INFORMATION TO USERS

This manuscript has been reproduced from the microfilm master. UMI films the text directly from the original or copy submitted. Thus, some thesis and dissertation copies are in typewriter face, while others may be from any type of computer printer.

The quality of this reproduction is dependent upon the quality of the copy submitted. Broken or indistinct print, colored or poor quality illustrations and photographs, print bleedthrough, substandard margins, and improper alignment can adversely affect reproduction.

In the unlikely event that the author did not send UMI a complete manuscript and there are missing pages, these will be noted. Also, if unauthorized copyright material had to be removed, a note will indicate the deletion.

Oversize materials (e.g., maps, drawings, charts) are reproduced by sectioning the original, beginning at the upper left-hand corner and continuing from left to right in equal sections with small overlaps. Each original is also photographed in one exposure and is included in reduced form at the back of the book.

Photographs included in the original manuscript have been reproduced xerographically in this copy. Higher quality 6" x 9" black and white photographic prints are available for any photographs or illustrations appearing in this copy for an additional charge. Contact UMI directly to order.

UMI

A Bell & Howell Information Company
300 North Zeeb Road, Ann Arbor MI 48106-1346 USA
313/761-4700 800/521-0600

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

MOLECULAR CLONING AND CHARACTERIZATION OF PHYTOENE DESATURASE cDNA AND LEUCINE-RICH REPEAT PROTEIN KINASE cDNA FROM MAIZE

by

ZHAOHUI LI

A dissertation submitted to the Graduate Faculty in Biochemistry in partial fulfillment of the requirements for the degree of Doctor of Philosophy, The City University of New York

1998

UMI Number: 9820555

UMI
300 North Zeeb Road
Ann Arbor, MI 48103

This manuscript has been read and accepted for the Graduate Faculty in Biochemistry in satisfaction of the dissertation requirement for the degree of Doctor of Philosophy.

Jan 23, 1998
Date

Chair of Examining Committee

January 30, 1998
Date

Executive Officer

Shirley Raps

Supervisory Committee

The City University of New York

Abstract

MOLECULAR CLONING AND CHARACTERIZATION OF PHYTOENE DESATURASE cDNA AND LEUCINE-RICH REPEAT PROTEIN KINASE cDNA FROM MAIZE

by

Zhaohui Li

Adviser: Professor Eleanore T. Wurtzel

In plants, carotenoids are major components of the photosynthetic systems as energy trapping pigments, as protectors against photooxidation and as precursors to abscisic acid. For humans, they are the major source for vitamin A and retinoid synthesis. To study regulation of the carotenoid biosynthetic pathway in maize, cDNAs encoding phytoene desaturase (PDS) were cloned and characterized. The protein expressed by one of the *Pds* cDNA clones catalyzed the desaturation of phytoene to ζ-carotene in *Escherichia coli*. The *Pds* cDNA mapped to the *viviparous5* locus, whose recessive alleles confer phytoene accumulation in leaf and endosperm of maize. Accumulation of *Pds* and of *Psy* (encoding phytoene synthase, PSY) transcripts in maize endosperm were compared using RT-PCR (reverse transcriptase-polymerase chain reaction). While *Psy* transcript accumulation was temporally regulated in the developing endosperm, *Pds* transcripts were constant. Results of transcript analysis in maize lines that vary in endosperm carotenoid content suggest that *Psy* transcript level may be the regulating point for carotenoid biosynthesis in this tissue.

In the course of analyzing *Pds* cDNA clones, a chimeric clone was found containing part of *Pds* cDNA and part of a novel receptor-like protein kinase cDNA. Using the partial cDNA sequence of this novel kinase with the combination of RACE (rapid amplification of cDNA ends) and RT-PCR techniques, a small gene family that encodes three different, but sequence-related, receptor-like protein kinases was cloned and characterized. The predicted amino acid sequence showed an extracellular leucine-rich repeat (LRR) domain, a transmembrane domain and an intracellular protein kinase domain and therefore, it was designated as a leucine-rich repeat transmembrane protein kinase (LTK). One of the genes, *Ltk1,* was mapped near *vp5,* and was found closely linked to *Pds*. While the expression of *Ltk1* was not tissue-specific, the expression of *Ltk2* and *Ltk3* was restricted to the endosperm. During the development of maize endosperm, *Ltk1* and *Ltk2* were temporally expressed at the RNA level, implying possible roles of LTK proteins in endosperm and seed development.

Acknowledgment

I thank the members of my advisory committee for their critical comments and suggestions. I am especially grateful to Dr. Eleanore Wurtzel, my mentor, for her constant advisement and help. I am also grateful to Dr. Manfred Philipp for his kindness and earlier reviews of this thesis. I thank Dr. Sharon Cosloy, Dr. Shirley Raps and Dr. Thomas Schmidt-Glenewinkel for their generosities to serve in my committee.

I wish to acknowledge every member in Dr. Wurtzel's laboratory for their friendship and help. I thank Arulmolee Yoganathan and Paul D. Matthews for critical reading of this thesis and help in the laboratory work. I thank Jia Yu for her friendship and help in the field work. I am especially grateful to Ruibai Luo, my classmate and friend, for her continuous help in HPLC analysis, field work and reading of this thesis. I am also grateful to Dr. Miguel Cervantes-Cervantes for his critical reading and revision. I thank Drs B. Meurer-Grimes and V. Upasani for assistance in HPLC analysis and Yingzhen Yang for assistance in genomic DNA library screening. I thank Dr. Benjamin Burr, Dr. F. Burr and Eileen Matz for the RFLP analysis. I thank Dr. N. Misawa for providing pACCRT-EB plasmid and S. Briggs, D. Robertson, and the Maize Coop for providing maize seeds. I thank Dr. T. Jensen for providing teaching assistantships, Ms. D. Vitanza for her kind help in administration and Mr. D. Cain for his help in growing corn in the field and green house. I thank Dr. Horst Schulz and Ms. S. Weintraub for their help through all these six and half years.

I am extremely grateful to my wife, Juan Zhang. Without her love, patience and support, it would have taken an even longer time to complete this thesis. For this, I thank her deeply. I thank my parents, Shiya and Shengting Li, for all their life long love, friendship and encouragement. I thank my sister, Xian Li and my brother-in-law, Huan Tao for taking care of our parents in China so that I could concentrate on my study in the USA. I thank my mother-in-law, Xiangfeng Wan; my sister-in-law, Yuan Zhang and my brother-in-law, Heng Lu for their encouragement and help.

This work was funded by grants from The Rockefeller Foundation International Rice Biotechnology Program; The National Institutes of Health-Minority Biomedical Research Support Program (5S06GM/AG08225-10); The National Science Foundation (RII-8805140);The City University of New York (CUNY-HEAT and PSCUNY research awards); and the Black Rock Forest Consortium.

Table of Contents

List of Abbreviations

aa	amino acid
AAP	Abridged Anchor Primer
ABA	absisic acid
AUAP	Abridged Universal Amplification Primer
bp	base pair
BSA	bovine serum albumin
cDNA	complementary DNA
cl	chlorophyll
CLA1	Clavata1
Clm	modifier of *cl*
Cpn	chaperonin
crtD	carotenoid gene D
crtI	carotenoid gene I
DAP	days after pollination
DEPC	diethyl pyrocarbonate
DMAPP	dimethylallyl pyrophosphate
DPA	diphenylamine
DTT	dithiothreitol
E. coli	*Escherichia coli*
ENOD40	early nodulation gene 40
ER	ERECTA
FAD	flavin adenine dinucleotide
FEN	allele sensitive to the insecticide fenthion
GPP	geranylgeranyl pyrophosphate
GGPPS	geranylgeranyl pyrophosphate synthase
GSP	gene specific primer
GUS	β-glucuronidase
HPLC	high-performance liquid chromatography
Hsp	heat shock protein
IPP	isopentenyl pyrophosphate
IPTG	isopropylthio-β-galactoside
IUPAC-IUB	International Union of Pure and Applied Chemistry-International Union of Biochemistry
kb	kilo base pair
kDa	kilo Dalton
LB	Luria-Bertani
LTK	leucine-rich repeat transmembrane protein kinase
LRR	leucine-rich repeat
lw	lemon white
MPDS	maize phytoene desaturase
Mum	mutator

NFZ	norflurazone
ORF	open reading frame
OsTMK	transmembrane kinase from *Oryza sativa*, L.
PCR	polymerase chain reaction
PDS	phytoene desaturase
PEST	hydrophilic region containing proline (P), glutamic acid (E) or aspartic acid (D), serine (S) or threonine (T).
PPPP	prephytoene pyrophosphate
PR5	pathogenesis-related protein 5
PRK1	pollen receptor-like kinase
PRR	proline-rich region
ps	pink scutellum
PSY	phytoene synthase
PTI	PTO-interacting protein
PTO	allele resistance to *Pseudomonas syringae* pv *tomato* strains expressing the avirulence gene *avrPto*
RACE	rapid amplification of cDNA ends
rbcL	ribulose bisphosphate carboxylase large subunit gene
RFLP	restriction fragment length polymorphism
RLK	receptor-like protein kinase
RPK1	receptor-like protein kinase one
RT-PCR	reverse transcriptase polymerase chain reaction
Rubisco	ribulose bisphosphate carboxylase
SERK	somatic embryogenesis receptor-like kinase
Sh	Shrunken
SRR	serine-rich region
TATA	conserved AT-rich sequence upstream of the transcription start site
TCTP	translationally controlled tumor protein
TMK1	transmembrane kinase one
TMKL1	transmembrane kinase-like one
Tn	transposon
TNFR	tumor necrosis factor receptor
TPDS	tomato phytoene desaturase
TRK	tyrosine receptor kinase
vp	viviparous
w	white
Wc	white cap
XA21	allele resistance to *Xanthomonas oryzae* pv. *oryzae* race 6
y	yellow
ZDS	ζ-carotene desaturase

List of Illustrations

Chapter 1: Background and introduction.

Carotenoids are polyene pigments ranging from colorless to yellow, pink and orange. They are prevalent in the plant kingdom, and in both photosynthetic and non-photosynthetic bacteria, but are less abundant in the animal kingdom, especially in humans, who need them for the synthesis of vitamin A and retinoids (Britton et al., 1995). Because of this imperative requirement by humans, the study of carotenoids has been pursued for more than 175 years, encompassing a variety of disciplines including chemistry, biochemistry, biology, genetics and recently, molecular biology.

To date, more than 600 kinds of carotenoids have been described. Some of them have been synthesized *in vitro*, especially β-carotene, which is the major active precursor for synthesizing vitamin A and retinoids (Britton et al., 1995). Plant physiological studies demonstrate that carotenoids, most of them derivatives of β-carotene, play important roles in photosynthesis as components of light harvesting centers and as protectors against photooxidation (Bartley and Scolnik, 1995). In the past twenty years, through biochemical, genetic and molecular biological approaches, scientists have made enormous progress in the study of carotenoid biosynthesis in bacteria, fungi, algae and higher plants. A core carotenoid biosynthesis pathway has been proposed and dozens of genes coding enzymes of the pathway have been cloned (for review, see (Bartley and Scolnik, 1995; Bramley, 1985; Sandmann, 1994)). Maize (*Zea mays*) has played an important role in the study of carotenoid biosynthesis because various mutations that affect its carotenoid biosynthetic pathway have been described (Robertson, 1975).

1.1 Structure, chemical features and biological function of carotenoids

Carotenoids are a class of hydrocarbons (carotenes) and their oxygenated derivatives (xanthophylls). A typical carotenoid is lycopene (Figure 1-1A), which consists of eight isoprenoid units joined together with a reverse at its center. Thus, the two central methyl groups are in a 1,6-positional relationship while the remaining non-terminal methyl groups are in a 1,5-positional relationship (as defined by the IUPAC-IUB (International Union of Pure and Applied Chemistry-International Union of Biochemistry) (IUPAC, 1972). In plants, other carotenoids can be considered as modifications of lycopene.

These modifications of the lycopene structure are (i) hydrogenation, (ii) dehydrogenation, (iii) cyclization, (iv) oxidation; or combinations of these processes. These modifications occur primarily at either one or both ends of the carotenoid skeleton, a polyene chain with conjugated carbon-carbon double bonds (Weedon and Moss, 1995). Stereochemically, carotenoids present different geometrical isomers (*trans* or *cis*; *Z* or *E*) and optical isomers (d or l), which account for more than 600 structures currently known (Eugester, 1995).

The carotenoid skeleton structure is a chromophore because of its light absorption properties, which give it color. Cyclization of the ends of the carotenoid skeleton intensifies or changes the characteristic light absorption of each carotenoid, a process that fills the world with colorful flowers and fruits, colorful fish skin and bird feathers. The color of flowers will attract insects for pollination, while the color of fruits will allure animals to eat them and spread the seeds they contain. The color of fish skin and bird feathers either provides attraction for mating or protection from predators (Purves et al., 1995).

The light absorption properties of carotenoids allow them to play an important role in light harvesting in photosynthetic systems. The major carotenoids in photosystem I and photosystem II of green leaves are β-carotene, lutein, neoxanthin, antheraxanthin, zeaxanthin and violaxanthin, whereas fucoxanthin is the major carotenoid of the photosynthetic membrane of brown algae (Siefermann-Harms et al., 1985). Most of these carotenoids are protein-bound in the photosynthetic apparatus and can receive energy from 400-500 nm light (Young, 1991). This energy will then be transferred to chlorophylls or bacteriochlorophylls by singlet-singlet excitation between carotenoid and chlorophyll or bacteriochlorophyll molecules (Young and Frank, 1996).

Besides harvesting light, carotenoids also provide protection for photosynthetic organisms from light-mediated stress. This photoprotection includes two mechanisms: (1) through either tracking down or preventing the formation of singlet-oxygen, and (2) through dissipation of excess excitation energy (Young, 1991). Singlet-oxygen is mainly generated by bacteriochlorophyll or chlorophyll triplets, which result from light illumination. Carotenoids can reduce the lifetime of the triplet-state of bacteriochlorophyll or chlorophyll and prevent the formation of singlet-oxygen that can destroy cell components by oxidation. When photosynthetic tissues are illuminated beyond the requirement of photosynthesis, excitation energy builds up excessively, potentially damaging these tissues by destroying the photosynthetic complex. To prevent this, the excess energy needs to be dissipated by photochemical and non-photochemical quenching into heat (Weis and Berry, 1988), fluorescence (Demming et al., 1990; Demming et al., 1989; Phillip et al., 1996), or chemical energy (ΔpH or phosphorylation) (Noctor et al., 1991; Rees et al., 1989). There is evidence that the xanthophyll zeaxanthin is involved in the non-photochemical or radiationless energy dissipation in the leaves (Demming et al., 1990).

In plants, certain carotenoids are precursors of abscisic acid (ABA), a plant hormone regulating plant differential growth, leaf abscission, embryo development and dormancy and stress response (Parry and Hogan, 1991; Zeevaart and Creelman, 1988). Some maize mutants that are blocked in the biosynthesis of carotenoids (see part 1.4. Carotenoid biosynthesis mutants in maize) are viviparous (do not undergo dormancy because of a presumed lack of ABA), which suggests that ABA is a by-product of carotenoid biosynthesis (Zeevaart and Creelman, 1988), Fong and Sandemann, 1984).

In animals, including humans, carotenoids can serve as precursors for the synthesis of retinoids or vitamin A, which is a breakdown product of β-carotene. People with vitamin

A deficiency, especially children, will develop either night blindness or even irreversible blindness. An investigation by Sommer (Sommer, 1988) estimated that each year, five million children in Southeast Asia develop xerophthalmia, which can lead to irreversible blindness. Vitamin A deficiency can also increase susceptibility to several potentially fatal illnesses, such as diarrhea, respiratory diseases and measles (Berman, 1991; Bhaskaram, 1995 ; Tomkins, 1991). Recent studies have shown that retinoids can serve as a signal during embryonic development (Bertram and Bortkiewicz, 1995) hence, the lack of retinoids will create severe developmental defects in humans (Elmazar et al., 1996; Morriss-Kay and Sokolova, 1996). Since humans cannot produce vitamin A *de novo* but must either absorb it directly from the diet (meat or liver stored vitamins A) or convert it from dietary precursors (vegetables containing β-carotene), a supplement of vitamin A or β-carotene in the diet can overcome vitamin A deficiency (Sommer, 1989).

Other experiments have suggested that vitamin A can also intensify the function of the human immune system (Ross, 1992). Bertram and his colleagues have further shown that carotenoids are chemopreventives for human cancer, through inhibiting carcinogen-induced neoplastic transformation, inhibiting plasma membrane lipid oxidation, and inducing up-regulation of the expression of *connexin 43*, a gene coding for the structural unit of gap junctions (Bertram and Bortkiewicz, 1995).

1.2 Carotenoid biosynthetic pathway and related enzymes

Studies from several groups have demonstrated that carotenoid biosynthesis, beginning with the synthesis of phytoene, takes place inside the plastids (plants and algae) or within a combination of organelle fractions (fungi) of eukaryotic cells, whereas it takes place in the cytosol and photosynthetic membrane of prokaryotic cells (Beyer et al., 1985; Bramley, 1994; Camara et al., 1982; Linden et al., 1993). As shown in Figure 1-2, carotenoid biosynthesis starts with condensation of two C_{20} molecules of geranylgeranyl pyrophosphate (GGPP, Figure 1-1B), a precursor for several terpenoid pathways, into one C_{40} molecule, phytoene (Figure 1-1D), the first specific precursor for the carotenoid biosynthetic pathway. This condensating reaction is catalyzed by phytoene synthase (PSY) and includes the formation of an intermediate molecule, prephytoene pyrophosphate (PPPP, Figure 1-1C). PSY is localized in the plastid stroma of higher plants and algae [Camara, 1982 #162, Schiff *et al.*, 1981); in the endoplasmic reticulum of fungi, as a peripheral membrane protein (Mitzka-Schmabel and Rau, 1981) and in the cytosol or membrane as a multienzyme complex in bacteria (Gregonis and Rilling, 1974). However, there is evidence, from pea chloroplast importing experiments, that imported PSY can either bind to the Cpn60-complex, remaining in the stroma, or associate with the thylakoid membrane [Bonk, 1997 #184].

Tomato has two different phytoene synthases encoded by two different genes, which are expressed in a tissue specific way (Bartley and Scolnik, 1993; Bartley et al., 1992). Other organisms examined contain only one *Psy* gene encoding one enzyme (Bucher, 1990; Dogbo et al., 1988; Karvouni et al., 1995; Scolnik and Bartley, 1993). Phytoene molecules produced by PSYs have predominantly 15-*cis* configurations from all carotenoid producing organisms although other carotenoids synthesized by these organisms are in all-*trans* configurations.

Following the synthesis of phytoene, four desaturation (dehydrogenation) steps are carried out to convert colorless phytoene into the pink-colored lycopene by the addition of four carbon-carbon double bonds onto the carotenoid skeleton. In the first step, one double bond between C11-C12 is added, thus forming phytofluene (Figure 1-1E); the second desaturation step between C11'-C12' forms ζ-carotene (Figure 1-1F). Two more desaturation steps (step three and four) add another pair of double bonds to the ζ-carotene skeleton at symmetric positions (C7-C8 and C7'-C8'), creating neurosporene and lycopene, respectively. Each desaturation step will produce two hydrogen ions and two electrons, which then can be transferred to the photosynthetic electron transport chain in the chloroplast, using plastoquinone as an electron carrier (Norris et al., 1995). In daffodil chromoplasts, oxygen was thought to be the electron acceptor (Beyer et al., 1989) but quinone compounds might also play a role in transfer of the electrons (Mayer et al., 1992). Along with the increasing number of conjugated double bonds, the chromophore feature of the carotenoid backbone is revealed as the color of these carotenoids changes from the colorless phytoene to the yellow ζ-carotene to the deep yellow neurosporene and to the pink lycopene.

In different organisms, desaturases catalyze these desaturation reactions differently. In oxygenic photosynthetic organisms, two carotenoid desaturases catalyze the four desaturation reactions, each one catalyzing a two-step desaturation (reviewed by (Bartley and Scolnik, 1995)). The first one, phytoene desaturase (PDS) uses phytoene or phytofluene as substrate, producing ζ-carotene. (Al-Babili et al., 1996; Armstrong et al., 1989; Bartley et al., 1991; Chamovitz et al., 1991; Hable and Oishi, 1995; Hugueney et al., 1992; Pecker et al., 1992; Scolnik and Bartley, 1993). The second one is ζ-carotene desaturase (ZDS), which can convert ζ-carotene into lycopene (Albrecht et al., 1995; Linden et al., 1994). Even though they use different substrates, PDS and ZDS have certain homology in their amino acid sequences, although this homology is lower than the homology between members of each enzyme. One common feature of these two enzymes is that, like other desaturases, each contains a dinucleotide-binding site at its amino terminus. Another common feature is that both proteins are longer in higher plants than in prokaryotic cells. This is due to the presence of a transit peptide in the eukaryotic versions of the enzymes.

In contrast to oxygenic photosynthetic organisms, a single desaturase, CRTI, catalyzes the desaturation of phytoene to either neurosporene (Armstrong et al., 1989) or lycopene (Misawa et al., 1990) in non-oxygenic photosynthetic organisms or nonphotosynthetic organisms. This kind of desaturase is designated as the *crtI*-type, which is different from the *Pds*-type from oxygenic photosynthetic organisms (Sandmann, 1994). The *crtI*-type enzymes can then be ascribed to two subclasses: those that catalyze three desaturation steps from phytoene to neurosporene (e.g. *Rhodobacter*), or those that catalyze four desaturations to lycopene (e.g. *Erwinia, Neurospora*). Protein sequence comparison reveals that the *crtI*-type and the *Pds*-type desaturase have little homology, except for the dinucleotide-binding site (reviewed by (Sandmann, 1994)). *In vitro* enzyme activity analysis suggested that both types of desaturase can use flavin adenine dinucleotide (FAD) as hydrogen ion and electron receptor. However, the activities of these enzymes are inhibited by different kinds of herbicides. The *crtI*-type enzyme can be inhibited by diphenylamine (DPA), whereas the *Pds*-type enzyme is inhibited by norflurazon (NFZ) (Sandmann, 1994).

When the amino acid sequence of ZDS from the cyanobacterium *Anabaena* was compared to the sequences of two types of phytoene desaturase, it was found that it had higher homology to the *crtI*-type than to the *Pds*-type (Linden et al., 1994). This finding was inconsistent with the proposal that carotenoid biosynthetic enzymes evolved independently in photosynthetic and non-photosynthetic organisms (Pecker et al., 1992). The *Anabaena* ZDS could be an exception, since non-photosynthetic organisms do not contain a second desaturase for the desaturation of phytoene to lycopene (Sandmann, 1994). However, when *Capsicum* ZDS was compared to the *Pds*-type PDS and *Anabaena* ZDS, the homology of the amino acid sequence between *Capsicum* ZDS and the *Pds*-type PDS was higher than that between the two ZDS proteins (Albrecht et al., 1995). The high homology between the *Anabaena* ZDS and the *crtI*-type desaturase was also consistent with the finding that, as it occurs with the *crtI*-type desaturase, *Anabaena* ZDS activity could not be inhibited by NFZ (Albrecht et al., 1995). Furthermore, the high homology between the *Capsicum* ZDS and the *Pds*-type desaturase suggests that, in higher plants, *Pds* and *Zds* may have evolved from the same ancestral gene.

Because phytoene and ζ-carotene are either hydrophobic or lipophilic, it is believed that enzymes catalyzing desaturation are intrinsic membrane proteins. This theory has been proven by the *in vitro* analysis of PDS activity with a membrane fraction of thylakoids isolated from *Synechocystis* (Serrano et al., 1990) and chloroplasts or chromoplasts separated from different plants (Beyer et al., 1985; Camara et al., 1982; Lutke-Brinkhaus et al., 1982). Using immunogold localization with an anti-PDS antibody, it was demonstrated that the majority of the PDS molecules are localized within the thylakoid membranes of higher plant chloroplasts (Linden et al., 1993). However, recent data from western analysis of PDS from the chromoplasts of daffodil (*Narcissus pseudonarcissus*) flowers show that the enzyme is not only membrane-bound but it is also present as a soluble protein bound to the Hsp70-complex in the plastid stroma (Al-Babili et al., 1996).

Recently, two different genes for lycopene cyclase, which catalyzes the cyclization of lycopene ends, have been cloned from *Arabidopsis* (Cunningham et al., 1996). One of them, the ε-cyclase, catalyzes the cyclization of lycopene with a single ε-ring forming δ-carotene; the other, β-cyclase, catalyzes the formation of β-carotene by adding two β rings at each end of lycopene (Figure 1-2). The action of both lycopene cyclases together, adding one β-ring at one end and one ε-ring at the other end of lycopene forms α-carotene. Lycopene cyclase genes of plants, algae and cyanobacteria have higher homology among them than those of non-photosynthetic bacteria (Cunningham et al., 1996; Hugueney et al., 1995; Pecker et al., 1996), which may be evidence that carotenogenic enzymes evolved independently in photosynthetic and non-photosynthetic organisms. In bacteria, lycopene cyclase is membrane-bound (Sandmann, 1994).

Most carotenoid-producing organisms share the pathway up to the step of β-carotene and α-carotene formation (Armstrong, 1994). The pathway diverges among different organisms or different tissues within the same organism. In leaf tissue, α-carotene may be converted to lutein by a yet unidentified α-carotene hydroxylase (Sandmann, 1994), whereas β-carotene is converted into zeaxanthin by β-carotene hydroxylase (Sun et al., 1996). Subsequently, zeaxanthin is converted into the epoxy-xanthophylls violaxanthin or neoxanthin by zeaxanthin epoxidase, whose gene was recently cloned from *Nicotiana plumbaginifolia* (Marin et al., 1996). These epoxy-xanthophylls are substrates of dioxygenases and can be cleaved into C_{25} apo-aldehydes and xanthoxin, the precursor for ABA biosynthesis. The recently cloned *Vp14* gene of maize encodes one of these dioxygenases; the function of VP14-specific oxidative cleavage of 9'-*cis*-epoxy-xanthophylls has been demonstrated *in vitro* (Schwartz et al., 1997). Because all carotenoids synthesized after phytoene are in an all-*trans* configuration, the *trans-cis* isomerization of epoxy-xanthophylls (Figure 1-2) may involve an isomerase.

1.3 The regulation of carotenoid biosynthesis in higher plants

The regulation of carotenoid biosynthesis in higher plants is a complex process: (1) the enzymes involved in the pathway are nuclear-encoded but function in plastids; nuclear-organelle interaction may influence the pathway; (2) the plastids in different tissues originate differently and may regulate the pathway differently; (3) considering that the precursor for phytoene synthesis, GGPP, is also a precursor for other terpenoids, the regulation of carotenoid biosynthesis will affect the regulation of the synthesis of other terpenoids, and *vice versa*; (4) the pathway can be regulated within itself by regulating the expression of enzymes that catalyze the synthesis of phytoene and thereafter at transcriptional, translational and posttranslational levels; (5) the pathway can be regulated upstream before the synthesis of phytoene. Because of these reasons, it is hard to categorize data on carotenoid biosynthesis regulation obtained from different species or tissues. However, results on regulation of carotenoid biosynthesis are summarized here into two categories: within and upstream of the pathway itself.

Likely points of regulation within the pathway are PSY and PDS because they are, respectively, the first and second enzymes specific to the pathway. Tomato, as mentioned before, contains two different *Psy* genes that are differentially regulated in leaf and fruit. It has been found that *Psy1* is predominantly expressed in seedlings and at the late stages of fruit development, whereas *Psy2* is expressed mostly in the mature leaf (Bartley and Scolnik, 1993). In tomato leaf, both *Psy1* and *Psy2* transcript levels increase more than four-fold than those in etiolated seedlings (Bartley and Scolnik, 1993) though *Psy1* expression is constitutive during photomorphogenesis, when chloroplasts develop from etioplasts in the presence of light (Giuliano et al., 1993). However, in tomato fruit, *Psy1* transcripts increase when chloroplasts become chromoplasts, while the *Psy2* transcripts remain constant throughout the fruit development (Bartley and Scolnik, 1993). Two different sizes of *Psy* transcripts have been observed in *Capsicum* leaf and fruit, although it is not known how these two different transcripts are generated (Romer et al., 1993). The induction of *Psy* expression at the RNA level was also observed during *Capsicum* fruit ripening (Romer et al., 1993). Similar upregulating of *Psy* was also observed during the ripening of melon fruits (Karvouni et al., 1995).

In daffodil flowers, the upregulation of carotenoid synthesis is exerted by increasing the expression of *Psy* at levels higher than those in leaves, both at the transcriptional and translational levels (Schledz et al., 1996). The increase in *Psy* expression, however, does not occur during the chloroplast-chromoplast transition but rather at an early stage (closed bulb) of flower development. In the tomato flower, the increment in *Psy1* expression at the RNA level happens in two-steps. The first step is similar to that in the daffodil flower in which the early developing flower contains higher amounts of *Psy1* transcripts with respect to the mature leaf. The second step parallels flower development where *Psy1* transcripts reach a peak before anthesis (Giuliano et al., 1993). Not surprisingly, this second step of upregulation of *Psy1* in tomato flower occurs during the chloroplast-chromoplast transition and the concomitant carotenoid accumulation. Furthermore, overexpression of *Psy* in transgenic plants led to increased production of carotenoids (Fray et al., 1995; Kumagai et al., 1995) whereas expression of antisense RNA of *Psy* in transgenic plants blocked carotenoid synthesis (Bramley et al., 1992).

