal were also found to regulate the expression of polyubiquitin genes (Garbarino *et al*, 1992; Genschick *et al*, 1992). The expression of an ubiquitin gene was induced *in Phytophthora infestans* during the colonization of potato leaves (Pieterse *et al*, 1991). Alternation of ubiquitin conjugate pattern in powdery mildew-infected barley was also observed during fungal development and sporulation (von Kampen *et al*, 1996). Transgenic tobacco plants expressing ubiquitin transgene showed either decreased virus replication in the case of sensitive plants or fewer but darker lesions in resistant plants, which suggested the involvement of ubiquitin system in the induction of the hypersensitive response and necrotic cell death (Becker *et al*, 1993).

Ubiquitin activating enzymes (E1) and genes have been characterised in a number of organisms. It has been reported that there are several E1 proteins and genes in animals (rabbits, mice and humans) (Ciechanover, 1982; Kay *et al.*, 1991; Handley *et al.*, 1991; Kudo *et al.*, 1991) and plants (wheat and *Arabidopsis*) (Hatfield and Vierstra, 1989 & 1992). In yeast only one essential E1 gene was demonstrated to exist (McGrath, 1991). *Arabidopsis* E1 genes AtUBA1 and 2 were expressed non-differentially in most of tissues and cells.

E1s catalyse the initial steps in the ubiquitin conjugation system, and therefore potentially play important role in the metabolic processes involving ubiquitin attachment.

A putative ubiquitin-activating enzyme cDNA clone (or cMeUba1) was identified from the cassava PPD-related cDNA library during the isolation of PAL cDNA clones. The cDNA was demonstrated to be a partial sequence of the gene by aligning its deduced amino acid sequences with those of wheat and *Arabidopsis thaliana* E1 genes (Fig.7.1). The E1 amino acid sequences among plants are very conserved. The identity and similarity in amino acid sequences between cassava E1 and other organisms are from 46% (67%) (human) to 80% (92%) (*Arabidopsis thaliana*).

The full length sequence could be obtained by 5'RACE. Genomic Southern blots could be carried out to estimate how many gene members in the E1 gene family. Other genes in the family may be isolated and the expression of these genes could be studied to understand their roles in cassava development and stress responses. If they are constitutively expressed as housekeeping genes, they could be used as controls in the expression study of other genes in cassava.


860 * 880 * 900
MEUBA1 : ----------SVDDALVIDELIIKLENCQQHLLPGFKMNPVQFE : 34
ATUBA1 : TDEKATTLTTASVDDAAVIDDLIAKIDQCRHNLSPDFRMKPIQFE : 881
ATUBA2 : TDEKATTLSTASVDDAAVIDELNAKLVRCRMSLQPEFRMKAIQFE : 878
NTUBA : TDEKATSMAASSIDDAAVINELVMKLETCRQKLPSGYKMNPIQFE : 881
TAUBA1 : TDEKATSLSSASVDDAAVIEELIAKLEEVSKTLPSGFHMNPIQFE : 852
TAUBA2 : TDEKASNLSSTSVDDVAVIEDLLAKLQEYAKMLLPGFQMKPIQFE : 854
SCUBA1 : VNDDDPDPNANAANGSDEIDQLVSSLPDPST--LAGFKLEPVDFE : 822
HSUBA1 : VSDQE-LQSANASVDDSRLEELKATLPSPDK--LPGFKMYPIDFE : 850
* 920 * 940
MEUBA1 : KDDDTNYHMDMIAGLANMRARNYGIPEVDKLKAKFIAGRIIPAIA : 79
ATUBA1 : KDDDTNYHMDVIAGLANMRARNYSIPEVDKLKAKFIAGRIIPAIA : 926
ATUBA2 : KDDDTNYHMDMIAGLANMRARNYSVPEVDKLKAKFIAGRIIPAIA : 923
NTUBA : KDDDTNYHMDLIAGLANMRARNYSIPEVDKLKAKFIAGRIIPAIA : 926
TAUBA1 : KDDDTNFHMDVIAGFANMRARNYSIPEVDKLKAKFIAGRIIPAIA : 897
TAUBA2 : KDDDTNFHMDLISGLANMRARNYSIPEVDKLKAKFIAGRIIPAIA : 899
SCUBA1 : KDDDTNHHIEFITACSNCRAQNYFIETADRQKTKFIAGRIIPAIA : 867
HSUBA1 : KDDDSNFHMDFIVAASNLRAENYDIPSADRHKSKLIAGKIIPAIA : 895
* 960 * 980 *
MEUBA1 : TSTALATGLVCLELYKVLDGGHKLEDYRNSFANLSLPLFSMAEPV : 124
ATUBA1 : TSTAMATGLVCLELYKVLDGGHKVEAYRNTFANLALPLFSMAEPL : 971
ATUBA2 : TSTAMATGFVCLEMYKVLDGSHKVEDYRNTFANLALPLFSMAEPV : 968
NTUBA : TSTAMATGLVCLELYKVLDGGHKVEDYRNTFANLALPLFSMAEPV : 971
TAUBA1 : TSTAMATGLVCLELYKALAGGHKVEDYRNTFANLAIPLFSIAEPV : 942
TAUBA2 : TSTAMATGLVCLELYKVIAGEHPVEDYRNTFANLALPLFSMAEPV : 944
SCUBA1 : TTTSLVTGLVNLELYKLIDNKTDIEQYKNGFVNLALPFFGFSEPI : 912
HSUBA1 : TTTAAVVGLVCLELYKVVQGHRQLDSYKNGFLNLALPFFGFSEPL : 940
1000 * 1020 *
MEUBA1 : PPKKIKHQDMSWT-VWDRWILKD------NPTLRELLDWLK-NKG : 161
ATUBA1 : PPKVVKHRDMAWT-VWDRWVLKG------NPTLREVLQWLE-DKG : 1008
ATUBA2 : PPKVVKHQDQSWT-VWDRWVMRG------NPTLRELLDWLK-EKG : 1005
NTUBA : PPKVVKHQDMNWT-VWDRWILKD------NPTLRELLQWLQ-NKG : 1008
TAUBA1 : PPKTIKHQELSWT-VWDRWTVTG------NITLRELLEWLK-EKG : 979
TAUBA2 : PPKVMKHKETSWT-VWDRWSVQG------NLTLAELLQWFA-DKG : 981
SCUBA1 : ASPKGEYNNKKYDKIWDRFDIKG------DIKLSDLIEHFEKDEG : 951
HSUBA1 : AAPRHQYYNQEWT-LWDRFEVQGLQPNGEEMTLKQFLDYFKTEHK : 984

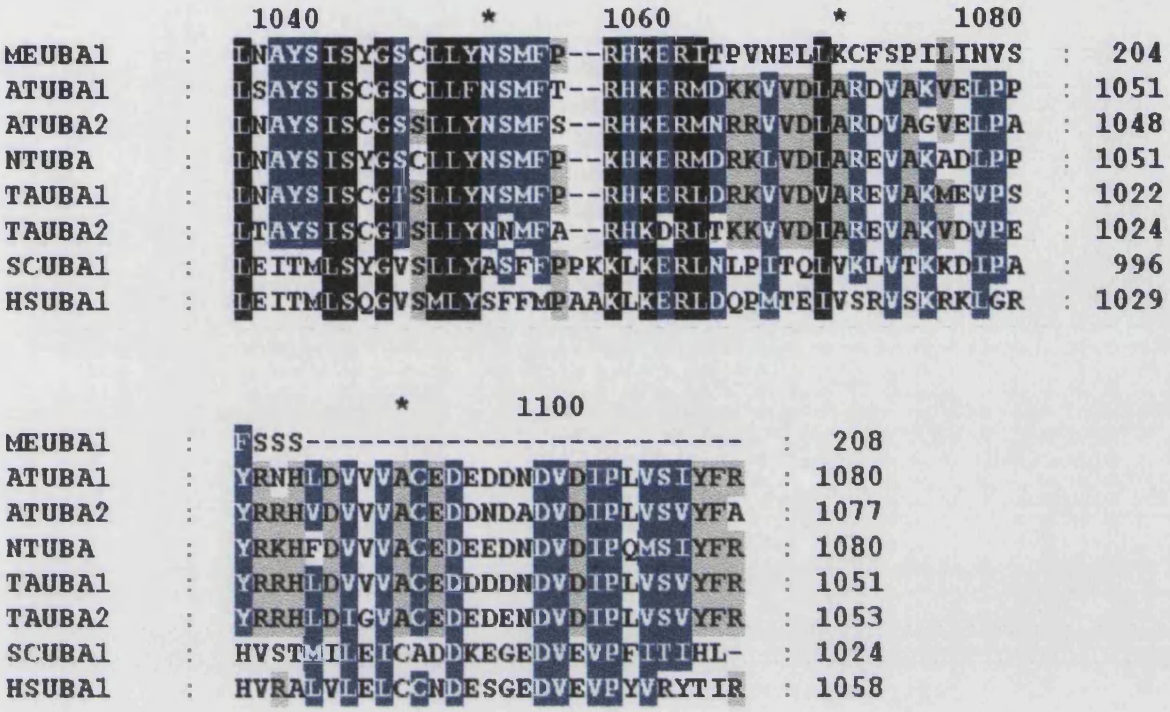


Fig.7.1 The multiple alignment of the deduced amino acid sequences of the genes encoding ubiquitin-activating enzymes in different organisms including cassava (MEUBA1), *Arabidopsis thaliana* (ATUBA1 & 2), Tobacco (NTUBA), wheat (TAUBA1 & 2), yeast (SCUBA1) and human (HSUBA1). The residues that were identical or similar among all the sequences compared are black-shaded. The residues that were identical or similar among no less than six sequences are blue-shaded. The residues that were identical or similar among five sequences are grey-shaded ones. The alignment before 860 is not shown as no nucleotide sequences of cassava E1were available in that region and the end of the alignment was the ends of the amino acid sequences for all the genes compared. The accession numbers of these genes are U80808 for ATUBA1, U40566 for ATUBA2, Y10804 for NTUBA, A38373 for TAUBA1, M90663 for TAUBA2, X55386 for SCUBA1 and Z28210 for HSUBA1.

7.2 A cDNA Clone with a Serine/threonine Kinase or Receptor-like Kinase Domain

Protein phosphorylation/de-phosphorylation is known to play crucial roles in signal transduction pathways in both prokaryotes and eukaryotes (Karin and Hunter, 1995). In plants, much effort has been put into the molecular analyses of protein kinases that may be involved in signal transduction pathways in recent years. More than one hundred protein kinase genes have been identified from plants. Receptor-like protein kinases (RLKs) are one class of these kinases and they appear to be serine/threonine protein kinases.

The RLKs are classified into three groups, S-domain kinases, kinases with leucine-rich repeats (LRR) and disease-resistance-related kinases. The S-domain class is characterized by an extracellular S-domain. The feature of the S-domain is an array of ten cysteine residues in combination with other conserved motifs, which was first described in the self-incompatibility-locus glycoproteins (SLGs) from Brassicaceae (Stein *et al*, 1991). S-domain RLKs have been identified in *Brassica* species, *Arabidopsis* and maize (Walker and Zhang, 1990; Stein *et al*, 1991; Tobias *et al*, 1992). The second RLK class includes kinases with leucine-rich repeats (LRR). LRR motif is found in many proteins that participate in protein-protein interactions. Rice *Xa21*, which confers resistance to *Xanthomonas oryzae* pv *oryzae* (xoo), is a typical RLK gene. A model for the resistance mechanism was proposed by Ronald (1997). There is a LRR domain in *Xa21*, which may bind a polypeptide produced by the pathogen (or plant cell) and the specific interaction may be mediated by a finite subset of amino acids in the LRR. Specific binding may lead to activation of *Xa21* kinase with subsequent phosphorylation on specific serine or threonine residues. Phosphorylated residues may then serve as binding sites for proteins that can initiate down stream responses (Ronald, 1997). Another group of kinases is the disease-resistance-related kinase such as Pto or herbicide-resistance-related kinase such as Fen. Unlike Pto (confers resistance to strains of the bacterial pathogen *Pseudomonas syringae* pathovar *tomato* that carry the avrPto gene), Fen is a serine/threonine kinase lacking the regulatory domain or the LRR domain.

During the screening for PAL clones from the PPD-related cDNA library, a false positive clone was identified (4.2.4) to be similar to serine/threonine kinase protein or receptor-like

kinase. The cDNA is 700bp in length, a partial sequence of the corresponding gene. The deduced amino acid sequence of the cDNA sequence contains residues (DFG) of serine/threonine kinase specificity and showed high similarity to the amino acid sequence of tomato *Fen* gene, tobacco serine/threonine kinase and rice *Xa21* gene (Fig. 7.2). The full length of the cDNA corresponding this gene needs to be cloned to know whether there is a regulatory domain LRR region, which may provide more information about the possible function of this gene.

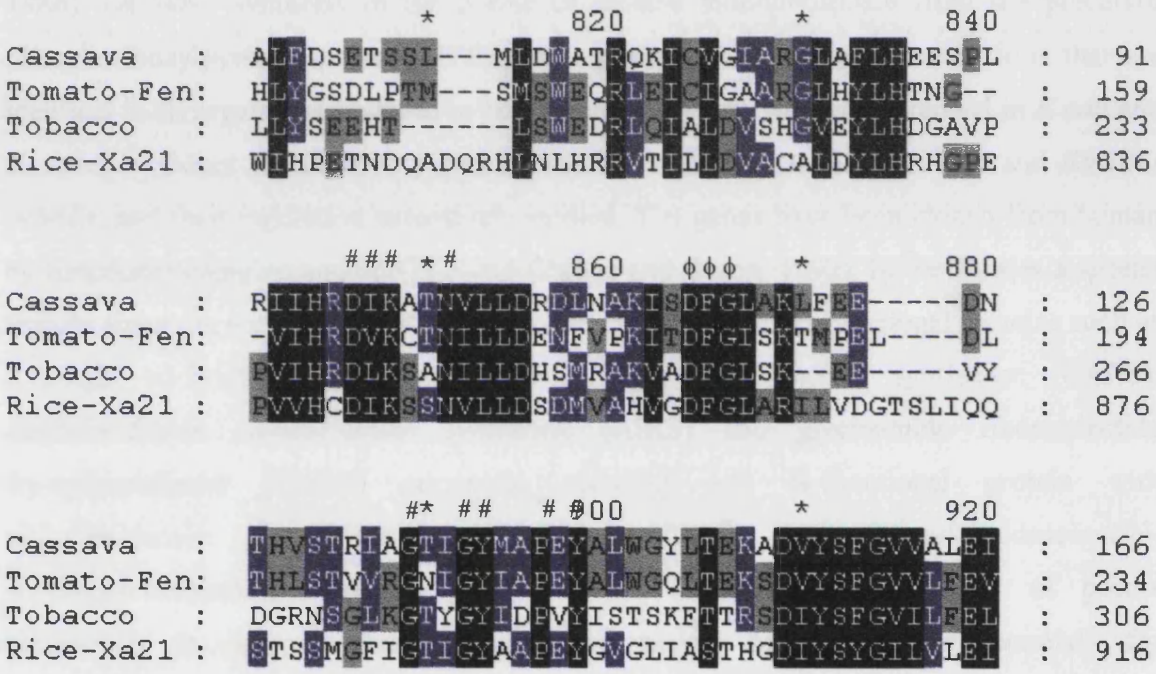


Fig.7.2 Multiple alignment of partial deduced amino acid sequences of serine-threonine protein kinase or receptor kinase-like protein genes. Accession number Xa21 / U59318, Fen / U59318 and tobacco serine-threonine kinase / D31737. # indicate residues of serine-threonine kinase specificity, ϕ conserved sub-domain VII in kinases (Hanks and Quinn, 1991).

7.3 A cDNA Encoding Glycinamide Ribonucleotide Synthetase (GARS)

Purines are required to produce DNA and RNA precursors for cell development and reproduction. In many plants, purine metabolism is a major component of the nitrogen assimilation pathway. In tropical legumes, fixed nitrogen goes through the purine biosynthesis pathway and produces inosine monophosphate (IMP). IMP is converted into uric acid and then ureides, which are transported throughout the plant. In other legumes, fixed nitrogen is converted into Gln and Asn, which are transported to different parts of the plant. The purine base is constructed from single nitrogen and carbon units from one-carbon metabolism, mainly from the amino acids Gly, Gln and Asp (Neuhard and Nygaard, 1986). *De novo* synthesis of the purine or inosine monophosphate from the precursor phosphoribosylpyrophosphate (PRPP) is completed *via* ten enzymatic reactions that are identical in all organisms analysed so far. The pathway was well-characterised in *E coli* and *B. subtilis.* Genes for *de novo* purine synthesis have been cloned in *E coli* and *Bacillus subtilis,* and their regulation extensively studied. The genes have been cloned from human by functional complementation in *E coli* (Zalkin and Dixon, 1992). In vertebrates a special feature about the pathway is that it is catalyzed by bi- or /and tri-functional proteins such as a single tri-functional polypeptide glycinamide ribonucleotide synthetase (GARS), aminoimidasole ribonucleotide synthetase (AIRS) and glycinamide ribonucleotide formyltransferase (GART) enzymatic activities and bi-functional protein with aminoimidasole ribonucleotide carboxylase (AIRC) and 5-aminoimidasole-4-N-succinocarboxyamide ribonucleotide synthetase (SAICARS). The study of purine metabolism in mammals has been intensive, mainly because of the potential for development of drugs against a variety of diseases and for detecting metabolic defects in *de novo* purine nucleotide synthesis (Zalkin and Dixon, 1992). In plants, the cDNAs encoding the first, second, third, fifth, sixth and seventh enzymes of the purine biosynthesis have been isolated and characterised. The cDNA encoding phosphoribosyl pyrophosphate amidotransferase (PRAT), the first enzyme, have been isolated both in soybean and *Vigna aconitifolia* (Kim *et al,* 1995). Three cDNAs encoding the second enzyme glycinamide ribonucleotide synthetase (GARS), the third enzyme aminoimidasole ribonucleotide synthetase (AIRS) and the fifth enzyme glycinamide ribonucleotide formyltransferase (GART) have been isolated from *A thaliana* (Schnorr *et al,* 1994; Senecoff and Meagher,

1993). The cDNAs of GARS and GART were also isolated from soybean (Schnorr *et al*, 1996). The cDNAs of the sixth enzyme AIRC and the seventh enzyme SAICARS have been cloned from *Vigna aconitifolia* (Chapman *et al*, 1994). The expression and regulation of these genes are under investigation.

During the isolation of PAL cDNA clones from the PPD-related cDNA library, one clone turn out to be chimeric, which contained a 938bp cDNA fragment exhibiting high homology to glycinamide ribonucleotide synthetase (GARS) or phosphoribosylamine-glycine ligase in deduced amino acid sequence. The cDNA was designated as cMeGARS (Fig. 7.3). Over the sequence of 170 amino acid residues, MeGARS shared 40% to 80% identity with GARS in yeast (S00652), human (X54199), mouse (U01024), *E coli* (X51950), *Bacillus subtilis* (Z99107), soybean (X96864) and *Arabidopsis thaliana* (X74766). The cMeGARS cDNA fragment is a partial sequence and it lacks the 3' end.

```
  1  TGTCCCTTTTCCTCTTCTTTTCCTCCTCTTCCTCCGGATTCTAGCTCCAGACGTGAAAAT   60
 61  TTTTTGAATTTCAGGGAGTATTTTGGGGGAAGAATAGAGCTATGTCTTGTGTTACTTTAA  120
                                          M  S  C  V  T  L  N
121  ACCTGGTACCTTCTATCAATCTCAATGGCAGAAACGCCAGTGTCAACTCTGCTCGATTAT  180
       L  V  P  S  I  N  L  N  G  R  N  A  S  V  N  S  A  R  L  S
181  CAAACAGTCTCTCTTGCGTGTTTGGCAACTCTTCTTCTTCGTCTTTCTCCTTTCTGGGAT  240
       N  S  L  S  C  V  F  G  N  S  S  S  S  S  F  S  F  L  G  Y
241  ACTTGAGTTCTAGTGATAGCAAAAATCGCTGTGATTCCCGCCGCGTAATCAAGGGCTGTA  300
       L  S  S  S  D  S  K  N  R  C  D  S  R  R  V  I  K  G  C  R
301  GGTCGTTTTCCTCTGTGTTCAAGTGTGTTTCCCAGAAATCAGAACCGTCAGTTTCGATTA  360
       S  F  S  S  V  F  K  C  V  S  Q  K  S  E  P  S  V  S  I  N
361  ATGCCCGTGGCAACGGTGCTTCTGAAGAGAGGGTGGTTGTGTTAGTTATTGGTGGTGGAG  420
       A  R  G  N  G  A  S  E  E  R  V  V  V  L  V  I  G  G  G  G
421  GAAGAGAACATGCACTTTGCTATGCCTTGCAACGATCACCATCCTGTGATGCTGTTTTCT  480
       R  E  H  A  L  C  Y  A  L  Q  R  S  P  S  C  D  A  V  F  C
481  GTGCTCCTGGCAATGCGGGTATTTCCAATTCAGGGAATGCTACTTGTATTCCAGACCTTG  540
       A  P  G  N  A  G  I  S  N  S  G  N  A  T  C  I  P  D  L  D
541  ACATCTCTGATAGCTCAGCTGTTATCTCTTTCTGCCGCCAATGGAATGTGGGATTGGTTG  600
       I  S  D  S  S  A  V  I  S  F  C  R  Q  W  N  V  G  L  V  V
601  TTGTTGGACCAGAGGCACCTCTAGTTGCTGGCCTTGCAAATGATCTAGGGAAGGCTGGAA  660
       V  G  P  E  A  P  L  V  A  G  L  A  N  D  L  G  K  A  G  I
661  TCCTTACTTTTGGCCCATCTGCAGAGGCTTCAGCTTTGGAAGGTTCAAAGAACTTTATGA  720
       L  T  F  G  P  S  A  E  A  S  A  L  E  G  S  K  N  F  M  K
```

```
721  AGAGCTTGTGTGACAAATATGGAATTCCAACTGCAAAGTACCAAACATTTACAGATCCAT   780
       S  L  C  D  K  Y  G  I  P  T  A  K  Y  Q  T  F  T  D  P  S
781  CTGCTGCAAAGCAATATATTCAGAACCAGGGAGCTCCTATAGTTATCAAAGCAGATGGAT   840
       A  A  K  Q  Y  I  Q  N  Q  G  A  P  I  V  I  K  A  D  G  L
841  TGGCTGCTGGGAAAGGGGTTATTGTTGCAATGACACTGGAGGAGGCATATGAAGCTGTGG   900
       A  A  G  K  G  V  I  V  A  M  T  L  E  E  A  Y  E  A  V  D
901  ATTCAATGCTTGTGAAAGGTGCTTTTGGTTCTGCTGGC  938
       S  M  L  V  K  G  A  F  G  S  A  G
```

Fig.7.3 The sequence of a cDNA fragment encoding a putative glycinamide ribonucleotide synthetase (GARS).

7.4 Identification of a cDNA Encoding Phosphoserine Aminotransferase

Serine can be produced through two different pathways from two different precursors in plants. One is the glycolate pathway with glycine as the precursor. Serine is formed from two molecules of glycine, which is catalyzed by glycine decarboxylase. This pathway was assumed as the major pathway in photosynthetic tissues (Saito *et al*, 1997). In another pathway, precursor D-3-phosphoglycerate is oxidized to phosphohydroxypyruvate by 3-phosphoglycerate dehydrogenase, and the phosphohydroxypyruvate is converted to phosphoserine catalysed by phosphoserine aminotransferase (PSAT, EC2.6.1.52). Phosphoserine is dephosphorylated by phosphserine phosphatase and serine is produced (Keys, 1980). This pathway plays a primary role in the supply of serine in rapid proliferating tissues of low photosynthetic activity (Cheung *et al*, 1968 & Reynolds *et al*, 1988). Phosphoserine aminotransferase gene in spinach showed preferential expression in the roots of green seedlings and in the suspension cells cultured under dark conditions (Saito *et al*, 1997).

Part of the antisense strand of cMeHRGP1 cDNA was identified to encode putative phosphoserine aminotransferase (section 5.2). The region was 445bp in length. Its deduced amino acid sequences were aligned with those of phosphoserine aminotransferase genes in *Arabidopsis thaliana*, spinach (*Spinacia oleracea*) and yeast (*Saccharomyces cerevisiae*) (Fig.7.4). The amino acid sequences were highly conserved in the plants in the regions compared. The identity and similarity between the sequences of cassava and *Arabidopsis thaliana* was 81% and 89%. The identity and similarity between the sequences of cassava

and spinach was 76% and 84%. The plant phosphoserine aminotransferases also shared relatively high similarity (59%) with that of yeast.