Based on these results, it has been suggested that *Psy* is a putative site for controlling carotenoid biosynthesis in higher plants. The increase of carotenoid content can be achieved by simply increasing the expression of the first carotenogenic enzyme, PSY, at the RNA and/or protein level. However, when *Psy* was transferred into rice and overexpressed in its endosperm, an amyloplast containing tissue, phytoene accumulated without the final product, carotenoids (Burkhardt et al., 1997). Since rice does not accumulate any carotenoid in its endosperm (Burkhardt et al., 1997), the failure of carotenoid overproduction by increasing *Psy* expression in the transgenic rice may due to a lack of functional enzymes of the carotenoid biosynthetic pathway in rice endosperm. Studies presented in Chapter 3 on *Psy* expression during maize endosperm development strongly support this hypothesis. However, recent data obtained from study of light-dependent regulation of carotenoid biosynthesis in *Sinapis alba* and *Arabidopsis* have indicated that elevating *Psy* at RNA level is necessary, but not sufficient, for increasing the carotenoid content during photomorphogenesis (von Lintig et al., 1997). Whether a similar regulatory mechanism of carotenoid biosynthesis is shared between rice amyloplast and *Arabidopsis* chloroplast needs more investigation.

Like *Psy*, *Pds* transcript level is upregulated in the developing tomato fruit along with the transition of chloroplast to chromoplast (Fraser et al., 1994; Giuliano et al., 1993; Pecker et al., 1992). As in tomato fruit, *Pds* expression increases at both the transcriptional and translational level during *Capsicum* fruit maturation (Hugueney et al., 1992). A similar result was obtained in the study of developing tomato flower, in which the *Pds* transcript level was elevated 9-fold (Giuliano et al., 1993). In contrast, the increase of *Pds* expression at the RNA level, like *Psy* expression, occurred at an early stage of development in daffodil flowers and remained constant during chromoplast development from chloroplast (Al-Babili et al., 1996). However, levels of PDS protein were elevated during daffodil flower development, which suggests a post-transcriptional regulation of *Pds* expression in developing daffodil flower (Al-Babili et al., 1996). Furthermore, Al-Babili also found that PDS in the daffodil flower chromoplast was present as an FAD-containing active form, bound to the membrane and as an inactive species, forming a soluble complex containing both chaperonin 60 (Cpn60) and heat-shock protein (Hsp70) in the stroma (Al-Babili et al., 1996; Bonk et al., 1996). The activation of PDS includes binding of FAD and association with the membrane. Though the mechanism of PDS activation in the chromoplast is still unknown, this process implies a posttranslational regulation of PDS expression. Recently, using an *in vitro* chloroplast importing system, Bonk *et al.* demonstrated that the imported PDS can also constitute a soluble form with Cpn60 in the stroma and a membrane-bound form in the thylakoid membrane of chloroplast (Bonk et al., 1997).

Unlike *Psy*, the expression of *Pds* at the RNA level cannot be increased by light, although carotenoid content has been shown to increase in development of chloroplast from etioplast in tomato leaf (Giuliano et al., 1993). This result again indicates a post-transcriptional regulation of *Pds* expression in tomato leaf. However, when a GUS reporter gene was fused with the tomato *Pds* promoter in transgenic tomato plants, it was found that PDS/GUS expression was higher in etiolated seedlings that contained fewer carotenoids than in light-grown seedlings (Corona et al., 1996). Inhibitors that block carotenoid biosynthesis could also induce the transcription of PDS/GUS. These results suggested an end-product (or feed-back) regulation of *Pds* expression in green tissues (Corona et al., 1996). Similar results were obtained with increasing expression of both *Psy* and *Pds* in tomato seedlings treated with the PDS inhibitor NFZ (Giuliano et al., 1993). Furthermore, blocking of a particular enzyme of the pathway can initiate the accumulation of a particular carotenoid. For example, the gene encoding lycopene cyclase is down regulated during chromoplast development and lycopene is the major carotenoid accumulated in tomato fruits (Pecker et al., 1996).

Regulation of the carotenoid biosynthetic pathway was also observed upstream, before phytoene synthesis. Albrecht and Sandmann showed that light can up-regulate carotenoid biosynthesis during etioplast to chloroplast transformation in maize by increasing the enzyme activity of isopentenyl pyrophosphate (IPP) isomerase (Albrecht and Sandmann, 1994). It was proposed that an increase in IPP isomerase activity could increase the amount of DMAPP (dimethylallyl pyrophosphate) the substrate of geranylgeranyl pyrophosphate (GGPP) synthase (Badillo et al., 1994), whose product, GGPP, is the substrate of PSY.

It is still unclear why increasing IPP isomerase activity will increase only the production of carotenoids, because GGPP is a universal precursor for terpenoid synthesis (McGarvey and Croteau, 1995). A recent study in green algae by radioactive labeling of precursors suggested a novel pathway of IPP synthesis through pyruvate/glyceraldehyde 3-phosphate rather than mevalonate (Schwender et al., 1996). Although the data from this study contradicted the proposed pathway of isoprenoid synthesis (McGarvey and Croteau, 1995), it still suggested that terpenoids were synthesized from the same IPP or GGPP precursors. It was found that the regulation of terpenoid biosynthetic pathways was associated with each other (Fraser et al., 1995) and that alteration of one pathway could affect another (Fray et al., 1995). These results suggest that the specific regulation of one pathway (i.e. carotenoid biosynthesis) may be accomplished by control within the pathway (i.e. regulation of PSY expression) or by altering other pathways (i.e. gibberellin biosynthesis).

1.4 Carotenoid biosynthesis mutants of maize

Maize has been a classic material of genetic study for more than a century. Many mutants that affect carotenoid biosynthesis have been identified and mapped on different chromosomes (for maize database, see http://www.agron.missouri.edu). In a review by Robertson (Robertson, 1975), more than 10 mutants that had phenotype of white or pale yellow endosperm and albino seedlings were reported with several alleles for each one. Some of them - *vp2*, *vp5*, *vp7*, *vp9*, *w3*, and *y9* - were viviparous and accumulated carotenoid precursors, while others, *lw*, *lw2*, *lw3*, *lw4*, *cl*, *w7748* and *y1* were not viviparous and did not accumulate any precursors. Other viviparous mutants found in maize are *vp1* (Ester, 1931), and the recently identified *vp12* (Maluf et al., 1997) and *vp14* (Schwartz et al., 1997). One of the interesting white mutants was *cl* whose albino seedling phenotype was exhibited only when a corrresponding dominant allele at the modifier *Clm* locus was absent (Robertson et al., 1966). A dominant mutaion, *Wc1*, can affect the carotenoid accumulation in maize endosperm with a dosage effect (Kulkarni, 1927). Analysis of the accumulation precursors provided strong evidence for the proposed carotenoid biosynthetic pathway (Robertson et al., 1978).

Maize mutants with impaired carotenoid biosynthesis not only define components of the pathway, but also provide material for molecular genetic studies. The first gene related to the carotenoid pathway in higher plants was cloned by transposable-element tagging from the maize *y1* mutant (Buckner et al., 1990). Although *y1* had been considered as a regulatory gene because of its dosage effect (Mangelsdorf and Fraps, 1931; Randolph and Hand, 1940), it was found, based on sequence homology, that the *Y1* gene encodes maize PSY (Buckner et al., 1996). The recessive homozygous mutant, *y1/y1*, showed a defect in phytoene synthesis and carotenoid accumulation in endosperm and sometimes in leaves (Robertson and Anderson, 1961). Therefore endosperms with the *y1/y1* genotype are albino or pale yellow, as compared to dominant homozygous (*Y1/Y1*) or heterozygous

(*Y1/y1*) endosperms. The leaves of most plants with the *y1* allele are green like those with the *Y1* allele (Robertson and Anderson, 1961), which has led to the speculation that there are two *Psy* genes in maize, as in tomato (Bartley and Scolnik, 1993).

Four distinctive maize mutants, *vp2*, *vp5*, *vp9* and *w3*, block the four desaturation steps of phytoene into lycopene. Genetic analyses have shown that all four are single, nonallelic recessive mutations, and they have been mapped to chromosomes 5, 1, 7, and 2, respectively (Robertson, 1975). Early biochemical analysis of the pigments revealed that recessive homozygous *vp2* and *vp5* accumulated both phytoene and phytofluene, while *w3* accumulated phytoene, phytofluene and ζ-carotene (Robertson et al., 1978). Using HPLC analysis, Neill *et al.* determined that endosperms and embryos from *vp2* and *vp5* mutants accumulated phytoene, while *vp9* accumulated ζ-carotene (Neill et al., 1986).

Block of carotenoid biosynthesis at the desaturation steps prevents conversion of the colorless phytoene to the pink lycopene, which otherwise would be cyclized and oxidized into other colorful carotenoids. Instead of accumulating yellow or orange carotenoids as in the wild type endosperm, the four mutants (*vp2*, *vp5*, *vp9* and *w3*) will accumulate either colorless or pale yellow carotenoids in their endosperms, which result in white or pale yellow kernels. In leaves of the young mutant seedlings, the lack of colored carotenoids (because of the blocking of phytoene desaturation) will diminish photoprotection, but surprisingly, might not affect the assembling of the light-harvesting center (see 1.1 Structure, chemical feature and biological function of carotenoids).

Cytological studies have shown that the plastids of phytoene desaturation mutants have plastid membrane structure similar to normal plastids when seedlings are grown in the dark, in dimlight or at the early stages after transfer to light (Robertson et al., 1978). When exposed to light, however, the mutant plastids cannot form normal grana structure due to the lack of photoprotection, usually provided by carotenoids; light will bleach chlorophyll and will make the plastid membrane structure unstable. This will lead to the albino seedling present in the four mutants with homozygous recessive genes.

As reviewed in 1.2 (Biosynthetic pathway of carotenoids and related enzymes), the plant hormone ABA, which maintains seed dormancy and prevents germination, is a by-product of carotenoid biosynthesis. Low ABA content was detected in embryos from *vp2*, *vp5*, *vp9* and *vp7* (whose recessive allele confers accumulation of lycopene); this is why these mutants are viviparous, or geminate before maturity (Neill et al., 1986). Other viviparous mutants found in maize are due to deficiency in ABA perception (such as *vp1*) (McCarty et al., 1991) or of ABA synthesis (such as *vp14*), although they synthesize carotenoids normally (Schwartz et al., 1997).

1.5 Objectives

1.5.1 Significance

To overcome the deficiency of vitamin A, we can chemically synthesize and provide it as a nutritional supplement. For developing countries, where most of children with vitamin

A deficiency reside, providing vitamin A would be an expensive process (Sommer, 1989). Rice, which contains no carotenoids in its endosperm, is the major food supply in many of these countries. Biological engineering is considered as an alternative way to overcome vitamin A deficiency in developing countries by conditioning carotenoid biosynthesis in rice endosperm. However, the cause of carotenoid biosynthesis defects in rice endosperm is unclear. To decipher this puzzle, studying a model system that can accumulate carotenoids in the endosperm will be significant and informative. Maize is an excellent candidate, since it is phylogenetically related to rice and it can synthesize and accumulate carotenoids in its endosperm. Also, the many mutants that block the carotenoid biosynthetic pathway in maize (cf. 1.4 Carotenoid biosynthesis mutants in maize) can provide useful information and material for molecular genetic analysis.

Study of the carotenoid biosynthetic pathway will not only benefit carotenoid biosynthesis research in rice, but will also help us understand the tissue-specific regulation of secondary metabolism in plants, the nuclear gene regulation of organelle biosynthetic pathways, and the evolutionary development of biosynthetic pathways in plastids.

1.5.2 Specific aims

As reviewed in section 1.3 two desaturases, PDS and ZDS, are involved in the desaturations from phytoene to lycopene in higher plants. However, these other studies were conducted in dicotyledonous plants. It is not known if maize (a monocotyledonous plant) contains four different genes encoding four desaturases or one desaturase with four subunits or two desaturases with two subunits each. However, since four mutants which block the four desaturation steps have been reported (section 1.4), it has been proposed that four different genes encoding four polypeptides catalyze these desaturation steps. Besides, genetic studies have shown that the phenotype of these desaturation mutants is different from that of *Psy* mutants. The *Pds* mutants have white seedlings and white kernels while the *Psy* mutants have, most of the time, white kernels only. This suggests that the expression of these genes is regulated differently. Understanding the differential regulation of *Psy* and *Pds* will further help us understand how the carotenoid biosynthetic pathway is regulated in maize.

More specifically, the questions addressed are:

(1) How many genes encode phytoene desaturase (PDS) in maize?

Cloning the cDNA of maize *Pds* and comparing the locus of *Pds* to genetic loci of mutants that can accumulate phytoene.

(2) How many desaturation steps are catalyzed by PDS?

Analyzing the function of the protein encoded by the cloned *Pds* cDNA.

(3) Is *Pds* regulated during carotenoid biosynthesis in maize endosperm?

Analyzing the expression of *Pds* at RNA level in the developing maize endosperm.

(4) Are *Psy* and *Pds* differentially regulated during maize endosperm development during carotenoid accumulation?

Comparing the expression of Psy and Pds genes at RNA level in the developing maize endosperm.

To answer these questions, cloning *Pds* gene will be a fundamental and crucial task. In Chapter 2, the successful cloning and characterization of *Pds* cDNA is reported. The continuing analysis of other *Pds* clones obtained by cDNA library screening, cloning of the *Pds* promoter region, and a comparison of *Pds* and *Psy* expression are presented in Chapter 3. Analysis of *Pds* clones led to the serendipitous discovery of a novel receptor-like protein kinase. The cloning of the gene encoding this novel receptor-like protein kinase and preliminary results of this new gene are described in Chapter 4. Finally, future studies of the *Pds* cDNA and promoter and the novel receptor-like kinase cDNAs are discussed in Chapter 5.

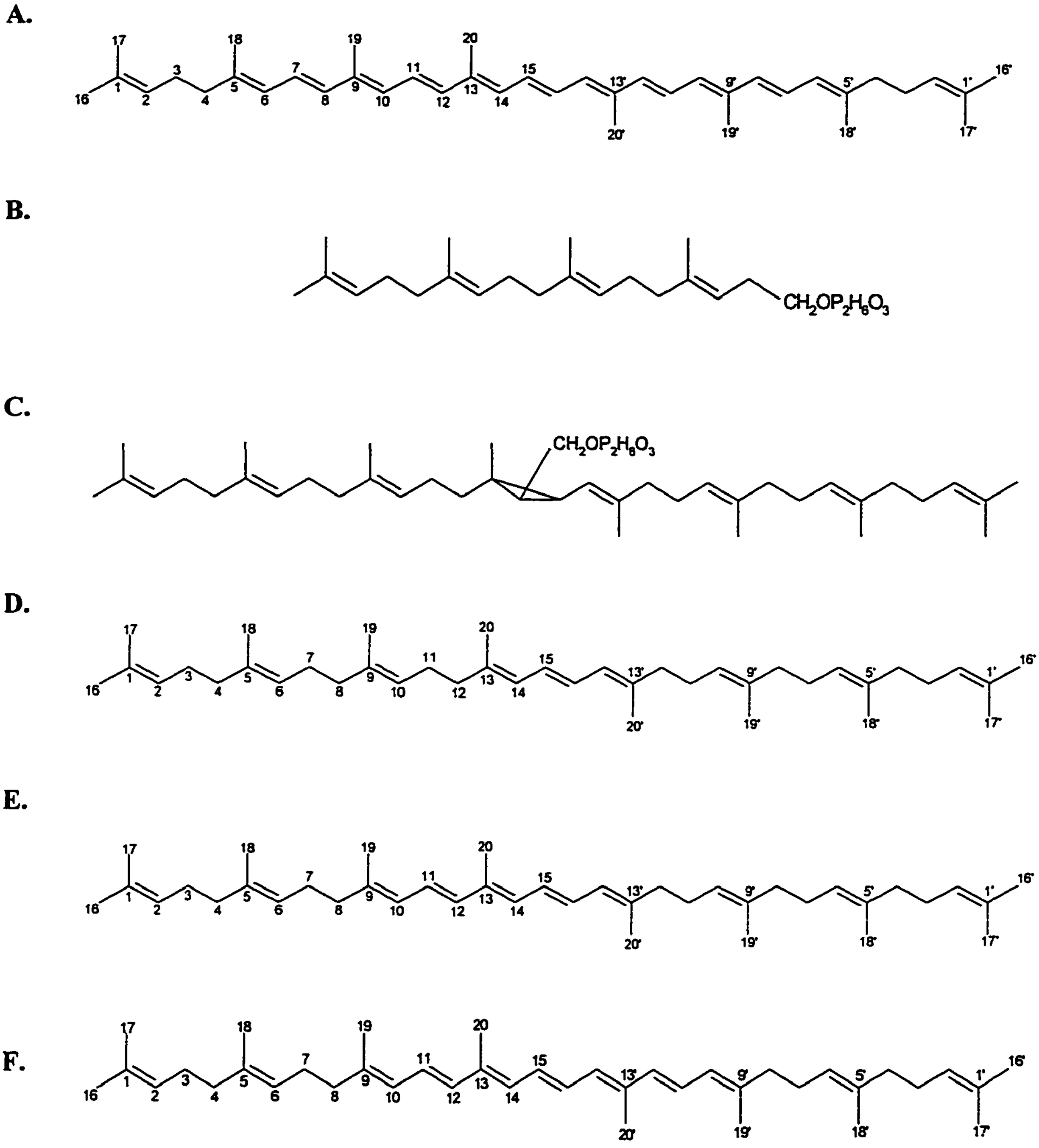


Figure 1-1. The chemical structure of several carotenoids. A, Lycopene; B, GGPP; C, PPPP; D, Phytoene; E, Phytofluene and F, ζ-carotene.

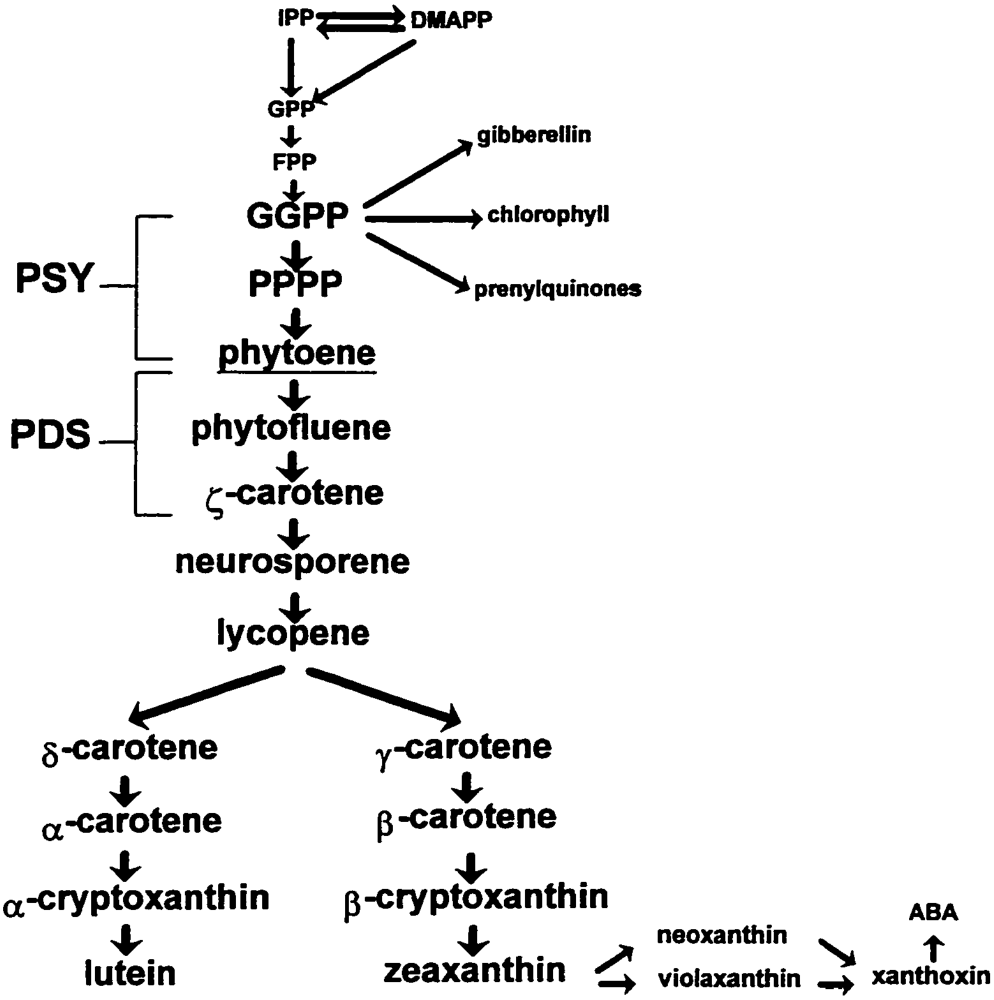


Figure 1-2. The carotenoid biosynthetic pathway with upstream and downstream products. The carotenogenic pathway is indicated with thick arrows, while upstream and downstream of the pathway are indicated with thin arrows. The first specific precursor of carotenoid biosynthesis, phytoene, is underlined. Genes of PSY (phytoene synthase) and PDS (phytoene desaturase) are studied in this dissertation. Other enzymes are described in the text.

Chapter 2: Cloning and characterization of a maize cDNA encoding phytoene desaturase, an enzyme of the carotenoid biosynthetic pathway[1].

2.1 Abstract

To study regulation of the plastid-localized maize carotenoid biosynthetic pathway, a cDNA encoding phytoene desaturase (PDS) was isolated and characterized. The DNA sequence of the maize *Pds* cDNA was determined and compared with available dicot *Pds* genes. The deduced PDS protein, estimated at 64.1 kD (unprocessed), had a dinucleotide binding domain and conserved regions characteristic of other carotene desaturases. Alignment of available PDS sequences from distantly related organisms suggests that *Pds* has potential as a phylogenetic tool. By use of heterologous complementation in *Escherichia coli*, maize PDS was shown to catalyze two desaturation steps converting phytoene to ζ-carotene. RFLP (restriction fragment length polymorphism) mapping was used to place *Pds* on chromosome 1S near *vp5*, and RT-PCR analysis indicated reduced *Pds* transcript in *vp5* mutant relative to normal endosperm. Other phytoene-accumulating mutant endosperms, *vp2* and *w3*, showed no difference in *Pds* transcript accumulation as compared with normal endosperm counterparts. RT-PCR analysis of *Pds* transcript accumulation in developing endosperm showed *Pds* was constitutively expressed.

---

[1] This Chapter is a revision of Li *et al*, (1996) Plant Mol. Biol. 30, 269-279 with the permission from the publisher.

Therefore, endosperm carotenogenesis is not regulated by increasing the level of *Pds* transcripts.

2.2 Introduction

Carotenoids serve multiple functions in plants; as accessory pigments in photosynthesis; as photoprotectors; and as precursors to the hormone, abscisic acid (ABA). In animals, carotenoids are essential precursors to Vitamin A and related compounds (reviewed in (Armstrong, 1994)). Carotenoids are synthesized and accumulated in plastids; these plastids include chloroplasts and nonphotosynthetic plastids, such as chromoplasts of fruits, flowers, some endosperms and roots (for reviews, see (Britton and Goodwin, 1982; Goodwin, 1976; Sandmann, 1991)).

Because of the varied roles and localization of carotenoids within plastids of different membrane architectures, I expect tissue-specific regulation of the pathway. Maize is an excellent model system to explore regulation of carotenoid biosynthesis because of the many mapped and biochemically characterized mutations blocking the pathway. These include recessive, dominant, and suppressor/modifier alleles (Robertson, 1975).

The synthesis of C_{40} carotenoids begins with condensation of two molecules of geranylgeranyl pyrophosphate to produce phytoene, a step catalyzed by phytoene synthase (PSY). Phytoene synthesis occurs in plastid stroma, whereas subsequent steps leading to synthesis of colored carotenoids occur on plastid membranes (Beyer et al., 1985; Kreuz et al., 1982; Lutke-Brinkhaus et al., 1982; Mayfield et al., 1986). This latter phase includes four sequential desaturations of phytoene. In dicots and cyanobacteria, these steps are catalyzed by two enzymes, PDS (phytoene desaturase) and ZDS (ζ-carotene desaturase), each mediating two steps (reviewed in (Bartley et al., 1994)). In other carotenogenic organisms, including fungi, nonphotosynthetic bacteria, and photosynthetic bacteria, one phytoene desaturase enzyme may catalyze up to four desaturation steps (reviewed in (Armstrong, 1994)). As shown in Figure 2-1, recessive alleles of four maize loci, *vp2*, *vp5*, *w3*, and *vp9*, block desaturation both in endosperm and plant (Neill et al., 1986; Treharne et al., 1966). It is presently unclear what functions are represented by these genes and whether these genes reflect an alternative array of overlapping or non-overlapping desaturation functions, each mediating a particular number of steps, or encode regulatory or ancillary functions such as pigment-binding proteins or oxidoreductases (Mayer et al., 1992).

With the exception of maize *Y1*, which has been shown by heterologous complementation to encode PSY [Buckner, 1990 #218; Yoganathan and Wurtzel, unpublished], there has been little characterization of genes encoding the biosynthetic enzymes in monocots. Genes encoding PDS have primarily been isolated from dicots (reviewed in (Bartley et al., 1994)) and the temporal regulation of their expression examined in chromoplasts (Fraser et al., 1994; Giuliano et al., 1993; Hugueney et al., 1992; Pecker et al., 1992). Here PDS transcript abundance appears to correlate with carotenoid accumulation.

Therefore, I decided to isolate the maize gene encoding PDS to characterize the temporal regulation of PDS expression in endosperm plastids (amyloplasts). Using RFLP mapping and analysis of steady-state levels of accumulated transcripts, I have associated a known maize carotenoid locus with a specific gene product, PDS. To demonstrate the function and determine the number of desaturation steps catalyzed by maize PDS, a heterologous complementation system was employed. Finally, the carotenoid biosynthetic pathway of maize endosperm was found not to be regulated by modulation of *Pds* transcript levels.

2.3 Materials and methods

2.3.1 Plant materials

For developmental studies, maize plants were grown under standard field conditions at the Black Rock Forest (Cornwall, New York). For other studies, plants were grown either in the field or in greenhouses at Lehman College, CUNY. Argentina high carotenoid line was obtained from Dr. S. Briggs (Pioneer Hy-Bred) and mutant lines were provided by Dr. D. Robertson (Iowa State University) and from the Maize Coop (University of Illinois, Urbana, IL). The *vp5* mutant used in this study was generated in a *Mutator* background, by Dr. Robertson, and is therefore designated *vp5-Mum*. By introduction of this mutation into another genetic background in which the *Mutator* element was inactive, a stable phenotype was obtained and revertant sectors no longer observed. *vp2* and *w3* are stable mutations. Developing endosperms were collected and frozen at -80^{0}C prior to use.

2.3.2 Amplification of the maize *Pds* gene

The CLUSTAL program (Higgins and Sharp, 1988) from PC/Gene software (Intelligenetics Inc., Mountain View, CA) was used to align *Pds* sequences from tomato (GenBank X59948; S36691), soybean (GenBank M64704), pepper (GenBank X68058) and arabidopsis (GenBank L16237). A region of high homology, corresponding to nt. 1905-2077 of tomato *Pds*, was shared by all available dicot *Pds* genes. Degenerate oligonucleotide primers, oligo 94, 5'-CCTGATGAAATC(T)TCG(A,T)GCG(A,T)GAC (T)CA-3' and oligo 95, 5'-ACAGCA(G)CCTTCCATG(T)GAAGCC(T)AA)-3', were used to amplify the corresponding region of maize *Pds* from maize B73 genomic DNA as follows: DNA, 0.1ug, in 20 ul PCR buffer (20 mM Tris-HCl, pH 8.4; 50 mM KCl; 1 mM $MgCl_2$; 160 µM each dNTP; 0.4 µM each primer; 1 µg/µl bovine serum albumin (BSA); 0.1 U/µl Taq polymerase (Gibco-BRL, Gaithersburg, MD) was incubated for 1 cycle at 94°C (3 min), followed by 40 cycles at 94°C (30 sec), 48°C (30 sec), 72°C (30 sec) and one cycle of 72°C (10 min).

2.3.3 Isolation of maize *Pds* cDNA clones

One to two million clones of a Lambda gt11 cDNA library (Fontes et al., 1991), prepared from RNA extracted from maize endosperm dissected at 14 days after pollination (DAP), were screened with the PCR-amplified maize *Pds* fragment (Sambrook et al., 1989). Seven positive clones were isolated, phage DNA extracted according to Sambrook *et al.* (Sambrook et al., 1989), and inserts amplified using primers, oligo 96, 5'-AGGCACAT GGCTGAATATCG-3' and oligo 97, 5'-CGGCAGTACAATGGATTTCC-3'. Lambda

DNA, 0.1 µg, in 20 µl PCR buffer (20 mM Tris, pH 8.2; 10 mM KCl; 2.5 mM $MgCl_2$; 1 mM $(NH_4)_2SO_4$; 200 µM each dNTP; 1 µM each primer; 1 µg/µl BSA; 0.1% Triton X-100; 0.025 U/µl *Pfu* DNA polymerase (Stratagene) was incubated for one cycle at 94°C (2.5 min), then 40 cycles at 94°C (30 sec), 55°C (30 sec), 72°C (2 min)] and one cycle of 72°C (10 min).

2.3.4 DNA sequence analysis

The *fmol*™ DNA Sequencing System (Promega, Madison, WI) was used for initial sequencing of phage DNA inserts using primers described for PCR. The plasmid deletion series prepared for the complementation analysis was used for making single-stranded templates for sequence analysis of the entire gene. Complete sequencing of maize *Pds* was carried out using the Sequenase™ Version 2.0 DNA Sequencing Kit (United States Biochemical, Cleveland, OH). Sequence analysis and homology comparisons were carried out using PCGene software (Intelligenetics). Alignments were carried out using the program CLUSTAL. *Pds* sequences used for comparisons shown in Figure 2-3 and 4 are as follows: maize, this Chapter (GenBank #U37285); for arabidopsis, pepper, soybean, tomato, see GenBank numbers listed above; *Synechocystis*, GenBank #X62574; *Synechococcus*, GenBank #X55289. Sequence encoding CRTI from *Erwininia herbicola*, GenBank #M87280; Sequence encoding CRTI and CRTD from *R. sphaeroides*, Genbank #X82458; Sequence encoding ZDS from anabaena, Genbank #S43324. Maize and pepper *rbcL* sequences for DNA and protein comparisons described in the discussion were GenBank #Z11973 and GenBank #U08610, respectively.

2.3.5 Subcloning, expression and functional complementation of maize *Pds*

Amplified Lambda gt11 inserts were purified by adsorption to Glass-Milk (The Geneclean II kit, BIO 101 Inc., Vista, CA) following the manufacturer's directions, treated at 37°C for 1 hour with Klenow fragment and 10 mM dNTP to create blunt ends, and then ligated to *Sma* I linearized vector, pBluescript II SK(-) (Stratagene). One clone, found to be in the sense orientation with respect to *lacz*, on the basis of sequencing and restriction mapping, was designated pMPDS3. This plasmid was purified by CsCl equilibrium density centrifugation according to Sambrook *et al.* (Sambrook et al., 1989). To create an in-frame fusion with *lacZ*, the plasmid was linearized with *Not* I and *Bst*X I, and subjected to 5'-end deletions using the Exonuclease III and Mung Bean Nuclease Deletion Kit (Stratagene, La Jolla, CA). Religated plasmids were transformed into *E. coli* JM101 containing plasmid pACCRT-EB, encoding GGPPS and PSY from *Erwinia uredovora* (Linden et al., 1991). Transformants containing both the deletion derivative of pMPDS3 and pACCRT-EB were selected by resistance to ampicillin and chloramphenicol and then grown in liquid culture (LB medium) with appropriate antibiotics. Expression of the LacZ-PDS fusion proteins was induced by addition of isopropylthio-β-D-galactoside (IPTG) (1 mM final concentration) during log phase. After growth to stationary phase, pigments were extracted and analyzed by HPLC.

2.3.6 Pigment extraction and HPLC analysis

Fifteen ml stationary phase cultures of *E. coli* cells were pelleted and resuspended in 20 ml methanol and carotenoids extracted according to Sandmann (Sandmann, 1993), except that 15 ml petroleum ether was used in place of diethyl ether. Pigments were separated by reverse phase HPLC, using a 25 cm x 4.6 mm Spherisorb ODS-1 5u C18 column (Phenomenex, Torrance, CA), and a solvent of acetonitrile/methanol/isopropanol (85:10:5) with flow rate of 1 ml/min using a Series 410 BIO LC Pump (Perkin Elmer, Norwalk, CT). Peaks were detected using an LC-480 Auto Scan photodiode array detector (Perkin Elmer). Alternatively, a Waters HPLC system with 600 controller and pump, a 996 photodiode array detector, and WISP 717 autosampler were used. Peaks were identified on the basis of co-migration and shared spectrophotometric profiles with known standards.

2.3.7 RNA Extraction & RT-PCR

Total RNA of maize was extracted from endosperms collected at various DAP and from leaves of young plants (2-3 leaf stage) (Logemann et al., 1987). RNA pellets were resuspended in DEPC (diethyl pyrocarbonate)-treated water, centrifuged 5 min at 14000 rpm in an Eppendorf centrifuge, and the supernatants collected. RNA concentration was estimated spectrophotometrically and total RNA (1ug) used as template for cDNA synthesized with the SuperScript™ Preamplification System (Gibco-BRL) for First Strand cDNA Synthesis. One fourth (5 µl) of product, approximately 1-4 ng cDNA, was used for PCR in a final volume of 25 ul. The amount of total RNA used was first tested to ensure linearity of response in the RT-PCR reaction (data not shown). Primers used for amplification were: *Pds*, 5'-GGAACTGTGAAACACTTCGC-3' (oligo 110) and 5'-GAAACCTTCGATAGGTGACC-3' (oligo 111); *Sh1*, 5'-ATCCCTGAGAAAGGC

AGAGG-3' (oligo 141) and 5'-AGTGACTCCCAACTTGTGCG-3' (oligo 142, GenBank X02382). The conditions for PCR were: 20 mM Tris-HCl, pH 8.4; 50 mM KCl; 2.5 mM $MgCl_2$; 100 μg/μl BSA; 2 mM DTT (dithiothreitol); 0.1 mM each dNTP; 0.4 μM each primer, 0.1 U/μl Taq DNA polymerase (BRL). The protocol for PCR was: one cycle of 94°C (2 min), followed by 40 cycles of 94°C (30 sec), 52°C (30 sec), 72°C (30 sec) and one cycle of 72°C (10 min). 10 μl of each PCR reaction was analyzed by electrophoresis on 1.8% agarose gels in 0.5 x TBE. The sizes of the PCR products for *Sh* and for *Pds* were 673 and 528 bp, respectively.

2.4 Results

2.4.1 Isolation of a maize *Pds* cDNA

Since dicot *Pds* genes hybridized poorly to maize sequences, I amplified maize *Pds* from genomic DNA using degenerate oligonucleotide primers designed by alignment of all available dicot *Pds* sequences. The maize PCR product was sequenced to verify homology to other *Pds* genes, and this PCR fragment was then used to screen a maize cDNA library of 1-2 million clones. Seven clones were obtained. Based on preliminary sequence analysis and alignment with the dicot *Pds* sequences, a 2.0 kb clone was chosen for further characterization.

2.4.2 Sequence analysis of maize *Pds* and comparison with other dicot *Pds* genes

The maize *Pds* cDNA was sequenced as shown in Figure 2. Based on the deduced amino acid sequence, maize PDS, including its putative transit peptide, was found to be 571 amino acid residues with a mass of 64.1 kD. The sequence determined here, is almost identical to a sequence of a *Pds* cDNA from another maize line, Funk F, except at 14 nucleotide positions, only four of which resulted in a change of amino acid sequence (residues 61-63, 68, and 555) (Hable and Oishi, 1995). However, at amino acid position 61-63, the maize *Pds* protein sequence reported here is identical to the dicot sequences shown in Figure 2-3A, whereas the Funk F *Pds* protein sequence is different due to a shift in reading frame.

Figure 2-3A shows the comparison of the N-terminal sequence of available dicot and cyanobacterial PDS proteins. Based on this comparison, I estimate that the maize PDS transit peptide is approximately 96 residues or 10.6 kD and therefore the plastid-localized PDS should be about 53.5 kD. The highest homology found between the dicot and monocot putative transit sequences corresponds to residues 59-96 of maize PDS. The comparison of the proposed dinucleotide binding domain, shared by carotene desaturases (PDS, CRTI, CRTD, ZDS) found in phylogenetically distant carotenogenic organisms, is shown in Figure 2-3B. Figure 2-3C shows the region of high homology at the C-terminus, which was used for design of the degenerate oligonucleotide primers used initially to amplify the maize *Pds* gene. Overall nucleotide homology between the maize and other dicot *Pds* genes ranges from 70.5-72%, whereas the amino acid homology based on identical or similar residues is about 77% and 82.8-84.2%, respectively. An alignment of all available PDS protein sequences was carried out and the results are shown in Figure 2-

4.

2.4.3 Functional analysis of maize PDS by heterologous complementation in *E. coli.*

To test the function of the maize *Pds* gene product, I subcloned the 2.0 kb insert into pBluescript II SK- and designated the clone pMPDS3. By creating progressive 5'-end deletions, I obtained several plasmids in which the insert was cloned in-frame with the *lacZ* gene. One of these, pMPDSd3-33, was introduced into *E. coli* cells carrying the plasmid pACCRT-EB encoding GGPPS (geranylgeranyl pyrophosphate synthase) and PSY from the nonphotosynthetic bacterium, *Erwinia uredovora* (Linden et al., 1991). The enzymes GGPPS and PSY will together catalyze the synthesis of phytoene in *E. coli.* Therefore, if maize PDS is a two-step desaturase, as is the case for the dicot PDS enzyme, transformants expressing both the *Erwinia* genes and maize *Pds* will accumulate ζ-carotene (refer to Figure 2-1). Phytoene and ζ-carotene are distinguished on the basis of retention times on reverse phase HPLC (High Pressure Liquid Chromatography) as well as by their unique spectrophotometric profiles. As shown in Figure 2-5A, *E. coli* cells containing only the genes encoding GGPPS and PSY, accumulated phytoene, detected at 285 nm at about 19 min. Whereas, in cells carrying genes for GGPPS, PSY and maize PDS, phytoene was converted to a mixture of ζ-carotene isomers, which were detected at 400 nm at about 14-15 min, as shown in Figure 2-5B. These results indicate that the cloned maize *Pds* encodes a two-step desaturase which catalyzes the desaturation of phytoene to ζ-carotene.

2.4.4 RFLP mapping of maize *Pds* to chromosome

As shown in Figure 2-1, mutant alleles of at least three genetic loci cause phytoene accumulation in both endosperms and green tissues, and therefore one of these loci might encode PDS. To identify the correct maize locus, RFLP analysis of a recombinant inbred family produced from a cross between T232 and CM37 was used (Burr et al., 1988). On the basis of strong hybridization to one fragment, maize *Pds* was mapped to chromosome 1S, near *vp5* (data not shown). Additional weak hybridization signals mapped to three other loci; 4L (1 map unit from *o2*); 2L (1 map unit from bnl 17.25); 1L (1 map unit from dup103).

2.4.5 RT-PCR analysis of *Pds* in phytoene-accumulating mutant endosperms

RT-PCR was used to examine *Pds* transcript levels in mutant endosperms accumulating phytoene. The low transcript abundance required the use of this sensitive technique over conventional Northern analysis. Endosperms of the genotypes *vp5*, *vp2*, and *w3* accumulate phytoene and were tested for *Pds* transcript accumulation in comparison to normal endosperm counterparts. As shown in Figure 2-6A, only *vp5* endosperms showed a visible difference in transcript accumulation as compared with transcripts accumulating in the normal endosperms. For normalization, I amplified *Sh1* (*Shrunken1*) sequences from the same cDNA; no differences in the amount of *Sh* amplification product were observed.

2.4.6 Determination of *Pds* transcript accumulation in leaves and developing endosperms.

RT-PCR analysis was used to assess the temporal pattern of *Pds* transcript accumulation in developing endosperms. Also the level of transcript accumulation in endosperm was compared to that in leaves. In comparison, I also examined the accumulation of the maize *Sh* transcript which has been previously studied in developing endosperm and in leaves using Northern analysis (Springer et al., 1986; Wurtzel et al., 1987). As shown in Figure 2-6B, the abundance of *Pds* transcripts in maize leaves and in endosperms were of comparable levels. Figure 2-6C shows that *Pds* was expressed in the unfertilized ear and transcript levels did not vary substantially (less than a 1.5-fold difference determined densitometrically) over the entire period of endosperm development, except at 40 DAP (days after pollination), when an almost three-fold reduction was observed. In contrast, accumulation of the *Sh* transcript appeared to be under temporal control; the *Sh* amplification product was first detected at 10 DAP, increased at 15 DAP and remained at a constant level between 15-30 DAP, at which point the level dropped, and no product was obtained from the 40 DAP endosperm. The temporal pattern of *Sh* transcript accumulation in developing endosperm detected by RT-PCR was consistent with previous results using Northern analysis (Wurtzel et al., 1987).

2.5 Discussion

A maize *Pds* cDNA clone was isolated, and by using a heterologous complementation system in *E.coli*, shown to encode a two-step desaturase, an enzyme catalyzing the desaturation of phytoene to ζ-carotene in two steps. The presence of a two-step desaturase in maize, a monocot, as well as in several dicots, suggests that all higher plants must encode such a two-step desaturase.

Recessive alleles of four unlinked loci, *vp2*, *vp5*, *w3*, and *vp9*, condition a block in the desaturation steps; the first three condition an accumulation of phytoene, the substrate of PDS. Therefore, one of these three loci might encode PDS. RFLP mapping results showed that *Pds* mapped near *vp5* on chromosome 1S. Furthermore, transcript analysis, using RT-PCR, showed that only *vp5* endosperms had lower levels of *Pds* transcripts accumulating in comparison to normal endosperms segregating on the same ear. Taken together, these results suggest that maize PDS may be encoded by the *vp5* locus. Consistent with these results is a previous genetic experiment showing *vp5* to encode a cell-autonomous product and not some diffusible regulator (Wurtzel, 1992). Since *vp2* and *w3* endosperms also accumulate phytoene, but do not affect *Pds* transcript accumulation as detected by RT-PCR (Figure 2-6A), it is unlikely that these loci encode transcriptional regulators. These other loci may encode a phylogenetically diverged *Pds*, such as found in the case of *Psy* genes of tomato (Bartley et al., 1994). However, DNA hybridization results obtained using maize *Pds* as a probe, suggest *Pds* is a single copy gene (not shown). Other weak hybridization signals did not map to loci associated with blocks in the desaturation steps. Therefore, *vp2* and *w3* are not as likely to be structural genes and might encode ancillary functions such as pigment binding proteins or oxidoreductases (Mayer et al., 1992). The *vp9* gene might encode or regulate expression of ZDS, since recessive alleles confer accumulation of ζ-carotene (see Figure 2-1).

Like RbcL (Ribulose bisphosphate carboxylase), which has been widely used for plant evolutionary studies, PDS shows high homology in comparing amino acid sequences of dicots with that of maize, a monocot. For example, amino acid identity and similarity for maize and pepper PDS proteins is 77.2 and 84.2%, respectively. In contrast, RbcL identity and similarity are 90.6 and 95%, respectively. The phylogenetic tree (Figure 2-4) produced by alignment of cyanobacterial, monocot, and dicot PDS amino acid sequences is consistent with current hypotheses of plant evolutionary relationships. However, unlike the chloroplast encoded *rbcL*, which is highly conserved at the nucleotide level between monocots and dicots, the nuclear encoded *Pds* nucleotide sequence is more variable. Nucleotide homology between maize and pepper *Pds* genes is 72%, whereas for *rbcL*, the homology is 85.4%.

The *rbcL* gene has been an important tool for plant evolutionary studies. However, its high degree of conservation limits its utility for evolutionary studies concerning lower ranked taxonomic groupings. In contrast, *Pds* is an essential, nuclear encoded gene with greater variability, suggesting that it holds great potential for studies at lower taxonomic ranks than shown in Figure 2-4.

A major question regarding control of carotenoid biosynthesis is whether the pathway is differentially regulated in various tissues, i.e. in different plastid types. During endosperm development, plastids triple in number (McCullough et al., 1989) and total colored carotenoids increase dramatically (Yu and Wurtzel, unpublished) during the period of 10-20 days after pollination. Using RT-PCR, I studied the expression of *Pds* transcripts in developing maize endosperm during the period of carotenoid accumulation. No marked change in *Pds* transcript level was found between 5-35 DAP, as compared with the temporally regulated expression of *Sh*. At 40 DAP, there was a reduction in *Pds* transcripts. However, this stage is late in endosperm development and well past the greatest period of carotenoid accumulation. This constitutive expression of *Pds* during the period of carotenoid accumulation in developing maize endosperm is in contrast to the temporal control of *Psy* and *Pds* transcript accumulation in developing tomato fruit; during development of tomato chromoplasts from chloroplasts, carotenoid accumulation is accompanied by a 25-fold increase in *Psy* transcripts and a 3-10 fold increase in *Pds* transcripts (Giuliano et al., 1993; Pecker et al., 1992). This difference is not unexpected, since carotenoid-containing plastids of endosperm (amyloplasts) and fruit (chromoplasts) are the products of different developmental processes (Kirk and Tilney-Bassett, 1978).

It is not surprising that maize endosperm carotenoid accumulation is not regulated by specific induction of *Pds* transcript accumulation. The induction of carotenoid accumulation in the endosperm may not necessarily be regulated at the level of transcript accumulation. Alternatively, the endosperm pathway may be regulated by transcriptional control of *Psy*. Furthermore, the endosperm pathway may not be regulated by controlling expression of the enzymes within the pathway, but alternatively by controlling the flow of substrates to the pathway. Such upstream control (above PSY, the first enzyme specific to carotenogenesis) has been previously documented. Albrecht and Sandmann (Albrecht and Sandmann, 1994) demonstrated by *in vitro* labeling experiments that the phytochrome mediated accumulation of carotenoids during the course of conversion of an etioplast to a chloroplast is regulated upstream of the pathway, via activation of IPP (isopentenyl pyrophosphate) isomerase. Another example of upstream regulation of carotenogenesis, occurs in developing pepper fruits; the abundance of transcripts encoding GGPP synthase show a concomitant increase, followed by an increase in enzyme activity, that is associated with carotenoid accumulation (Kuntz et al., 1992). Both IPP and GGPP are precursors to a variety of terpenoid pathways (Chappell, 1995). Therefore, evidence is mounting that the pathway is not only regulated with respect to tissue-specificity, but is controlled both within the pathway as well as upstream of the pathway.

The genetically identified mutant alleles affecting carotenoid synthesis in maize endosperm, but not leaves, will be useful for determining how the carotenoid biosynthetic pathway is regulated in different plastid/tissue types, where the role and localization of carotenoids varies. Furthermore, an understanding of the molecular regulation of carotenoid biosynthesis in endosperm is of great value for engineering the pathway in endosperms of other cereal crops that are otherwise poor nutritional sources of carotenoids. The stable expression of *Pds* transcripts in developing endosperm may also serve as a useful internal experimental control for future studies of endosperm gene expression, including other genes involved in carotenogenesis.

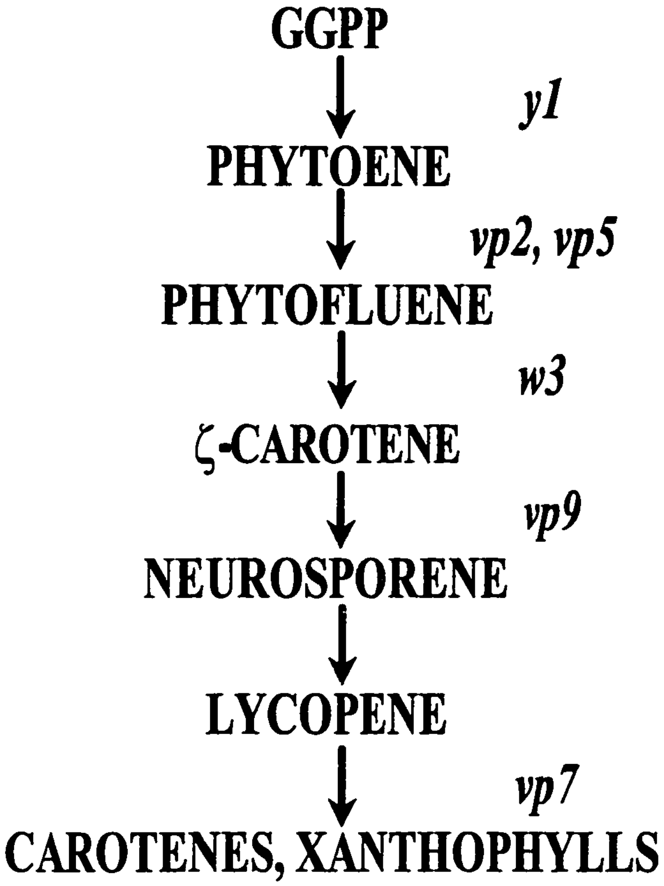


Figure 2-1. Lethal recessive mutations blocking the carotenoid biosynthetic pathway of maize. *w3* results in accumulation of both phytoene and phytofluene. Many, but not all, *yl* mutations affect only endosperm and are not lethal to the plant.

1 ATGGACACTGGCTGCCTGTCATCTATGAATATTACTGGAGCTAGCCAG
1 M D T G C L S S M N I T G A S Q
49 ACAAGATCTTTTGCGGGGCAACTTCCTCCTCAGAGATGTTTTGCGAGT
17 T R S F A G Q L P P Q R C F A S
97 AGTCACTATACAAGCTTTGCCGTGAAAAAACTTGTCTCAAGGAATAAA
33 S H Y T S F A V K K L V S R N K
145 GGAAGGAGATCACACCGTAGACATCCTGCCTTGCAGGTTGTCTGCAAG
49 G R R S H R R H P A L Q V V C K
193 GATTTTCCAAGACCTCCACTAGAAAGCACAATAAACTATTTGGAAGCT
65 D F P R P P L E S T I N Y L E A
241 GGACAGCTCTCTTCATTTTTTAGAAACAGCGAACGCCCCAGTAAGCCG
81 G Q L S S F F R N S E R P S K P
289 TTGCAGGTCGTGGTTGCTGGTGCAGGATTGGCTGGTCTATCAACAGCG
97 L Q V V V A **G A G L A G L S T A**
337 AAGTATCTGGCAGATGCTGGCCATAAACCCATATTGCTTGAGGCAAGA
113 **K Y L A D A G H K P I L L E A R**
385 GATGTTTTGGGTGGAAAGGTAGCTGCTTGGAAGGATGAAGATGGAGAT
129 **D V L G G** K V A A W K D E D G D
433 TGGTACGAGACTGGGCTTCATATATTTTTTGGAGCTTATCCCAACATA
145 W Y E T G L H I F F G A Y P N I
481 CAGAATCTGTTTGGCGAGCTTAGGATTGAGGATCGTTTGCAGTGGAAA
161 Q N L F G E L R I E D R L Q W K
529 GAACACTCTATGATATTCGCCATGCCAAACAAGCCAGGAGAATTCAGC
177 E H S M I F A M P N K P G E F S
577 CGGTTCGATTTCCCAGAAACTTTGCCAGCACCTATAAATGGGATATGG
193 R F D F P E T L P A P I N G I W
625 GCCATATTGAGAAACAATGAAATGCTTACTTGGCCGGAGAAGGTGAAG
209 A I L R N N E M L T W P E K V K
673 TTTGCAATCGGACTTCTGCCAGCAATGGTTGGTGGTCAACCTTATGTT
225 F A I G L L P A M V G G Q P Y V
721 GAAGCTCAAGATGGCTTAACCGTTTCAGAATGGATGAAAAAGCAGGGT
241 E A Q D G L T V S E W M K K Q G
769 GTTCCTGATCGGGTGAACGATGAGGTTTTTATTGCAATGTCCAAGGCA
257 V P D R V N D E V F I A M S K A
817 CTCAATTTCATAAATCCTGATGAGCTATCTATGCAGTGCATTTTGATT
273 L N F I N P D E L S M Q C I L I
865 GCTTTGAACCGATTTCTTCAGGAGAAGCATGGTTCTAAAATGGCATTC
289 A L N R F L Q E K H G S K M A F
913 TTGGATGGTAATCCGCCTGAAAGGCTATGCATGCCTATTGTTGATCAC
305 L D G N P P E R L C M P I V D H
961 ATTCGGTCTAGGGGTGGAGAGGTCCGCCTGAATTCTCGTATTAAAAAG
321 I R S R G G E V R L N S R I K K
1009 ATAGAGCTGAATCCTGAT**GGAACTGTAAAACACTTCG**CACTTAGTGAT
337 I E L N P D G T V K H F A L S D
1057 GGAACTCAAATAACTGGAGATGCTTATGTTTGTGCAACACCAGTCGAT
353 G T Q I T G D A Y V C A T P V D
1105 ATCTTCAAGCTTCTTGTACCTCAAGAGTGGAGTGAAATTACTTATTTC
369 I F K L L V P Q E W S E I T Y F
1153 AAGAAACTGGAGAAGTTGGTGGGAGTTCCTGTTATCAATGTTCATATA
385 K K L E K L V G V P V I N V H I
1201 TGGTTTGACAGAAAACTGAACAACACATATGACCACCTTCTTTTCAGC
401 W F D R K L N N T Y D H L L F S
1249 AGGAGTTCACTTTTAAGTGTCTATGCAGACATGTCAGTAACCTGCAAG
417 R S S L L S V Y A D M S V T C K
1297 GAATACTATGACCCAAACCGTTCAATGCTGGAGTTGGTCTTTGCTCCT
433 E Y Y D P N R S M L E L V F A P
1345 GCAGACGAATGGATTGGTCGAAGTGACACTGAAATCATCGATGCAACT
449 A D E W I G R S D T E I I D A T
1393 ATGGAAGAGCTAGCCAAGTTATTT**CCTGATGAAATTGCTGCTGATCAG**
465 M E E L A K L F P D E I A A D Q
1441 AGTAAAGCAAAGATTCTTAAGTATCATATTGTGAAGACACCGAGATCG
481 S K A K I L K Y H I V K T P R S
1489 GTTTACAAAACTGTCCCAAACTGTGAGCCTTGCCGGCCTCTCCAAA**GG**
497 V Y K T V P N C E P C R P L Q R
1537 **TCACCTATCGAAGGTTTC**TATCTAGCTGGTGATTACACAAAGCAGAAA
513 S P I E G F Y L A G D Y T K Q K
1585 TAC**CTGGCTTCTATGGAAGGTGCAGT**CCTATCCGGGAAGCTTTGTGCC
529 Y L A S M E G A V L S G K L C A
1633 CAGTCCATAGTGCAGGATTATAGCAGGCTCGCACTCAGGAGCCAGAAA
545 Q S I V Q D Y S R L A L R S Q K
1681 AGCCTACAATCAGGAGAAGTTCCCGTCCCATCTTAGTTGTAGTTGGCT
561 S L Q S G E V P V P S *
1729 TTAGCTATCGTCATCCCCACTGGGTGCTATCTTATCTCCTATTTCAAT
1777 GGGAACCCACCCAATGGTCATGTTGGAGACAACACCTGTTATGGTCCT
1825 TTGACCATCTCGTGGTGACTGTAGTTGATGTCATATTCGGATATATAT
1873 GTAAAAGGACCTGCATAGCAATTGTTAGACCTTGGAAAAAAA

Figure 2-2. Nucleotide and amino acid sequence of the maize *Pds* cDNA of pMPDS3. Deduced amino acid sequence is shown as single letters below the nucleotide sequence. Bold letters indicate the putative dinucleotide binding domain in the protein sequence and the oligonucleotide primers used for PCR or RT-PCR in the nucleotide sequence. The sequence as shown represents the entire sequence available from the cloned cDNA (U37285).

A

```
                    10            20                30         40
ZM      M---DTGCLSSMNITGASQ----TRSFAGQLPPQR--------CFASSHYTSFAVKKLVS
AT      MVVFGNVSAANL------PYQNG------FLEA---------LSSGGCELMGHSFRVPTS
CA      MPQIGLVSAVNLRVQGNSAYLWSSRSS-LGTDSQDGCSQRNSLCFGGSDSMSHRLKIRNP
GM      MAACGYISAANF------NYLVGARNISKFASSDATI----SFSFGGSDSMGLTLRPAP-
LE      MPQIGLVSAVNLRVQGSSAYLWSSRSSSLGTESRDGCLQRNSLCFAGSESMGHKLKIRTP
        *   .  ..                                   .....  ..   .

            50        60         70         80        90        100
ZM      RNKGRRSHRRH--PALQVVCKDFPRPPLESTINYLEAGQLSSFFRNSERPSKPLQVVVA
AT      QALKTRTRRRSTAGPLQVVCVDIPRPELENTVNFLEAASLSASFRSAPRPAKPLKVVIA
CA      HSITRRLAK--DFRPLKVVCIDYPRPELDNTVNYLEAAFLSSSFRSSPRPTKPLEIVIA
GM      ----IRAPKRNHFSPLRVVCVDYPRPELENTVNFVEAAYLSSTFRASPRPLKPLNIVIA
LE      HATTRRLVK--DLGPLKVVCIDYPRPELDNTVNYLEAAFLSSTFRASPRPTKPLEIVIA
SS                                                           MRVVIA
SY                                                           MRVAIA

             *  .     .*.*** * ***.*..*.*..**. **. ** ..** *** ...*
```

B

```
          103                           133
ZM        GAGLAGLSTAKYLADAGHKPILLEARDVLGG
AT        GAGLAGLSTAKYLADAGHKPLLLEARDVLGG
CA        GAGLGGLSTAKYLADAGHKPILLEARDVLGG
GM        GAGLAGLSTAKYLADAGHKPILLEARDVLGG
LE        GAGLGGLSTAKYLADAGHKPILLEARDVLGG
SS        GAGLAGLACAKYLADAGFTPVVLERRDVLGG
SY        GAGLAGLSCAKYLADAGHTPIVYERRDVLGG
-----------------------------------------
EH(CRTI)  GAGFGGLALAIRLQAAGIPVLLLEQRDKPGG
RS(CRTI)  GSGLGGLAAAMRLGAKGWRVTVIDKLDVPGG
RS(CRTD)  GAGMGGLASAIRLARAGCEVTLLEAREAPGG
AN(ZDS)   GAGPGGLATAIRLAGLGYQVEIFEAAERVGG
          *.* .**  * .*   *  .   .  .  **
```

C

```
          521                  542
ZM        AGDYTKQKYLASMEGAVLSGKL
AT        AGDYTKQKYLASMEGAVLSGKF
CA        AGDYTKQKYLASMEGAVLSGKL
GM        AGDYTKQKYLASMEGAVLSGKL
LE        AGDYTKQKYLASMEGAVLSGKL
SS        AGDFTMQKYLGSMEGAVLSGKQ
SY        TGDYTMQRYLASMEGAVLSGKL
          .**.*.*.**.**********
```

Figure 2-3. PDS amino acid sequence alignments. A, N-terminus and transit sequence; B, Putative dinucleotide-binding domain: sequences above the line are from *Pds*-type desaturase and sequences under the line are from *CrtI*-type or *Zds*; C, Highly conserved region used for design of degenerate oligonucleotide primers used for amplification of the maize gene. Asterisks indicate identical residues and dots indicate similar residues. ZM, maize; AT, arabidopsis; CA, pepper; GM, soybean; LE, tomato; SS, *Synechocystis*, SY, *Synechococcus*; EH, *Erwinia herbicola*; RS, *Rhodobacter sphaeroides*; AN, *Anabaena*.