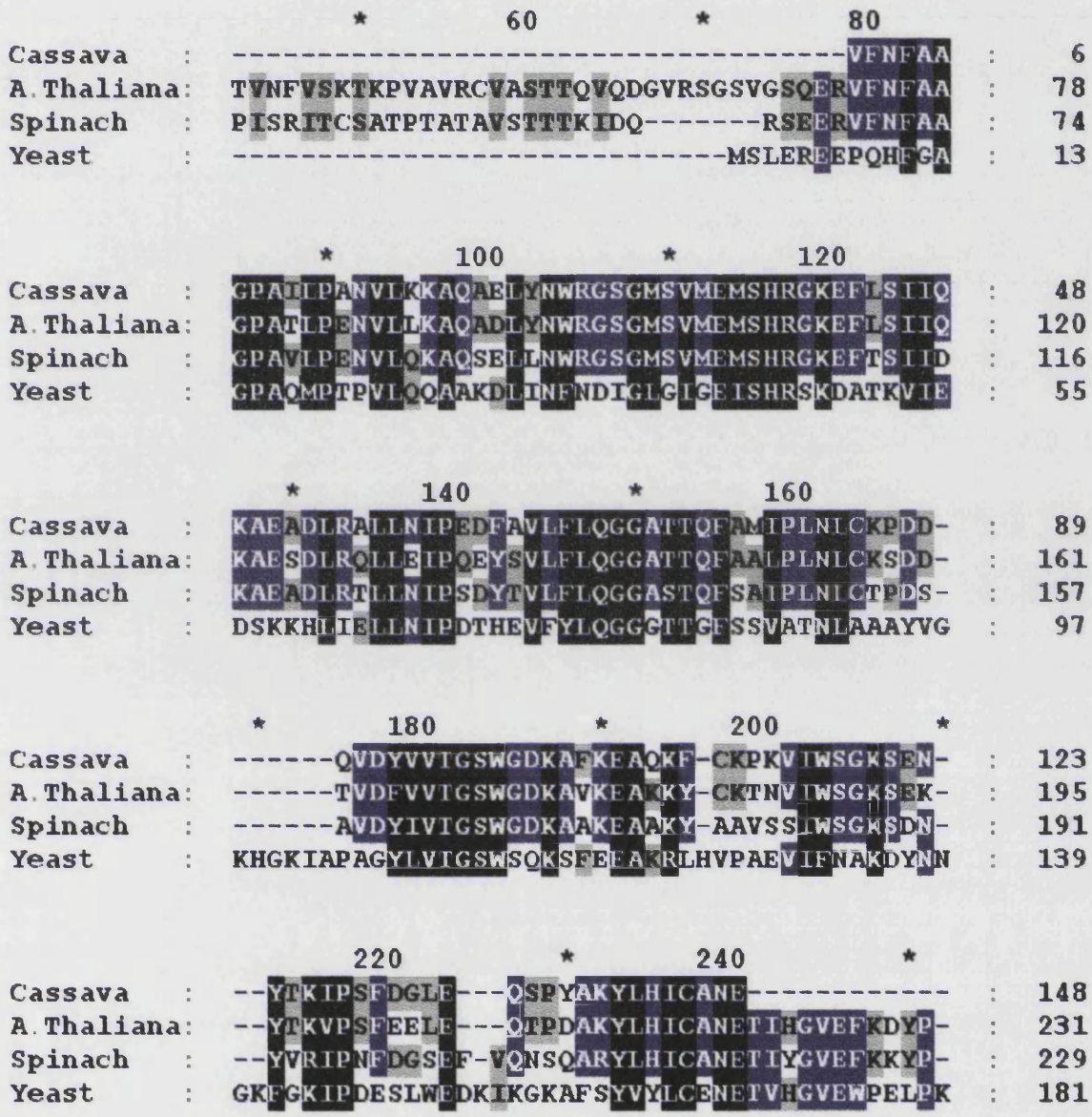


Fig.7.4 Multiple alignment of deduced amino acid sequences of phosphoserine aminotransferase genes in cassava, *Arabidopsis thaliana*, spinach and yeast (*Saccharomyces cerevisiae*). The full length sequences of PSAT in *A. thaliana* and spinach were 430 residues. The nucleotide sequence accession numbers are D88541 for *A. thaliana*, D84061 for spinach and L20917 for yeast. The residues with black shading are identical or three of them are identical and the other one is similar in biochemical properties among the four sequences. Purple-shaded residues indicate identical or similar among three sequences. Grey-shaded residues are identical or shows similar residues between two sequences.

7.5 Identification of a cDNA Clone Encoding an ACC Oxidase

The biological functions of ethylene have been extensively studied and it has been established that ethylene is involved in plant growth, development and senescence (for review see Reid, 1995).

The role of ethylene in promoting fruit ripening is well known. Ethylene was shown to be involved in regulating the expression of several ripening-related genes (Grierson *et al*, 1986). Manipulation of fruit ripening by ethylene treatment or inhibiting either ethylene biosynthesis or its action has been a common practice (Yang, 1987; Theologis, 1992). Aminoethoxyvinylglycine (AVG), inhibitor of ethylene biosynthesis, and 2,5-norbornadiene (NBD) and silver ions, inhibitors of ethylene action have been used for delaying fruit ripening.

Wounding is usually accompanied with increased ethylene production. Mechanical wounding of iceberg lettuce (*Lactuca sativa L.*) by cuts or punctures caused an increase in ethylene production (Ke and Saltveit, 1989). It was observed in lettuce that the changes in the content of phenolic compounds during the first three days of ethylene exposure appeared to follow the same pattern as the synthesis of the phenolic compounds during wound induction (TomasBarberan *et al*, 1998). ACC oxidase activity in tomato leaves increased 5-fold in response to wounding (B. Blume, unpublished results). Some plants respond to wounding with secretory processes such as production of latex in rubber tree (*Hevea*) and cassava. Latex production in rubber was shown to be stimulated by application of ethephon (Reid, 1995).

Similar to the wounding response, rapid accumulation of ethylene also occurs during plant defence against pathogens and a number of wound-related or / and defence-related genes were regulated by ethylene. It has been shown that ethylene is required for proteinase inhibitor (Pin2) gene expression upon wounding (Bowles, 1997). Down-regulation of ACO gene by antisense techniques led to the absence of Pin2 gene expression upon wounding. Ethylene has also been shown to be involved in the regulation of β-1,3-glucanases, HRGP

and PAL genes. In tobacco, ethylene treatment and wounding strongly induced basic glucanase genes but had little effect on acidic glucanase genes (Vanderhee *et al*, 1993). GUS driven by a β-1,3-glucanase promoter was induced in tobacco leaves by ethylene treatment (Vogelilange *et al*, 1994). Ethephon promoted the activities of acidic glucanase and chitinase in chickpea stems but not in roots, similar to the induction of their neutral isoforms in stems and roots by wounding (Cabello *et al*, 1994). Ethylene production was correlated with the induction of HRGP in melon during fungal infection (Roby *et al*, 1985). In maize seedlings, application of 100ppm ethylene resulted in a dramatic increase in the expression of HRGP (Tagu *et al*, 1992). In maize, ethylene appeared to be involved in the transduction pathway of HRGP gene activation by fungal elicitor (*F. moniliforme*) (GarciaMuniz *et al*, 1998). Fungal elicitor or ACC (ethylene precursor) or α-AB (α-aminobutyric acid, an inducer of ethylene production) enhanced significantly the expression of HRGP; while aminooxyacetic acid (AOA, inhibitor of ethylene biosynthesis) and silver thiosulfate (STS, an inhibitor of ethylene action) inhibited the effect of the fungal elicitor. A tobacco extensin gene was induced in leaves by wounding and ethylene (Memelink *et al*, 1993). An ethylene responsive element AGCCGCC was identified in the promoter of PRB-1b gene (basic-type pathogenesis-related protein) and PR-5 (osmotin-like protein) gene. Mutation of the sequence inhibited the induction of GUS expression driven by these promoters (Sessa *et al*, 1995; Sato *et al*, 1996). However, some defence-related genes seem to be regulated by an ethylene-independent wounding signal. A melon fruit PAL is transcriptionally induced both in response to fruit ripening and wounding. PAL gene expression was coincident with the kinetics of expression of ethylene biosynthesis genes during fruit development. However, in response to wounding, there was no such concomitant expression of these genes (Diallinas and Kanellis, 1994), indicating that this PAL gene was induced by an ethylene-independent wound signal.

ACC oxidase catalyses the last reaction of ethylene biosynthesis. Its importance in controlling ethylene production was demonstrated by the reduced ethylene production and the delayed fruit ripening in transgenic tomato plants containing antisense ACC oxidase gene (Hamilton *et al*, 1990).

Mechanical wounding of leaves or fruits also enhanced the expression of ACC oxidase genes. The expression of ACC oxidase genes in peach leaves increased within 1 to 4 hrs

after wounding, resulting in a dramatic increase in the evolution of ethylene (Callahan *et al*, 1992). A similar response was also observed in wounded peach fruits but the increase in the expression of ACC oxidase gene and ethylene production was not so significant (Callahan *et al*, 1992). Two ACC oxidase genes were accumulated in broccoli florets in different patterns (Pogson *et al*, 1995). One of them was induced within two hours post-harvest, while the other was expressed before harvest but its expression was enhanced dramatically post-harvest. It was not clear whether the expression patterns were a pure wounding response or an enhanced senescence by wounding during harvest.

Evolution of ethylene from cassava root slices increased within 6 to 16 hours after wounding (Plumbley *et al*, 1981; Hirose *et al*, 1984). Exogenous application of ethylene to cassava roots, however, did not affect the development of PPD (Hirose *et al*, 1984). The role of ethylene during PPD could not be deduced from these results. Therefore it was decided to clone ACC oxidase genes in cassava, so that the expression of ACC oxidase genes during PPD could be analysed, ACC oxidase might be down-regulated by transforming cassava with antisense ACC oxidase gene and the possible link between ethylene production and PPD might be elucidated.

The PPD-related cDNA library was screened using a probe made from ACO1 (ACC oxidase cDNA pTOM13, a kind gift from Professor Don Grierson) (Slater *et al*, 1985). One positive cDNA clone was isolated. The cDNA was then subcloned into pBS KS II(+) and sequenced (Fig.7.5). The cDNA, designated as cMeACO1 was 1006bp, similar to the size (about 1kb) of most ACC oxidase cDNAs in other plant systems. The deduced peptide from the cDNA sequence consisted of 324 amino acids. Homology search in NCBI data base with the sequence revealed that its deduced amino acid sequence had high similarity (up to 85%) to ACC oxidase genes in other plants such as tomato (LeACO1), tobacco (NtACO) and plume (PpACO). The amino acid sequence of cMeACO1 was compared with those of tomato LeACO1, tobacco NtACO and plume PpACO by multiple alignment (Fig.7.6). As it can be seen from the alignment, cMeACO1 is similar to the other ACO amino acid sequences in most of the regions except for the C terminals.

```
                                                        GAGAGAGAGA G    11

 12  ATGGAGTTCCCAGTCATCAACCTTGAGAAGCTTAATGGTGAGGAGAGAGCTGCCACCATG    71
     M  E  F  P  V  I  N  L  E  K  L  N  G  E  E  R  A  A  T  M

 72  GCTAAGATCAAAGATGCCTGTGAAAATTGGGGATTCTTTGAGTTGCTGAACCATGGAATA   131
     A  K  I  K  D  A  C  E  N  W  G  F  F  E  L  L  N  H  G  I

132  GAGCCAGAGTTCTTGGACAGAGTTGAGAGTATGACAAAGGGTCACTACAGAAAATGCATG   191
     E  P  E  F  L  D  R  V  E  S  M  T  K  G  H  Y  R  K  C  M

192  GAGCAAAGATTCAAAGAAATGGTGGCCAACAAAGGCCTCGACGCCGTCCAAACTGAGATC   251
     E  Q  R  F  K  E  M  V  A  N  K  G  L  D  A  V  Q  T  E  I

252  AAAGATATGGACTGGGAGAGCACCTTCTTCATCCGTCACCTCCCTGACTCAAATCTTGCT   311
     K  D  M  D  W  E  S  T  F  F  I  R  H  L  P  D  S  N  L  A

312  CAGCTTCCTGATCTCGATGATGAACACAGGGCTGTGATGAAGGAATTTGCAGCAAAGCTG   371
     Q  L  P  D  L  D  D  E  H  R  A  V  M  K  E  F  A  A  K  L

372  GAGAAACTGGCGGAGGATCTTTTGGACCTGTTGTGTGAGAATCTTGGGCTCGAGAAAGGT   431
     E  K  L  A  E  D  L  L  D  L  L  C  E  N  L  G  L  E  K  G

432  TACCTGAAGAAGGCGTTCTATGGGTCCAGGGGTCCAACCTTTGGCACCAAGGTTAGCAAC   491
     Y  L  K  K  A  F  Y  G  S  R  G  P  T  F  G  T  K  V  S  N

492  TACCCACCATGTCCCAAGCCAGACCTGATCAAGGGACTCAGAGCCCACACAGACGCTGGT   551
     Y  P  P  C  P  K  P  D  L  I  K  G  L  R  A  H  T  D  A  G

552  GGCATCATCTTGCTATTTCAAGATGACAGGGTCAGTGGCCTTCAACTCCTGAAAGATGGG   611
     G  I  I  L  L  F  Q  D  D  R  V  S  G  L  Q  L  L  K  D  G

612  CAATGGATTGATGTGCCTCCTATGCGCCACTCCATTGTTGTTAACCTTGGAGACCAGCTT   671
     Q  W  I  D  V  P  P  M  R  H  S  I  V  V  N  L  G  D  Q  L

672  GAGGTGATTACGAATGGCAAGTACAAGAGTGTCGAGCACAGAGTTGTCGCTCAGACAGAC   731
     E  V  I  T  N  G  K  Y  K  S  V  E  H  R  V  V  A  Q  T  D

732  GGTACCAGGATGTCATTAGCTTCATTCTATAACCCTGGAAGTGATGCAGTGATCTACCCT   791
     G  T  R  M  S  L  A  S  F  Y  N  P  G  S  D  A  V  I  Y  P

792  GCTCCAGCTTTGGTGGAGAAAGAAGCAGAAGAGAAGAAGCAAGTCTACCCGAAATTTGTG   851
     A  P  A  L  V  E  K  E  A  E  E  K  K  Q  V  Y  P  K  F  V

852  TTTGAAGACTACATGAAGCTCTATGTTGGGTTGAAATTTCAGGCGAAGGAGCCAAGATTT   911
     F  E  D  Y  M  K  L  Y  V  G  L  K  F  Q  A  K  E  P  R  F


912  GAAGCCATGAAGGCTGTGGAAAATAATGTCAATTTGGGTCCCAATTGCTACTGCTTAATT   971
     E  A  M  K  A  V  E  N  N  V  N  L  G  P  N  C  Y  C  L  I

972  ATAAATTATTATTAATATTATTTTAATTTTAGGGG  1006
     I  N  Y  Y  *
```

Fig.7.5 Nucleotide sequence of an ACC oxidase cDNA isolated from the PPD-related cDNA library in cassava. * indicates translation stop codon.

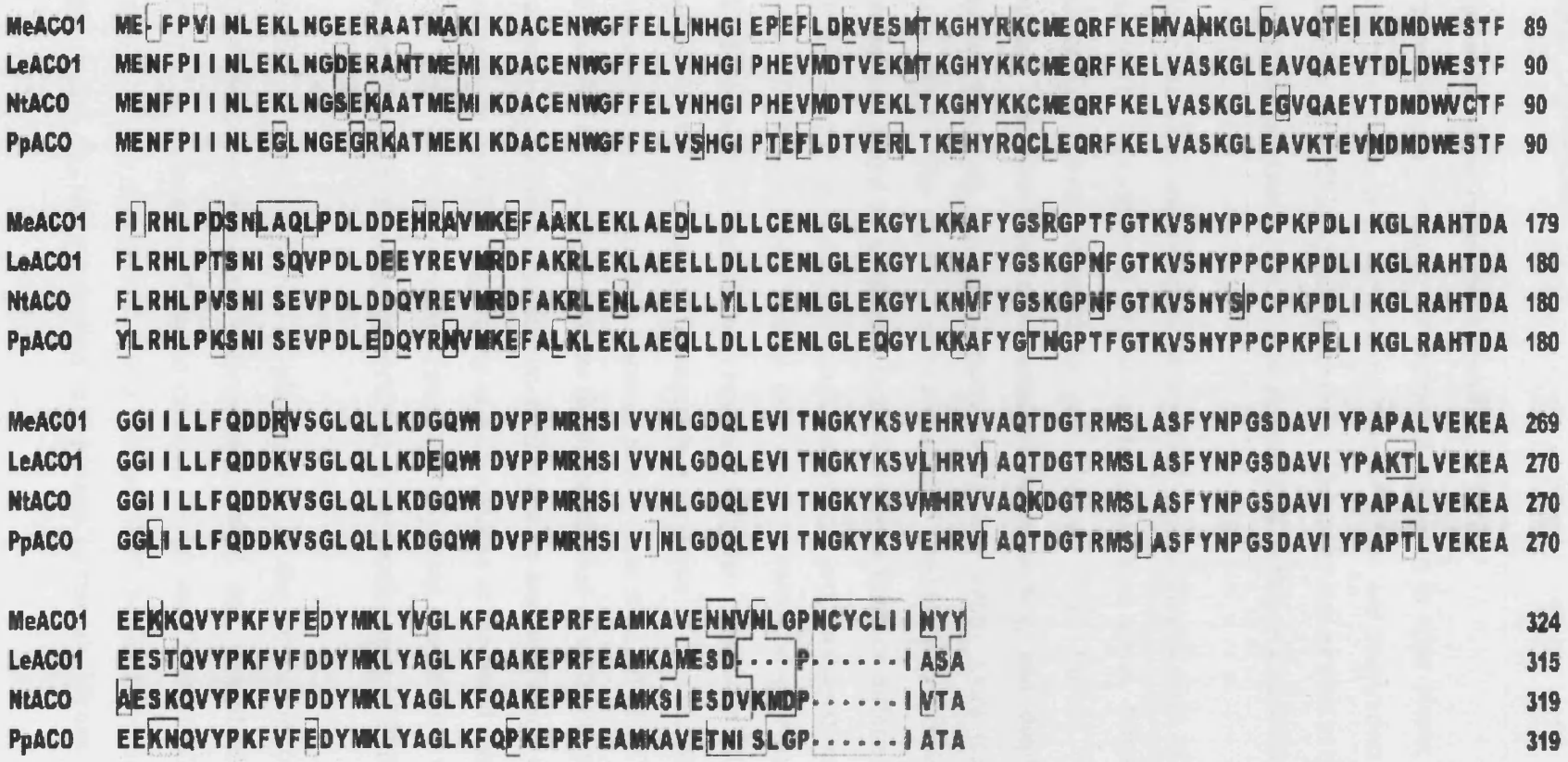


Fig. 7.6 Multiple alignment of the deduced amino acid sequences of cMeACO1, LeACO1, NtACO and PpACO. Residues which are different among these sequences are indicated by boxes. Access No: LeACO-A35021, NtACO-X98493, PpACO-X77232.

Chapter Eight

General Discussion

8.1 Towards a Molecular Understanding of PPD

PPD appears to be similar to the wounding responses in other plants; for example, there is an associated increase in the production of ethylene and respiration rate. Yet the wound response during PPD is not localized at the wound site but spreads to the whole root, first to xylem vessels and then to the storage parenchyma. Is this related to the wound signals?

Ethylene has been suggested to be involved in the transmission of the wound stimulus leading to the expression of wound- or defense-related genes. Ethylene production from stored cassava roots increased to peak 16-18hr after harvest (Hirose, 1986). PPD susceptible cultivars were more sensitive to wounding, and displayed more ethylene production than less susceptible cultivars (Hirose, 1986). Does the high production of ethylene trigger the cascade of non-localized wounding response? Ethylene at high concentrations could be a negative factor for regeneration or differentiation. It is easier to obtain regenerated plantlets from tomato callus with suppressed expression of ACC oxidase gene than from callus with normal expression or enhanced expression of ACC oxidase genes (Han *et al.*, unpublished results). Ethylene has been shown to affect cell differentiation or regeneration; for example, expression of an antisense ACC oxidase gene stimulated shoot regeneration in *Cucumis melo* (Amor *et al.*, 1998). There is the possibility that high ethylene production affects the differentiation of cells into periderm during PPD. The effect of ethylene production on PPD could be analysed using inhibitors of ethylene action (such as $AgNO_3$). The activity of ACC oxidase in harvested roots could be compared with the activity in non-harvested (still attached) roots after similar wounding in order to understand if ethylene production is related to the development of PPD.

Ethylene has been suggested to play a role in the activation of genes involved in phenylpropanoid metabolism, pathogenesis-related and wound-related genes such as chitinase, β-1,3-glucanase, Pin2 and extensin (Ecker and Davis, 1987; Boller *et al.*, 1983; Bol *et al.*, 1996; Takeuchi *et al.*, 1990; Bowles, 1997; Memelink *et al.*, 1993). In wounded cassava roots, ethylene production in the roots increased (Hirose, 1986), PAL activity

increased (Tanaka *et al.*, 1983; Uritani et al., 1984) and β-1,3-glucanase genes were activated as shown in chapter 6 of this thesis. But it is not clear whether the activation of these genes is related to the increased production of ethylene. The activation of β-1,3-glucanase gene cMeGLuc1 was very late, 72hrs after harvest in cultivar M Col22, compared with 3-6hrs needed for induction of glucanases by ethylene and wounding in potato leaves (Beerhues and Kombrink, 1994). If β-1,3-glucanases function in a similar way during wounding responses as during defense response against fungal pathogens, the late induction of β-1,3-glucanase gene expression may lead to the release of signals too late to induce further responses for wound healing.

The development of PPD is closely associated with wounding and water loss (Booth, 1976; Marriott *et al.*, 1978). ABA is involved in the response of plants to water loss and wounding response (Cornish and Zeevaart 1985; Peña-Cortês *et al.*, 1996). The possible role of ABA in PPD has not been investigated. However, it is interesting to note that two regions in the promoter of gMePAL2 are similar to the sequence RYACGTGGYR (R refers to purine-containing base and Y to pyrimidine-containing base) of an ABA responsive element (Bray, 1997; Shen *et al.*. 1995). If gMePAL2 is induced during PPD, its ABA-responsive element in its promoter may provide a hint that ABA may be involved in PPD.

With the construction of the first PPD-related cDNA library, and the subsequent isolation of several wound- or defense-related cDNA clones from the library, PPD can begin to be understood at a molecular level. The isolated cDNA clones from the library encoded PAL, HRGPs, β-1,3-glucanase, ACC oxidase, cysteine proteinase inhibitor, putative serine/threonine protein kinase and catalase, which were isolated by the author of this thesis and Reilly (1998) respectively. Using cDNA-AFLP, hundreds of PPD-specific genes were recently shown to be induced during PPD (Huang *et al.*, 1998).

Four PAL genes were cloned in cassava and three of which were expressed during PPD. However, the expression patterns of two PAL genes MePAL and MePAL1 were different from the pattern of PAL activity during PPD, indicating that the other PAL genes may be expressed in a different pattern from MePAL and MePAL1. The distribution of PAL activity among different PAL genes needs to be clarified. It is not clear which PAL gene

contributes most to the phenolics which accumulate and are responsible for the discoloration during PPD. The PAL gene family in cassava needs to be fully explored, by cloning the full length sequences of the genes and analyzing their expression patterns and the promoter acitivities. If it could be established that one PAL gene is responsible for the production of scopoletin by comparing the expression patterns of PAL genes and the accumulation pattern of scopoletin during PPD, antisense construct of this gene may be transformed into cassava to test its effect on PPD.

A cDNA with a serine/threonine protein kinase domain was isolated from the PPD-related cDNA library. The determined partial sequence of this clone is of high similarity to serine/threonine protein kinase or receptor-like kinase, such as *Xa21* and Fen, which are thought to be involved in signal transduction (Ronald, 1997). Isolation and characterization of the full length cDNA of this clone could shed more light on the possible function of this gene during PPD.

Genes of HRGPs, a frequent component of wound healing, were expressed during PPD. Comparison of the expression of HRGP genes upon wounding in attached roots undergoing localised healing and detached roots undergoing PPD could reveal whether sufficient expression of HRGPs for healing is induced during PPD or not. HRGPs function *via* cross-linking by peroxidases to seal off the wounding sites. But no data are available to tell if this oxidation during PPD is enough for cross-linking HRGPs to seal the wound. The oxidation of large amount of scopoletin and other phenolic compounds produced during PPD could place very high demands on ROS. If these phenolic compounds are oxidized in preference to other substances such as proline-rich cell wall proteins, insolubilisaition of these proteins would be inhibited and wound healing may be affected.

β-1,3-glucanases are proposed to function by releasing phytoalexin elicitors by hydrolysing glucans in fungal pathogens to stimulate further responses. It is not known whether they function in a similar way during wounding by breakdown of plant cell walls to induce further responses. β-1,3-glucanase–deficient plants have been generated by antisense transformation to assess the effect on susceptibility to disease (Beffa and Meins, 1996) but no research has been done on its effect on wounding response. The special wounding

response (with no or very limited healing response) during PPD provides the opportunity to investigate the possible functions of β-1,3-glucanase and other wound-inducible genes in wound healing. β-1,3-glucanase gene(s) can be over-expressed in the storage root or constitutively in the plant to analyse its effect on PPD. It may be argued that one single gene may not have a significant impact on wound healing as the healing process is most likely to be a complex process involving several or many genes. However, several genes may be transferred into cassava to test their roles in wound healing. Recently, it was reported that up to 14 genes have been transferred into rice in one transformation event and all of them were expressed (Chen *et al.*, 1998)

Catalase activity was detected by tissue printing in the parenchyma tissue but not in the cortex, and the activity increased to a peak at 48-72hr after harvest in cultivar Mcol 22 (Reilly *et al.*, 1998). Using a similar method, Reilly *et al.* revealed that peroxidase activity was initially predominant around the periphery of the tuber and xylem vessels. Rapid spread of peroxidase activity throughout the tissue was observed in cultivars with high PPD response A cassava cDNA clone showing 91% amino acid similarity to *Ricinus communis* Catalase 2 protein was isolated from the PPD-related cDNA library and its expression pattern is being analysed (Reilly *et al.*, 1998).

Two cDNA clones of cysteine protease inhibitor were isolated from the PPD-related cDNA library (Reilly *et al.*, 1998). Proteinase inhibitors are an important element of the plant defense against insect attack. Cysteine protease is one of the major digestive proteolytic enzymes of insects. The gene expression of protease inhibitors such as serine protease inhibitor and cysteine protease inhibitor can be induced by mechanical wounding (Pena-Cortes *et al.*, 1995; Botella *et al.*, 1996). It is not clear yet whether the cassava cysteine proteinase inhibitor genes are wound inducible or not. If they are wounding specific, they could be used for dissecting the possible signals involved in the special wounding response during PPD.

A cDNA clone isolated from the PPD-related cDNA library showed high homology to other plant aspartic proteases (Reilly *et al.*, 1998), which are suggested to be involved in programmed cell death (Koonberg-Roos *et al.*, 1991). During storage, cassava roots

probably undergo senescence. The expression of the putative aspartic protease gene during PPD and senescence of leaves could be analysed to understand the possible functions of the gene.

The factors affecting PPD, the responses (investigated so far) of the storage root during PPD and the relevant cDNA clones isolated from the PPD-related cDNA library are illustrated diagrammatically in Fig.8.1.

8.2 Identification of Genes for the Genetic Manipulation of PPD

One strategy to identify the key genes involved in PPD has been to isolate genes expressed during PPD and analyse their roles in PPD. One approach has been to screen a PPD-related cDNA library with probes made from wound- or defense response-related genes from other plant systems. By this approach, as used in this project, a number of genes have been cloned and characterized. These genes include the clones described in chapters 4, 5, 6 and 7 of this thesis and those isolated by Reilly *et al.* (1998) such as catalase, aspartic protease and cysteine proteinase inhibitor. Another approach has been to clone genes specifically expressed in PPD and analyze their roles in PPD. cDNA-AFLP has been used to compare the cDNA fingerprints of physiologically deteriorating roots and freshly harvested roots, which revealed that hundreds of genes were specifically expressed during PPD. The first approach may lead to the understanding of whether the wound- or defense-related genes are expressed during PPD or not. For example, the isolation of cMeHRGPs and cMeGlucs indicated that there was a wound healing process and defense-related response during PPD. The second approach can lead to the identification of the key genes leading to the development of PPD. However, PPD may be the result of lack of expression of some wound healing-related genes or defense-related genes; these genes would not be identified with the strategy above as it was based on the deteriorating material.

Wound healing occurs when the storage roots are wounded but still attached to the plant or when the harvested roots are stored in high humidity and high temperature (80-90%RH, 35°C) or curing condition (Booth, 1976). By running differential hybridization between pools of mRNAs from the root undergoing healing process and those from deteriorating roots, healing specific genes and PPD specific genes may be identified. cDNA-AFLP

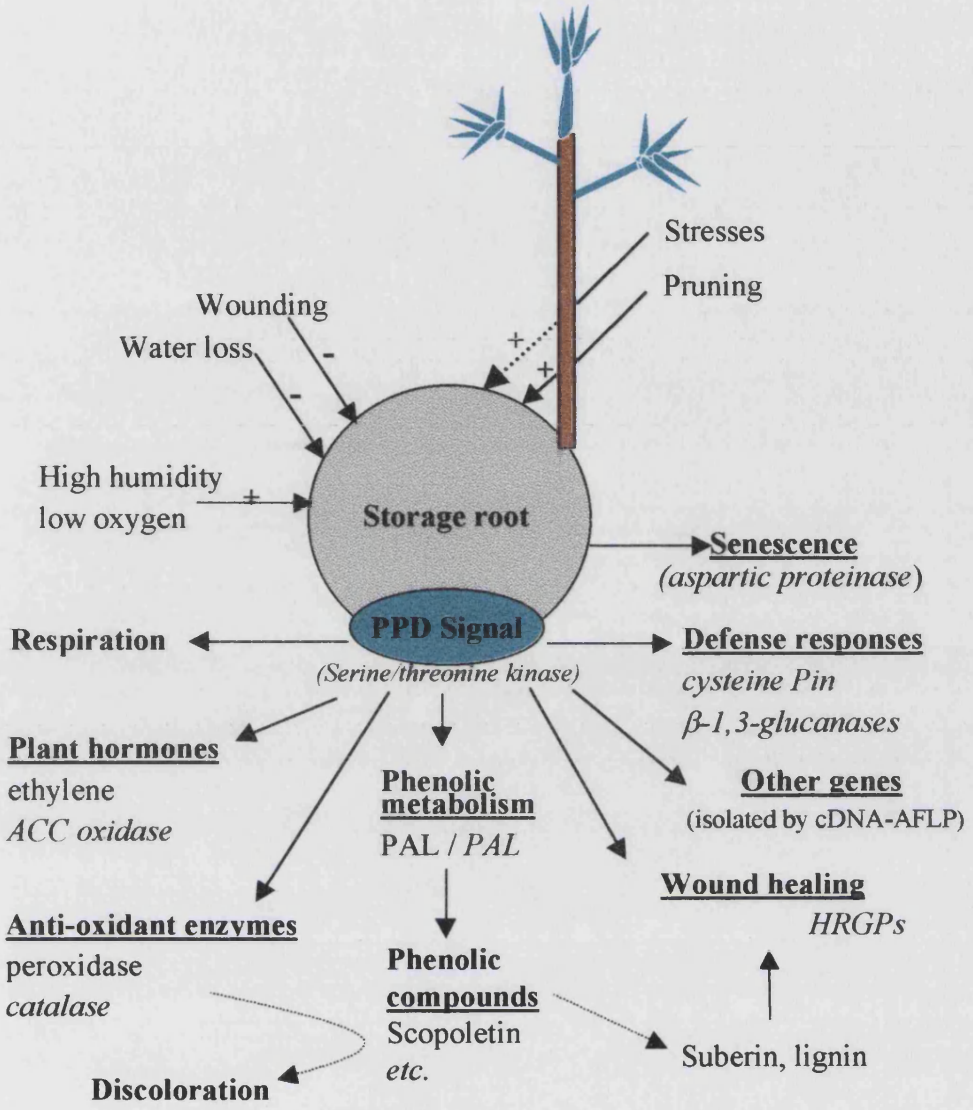


Fig.8.1 Schematic diagram of the factors affecting PPD and the responses occurring during PPD. PPD is closely related to wounding and aafected by water loss. Storage at high RH and low O_2 can induce the formation of periderm and delay the development of PPD. Stresses during the cultivation of plants can reduce the susceptibility of the root to PPD. Pruning 30cm above the stem base of the plant two to three weeks before the harvest can also delay the development of PPD. The signals in response to the stresses after harvest lead to increased respiration, production of ethylene (expression of ACC oxidase gene), biosynthesis of phenolic compounds such as scopoletin (expression of PAL genes), increased activity of peroxidase and catalase. Vascular streaking may be due to the oxidation of phenolic compounds and subsequent reaction with proteins or amino acids, which leads to the formation of colored pigments. Defense-related gene such as β-1,3-glucanase genes and cysteine proteinase inhibitor genes are expressed. HRGP genes are also expressed during PPD.

Note: cDNA clones isolated from PPD related library are in italics; possible processes are indicated with dotted arrows; '+' and '-' indicate positive and negative effects respectively on PPD.

would also be an efficient way to identify such genes. Pruning one to two weeks before harvest (Lozano *et al.*, 1978; Tanaka *et al.*, 1984) and storing the roots under curing conditions (Booth, 1976) can delay the PPD. However, these are probably due to different mechanisms, as the former treatment reduced the sensitivity to mechanical wounding whereas the latter did not. Comparison between the pools of mRNA from the roots treated in these two ways or between the pools of mRNA from either of them and normal deteriorating roots may also lead to the isolation of important genes for understanding PPD and possibly for genetic manipulation.

8.3 Possibility of Genetic Manipulation of PPD Using the Isolated Genes

8.3.1 PAL genes

Scopoletin appearred to be highly related to PPD. Roots from pruned plants contained much less scopoletin (about 10% of the unpruned control) (Wheatley and Schwabe, 1985) and PPD was delayed and slowed down (Lozano *et al.*, 1978; Tanaka *et al.*, 1984). Exogenous application of scopoletin to the roots of unpruned plants stimulated the development of PPD (Wheatly and Schwabe, 1985). Therefore, PPD may be controlled if the synthesis of scopoletin in the roots could be manipulated. It might be proposed to inhibit PAL activity or knock-out PAL gene expression using antisense techniques to reduce the production of scopoletin in order to inhibit PPD. However, as PAL is the first enzyme in the phenylpropanoid pathway which leads to several branch pathways for a broad spectrum of products, the constitutive inhibition of PAL could lead to the reduction in anti-microbial phenolic compounds, lignin and suberin production in organs besides storage roots, which could result in increased susceptibility to pathogen and impaired wound responses. Transgenic tobacco plants with suppressed level of PAL activity and correspondingly low level of chlorogenic acid exhibited more rapid and extensive lesion development than wild-type plants after infection by fungal pathogen *Cercospora nicotianae* (Maher *et al.*, 1994)

Reduced phenylpropanoid synthesis also influences systemic acquired resistance in transgenic tobacco. Salicylic acid (SA) levels increased on primary infection in tobacco and several other species, which might act as a signal to establish SAR (Klessing and Malamy,

1994; Gaffney *et al.*, 1993). SA is synthesized from trans-cinnamic acid, the product of deamination of phenylalanine catalyzed by PAL. A reduced level of SA was found in PAL-suppressed tobacco, which was still able to respond to TMV with hypersensitive resistance but unable to establish SAR. Exogenous application of SA partially restored the SAR (Pallas *et al.*, 1996)

Suppressing PAL activity may also interfere with other aspects of plant development. A wide range of abnormal plant development was observed in transgenic tobacco containing a heterologous PAL gene from bean (Maher *et al.*, 1994). It was proposed that the transgene disrupted PAL regulation and that some of the abnormal phenotypes might reflect interference with putative signals dependent on phenylpropanoid biosynthesis.

On the other hand, antisense or sense suppression of PAL expression also showed the potential of genetic manipulation of lignin content. Reduced lignin content was achieved in tobacco by suppression of PAL gene expression with a PAL transgene (Sewalt *et al.*, 1997).

Over-expression of a bean PAL gene in tobacco led to a flow of phenylpropanoid production into chlorogenic acid (Howles *et al.*, 1996). Suppression of PAL in tobacco with the PAL transgene from bean led to decreased chlorogenic acid. This means that the bean PAL gene in tobacco is responsible for diverting the metabolic flux into production of chlorogenic acid, indicating that PAL is a control point directing its products to a certain pathway(s). As PAL is encoded by a multigene family in most plants, it may be possible that different members of the gene family may control different pathways or one PAL gene may direct part of the PAL activity to a certain pathway.

By transforming plants with antisense genes for different PAL genes and studying their effects on the phenolic products or comparing the expression patterns of PAL genes with the profiles of phenolic compounds, clues to the possible links between different PAL genes and different branch pathways from general phenylpropanoid pathway could be obtained. The abnormal phenotypes of transgenic plants with suppressed or over-expressed PAL activity may be avoided by using specific promoters or wound-inducible promoters to drive the transgene rather than a constitutive promoter. Since the PAL genes in cassava are

of high similarity, gene-specific regions may have to be used to suppress the target PAL genes.

8.3.2 HRGP Genes

Transgenic tobacco plants over-expressing or co-suppressing extensin genes showed different levels of total hydroxyproline and intensity of immunolabeling in certain tissues (Memelink *et al.*, 1993). Plants over-expressing the extensin gene showed four times more hydroxyproline in the leaves and 1.5 times more in stems and roots than did control plants. Plants with down-regulated extensin gene expression by introducing antisense extensin gene showed reduced levels of hydroxyproline. Immunolabelling with an extensin antibody revealed that over-expressing the extensin gene led to consistently higher intensity of the labelling than in control plants and antisense down-regulated plants. Yet the effect of over-expression or down-regulation of the extensin gene on wound responses or wound healing was not investigated. However, it is encouraging to know that over-expression of the extensin gene did not alter plant growth and development. Cassava HRGP genes may be over-expressed to analyse their effects on PPD. Or alternatively, the HRGP genes of cassava may be over-expressed in tobacco to observe the effect of over-expression on wound healing, since tobacco transformation is a rapid and routine practice compared with cassava transformation.

8.3.3 ACC Oxidase Gene

Down-regulation of the expression of the genes controlling the biosynthesis of ethylene production by genetic engineering has provided some of the approaches for controlling fruit-ripening and improvement of shelf-life of fruit (Picton *et al.*, 1995). During PPD a susceptible cultivar showed higher ethylene production than a cultivar with lower susceptibility (Hirose, 1986), this needs to be further investigated as described in Section 7.2 to test the correlation between ethylene production and PPD. There is a possibility that down-regulation of ACC oxidase gene expression may lead to reduced PPD. In addition to the antisense technique for down-regulation of gene expression, sense silencing with inverted repeats introduced in the 5' UTR of the sense ACO transgene has been shown recently to be very efficient for co-suppression of the endogenous and transgene expression

(Hamilton *et al.*, 1998). Down-regulation with these techniques may be applied to explore the possibility of genetic manipulation of PPD.

8.4 The Search for Promoters for Genetic Manipulation of PPD

Choice of a suitable promoter is a crucial element for the success of genetic manipulation of PPD. A variety of promoters have been tested for tissue specific activity in cassava using transient assays and a low frequency of transient gene expression was observed in the storage roots (Arias-Garzon and Sayre, 1993). For example, the CaMV 35S promoter and soybean root-specific glutamine synthetase promoter had high activities in cassava leaves but had low frequency of transient expression in the storage roots (Arias-Garzon and Sayre, 1993).

Several promoters have been used in cassava transformation, but none of them was isolated from the cassava genome. CaMV 35S and cassava vein mosaic virus (CsVMV) constitutive promoters are the ones that so far have been the most used to drive transgenes in cassava transformation (Taylor *et al.*, 1999). CsVMV promoters and their derivatives showed constitutive, root- and phloem-specific activities in tobacco and rice (Verdaguer *et al.*, 1998). These promoters were fused to reporter gene uidA and have been transferred into cassava to test their expression (Taylor *et al.*, 1999). A senescence specific promoter *sag* from *Arabidopsis* was fused with a gene encoding cytokinin biosynthesis (or *ipt* gene) from *A. tumefaciens* and transformed into cassava to prolong the leaf life (Li *et al.*, 1998).

There are only a few gene promoters isolated from cassava. The promoter of cassava MeEF1, a gene encoding protein synthesis elongation factor, was isolated. Transient expression assay demonstrated that the promoter was able to drive GUS gene expression in leaves of cassava, *Arabidopsis*, radish and peas (Suhandono *et al.*, 1998). A PAL promoter of cassava has been isolated. It contains three putative *cis*-acting elements, which have been shown in other plants to be necessary for elicitor- and light-mediated expression (Logemann *et al.*, 1995). What is exciting about the promoter is that two putative G-box motifs are located in it. G-box (ACGTG) has been found in a wide variety of plant gene promoters and have been demonstrated to play a role in the control of gene expression in respond to a variety of environment stimuli (Menkens *et al.*, 1995). A 16bp motif

(GTACGTGTTATAAAACGTGT) in the promoter of extA extensin gene from *Brassica napus* controls the activation of the gene expression upon wounding (Shirsat, 1998). The two direct repeats (ACGTG) are similar to the G-box motif. When a G-box like motif CACGTG was deleted from the promoter of a horse-radish peroxidase prxC2 gene, the promoter failed to response to wounding (Kawaka *et al.*, 1994). Further analysis of the cassava PAL promoter is underway using cassava transformation with the PAL promoter fused with GUS.

PPD is an unusual response, a wounding response with minimal healing processes. PPD-specific or wound-inducible or root wounding-specific promoters should be isolated to identify efficient promoters for genetic manipulation of PPD. PPD-specific promoters can be isolated by identifying PPD-specific cDNA clones and isolating the corresponding genomic clones. cDNA-AFLP cDNA pools from physiologically deteriorating roots were compared with the cDNA pools from freshly harvested roots and hundreds of cDNAs appeared to be PPD-specific. These cDNAs are being cloned, analysed, and used for the isolation of PPD-specific promoters (Huang *et al.*, 1998).

Wound-specific or wound-inducible promoters have been used to drive candidate genes for wound-specific expression. The Pin II promoter was shown to be able to direct the expression of *ipt* gene (bacterial isopentenyl transferase, involved in cytokinin biosynthesis) at wound sites (Smigocki *et al.*, 1993). Transcripts of *ipt* in excised leaf discs increased dramatically within 24hr after wounding. Two cassava cDNA clones of a cysteine protease inhibitor have been isolated from the PPD-related cDNA library (Reilly *et al.*, 1998). Cysteine proteinase inhibitor genes in other plants showed constitutive expression or inducible expression in response to wounding (Botella *et al.*, 1996). The expression patterns of the two cassava genes of cysteine proteinase inhibitor in cassava would reveal whether the genes are wound-specific or wound inducible. If so, their promoters could be used to drive candidate genes for genetic manipulation of PPD.

Attempts are being made to isolate root-specific promoters. Using differential hybridization about 80 storage root specific (*vs* leaf) cDNA clones were identified. These clones were sequenced and tested for their root-specific expression before used in isolation of the

corresponding promoters (Bohl-Zenger *et al.*, 1998). A β-glycosidase promoter was isolated from cassava and proved to be root-specific by analysing the gene expression and the transient activity of the promoter in different tissues (Liddle *et al.*, 1997).

8.5 The Roles of the Isolated Genes in the Plant Response to Pathogen Infection can be Explored

Defense responses are similar to wounding responses in many aspects as described in Chapter1. Many genes are involved in both defense and wound responses. The isolated cDNA clones such as PAL, glucanases and HRGP, which have been associated with defense responses in many plants, may therefore be used to understand their possible roles in the interaction between cassava and its pathogens.

Following PPD, microbial deterioration occurs within few days. It may not be surprising that cassava roots fail in both wound healing and defense responses against pathogens. It is likely that genetic manipulation of PPD could also lead to reduced susceptibility to microbial deterioration at the same time if an approach is taken to enhance the wound healing response. Over-expression of HRGPs may help to seal off wound sites and establish a physiological barrier to pathogens. Other genes such as β-1,3-glucanase isolated from the PPD-related cDNA library may also be used to analyze the interaction between cassava and pathogens during microbial deterioration to understand the pathology and investigate the possibility of their usage in genetic engineering. Thus was observed in transgenic tobacco or alfalfa with constitutive expression of a beta-1,3-glucanase cDNA, these plants showed increased resistance to certain fungal pathogens (Lusso and Kuc, 1996; Yoshikawa *et al.*, 1993; Masoud *et al.*, 1996) (as described in Chapter 1).

8.6 Possible Mechanism of the Effect of Pruning on PPD

Pruning or cutting off the aerial part of a cassava plant 20 to 30cm above the soil 1 to 2 weeks before harvest can delay the development of PPD. This effect may be related to the signalling from the remaining part of the stem, the change of the sink/source relationship of the root, and the change of the state of hormone composition.

There is little research on the transduction of stress signals from shoot to roots; there has been substantial research on the hormones from the roots as signals for the shoots of the stressed plants (reviewed by Jackson, 1997). Signals were sent to shoots from roots as early warning of stress conditions such as drought, flooding, nutrient deficiency and salinity to induce changes in the shoots that may increase stress tolerance. The most convincing example of root to shoot communication involving a hormone as Jackson (1997) pointed out is that of the signaling molecule ACC (ethylene precursor) from flooded roots which promoted epinastic leaf growth. Similarly, cutting off the top of the cassava plant 1 to 2 weeks before harvest could release wound signals and turgor signals, which could be transducted to roots to induce responses to the stresses. Does pruning slow down PPD by inducing adaptive responses in roots or do the early stresses caused by pruning help the roots prepare for further stresses such as PPD?

Wounding of one leaf of young rice plants caused a strong and transient accumulation of JA, followed by induction of a number of pathogenesis-related-genes in the wounded leaf (Schweizer *et al.*, 1998). The non-wounded leaf which emerged after the wounding, was more resistant to challenge infection by the rice blast fungus *Magnaporthe grisea* (Hebert) Barr. The non-wounded leaf also showed a transient but delayed accumulation of JA. Unlike the wounded leaf, there was no accumulation of PR mRNAs in the non-wounded leaf. Schweizer *et al.* (1998) also showed that JA resulted in a similar degree of systemic disease resistance as wounding. Similarly, it might be argued that pruning could cause a systemic response in the storage roots, which enables the roots to show less susceptibility to further wounding or PPD, or pruning may lead to an adaptive response. As the pruning is applied one to two weeks before the harvest, could the systemic response last sufficiently long as to be effective upon further wounding during harvest? It seems unlikely. The induced expression of wound- or JA-inducible genes such as Jr1 (jasmonate-responsive gene 1), Wr3 (wound-responsive gene 3) and Vsp (vegetative storage protein gene) in leaves affected systemically after different leaves were wounded lasted just less than 24hrs in *Arabidopsis*, and systemic Jr2 expression lasted longer than 48hrs (Titarenko *et al.*, 1997). As JA can mimic the wounding response, exogenous JA could be applied to cassava plants at periods before harvest to test its effect on PPD, which may reveal if a systemic response caused by wounding could affect the development of PPD.

Pruning also results in the change of sink and source relationship. Leaves and storage roots are the source and sink of the plant respectively. The apical dominance exerted by the auxin from apical buds is eliminated by pruning off the top of the plant and the budding from the remaining nodes on the stem is activated. The storage roots start to supply nutrients needed for budding from the remaining stem, so they then serve as a source. Within 1 to 2 weeks after the pruning, the growth of the shoots relies fully on the roots until functional leaves or the source appear about 3 weeks after the pruning. The change of the storage root from a sink to a source means that breakdown of the storage components such as starch may exceed their synthesis. Indeed, the soluble sugar content in roots from pruned plants was about twice of that in roots of non-pruned plants (Data *et al.*, 1984; Hirose, 1986). Therefore, it is probable that pruning leads to the increased ability of the roots to retain water when the roots are harvested, as the switch from sink to source could increase the concentration of osmolytes such as sugars and amino acids. Since PPD is strongly associated with water loss, the possible reduced water loss due to pruning may explain partly why pruning can slow down the development of PPD. To test whether it is the source or sink state or the possible systemic response that affects the rate of PPD after pruning, the comparison between the effect of pruning from a few nodes up and from the base of the stem on PPD could be made.

8.7 Water loss or drought and PPD

PPD has been shown to be closely associated with water loss (Marriott *et al.*, 1978). The responses to water loss are affected by different factors. The responses of plants to water loss are dependent on the extent and rate of the water loss, the stage of development, organ and cell types. A slow water loss may permit acclimatization to the change and limit the extent of injury, while a rapid rate of water loss may preclude acclimatization (Leone *et al.*, 1996). At high relative humidity, water loss was slow (Marriott *et al.*, 1978), which may allow the storage roots of cassava to acclimatize and lead to healing process at the wound sites. When non-harvested roots are wounded, wound healing occurs, which may be related to hormone signalling, or may be partly related to the possible high humidity condition in

the ground. Even if harvested roots are stored in the ground, PPD is delayed (Ravi *et al.*, 1996), which also suggests that water loss rate is important to the development of PPD.

Water loss is not only affected by the storage conditions but could also be influenced by the state of metabolism of the plants. Roots from pruned and non-pruned plants, may be differ in hormone composition, respiration and soluble sugars, which could affect their response to water loss. The contents of possible osmolytes such as amino acids (eg prolines), sugar alcohols (eg.pinitol), other sugars (eg. fructans) in pruned and non-pruned roots should be compared during different stage of PPD to understand to what extent they may affect PPD. Preliminary analysis of post-harvest water loss rate (Table. 8.1) was carried out in the roots harvested from untreated plants (control), from plants pruned from the base of the stem 1 week before harvest and from plants pruned 20cm from the base of the stem with a few nodes left. One week after pruning, 3cm long shoots emerged from the top three nodes. Within the first hour after harvest, the water loss rate was 1.09%(w/w) from the roots of plants pruned 20cm from the base of the stem, 1.17% in the roots from plants pruned from the base of the stem and 1.22% in the roots from control. During 4, 8 and 22 hrs incubation after harvest, water loss of roots from pruned plants were similar to the control or higher than control. It seems that pruning 20cm from the stem base one week before harvest may reduce the rate of water loss in the first hour after harvest. Further experiments should be done on a larger scale, ideally in the field. Effects of pruning two or three weeks and longer before harvest should be tested to assess the effects of source-sink switch on the development of PPD. It would also be interesting to investigate the difference in osmolytes and water retaining ability among different cultivars with different susceptibility to PPD.