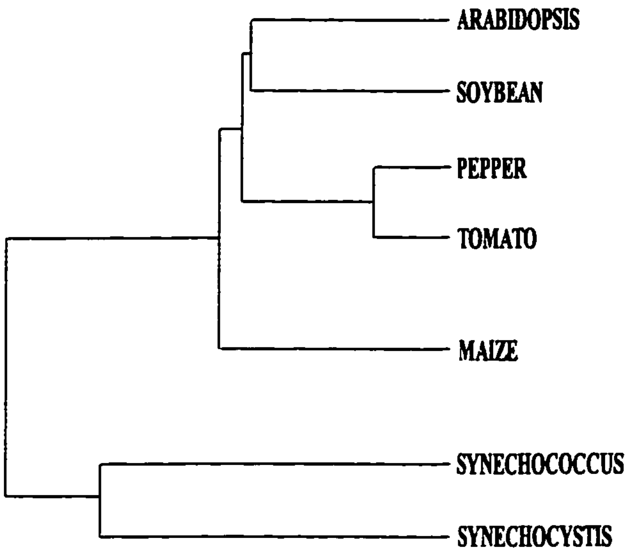


Figure 2-4. Phylogenetic tree based on alignment of PDS amino acid sequences in monocots, dicots, and cyanobacteria. The program used to generate this dendrogram is CLUSTAL developed by Higgins and Sharp (Higgins and Sharp, 1988).

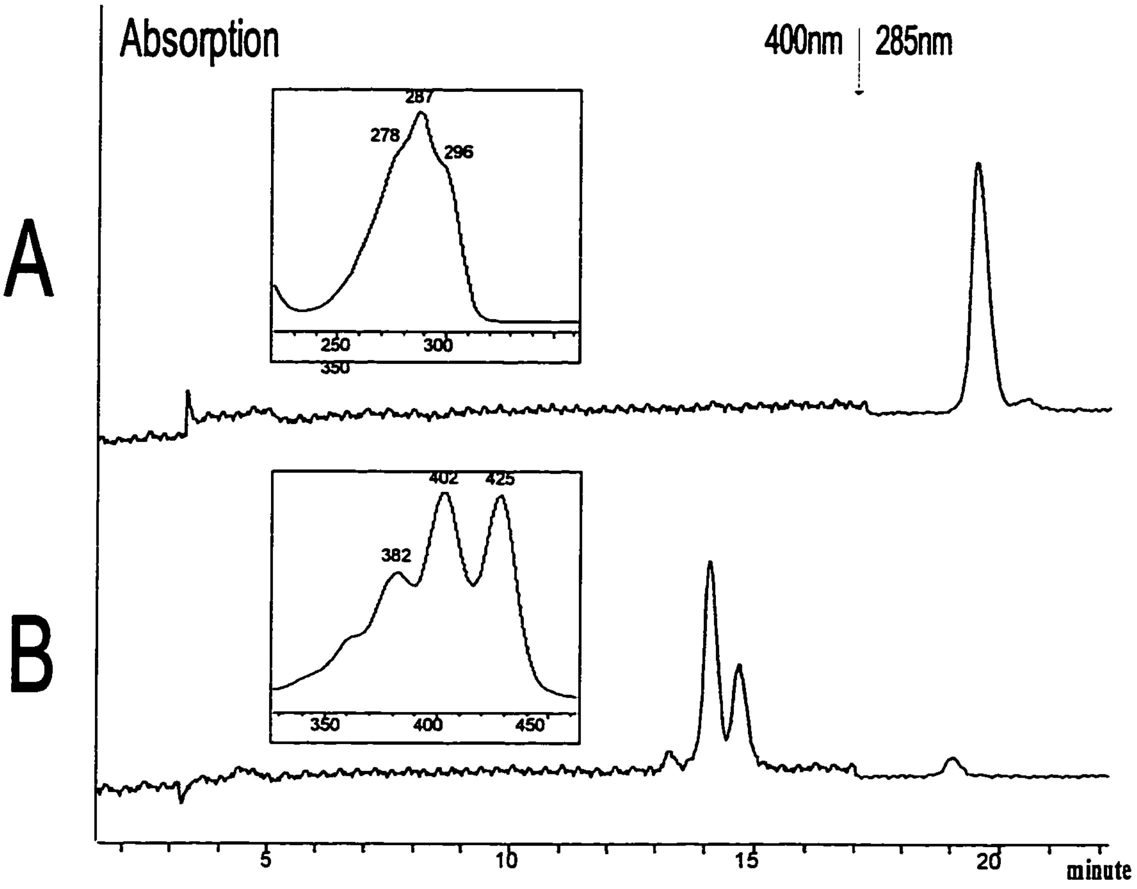


Figure 2-5. HPLC analysis of products in *E.coli* heterologous complementation system. Pigments were extracted from *E.coli* cells transformed with either (A) pACCRT-EB only; (B) pACCRT-EB and pMPDSd3-33. The insets show spectrophotometric profiles of the major peaks. Phytoene has a retention time of 19 minutes and ζ-carotene, 14-15 min.

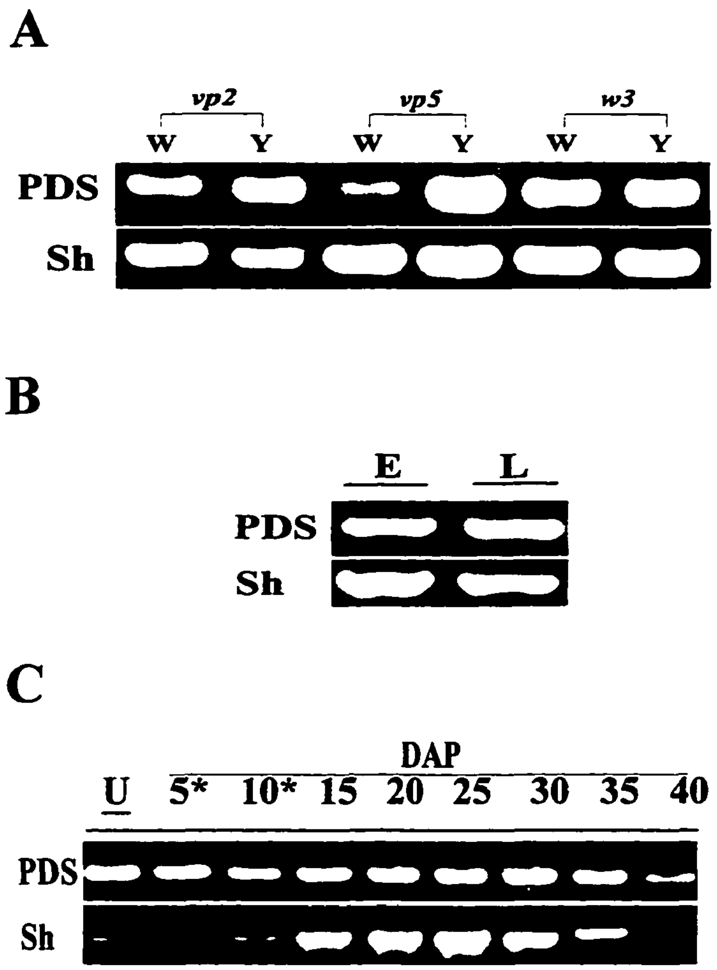


Figure 2-6. RT-PCR analysis of *Pds* and *Sh* transcripts. RNA for RT-PCR was extracted from A, white carotenoid mutant (W) and normal yellow endosperms (Y) segregating on maize ears harvested at 20 DAP; B, endosperms harvested at 20 DAP and young leaves; C, developing endosperms harvested at varying DAP (as shown by corresponding numbers) and unfertilized ear (U). * indicates that some maternal tissue may also be present in the endosperm sample. Identity of amplification products are shown at left of bands in each panel.

Chapter 3: Comparison of *Pds* and *Psy* expression and further examination of *Pds* clones.

3.1 Introduction

In Chapter 2, RT-PCR analysis of *Pds* expression at the RNA level in the endosperm of an Argentina high carotenoid line of maize is reported. Compared to *Shrunken*, whose RNA level increases between 10 and 20 DAP (days after pollination), *Pds1* expression at the RNA level is constant, even when the carotenoid level increases. In this chapter, similar experiments were performed in two other maize inbred lines and similar results were obtained. In order to analyze which enzyme of the pathway is regulated in maize endosperm during carotenoid biosynthesis, *Psy* expression at the RNA level was analyzed by RT-PCR and compared to *Pds1* expression. Data of this study are presented.

Along with pMPDS3 (Chapter 2), six other *Pds* clones were subcloned from the λgt11 clones obtained by cDNA library screening. All seven clones were analyzed; some inserts were close to or longer than 2 kb while others were less than 1 kb. Thus, besides pMPDS3, which was characterized in Chapter 2, three other longer clones were further analyzed and the results are reported in this chapter.

Out of these three longer clones, pMPDS7 was about 3 kb, which was the longest of the seven *Pds* clones and longer than pMPDS3. Since pMPDS3 did not contain the 5' untranslated region, pMPDS7 was a proper candidate for analyzing the 5' untranslated region of *Pds*. Consequently, a genomic library of maize B73 line was screened with the 5' end portion of pMPDS7 insert to isolate the *Pds* promoter and preliminary results are described. The other two (pMPDS2 and pMPDS10) were identical and by Southern analysis proved to be different from pMPDS3, though they have a similar size of inserts to that of pMPDS3. This difference led to the discovery of a group of novel receptor-like protein kinases of higher plants (see Chapter 4 for details).

3.2 Materials and Methods

3.2.1 Plant materials

Three inbred lines, B73 (96004sib), A632 (96010sib) and an Argentina high carotenoid line (96016sib), were grown in a field at Pelham Bay Park, Bronx, New York. Kernels of inbred lines were harvested at 5, 10, 15, 20, 25 and 30 days after pollination (DAP) and the endosperm was separated from the embryo. Unfertilized ovules were separated from unfertilized ears. Endosperms, embryos and unfertilized ovules were quickly frozen in liquid nitrogen. *Y1* (94047sib) and *y1* (94048sib) were grown in a field at Black Rock Forest (Cornwall, New York). Kernels were harvested at 20 DAP and endosperms were separated. Leaf tissues were obtained from seedlings (3-leaf stage) grown in the greenhouse at Lehman College, CUNY. All tissues were stored at -80°C.

3.2.2 RT-PCR analysis of *Psy*, *Pds* and *Sh* expression

Primers for *Psy* RT-PCR were designed according to the maize Y1 gene sequence (U32636) to flank three introns. The forward primer (oligo 106) is 5'-TGTGAGGAGTAT GCCAAGACG-3' (from 2309 to 2329) and the reverse primer (oligo 109) is 5'-CAGGTA CGCTCATTAACCCG-3' (from 4286 to 4267). cDNAs were synthesized from 1 µg total RNA and one quarter used in PCR reactions as described in Chapter 2. The linearity of product formation in the RT-PCR reaction was also tested as previously described, except that the following procedure for PCR was used. The conditions for PCR were: 1-4 ng cDNA in 20 mM Tris-HCl, pH 8.4; 50 mM KCl; 2.5 mM $MgCl_2$; 100 mg/µl BSA; 2 mM DTT (dithiothreitol); 0.1 mM each dNTP; 0.4 mM each primer, 0.1 U/ml Taq DNA polymerase (Gibco-BRL). The protocol for PCR was one cycle of 94°C (2 min), followed by 40 cycles of 94°C (30 sec), 52°C (30 sec), 72°C (30 sec) and one cycle of 72°C (10 min). Ten µl of each PCR reaction was analyzed by electrophoresis on 1.8% (w/v) agarose gels in 0.5 x TBE. RT-PCR analysis of *Sh* and *Pds* was performed as in Chapter 2.

3.2.3 Genomic library screening

A genomic library (Cat. # FL 1032D) from the B73 line of maize was purchased from CLONTECH Laboratories, Inc. (Palo Alto, CA) and 7×10^6 phage clones were screened with the radiolabeled 500 bp *Eco*R I fragment of the 5' end of pMPDS7. Sixty-four positive clones were isolated and phage DNA of one positive clone (phage clone 624) was isolated according to Li *et al.* (Li et al., 1996). The phage DNA of clone 624 was digested with *Sal* I and a 12.2 kb insert was separated from the vector in a low melting point agarose gel and then ligated into pBluescript II SK (-) at the *Sal* I site, after Li *et al.* (Li et al., 1996). This plasmid was designated as p624.

Plasmid DNA of p624 was isolated using the Wizard *Plus* Miniprep DNA Purification System (Promega, Madison, WI) and incubated with *Hin*d III. The digested plasmid DNA was then separated on a 1% (w/v) agarose gel and blotted onto nitrocellulose. The filter was first probed by *Pds* 5'-end, which could hybridize to the 5' end of *Pds* and its promoter. The same filter was then stripped and probed again by the receptor-like protein kinase gene. Both hybridizations were performed following Studier hybridization (Simon and Studier, 1973). A 5.8 kb *Hin*d III fragment could be probed by both probes and was ligated into pBluescript II SK (-) at the *Hin*d III site. This clone was designated as pH58.

3.2.4 DNA Sequencing

DNA sequencing was carried out using the Sequenase™ Version 2.0 DNA Sequencing Kit (United States Biochemical, Cleveland, OH) with plasmid DNA isolated by the Wizard *Plus* Miniprep DNA Purification System as template. Alternatively, automatic sequencing was used to sequence both strands of pH58 by primer walking.

3.2.5 Southern analysis

Genomic DNA of rice was digested by *Eco*R I and *Hin*d III and then separated by electrophoresis on a 0.6% agarose gel. A Southern blot was done following Wurtzel *et al.*

(Wurtzel et al., 1987) and filters were probed by with the insert of pMPDS2 or pMPDS3 respectively. Washing conditions were as described in Wurtzel *et al.* (Wurtzel et al., 1987).

3.3 Results

3.3.1 Comparison of *Psy* and *Pds* expression at the RNA level in maize endosperm

The expression of *Pds* at the RNA level was first analyzed in endosperm of the maize Argentina high carotenoid line (Chapter 2) and found to be constant during endosperm development, especially during the initial period of carotenoid-accumulation (10-20 DAP) (Yu and Wurtzel, unpublished data). The same analysis was conducted in two other maize inbred lines, B73 and A632 and similar results were obtained (Figure 3-1); and a conclusion was drawn that *Pds1* was constantly expressed at the RNA level during maize endosperm development. This conclusion suggested that the carotenoid biosynthetic pathway in maize endosperm is not regulated through *Pds* expression at the RNA level. As a control, *Sh* expression was also analyzed by RT-PCR. *Sh* transcripts accumulated in a pattern consistent with a previous study showing temporal regulation during maize endosperm development.

Data from several groups show that PSY is the regulated enzyme of carotenoid biosynthesis in other plants (Fraser et al., 1994; Giuliano et al., 1993; Karvouni et al., 1995). To test whether a similar mechanism operated in maize endosperm carotenogensis, RT-PCR was performed with total RNA isolated from a developing maize endosperm series. As shown in figure 3-1A, the *Psy* mRNA level was constant in the Argentina line, which has a higher carotenoid content than the other two inbred lines (B73 and A632), in which *Psy* mRNA is temporally regulated parallel to the carotenoid accumulation profile (data not shown). This result indicates that early expression (10 DAP) of *Psy* at the RNA level may result in high carotenoid accumulation in the endosperm. Therefore, as in other plants, PSY could be the regulated enzyme for carotenogenesis in the maize endosperm, this regulation being at the RNA level.

In other plants, *Psy* was also found to be tissue-specifically regulated (Fraser et al., 1994; Giuliano et al., 1993; Karvouni et al., 1995). A similar result is shown in Figure 3-1B in which little or no *Psy* transcripts were detected in the *y1* homozygous endosperm compared to that of embryo and leaf tissue. In contrast, *Psy* transcripts were detected in all three tissues (endosperm, embryo and leaf) of *Y1* homozygous background. As a control, *Pds* transcripts were found in these three tissues in both *YY* and *yy* backgrounds. This result suggests that maize *Psy* is regulated at the RNA level in endosperm, but differently in embryo and leaf.

3.3.2 The full length cDNA sequence of *Pds1*

To analyze the seven positive *Pds* clones isolated from the screening of the maize cDNA library (Chapter 2), *Eco*R I digestions and sequencing the inserts ends were used. The results of these analyses were summarized in Figure 3-2. Four of them, pMPDS2, pMPDS3, pMPDS7 and pMPDS10, were found to contain inserts about 2 kb in length or longer. These clones were further analyzed.

It was noticed that one of them, pMPDS7 contained the longest insert (about 3 kb). Sequencing of the 5' end of pMPDS7 insert showed that it had a longer sequence than that of pMPDS3 (Chapter 2) at the 5' end. When pMPDS7 was translated into its amino acid sequence, which was identical to that of pMPDS3, it was found to contain three stop codons upstream of the start codon of *Pds* as suggested in Chapter 2 (Figure 3-3). This result suggested that pMPDS7 comprised both the ORF (open reading frame) and the 5' untranslated region of *Pds*. Preliminary sequence of the 3' end of pMPDS7 insert revealed that it contained *Pds* cDNA and a ~700 nt cDNA encoding a TCTP (translationally controlled tumor protein)-like protein (Pay et al., 1992). Compared to the *Eco*R I restriction map of pMPDS3, this TCTP-like protein cDNA within pMPDS7 was attached to the 3' end of the *Pds* cDNA through an *Eco*R I site (Figure 3-2).

Adding the 5' untranslated region to the ORF of pMPDS3 (containing 1914 bp insert) resulted in a 2264 nt cDNA which was the insert of pMPDS7 without TCTP. Compared to the mRNA length revealed by Northern analysis with a *Pds* probe (data not shown), the pMPDS7 insert was less than 50 nt shorter and therefore was considered as a full-length cDNA of *Pds*. The 5' untranslated sequence in clone pMPDS7 was added to the *Pds* sequence (U37285) determined from clone pMPDS3, as described in Chapter 2.

3.3.3 Cloning the promoter of *Pds1*

A 5' end fragment of the clone pMPDS7 containing the untranslated region was used to screen a genomic DNA library of the maize B73 inbred line. The 12.2 kb insert of one positive lambda clone was subcloned into pBluescript II SK (-) at the *Sal* I site and this clone was designated as p624. A restriction map of p624 was generated by *Hin*d III and *Eco*R I analysis. Southern analysis of p624 revealed that a 5.8 kb *Hin*d III fragment could be probed by the 5' fragment of pMPDS7 (for a diagram, see Figure 3-4A). Surprisingly, this fragment could also be probed by a novel receptor-like protein kinase cDNA (Chapter 4).

This 5.8 kb *Hin*d III fragment was then subcloned into pBluescript II SK (-) and was designated as pH58. Its insert was sequenced using primer walking from both directions. The sequence data revealed that this fragment possessed the first exon and part of the first intron of *Pds1* (Figure 3-4B, GenBank # AF039585). However, the overlapped region between the first exon sequence of *Pds1* promoter and the *Pds1* untranslated region (5' end of pMPDS7 insert) was not identical (data not shown). The possible reason is that these two sequences were isolated from different maize lines. As shown in Figure 3-4B, a putative promoter region with the TATA box and a related Cap site for the *Pds* mRNA were identified according to the PCGene program. This fragment also contained part (1.34 kb) of transposon 10 (Tn10) and the last exon of a novel protein kinase (see Chapter 4 and Appendix 3). The Tn10 sequence starts at 1125 nt and ends at 2460 nt. The last exon of the novel protein kinase starts at 357 nt and ends at 744 nt.

3.3.4 The chimeric cDNA clone of the novel protein kinase and phytoene desaturase

Southern analysis using *Pds* cDNA clones as probes revealed that insert of pMPDS2 hybridized to unique bands that could not be probed by pMPDS3 (Figure 3-5). Since the inserts of these three clones were similar in size, at about 2 kb, it was presumed that pMPDS2 contained an insert with a different sequence from that of pMPDS3. Sequencing the ends of the inserts of pMPDS2 and pMPDS10 showed that these two clones contained the same sequences, suggesting that these two clones were identical. Therefore, one of them, pMPDS10 was fully sequenced and the sequence was compared to the insert of pMPDS3 (*Pds1*), as shown in Figure 3-6.

As predicted from Southern analysis, pMPDS10 had a different sequence from that of pMPDS3, which encodes PDS1 (Chapter 2). pMPDS10 contained the 3' end of the *Pds* cDNA but lacked the 5' end which encodes the dinucleotide binding site. As shown in Figure 3-7, pMPDS10 had a unique 5' end encoding a polypeptide in frame with the carboxyl terminal of PDS1. A GenBank search using this unique sequence showed that it had homology to sequences encoding the ATP binding site of protein kinases. Since pMPDS10 possessed a truncated *Pds1* cDNA, it was proposed that the protein encoded by the chimeric cDNA did not have PDS activity. To demonstrate this, the insert of pMPDS10 was made in frame with *lacZ* as a translational fusion and expressed in a cell line containing the pACCRT-EB plasmid as described in Chapter 2. Pigments were isolated and analyzed spectrophotometrically: no ζ-carotene could be detected (data not shown). This result strongly suggested that pMPDS10 was composed of a chimeric cDNA that encoded a PDS protein lacking activity.

3.4 Discussion

Like other biosynthetic pathways, the regulation of carotenoid biosynthesis occurs through the regulation of expression or activity of biosynthetic enzymes of the pathway. Studies in tomato reveal that, in different tissues, the regulation of carotenoid biosynthesis is different and is developmentally regulated, especially in tissues containing chromoplasts (Giuliano et al., 1993). With the maturation of the tomato fruit, *Psy* transcripts elevated 20-fold, but a lesser change in *Pds* transcripts (3-fold) was observed (Giuliano et al., 1993). Similar results were obtained from studying developing endosperm of two maize lines (B73 and A632) in which *Psy* transcript accumulation increased whereas *Pds* was constantly expressed. However, in a high carotenoid-containing line (Argentina), the steady-state *Psy* transcript was constant during the initial carotenoid accumulation period (10-20 DAP). The high expression of *Psy* at the RNA level at the early stage (10 DAP) of endosperm development may contribute to the high carotenoid content of its endosperm. This phenomenon is similar to the expression pattern of the *Psy* steady-state transcript observed during flower development in the daffodil (*Narcissus pseudonarcissus*) (Schledz et al., 1996). Whereas the carotene content of the daffodil flower increases 6-fold from green bulbs to fully developed, open flower, the *Psy* transcript level remains constant.

It is reasonable to expect that regulation of carotenoid biosynthesis occurs through modulation of *Psy* expression, since PSY is the first enzyme specific to the pathway and its product, phytoene, is the first unique precursor of the pathway. In tomato, two *Psy* genes have been cloned with two different transcripts which are differentially expressed in tissues containing chloroplast or chromoplast containing tissues (see 1.4 The regulation of carotenoid biosynthesis in higher plants.). In maize, a single transcript has been cloned, and it is transcribed from the *Y1* gene (Buckner et al., 1996). However, the *Y1* gene is differentially transcribed in leaf (chloroplast-containing tissue) and endosperm (amyloplast-containing tissue). The recessive mutation, *y1*, can only impair *Psy* expression at the RNA level in endosperm but not in leaf or embryo, etioplast-containing tissue. This result suggested that tissues containing different types of plastid may regulate *Psy* expression differently. However, in maize developing endosperm that contains amyloplasts developed from proplastids, the expression of *Psy* is similar to that in developing tomato or *Capsicum* fruit that contains chromoplasts developed from chloroplasts (Giuliano et al., 1993; Romer et al., 1993).

Sequencing pMPDS7 revealed that it contained a full-length cDNA of *Pds1*. The 5' untranslated sequence (350 nt) has been added to the published *Pds1* sequence (Figure 3-3 and GenBank # U37285) which, based on the sequence of pMPDS3, started at the putative start codon. This full length *Pds* sequence is also about 280 nt longer than another published maize *Pds* cDNA sequence, GenBank # L39266 (Hable and Oishi, 1995). Furthermore, the fragment containing the 5' untranslated region was used as a probe to isolate a genomic clone possessing a putative promoter of *Pds1* from the maize B73 line genomic DNA library. This finding will facilitate the further analysis of the regulation of *Pds1* expression not only in the endosperm but also in other maize tissues (discussed in Chapter 5).

It was found that pMPDS7 also contains a ~700 bp fragment encoding a TCTP-like protein. EcoR I digestion showed that the pMPDS7 insert had an extra *Eco*R I fragment (~ 700 bp) as compared to the *Pds* cDNA digested with *Eco*R I (data not shown). These results suggested that the cDNA encoding the TCTP-like protein was co-ligated into the λgt11 vector at the *Eco*R I site with *Pds1* cDNA. In light of the high expression of the TCTP-like protein in the cell of other organisms including plants (Pay et al., 1992), it is not surprising that its cDNA would be co-cloned with the *Pds1* cDNA into the vector.

In the process of cloning the promoter of *Pds1*, a clone (pH58) hybridized by both the *Pds1* 5' end and the 3' end of a novel protein kinase gene was obtained. This result demonstrated that these two genes are closely linked, which was suggested by RFLP mapping (see Chapter 4). It has been reported that 50% of the maize genome is composed of repeat sequences (SanMiguel et al., 1996). When the insert of p58H was used to probe a maize genomic DNA filter, only a single band was observed, which suggests that this region contains no repeat sequence (Figure 3-8). Searching GenBank with sequence from this region showed no repeat sequences from plants but did reveal part of a bacterial transposon (Tn10) inserted between these two genes (Appendix 3). It is reasonable that this inserted transposon was an artificial product of genomic DNA library construction since the transposon has only been found in *E. coli*. If the transposon sequence is eliminated, the distance between the last nucleotide upstream of the polyA tail of the novel protein kinase gene and the first nucleotide of *Pds1* TATA box is only about 3.6 kb.

Sequencing the ends of the inserts of pMPDS2 and pMPDS10 showed that these two clones were identical. Fully sequencing of pMPDS10 showed that, like pMPDS7, it contained a partial *Pds1* cDNA and a partial sequence of another gene. This new gene could encode a protein kinase because the unique region at the 5' end of the pMPDS10 insert encoded a protein kinase ATP binding site. However, unlike pMPDS7, which contained a 2.2 kb Pds and 0.7 kb TCTP cDNA joined at an *Eco*R I site, pMPDS10 contained a chimeric cDNA, whose components were not joined at an *Eco*R I site. This result implies that this chimeric cDNA was not an artificial product of cDNA library construction but the product of a chimeric transcript present in the endosperm of the maize line used to isolate mRNA. Two possibilities regarding the production of this chimeric mRNA can be suggested. First, this chimeric mRNA could be a product of RNA splicing. Since *Pds1* and the novel protein kinase gene were less than 4 kb apart and the novel protein kinase gene was upstream of *Pds1* and transcribed in the same direction, a chimeric hnRNA composed of both genes might be produced. This hnRNA could then be spliced into two mature mRNA species encoding PDS and a novel protein kinase or alternatively, a chimeric mRNA of these two genes could be generated. Second, this chimeric mRNA could be the product of an impaired chimeric gene of the novel kinase and *Pds1*. The maize line used to isolate endosperm mRNA for cDNA library construction was a heterozygous mutant of *vp5*. Results from Chapter 2 strongly suggested that *vp5* encoded PDS1, the product of *Pds1*. Therefore, it is possible that the recessive mutant of *Pds1* might contain a short deletion between these two genes and a chimeric gene including both *Pds1* and the novel protein kinase gene could be generated. The hnRNA of this chimeric gene would then be processed into the chimeric mRNA including the 5' end of the novel protein kinase gene and the 3' end of *Pds1*.