Table 8.1 Water loss rate (% FW) in roots from pruned and non-pruned cassava plants

	1hr	2hr	4hr	8hr	22hr
Non-pruned	1.22±0.10	1.97±0.14	3.29±0.13	5.37±0.43	12.06±1.05
Prune A	1.17±0.17	2.37±0.65	3.52±0.37	6.03±0.49	13.86±0.92
Prune B	1.09±0.16	1.85±0.24	3.25±0.22	6.05±0.56	12.69±1.42

Note: Cassava plants were pruned 1 week before harvest. Two plants were pruned from the base of the stem without leaving any node, Prune A. Two plants were pruned 20cm above the stem base leaving some nodes on, Prune B. Roots were harvested and sliced into 2 cm thick slices and incubated at 27±3°C.

Roots from plants grown under stress conditions are less susceptible to PPD. The adaptation of the plant to stress before harvest might play a role in the reduced susceptibility to PPD (Wheatley, 1980). Higher levels of soluble sugars were produced in roots after drought than in roots from plants under favorable cultivation conditions (observed by Thailand scientist, Sriroth K, personal communication). The high soluble sugars might enable the roots to maintain a low rate of water loss, possibly leading to less susceptibility to PPD.

8.8 Conclusion

A physiological deterioration-related cDNA library, the first of its kind, was constructed from cassava storage roots. The library has been providing an efficient way for isolating PPD-related genes to understand the molecular processes of PPD. From the cDNA library, a number of PPD-related genes have been isolated, including all the cDNA clones isolated in this project and other clones such as catalase and cysteine protease inhibitor (Reilly *et al.*, 1998).

Two PAL cDNA clones, cMePAL1 and cMePAL3, were isolated from the cDNA library using a probe prepared from bean PAL gene gPAL2 (Cramer *et al.*, 1995). They were 2253bp and 1126bp (containing only the 3' region of the gene) respectively, and their deduced amino acid sequences shared very high similarity (93%). Comparison among the cassava PAL genes including cMePAL1, gMePAL2 (Hongying *et al.*, unpublished results), cMePAL3 and MePAL (Periera *et al.*, 1999) showed that they shared high similarity (over 90%) in deduced amino acid sequences. It was noted that cMePAL3 and gMePAL2 were of extremely high similarity over the known sequence of cMePAL3, 98% in DNA sequence and about 100% in amino acid sequence. Investigation was carried out to identify whether the difference was due to the fact that these two genes from two different cultivars, which demonstrated that there were these two genes in one cultivar. Cassava PAL genes shared high similarity in deduced amino acid sequences with PAL genes in other plant species, especially dicotyledonous tree plants such as poplar and lemon (93% and 92% respectively). Southern analysis indicated that there was a multiple PAL gene family in cassava, with at least four genes in the family. MePAL1 was expressed in young leaves, stems and vegetative roots. It was not expressed in unwounded storage roots but was

induced within 8hr after harvest; its expression stopped for at least 20 hrs and was induced again 40 hr post-harvest. The double peak expression pattern of MePAL1 was also observed with MePAL (Periera *et al.*, 1999). The first peak may be the initial response to the wounding and the second may be caused by PPD.

Nine putative HRGP clones with inserts ranging from 100bp to 1900bp were isolated from the cDNA library using a heterologous probe. One cDNA, cMeHRGP1, was subcloned and fully sequenced. Its deduced amino acid sequence consisted of repetitive motifs such as $Ser(Pro)_4$ and $(Tyr)_3$. It also had large and tandem repetitive motifs such as $Ser(Pro)_4$-$(Tyr)_3$-His-$Ser(Pro)_4$-Val-Lys and $Ser(Pro)_4$-$(Tyr)_3$-His-$Ser(Pro)_4$-Ser-Pro, which was a combination of motifs present in dicot and monocot HRGPs (Sommer-Knudsen, 1998). Partial sequences of other positive clones showed that their deduced amino acid sequences had repetitive motifs and palindromes. Three of the clones had high similarity to cMeHRGP1 in nucleotide or deduced amino acid sequences, and two of them showed high similarity to the sequences of the probe used for their isolation. The isolation of these HRGP cDNA clones suggested the presence of wound healing components during PPD.

Two putative β-1,3-glucanase cMeGLUC1 and cMeGLUC2 among the eight positive clones isolated from the cDNA library were analysed. cMeGLUC1 corresponded to 3' part of cMeGLUC2 with only one base difference. cMeGLUC2 shared high identity with with PR-Q', a possible elicitor-releasing glucanase in tobacco. Northern analysis showed that β-1,3-glucanase genes were expressed 72 hr post-harvest.

A putative ACC oxidase cDNA clone was isolated using tomato ACC oxidase cDNA pTOM13 as a probe. Non-preconceived target genes including those encoding putative ubiquitin-activating enzyme, serine/threonine kinase, glycinamide ribonucleotide synthetase (GARS) and phosphoserine aminotransferase (PSAT) were also identified from the cDNA library during screening for target genes.

With the isolation and characterisation of the PPD-related genes from the cDNA library, the molecular mechanism of PPD begins to be explored. With the establishment of cassava transformation techniques, the roles of these genes in the development of PPD can be

investigated by knock-out techniques such as antisense and sense gene suppression. However, to understand whether the development of PPD is due to absence of some normal components of wound response, a comparative study of the genes expressed in deteriorating roots and roots undergoing wound healing should be carried out, using techniques such as differential screening or cDNA-AFLP.

The delayed PPD by pruning two or three weeks before harvest may be due to the systemic wound response caused by pruning, but more likely due to the change of sink and source relationship between storage roots and the top, and possible subsequent reduction of water loss at the initial stage of storage.

Reference

Ahn JH, Choi Y, Kim SG, Kwon YM, Choi YD and Lee JS 1998 Expression of a soybean hydroxyproline-rich glycoprotein gene is correlated with maturation of roots. Plant Physiology 116(2): 671-679

Ahn JH, Choi Y, Kwon YM, Kim SG, Choi YD and Lee JS 1996 A novel extensin gene encoding a hydroxyproline-rich glycoprotein requires sucrose for its wound-inducible expression in transgenic plants. The Plant Cell 8: 1477-1490

Akano AO and Dixon AGO 1998 Reactions of four F1 mapping populations to cassava mosaic disease and cassava bacterial blight. The Fourth International Scientific Meeting Cassava Biotechnology Network, Brazil

Allem AC 1987 *Manihot esculenta* is a native of the neotropics. Plant Genetic Resource Newsletter (International Board Plant Genetic Resource) 71: 22-24

Allem AC, Mendes RA, Salomao AN and Burle ML 1998 The value of aggregation of *Manihot* genetic resources. The Fourth International Scientific Meeting Cassava Biotechnology Network, Brazil

Allem AC, Roa AC, Mendes RA, Salomao ML, Burle ML, Second G, de Carvalho PCL and Cavalcanti J 1998 The primary genepool of cassava (*Manihot esculenta* Crantz). The Fourth International Scientific Meeting Cassava Biotechnology Network, Brazil

Amor MB, Guis M, Latche A, Bouzayen M, Pech JC and Roustan JP 1998 Expression of an antisense ACC oxidase gene stimulates shoot regeneration in *Cucumis melo*. Plant Cell Report 17(6-7): 586-589

Arias-Garzon DI and Sayre RT 1993 Tissue specific inhibition of transient gene expression in cassava (*Manihot esculenta* Crantz). Plant Science. 93(1-2): 121-130

Arias-Garzon DI and Sayre RT 1998 Genetic engineering approaches to reducing the cyanide toxicity in cassava (*Manihot esculenta*, Crantz). The Fourth International Scientific Meeting Cassava Biotechnology Network, Brazil

Averre CW 1967 Vascular streaking of stored cassava roots. Proceedings of 1st International Symposium on Tropical Root Crops. Trinidad 2(4): 31-35

Awoleye F, van Duren M, Dolezel J and Novak FJ 1994 Nuclear DNA content and *in vitro* induced somatic polyploidization cassava (*Manihot esculenta* Crantz) breeding. Euphytica 76: 195-202

Awoyinka AF, Abegunde VO and Adewusi SRA 1995 Nutrient content of young cassava leaves and assessment of their acceptance as a green vegetable in Nigeria. Plant Food for Human Nutrition 47: 21-28

Baldwin TC, Coen ES, Dickinson HG 1992 The *ptl1* gene expressed in the transmitting tissue of *Antirrhinum* encodes an extensin-like protein. *Plant Journal* 2:733-739

Becker F, Buschfeld E, Schell J and Bachmair A 1993 Altered response to viral infection by tobacco plants perturbed in ubiquitin system. Plant Journal. 3: 875-881

Beeching JR, Dodge AD, Moore KG and Wenham JE 1995 Physiological deterioration in cassava: An incomplete wound response? In: Thro AM & W Roca, The Cassava Biotechnology

Network: Proceedings of the Second International Scientific Meeting. CIAT Cali, Colombia. Pp 729-736

Beerhues L and Kombrink E 1994 Primary structure and expression of mRNAs encoding basic chitinase and β-1,3-glucanase in potato. Plant Molecular Biology 24: 353-367

Beffa R and Meins JF 1996 Pathogenesis-related functions of plant β-1,3-glucanases investigated by antisense transformation – a review. Gene 179: 97-103

Benhamou N, Mazau D, Grenier Jm and Esquerretugaye MT 1991 Time-course study of the accumulation of hydroxyproline-rich glycoproteins in root-cells of susceptible and resistant tomato plants infected by *Fusarium-oxysporum F sp Radicis-lycopersici*. Planta 184(2):196-208

Bennet RN and Wallsgrove RM 1994 Secondary metabolites in plant defence mechanisms. New Phytology 127: 617-633

Benz A and Spring D 1995 Identification and characterization of an auxin-degrading enzyme in downy mildew infected sunflower. Physiological and Molecular Plant Pathology 46(3):163-175

Bernards MA and Lewis NG 1992 Alkyl ferulates in wound-healing potato tubers. Phytochemistry 31: 3409-3412

Bernards MA and Lewis NG 1998 The macromolecular aromatic domain in suberized tissue: a changing paradigm. Phytochemistry 47(6): 915-933

Bernards MA, Fleming WD, Llewellyn DB, Priefer R, Yang XL, Sabatino A and Plourde GL 1999 Biochemical characterization of the suberization-associated anionic peroxidase of potato. Plant Physiology 121(1): 135-145

Birkenmeier GF and Ryan CA 1998 Wound signaling in tomato plants: Evidence that ABA is not a primary signal for denfense gene activation. Plant Physiology 117: 687-693

Bohl-Zenger S, Potrykus I, Puonti-Kaerlas J 1998 Isolation of root specific promoters from cassava. The Fourth International Scientific Meeting Cassava Biotechnology Network, Brazil

Bol JF, Buchel AS, Knoester M, Baladin T, VanLoon LC, Linthorst HJM 1996 Regulation of the expression of plant defence genes. Plant Growth Regulation 18(1-2): 87-91

Boller T 1988 Ethylene and the regulation of antifungal hydrolases in plants. Oxford survey plant molecular cell biology 5: 145-174

Bolwell GP and Wojtaszek P 1997 Mechanisms for the generation of reactive oxygen species in plant defence – a broad perspective. Physiological and Molecular Plant Pathology 51: 347-366

Bolwell GP, Davies DR, Gerrish C, Auh C-K and Murphy TM 1998 Comparative biochemistry of the oxidative burst produced by rose and French bean cells revealed two distinct mechanisms. Plant Physiology 116:1379-1385

Bonnelye E and Laudet V 1994 Overlapping genes. M S-Medecine Sciences 10(8-9): 805-816

Booth RH 1976 Storage of fresh cassava I. Post-harvest deterioration and its control. Experimental Agriculture 12:103-111

Booth RH 1977 Storage of fresh cassava (*Manihot esculenta*) II. Simple storage techniques. Experimental Agriculture 13: 119-128

Botella MA, Xu Y, Prabha TN, Zhao Y, Narasimhan ML, Wilson KA, Nielsen SS, Bressan RA and Hasegawa PM 1996 Differential expression of soybean cysteine proteinase inhibitor genes during development and in response to wounding and methyl jasmonate. Plant Physiology 112(3): 1201-1210

Boudet AM 1998 A new view of lignification. Trends in Plant Science 3(2): 67-71

Boudet AM, Lapierre C and Grima-Pettenatti J 1995 Biochemistry and molecular biology of lignification. New Phytology 129: 203-236

Bowles DJ 1997 The wound response of tomato plants: analysis of local and long-range signalling events. Essays in Biochemistry 32: 161-169

Bowles DJ 1990 Defense-related proteins in higher plants. Annual Review Biochemistry 59: 873-907

Bown DP, Bolwell GP and Gatehouse JA 1993 Characterisation of *potato (Solanum tuberosum* L.) extensins: a novel extensin-like cDNA from dormant tubers. Gene 134: 229-233

Boyle JS and Lew AM 1996 An inexpensive alternative to glassmilk for DNA purification. Technical tips online T00003

Bradley DJ, Kjellbom P and Lamb CJ 1992 Elicitor- and wound-inducible oxidative cross-linking of a proline-rich plant cell wall protein. Cell 70: 21-30

Bray EA 1997 Plant responses to water deficit. Trends in Plant Science 2(2): 48-54

Brederode FT, Linthorst HJ and Bol JF 1991 Differential induction of acquired resistance and PR gene expression in tobacco by virus infection, ethephone treatment, UV light and wounding. Plant Molecular Biology 17(6): 1117-1125

Breton F, Sanier C and dAuzac J 1997 Scopoletin production and degradation in relation to resistance of *Hevea brasiliensis* to *Corynespora cassiicola*. Journal of Plant Physiology 151(5): 595-602

Brett C and Waldron K 1996 Topics in plant functional biology: Physiology and biochemistry of plant cell walls, second edition, Chapman & Hall.

Brisson LF, Tenhaken R and Lamb C 1994 Function of oxidative cross-linking of cell wall structural proteins in plant disease resistance. The Plant Cell 6: 1703-1712

Brown JWS 1986 A catalogue of splice junction and putative branch point sequences from plant introns. Nucleic Acid Research 14: 9549-9554

Brownleader MD, McNally PE, Davies GE, Trevan M and Dey PM 1997 Elicitor-induced extensin insolubilization in suspension-cultured tomato cells. Phytochemistry 46(1): 1-9

Bucciaglia PA and Smith AG 1994 Cloning and characterisation of *Tag*1, a tobacco anther β-1,3-glucanase expressed during tetra dissolution. Plant Molecular Biology 24: 903-914

Butland SL, Chow ML and Ellis BE 1998 A diverse family of phenylalanine ammonia-lyase genes expressed in pine trees and cell cultures. Plant Molecular Biology 37: 15-24

Cabello F, Jorrin JV and Tena M 1994 Chitinase and β-1,3-glucanase activities in chickpea (*Cicer arietinum*) – induction of different isoenzymes in response to wounding and ethephon. Physiologia Plantarum 92(4): 654-660

Cabral GB, Carvalho LJCB and Schaal BA 1998 Relationship analysis of closely related species to cassava (*Manihot esculenta* Crantz) based on microsatellite-primed PCR. The Fourth International Scientific Meeting Cassava Biotechnology Network, Brazil

Callahan AM, Morgens PH, Wright P and Nichols KE Jr 1992 Comparision of Pch313 (pTOM13 homolog) RNA accumulation during fruit softening and wounding of two phenotypically different peach cultivars. Plant Physiology 100: 482-488

Callis J 1995 Regulation of protein degradation. Plant Cell 7: 845-857

Calvert LA, Cuervo MI, Ospina MD, Fauquet CM and Ramirez BC 1996 Characterization of common mosaic virus and a defective RNA species. Journal of General Virology 77: 525-530

Carvalho LJCB, Cabral GB and Schaal BA 1998 A starchless clone of cassava. The Fourth International Scientific Meeting Cassava Biotechnology Network, Brazil

Cassab GI and Varner JE 1988 Cell wall proteins. Annual Review Plant Physiology 39: 321-353

Cattivelli L and Bartels D 1990 Molecular cloning and characterisation of cold-regulated genes in barley. Plant Physiology 93: 1504-1510

Cebrat S, Mackiewicz P and Dudek MR 1998 The role of the genetic code in generating new coding sequences inside existing genes. Biosystems 45(2): 165-176

Chapman KA, Delauney AJ, Kim JH and Verma DPS 1994 Structural organization of *de novo* purine biosynthesis enzymes in plants: aminoimidasole ribonucleotide carboxylase and 5-aminoimidasole-4-N-succinocarboxyamide ribonucleotide synthetase cDNAs from *Vigna aconitifolia*. Plant Molecular Biology 24: 389-395

Chavarriaga P, Mancila L, Segovia V and Roca W 1998 Developing cassava transformation strategies against stem borers. The Fourth International Scientific Meeting Cassava Biotechnology Network, Brazil.

Chen CG, Cornish EC and Clarke AE 1992 Specific expression of an extensin-like gene in the style of *Nicotiana alata.* Plant cell 4(9): 1053-1062

Chen J and Varner JE 1985 An extracellular matrix protein in plants: Characterisation of a genomic clone for carrot extensin. EMBO J. 4: 2145-2151

Chen L, Fincher GB and HØj PB 1993 Evolution of polysaccaride hydrolase substrate specificity. The Journal of Biological Chemistry 268(18): 13318-13326

Chen L, Marmey P, Taylor NJ, Brizard J-P, Espinoza C, D'Cruz P, Huet H, Zhang S, de Kochko A, Beachy RN and Fauquet CM 1998 Expression and inheritance of multiple transgenes in rice plants. Nature Biotechnology 16: 1060-1064

Cheung GP, Rosenblum IY and Sallach HJ 1968 Comparative studies of enzymes related to serine metabolism in higher plants. Plant Physiology 43:1813-1820

Chye M-L and Cheung K-Y 1995 β-1,3-glucanase is highly expressed in laticifers of *Hevea brasiliensis*. Plant Molecular Biology 19: 397-402

CIAT 1992 Cassava programme 1987-1991. Working document 116

Ciechanover A and Schwartz A 1994 The ubiquitin-mediated proteolytic pathway: Mechanisms of recognition of the proteolytic substrate and involvement in the degradation of native cellular proteins. FASEB Journal 8: 182-191

Ciechanover A, Finley D and Varshavsky A 1984 Ubiquitin dependence of selective protein degradation demonstrated in the mammalian cell cycle mutant ts85. Cell 37: 57-66

Cock JH 1985 Cassava: new potential for a neglected crop. London, Westview Press.

Collinge DB and Slusarenko AJ 1987 Plant gene expression in response to pathogens. Plant Molecular Biology 9: 389-410

Constabel CP, Bergey DR and Ryan CA 1995 Systemin activates synthesis of wound-inducible tomato leaf polyphenol oxidase via the octadecanoid defense signalling pathway. Proceedings of the National Academy of Sciences of USA 92: 407-411

Cooke R and Cock JH 1989 Cassava crops up again. New Scientists 17: 63-68

Corbin DR, Sauer N and Lamb CJ 1987 Differential regulation of a hydroxyproline-rich glycoprotein gene family in wounded and infected plants. Molecular Cell Biology 7: 4337-4334

Cordereiro CM, Abadie T, Fukuda WMG, Barreto JF and Burtle ML 1998 The Brazilian core collection of cassava. The Fourth International Scientific Meeting Cassava Biotechnology Network, Brazil.

Cornish K and Zeevaart JAD 1985 Abscisic acid accumulation by roots of *Xanthium strumarium* Mill. in relation to water stress. Plant Physiology 79:653-658

Cottle W and Kolattukudy PE 1982 Biosynthesis, deposition, and partial characterisation of potato suberin phenolics. Plant Physiology 69:393-399

Coursey DG and Booth RH 1977 Root and tuber crops. Food Crops of the Lowland Tropics. Edited by Leakley and Wills JB. Oxford Univerisity Press. 75-83

Cramer CL, Edwards K, Dron M, Liang X, Dildine SL, Bolwell GP, Dixon RA, Lamb CJ and Schuch W 1989 Phenylalanine ammonia-lyase gene organisition and structure. Plant Molecular Biology 12: 367-383

Creelman RA and Muller JE 1997 Biosynthesis and action of jasmonates in plants. Annual Review of Plant Physiology and Plant Molecular Biology 48: 355-381

Data ES and Quevedo MA 1984 Pruning techniquess affecting the root quality of cassava at harvest and subsequent storage. In: Tropical Root Crops: Post-harvest physiology and proceeding, pp127-143

de Loose M, Alliotte A, Gheysen G, Genetello C, Gielen J, Soetaert P, Van Montagu M and Inze D 1988 Primary structure of a hormonally regulated β-1,3-glucanase of *Nicotiana plumbaginifolia*. Gene 70: 13-23

Diallinas G and Kanellis AK 1994 A phenylalanine ammonia-lyase gene from melon fruit-cDNA cloning, sequence and expression in response to development and wounding. Plant Molecular Biology 26(1): 473-479

Dixon RA and Lamb CJ 1990 Molecular communication in interactions between plants and microbial pathogens. Annual Review Plant Physiology and Plant Molecular Biology 41: 339-367

Dixon RA, Maria JH and Paiva NL 1995 The isoflavonoid phytoalexin pathway: from enzymes to genes to transcription factors. Physiologia Plantarum 93: 385-392

Dixon RA and Paiva NL 1995 Stress-induced phenylpropanoid metabolism. Plant cell 7: 1085-1097

Domingo C, Conejero V and Vera P 1994 Genes encoding acidic and basic class III β-1,3-glucanase are expressed in tomato plants upon viroid infection. Plant Molecular Biology 24: 725-732

Dorey S, Baillieul F, Pierrel MA, Saindrenan P, Fritig B and Kauffmann S 1997 Spatial and temporal induction of cell death, defense genes, and accumulation of salicylic acid in tobacco leaves reacting hypersentitively to a fungal glycoprotein elicitor. Molecular Plant-microbe Interaction 10(5): 646-655

Driscoll J and Finley D 1992 A controlled breakdown: antigen processing and the turnover of viral proteins. Cell 68: 823-825

Ebener W, Fowler TJ, Suzuki H, Shaver J and Tierney ML 1993 Expression of DcPRP1 is linked to carrot storage root formation and induced by wounding and auxin treatment. Plant Physiology 101:259-265

Ecker JR and Davis RW 1987 plant defence genes are regulated by ethylene. Proceedings of the National Academy of Sciences of USA 84: 5202-5206

Edwards K, Cramer CJ, Bolwell GP, Dixon RA, Schuch W and Lamb CJ 1985 Rapid transient induction of phenylalanine ammonia-lyase mRNA in elicitor-treated bean cells. Proceedings of the National Academy of Sciences of USA. 82: 6731-6735

Edwards R, Stones SM, GutierrrezMellado MC and Jorrin J 1997 Characterization and inducibility of a scopoletin-degrading enzyme from sunflower. Phytochemistry 45(6):1109-1114

Eferl R, Lehner M, Kenner L, Kapfer I, Gurtl B and Zatloukal K 1997 Evalution of different RNA standards for quantitative reverse transcription PCR: possible pitfalls and strategies for avoidance. Technical Tips Online. T01214.

Eichholz R, Harper J, Felix G and Mains JF 1983 Evidence for an abundant 33,000-dalton polypeptide regulated by cytokinins in cultured tobacco tissues. Planta 158: 410-415

Epstein L and Lamport DTA 1984 An intramolecular linkage involving isodityrosine in extensin. Phytochemistry 23: 1241-1246

Escobar RH, Palacio JD, Rangel MP and Roca WM 1998 Cassava cryopreservation of shoot tips through encapsulation-dehydration technique. The Fourth International Scientific Meeting Cassava Biotechnology Network, Brazil.

Esquerre-Tugaye MT, Laffite C, Mazare D, Toppan A and Touze A 1979 Cell surfaces in plant -microorganism interactions: evidence for the accumulation of hydroxyproline-rich glycoproteins in the cell wall of diseased plants as a defense mechanism. Plant Physiology 64: 320-326

Evans IM, Gatehouse LN, Gatehouse JA, Yarwood JN, Boulter D and Croy RRD 1990 The extensin gene family in oilseed rape-characterisation of sequences of representitive members of the family. Molecular and General Genetics 223-273

ExPASy PROSITE 1998 http://expasy.hcuge.ch/

Faktor O, Kooter JM, Loake GJ, Dixon RA and Lamb CJ 1997 Differential utilization of regulatory *cis*-elements for stress-induced and tissue-specific activity of a French bean chalcone synthase promoter. Plant Science 124(2): 175-182

FAO 1997 FAOSTAT Database http://apps.fao.org/

Farmer EE 1994 Fatty acid signalling in plants and their associated microorganisms. Plant Molecular Biology 26: 1423-1437

Farmer EE and Ryan CA 1990 Interplant communication: Airborne methyl jasmonate induces synthesis of proteinase inhibitors in plant leaves. Proceedings of the National Academy of Sciences of USA 87: 7713-7716

Felix G and Mains JF 1985 Purification, immunoassay and characterisation of an abundant, cytokinin-regulated polypeptide in cultured tobacco tissues. Planta 164: 423-428

Fregene M, Angel F, Suarez MC, Janneth G, Chavarriaga P, Roca W, Bonierbale M and Tohme J 1998 Genome mapping in cassava improvement: Challenges, achievements and opportunities. The Fourth International Scientific Meeting Cassava Biotechnology Network, Brazil.

Fujinami K, Tanahashi N, Tanaka K, Ichihara A, Cejka Z, Baumeister W, Miyawaki M, Sato T and Nakagawa H 1994 Purification and characterisation of the 26S protesome from spinach leaces. Journal of Biological Chemistry. 269: 25905-10

Gaffney T, Friedrich L, Vernooij B, Negrotto D, Nye G, Uknes S, Ward E, Kessmann H and Ryals J 1993 Requirement of salicylic acid for the induction of systemic acquired resistance. Science 261: 754-756

Garbarino JE, Rockhold DR and Belknap WR 1992 Expression of stress-responsive ubiquitin genes in potato tubers. Plant Molecular Biology 20: 235-244

GarciaMuniz N, MartinezIzquierdo JA and Puigdomenech P 1998 Induction of mRNA accumulation corresponding to a gene encoding a cell wall hydroxyproline-rich glycoprotein by fungal elicitors. Plant Molecular Biology 38(4): 623-632

Genschik P, Jamet E, Philipps G, Parmentier Y, Gigot C and Fleck J 1994 Molecular characterization of a b-type proteasome subunit from Arabidopsis thaliana co-expressed at a high level with an a-type subunit early in the cell syvle. Plant Journal 6: 537-546

Gloria LA and Uritani I 1984 Changes in β-carotene content of golden yellow cassava in relation to physiological deterioration. In: Tropical Root Crops: Postharvest Physiology and Processing. I Uritani, ED Reyes, eds. JSSP, Tokyo. pp. 163-168

Glotzer M, Murray AW and Kirschner MW 1991 Cyclin is degraded by the ubiquitin pathway. Nature 349: 132-138

Goldman MH de S, Pezzotti M, Seurinck J and Mariani C 1992 Developmental expression of tobacco pistil-specific genes encoding novel extensin-like proteins. Plant Cell 4:1041-1051

Gomez-Vasquez R, Day R, Beeching JR and Cooper RM 1998 Biochemical components of disease resistance of cassava. The Fourth International Scientific Meeting Cassava Biotechnology Network, Brazil.

Gottschalk TE, Mikkelsen JD, Nielsen JE, Nielsen KK and Brunstedt J 1998 Immunolocalization and characterization of a β-1,3-glucanase from sugar beet, deduction of its primary structure and nucleotide sequence by cDNA and genomic cloning. Plant Science 132: 153-167

Grierson D, Maunders MJ, Slater A, Ray J, Bird CR, Schuch W, Holdsworth ML, Tucker GA and Knapp JE 1986 Gene expression during tomato ripening. Philosophical Transactions of the Royal Society. Series B 314: 399-410

Guo Y, Delseny M and Piugdomenech P 1994 mRNA accumulation and promoter activity of the gene coding for a hydroxyproline-rich glycoprotein in *Oryza sativa*. Plant Molecular Biology 25: 159-165

Guypers B, Schmelzer E and Hahlbrock K 1988 In situ localization of rapidly accumulated phenylalanine ammonia-lyase mRNA around penetration sites of *Phytophthora infestans* in potato leaves. Molecular Plant-Microbe interaction 1:157-160

Halbrock K and Scheel D 1989 Physiology and molecular biology of phenylpropanoid metabolism. Annual Review Plant Physiology and Plant Molecular Biology 40:347-369

Ham KS, Kauffmann S, Albersheim P and Davill AG 1991 Host-pathogen interactions.39. A soybean pathogenesis-related protein with beta-1,3-glucanase activity releases phytoalexin elicitor-active heat-stable fragments from fungal walls. Molecular plant-microbe interactions 4(6): 545-552

Hamilton AJ, Brown S, Han YH, Ishizuka M, Lowe A, Solis AGA and Grierson D 1998 A transgene with repeated DNA causes high frequency, post-transcriptional suppression of ACC-oxidase gene expression in tomato. Plant Journal 15(6): 737-746

Hanks SK and Quinn AM 1991 Protein kinase catalytic domain sequence data base: identification of conserved features of primary structure and classification of family members. Methods Enzymology 200: 38-62

Handley PM, Mueckler M, Siegel NR, Ciechanover A and Schwartz AL 1991 Molecular-cloning, sequence, and tissue distribution of the human ubiquitin-activating enzyme-E1. Proceedings of the National Academy of Sciences of the USA 88(1): 258-262

Harrison BD, Zhou X, Otim-Nape GW, Liu Y and Robinsson DJ 1997 Role of novel type of double infection in the germinivirus-induced epidemic of severe cassava mosaic in Uganda. Annals of Applied Biology 131: 437-448

Hass AL and Rose IA 1982 The mechanism of ubiquitin activating enzyme. Journal of Biological Chemistry. 257: 10329-37

Hatfield PM, Vierstra RD 1989 Ubiquitin-dependent proteolytic pathway in wheat-germ - isolation of multiple forms of ubiquitin-activating enzyme, E1. Biochemistry 28(2): 735-742

Hatfield PM, Vierstra RD 1992 Multiple forms of ubiquitin-activating enzyme E1 from wheat. Identification of an essential cysteine by in vitro mutagenesis. Journal of Biological Chemistry 267: 14799-14803

Hedrick SA, Bell JN, Boller T and Lamb CJ 1988 Chitinase cDNA cloning and mRNA induction by fungal elicitor, wounding and infection. Plant Physiology 86: 182-186

Henderson SJ and Friend J 1979 Increase in PAL and lignin-like compounds as race-specific resistance responses of potato tubers to *Phytophthora infestans*. Journal of Phytopathology 94: 323-34

Hershey C, Iglesias C, Iwanaga M and J Tohme 1992 Definition of a core collection for cassava. Proceedings of the First International Meeting Cassava Biotechnology Network. pp145-156

Hildmann T, Ebneth M, Pena-Cortez H, Sanchez-Serrano J, Willmitzer L and Prat S 1992 General roles of abscisic and jasmonic acids in gene activation as a result of mechanical wounding. Plant Cell 4: 1157-1170

Hippesanwald S, Marticke KH, Kieliszewski MJ and Somerville SC 1994 Immunogold localization of THRGP-like epitopes in the haustorial interface of obligate, biotrophic fungi on monocots. Protoplasma 178(3-4):138-155

Hirose S 1986 Physiological studied on post-harvest deterioration of cassava roots. Japanese Agriculture Research Quarter 19(4): 241-252

Hirose S and Data ES 1984 Physiology of post-harvest deterioration of cassava roots. In: Tropical Root Crops. Post-harvest Physiology and Processing. Uritani I and Reyes ED eds, Tokyo, Japan Scientific Societies Press. pp31-51

Hirose S, Data ES and Quevedo MA 1984 Changes in respiration and ethylene production in cassava roots in relation to post-harvest deterioration. In: Tropical Root Crops. Post-harvest Physiology and Processing. Uritani I and Reyes ED eds, Tokyo, Japan Scientific Societies Press. Pp83-98

Hirsinger C, Permentier Y, Durr A, Fleck J and Jamet E 1997 Characterization of a tobacco extensin gene and regulation of its gene family in healthy plants and under various stress conditions. Plant Molecular Biology 33: 279-289

Hochstrasser M, Ellision MJ, Chau V and Varshavsky A 1991 The short-lived MATα2 transcriptional regulator is ubiquitinated in vivo. Proceedings of the National Academy of Sciences of USA. 88:4606-4610

Hong JC, Cheong YH, Nagao RT, Bahk JD, Cho MJ and Key JL 1994 Isolation and characterization of three soybean extensin cDNAs. Plant Physiology 104(2): 793-796

Howles PA, Sewalt VHJ, Paiva NL, Elkind Y, Bate NJ, Lamb C and Dixon RA 1996 Over-expression of phenylalanine ammonia-lyase in transgenic tobacco plants reveals control points for flux into phenylpropanoid biosynthesis. Plant Physiology 112: 1617-1624

Huang J, Bachem C, Jacobsen E and Visser R 1998 Molecular analysis of post-harvest deterioration in cassava. The Fourth International Scientific Meeting Cassava Biotechnology Network, Brazil.

Hughes J, Carvalho F and Hughes M 1994 Purification, characterization, and cloning of a α-hydroxynitrile lyase from cassava (*Manihot esculenta* Crantz). Archives Biochemistry and Biophysics 311: 496-502

Hughes MA, Brown K, Pancoro A, Murray BS, Oxtoby E and Hughes J 1992 A molecular and biochemical analysis of the structure of the cyanogenic β-glucosidase, linamarase, from cassava (*Manihot esculenta*, Crantz). Archives of Biochemistry and Biophysics 295: 273-279

Hunt LA, Wholey DW and Cock JH 1977 Growth physiology of cassava (*Manihot esculenta* Crantz). Field Crop Abstract 30(2): 77

Ishizuka M, Yamada F, Tanaka Y, Takeuchi Y and Imaseki H 1991 Sequential induction of messenger-RNAs for phenylalanine ammonia-lyase in slices of potato-tuber. Plant and Cell Physiology 32(1): 57-64

Jach J, Gornardt B, Mundy J, Logemann J, Pinsdorf P, Leah R, Schell J and Maas C 1995 Enhanced quantitative resistance against fungal disease by combination of different barley antifungal proteins in transgenic tobacco. Plant Journal 8(1): 97-109

Jackson M 1997 Hormones from roots as signals for the shoots of stressed plants. Trends in Plant Science 2(1):22-28

Jangaard NO 1974 The characterisation of phenylalanine ammonia-lyase from several plant species. Phytochemistry 13: 1765-1768

Jones DA 1998 Why are so many food plants cyanogenic? Phytochemistry 47(2): 155-162

Jones DH 1984 Phenylalanine ammonia-lyase: regulation of its induction, and its role in plant development. Phytochemistry 23(7): 1349-1359

Joos H-J and Hahlbrock K 1992 Phenylalanine ammonia-lyase in potato. European Journal of Biochemistry 204: 621-629

Joshi CP, Zhou H, Huang X and Chiang V 1997 Context sequences of translation initiation codon in plants. Plant Molecular Biology 35: 993-1001

Kahn TL, Fender SE, Bray EA and Oconnell MA 1993 Characterisation of expression of drought- and ABA-regulated tomato genes in the drought-resistant species *Lycopersicon pennellii*. Plant Physiology 103:597-605

Karin M and Hunter T 1995 Transcriptional control by protein phosphorylation: signal transduction from the cell surface to the nucleus. Current Biology 5:747-757

Karrer EE, Beachy RN and Holt CA 1998 Cloning of tobacco genes that elicit the hypersensitive response. Plant Molecular Biology 36(5): 681-690

Kawaka A, Kawamoto T, Ohta H, Sekine M, Takano M and Shinmgo A 1994 Wound induced expression of horse radish peroxidase gene. Plant Cell Report 13: 149-154

Kawalleck P, Schmeltzer E, Hahlbrock K and Somassich IE 1995. Two pathogen-responsive genes in parsley encode a tyrosine-rich hydroxyproline-rich glycoprotein(HRGP) and an anionic peroxidase. Molecular general genetics 24: 444-452

Kay GF, Ashworth A, Penny GD, Dunlop M, Swift S, Brockdorff N and Rastan S 1991 A candidate spermatogenesis gene on the mouse y-chromosome is homologous to ubiquitin-activating enzyme-E1. Nature 354(6353): 486-489

Ke D and Saltveit ME 1989 Wound-induced ethylene production, phenolic mechanism and susceptibility to russet spotting in iceberg lettuce. Physiologia Plantarum 76: 412-418

Keefe D, Hinz U and Meins F 1990 The effect of ethylene on the cell-type-specific and intracellular localization of β-1,3-glucanase and chitinase in tobacco leaves. Planta 182: 43-51

Keen NT and Yoshikawa M 1983 β-1,3-glucanase from soybean releases elicitor-active carbohydrates from fungus cell walls. Plant Physiology 71: 460-465

Kernan A and Thornburg RW 1989 Auxin levels regulate the expression of a wound-inducible proteinase-inhibitorII-chloramphenicol acetyl transferase gene fusion *in vitro* and *in vivo*. Plant Physiology 91:73

Keys AJ 1980 Synthesis and interconversion of glycine and serine. In Amino acids and derivatives. Miflin BJ (ed), Academic Press, New York. pp359-374

Kieliszewski, MJ and Lamport, TA 1994. Extensin: repetitive motifs, functional sites, post-translational codes, and phylogeny The Plant Journal 5(2): 157-172

Kim JH, Delauney AJ and Verma DPS 1995 Control of de novo purine biosynthesis genes in ureide-producing legumes: induction of glutamine phosphoribosylpyrophosphate amidotransferase gene and characterisation of its cDNA from soybean and *Vigna aconitifolia*. Plant Journal 7: 77-86

Kimelman D and Kirschner MW 1989 An antisense mRNA directs the covalent modification of the transcript encoding fibroblast growth factor in *Xenopus oocytes*. Cell 59: 687-696

King NLR and Bradbury JH 1995 Bitterness of cassava: Identification of a new apicosyl glucoside and other compounds that affect its bitter taste. Journal of Scientific Food Agriculture 68: 223-230

Klessing DF and Malamy J 1994 The salicylic acid signal in plants. Plant Molecular Biology 26: 1439-1458

Knee R and Murphy PR 1997 Regulation of gene expression by natural antisense RNA transcripts. Neurochemistry International 31(3): 379-392

Kolattukudy PE 1980 Biopolyester membranes of plants: cutin and suberin. Science 208: 990-1000

Koonberg-Roos P, Tormkangas K and Ostman A 1991 Primary structure of a barley grain aspartic proteinase. A plant aspartic proteinase resembling cathepin D. European Journal of Biochemistry 202(3): 1021-1027

Kudo M, Sugasawa K, Hori T, Enomoto T, Hanaoka F, Ui M 1991 Human ubiquitin-activating enzyme (E1) - compensation for heat- labile mouse E1 and its gene localization on the x-chromosome. Experimental Cell Research 192(1): 110-117

Lamb C and Dixon RA 1997 The oxidative burst in plant disease resistance. Annual Review Plant Physiology and Plant Molecular Biology 48: 251-275

Lamb CJ 1977 Trans-cinnamic acid as a mediator of the light-stimulated increase in hydroxycinnamoyl-CoA:quinate hydroxycinnamoyl transferase. FEBS Letters 75: 37-40

Lamport DTA 1967 Hydroxyproline-*O*-glycosidic linkage of the plant cell wall glycoprotein extensin. Nature 216: 1322-1324

Lamport DTA 1977 Structure, biosynthesis and significance of cell wall glycoproteins. Recent Advances Phytochemistry 11: 79-115

Lamport, DTA 1965 cell wall proteins. Advanced Botany Research 2: 251-281

Lawton K, Ward E, Payne G, Moyer M and Ryals J 1992 Acidic and basic class III chitinase mRNA accumulation in response to TMV infection of tobacco. Plant Molecular Biology 19: 735-743

Lawton MA and Lamb CJ 1987 Transcriptional activation of plant defense genes by fungal elicitor, wounding, and infection. Molecular Cell Biology 7: 335-341

Leach JE, Cantrell MA and Sequeira L 1982 A hydroxyproline-rich bacterial agglutinin from potato: extraction, purification and characterisation. Plant Physiology 92: 327-333

Lee MY and Kim SS 1998 Characterization of a low molecular weight minor anionic isoperoxidase A(3n) from radish. Journal of Biochemistry and Molecular Biology 31(6): 548-553

Lee S, Nazar RN, Powell DA and Robb J 1992 Reduced PAL gene suppression in *Verticillium*-infected resistant tomato. Plant Molecular Biology 18: 345-352

Legris G, Potrykus I and Puonti-Kaerlas J 1998 Engineering insect resistance in cassava. The Fourth International Scientific Meeting Cassava Biotechnology Network, Brazil

Leja M, Mareczek A, Wojciechowska R and Rozek S 1997 Phenolic metabolism in root slices of selected carrot cultivars. ACTA Physiological Plantarum 19(3): 319-325

LeJohn HB, Cameron LE, Yang B and Rennie SL 1994 Molecular characterisation of an NAD-specific glutamate dehydrogenase gene inducible by L-glutamine. The Journal of Biological Chemistry 269(6): 4523-4531

Leone A, Costa A, Grillo S, Tucci M, Horvath I and Vigh L 1996 Acclimation to low water potential determines changes in membrane fatty acid composition and fluidity in potato cells, Plant Cell Environment 19:1103-1109

Li H, Han Y and Beeching JR 1998 Phenylalanine ammonia-lyase gene organisation, structure and activity in cassava. The Fourth International Scientific Meeting Cassava Biotechnology Network, Brazil.

Li HQ, Santter C, Potrykus I and Puonti-Kaerlas J 1996 Genetic transformation of cassava (*Manihot esculenta* Crantz) Nature Biotechnology 14(6): 736-740

Li HQ, Potrykus I and Puonti-Kaerlas J 1998 Engineering leaf life length in cassava. The Fourth International Scientific Meeting Cassava Biotechnology Network, Brazil.

Li X-B, Kieliszewski M, and Lamport DTA 1990 A chenopos extensin lacks repetitive tetrahydroxyproline blocks. Plant Physiology 92:327-333

Liang XW, Dron M, Cramer CL, Dixon RA and Lamb CJ 1989 Differential regulation of phenylalanine ammonia-lyase genes during plant development and by environmental cues. Journal of Biological Chemistry 264(24): 14486-14492

Liddle S, Hughes J and Hughes MA 1997 Analysis of a cassava root specific β-glycosidase promoter. African Journal of Root and Tuber Crops 2(1-2):158-162

Lindsey JS 1991 Self-assembly in synthetic routes to molecular devices. Biological principles and chemical perspectives: A review. New Journal of Chemistry 15: 153-180

Linthorst HJM, Melchers LS, Mayer A, van Roekel JSC, Cornelissen BJC and Bol JF 1990 Analysis of gene families encoding acidic and basic β-1,3-glucanases. Proceedings of the National Academy of Sciences of USA 87: 8756-8760

LoaizaVelarde J, TomasBarberan FA and Saltveit ME 1997 Effect of intensity and duration of heat-shock treatment on wound-induced phenolic metabolism in iceberge lettuce. Journal of the American Society for Horticultural Science 122 (6): 873-877

Logemann E, Parniske M and Hahlbrock K 1995 Modes of expression and common structural features of the complete phenylalanine ammonia-lyase gene family in parsley. Proceedings of the National Academy of Sciences of USA 92: 5905-5909

Lozano LC, Cock JH and Castano J 1978 New developments in cassava storage. CIAT Series CE-14: 135-141

Ludford PM 1995 Postharvest hormone changes in Vegetables and Fruit. In Davies PJ ed. Plant Hormones. Physiology, Biochemistry and Molecular Biology, 2ed. Kluwer Academic Publishers, pp725-750

Lulai EC and Corsini DL 1998 Differential deposition of suberin phenolic and aliphatic domains and their roles in resistance to infection during potato tuber (*Solanum tuberosum* L.) wound-healing. Physiological and Molecular Plant Pathology 53(4): 209-222

Lusso M and Kuc J 1996 The effect of sense and antisense expression of the PR-N gene for β-1,3-glucanase on disease resistance of tobacco to fungi and viruses. Physiological and Molecular Plant Pathology 49: 267-283

Maeshima M, Sasaki T and Asahi T 1985 Characterisation of major proteins in sweet potato tuberous roots. Phytochemistry 24(9): 1899-1902

Maher EA, Bate NJ, Ni W, Elkind Y, Dixon RA and Lamb CJ 1994 Increased disease susceptibility of transgenic tobacco plants with suppressed levels of preformed phenylpropanoid products. Proceedings of the National Academy of Sciences of USA 91: 7802-7806

Marriott J, Been BO and Perkins C 1978 The aetiology of vascular discoloration in cassava roots after harvesting: association with water loss from wounds. Physiologia Plantarum 44: 38-42

Mason HS and Mullet JE 1990 Expression of 2 soybean vegetative storage protein genes during development and in response to water deficit, wounding, and jasmonic acid. Plant Cell 2(6): 569-579

Masoud SA, Zhu Q, Lamb CJ, Dixon RA 1996 Constitutive expression of an inducible beta-1,3-glucanase in alfalfa reduces desease severity caused by the oomycete pathogen *phytophthora megasperma f sp. medicaginis*, but does not reduce disease severity of chitin-containing fungi. Transgenic research 5(5): 313-323

Mauch F, Meehl JB and Staehelin LA 1992 Ethylene-induced chitinase and β-1,3-glucanase accumulate specifically in the lower epidermis and along vascular strands of bean leaves. Planta 186(3): 367-375

Mauch f and Staehelin LA 1989 Functional implications of the subcellular-localization of ethylene-induced chitinase and beta-1,3-glucanase in bean-leaves. Plant Cell 1(4): 447-457

Mcgrath JP, Jentsch S and Varshavsky A 1991 UBA1 - an essential yeast gene encoding ubiquitin-activating enzyme. EMBO Journal 10(1): 227-236

McGurl B, Orozoco-Cardens M, Pearce G and Ryan CA 1994 Overexpression of the prosystemin gene in transgenic tomato plants generates a systemic signal that constitutively induces proteinase inhibitor synthesis. Proceedings of the National Academy of Sciences of USA 1191(21): 9799-9802

McGurl B, Pearce G, Orozco-Cardenas M and Ryan CA 1992 Structure, expression, and antisense inhibition of the systemin precursur gene. Science 255: 1570-1573

McMahon J, White W and Sayre RT 1995 Cyanogenic glycosides in cassava (*Manihot esculenta* Crantz). Journal of Experimental Botany 46:731-741

McNeil M, Darvill AG, Fray SC, and Albersheim P 1984 Structure and function of the primary cell walls of plants. Annual Review Biochemistry 53: 625-663

Mehdy MC and Lamb CJ 1987 Chalcone isomerase cDNA cloning and mRNA induction by fungal elicitor, wounding, and infection. EMBO Journal 6: 1527-1533

Meins JF, Nuehaus JM, Sperisen C and Ryals J 1992 The primary structure of plant pathogenesis-related glucanohydrolases and their genes. In: Boller T and Meins JF (eds.), Genes Involved in Plant Defense. Springer, Vienna, pp.245-282

Melchers LS, Selabuurlage MB, Vloemans SA, Woloshuk CP, Vanroekel JSC, Pen J, Vandenelzen PJM and Cornelissen BJC 1993 Extracellular targeting of the vacuolar tobacco proteins-AP24, chitinase and β-1,3-glucanase in transgenic plants. Plant Molecular Biology 21(4): 583-593

Mellon JE and Helgeson JP 1982 Interaction of a hydroxyproline-rich glycoprotein from tobacco callus with potential pathogens. Plant Physiology 70: 401-405

Memelink J, Swords KMM, de Kam RJ, Schilperoort RA, Hoge JHC and Staehelin LA 1993 Structure and regulation of tobacco extensin. The Plant Journal 4(6): 1011-1022

Mendes RA, de Goes M, Labuto LBD and Cardoso LD 1998 *In vitro* conservation of cassava germplasm at Cenargen. The Fourth International Scientific Meeting Cassava Biotechnology Network, Brazil.

Menkens AE, Schindler U and Cashmore AR 1995 The G-box: A ubiquitous regulatory DNA element in plants bound by the GBF family of bZIP proteins. Trends in Biochemistry Science 20:506-510

Menossi M, Martinez-izquierdo JA and Puigdomenech P 1997 Promoter tissue specific activity and ethylene control of the gene coding for the maize hydrocyproline-rich glycoprotein in maize cells transformed by particle bombarment. Plant Science 125:189-200

Mizutani M, Ohta D and Sato R 1997 Isolation of a cDNA and a genomic clone encoding cinnamate-4-hydroxylase from Arabidopsis and its expression manner in planta. Plant Physiology 113(3): 755-763

Montaldo A 1973 Vascular streaking of cassava root tubers. Tropical Science 15(1): 41-46

Munyikwa TRI, Raemakers CCJM, Schreuder M, Kreuze J, Suurs L, Kok R, Rozeboom M, Jacobsen E and Visser RGF 1998 Introduction and expression of antisense ADPG-pyrophos-phorylase of cassava leads to decreased levels of starch and increased levels of sugars. The Fourth International Scientific Meeting Cassava Biotechnology Network, Brazil.