A study by Waegemann and Soll proposed that phosphorylation of the transit sequence of chloroplast precursor proteins might play a role in protein-specific sorting and productive intracellular translocation in plants (Waegemann and Soll, 1996). If the chimeric cDNA indeed encodes part of a protein kinase, and if the protein kinase gene is closely linked to and differentially spliced with *Pds1* whose products need to be translocated into the chloroplast, I thought it was important to clone the cDNA encoding this novel protein kinase to test whether the kinase could phosphorylate the PDS transit sequence. Therefore, the unique region of the chimeric cDNA was then used to design a primer for 5' and 3' RACE (rapid amplification of cDNA ends) and sequence analysis of the RACE product led to the discovery of a novel receptor-like protein kinase. The characterization of this novel receptor-like protein kinase is described in the next chapter.

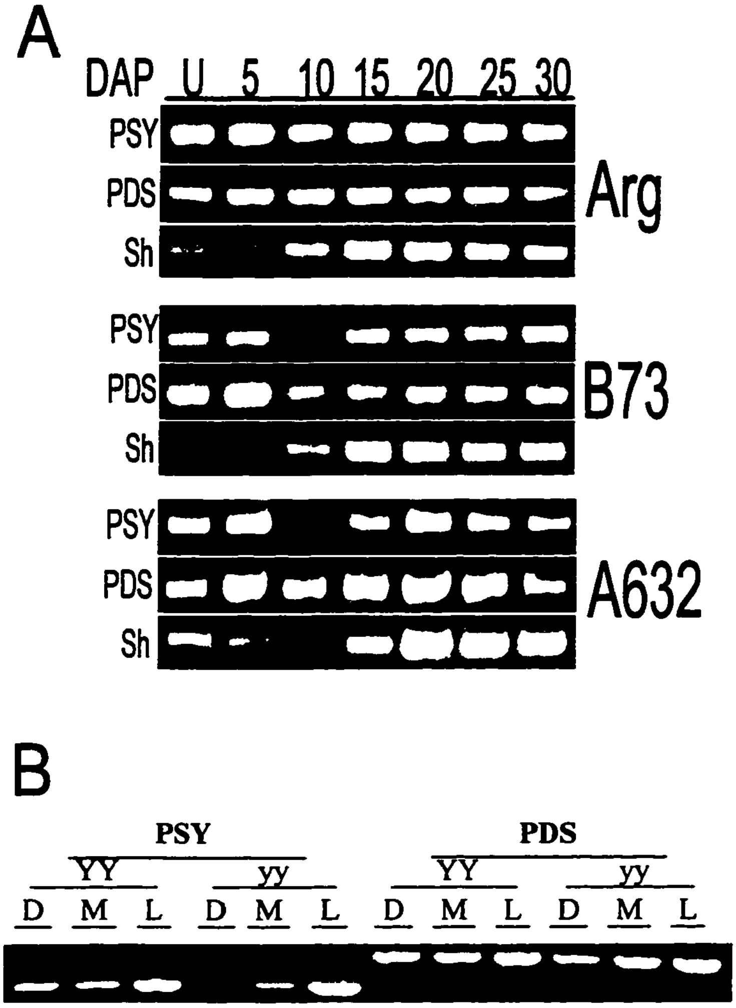


Figure 3-1. Comparison analysis of *Psy* and *Pds* expression by RT-PCR. (A) Developing series of maize endosperm. DAP, days after pollination. U, unfertilized ear. 5 DPA may include some maternal tissue. Arg, high carotenoid Argentina line. B73 and A632 are standard maize inbred lines. (B) *Psy* and *Pds* expression in different tissues of *Y1* or *y1* homozygous. YY stands for *Y1* homozygous while yy stands for *y1* homozygous. D, endosperm; M, embryo; L, leaf.

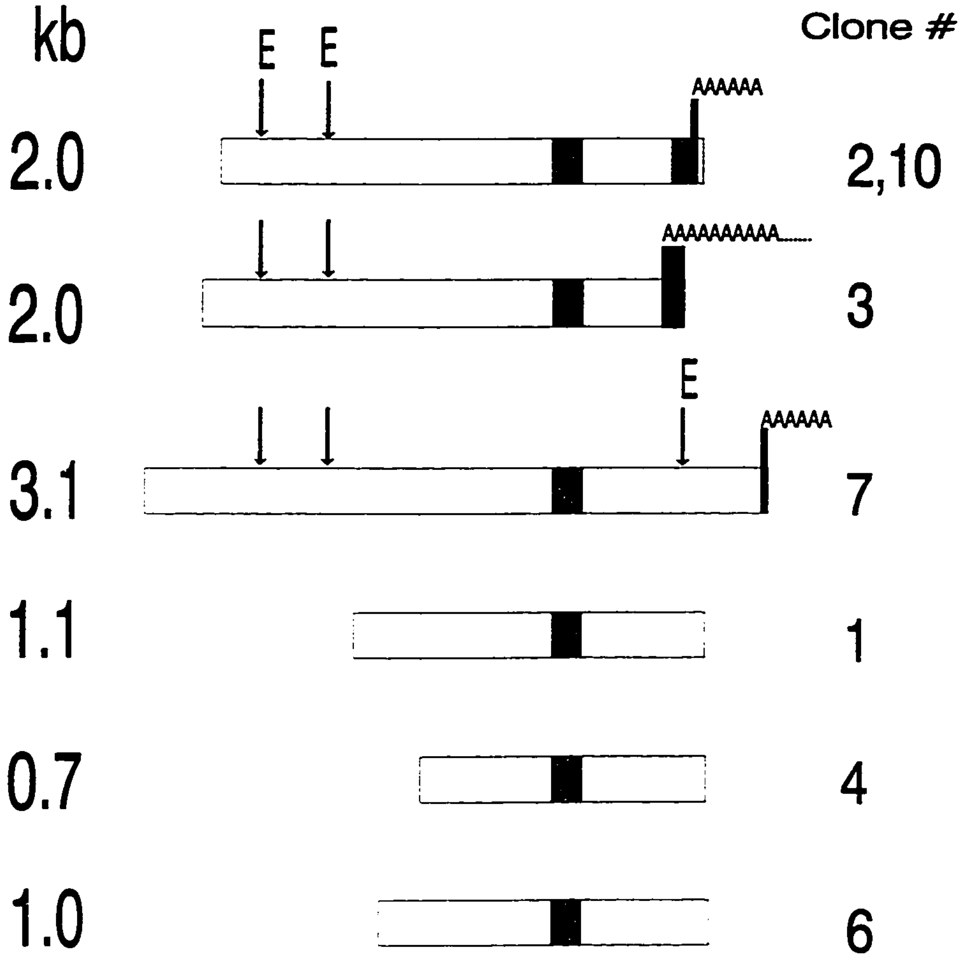
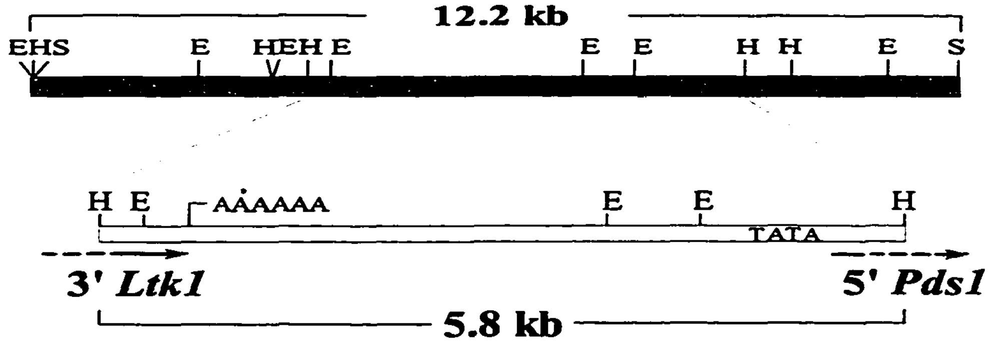


Figure 3-2. A diagram of the structure of seven *Pds* clones. Clone numbers are shown at the right of each structure and the length of each clone is indicated in kb (kilo base pair) at the left of each structure. Cross line bar indicates the binding site of the probe used to screen cDNA library. Solid bar with "AAA" shows the relative polyA tail length and the straight line bar shows the different sequence between clone 2/10 and 3 before polyA tail. E, *Eco*R I site.

```
   1 CTCCAAATGCGGAGGTCTCGACTCTTCTCTCTTCCTCCATCTTTATCATCGCCCCACGTA
  61 CACACCCAATTCCTCGCAACTGGGCTCCCCCGCCTCCACGACACTGCCCCCCCGTCTCAAG
 121 TCCGCCGCCTCCATTCTTCAGCTCTCCTATCCTCCGCCTAGAATATCTTCATCGGTATTT
 181 TACCAACCTGGATCAATTTACTCACGATACTCTGAAGCGTATACATATGCCATATGGGAA
 241 ATGACTTCATAGCTGTGGGTTGTCTTATGGCTCCTTGAATTTGCAGTAGTCTGCCTGTAC
 301 CTATTGGCTGAAGCAGAGCTGACCCCCACTTTATCAAGAGTTGCTCAACGATGGACACTG
   1                                                   M  D  T
 361 GCTGCCTGTCATCTATGAATATTACTGGAGCTAGCCAGACAAGATCTTTTGCGGGGCAAC
   4 G  C  L  S  S  M  N  I  T  G  A  S  Q  T  R  S  F  A  G  Q
 421 TTCCTCCTCAGAGATGTTTTGCGAGTAGTCACTATACAAGCTTTGCCGTGAAAAAACTTG
  24 L  P  P  Q  R  C  F  A  S  S  H  Y  T  S  F  A  V  K  K  L
 481 TCTCAAGGAATAAAGGAAGGAGATCACACCGTAGACATCCTGCCTTGCAGGTTGTCTGCA
  44 V  S  R  N  K  G  R  R  S  H  R  R  H  P  A  L  Q  V  V  C
 541 AGGATTTTCCAAGACCTCCACTAGAAAGCACAATAAACTATTTGGAAGCTGGACAGCTCT
  64 K  D  F  P  R  P  P  L  E  S  T  I  N  Y  L  E  A  G  Q  L
 601 CTTCATTTTTTAGAAACAGCGAACGCCCCAGTAAGCCGTTGCAGGTCGTGGTTGCTGGTG
  84 S  S  F  F  R  N  S  E  R  P  S  K  P  L  Q  V  V  V  A  G
 661 CAGGATTGGCTGGTCTATCAACAGCGAAGTATCTGGCAGATGCTGGCCATAAACCCATAT
 104 A  G  L  A  G  L  S  T  A  K  Y  L  A  D  A  G  H  K  P  I
 721 TGCTTGAGGCAAGAGATGTTTTGGGTGGAAAGGTAGCTGCTTGGAAGGATGAAGATGGAG
 124 L  L  E  A  R  D  V  L  G  G  K  V  A  A  W  K  D  E  D  G
 781 ATTGGTACGAGACTGGGCTTCATATATTTTTTGGAGCTTATCCCAACATACAGAATCTGT
 144 D  W  Y  E  T  G  L  H  I  F  F  G  A  Y  P  N  I  Q  N  L
 841 TTGGCGAGCTTAGGATTGAGGATCGTTTGCAGTGGAAAGAACACTCTATGATATTCGCCA
 164 F  G  E  L  R  I  E  D  R  L  Q  W  K  E  H  S  M  I  F  A
 901 TGCCAAACAAGCCAGGAGAATTCAGCCGGTTCGATTTCCCAGAAACTTTGCCAGCACCTA
 184 M  P  N  K  P  G  E  F  S  R  F  D  F  P  E  T  L  P  A  P
 961 TAAATGGGATATGGGCCATATTGAGAAACAATGAAATGCTTACTTGGCCGGAGAAGGTGA
 201 I  N  G  I  W  A  I  L  R  N  N  E  M  L  T  W  P  E  K  V
1021 AGTTTGCAATCGGACTTCTGCCAGCAATGGTTGGTGGTCAACCTTATGTTGAAGCTCAAG
 221 K  F  A  I  G  L  L  P  A  M  V  G  G  Q  P  Y  V  E  A  Q
1081 ATGGCTTAACCGTTTCAGAATGGATGAAAAAGCAGGGTGTTCCTGATCGGGTGAACGATG
 241 D  G  L  T  V  S  E  W  M  K  K  Q  G  V  P  D  R  V  N  D
1141 AGGTTTTTATTGCAATGTCCAAGGCACTCAATTTCATAAATCCTGATGAGCTATCTATGC
 261 E  V  F  I  A  M  S  K  A  L  N  F  I  N  P  D  E  L  S  M
1201 AGTGCATTTTGATTGCTTTGAACCGATTTCTTCAGGAGAAGCATGGTTCTAAAATGGCAT
 281 Q  C  I  L  I  A  L  N  R  F  L  Q  E  K  H  G  S  K  M  A
1261 TCTTGGATGGTAATCCGCCTGAAAGGCTATGCATGCCTATTGTTGATCACATTCGGTCTA
 301 F  L  D  G  N  P  P  E  R  L  C  M  P  I  V  D  H  I  R  S
1321 GGGGTGGAGAGGTCCGCCTGAATTCTCGTATTAAAAAGATAGAGCTGAATCCTGATGGAA
 321 R  G  G  E  V  R  L  N  S  R  I  K  K  I  E  L  N  P  D  G
1381 CTGTAAAACACTTCGCACTTAGTGATGGAACTCAAATAACTGGAGATGCTTATGTTTGTG
 341 T  V  K  H  F  A  L  S  D  G  T  Q  I  T  G  D  A  Y  V  C
1441 CAACACCAGTCGATATCTTCAAGCTTCTTGTACCTCAAGAGTGGAGTGAAATTACTTATT
 361 A  T  P  V  D  I  F  K  L  L  V  P  Q  E  W  S  E  I  T  Y
1501 TCAAGAAACTGGAGAAGTTGGTGGGAGTTCCTGTTATCAATGTTCATATATGGTTTGACA
 381 F  K  K  L  E  K  L  V  G  V  P  V  I  N  V  H  I  W  F  D
1561 GAAAACTGAACAACACATATGACCACCTTCTTTTCAGCAGGAGTTCACTTTTAAGTGTCT
 401 R  K  L  N  N  T  Y  D  H  L  L  F  S  R  S  S  L  L  S  V
1621 ATGCAGACATGTCAGTAACCTGCAAGGAATACTATGACCCAAACCGTTCAATGCTGGAGT
 421 Y  A  D  M  S  V  T  C  K  E  Y  Y  D  P  N  R  S  M  L  E
1681 TGGTCTTTGCTCCTGCAGACGAATGGATTGGTCGAAGTGACACTGAAATCATCGATGCAA
 441 L  V  F  A  P  A  D  E  W  I  G  R  S  D  T  E  I  I  D  A
1741 CTATGGAAGAGCTAGCCAAGTTATTTCCTGATGAAATTGCTGCTGATCAGAGTAAAGCAA
 461 T  M  E  E  L  A  K  L  F  P  D  E  I  A  A  D  Q  S  K  A
1801 AGATTCTTAAGTATCATATTGTGAAGACACCGAGATCGGTTTACAAAACTGTCCCAAACT
 481 K  I  L  K  Y  H  I  V  K  T  P  R  S  V  Y  K  T  V  P  N
1861 GTGAGCCTTGCCGGCCTCTCCAAAGGTCACCTATCGAAGGTTTCTATCTAGCTGGTGATT
 501 C  E  P  C  R  P  L  Q  R  S  P  I  E  G  F  Y  L  A  G  D
1921 ACACAAAGCAGAAATACCTGGCTTCTATGGAAGGTGCAGTCCTATCCGGGAAGCTTTGTG
 521 Y  T  K  Q  K  Y  L  A  S  M  E  G  A  V  L  S  G  K  L  C
1981 CCCAGTCCATAGTGCAGGATTATAGCAGGCTCGCACTCAGGAGCCAGAAAAGCCTACAAT
 541 A  Q  S  I  V  Q  D  Y  S  R  L  A  L  R  S  Q  K  S  L  Q
2041 CAGGAGAAGTTCCCGTCCCATCTTAGTTGTAGTTGGCTTTAGCTATCGTCATCCCCACTG
 561 S  G  E  V  P  V  P  S  *
2101 GGTGCTATCTTATCTCCTATTTCAATGGGAACCCACCCAATGGTCATGTTGGAGACAACA
2161 CCTGTTATGGTCCTTTGACCATCTCGTGGTGACTGTAGTTGATGTCATATTCGGATATAT
2221 ATGTAAAAGGACCTGCATAGCAATTGTTAGACCTTGGAAAAAAA
```

Figure 3-3. The full-length sequence of *Pds*, deduced from the 5' end of the insert of pMPDS7 combined with the Pds sequence of pMPDS3. Deduced amino acid sequence is shown in single letter code below the nucleotide sequence. Stop codons flanking and in frame with the ORF are bold and the start codon is bold and underlined. The first residue, M, is bold and the stop position is shown with an asterisk. The original *Pds* sequence (insertion of pMPDS3, GenBank # U37285) starts at the AUG codon (underlined). GenBank # U37285 was updated to reflect the 5' untranslated region determined from pMPDS7.

A



B

```
   1 ATGAGAGTTTGACGAGCCATTCAATTTTAACAAAGTTATATACTGCACAAGTGCGTGGTT
  61 CACTTACTTTGAAGATAATAAGCAACAAGATGCCCCCACTTACAGCAGTCCATTGTTTTT
 121 TCTTTCTTCTGGATTTTCATACAGTAATGACATGCAGTTCTTTTTATCATATACATAGTT
 181 TAACCCTCTGTGTTTGTTTTAAAGACGTTGAAAAAAGGCCCATGAAGCCTGCACCGCTCG
 241 TATCATTTACATGCACGAAAAGAGGTGTAGGCCATCGTTGCCATGTTTTTGTTTTTATGA
 301 CCATGCTCATGGTCATAATCAGTTGGAGGGGGTGATTCAGGATGTTGGAGGGGTGAGGGA
 361 CTCGATTTCTTGCAGGCATGCAATTGTTAACAAATGTTGAGCTACTGAAGATGCCAAGCA
 421 TGGTATATTTACACCAATATTATTATACTGGAGCAAATATTTATCCACTAGAAAAGGTCC
 481 GCCTTGTTACGGTTGACAGTCACCGACAGCATGTCACCCTAATGTTTTGGTGAGTATGGG
 541 AGATGCGGCCAGAGTTCTAGGCCACTTTACCTTCATTTCATGGCGACTGAAGATAAACCT
 601 GCGTACTACGTCTCAAAAGGAGGTGAGCTGGGAGGAAAGACATCCATCTTTTATCTTTGT
 661 GCGGGGTGGGTCTGAAACAATAATAATTATTATTTTACCAAATCTGGCTACTATGTATTG
 721 ATGTTAACTATTATTCCATTAAGATGTACTCGCTCTGTACTAAAATAGTAATCATTTTTG
 781 TCTGTTATTTTTATGTCTATATTCAGACAGATGATAATAAATCTAGACACATATACCAAG
 841 TAATGAATAAACTCATTAATTATCTAAAACGAATTTTAATTTAGGATATAGGGAGTATGA
 901 ACAACTGTTTACAATTAGTTCCCGTATTGATAATATTTATTCAACATAAATTTTTAATAT
 961 ATTTGGATCGACGTATTTATAATAATACTATATACTGTTCTATATCTATATTTAATTATG
1021 AAGTTTATCGTCACTATTGTAGTCACTCCGTCTGTTGCTAAGGTAATATGTTAGTTCGTC
1081 ACTATGTATTTCCGCGGCATCACACGGGCTAGGAAAAAAAAAAAAACAGCATAGATTATCG

1141 GAGTACTCCGGGTTAGCGGTATATACCACGGAAACGTAGAGCGCGCCAGGCGCCAAAAGG

1201 CATCCTCCACCCGCACATCTTCCTCATCCCGGTCTCGTCGTGCGCTTGTCCCTTTCTACG
      ↓
1261 GCTCCAAACAAATGCGGAGGTCTCCACTCTTTCTC--TTCCTCCATCATTATCATCGCCC
      |||||||    |||||||||||| |||||| |||  |||||||||||||||||||||||
   1  CTCCAAA----TGCGGAGGTCTCGACTCTT-CTCTCTTCCTCCATCTTTATCATCGCCC
1319 -ACGATAAACACGTCCAATTGCCTCGCAACTGGGCTCCCCCGCCTCCACGTACACTGCCC
      |||     |||  |||||| |||||||||||||||||||||||||||| ||||||||||
  55 CACGTA---CACACCCAATT-CCTCGCAACTGGGCTCCCCCGCCTCCACG-ACACTGCCC
1378 CCCGCTTCAAGTCCGCCGACTTCCATTCTTCAGGTCTCTCTATCGTCCGGTGAGTAGCCC
     ||||  |||||||||||| || ||||||||||| ||||  |||| ||||
 100 CCCGTCTCAAGTCCGCCGCCT-CCATTCTTCAGCTCTCC-TATCCTCCG
1438 CCTCCGCAGAATCTCTCTCTCTCTCGCCGATCGGATATCGTCGCATTCGCACTCGCACGG
1598 CGCATCGTCGCCTGCTTGATGCGCGGGAGTCCGCGCTCGTCTTCCACCATCGCGCCGGCT
1558 GCACCGCTCGTCCCGCCGGAGTTATCTTCTTTCCCATCTCCGGCGGTCTCGTGAAGGGCC
1618 GATCGGGGGAGGCATGTTGGCAGTAGGCTTGGCGGATTGGGTGGGCGGCCGGGCAGTTGG
1678 TTTGGGCAGGGGTGACCACTGGCCATTGCTGCTGCGGGGATGGAAGGAGGTCGCCGCGGC
1738 TGCGGCGTCGGCCGGGCCCGTTGGACGCTGGGGGCGGCAACCGGGGGCTGATGCCGGGGA
1898 CCGGTTAATTTTGATCGGGCCTTTAAGCTT
```

Figure 3-4. The putative promoter of *Pds* from B73 line.

(A) Restriction map of p624 and pH58. Shaded bar represents p624 while open bar represents pH58. E, *Eco*R I; H, *Hin*d III; S, *Sal* I. Dashed arows indicate the transcription direction of *Ltk1* and *Pds1* repectively. TATA, TATA box; AAA, polyA tail.

(B) Sequence of the *Pds* putative promoter. The putative TATA box is shown in bold within the box with the conserved TATA sequence underlined. The putative Cap signal is underlined and the first two nucleotides of the first intron are italicized and underlined. The full length *Pds* sequence of pMPDS7 (Figure3-3) starts at C indicated with ı. The the first exon of *Pds* is aligned with nucleotide 1 to 156 of the full length *Pds* sequence. Idetical nucleotides are indicated with | .

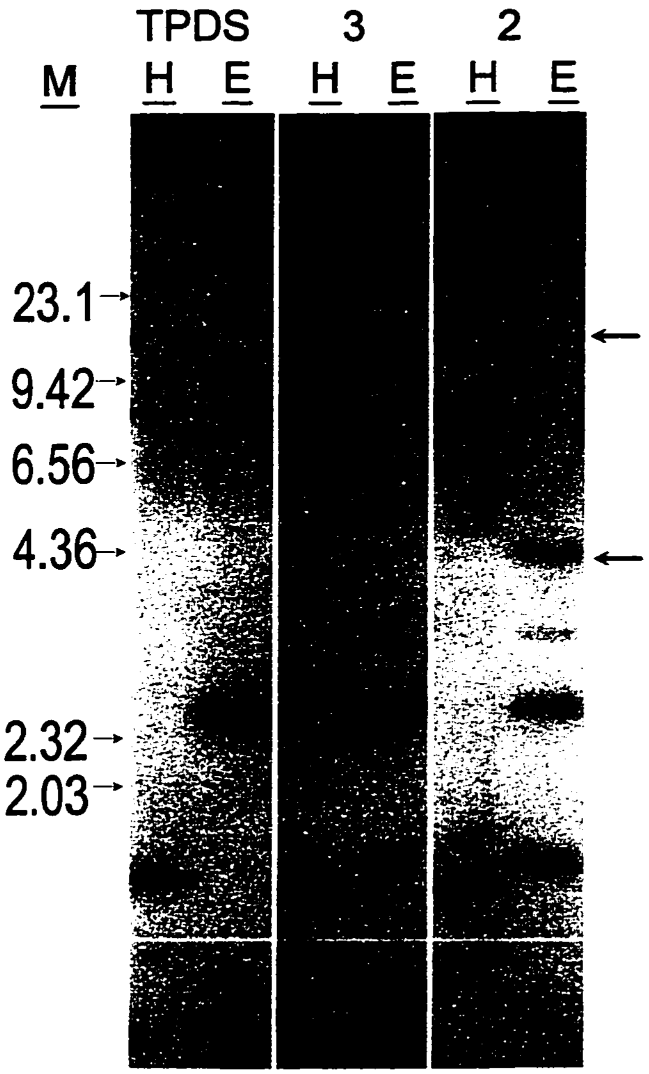


Figure 3-5. Southern analysis of maize *Pds* clones. Probes used were TPDS, tomato *Pds* cDNA; 3, pMPDS3; and 2, pMPDS2. Unique bands probed with pMPDS2 are indicated by large arrows. H, *Hind* III; E, *Eco*R I; M, molecular marker in kb.

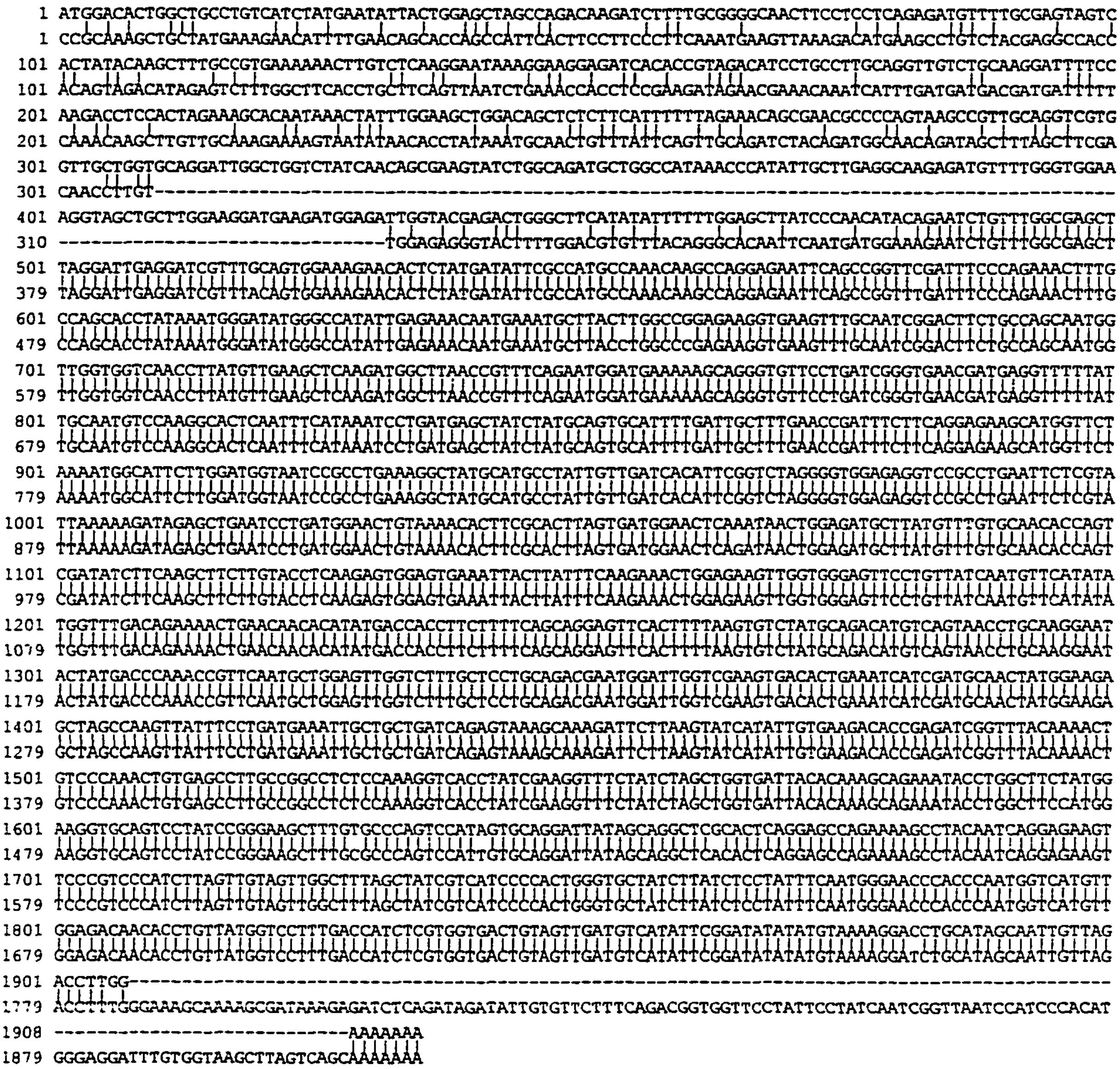


Figure 3-6. Comparison of nucleotide sequences of pMPDS3 and pMPDS10. Identical nucleotides are shown with |. The top sequence is that of pMPDS3, and the bottom one is of pMPDS10.

```
ZMPDS3      MDTGCLSSMNITGASQTRSFAGQLPPQRCFASSHYTSFAVKKLVSRNKGR 50
            |          |                      . |  || .
ZMPDS10     MKEHFEQHQPFTS---------------------FPSNEVKDMKPVYEAT 29

ZMPDS3      RSHRRHPALQVVCKDFPRPPLESTINYLEAGQLSSFFRNSERPSKPLQVV 100
                   |        | | .|    . .    |           |..
ZMPDS10     TVDIESLASPASVNLKPPPKIERNKSFDDDDDFSNKLVAKKSNITPINAT 79

ZMPDS3      VAGAGLAGLSTAKYLADAGHKPILLEARDVLGGKVAAWKDEDGDWYETGL 150
            |       ..|  .  |                              |
ZMPDS10     VYSVADLQMATDSFSFD------------------------NLVGEGTFG 105

ZMPDS3      HIFFGAYPNIQNLFGELRIEDRLQWKEHSMIFAMPNKPGEFSRFDFPETL 200
             ..   .    |||||||||||||||||||||||||||||||||||||||
ZMPDS10     RVYRAQFNDGKNLFGELRIEDRLQWKEHSMIFAMPNKPGEFSRFDFPETL 155

ZMPDS3      PAPINGIWAILRNNEMLTWPEKVKFAIGLLPAMVGGQPYVEAQDGLTVSE 250
            ||||||||||||||||||||||||||||||||||||||||||||||||||
ZMPDS10     PAPINGIWAILRNNEMLTWPEKVKFAIGLLPAMVGGQPYVEAQDGLTVSE 205

ZMPDS3      WMKKQGVPDRVNDEVFIAMSKALNFINPDELSMQCILIALNRFLQEKHGS 300
            ||||||||||||||||||||||||||||||||||||||||||||||||||
ZMPDS10     WMKKQGVPDRVNDEVFIAMSKALNFINPDELSMQCILIALNRFLQEKHGS 255

ZMPDS3      KMAFLDGNPPERLCMPIVDHIRSRGGEVRLNSRIKKIELNPDGTVKHFAL 350
            ||||||||||||||||||||||||||||||||||||||||||||||||||
ZMPDS10     KMAFLDGNPPERLCMPIVDHIRSRGGEVRLNSRIKKIELNPDGTVKHFAL 305

ZMPDS3      SDGTQITGDAYVCATPVDIFKLLVPQEWSEITYFKKLEKLVGVPVINVHI 400
            ||||||||||||||||||||||||||||||||||||||||||||||||||
ZMPDS10     SDGTQITGDAYVCATPVDIFKLLVPQEWSEITYFKKLEKLVGVPVINVHI 355

ZMPDS3      WFDRKLNNTYDHLLFSRSSLLSVYADMSVTCKEYYDPNRSMLELVFAPAD 450
            ||||||||||||||||||||||||||||||||||||||||||||||||||
ZMPDS10     WFDRKLNNTYDHLLFSRSSLLSVYADMSVTCKEYYDPNRSMLELVFAPAD 405

ZMPDS3      EWIGRSDTEIIDATMEELAKLFPDEIAADQSKAKILKYHIVKTPRSVYKT 500
            ||||||||||||||||||||||||||||||||||||||||||||||||||
ZMPDS10     EWIGRSDTEIIDATMEELAKLFPDEIAADQSKAKILKYHIVKTPRSVYKT 455

ZMPDS3      VPNCEPCRPLQRSPIEGFYLAGDYTKQKYLASMEGAVLSGKLCAQSIVQD 550
            ||||||||||||||||||||||||||||||||||||||||||||||||||
ZMPDS10     VPNCEPCRPLQRSPIEGFYLAGDYTKQKYLASMEGAVLSGKLCAQSIVQD 505

ZMPDS3      YSRLALRSQKSLQSGEVPVPS 571
            ||||.||||||||||||||||
ZMPDS10     YSRLTLRSQKSLQSGEVPVPS 526
```

Figure 3-7. Comparison of deduced amino acid sequences of pMPDS3 and pMPDS10. The transit sequence of PDS is italicized and the dinucleotide binding site of PDS is underlined. The putative ATP binding site of a novel protein kinase is bold and underlined. Identical residues are indicated with | and similar residues are indicated with a dot (•).