Mur LA, Bi YM, Darby RM, Firek S and Draper J 1997 Compromising early salicylic acid accumulation delays the hypersensitive response and increases viral dispersal during lesion establishment in TMV-infected tobacco. Plant Journal 12(5): 1113-1126

Neuhard J and Nygaard P 1986 Purine and Pyrimidines. In FC Neidhardt, ed, Escherichia coli and Salmonella typhimurium: Cellular and Molecular Biology. American Society of Microbiology, Washington, DC, pp 445-473

Neuhaus JM, Sticher L, Meins F and Boller T 1991 A short C-terminal sequence is necessary and sufficient for the targeting of chitinase to the plant vacuole. Proceedings of the National Academy of Sciences of USA 88(22): 10362-10366

Niderman T, Genetet I, Bruyere T, Gees Rene, Stintzi A, Legrand M, Fritig B and Mosinger E 1995 Pathogen-related PR-1 proteins are antifungal: isolation and characterisation of three 14Kda proteins of tomato and of a basic PR-1 of tobacco with inhibitory activity against *Phytophthora infestans*. Plant Physiology 108: 17-27

Nielsen KK, Bojsen K, Collinge DB and Mikkelsen JD 1994 Induced resistance in sugar beet against *Cercospora beticola*: induction by dichloroisonicotinic acid is independent of chitinase and β-1,3-glucanase transcript accumulation. Physiological and Molecular Plant Pathology 45: 89-99

Nolte P, Sear GA, Guanestad NC and Nenningson PJ 1993 Detection and identification of fluorescent compounds in potato-tuber tissue with corky patch syndrome. American Potato Journal 70(9):649-666

Noon RA and Booth RH 1977 Nature of post-harvest deterioration of cassava roots. Transaction of British Mycology Society 69(2): 287-290

O'Donnell PJ, Calvert C, Atzorn R, Wasternack A, Leyser HMO and Bowels DJ 1996 Ethylene as a signal mediating the wound response of tomato plants. Science 274: 1914-1917

Olson MM and Roseland CR 1991 Induction of the comarins scopoletin and ayapin in sunflower by insect feeding stress and effects of coumarin on the feeding of sunflower beetle (*Coleoptera chrysomelidae).* Environmental Entomology 20(4):1166-1172

Olsen KM and Schaal B 1999 Evidence on the origin of cassava: Phylogeograghy of Manihot esculenta. Proceedings of the National Academy of Sciences of USA 96: 5586-5591

Onwueme IC 1978 The Tropical Tuber Crops: Yam, Cassava, Sweet potato, Cocoyams. John Wiley and Sons, Chichester.

Oosterhaven K, Hartmans K, Scheffer JJC and Van Der Plas LHW 1995 S-Carvone inhibits PAL activity and suberization during wound healing of potato tubers. Journal of Plant Physiology 146:288-294

Ori N, Sessa G, Lotan T, Himmelhoch S and Fluhr R 1990 A major stylar matrix polypeptide (sp41) is a member of the pathogenesis-related proteins superclass. EMBO J 9(11): 3429-3436

Orozco-Cardenas M and Ryan CA 1999 Hydrogen peroxidase is generated systemically in plant leaves by wounding and systemin via the octadecanoid pathway. Proceedings of the National Academy of Sciences of USA 96: 6553-6557

Oyenuga VA 1968 Nigeria's Foods and Feeding Stuffs. University of Ibadan Press, Ibadan, Nigeria pp99

Pallas JA, Paiva NL, Lamb C and Dixon RA 1996 Tobacco plants epigenetically suppressed in phenylalanine ammonia-lyase expression do not develop systemic acquired resistance in response to infection by tobacco mosaic virus. The Plant Journal 10(2): 281-293

Pancoro A and Hughes MA 1992 *In situ* localization of cyanogenic glucosidase (linamarase) gene expression in leaves of cassava (*Manihot esculenta* Crantz) using non-isotopic riboprobes. Plant Journal 2: 821-827

Parmentier Y, Durr A, Marbach J, Hirsinger C, Criqui MC, Fleck J and Jamet E 1995 A novel wound-inducible extensin gene is expressed early in newly isolated protoplasts of *nicotiana-sylvestris*. Plant Molecular Biology 29(2): 279-292

Pautot V, Holzer FM, Reisch B and Walling L 1993 Leucine aminopeptidase: an inducible component of the defense response in tomato. Proceedings of the National Academy of Sciences of USA 90: 9906-9910

Pawlowski K, Kunze R, De Vries S and Bisseling T 1994 Isolation of total, poly(A) and polysomal RNA from plant tissue. Plant Molecular Biology Manual (D5). Kluwer Academic Publishers, pp1-13

Payne G, Ahl P, Moyer M, Happer A, Beck J, Meins F and Ryals J 1990 Isolation of complementary-DNA clones encoding pathogenesis-related protein-P and Protein-Q, 2 acidic chitinase from tobacco. Proceedings of the National Academy of Sciences of USA 90

Peña-Cortês H, Fisahn J and Willmitzer L 1995 Signals involved in wound-induced proteinase inhibitor II gene expression in tomato and potato plants. Proceedings of the National Academy of Sciences of USA 92: 4106-4113

Peña-Cortês H, Sanchez-Serrano J, Rocha-Sosa M and Willmitzer L 1988 Systemic induction of proteinase inhibitor II gene expression in potato plants by wounding. Planta 174:84-89

Peña-Cortês H, Prat S, Atzorn R, Wasternack C and Willmitzer L 1996 Abscisic acid-deficient plants do not accumulate proteinase inhibitor II following systemin treatment. Planta 198: 447-451

Pellegrini L, Rohfritsch O , Fritig B and Legrand M 1994 Phenylalanine ammonia-lyase in tobacco: Molecular cloning and gene expression during the hypersensitive reaction to tobacco mosaic virus and the response to a fungal elicitor. Plant Physiology 106: 877-886

Pendharker MB and Nair PM 1987 Potato Research 30: 589-601

Pereira LF, Goodwin PH and Erickson L 1999 The role of a phenylalanine ammonia lyase gene during cassava bacterial blight and cassava bacterial necrosis. Journal of Plant Research 112(1105): 51-60

Picton S, JE Gray and Grierson D 1995 Ethylene genes and fruit ripening. In: Plant Hormones-Physiology, Biochemistry and Molecular Biology. 2nd edition. Ed by PJ Davies. Kluwer Academic Publishers. pp372-394

Pieterse CMJ, Risseeuw EP and Davidse LC 1991 An in planta induces gene of Phytophthora infestans codes for ubiquitin. Plant Molecular Biology. 17: 799-811

Plumbley RA and Rickard JE 1991 Post-harvest deterioration of cassava. Tropical Science 31: 295-303

Plumbley RA, Hughes PA and Marriott J 1981 Studies on peroxidases and vascular discoloration in cassava root tissue. Journal of Scientific Food Agriculture 32:723-731

Pogson BJ, Downs CG and Davies KM 1995 Differential expression of 2 1-aminocyclopropane-1-carboxylic acid oxidase genes in Broccoli after harvest. Plant Physiology 108(2): 651-657

Puonti-Kaerlas J 1998 Cassava biotechnology. Biotechnology and Genetic Engineering Review 15:329-362

Ravi V, Aked J and Balagopalan C 1996 Review on tropical root and tuber crops I. Storage methods and quality changes. Critical reviews in Food Science and Nutrition 36(7): 661-709

Reid MS 1995 Ethylene in plant growth, development, and senesence. In: Plant Hormones, Physiology, Biochemistry and Molecular Biology. 2nd edition. Ed by Davies PJ. P486-508

Reilly K, Han Y, Iglesias C and Beeching JR 1998 Oxidative stress related genes in cassava post harvest physiological deterioration. The Fourth International Scientific Meeting Cassava Biotechnology Network, Brazil.

Reynolds PH and Blevis DG 1988 Phosphoserine aminotransferase in soybean root nodules. Plant Physiology 81: 293-293

Rickard JE 1981 Biochemical changes involved in the post-harvest deterioration of cassava roots. Tropical Science 23: 235-237

Rickard JE 1983 The development of occlusions in cassava root xylem vessels. Annal of Botany 52:811-821

Rickard JE 1985 Physiological deterioration of cassava roots Journal of Sciences in Food and Agriculture 36: 167-176

Rickard JE and Coursey DG 1981 Cassava storage. Part 1: Storage of fresh cassava roots. Tropical Science 23: 235-237

Rickard JE, Marriott J and Gahan PB 1979 Occlusions in cassava xylem associated with vascular discoloration. Annal of Botany 43: 523-526

Roa AC, Maya MM, Duque MC, Tohme J, Allem AC and Bonierbale MW 1997 AFLP analysis of relationships among cassava and other *Manihot* species. Theoretical and Applied Genetics 95: 741-750

Roby D, Toppan A and Esquerre-Tugaye MT 1985 Cell surfaces in plant microorganism interactions. V. Elicitors of fungal and plant origin trigger the synthesis of ethylene and hydroxyproline-rich glycoproteins in plants. Plant Physiology 77: 700-704

Rodriguez MX, Buschmann H, Iglesias C and Beeching JR 1998 Production of phytoalexins in cassava root during post-harvest physiological deterioration. The Fourth International Scientific Meeting Cassava Biotechnology Network, Brazil.

Rogers DJ and Appan SG 1973 *Manihot, Manihotoides (Euphorbiaceae)*. Flora Neotropica (Monogragh No. 13) Hafner Press, New York

Roggen HP and Stanley RG 1969 Cell wall hydrolysing enzymes in wall formation as measured by pollen-tube extension. Planta 84: 295-303

Ronald PC 1997 The molecular basis of disease resistance in rice. Plant Molecular Biology 35: 179-186

Rosler J, Krekel F, Amrhein N and Schmid J 1997 Maize phenylalanine ammonia-lyase has tyrosine ammonia-lyase activity. Plant Physiology 113(1): 175-179

Rushton PJ, Torres JT, Parniske M, Wernert P, Hahlbrock K and Somssich IE 1996 Interaction of elicitor-induced DNA-binding proteins with elicitor response elements in the promoter of parsley PR1 gene. EMBO J 15:5690-5700

Ryals J, Uknes S and Ward E 1994 Systemic acquired resistance. Plant Physiology 104:1109-1112

Ryder TB, Hedrick SA, Bell JN, Liang X, Clouse SD and Lamb CJ 1987 Organisation and differential activation of a gene family encoding the plant defense enzyme chalcone synthase in *Phaseolus vulgaris*. Molecular General Genetics 210: 219-233

Sato F, Kitajima S, Koyama T and Yamada Y 1996 Ethylene-induced gene expression of osmotin-like protein, a neutral isoform of tobacco PR-5, is mediated by the AGCCGCC cis-sequence. Plant and Cell Physiology 37: 249-255

Saito K, Takagi Y, Ling HC, Talahashi H and Noji M 1997 Molecular cloning, characterization and expression of cDNA encoding phosphoserine aminotransferase involved in phosphorylated pathway of serine biosynthesis from spinach. Plant Molecular Biology 33:359-366

Sambrook J, Fritsch and Maniatis T 1989 Molecular Cloning, A Laboratory Manual. 2nd edition. Cold Spring Harbor Laboratory Press

Sanchez O, Restrepo S, Duque MC, Fregene M, Bonierbale M and Verdier V 1998 Application of AFLP for genetic characterization of cassava. The Fourth International Scientific Meeting Cassava Biotechnology Network, Brazil.

Sauer N, Corbin DR, Keller B and Lamb CJ 1990 Cloning and characterisation of a wound-specific hydroxyproline-rich glycoprotein in Phaseolus vulgaris. Plant, Cell and Environment 13: 257-266

Scharer-Hernandez NG, Frey PM, Potrykus I and Puonti-Kaerlas J 1998 Towards African cassava mosaic virus resistance in cassava. The Fourth International Scientific Meeting Cassava Biotechnology Network, Brazil.

Scheffner M, Werness BA, Huibregtse JM, Levine AJ and Howley PM 1990 The E6 oncoprotein encoded by human papillomavirus types 16 and 18 promotes the degradation of p53. Cell 63: 1129-1136

Schmelzer E, Kruger-Lebus S and Hahlbrock K 1989 Temporal and spatial patterns of gene expression around sites of attempted fungal infection in parsley leaves. Plant Cell 1: 993-1001

Schnorr KM, Laloue M and Hirel B 1996 Isolation of cDNAs encoding two purine biosynthetic enzymes of soybean and expression of the corresponding transcripts in roots and root nodules. Plant Molecular Biology 32: 751-757

Schnorr KM, Nygaard P and Laloue M 1994 Molecular characterization of Arabidopsis thaliana cDNAs encoding three purine biosynthesis enzymes. Plant Journal 6: 113-121

Schopke C, Taylor NJ, Carcamo R, Konan NK, Marmey P, Henshaw GG, Beachy RN and Fauguet CM 1996 Regeneration of transgenic cassava plants (Mnihot ezculenta Crantz) from microbombardment embrogenic suspension cultures. Nature Biotechnology 14(6): 731-735

Schopke C, Taylor NJ, Masona MV, Carcamo R, Ho T, Beachy RN and Fauguet CM 1998 Characterization of transgenic cassava plants containing the coat protein gene of cassava common mosaic virus. The Fourth International Scientific Meeting Cassava Biotechnology Network, Brazil.

Schuster B and Rétey J 1994 Serine-202 is the putative precursor of the active site dehydroalanine of phenylalanine ammonia-lyase: Site-directed mutagenesis studies on the enzyme from parsley (*Petroselinum crispum* L.). FEBS Letters 349: 252-254

Schweizer P, Buchala A, Dudler R and Metraux JP 1998 Induced systemic resistance in wounded rice plants. Plant Journal 14(4):475-481

Senecoff JF and Meagher RB 1993 Isolating the Arabidopsis thaliana genes for the de novo purine synthesis by suppression of *E. coli* mutants. Plant Physiology 102: 387-399

Sessa G, Meller Y and Fluhr R 1995 A GCC element and a G-box motif participate in ethylene-induced expression of the PRB-1b gene. Plant Molecular Biology 28:145-153

Sewalt VJH, Ni W, Blount JW, Jung HG, Masoud SA, Howles PA, Lamb C and Dixon RA 1997 Reduced lignin content and altered lignin composition in transgenic tobacco down-regulated in expression of phenylalanine ammonia-lyase or cinnamic 4-hydroxylase. Plant Physiology 115: 41-50

Shah J and Klessing DF 1996 Identification of a salicylic acid-responsive element in the promoter of the tobacco pathogenesis-related β-1,3-glucanase gene PR-2d. Plant Journal 10:1089-1101

Shanklin J, Jabben M and Vierstra RD 1987 Red light-induced formation of ubiquitin-phytochrome conjugates: identification of possible intermediates of phytochrome degradation. Proceedings of the National Academy of Sciences of USA 84: 359-363

Sharan M, Taguchi G, Gonda K, Jouke T, Shimodaka M, Hayashida N and Okazaki M 1998 Effects of methyl jasmonate and elicitor on the activation of PAL and the accumulation of scopoletin and scopolin in tobacco cell cultures. Plant Science 132(1): 13-19

Shen Q, Zhang PN and Ho THD 1995 Functional dissection of an abscisic acid (ABA)-inducible gene reveals two independent ABA-responsive complexes each containing a G-box and a novel cis-acting element. Plant Cell 7: 295-307

Shiraishi T, Yamada T, Nicholson RL and Kunoh H 1995 Pheylalanine ammonia-lyase in barley: activity enhancement in response to *Erysiphe graminis f.sp. hordei* (race 1) a pathogen, and *Erysiphe pisi*, a nonpathogen. Physiology and Molecular Plant Pathology 46: 153-162

Shirasu K, Nakajima H, Rajasekhar VK, Dixon RA and Lamb C 1997 Salicylic acid potentiates an agonist-dependent gain control that amplifies pathogen signals in the activation of defense mechanisms. Plant Cell 9(2): 261-270

Shirsat AH 1998 Promoter regions of the extA extensin gene from *Brassica napus* control activation in response to wounding and tensile stress. Plant Molecular Biology 37:675-687

Showalter AM and Rumeau D 1990 Molecular biology of the plant cell wall hydroxyproline-rich glycoproteins. In : Arntzen CJ and Ryan C (eds) Molecular Strategies for Crop Protection. pp. 375-392

Showalter AM, Bell JN, Cramer CL, Bailey JA and Varner JE 1985 Accumulation of hydroxyprolinerich glycoprotein mRNAs in response to fungal elicitor and infection. Proceedings of the National Academy of Sciences of USA 82: 6551-6555

Showalter AM, BellJN, Cramer CL, Bailey JA, Varner JE and Lamb CJ 1985 Accumulation of hydroxyproline-rich glycoprotein mRNAs in response to fungal elicitor and infection. Proceedings of the National Academy of the Scientific USA 82: 6551-6555
Showalter AM, Butt AD and Kim S 1992 Molecular details of tomato extensin and glucine-rich protein gene expression. Plant molecular Biology 19: 205-215

Showalter AM 1993 Structure and function of plant cell wall proteins. The Plant Cell 5:9-23

Shufflebottom D, Edwards K, Schuch W and Bevan M 1993 transcription of two members of a gene family encoding phenylalanine ammonia-lyase leads to remarkably different cell specificities and induction patterns. The Plant Journal 3(6): 835-845

Siefert F, Langebartels C, Boller T and Grossmann K 1994 Are ethylene and ACC involved in the induction of chitinase and β-1,3-glucanase activity in sunflower cell-suspension cultures. Planta 192(3): 431-440

Silke J 1997 The majority of long non-stop reading frames on the antisense strand can be explained by biased codon usage. Gene 194(1): 143-155

Simons CR, Litts JC, Huang N and Rodriguez RL 1992 Structure of a rice beta-glucanase gene regulated by ethylene, cytokinin, wounding, salicylic acid and fungal elicitors. Plant Molecular Biology 18(1): 33-45

Slakeski N, Baulcombe DC, Devos KM, Ahluwalia B, Doan DN and Fincher GB 1990 Structure and tissue-specific regulation of genes encoding barley (1-3, 1-4)-β-glucan endohydrolases. Molecular General Genetics 224: 437-449

Slater A, Maunders MJ, Edwards K, Schuch W and Grierson D 1985 Isolation and characterisation of cDNA clones for tomato polygalacturonase and other ripening-related proteins. Plant Molecular Biology 5: 137-147

Smigocki A, Neal JW, McCamma I and Douglas L 1993 Cytokinin-mediated insect resistance in *Nicotiana* plants transformed with the *ipt* gene. Plant Molecular Biology 23(2):325-335

Smith JJ, Muldoon EP, Willard JJ and Lamport DTA 1986 Tomato extensin precursur P1 and P2 are highly periodic structures. Phytochemistry 25: 1021-1030

Sommer-Knudsen JA, Bacic A and Clarke AE 1998 Hydroxyproline-rich plant glycoproteins. Phytochemistry 47: 483-497

Stein JC, Howlett B, Boyes DC, Nasrallah ME and Nasrallah JR 1991 Molecular cloning of a putative receptor protein kinase gene encoded at the self-incompatibility locus of *Brassica oleracea.* Proceedings of the National Academy of Sciences of USA 88: 8816-8820

Stintzi A, Heitz T, Prasad V, Wiedemannmerdinoglu S, Kauffmann S, Geoffroy P, Legrand M and Fritig B 1993 Plant pathogenesis-related proteins and their role in defense against pathogens. Biochimie 75(8): 687-706

Street PFS, Robb J and Ellis BE 1986 Secretion of vascular coating components by xylem parenchyma cells of tomatoes infected with *Verticillium albo atrum.* Protoplasma 132(1-2): 1-11

Subramanian R, Reinold S, Molitor Ek and Douglas CJ 1993 Structure, inheritance, and expression of hybrid poplar (*Populus tricocarpa x Populus deltoides*) phenylalanine ammonia-lyase genes. Plant Physiology 102: 71-83

Suhandono S, Hughes J, Brown K, Sirju-Charan G and Hughes MA 1998 Characterization of an elongation factor-la gene from cassava: a new promoter for biotechnology. The Fourth International Scientific Meeting Cassava Biotechnology Network, Brazil.

Suzuki H, Wanger T and Tierney ML 1993 Differential expression of 2 soybean (glycine-max L) proline-rich protein genes after wounding. Plant Physiology 101(4): 1283-1287

Tagu D, Walker N, Ruiz-Avila L, Burgess S, Martinez-Izquierdo JA, Leguay JJ, Netter P and Puigdomenech P 1992 Regulation of the maize HRGP gene expression by ethylene and wounding: mRNA accumulation and qualitative expression analysis by microprojectile bombardment. Plant Molecular Biology 20: 529-538

Takeda J, Ozeki Y and Yoshida K 1997 Action spectrum for induction of promoter activity of phenylalanine ammonia-lyase gene by UV in carrot suspension cells. Photochemistry & Photobiology 66(4): 464-470

Takeuchi Y, Yoshikawa M, Takeba G, Tanaka K, Shibata D and Horino Osamu 1990 Molecular cloning and ethylene induction of mRNA encoding a phytoalexin elicitor-releasing factor, β-1,3-glucanase, in soybean. Plant Physiology 93: 673-682

Tanaka Y, Data ES, Hirose S, Taniguchi T and Uritani I 1983 Biochemical changes in secondary metabolites in wounded and deteriorated cassava roots. Agricultural Biology and Chemistry 47(4): 693-700

Tanaka Y, Data ES, Lape VG, Villegas-Godoy CDV, Gorgonio MA and Hirose S 1984 Effect of pruning treatment on physiological deterioration and related biological changes in cassava roots. In Tropical Root Crops: Postharvest Physiology and Processing pp119-125

Taniguchi T, Data ES, Burden OJ, Uritani I and Gorgonio M 1984 The appearance of antifungal activity in cassava root tissue in response to physiological and microbial deterioration. Annals of Phytopathology Society of Japan 50: 286-288

Taylor NJ, Schopke C, Masona MV, CaRcamo R, Ho T, Beachy RN and Fauquet CM 1998 Production of transgenic cassava plants with marker genes of interest by microparticle bombardment and agrobacterium-mediated gene transfer. The Fourth International Scientific Meeting Cassava Biotechnology Network, Brazil.

Taylor RG, Lambert MA, Sexsmith E, Sadler SJ, Ray PN, Mahuran DJ and McInnes RR 1990 Cloning and expression of rat histidase. The Journal of Biological Chemistry 265(30): 18192-18199

Tayloy NJ, Masona MV, C Shopke and Fanquet CM 1999 Transgenic cassava for food security and economic development. In Transgenic Foods. In Press

Theologis A 1992 One rotten apple spoils the whole bushel: the role of ethylene in fruit ripening. Cell 70: 181-184

Thomma BPHJ, Eggermont K, Penninckx IAMA, Mauch-Mani B, Vogelsang R, Cammue BPA and Broekaert WF 1998 Separate jasmonate-dependent and SA-dependent defense-response pathways in arabidopsis are essential for resistance to distinct microbial pathogens. Proceedings of the National Academy of Sciences of USA 95(25): 15107-15111

Thompson AK and Marinia AL 1977 Storage and marketing of cassava in plastic films. Proceedings of the Tropical region. American Society of Horticultural Science 21:30-33

Tierney ML, Weichert J and Pluymers D 1988 Analysis of the expression of extensin and p33 related cell wall proteins in carrot and soybean. Molecular General Genetics 211: 393-399

Tire C, Derycke R, de Loose M, Inze D, Vanmontagu M and Engler G 1994 Extensin gene-expression is induced by mechanical stimuli leading to local cell wall strengthening in *Nicotiana-plumbaginifolia*. Planta 195(2):175-181

Titarenko E, Rojo E, Leon J and Sanchez-Serrano JJ 1997 Jasmonic acid-dependent and -independent signalling pathways control wound-induced gene activation in Arabidopsis thaliana. Plant Physiology 115:817-826

Tobias CM, Howlett B and Nasrallah JB 1992 An *Arabidopsis thaliana* gene with sequence similarity to the S-locus receptor kinase *Brassica oleracea*. Plant Physiology 99: 284-290

TomasBarberan FA, LoaizaVelarde J, Bonfanti A and Saltveit ME 1997 Early wound- and ethylene-induced changes in phenylpropanoid metabolism in harvested lettuce. Journal of the American Society for Horticultural Science 122 (3): 399-404

Uritani I, Data ES, Villegas RJ, Flores P and Hirose S 1983 Relationship between secondary metabolism changes in cassava root tissue and physiological deterioration. Agricultural Biology and Chemistry 47(7): 1591-1598

Uritani I, Data ES and Tanaka Y 1984 Biochemistry of postharvest deterioration of cassava and sweet potato roots. In Uritani I & Reyes ED (Eds), Tropical Root Crops: Postharvest Physiology and Proceeding, Japanese Scientific Societies Press, Tokyo. Pp 61-75

Ussuf KK, Laxmi NH and Nair PM 1996 Possible role of calcium dependent protein phosphorylation in the modulation of wound induced HRGP gene activation in potatoes after gamma irradiation. Indian Journal of Biochemistry & Biophysics 33: 484-490

Van de Peer Y and De Wachter R 1993 TREECON: a software package for the construction and drawing of evolutionary trees. Computer Application in Bioscience 9: 177-182

Van den Bulcke M, Bauw G, Castresana C, Van Montagu M and Vanderchkove J 1989 Characterisation of vacuole and extracellular β-1,3-glucanases of tobacco; evidence for a strictly compartmentalized plant defense system. Proceedings of the National Academy of Sciences of USA 86: 2673-2677

Vanderhee MD, Lemmers R and Bol JF 1993 Analysis of regulatory elements involved in stress-induced and organ-specific expression of tobacco acidic and basic β-1,3-glucanase genes. Plant Molecular Biology 21(3): 451-461

Varghese JN, Garrett TPJ, Colman PM, Chen L, Hoj PB and Fincher GB 1994 Three dimensional structures of two plant β-glucan endohydrolases with distinct substrate specificities. Proceedings of the National Academy of Sciences of USA 91: 2785-2789

Varner JE and Lin L-S 1989 Plant cell wall architecture. Cell 56: 231-239

Verdaguer B, de Kochko A, Fux CI, Beachy RN and Fauquet C 1998 Functional organization of the cassava vein mosaic virus (CsVMV) promoter. Plant Molecular Biology 37: 1055-1067

Vierheilig H, Alt M, Lange J, Gutrella M, Wiemken A and Boller J 1995 Colonization of transgenic tobacco constitutively expressing pathogenesis-related proteins by the vesicular-arbuscular mycorrhizal fungus Glomus-mosseae. Applied and environmental microbiology 61(8): 3031-3034

Viestra RD 1993 Protein degradation in plants. Annual review of Plant Physiology. Plant Molecular Biology. 44: 385-410

Vogelilange R, Frundt C, Hart CM, Nagy F and Meins F 1994 Development, hormonal, and pathogenesis-related regulation of the tobacco class-I β-1,3-glucanase-B promoter. Plant Molecular Biology 25(2): 299-311

von Kampen J and Wettern M 1995 Expression of ubiquitin genes in *Chlamydomonas reinhardtii*: involvement in stress response and cell cycle. Planta. 197: 528-534

Von Kampen J, Wettern M and Schulz M 1996 The ubiquitin system in plants. Physiologia Plantarum. 97: 618-624

Walker JC and Zhang R. 1990 Relationship of a putative receptor protein kinase from maize to the S-locus glycoproteins of Brassica. Nature 345: 743-746

Walker-Simmons M, and Ryan CA 1977 Wound-induced peptidase activity in tomato leaves. Biochemistry Biophysics Research Communication 74: 411-416

Wanner LA, Li G, Ware D, Somssich IE and Davis KR 1995 The phenyalanine ammonia-lyase gene family in *Arabidopsis thaliana*. Plant Molecular Biology 27: 327-338

Ward ER, Payne GB, Moyer MB, Williams SC, Dincher SS, Sharkey KC, Beck JJ, Taylor HT, Ahlgoy P, Maing F and Ryals JA 1991 Differential regulation of beta-,3-glucanase messenger-RNAs in response to pathogen infection. Plant Physiology 96(2): 390-397

Wenham JE 1995 Post-harvest deterioration of cassava. A biotechnological Perspective, FAO, Rome.

Wessels JGH and Sietsma JH 1981 Fungal cell walls: a survey. Encyclopedia Plant Physiology NS 13B: 352-394

Wheatley C 1980 Studies related with the nature of post-harvest physiological deterioration in cassava roots. CIAT seminarios Internos SE-16-80

Wheatley C 1982 Studies on cassava root post-harvest physiological deterioration. PhD thesis, London University

Wheatley C, Fernandez FO and Medina LML 1989 Conservation of cassava roots in polythene bags. CIAT, Cali Colombia pp33

Wheatley CC and Schwabe WW 1985 Scopoletin involvement in post-harvest physiological deterioration of cassava root (*Manihot esculenta* Crantz). Journal of Experimental Botany 36: 783-791

Whetten RW and Sederoff RR 1992 Phenylalanine ammonia-lyase from loblolly pine: purification of the enzyme and isolation of complementary DNA clones. Plant Physiology 98: 380-386

White WLB, Arias-Garzon DI, McMahon JM and Sayre RT 1998 Cyanogenesis in cassava-The role of hydroxynitrile lyase in root cyanide production. Plant Physiology 116: 1219-1225

Whitmore FW 1978 Lignin-protein complex catalyzed by peroxidase. Plant Science letters 13:241-245

Wickham LD and Wilson LA 1988 Quallity changes during long-term storage of cassava roots in moist media. Tropical Science 28:79

Woodward JR and Fincher GB 1982 Purification and chemical properties of two 1,3; 1,4-β-glucan endohydrolases from germinating barley. European Journal of Biochemistry 121: 663-669

Worrall D, Hird DL, Hodge R, Paul W, Draper J and Scott R 1992 Premature dissolution of the microsporocyte callose wall causes male sterility in transgenic tobacco. Plant Cell 4: 759-771

Wycoff KL, Powell_PA, Gonzales RA, Corbin DR, Lamb C and Dixon RA 1995 Stress activation of a bean hydroxyproline-rich glycoprotein promoter is superimposed on a pattern of tissue-specific developmental expression. Plant Physiology 109: 41-52

Yang SF 1987 The role of ethylene and ethylene synthesis in fruit ripening. In Plant senescence: its biochemistry and physiology. Eds. WW Thompson, EA Nothnagel and RC Huffaker. American Society of Plant Physiologists. pp 156-166

Ye ZH and Varner JE 1991 Tissue specific expression of cell wall proteins in developing soybean tissues. Plant Cell 3: 23-37

Yoshikawa M, Matama M and Masago H 1981 Release of a soluble phytoalexin elicitor from mycelial walls of *Phytophthora megasperma var. sojae* by soybean tissues. Plant Physiology 67: 1032-1035

Yoshikawa M, Tsuda M and Takeuchi Y 1993 Resistance to fungal diseases in transgenic tobacco plants expressing the phytoalexin elicitor-releasing factor, beta-1,3-endoglucanase, from soybean. Natuwissenschaften 80:417-420

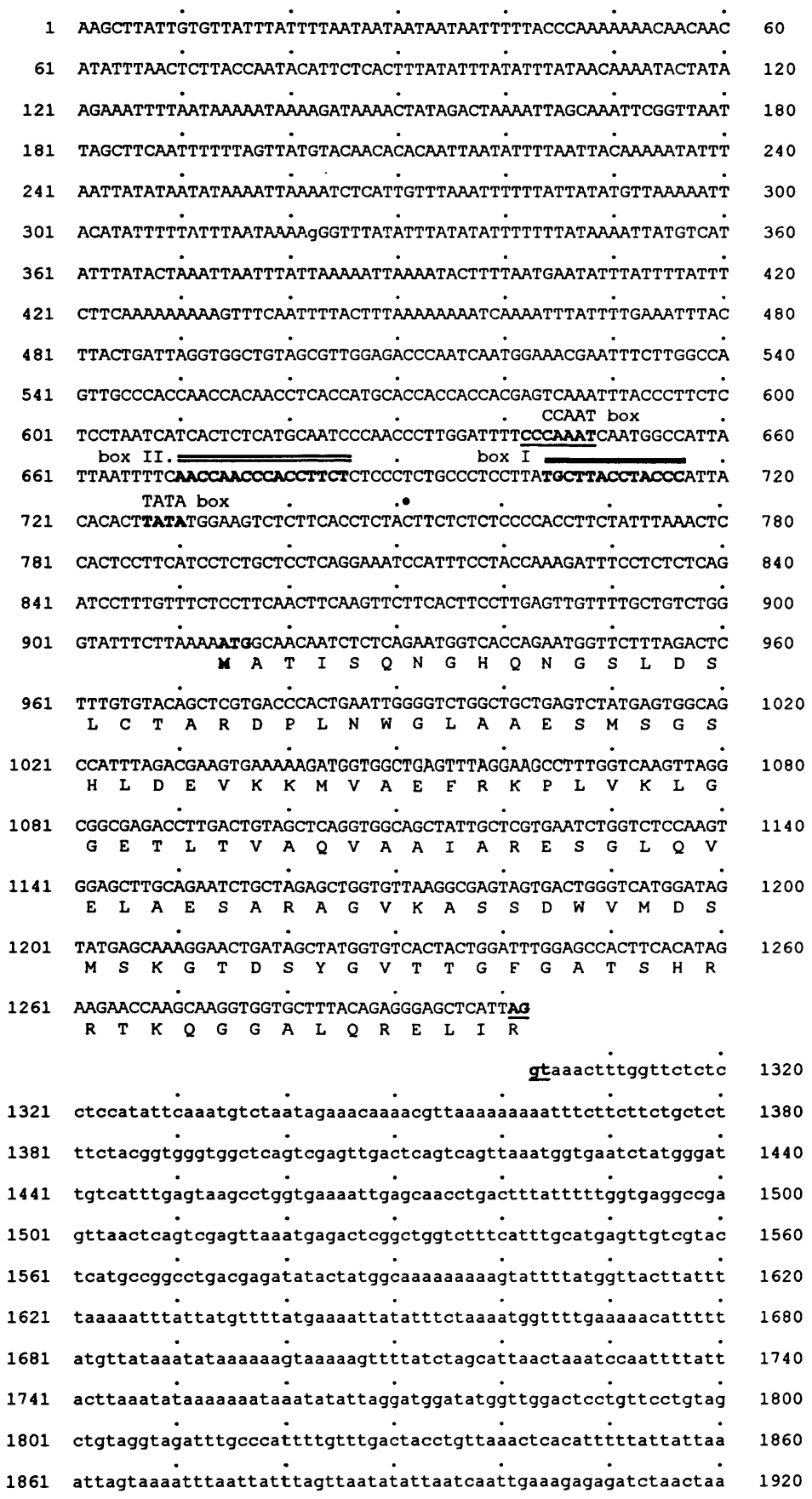
Zalkin H and Dixon JE 1992 *De novo* purine nucleotide biosynthesis. Progress in Nucleic Acid Research and Molecular Biology. Ed by Cohn WE and Molddave K. Academic Press Inc. 42: 259-287

Zhu Q, Dabi T, Beeche A, Yamamoto R, Lawton MA and Lamb C 1995 Cloning and properties of a rice gene encoding phenylalanine ammonia-lyase. Plant Molecular Biology 29: 535-550

Zhu Q, Maher ER, Masoud S, Dixon RA and Lamb CJ 1994 Enhanced protection against fungal attack by constitutive coexpression of chitinase and glucanase genes in transgenic tobacco. Biotechnology 12(8): 807-812

Appendix A1

The Sequence of gMePAL2



```
1921  atgatacttttggcagccttagaattcatattttcctagaaaattcatccaaaacagtaa  1980

1981  ctaaacatgttgtatgttactttgtttctgcag

                                                    ATTCCTGAATGCTGGGATTTTTGGCAA  2040
                                                     F  L  N  A  G  I  F  G  N

2041  TAAGACAGAATCGTGTCACACTTTGTCACACTCTGCAACAAGAGCAGCAATGCTAGTGAG  2100
       K  T  E  S  C  H  T  L  S  H  S  A  T  R  A  A  M  L  V  R

2101  GATCAACACTCTTCTCCAAGGTTACTCAGGCATTAGATTTGAAATCTTGGAAGCTATCAC  2160
       I  N  T  L  L  Q  G  Y  S  G  I  R  F  E  I  L  E  A  I  T

2161  CAAGCTCCTCAACCACAATATTACTCCTTGCTTGCCGCTGAGAGGCACAATCACTGCTTC  2220
       K  L  L  N  H  N  I  T  P  C  L  P  L  R  G  T  I  T  A  S

2221  AGGGGATTTAGTTCCATTGTCCTACATTGCTGGATTGCTCACCGGCCGGCCTAATTCCAA  2280
       G  D  L  V  P  L  S  Y  I  A  G  L  L  T  G  R  P  N  S  K

2281  GGCTGTTGGCCCTAATGGAGAATCCCTAGATGCCCAGCAAGCCTTTCACTCTGCTGGTAT  2340
       A  V  G  P  N  G  E  S  L  D  A  Q  Q  A  F  H  S  A  G  I

2341  TGATTCTGGCTTCTTTGAGTTGCAGCCTAAAGAAGGCCTTGCTCTGGTTAATGGCACTGC  2400
       D  S  G  F  F  E  L  Q  P  K  E  G  L  A  L  V  N  G  T  A

2401  TGTTGGTTCTGGCTTGGCTTCCATGGTTCTCTTTGAGGCAAATGTTCTTGCTGTTTTATC  2460
       V  G  S  G  L  A  S  M  V  L  F  E  A  N  V  L  A  V  L  S

2461  AGAAGTCTTATCAGCTATTTTCGCCGAAGTTATGAATGGAAAACCTGAGTTTACTGATCA  2520
       E  V  L  S  A  I  F  A  E  V  M  N  G  K  P  E  F  T  D  H

2521  CTTGACTCATAAGTTGAAGCACCATCCAGGCCAAATAGAGGCTGCAGCTATAATGGAGCA  2580
       L  T  H  K  L  K  H  H  P  G  Q  I  E  A  A  A  I  M  E  H

2581  TATTTTAGATGGTAGCTCTTATATTAAAGCAGCTAAGAAGTTGCATGAAATTGATCCATT  2640
       I  L  D  G  S  S  Y  I  K  A  A  K  K  L  H  E  I  D  P  L

2641  GCAGAAACCAAAGCAAGATCGATATGCTCTCAGAACTTCCCCACAATGGCTAGGTCCTCA  2700
       Q  K  P  K  Q  D  R  Y  A  L  R  T  S  P  Q  W  L  G  P  Q

2701  GATTGAAGTTATCAGATTCTCCACAAAATCGATCGAAAGAGAGATTAATTCAGTCAATGA  2760
       I  E  V  I  R  F  S  T  K  S  I  E  R  E  I  N  S  V  N  D

2761  CAACCCTTTGATTGATGTTTCTAGGAACAAGGCCTTGCATGGTGGAAATTTCCAGGGGAC  2820
       N  P  L  I  D  V  S  R  N  K  A  L  H  G  G  N  F  Q  G  T

2821  CCCAATTGGAGTCTCAATGGATAATGCACGTTTGGCCATTGCATCAATAGGAAAGCTCAT  2880
       P  I  G  V  S  M  D  N  A  R  L  A  I  A  S  I  G  K  L  M

2881  GTTTGCTCAGTTCAGTGAGCTTGTAAATGATTTTTACAACAATGGGTTGCCATCAAATCT  2940
       F  A  Q  F  S  E  L  V  N  D  F  Y  N  N  G  L  P  S  N  L

2941  CACAGCCAGCAGGAATCCAAGCTTGGATTATGGCTTCAAGGGAGCTGAAATTGCAATGGC  3000
       T  A  S  R  N  P  S  L  D  Y  G  F  K  G  A  E  I  A  M  A

3001  TTCTTACTGTTCTGAGCTCCAATACCTTGCAAATCCAGTCACCAGCCATGTACAAAGTGC  3060
       S  Y  C  S  E  L  Q  Y  L  A  N  P  V  T  S  H  V  Q  S  A

3061  AGAGCAGCACAATCAAGATGTAAATTCATTGGGGCTAATTTCTTCAAGAAAGACAGAAGA  3120
       E  Q  H  N  Q  D  V  N  S  L  G  L  I  S  S  R  K  T  E  E

3121  AGCTGTAGACATCTTGAAGCTCATGTCCACGACTTTCTTAGTAGCACTTTGTCAAGCTAT  3180
       A  V  D  I  L  K  L  M  S  T  T  F  L  V  A  L  C  Q  A  I

3181  TGACTTGAGGCATTTGGAGGAGAACTTGAAGCACGCAGTCAAAAACACAGTAAGCCAAGT  3240
       D  L  R  H  L  E  E  N  L  K  H  A  V  K  N  T  V  S  Q  V

3241  AGCTAAGAGGATTCTAACTACAGGAGCTAGTGGAGAACTTCACCCATCAAGATTCTGCGA  3300
       A  K  R  I  L  T  T  G  A  S  G  E  L  H  P  S  R  F  C  E

3301  GAAGGACTTGCTCAAAGTGGTGGATCGCGAGCAAGTCTTCTCTTATGTCGATGACGCCTG  3360
       K  D  L  L  K  V  V  D  R  E  Q  V  F  S  Y  V  D  D  A  C

3361  CAGTGCTACCTATCCATTGATGCAAAAACTAAGGCAAGTTCTCGTGGACCATGCCTTGGC  3420
       S  A  T  Y  P  L  M  Q  K  L  R  Q  V  L  V  D  H  A  L  A
```

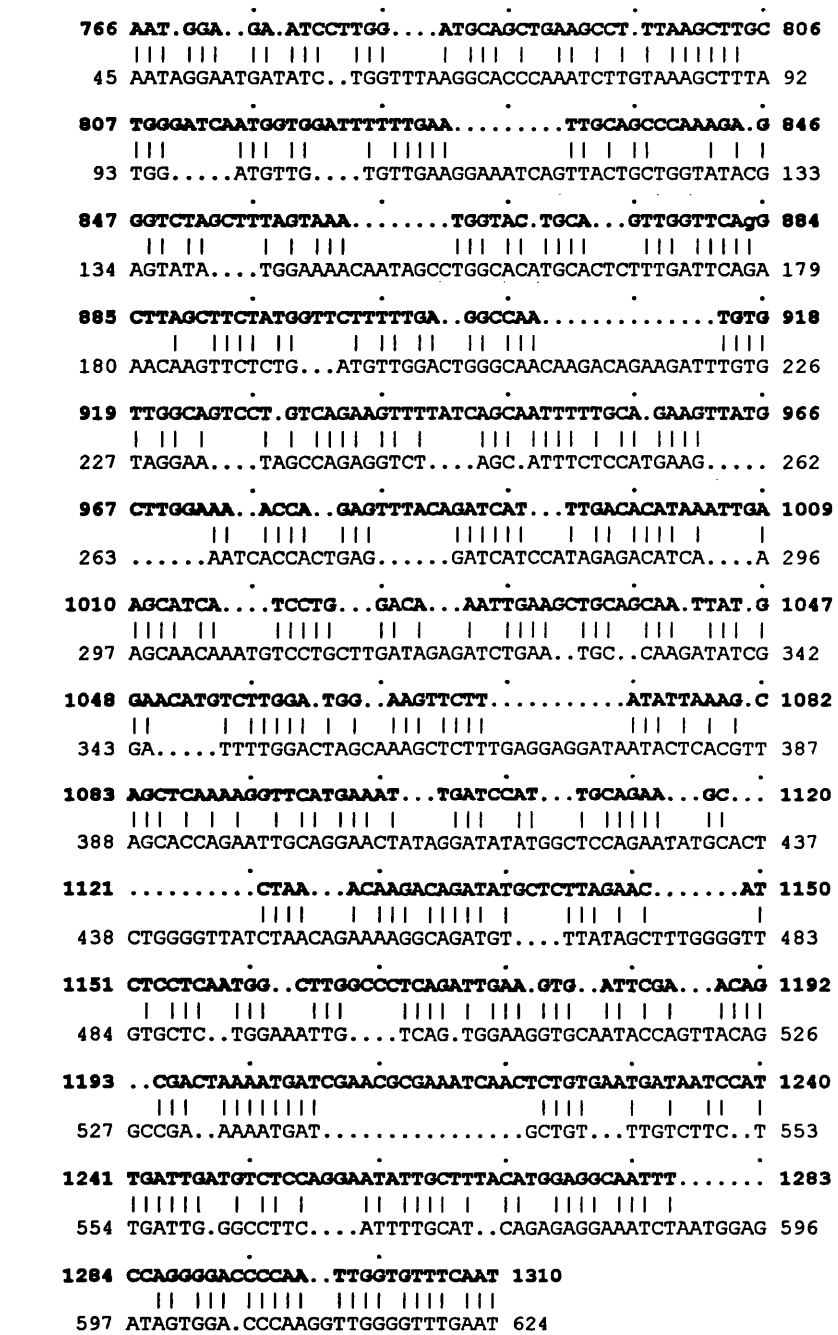
```
3421  AAATGGCGAGAGTGAGAAGAATGCCAGCACTTCAATCTTCCAAAAGATCAGAGCTTTTGA   3480
       N  G  E  S  E  K  N  A  S  T  S  I  F  Q  K  I  R  A  F  E
3481  GGAAGAATTGAAAGCCCTTTTGCCTAAAGAAGTTGAGAGTGCAAGAGAGGCATACGAGAA   3540
       E  E  L  K  A  L  L  P  K  E  V  E  S  A  R  E  A  Y  E  N
3541  CGGGAATCCAGCAATTGCCAACAAGATCAAGGAATGCAGATCTTACCCATTGTATAAGTT   3600
       G  N  P  A  I  A  N  K  I  K  E  C  R  S  Y  P  L  Y  K  F
3601  TGTGAGAGAGGAAATAGGAACTGGGTTGCTCACCGGAGAAAAGATCCGGTCACCGGGAGA   3660
       V  R  E  E  I  G  T  G  L  L  T  G  E  K  I  R  S  P  G  E
3661  GGAATTTGATAAGGTTTTCACTGCTATGTGCCAAGGAAAGATCATTGATCCAATGCTGGA   3720
       E  F  D  K  V  F  T  A  M  C  Q  G  K  I  I  D  P  M  L  D
3721  TTGTCTCAAAGAGTGGAATGGTGCCCCTCTTCCAATATGTTAAACTGTAACTTTCTTGTT   3780
       C  L  K  E  W  N  G  A  P  L  P  I  C  *
3781  TTGTTTACACTTCAAGATTTGTTTTCCAATTGCTTTTTATGTACTTATAATTTGTGATGT   3840
3841  AAAAAATCTGTAATGCATTTCTTTTAAATGTTCAATTGTTATCTTCTCACTTTTGTGCTG   3900
3901  GAATTGAAGGCAGAATAGCAATGGTAATTACTTCAAGATTACTGAATTGAAATTTTTTGA   3960
3961  GTTATTCATTAATCTAAGTGTTTATTTAAAAAAATTTAAGAGAATTTTAATTGACAGAAA   4020
4021  GAAGTGTATCCATATATTTTCGAGGTCGAGAATCCATGAACGTGTAGCGGTTGGCGCTGT   4080
4081  GCTGTTCCTCTCCTTGACCTTGAGCTTATTGCTTGGCAGCTTGAGAGCCTCTCCTTGAAG   4140
4141  GGAAGACGAGAGATGTACCTCTTCTTGTTTGGATATATGGGGCCCATTAGATTTCCCAAA   4200
4201  ATTAATGTTTCGTTGTTTCTACGTACTTCTGTTTTCACCCCAGATGTTGATGCTTAACAA   4260
4261  ACTCCAAATTTCAATTGAAATTAATTTTTTTTATTGTGCATATATGGTAAAAAAAAAGAA   4320
4321  TATAAAAAAATAAAATTAGAATATAATTAAATAAGAAGTTTTTTTTGAAATTGAAAATTT   4380
4381  TACTATATCTATCATTAATAATGATTTGGTCTTTGTGTTTTAAAACATTAGACTGTTTAG   4440
4441  TTCCTATTTTCTTTTGCCATACACTTGGATCGTTTTATCCATTTTTCCTTTAGTATGnTA   4500
4501  ATATAAAATTATATTTAGTCCATAGATATTTATTTAAAATTCCTCATAATTATTTACAAT   4560
4561  TTCATATATTAATTCTTCTTAGAGCATAGAGATATATAAAAAGAGGGAGAGAAAATTAAG   4620
4621  AGAGGAAAAAGACATAAATATGAGATATGAAAGTATnACTGATTATTGGGGACAAATTAC   4680
4681  CAAATAACTGGAATATTCAAGAATTnACTAGTAAATTTAATTTAAAAACTAATTATAAAA   4740
4741  TTAATnGGAAAAAGTTTCAATGTAAGCAAAAATTAGGAGGACTTAAATAGT  4791
```

Translation initiation and termination sites are at position 914 and 3763 in bold. The intron region (size 711bp) is in lower case. The exon/intron boundaries are in bold and underlined. Putative CCAAT box and TATA box are in bold. Two putative regulatory element regions (Hatton et al, 1995) are overlined with single bold line and double lines respectively.