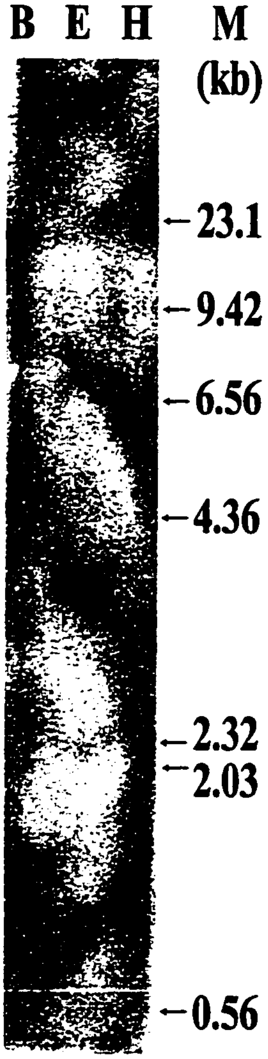


Figure 3-8. Southern analysis of maize genome (B73 line) with a 5.8 kb *Hin*d III fragment, insert of pH58. B, *Bam*H I; E, *Eco*R I; H, *Hin*d III. M, molecular marker in kb.

Chapter 4: The *Ltk* gene family encodes novel receptor-like kinases with temporal expression in developing maize endosperm[2].

4.1 Abstract

Maize cDNAs that are transcribed from a small gene family and encode a novel group of receptor-like kinases (RLKs) were isolated and characterized. The distinctive extracellular domain of these novel RLKs includes a unique number and arrangement of leucine-rich repeats (LRRs), a proline-rich region (PRR), a putative protein degradation target sequence (PEST), and a serine-rich region (SRR). The intracellular domain contains a putative serine/threonine protein kinase. To distinguish them from other reported RLKs, these novel RLKs were termed leucine-rich repeat transmembrane protein kinases (LTKs). Based on analysis of available deduced protein sequences, LTK1 and LTK2 were predicted to be 92.1% identical, while LTK2 and LTK3 were predicted to be 97.5% identical. Though the three LTK proteins showed high homology, the region that most distinguished LTK1 from LTK2 and LTK3 was found in the extracellular domain, in the SRR. To differentiate between expression of the individual *Ltk* genes, RT-PCR was used in combination with restriction enzyme analysis. While *Ltk1* transcripts were constantly present in all tissues tested, *Ltk2* and *Ltk3* transcripts were only detected in the endosperm. Furthermore, transcript levels for both *Ltk1* and *Ltk2* showed modulation

---

[2] The revised version of this Chapter has been accepted for publication (Li and Wurtzel, in press).

during endosperm development, peaking at 20 days after pollination. These results suggest that members of the *Ltk* gene family mediate signals associated with seed development and maturation.

4.2 Introduction

Cell surface receptors of higher plants transduce primary signals leading to development of reproductive organs, growth of vegetative tissue, cell differentiation, and disease resistance. In the last few years a number of genes encoding higher plant receptor-like protein kinases (RLKs) have been discovered, although the receptor ligands or signals are largely unknown. These RLKs have a tripartite structure: an extracellular ligand binding domain, coupled to a transmembrane motif and an intracellular serine/threonine protein kinase (Walker, 1994). The higher plant RLKs have been classified based on structural characteristics of the extracellular putative ligand binding domain. Plant RLKs of the S type, have an extracellular domain similar to that of the self-incompatibility locus glycoprotein of *Brassica*. Another major class of RLKs have extracellular domains distinguished by a variable number of leucine-rich repeats (LRRs). A third class, represented by only one member from *Arabidopsis*, has an epidermal growth factor-like extracellular domain (Kohorn et al., 1992). Recently, several new RLKs have been isolated, including a TNFR (tumor necrosis factor receptor)-like receptor kinase of maize (Becraft et al., 1996), a PR5 (pathogenesis-related protein 5)-like receptor kinase of

Arabidopsis (Wang et al., 1996), a lectin-like receptor kinase of *Arabidopsis* (Herve et al., 1996), and a novel RLK of *Catharanthus roseus* (Schulze-Muth et al., 1996).

Because of its multiple roles, the LRR-containing class is particularly intriguing. Among the known functions for LRR-containing RLKs, are several significant physiological roles including cell differentiation (Mu et al., 1994), plant development (Clark et al., 1997; Torii et al., 1996) and disease resistance (Song et al., 1995). The LRR class of receptor kinases is found almost exclusively in plants (Chang et al., 1992; Clark et al., 1997; Hong et al., 1997; Mu et al., 1994; Schmidt et al., 1997; Song et al., 1995; Torii et al., 1996; Valon et al., 1993; van der Knaap et al., 1996; Walker, 1993) with the exception of the *trk* tyrosine kinase receptors of animals which have only three LRRs in the extracellular domain (Schneider and Schweiger, 1991). The LRR class may represent an extracellular-signal transmitting system unique to higher plants and which has evolved independently of the well-known tyrosine kinase signal transduction system of animals.

Members of the LRR class share a conserved intracellular serine/threonine protein kinase domain but a less conserved extracellular domain, varying in the number, length and arrangement of LRR repeats; therefore, it is possible that each member of this class transfers a different signal into the cell. These signals are likely to be polypeptides as suggested by several studies, including crystal structure analysis, implicating LRR motifs in protein-protein interaction (Kobe and Deisenhofer, 1995; Krantz and Zipursky, 1990).

Although specific peptide ligands have yet to be found, examples of peptide signal molecules, such as ENOD40 and systemin, have recently been found in plants (Schaller and Ryan, 1996; van de Sande et al., 1996). Alternatively, LRRs may function in cell adhesion, mediating the orientation of cells during development, as is the case for the *Drosophila* proteins Toll and chaoptin (Keith and Gay, 1990; Krantz and Zipursky, 1990). About ten LRR-containing RLKs have been found in higher plants, but neither the receptor ligand, nor the downstream substrate of the intracellular protein kinase domain has been identified.

Further identification and characterization of these receptor kinases are central to understanding signal transduction in plant growth and development and environmental responses. In maize, only two RLK genes have been identified. The first RLK isolated (Walker and Zhang, 1990) was of the S type. Recently, an RLK involved in maize epidermal differentiation was cloned by transposon tagging (Becraft et al., 1996). Here, the isolation and characterization of cDNAs encoding a novel group of receptor-like kinases are reported. The unique extracellular domain of this group of RLKs contains an LRR region that may function as a ligand binding domain, a proline-rich region (PRR) and serine-rich region (SRR) that might serve a structural role in the ligand binding domain, and a novel PEST sequence that might play a role in turnover of these novel RLKs. These cDNAs, encoded by a small gene family, exhibited a unique profile of expression in developing endosperm. Therefore, a possible role of these novel RLKs in endosperm development is discussed.

4.3 Materials and Methods

4.3.1 Plant materials

Maize plants (B73 inbred line) grown in a growth chamber at 25°C (12 hrs light/12 hrs dark), were harvested after twenty days when plants were at the two to three-leaf stage. Harvested plants were separated into shoots and roots. For collection of endosperm and embryo samples, B73 maize was grown in our experimental field at Pelham Bay Park, Bronx, New York. Ears were harvested at 10, 15, 20 and 25 days after pollination (DAP) and the endosperm separated from the embryo. Unfertilized ovules were separated from unfertilized ears. All tissues were frozen in liquid nitrogen and stored at -80°C.

4.3.2 Cloning *Ltk1* cDNA by 5' and 3' rapid amplification of cDNA ends (RACE)

An *Ltk1* cDNA was produced by using 5' and 3' RACE to extend sequences present in pMPDS10, a 1.9 kb chimeric cDNA clone obtained in screening for cDNAs encoding phytoene desaturase (PDS) (Li et al., 1996), that contained 357 bp of *Ltk1* (nt 861-1218) upstream of 1.5 kb of *Pds1* (starting at nt 481, GenBank U37285). The 5' and 3' RACE reactions were carried out by using the 5' RACE system for Rapid Amplification of cDNA Ends, version 2.0 and the 3' RACE System for Rapid Amplification of cDNA Ends (Gibco BRL, Gaithersburg, MD, USA), respectively. For 5' RACE, first strand cDNA was synthesized from total leaf RNA using primer 189 as GSP1 (gene specific primer one), 5'-CGTCCAAAAGTACCCTCT-3', and tailing with polyG following the vendor's procedure. The first round PCR amplification was carried out with primer 208 as GSP2 (5'-AAGACTCTATGTCTACTGTGG-3') and AAP (5' RACE Abridged Anchor Primer, 5'-GGCCACGCGTCGACTAGTACGGGIIGGGIIGGGIIG-3', vendor-provided) in 20 mM Tris-HCl, pH 8.4; 50 mM KCl; 2.5 mM $MgCl_2$; 0.2 mM each dNTP; 0.4 mM each primer, and 0.05 U/ml Taq DNA polymerase (Gibco-BRL). The PCR protocol was: one cycle of 94°C (2 min); followed by 35 cycles of 94°C (30 sec), 66°C (30 sec), 72°C (2 min); and one cycle of 72°C (10 min). The first-round amplification product was diluted 1:20 and then used as a template for a second-round of amplification with GSP2 and AUAP (Abridged Universal Amplification Primer, 5'-GGCCACGCGTCGACTAGTAC-3', vendor-provided). The reaction conditions and protocol for second-round PCR were the same as for the first, except that the annealing temperature was 62°C. The 990 bp PCR product obtained from the second-round amplification using primers GSP2 and AUAP, was treated with Klenow and subcloned into pBluescript II SK(-) at the *Sma* I site, as in Li *et al.* (Li et al., 1996). This clone was designated as p5R900.

For 3' RACE, the first strand cDNA was synthesized from total endosperm RNA using AP (Adapter Primer, 5'-GGCCACGCGTCGACTAGTACTTTTTTTTTTTTTTTTT-3', vendor- provided). PCR was carried out using primer 219 as GSP (5'-GGGCATTAC TTGCTCAGGAT-3') and AUAP for first and second round amplifications. Conditions for both PCR rounds were identical: reactions contained 20 mM Tris-HCl, pH 8.4; 50 mM KCl; 2.0 mM $MgCl_2$; 0.2 mM each dNTP; 0.4 mM each primer, and 0.05 U/ml Taq DNA polymerase (Gibco BRL). The PCR protocol was: one cycle of 94°C (2 min); followed by 35 cycles of 94°C (30 sec), 62°C (30 sec), 72°C (2 min); and one cycle of 72°C (10 min). The 2.2 kb PCR product was treated with Klenow and subcloned into pBluescript II SK (-) at the *Eco*R V site, as in Li *et al* (Li et al., 1996). This clone was designed as p3R41, and contained a 2.2 kb segment (nt 113-2310) of *Ltk1*. The 990 bp 5' RACE and 2.2 kb 3' RACE products were sequenced and found to overlap in an 865 nt region having 100% homology.

Prior to combining the 5' and 3' RACE products, the 5' RACE product was first subcloned into pMPDS10. To extend the 5' end of *Ltk1* contained within pMPDS10, a 112 nt *Spe* I (vector site)-*Ahd* I fragment was removed and replaced by a 985 nt *Spe* I (restriction site in the AUAP adaptor sequence)-*Ahd* I fragment from p5R900 that contained nt 1-971 of *Ltk1*. The resulting clone, p10F(E) contained nt 1-1218 of *Ltk1* followed by the 3' end of *Pds1*.

Afterwards, the 3' RACE product was inserted in place of the *Pds* sequence. The 3' RACE product (p3R41) and p10F(E) overlapped between nt 113 and 1218 of *Ltk1*. This region also contained a unique *Bsp*M I site, thereby providing a means to replace the *Pds1* sequence with the 3' RACE product. The insert from p3R41 was released by digestion at the common *Bsp*M I site, and to create a blunt end, digested at a unique *Sma* I site within the vector. This fragment (containing nt 995-2310 of *Ltk1*) was then combined with p10F(E) which had been digested by *Bsp*M I and *Eco*R V, creating a blunt end from the vector. This clone, designated as pLTK1, contained nt 1-2310 of *Ltk1*.

4.3.3 Cloning the *Ltk2* cDNA by RT-PCR

Based on Northern analysis, the *Ltk1* cDNA appeared to be about 200 bp shorter than expected. The following, which was carried out to recover sequence information further upstream of *Ltk1*, instead led to amplification of the *Ltk2* cDNA. A maize B73 genomic DNA library (CLONTECH Laboratories, Inc., Palo Alto, CA, Cat. # FL 1032D), containing 7×10^6 phage clones, was screened with the insert of p5R900 and sixty-four positive clones were isolated. Eight clones were randomly chosen for further analysis, the purified λ DNA digested by *Eco*R I, *Sal* I, or *Bam*H I, and hybridized by a probe corresponding to the very 5' end of the 5' RACE product. This probe was a PCR product, approximately 300 bp, generated from clone p10F(E) using primer 227, 5'-GGGAAGA TTGTACTGTACTTG-3' and primer 231, 5'-CGCAATAACCCTCACTAAAGG-3', derived from the vector, pBluescript II SK (-). The PCR reaction was identical to that for 3' RACE except for using 1 ng of plasmid DNA as template. The PCR protocol was: one cycle of 94°C (2 min); followed by 35 cycles of 94°C (30 sec), 65°C (30 sec), 72°C (30 sec); and one cycle of 72°C (10 min). Hybridizing fragments were subcloned into pBluescript II SK(-) and sequencing was attempted with primer 226, 5'- CCCTGATC

CTGACAAGTAAT-3', which was 119 nt from the very 5' end of the 5' RACE product. Although several fragments hybridized to the 5' end probe, primer 226 provided sequence from only one clone, p531, containing a 16.6 kb *Sal* I insert from the genomic clone, λ531. The *Eco*R I fragments of p531 were subcloned into pBluescript II SK(-), since the 5' RACE product contained an *Eco*R I site 45 nt from the 5' end, and it was expected that one of the *Eco*R I fragments would include the missing 5' sequence. One clone, p531E2, containing a 5.07 kb *Eco*R I fragment, provided the expected sequence adjoining the *Eco*R I junction site of the 5' RACE product. However, this *Eco*R I fragment contained an intron between the corresponding nt 15 and 16 of the 5' RACE sequence. To gain additional exonic sequence at the 5' end of the gene, without sequencing through the intron, this clone was then sequenced by primer 240, 5'-ACGATCGTTGGGGTC**AA**-3', consisting of the 15 nucleotides preceding the intron combined with an additional two nucleotides (shown in bold) that could pair with the conserved 5'-GT-3' sequence of introns. Primer 240 provided limited sequence which was then used to design primer 243, 5'-TGCACTTCACTTGTCAACAG-3', which allowed sequencing back into the intron. Based on this additional sequence, primer 246, 5'-GCTCTTCCTTATCGCCAT-3', was designed, and used to obtain upstream sequence that contained a putative promoter and the first exon, as analyzed by the PCGene program (Bucher, 1990).

To confirm the position of the promoter and the first exon encoding the 5' end of the *Ltk* mRNA, primer 247, 5'-CGGATTTGGAGGAGTCGAT-3', was designed according to the putative 5' exon sequence of p531E2, and used with primer 227 (see Figure 4-1), to amplify the 5' end of the *Ltk* mRNA by RT-PCR. One microgram total endosperm RNA was used to synthesize cDNA according to the procedure of the SuperScript™ Preamplification System for First Strand cDNA Synthesis kit (Gibco BRL). One fourth of the cDNA was used as template to perform PCR with the same reaction conditions as for 3' RACE. The PCR protocol was: one cycle of 94°C (2 min); followed by 40 cycles of 94°C (30 sec), 58°C (30 sec), 72°C (30 sec); and one cycle of 72°C (10 min). Based on Northern analysis and available sequence, a 500 bp amplification product was expected and a 490 bp product was obtained. Although the expected size of the RT-PCR product was obtained, the sequence slightly differed from available *Ltk1* sequence, implying that the promoter contained by p531E2 did not belong to *Ltk1* and that the 490 bp product represented the 5' end of another *Ltk* transcript, which I named *Ltk2*.

To amplify the complete *Ltk2* by RT-PCR, the genomic DNA-based primer 247 was used with primer 232 (5'-GGGAAGATTGTACTGTACTTG-3', sequence from the 3' RACE product). One microgram total endosperm RNA was used to synthesize cDNA according to the procedure of the SuperScript™ Preamplification System for First Strand cDNA Synthesis kit (Gibco BRL). One fourth of the cDNA was used as template to perform PCR with the same reaction conditions as for 3' RACE. The PCR protocol was: one cycle of 94°C (2 min); followed by 35 cycles of 94°C (30 sec), 54°C (30 sec), 72°C (2 min); and one cycle of 72°C (10 min). The PCR product was treated with Klenow and subcloned into pBluescript II SK (-) at the *Sma* I site, as in Li *et al.* (Li et al., 1996). Several clones of identical size were obtained and one of them, containing a 2388 bp insert, was designated as pLTK2.

4.3.4 DNA Sequencing

Plasmid DNA templates were isolated by the Wizard *Plus* Miniprep DNA Purification System (Promega, Madison, WI) and DNA sequencing was carried out using the Sequenase™ Version 2.0 DNA Sequencing Kit (United States Biochemical, Cleveland, OH). Alternatively, automatic sequencing of double stranded DNA templates was used. Both strands of *Ltk1*, *Ltk2*, and *Ltk3* cDNAs were sequenced by primer walking.

4.3.5 DNA extraction and Southern analysis

Genomic DNA was extracted according to Wurtzel *et al.* (Wurtzel et al., 1987) from 20 DAP maize B73 endosperm. Following restriction enzyme digestion with *Bam*H I, *Eco*R I or *Pvu* II, Southern hybridization with probes A, B, or C, was carried as in Wurtzel *et al.* (Wurtzel et al., 1987). Filters were washed in 0.1% SDS and 0.1xSSC for 15 min at r.t. and twice for 15 min at 55°C, followed by a 2-12 hr exposure to a Storage Phosphor Screen, which was then scanned on a 445SI Molecular Dynamics Phosphorimager. Probe A was the 2.1 kb *Ltk1* 3' RACE product described above. Probe B was a 307 nt PCR product (nt 861-1168 of *Ltk1*) amplified from pMPDS10 using primer 231 (see cloning *Ltk2*) and 229, 5'-CAAGGTCGACGAAGCTAAAGC-3'. The PCR conditions were the same as for PCR with primers 231 and 227 in *Ltk2* cloning. To prepare probe C, a 490 bp RT-PCR product of *Ltk2*, total endosperm RNA was used as template for first-strand cDNA synthesis and PCR amplified using primers 247 and 227 (see cloning *Ltk2* and Figure 4-1).

4.3.6 RNA extraction and Northern analysis

Total RNA was extracted from embryo, leaf, or roots following Logemann *et al.* (Logemann et al., 1987) and from endosperm or unfertilized ovules following Wurtzel *et al.* (Wurtzel et al., 1987). RNA concentration was determined spectrophotometrically. Electrophoresis of total RNA on 1.4% (w/v) agarose gels containing 6% (v/v) formaldehyde, transfer to nitrocellulose, and hybridization conditions were as in Wurtzel *et al.* (Wurtzel et al., 1987). Filters were exposed for 24 hr to a Storage Phosphor Screen which was then scanned on a 445SI Molecular Dynamics Phosphorimager.

4.3.7 RT-PCR and restriction enzyme diagnosis of *Ltk* transcript accumulation

Total RNA was used as template for cDNA synthesis as described in the synthesis of *Ltk2* cDNA. Using one fourth of the cDNA, PCR was performed with forward primer 276, 5'-TCACAGGTTGGCAGGCGAAT-3' and reverse primer 207, 5'-CCATCTGTAGATCTG CAACTGAA-3' (see Figure 4-1) in a reaction containing 20 mM Tris-HCl, pH 8.4; 50 mM KCl; 4.0 mM $MgCl_2$; 0.2 mM each dNTP; 0.4 mM each primer; and 0.05 U/ml Taq DNA polymerase (Gibco-BRL). The PCR protocol was: one cycle of 94°C (2 min);

followed by 40 cycles of 94°C (30 sec), 54°C (30 sec), 72°C (1 min); and one cycle of 72°C (10 min). To ensure linearity of the reactions, conditions were similar to those described in Li *et al.*, 1996 (Li et al., 1996). From a 25 µl reaction, 15 µl PCR product was incubated with restriction endonuclease *Hae* III for 2 hr following the vendor's specifications (Gibco-BRL) and then analyzed by electrophoresis on a 1.6% (w/v) agarose gel in comparison with 10 µl of undigested DNA remaining from the PCR reaction. Control RT-PCR reactions for amplification of *Pds* and *Sh* transcripts were performed as in (Li et al., 1996).

4.4 Results

4.4.1 Cloning of *Ltk1* and *Ltk2* cDNAs

A chimeric cDNA clone of unusual structure was found while screening a maize endosperm cDNA library for cDNAs encoding phytoene desaturase (PDS). In addition to isolating a cDNA clone encoding PDS, as described in Li *et al.* (Li et al., 1996), one clone (pMPDS10) appeared to be chimeric; the 3' end encoded most of PDS with the exception of an essential dinucleotide binding domain and transit sequence, in place of which was an ATP binding site and partial protein kinase sequence (data not shown). I decided to use 5' RACE and 3' RACE to reconstruct a cDNA, from a maize B73 inbred line, represented by the protein kinase sequence at the 5' end of the chimeric cDNA. The 990 bp 5' RACE and 2.2 kb 3' RACE products were sequenced and found to overlap in an

865 nt region having 100% homology. Therefore, the RACE products were combined to produce a plasmid containing a reconstructed cDNA of 2310 bp, designated as pLTK1. Hybridization of pLTK1 to B73 leaf RNA detected a transcript of about 2560 nt, suggesting that pLTK1 was missing about 200 bp at the 5' end.

In an effort to isolate the missing 5' end of *Ltk1*, a maize B73 genomic DNA library was screened. Using a primer close to the very 5' end of the RACE product, I identified one clone, p531E2 subcloned from p531, with sequence (GenBank #AF023267) contiguous to that of the *Ltk1* 5' RACE product and which presumably contained part of the *Ltk1* promoter (Figure 4-2A). Sequence analysis of the putative promoter region revealed a TATA box and mRNA Cap signal as shown in Figure 4-2B. To amplify the missing 5' end of *Ltk1* by RT-PCR, I designed a forward primer (primer 247) from the genomic DNA sequence and a reverse primer (primer 227) from the truncated *Ltk1* cDNA in pLTK1 (Figure 4-1). According to the predicted site of the putative Cap signal, amplification with these primers was expected to produce a 490 bp fragment which would include approximately 200 bp upstream of the 5' RACE product. As predicted, a 490 bp RT-PCR product (probe C, Figure 4-1) was obtained. When this 490 bp fragment was sequenced, the first 215 nt matched the first exon sequence, predicted from the *Ltk* genomic clone (Figure 4-2B). However, comparison of nt 216-490 of the RT-PCR product with that of the overlapping 5' RACE product showed the RT-PCR sequence to be slightly different. This result suggested that the RT-PCR product represented a second related gene, designated *Ltk2*, and that the genomic clone actually contained the promoter region of *Ltk2* and not that of *Ltk1*. To amplify a full-length cDNA encoding this second LTK, RT-PCR was performed with primers 247 and 232 (Figure 4-1), and the 2388 bp product was cloned. This second *Ltk* cDNA clone, designated as pLTK2, shared an identical sequence as that of the RT-PCR product, but a different DNA sequence from that of *Ltk1*.

4.4.2 Comparison of *Ltk1* and *Ltk2* cDNAs and their predicted gene products

The inserts of pLTK1 and pLTK2 were fully sequenced from both directions and the sequences submitted to GenBank with accession numbers AF023164 and AF023165. Searching of GenBank with the nucleotide or predicted protein sequences of either cDNA revealed no obvious homologs. However, the predicted C termini showed similarity to protein kinases and the predicted N termini, found later to contain eight leucine-rich repeats, showed similarity of these repeats to repeats contained within proteins having various numbers of leucine-rich repeats. I also identified motifs corresponding to a signal sequence and transmembrane region. Since these proteins had the general structure of receptor-like kinases with leucine-rich repeats in the putative extracellular domain, I named them LTK proteins, for leucine-rich repeat transmembrane protein kinase.

The 2388 bp *Ltk2* cDNA (pLTK2) contained a complete open reading frame (ORF), preceded by a 5' untranslated region of 99 bp and followed by a 3' untranslated region of

114 bp. As this was an RT-PCR product, the region adjacent to and including the poly A tail was missing (when compared to *Ltk1*, this was estimated as 159 bp, assuming a similar polyadenylation site to *Ltk1* on the basis of a single band detected by Northern analysis). The pLTK2 ORF encoded a polypeptide of 725 residues with a calculated molecular weight of 79.15 kD. The 2310 bp *Ltk1* cDNA (pLTK1) encoded 684 residues in the first 2054 bp followed by a 256 bp 3' untranslated region, including the poly A tail. The *Ltk1* and *Ltk2* cDNA sequences were compared and showed 91.2% identity at the nucleotide level. A comparison of the predicted protein sequences (Figure 4-3) showed 92.1% identity and 94.5% similarity.

As shown in Figures 4-1 and 4-3, the structure of the predicted LTK2 polypeptide showed it to include a signal peptide and leucine-rich repeats (LRRs) at the amino terminus, a transmembrane region, and a protein kinase domain in the carboxyl terminus. The unique number and organization of the LRRs, combined with other peculiar sequence hallmarks, as described below, indicated that the LTK proteins were new members of the class of RLKs having LRRs in their extracellular domains. Two hydrophobic motifs were predicted by sequence analysis of LTK2 using the SOAP program of PCGene (Kyte and Doolittle, 1982) (Figure 4-3). One region, residues 11 to 27 at the amino terminus, could serve as a signal peptide to target LTK2 to the plasma membrane. The other, residues 297 to 322, a putative transmembrane domain, separated the N-terminal LRR domain from the C-terminal protein kinase domain. Though the *Ltk1* cDNA was missing the 5' sequence needed to encode a signal peptide, it encoded, as *Ltk2*, the transmembrane motif from residues 257 to 282 (Figure 4-3). The transmembrane motifs of both LTK1 and LTK2, were followed by positively charged residues KRRKR(H)K that could function as stop-transfer sequences (Weinstein et al., 1982). During insertion of LTK proteins into the plasma membrane, the stop-transfer sequence would cause the protein kinase domain to remain inside the cell, while the LRR portion of LTK would be extracellular.

As predicted from the translated ORF, the extracellular domain of LTK proteins was found to consist of eight LRRs, the first of which was separated from the other seven tandemly repeated LRRs (Figures 4-3 and 4-4A). Each repeat ranged in length from 18 to 24 amino acids and contained a conserved core region, LXXLXLXXN, found in LRRs of other RLKs (Figure 4-4B). In a hydrophilic environment, this conserved core sequence is proposed to form a b-sheet structure that binds the protein ligand (Jiang et al., 1995; Kobe and Deisenhofer, 1995).The region containing the eight LRRs showed highest similarity (about 45-64%) to various stretches of leucine-rich repeats found in the Cf-2 proteins, a pair of membrane anchored disease resistance proteins from tomato having thirty-eight LRRs, but no kinase domain (Dixon et al., 1996).A similar comparison of the eight LRRs in the LTK proteins with the twenty LRRs of ERECTA, an *Arabidopsis* RLK involved in plant development, showed a maximum similarity of 60%. However, the number and arrangement of LRRs in LTK proteins was different from that of all known LRR-type

RLKs, implying a unique ligand specificity for the LTK extracellular domains.

Unlike other LRR-type RLKs, three novel regions were found in the predicted extracellular domains of LTK1 and LTK2. One region, located between the LRR and transmembrane domain, was a putative protease target site or PEST sequence. PEST sequences are defined as hydrophilic regions of at least 12 amino acid residues, containing at least one P (proline), one E (glutamic acid) or D (aspartic acid), and one S (serine) or T (threonine), but no positively charged residues, although the region may be flanked by K(lysine), R (arginine), or H (histidine) (Rechsteiner, 1990). The PEST sequences found were "INSLQTDGNSWSTGPAPPPPPYTAPPPPPN," aa 210 to aa 239 in LTK1, and "EINNLQTDGNSWSTGPAPPPPPYTAPPPPN," aa 248 to aa 277 in LTK2 (as analyzed by the PEST-FIND program of PCGene (Rogers et al., 1986)). These sequences received PEST scores of 8.7 and 8.1, respectively, where a score greater than five indicates that the region is most likely to be degraded by proteases and that the protein will have a fast turnover rate.

The second distinct region was a proline-rich region (PRR), which actually comprised the carboxyl terminus of the PEST sequence, from aa 224 to aa 238, PAPPPPPYTAPPPPP in LTK1 and from aa 263 to aa 276, PAPPPPPYTAPPPP, in LTK2. PRRs form an extended flexible structure because of the unique structure of proline, whose side group bonds to the backbone amide position and forms a rigid structure (Williamson, 1994). For example, in Calcineurin A, a stretch of 11 contiguous prolines have been proposed to be involved in PRP (proline-rich protein)-protein interaction (Guerini, 1997). Nonrepetitive PRR sequences may also mediate protein-protein interaction as seen in the Sos PRR site that mediates binding to SH3 (*src* homology 3) (Ren et al., 1993). In the case of LTK proteins, the PRR sites were noncontiguous; in LTK1, the PRR consisted of 11 prolines in a region of 15 residues, while the PRR in LTK2 consisted of 10 prolines in a region of 14 residues. Because of its position between the LRR and transmembrane domain of LTK, the PRR may function either as a "hinge" to increase the flexibility of the extracellular ligand binding domain, or as a "linker" to extend the protein ligand binding site.

The third unique region noted was a serine-rich region (SRR) found immediately downstream of the proline-rich region. Surprisingly, this SRR most differentiated LTK1 from LTK2, proteins that were otherwise 94.5% similar. LTK1 (residues 243-256), contained ADGSSSSSSSGGRS, while LTK2 (residues 281-296) contained GAGQNDDGSSSSGGRP. The SRR sequences, which differed in number of the polar serine residues, were flanked by negatively and positively charged residues. If the SSR plays any role in ligand binding, then LTK1 and LTK2 might also differ in ligand specificity and/or affinity.

Both LTK proteins contained in their putative intracellular domains, the eleven subdomain catalytic site characteristic of protein kinases (Figure 4-3). Specific features of subdomains VIb and VIII marked the kinase domains of the LTK proteins as serine/threonine protein kinases, similar to all other higher plant RLKs, and distinct from the animal receptor kinases that are either tyrosine or serine/threonine kinases (Hanks and Quinn, 1991; Walker, 1994). Consensus sequences for subdomains VIb and VIII, respectively, are DLKPEN and GTPXYIAPE for serine/threonine kinases and DLAARN and FPIKWMAPE for tyrosine kinases. The observed sequence homology suggests that, like the other LRR-type RLKs, the LTKs are capable of autophosphorylation or phosphorylation of substrate polypeptides at serine and/or threonine residues (Walker, 1994). In contrast to the variable extracellular LRR domain, the intracellular protein kinase domain was well conserved as compared with protein kinase domains of other LRR-type RLKs and with other cytosolic protein kinases. A comparison of protein kinase domains of all LRR protein kinases and three cytosol protein kinases was carried out using the Clustal program of PCGene (Higgins and Sharp, 1988) and the result is shown in Figure 4-4C. The kinase domain of TRK, the only LRR type receptor kinase in animals was also included in the analysis, but as this is a tyrosine kinase, it is not surprising that it was the least related. Interestingly, the LTK1 and LTK2 protein kinase domains showed highest homology to Pti, a defense-related cytoplasmic protein kinase from tomato (Zhou et al., 1995).

4.4.3 The *Ltk* gene family of maize

The *Ltk* gene copy number was investigated by Southern analysis. When maize B73 genomic DNA was probed by an almost full length, 2.1 kb cDNA probe (probe A, Figure 4-1), multiple bands were observed for three different restriction enzyme digestions (Figure 4-5A). These probed fragments might represent multiple *Ltk* genes. Alternatively, these bands might represent numerous protein kinase genes, since the C-terminal protein kinase region of *Ltk* has high homology to those of other protein kinases. Therefore, an identical filter was probed by B (Figure 4-1), a 280 bp probe corresponding to a region upstream of the kinase domain, but downstream of the region encoding the transmembrane domain. As shown in Figure 4-5B, fewer bands hybridized to probe B, as compared with the number of bands that had hybridized to probe A (Figure 4-5A). In the *Pvu* II digestion, only two fragments, 7.8 kb and 6.0 kb, were detected by probe B, while in *Eco*R I or *Bam*H I digestions, three major bands were detected by probe B. Since the 280 bp probe B was a short probe, these results suggest that there is more than one copy of *Ltk* in the maize genome.

To eliminate the possibility that multiple long introns in the region hybridized by B caused multiple bands in Southern analysis, probe B was removed and the filter hybridized by probe C, DNA encoding part of the LRR extracellular domain (Figure 4-1). Probe C was 490 bp and located 605 bp upstream of probe B. If the 7.8 kb and 6.0 kb *Pvu* II fragments probed by B originated from a single *Ltk* gene, then probe C would have hybridized only to one but not to both bands. If the two *Pvu* II fragments probed by B were from two or more *Ltk* copies, C would hybridize to both bands. As shown in Figure 4-5C, probe C did hybridize to both *Pvu* II fragments, demonstrating that more than one or at least two copies of *Ltk* are present in the maize genome. The same deduction could be applied to the three bands generated by *Eco*R I or *Bam*H I digestion and hybridized by B. If these three fragments were from only two different genes and could be probed by B, then probe C would have hybridized to two but not all fragments. Nonetheless, as shown in Figure 4-5C, probe C hybridized to all three bands previously probed by B, both in the *Bam*H I and *Eco*R I digestions. This suggests that there may be three copies of *Ltk* in the maize genome.

4.4.4 Expression of *Ltk* at the RNA level

Although three *Ltk* genes were suggested by Southern analysis, only a single 2560 nt transcript was detected in total RNA extracted from roots, leaves, 20 DAP endosperm, or 20 DAP embryo of the maize B73 inbred line (Figure 4-6). Also, the transcript level varied amongst tissues; four times as much RNA was required to detect comparable levels of *Ltk* transcript in leaves and roots, as compared with that of endosperm and embryo samples. When 10 μg total RNA was loaded for each tissue, a signal could only be detected for endosperm and embryo, but not for root or leaf samples (left panel, Figure 4-6). To detect a comparable signal, root and leaf RNA samples were increased to 40 μg (right panel, Figure 4-6). Also, the role of light in modulating leaf *Ltk* transcript level was tested, but found not to have any effect (data not shown).

Although one band was detected by Northern analysis, it was difficult to determine which *Ltk* gene was transcribed, especially if gene family members encoded transcripts of similar sizes. To differentiate among *Ltk* transcripts, another approach was needed. Since *Ltk1* and *Ltk2* cDNAs exhibited nucleotide sequence differences, I used RT-PCR, in combination with analysis of restriction enzyme site polymorphisms, to differentiate transcripts of the *Ltk* genes. Primers 276 and 207 (shown in Figure 4-1) were used to amplify transcripts from leaf, root, embryo (20 DAP) and endosperm (20 DAP). Only when reverse transcriptase was added to the cDNA synthesis reaction was an approximately 1070 bp RT-PCR product obtained (Figure 4-7A, panels 1 and 2). This demonstrated that the amplification product was derived from RNA and not from contaminating genomic DNA. Furthermore, the RT-PCR product was of the expected size based on the *Ltk* cDNA sequences; the predicted sizes for *Ltk1* and *Ltk2* were 1074 and 1077 bp, respectively, which could not be differentiated by gel electrophoresis. Therefore, the restriction enzyme *Hae* III was used to distinguish among the *Ltk* transcripts since *Hae* III digestion of the 1070 bp RT-PCR product was expected to yield a major fragment of 771 bp for *Ltk1*, but an 887 bp fragment for *Ltk2* (see Figure 4-1). As shown in Figure 4-7A (panel 3), only one 771 bp fragment resulted from *Hae* III digestion of the 1070 bp RT-PCR product amplified from leaf, root or 20 DAP embryo RNA using primers 276 and 207. Thus, *Ltk1* was the only *Ltk* transcript expressed in these three tissues. However, when 20 DAP endosperm RNA was used, three restriction fragments, 771 bp, 887 bp and a faint 957 bp fragment, were obtained. The 771 bp and the 887 bp fragments implied expression of *Ltk1* and *Ltk2*, respectively. The presence of the 957 bp fragment was unexpected. These three fragments suggested that there were three genes expressed in the endosperm, *Ltk1*, *Ltk2*, and a third, as yet unidentified, *Ltk* gene. Alternatively, the 957 bp fragment might have been a result of partial digestion of the *Ltk1* or *Ltk2* RT-PCR products. If three different transcripts were indeed present, then the 1070 bp RT-PCR product obtained with primers 276 and 207 would be predicted to contain, not two different products, but three. To test the hypothesis that there were three different endosperm transcripts, the 1070 bp RT-PCR products were cloned and as predicted, three different types of clones were isolated. The insert of each clone type was sequenced; on the basis of sequence homology, two types corresponded to *Ltk1* and *Ltk2* were identified (data not shown). The third type, designated as *Ltk3* (GenBank Accession #AF23166, Figure 4-1), showed 98.6% identity at the nucleotide level and 97.5% identity at the amino acid level, when compared with the corresponding sequences of the *Ltk2* cDNA and deduced protein. Based on the DNA sequence of the *Ltk3* RT-PCR product, it could generate a 957 bp *Hae* III fragment as observed in Figure 4-7A (panel 3-D).

To further investigate *Ltk* expression during maize endosperm development, RT-PCR with primers 276 and 207 was performed with total RNA extracted from endosperm at various stages of development and from the maternal, unfertilized ovule tissue. As shown in the top panel of Figure 4-7B, *Ltk* expression was relatively elevated in the maternal ovule tissue, while *Ltk* transcript levels were lower at 10 DAP. From about 15 DAP, *Ltk* transcript levels began to rise, reaching a peak at 20 DAP, and then decreasing as observed at 25 DAP. To determine more specifically which *Ltk* transcripts were being expressed, individual *Ltk* transcript levels were assayed by *Hae* III digestion of the RT-PCR product. The only detectable *Ltk* transcript accumulated in unfertilized maternal ovule tissue was *Ltk1*, as evidenced by the 771 bp *Hae* III fragment (Figure 4-7B). However, the *Ltk* transcripts undergoing temporal modulation in endosperm development were a combination of both *Ltk1* and *Ltk2*, as indicated by the presence of the 771 and 887 *Hae* III fragments, respectively. The absence of the *Ltk3 Hae* III product suggested that the corresponding transcript was of low abundance, as also suggested by the faint band seen in Figure 4-7A (panel 3-D). In comparison, RT-PCR was used to show transcripts levels for two other genes, *Pds* and *Sh*, which are known to be expressed in developing endosperm. As previously demonstrated, *Pds* transcript levels were constant, whereas *Sh* transcript levels increased from 10 to 15 DAP and remained constant thereafter (Li et al., 1996).

4.5 Discussion

The serendipitous isolation of a chimeric *Ltk-Pds* cDNA led to the discovery of a new class of maize receptor-like kinases containing leucine-rich repeats in their extracellular domain. These particular LRR type RLKs were termed LTK proteins. From Southern analysis, it appeared that there were three *Ltk* copies in the genome; this was later confirmed by cloning of three different RT-PCR products. Therefore, LTK proteins are encoded by a small gene family consisting of three members, *Ltk1*, *Ltk2*, and *Ltk3*. From sequences of the isolated *Ltk* cDNAs, it was determined that *Ltk1* and *Ltk2* shared 91.2 % nucleotide homology, while *Ltk2* and *Ltk3* shared 98.6 % homology. However, from RFLP mapping in two different families (T232 X CM37 and CO159 X Tx303) (Burr et al., 1988), only two *Ltk* loci have been detected; one locus is 0.58 cM distal to *Pds1* on chromosome 1S while the other is 1.19 cM proximal to bngl619, an RFLP marker on chromosome 9L (B. Burr, personal communication). From isolation and sequencing of a 5.8 kb genomic DNA fragment containing the 3' end of *Ltk1* and *Pds*, I know that *Ltk1* is the *Ltk* locus linked to *Pds1* and transcribed in the same direction. Furthermore, this close proximity is probably related to the origin of the *Ltk-Pds* chimeric clone, which, based on sequence identity, was actually a fusion of *Ltk1* and *Pds1* (data not shown). Since available sequence for *Ltk2* and *Ltk3* showed 98.6% homology, these loci might represent a recent gene duplication whereby the two genes are still closely linked and cannot be differentiated by RFLP mapping. *Ltk2* and/or *Ltk3* might be located on chromosome 9L.

From the complete ORF available for *Ltk2*, I predicted that it encodes a 79.15 kD protein with eight leucine-rich repeats in an amino terminal extracellular domain and a serine/threonine protein kinase in a carboxyl terminal intracellular domain. The kinase catalytic domain showed homology to other serine/threonine kinases, particularly to tomato Pti1, a soluble serine/threonine protein kinase involved in plant defense (Zhou et al., 1995). Also, the putative extracellular LRR domain showed some homology to Cf-2 proteins, a pair of membrane anchored defense proteins from tomato (Dixon et al., 1996). However, the LRR domain of LTK proteins was unique in number and arrangement of LRRs as compared to any other known LRR-type RLK, implying a unique ligand specificity for the LTK extracellular domains. Despite the distinctive architecture of the LRR domain, the leucine-rich repeats do contain the consensus core sequence thought to be involved in protein-protein interaction and/or ligand binding.

In addition to the unique LRR domain, LTK proteins contained some novel motifs in the extracellular domain. A PEST sequence, thought to be involved in protein degradation, was found downstream of the LRRs. PEST sequences have been found not only in metabolic enzymes but also in proteins involved in gene expression, signal transduction and cell-cycle regulation (Rechsteiner and Rogers, 1996). Clearly, from Northern and RT-PCR analysis, the transcript levels for the various family members were under both tissue-specific and developmental regulation; the PEST sequences may provide an additional level of post-transcriptional control of *Ltk* expression. Nearby the PEST sequence was a PRR, which can play a role in protein structure *vis a vis* ligand binding. Another novel region found in the extracellular domain, was an SRR. Though the function of this serine-rich region is unclear, its proximity to the LRR and PRR may suggest some role in ligand binding. Moreover, this region most distinguished LTK1 from LTK2, suggesting that not only are LTK proteins very different from all other LRR-type RLKs, but may differ from each other regarding ligand specificity and/or affinity. In addition to variation in the SRR domain, other minor differences were observed in the predicted intracellular regions external to the kinase catalytic domain. Intracellular segments external to the putative protein kinase catalytic domain have been proposed to affect substrate specificity (Walker, 1994). Minor residue differences between LTK1 and LTK2 intracellular domains may be interpreted to influence substrate specificity and/or affinity of the corresponding kinases. Subtle amino acid sequence differences between LTK types can have profound effects on extracellular ligand binding and/or intracellular signal transduction.

From Northern analysis, only one size transcript could be detected; however, use of RT-PCR combined with restriction enzyme digestion allowed monitoring of *Ltk* gene-specific expression. Although *Ltk* transcripts were found to be of low abundance, transcript levels were higher in endosperm and embryo when collected at 20 DAP, as compared to leaves and roots of young seedlings. Of the three *Ltk* transcripts, *Ltk1* was the only one expressed in all tissues examined, including roots, leaves, endosperm, embryo, and unfertilized ovules; *Ltk2* and *Ltk3* were only expressed in the endosperm. However, *Ltk3* was difficult to detect routinely, although an *Ltk3* cDNA could be cloned. While *Ltk2* was undetectable in unfertilized ovules, it was detected in developing endosperm, even at 10

DAP. This suggests that fertilization triggers increased *Ltk2* transcript levels in developing endosperm cells. Furthermore, endosperm development was also accompanied by modulation of the *Ltk1* transcript level. However, at 20 DAP, when both *Ltk1* and *Ltk2* were expressed in the endosperm, only *Ltk1* was expressed in the embryo.

Although the LRR and kinase domains of LTK showed similarity to corresponding regions of defense-related proteins, the observation that *Ltk1* and *Ltk2* transcript levels were modulated during endosperm development suggests that LTK proteins may play a role in endosperm development. Other LRR type RLKs have been implicated in control of plant development; in Arabidopsis, ERECTA (ER) is involved in determination of organ shape and vegetative growth, although the biochemical mechanism is unknown (Torii et al., 1996) and Clavata1 (CLA1), expressed in the inflorescence, may regulate differentiation of apical meristem cells (Clark et al., 1997).

In maize, the process of double fertilization establishes initiation of two parallel developmental programs, one leading to a mature embryo and the other leading to the endosperm. As described by Kiesselbach (Kiesselbach, 1949), immediately after fertilization, endosperm nuclei divide mitotically without formation of cell walls; after about two days, the free nuclei continue dividing while forming cell walls only on one side until the endosperm is completely cellular at about 4 DAP; at 5 DAP, cell division occurs throughout the endosperm; at about 6 DAP the endosperm is solidifying and surrounding nucellar tissue is being digested away; also, the basal endosperm cells differentiate to form placental tissue; between 10-20 DAP, most cell division is confined to the outer region of the endosperm, away from the embryo; after about 20 DAP, cell divisions are limited to the periphery several cells deep, where surface cells eventually differentiate into the aleurone layer. It is unclear what signals control the repression of cell division in the endosperm after 10 DAP. Interestingly, the change in *Ltk* gene expression correlated with this temporal and regional control of cell division in developing endosperm. When endosperm cell division becomes more restricted between 10-20 DAP, *Ltk2* and *Ltk1* transcript levels are steadily increasing. Since the LRR domain may interact with polypeptide ligands or cell surface proteins, LTK proteins could transduce signals involved in regulation of cell division. Similarly, LTK1 may have some role in control of cell division in other parts of the plant, as its transcripts are constantly present.

Alternatively, the increased transcript levels of *Ltk1* and *Ltk2* might be significant in regulating later stages of endosperm development associated with establishment of dormancy. After 20 DAP, the kernel proceeds towards the maturation phase and dormancy where only the embryo and aleurone layer of the endosperm develop tolerance to desiccation associated with this stage. More defined analysis of *Ltk* transcript localization in specific cells of developing endosperm and in other tissues, combined with isolation of *Ltk*-specific mutations, will help elucidate the role of each LTK type in maize growth and development. The LTK proteins and their corresponding genes represent new tools for probing signal transduction associated with the intriguing process of endosperm and seed development.

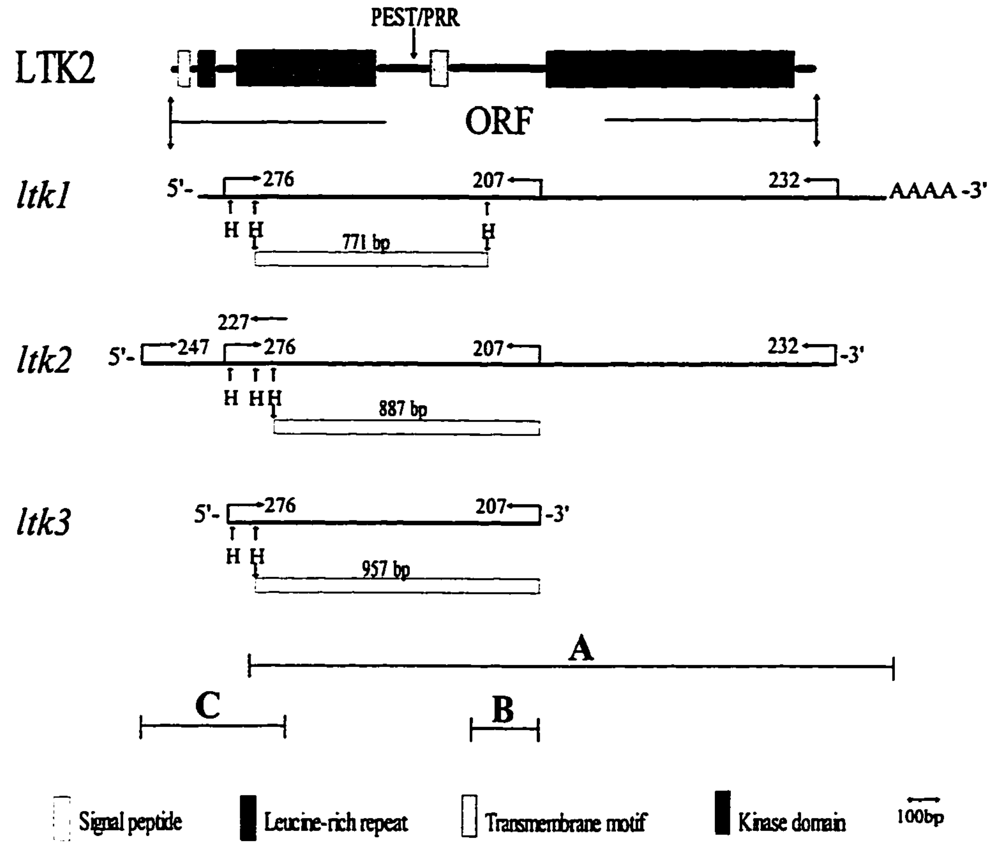


Figure 4-1. LTK protein and cDNA structures. ORF, open reading frame; PRR, proline rich region; PEST, putative protein degradation site. Numbers 207, 227, 232, 247 and 276 are primers used for RT-PCR. The open boxes, above which are DNA sizes in base pairs, represent the largest fragments produced by *Hae* III (H) digestion of the RT-PCR products (generated with primers 276/207). *Hae* III sites are shown only in the region between primers 276 and 207. A, B, and C correspond to regions used as hybridization probes (further detailed in materials and methods).

```
  1 AAATAAGCAAATAATGCCAGAATAATTACCGGATTTTTGCAAAGGCAGGCAAAGCAATTG
 61 ACCTCACCGCACCCGGGTTCTTGTCCCGGCGTTCTGCTGCCAGCTGCTGCTGTGCTAGCT
121 TCCAATGCCGGACGCCGCACACGCTATCCATGAATCCATCCATCCGTCTTGAAAAGCGGC
181 AGCGTGATCTGTAAAAAACAGCGGGTAAAATCAGTCGGATTTGGAGGAGTCGAATAAACA
                                                            →247
241 AGGAAACTCCTCCCGGCCCAGAAAGCACTTCACTTGCAACAGTCCCCGTCAGGGAGGGAG
                                                    →243
301 CATTGCTGTGCGGGATGGCGATAAGGAAGAGCGGAGGCGGTGGCGGGGTGGGGGGTGTGG
           246←
361 CGCGGCGGCTGCTGCTTCTGGCCTGCTGCTGCATTTGGCCTGGTGGACGGATCTTCACTG
421 CTGCGGACACTGACCCCAACGATCGTGANTTGGAAACTTCTGCTTTCATTGTGCTTCTTC
       240←
481 TTCCGTGTAAATTTCTTTATTATGGGATCGCTCTTT
```

Figure 4-2. Putative promoter sequence of *Ltk2*.

(A) The restriction map of genomic clone p531. I, *Eco*R I; V, *Eco*R V and S, *Sal* I. E1-4 represents the *Eco*R I fragments subcloned as p531E1, p531E2, p531E3 and p531E4. Numbers in the box indicate the distance between two restriction enzyme in kb. Arrow indicates direction of transcription. The direction of transcription was determined by partial sequencing of fragments E2 and E1 and comparison with *ltk* cDNA.