Hatton D, Sablowski R, Yung M-H, Smith C, Schuch W and Bevan M 1995 Two classes of *cis* sequences contributed to tissue-specific expression of a *PAL2* promoter in transgenic tobacco. The Plant Journal 7(6): 859-876

Appendix A.2

Comparison of Sequences of cMePAL1 and Putative Receptor-like Protein Kinase or Serine-threonine Protein Kinase



The sequence of cMePAL1 is in bold. The criteria for the comparison and summary of the results:

Gap Weight:	2.000	Average Match:	1.000
Length Weight:	0.300	Average Mismatch:	-0.900
Quality:	54.2	Length:	679
Ratio:	0.099	Gaps:	65
Percent Similarity:	77.354	Percent Identity:	77.354

Appendix A.3

Comparison of Sequences of cMePAL1 and cDNA Encoding Putatitive Ubiquitin-activating Enzyme

```
                .         .         .         .         .
 936 AGTTTTATCAGCAAT...TTTTGCAGAAGTTATGCTT.GGAAAACCAG.. 979
     |||||||| |  |||   |||| |||  |||||||||  | | || ||
1107 AGTTTTATTA..AATTTCTTTTTCAG..GTTATGCTTCTGCAGACAAGTA 1152
                .         .         .         .         .
 980 ......AGTTTACAGATCATTTGACACATAAATTGAAGCATCATCCTGGA 1023
           || | | ||| |   ||  | | ||| ||  | | |  ||||
1153 TGTTAAAGCTGAAAGAGCCGATGTAAGAGAAACTGCCGAAGCTGCCTGCC 1202

1024 CAAA 1027
      |||
1203 AAAA 1206
```

The cMePAL1 sequence is in bold. The criteria for the comparison and summary of the results:

```
        Gap Weight:  3.000      Average Match:  1.000
     Length Weight:  0.300   Average Mismatch: -0.900

           Quality:   16.9             Length:    104
             Ratio:  0.184               Gaps:      5
Percent Similarity: 69.318   Percent Identity: 69.318
```

Appendix A.4

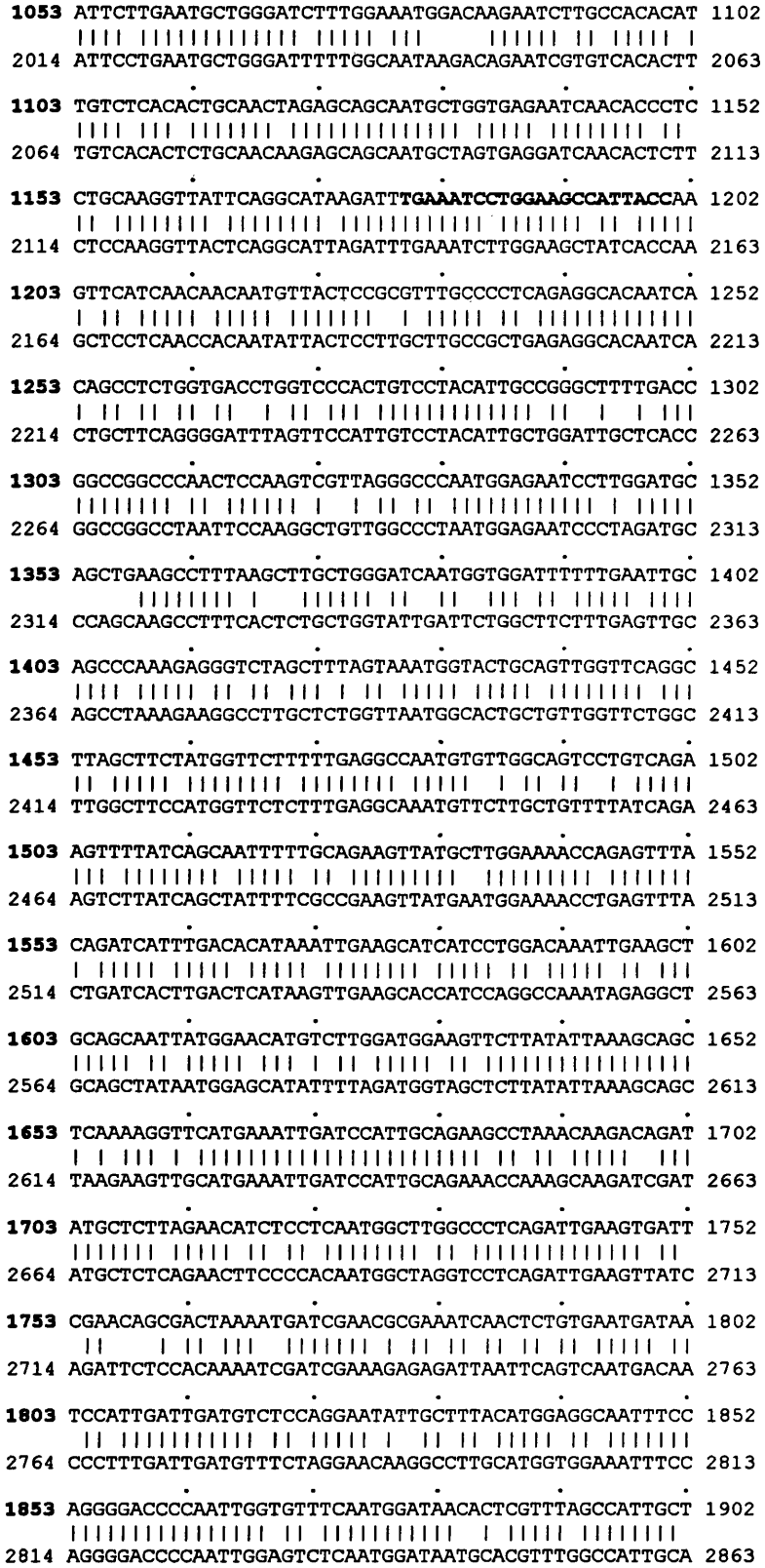
Comparison of Exon1 Nucleotide Sequences between MePAL1 and MePAL2

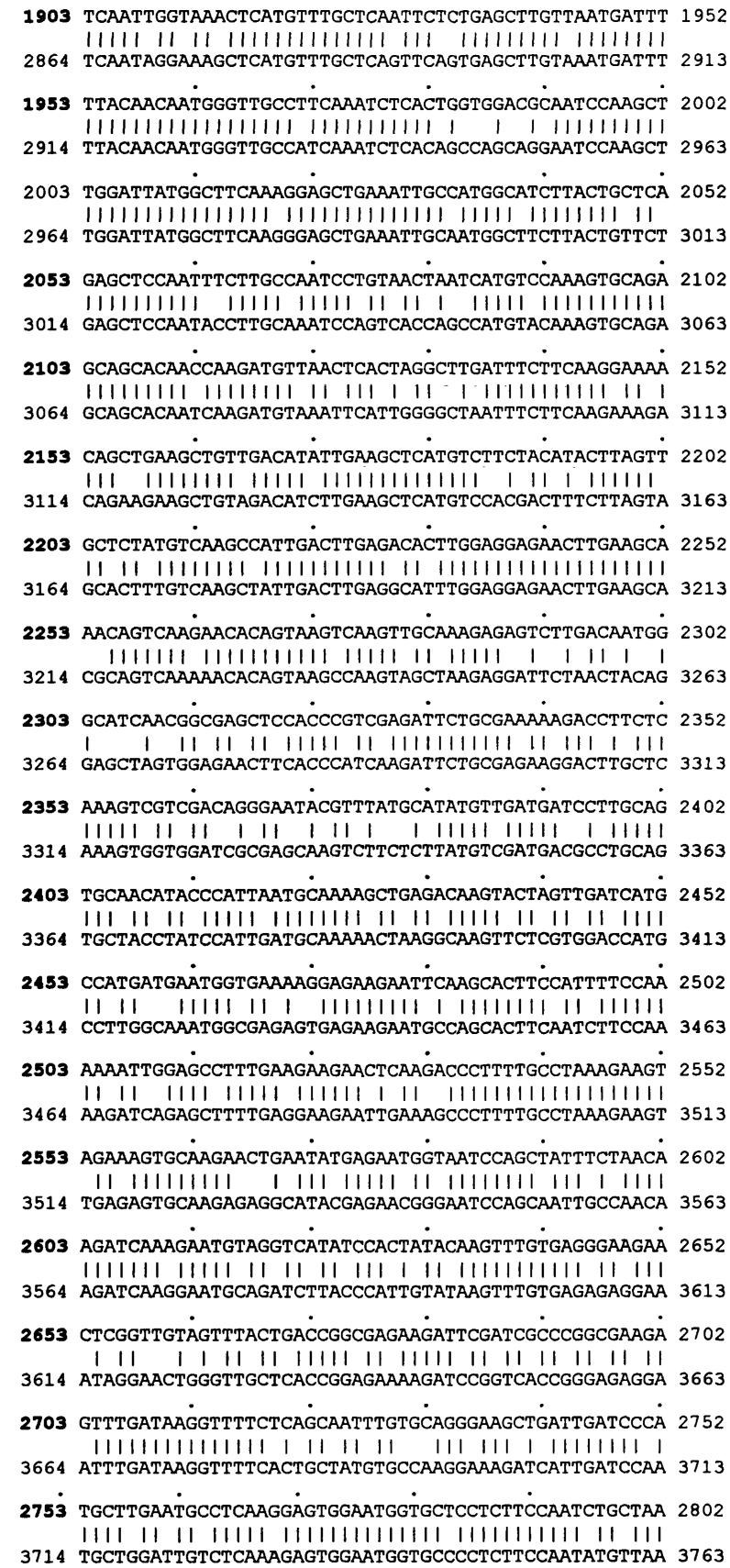
```
          .         .         .         .         .
 127 CTGCTGACCCATTGAACTGGGGCATGGCTGCAGAGTCACTGAAGGGCAGC 176
     ||  ||||||| |||| |||||  ||||||| |||||  |||  ||||||
 972 CTCGTGACCCACTGAATTGGGGTCTGGCTGCTGAGTCTATGAGTGGCAGC 1021
          .         .         .         .         .
 177 CACCTTGATGAGGTGAAGCGCATGGTGGATGAGTACAGGAAGCCTGTGGT 226
     ||  | || || |||||    ||||||| |||||  ||||||||| ||||
1022 CATTTAGACGAAGTGAAAAAGATGGTGGCTGAGTTTAGGAAGCCTTTGGT 1071
          .         .         .         .         .
 227 GAGGCTAGGGGGTGAGACCCTGACTATAGCCCAAGTTACAGCAATTGCGA 276
      | | |||| || |||||| ||||| |||| || ||  |||| |||||
1072 CAAGTTAGGCGGCGAGACCTTGACTGTAGCTCAGGTGGCAGCTATTGC.. 1119
          .         .         .         .         .
 277 ACCATGACTCAGGTGTCAAGGTTGAGCTGTCTGAGGAGGCTCGAGCTGGG 326
       | ||| || ||| || | || |||||  | ||    ||| |||||||
1120 .TCGTGAATCTGGTCTCCAAGTGGAGCTTGCAGAATCTGCTAGAGCTGGT 1168
          .         .         .         .         .
 327 GTCAAGGCCAGTAGTGACTGGGTTCTTGATTCCATGAATAAAGGAACCGA 376
     || ||||| ||||||||||||||  | |||   ||||  |||||||| ||
1169 GTTAAGGCGAGTAGTGACTGGGTCATGGATAGTATGAGCAAAGGAACTGA 1218
          .         .         .         .         .
 377 TAGTTATGGTGTCACCACTGGCTTTGGTGCAACCTCCCATAGAAGAACCA 426
     ||| |||||||||| ||||| ||||| || || || ||||||||||||
1219 TAGCTATGGTGTCACTACTGGATTTGGAGCCACTTCACATAGAAGAACCA 1268
          .         .         .
 427 AGCAGGGGGGTGCCCTTCAGAGAGAACTCATAAG 460
     |||| || |||||  | ||||| || ||||| ||
1269 AGCAAGGTGGTGCTTTACAGAGGGAGCTCATTAG 1302
```

The exon1 sequences of MePAL1 and MePAL2 were compared using Bestfit program in GCG. In the first 50bp of the exon1 the sequences are very converged between these two genes. Over the 334bp down-stream the similarity between these two genes is 76.13%. The sequences in bold were the forward primer used for PCR amplification of genomic DNA corresponding to cMePAL1 (see 4.2.2)
Note: The lines headed with bold numbers are the sequences of MePAL1 exon1.

Appendix A.5

Comparison of Exon2 Nucleotide Sequences between MePAL1 and MePAL2

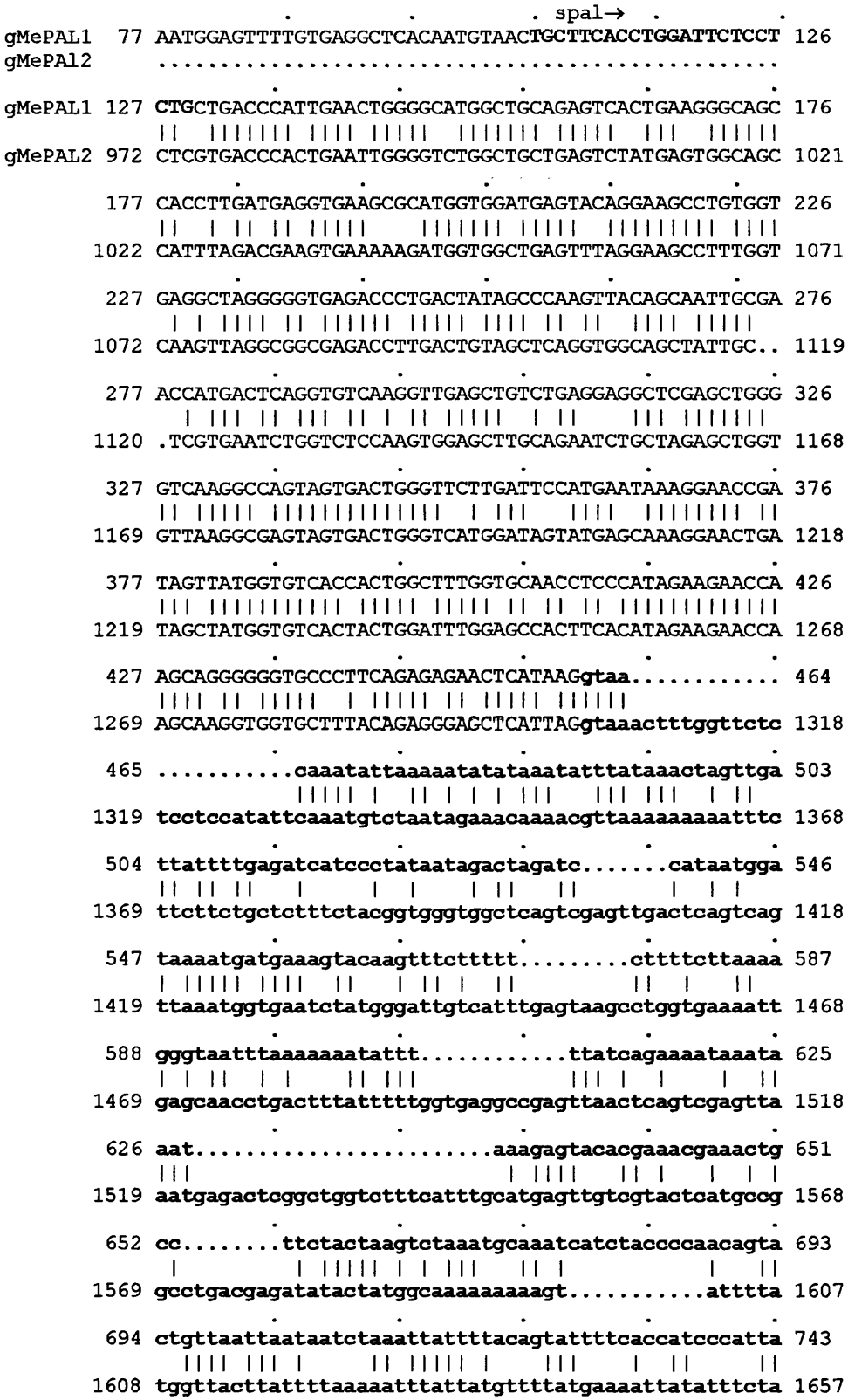




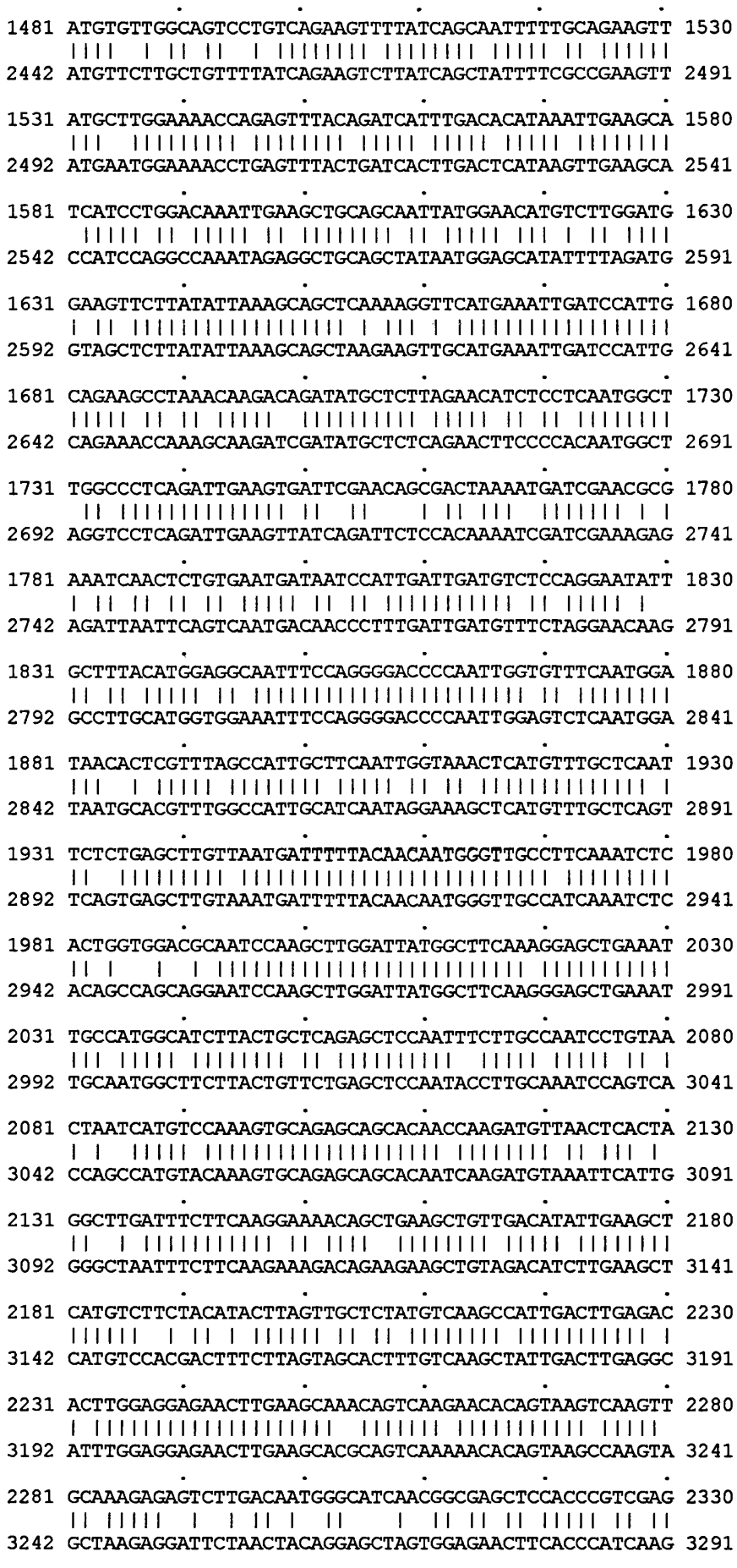
The exon2 sequences of MePAL1 and MePAL2 were compared using Bestfit program in GCG. Over the 1750bp of the exon2 the similarity between these two genes is 79.31%. The sequences in bold were the reverse primer used to amplify genomic DNA corresponding to cMePAL1 (see 4.2.2).
Note: The lines headed with bold numbers are the sequences of MePAL1 exon2.

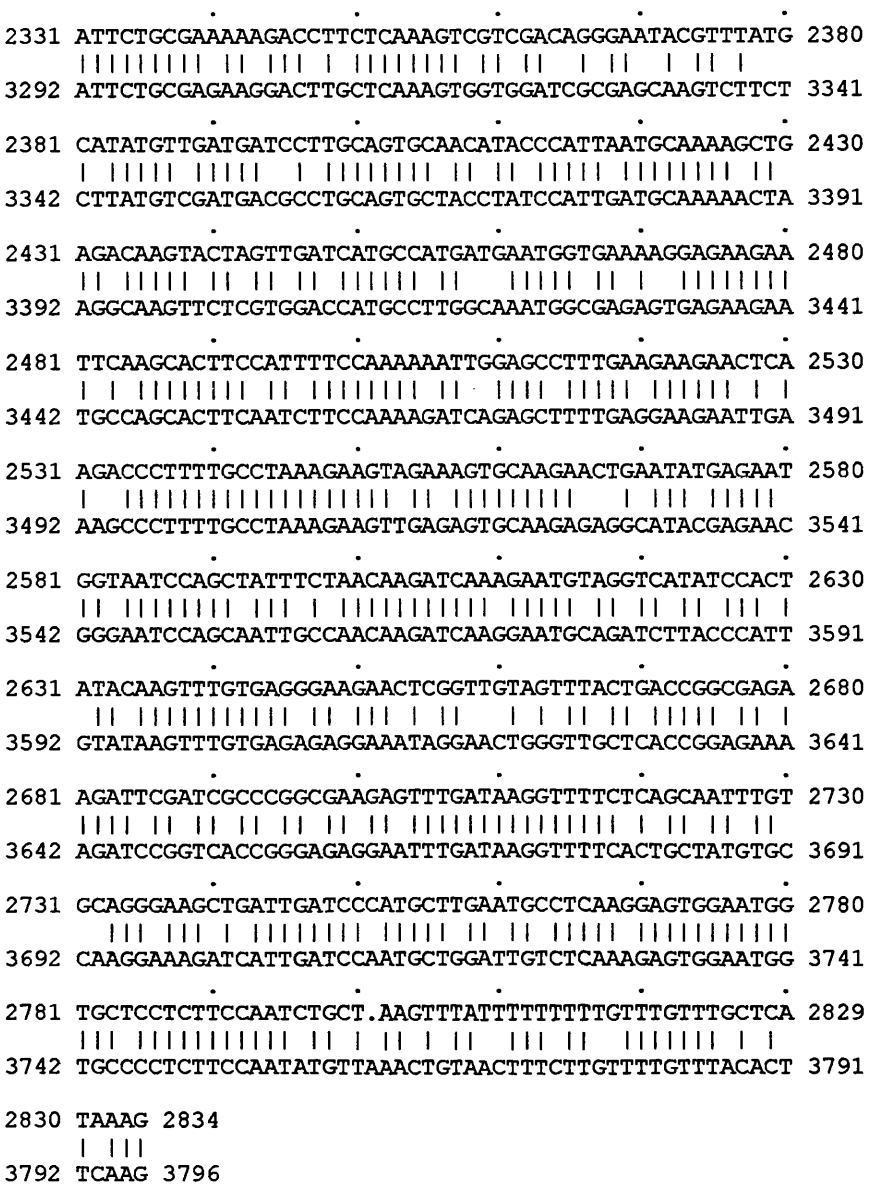
Appendix A.6

Comparison of gMePAL1 and gMePAL2 Sequences and the Location of RT-PCR Primers




744 aaatagat...atattcattttt...ttatataatttaaaaaatataa.. 785
1658 aaatggttttgaaaaacatttttatgttataaatataaaaaagtaaaaag 1707
786 ....ttaaaaataatacatttatattgccttttttagtaatata 824
1708 ttttatctagcattaactaaatccaattttattacttaaatataaaaaaa 1757
825 tacat.....tactcatattactcaa..... 845
1758 taaatatattaggatggatatggttggactcctgttcctgtagctgtagg 1807
846 .ttattaaatcatttttt.....aaattgataaatttttatttttattat 889
1808 tagatttgcccatttttgtttgactacctgttaaactcacatttttattat 1857
890 atattaaataagggtatatta.....aaaatattaaatgatagattca 932
1858 taaattagtaaaatttaattatttagttaatatattaatcaattgaaaga 1907
933 tataaatagctaattccaacaaatctgaaattaaaaatt..taattgact 980
1908 gagatctaactaaatgatacttttggcagccttagaattcatattttcct 1957
981 tgagcattgtttcttgaacaaatggttagtaatttaaatgtttttcttgg 1030
1958 agaaaattcatcc.....aaaacagtaactaa...acatgttgt..... 1993
1031 aaatttttgttttattctccagATTCTTGAATGCTGGGATCTTTGGAAAT 1080
1994 ..atgttactttgtttctgcagATTCCTGAATGCTGGGATTTTTGGCAAT 2041
gagtgtgacgttgatctcgtcgt←aspal
1081 GGACAAGAATCTTGCCACACATTGTCTCACACTGCAACTAGAGCAGCAAT 1130
2042 AAGACAGAATCGTGTCACACTTTGTCACACTCTGCAACAAGAGCAGCAAT 2091
1131 GCTGGTGAGAATCAACACCCTCCTGCAAGGTTATTCAGGCATAAGATTTG 1180
2092 GCTAGTGAGGATCAACACTCTTCTCCAAGGTTACTCAGGCATTAGATTTG 2141
1181 AAATCCTGGAAGCCATTACCAAGTTCATCAACAACAATGTTACTCCGCGT 1230
2142 AAATCTTGGAAGCTATCACCAAGCTCCTCAACCACAATATTACTCCTTGC 2191
1231 TTGCCCCTCAGAGGCACAATCACAGCCTCTGGTGACCTGGTCCCACTGTC 1280
2192 TTGCCGCTGAGAGGCACAATCACTGCTTCAGGGGATTTAGTTCCATTGTC 2241
1281 CTACATTGCCGGGCTTTTGACCGGCCGGCCCAACTCCAAGTCGTTAGGGC 1330
2242 CTACATTGCTGGATTGCTCACCGGCCGGCCTAATTCCAAGGCTGTTGGCC 2291
1331 CCAATGGAGAATCCTTGGATGCAGCTGAAGCCTTTAAGCTTGCTGGGATC 1380
2292 CTAATGGAGAATCCCTAGATGCCCAGCAAGCCTTTCACTCTGCTGGTATT 2341
1381 AATGGTGGATTTTTTGAATTGCAGCCCAAAGAGGGTCTAGCTTTAGTAAA 1430
2342 GATTCTGGCTTCTTTGAGTTGCAGCCTAAAGAAGGCCTTGCTCTGGTTAA 2391
1431 TGGTACTGCAGTTGGTTCAGGCTTAGCTTCTATGGTTCTTTTTGAGGCCA 1480
2392 TGGCACTGCTGTTGGTTCTGGCTTGGCTTCCATGGTTCTCTTTGAGGCAA 2441





Intron region is in bold; primer spal and aspal used for RT-PCR are in blue.