(B) The sequence of the *Ltk2* promoter deduced from p531E2. The boxed sequence is the putative TATA box; the putative Cap signal is underlined. The first exon, having an unidentified starting site, ends at nt 444, and contains a putative translation start codon at nt 315-317, shown bolded and underlined. The italicized sequence indicates the 5' end of the first intron; the first two conserved nucleotides are shown in bold. Numbers indicate sequences used for designing RT-PCR or sequencing primers and arrows indicate direction of synthesis.

```
LTK1                                            DPNDLNVLNTL11
                                                |||||||||||
LTK2 MAIRKSGGGGGVGGVARRLLLLACCCIWPGGRIFTAADTDPNDLNVLNTL50
                                                 ----------

LTK1 FTSLNSPGQLTGWQANGGDPCGQSWKGITCSGSGVTKIQLPNLSLTGNLA61
     ||||||||||  | ||||||||||| |||||||||||| |||||||||||
LTK2 FTSLNSPGQLRAWRANGGDPCGQSWQGITCSGSGVTKILLPNLSLTGNLA100
     - 1 --------                      -------- 2 -----
LTK1 YNMNNLGSLVELDMSQNNLGGGGQVQYNLPNMKLEKLNLAGNQFGGNLPY111
     |||||||||||||.||||||||||.||||||.||||||||||||||||||
LTK2 YNMNNLGSLVELDLSQNNLGGGGQIQYNLPNVKLEKLNLAGNQFGGNLPY150
     - -------- 3 --------- --------- 4 ---------- -
LTK1 SISTMPNLKYLNLNHNQLQGNISDVFSNLYSLSELDLSFNSLTGDLPQSF161
     ||||||||||||||||||||||.|||||||||||||||||||||||||||
LTK2 SISTMPNLKYLNLNHNQLQGNITDVFSNLYSLSELDLSFNSLTGDLPQSF200
     ------ 5 ------------- -------- 6 ---------- --
LTK1 TGLSSLKKVYLQNNQFTGNINVLANLPLETLNVANNHFTGWIPSQLKKIN211
     |||||||..||||||||| |||||||||||||| ||||||||||||| ||
LTK2 TGLSSLKRMYLQNNQFTGYINVLANLPLETLNVGNNHFTGWIPSQLKEIN250
     ---- 7 ---------- --------- 8 -------
LTK1 SLQTDGNSWSTGPAPPPPPYTAPPPPPNHWNADG--SSSSSSSGGRSGIG259
      ||||||||||||||||||||||||| |||| |     ||||||| |||
LTK2 NLQTDGNSWSTGPAPPPPPYTAPPPP-NHWNGAGQNDDGSSSSGGRPGIG299
                                    ===============

LTK1 GGGVAGIIISLLVVGSVVAFLVIKRRKRKAAMKEHFEQHQPFTSFPSNEV399
     ||||||||||||.||||||| .||||| || | | |||||||||||||||
LTK2 GGGVAGIIISLLIVGSVVAFFLIKRRKHKAIMEEQFEQHQPFTSFPSNEV349

LTK1 KDMKPVCEATTVDIESLASPASVNLKPPPKIERNKSFDDDDDFSNKLVAK359
      ||||. |.|||||||||||||.||||||||| |||||||||||||  | 
LTK2 NDMKPIYESTTVDIESLASPASINLKPPPKIEQNKSFDDDDDFSNKTAAN399

LTK1 KSNITPINATVYSVADLQMATDSFSFDNLVGEGTFGRVYRAQFNDGKVLA409
     .|||||. |||||||||||||||||||||||||||||||||||| |||||
LTK2 RSNITPMKATVYSVADLQMATDSFSFDNLVGEGTFGRVYRAQFNGGKVLA449
                             --------I---------  ------
LTK1 IKKLDSTVMPFQSSDDFAELVSNISKLHHPNLNELVGYCMEHGQHLLVYD459
     ||||| ||||||||||||||||||||||||||||||||||||||||||||
LTK2 IKKLDITVMPFQSSDDFAELVSNISKLHHPNLNELVGYCMEHGQHLLVYD499
     II---               ---III---  -------IV------  -----
LTK1 FHRNGSLHDLLHLSDEYSKPLSWNTRIKIALGSARALEYLHEICSPSIIH509
     |||||||||| |||||||| ||||.||| |||||||||||||||||||||
LTK2 FHRNGSLHDLRHLSDEYSKALSWNSRIKFALGSARALEYLHEICSPSIIH549
     ---V--------        ----------VIa-----------  ---
LTK1 KNFKSSNILLDSEFNPHLSDAGLASFIPDAEFQAAEQSAGYTAPEVDMTG559
     ||||.|||||||||||||||.||||||| |||||||||||||||||||||
LTK2 KNFKASNILLDSEFNPHLSDTGLASFIPGAEFQAAEQSAGYTAPEVDMTG599
     --VIb------     ---VII---            ------VIII----
LTK1 QYTFKSDVYSFGVVMLELLTGRRPFDSSRPRSEQSLVRWATPQLHDIDAL609
     ||| ||||||||||||||||||||||||||||||||||||.|||||||||
LTK2 QYTLKSDVYSFGVVMLELLTGRRPFDSSRPRSEQSLVRWAAPQLHDIDAL649
       ---------IX---------                -----------
LTK1 DRMVDPALKGLYPAKSLSRFADVLALCVQPEPEFRPPMSEVVQALVRLVQ659
     ||||||||||||||||||||||||||||||||||||||||||||||||||
LTK2 DRMVDPALKGLYPAKSLSRFADVLALCVQPEPEFRPPMSEVVQALVRLVQ699
     -X---------                ---------XI---------
LTK1 RANMTKRMLDG-DTSRRPDDLDQDFI684
     ||||||||||| |||| ||| || ||
LTK2 RANMTKRMLDGGDTSRGPDDQDQYFI725
```

Figure 4-3. Protein sequence comparison of LTK1 and LTK2. The leucine-rich repeats (LRRs) are underlined and numbered 1 to 8; conserved residues are shown in bold. Residues contained within the putative signal peptide are bold and within a box; residues comprising the transmembrane domain are bold. Residues within the PEST motif are italicized, while those in the PRR are italicized and bold. A double underline indicates residues in the SRR. The eleven subdomains of the protein kinase catalytic site are underlined and numbered from I to XI; conserved residues are shown in bold. Horizontal lines and dots indicate identical and similar residues, respectively. Dashes in the amino acid sequence are gaps inserted to produce an optimal alignment.

A

```
LTK1    2  PNDLNVLNTLFTSLNSPGQLTG   24  ]1
LTK2   41  PNDLNVLNTLFTSLNSPGQLRA   63
LTK1   44  GVTKIQLPNLSLTGNLAYN      63  ]2
LTK2   83  GVTKILLPNLSLTGNLAYN     102
LTK1   64  MNNLGSLVELDMSQNNLGGGGQV 86  ]3
LTK2  103  MNNLGSLVELDLSQNNLGGGGQI125
LTK1   87QYNLPNMKLEKLNLAGNQFGGNLP 110  ]4
LTK2  126QYNLPNVKLEKLNLAGNQFGGNLP 149
LTK1  111YSISTMPNLKYLNLNHNQLQGNIS 134  ]5
LTK2  150YSISTMPNLKYLNLNHNQLQGNIT 173
LTK1  135DVFSNLYSLSELDLSFNSLTGDLP 158  ]6
LTK2  174DVFSNLYSLSELDLSFNSLTGDLP 197
LTK1  159QSFTGLSSLKKVYLQNNQFTGNIN 182  ]7
LTK2  198QSFTGLSSLKRMYLQNNQFTGYIN 221
LTK1  183  VLANLPLETLNVANNHFTGWIP 204  ]8
LTK2  222  VLANLPLETLNVGNNHFTGWIP 243
```

B

Protein	Consensus Sequence	No. of Repeats
LTK1&2	xxaxxaxxLxxLxLxxNxaxGxxx	8
XA21	xxaxxLxxLxxLxLxxNxLxGxaP	23
CLV1	xxaxxaxxLxxaxxxxNxaxGxIP	23
TMK1	xxLxxLxxLxxLxLxxNxaxGxaP	11
ER	xxLGxLxxLxxLxLxxNxLxGxIP	20
RLK5	xxLxxLxxLxxLxLxxNxLSGxIP	21
TMKL1	xxLxxxxxLxSLxLxxNxLSGxIP	7
PRK1	LxxLxxaxxLxLxNNxxxGxIP	5
SERK	xxLxxLxxLxxLxLxNNxLSGPIP	5

←→ β Sheet

C

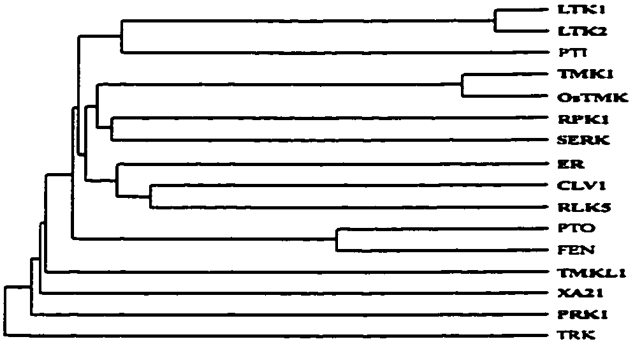


Figure 4-4. Comparison of LTK structural motifs with those of other RLKs. (A) Comparison of the 8 LRRs in LTK1 and LTK2; conserved residues (either identical or similar) are shown in bold; sequences are flanked by residue numbers for first and last residue. (B) Comparison of LRR consensus sequences for LRR type RLKs. a, hydrophobic residues; x, any residue. (C) Dendrogram of the alignment of the protein kinase domain of LRR type RLKs (bold), three cytosolic protein kinases and TRK (underlined). The GenBank accession number for each protein is: from *Arabidopsis*, TMK1 (L00670), TMKL1 (X72863), RLK5 (M84660), ER (U47029), CLV1 (U96879), RPK1 (U55875); from carrot, SERK (U93048); from maize, LTK1 (AF023164), LTK2 (AF023165); from rice, XA21 (U37133), OsTMK (Y07748); from tomato, PTI (U28007), PTO (U02271), FEN (U13923); from *Petunia inflata*, PRK1 (L27341) and from human, TRK (M23102). The program used to generate this dendrogram is CLUSTAL developed by Higgins and Sharp (Higgins and Sharp, 1988). The length horizontal lines indicate the relative aa difference between two protein sequences.

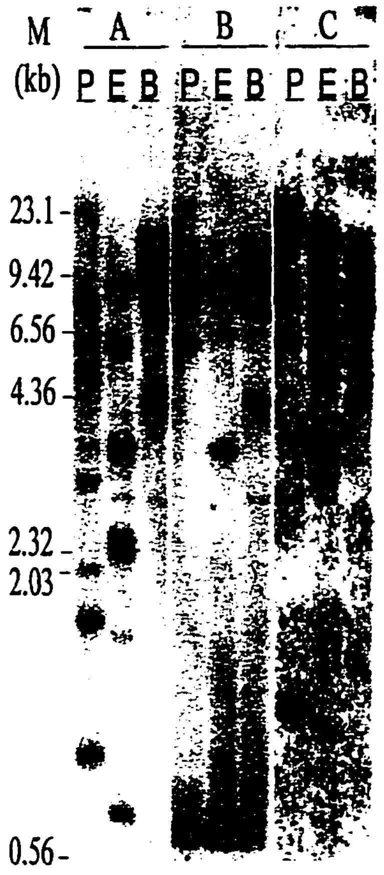


Figure 4-5. Southern analysis of the *Ltk* gene family in maize. A, B, and C are probes shown in Figure 4-1. P, *Pvu* II; E, *Eco*R I; B, *Bam*H I; M, molecular weight markers.

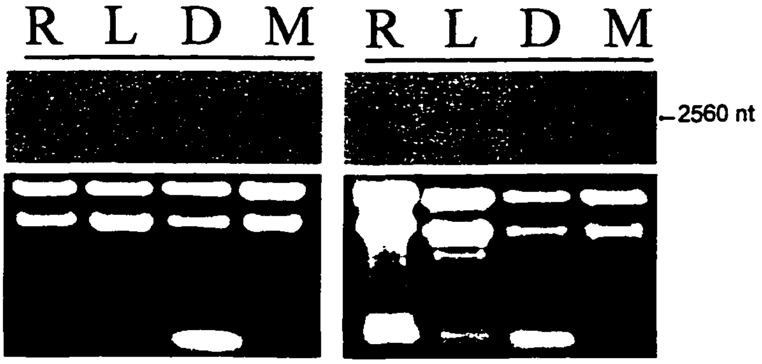


Figure 4-6. Northern analysis of *Ltk* expression. Top panels show transcripts hybridized by probe A (Figure 4-1). Bottom panels show ethidium bromide stained RNA electrophoretic gels used for hybridizations. Ten micrograms of total RNA were loaded per lane, except for R and L lanes in right panel, which contain 40 μg total RNA. R, roots; L, leaf; D, 20 DAP endosperm; M, 20 DAP embryo.

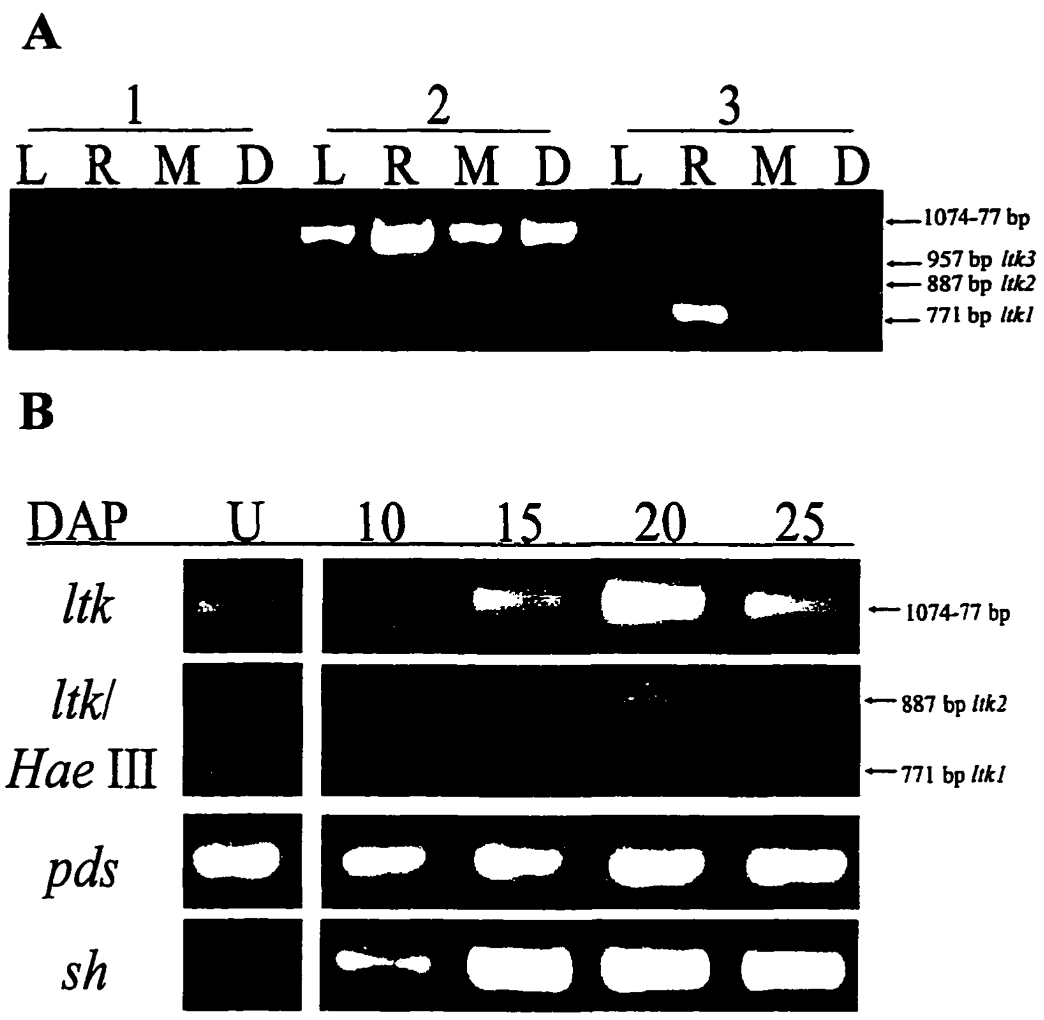


Figure 4-7. RT-PCR analysis of *Ltk* expression. (A) *Ltk* transcripts amplified with primers 276 and 207; 1, without reverse transcriptase; 2, with reverse transcriptase; or 3, with reverse transcriptase but products digested by *Hae* III. L, leaf; R, roots; M, embryo (20 DAP); D, endosperm (20 DAP). (B) Total RNA from unfertilized ovule (U) and developing endosperm at indicated DAP. Panels top to bottom show RT-PCR products for amplification of *Ltk* transcripts; the same products digested by *Hae* III; RT-PCR products for amplification of *Pds* and *Sh* transcripts, respectively.

Chapter 5: Discussion and perspectives.

Using molecular biology techniques and a well-known classic genetic material, maize, I have isolated and characterized two kinds of genes unique to plants. The first is *Pds1*, which encodes phytoene desaturase, with proven function in carotenoid biosynthesis. The second is *Ltk*, which represents a family of genes encoding novel receptor-like protein kinases, whose predicted function is in signal interpretation and transduction. Both the *Ltk1* and *Pds1* genes mapped near the *vp5* locus and are less than 4 kb apart. In this chapter, I will discuss the importance of cloning these genes and I will contemplate future studies of them.

5.1 Phytoene Desaturase

The study of carotenoid biosynthesis is such an active field, that in recent years dozens of genes encoding enzymes of the pathway have been cloned from various organisms. Though several studies have shown that *Psy* expression correlates with carotenoid content and that this is the rate-limiting step of the pathway in plants, the mechanism of how *Psy* is regulated remains unclear. In contrast, *Pds* expression appeared to be constitutive in some of these studies. In maize, a recessive *Pds* mutation blocks carotenoid biosynthesis in most tissues, while *Psy* mutants are impaired in endosperm carotenoid biosynthesis. Therefore, the comparative analysis of *Psy* and *Pds* promoters will shed light on how these two genes are differentially regulated. Cloning *Pds* cDNA and consequently the *Pds* promoter of maize places a foundation for this analysis since *Y1*, which encodes *Psy* of maize, has already been cloned.

Though *Pds* may not be the regulating point for the entire carotenoid biosynthetic pathway in maize, the desaturation of phytoene is an important and complex step, especially in higher plants including maize. Experiments with *Arabidopsis* mutants have shown that quinone biosynthesis is involved because quinones are electron acceptors and donors, which can transfer the electrons generated by dehydrogenation of phytoene. Mutants blocking phytoene desaturation are not necessarily mutants of *Pds* (Norris et al., 1995). Four non-allelic loci (*vp2*, *vp5*, *w3* and *vp9*) may not only encode PDS and ZDS, enzymes for carotenoid biosynthesis, but may also encode regulatory proteins or enzymes related to quinone biosynthesis. However, RT-PCR analysis of *Pds* expression in *vp2*, *vp5* and *w3* showed that it differed from that of *vp5*, which decreases the *Pds* expression at the RNA level; *vp2* and *w3* had no effect on steady-state levels of *Pds* transcripts. This result suggested that *vp2* and *w3* did not encode regulatory proteins controlling *Pds* expression at the RNA level. Preliminary results from Luo and Wurtzel have shown that *vp2*, as in *Arabidopsis*, might encode one of the enzymes in the quinone biosynthetic pathway (Luo and Wurtzel, pers. comm. and unpublished). Obviously, cloning of *Pds* cDNA advanced the understanding of phytoene desaturation mutants (*vp2*, *vp5*, *w3* and *vp9*) in maize.

Though cytobiochemical localization has shown that PDS is primarily located on the plastid membrane as previously suggested, it has also been found in the plastid stroma. Recently published work by Al-Babili *et al.* suggested that PDS had two forms, one bound to membrane as an active form and a second one remaining in the stroma as an inactive form, complexed with Hsp70 and chaperonin (Al-Babili et al., 1996; Bonk et al., 1996). This observation was made in the daffodil chromoplast and whether it is universal in all different types of plastids is still unknown. Cloned *Pds* cDNA can be used to produce maize PDS proteins in *E. coli* and then generate an antiserum. The PDS antibody will allow localization of PDS in the plastids, especially in the amyloplast of endosperm. The use of maize PDS antiserum in protein analysis will provide information on how cytosol-synthesized proteins are regulated in plastids. Compared with PDS isolated from prokaryotic cells, maize PDS, as PDS of other higher plants, contains a relatively long transit peptide (96 aa) (see Chapter 2). Understanding the function of the PDS transit peptide in assisting its plastid membrane integration will benefit the functional expression of PDS in the endosperms of transgenic plants and of other proteins targeted to the plastid membrane. The PDS antibody can be also used to analyze *Pds* expression in the developing endosperm at the protein level. Besides generating antiserum of PDS, the cDNA clone of *Pds* can be also employed to study the enzyme active site of PDS by mutagenesis.

One long-term goal of the study of carotenoid biosynthesis is to genetically engineer the rice genome to produce enzymes and proteins that can synthesize carotenoids in its endosperm. Like other plants, rice can produce carotenoids in its leaf though it cannot accumulate carotenoids or their precursors in its endosperm. Why carotenoid accumulation is diminished in the rice endosperm is still unknown. One possibility is that the whole carotenoid biosynthetic pathway is shut off in the rice endosperm. The other one is that the pathway is expressed in the endosperm, but the products or intermediates are not abundant enough to be detected. Recently results showed that both *Psy* and *Pds* were expressed in the endosperm at the RNA level indicated that the carotenoid biosynthetic pathway was turned on at least at the RNA level (Wurtzel et al., 1996). However, over-expression of daffodil PSY in the rice endosperm through transgenic techniques only produced an accumulation of phytoene without any carotenoid (Burkhardt et al., 1997). This result suggests that the carotenoid biosynthetic pathway is not functional in rice endosperm.

If genes of carotenoid biosynthetic enzymes are transcribed in the rice endosperm with no carotenoid detected, either the translation of *Psy* and *Pds* transcripts is blocked or PSY and PDS are translated with an extremely low enzyme activity. However, a simple possibility is that rice PSY and PDS cannot function well in the rice amyloplast, although they are active in the maize amyloplasts, which accumulate carotenoids. Comparative analyses of maize and rice *Pds* genes will advance the understanding of how *Pds* is expressed in the maize endosperm and why PDS cannot function in the rice endosperm.

Hopefully, the conclusion of these analyses will provide a unique means to overcome carotenoid deficiency of rice endosperm by transferring maize *Psy* and *Pds* genes into the rice genome.

5.2 Leucine-rich Repeat Transmembrane Protein Kinase

Cloning of the *Ltk* gene was serendipitous, since the original project was to study gene regulation of carotenoid biosynthesis in the maize endosperm. The finding of the close linkage of *Pds1* and *Ltk1* and of the possibility that these genes are differentially spliced from the same transcript (see Chapter 3) implies that their encoded proteins could be functionally related or that LTK1 could play a role in controlling PDS1 plastid import by phosphorylation of the PDS transit sequence. The cloned *Ltk* cDNAs can be transcribed and translated in an *in vitro* system and the function of LTK on phosphorylation of the PDS transit sequence can be tested. However, the function of LTK proteins proposed in Chapter 3 suggests that LTK may play a role in the development of maize. Also, it is hard to imagine that a plastid imported protein (PDS) has to be phosphorylated by a plasma membrane receptor (LTK).

Although genes encoding several LRR containing protein kinases have been cloned from higher plants, the structures of these proteins are different, especially at their N-terminal LRR region (Figure 5-1). This observation indicates that they can bind to different signal ligands and may induce different signal transduction pathways. As stated in Chapter 3, *Ltk* encoded proteins might play an important role in maize endosperm development as well as in whole plant growth. I certainly believe that cloning of *Ltk* cDNAs will benefit the study of development not only of maize but also of other plants because homologous genes will be found eventually. In fact, when rice DNA was probed by the maize *Ltk* cDNA, even at high stringency, it showed that an *Ltk* homolog was present in the rice genome (Figure 3-5 and data not shown). Consistent with this finding, database searching with the *Ltk* nucleotide sequence showed 89% homology to a rice EST clone sequence. These data suggest that *Ltk* not only is present in plants other than maize but is also well conserved in different species. However, when genomic DNA of two dicot plants (tomato and *Arabidopsis*) were probed by maize *Ltk* gene, no fragments hybridized (data not shown). This preliminary result implied that *Ltk* might be a gene unique to monocot plants or that the homology between monocots and dicots was so low that under stringent conditions the monocot *Ltk* could not hybridize to its dicot homolog. If this is correct, one could hypothesize that *Ltk* might play a significant role in controlling plants developing into monocot but not dicot and *vice versa*, into dicot but not monocot.

Sequence comparison of *Ltk3* to *Ltk1* or *Ltk2* showed that the cloned *Ltk3* region had a higher homology to *Ltk2* than to *Ltk1* (Chapter 3). However, the differential nucleotides of *Ltk3* were primarily identical either to *Ltk1* or to *Ltk2*. As shown in Figure 5-2, out of 85 different nucleotides among the three *Ltk* cDNAs, *Ltk3* has 70 nt different from *Ltk1* (number 2), 11 nt different from *Ltk2* (number 1) and only 3 nt different from both *Ltk1* and *Ltk2* (number 3). Furthermore, the 11 nt difference between *Ltk2* and the other *Ltk* genes clusters at the 5' ends, whereas the 70 nt difference between *Ltk1* and the other two genes was mostly at the 3' end. These results suggest that *Ltk3* might be a duplicated recombinant gene of *Ltk1* and *Ltk2* and that the enzyme encoded by *Ltk3*, LTK3, may have an identical or similar LRR domain to that of LTK1 but an identical or similar kinase domain to that of LTK2 (Figure 5-3). How and why *Ltk3* evolved from *Ltk1* and *Ltk2* will be valuable in studying gene origin.

To dissect the function of LTK proteins, lines carrying a *Mutator* insertion in each of the *Ltk* genes will be needed. In cooperation with Pioneer Hi-bred International, Inc. using the TUSC (Trait Utility System for Corn) system, several primers were designed according to the *Ltk1* sequence and tested in PCR reaction using maize B73 genomic DNA as template. Those primers that could generate PCR products were verified by Southern analysis. They were then used in combination with a primer designed based on a *Mutator* sequence to screen *Mutator*-containing maize lines by PCR. Preliminary results of the screening showed that several lines might contain a *Mutator* inserted into *Ltk* genes. These lines will be screened in the field for both *Mutator*-inserted *Ltk* genes and for mutant phenotypes. Using this technique, mutants of *Ltk* can be obtained and subsequently analysis of *Ltk* expression at both RNA and protein levels will provide information on the function and localization of *Ltk* products. Since LTKs are most likely involved in plant development as suggested in Chapter 3, phenotypes of aberrant growth are expected.

Signal transduction in the animal kingdom has been studied extensively through cloning genes for both receptor and ligand. In higher plants, however, the understanding of external signal perception is limited. Until now, no paired ligand and receptor genes have been cloned from plants, although dozens of receptor genes and several proteins or small peptide ligand genes have been cloned. The newly cloned *Ltk* cDNAs will provide another opportunity to isolate the protein ligands and to clone the ligand genes because their LRR binding domains will bind specifically to polypeptide ligands. A yeast two-hybrid system with the LRR domain as a probe may be used to identfy a polypeptide ligand expressed and secreted from maize cells. The same technique can also be used to clone genes encoding protein substrates of the LTK protein kinase domains. In these cases, the protein kinase domain will be utilized as a probe.

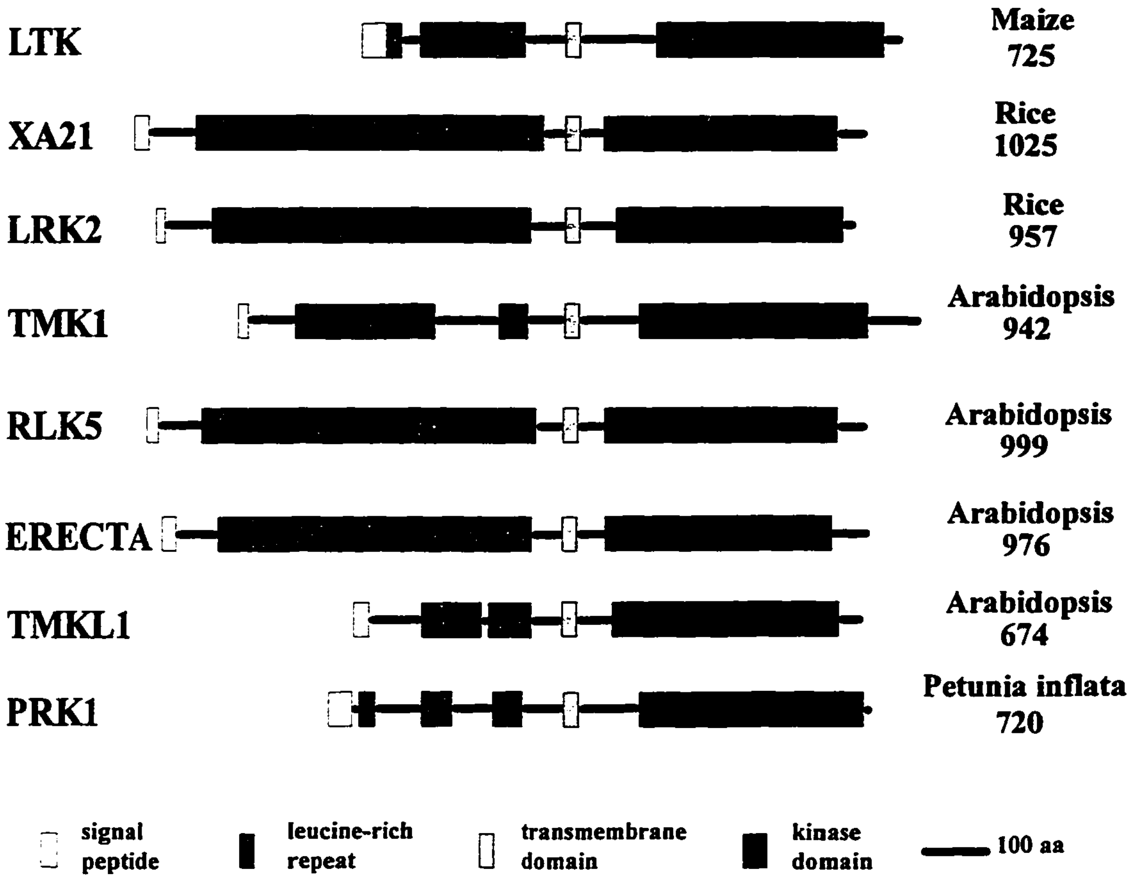


Figure 5-1. A diagram of several LRR-containing receptor-like protein kinases.

```
Ltk1 TCAGAGCTTGGCAGGCGAATGGGGGTGATCCTTGTGGCCAGTCATGGAAGGGCATTACTTGCTCAGGATCAGGGGTCACT 60
Ltk2 TCAGAGCTTGGCAGGCGAATGGCGGTGATCCTTGTGGCCAGTCATGGCAGGGCATTACTTGCTCAGGATCAGGGGTCACT 286
Ltk3 TCAGAGCTTGGCAGGCGAATGGGGGTGATCCTTGTGGCCAGTCATGGAAGGGCATTACTTGCTCAGGATCAGGGGTCACT 80
     **********************1************************1********************************

Ltk1 AAAATCCAATTACCTAACTTGTCACTCACTGGAAATTTGGCCTACAACATGAATAACTTGGGTTCATTAGTTGAGCTTGA 140
Ltk2 AAAATCCTATTACCAAACTTGTCACTCACCGGAAATCTGGCCTACAACATGAATAACTTGGGTTCATTAGTTGAGCTTGA 366
Ltk3 AAAATCCAATTACCAAACTTGTCACTCGCTGGAAATTTGGCCTACAACATGAATAACTTGGGTTCATTAGTTGAGCTTGA 160
     *******1******2*************3*1******1*******************************************

Ltk1 CATGAGCCAAAATAACCTTGGCGGTGGAGGTCAAGTACAGTACAATCTTCCCAATATGAAGCTTGAGAAACTCAATCTTG 220
Ltk2 CTTGAGCCAAAATAACCTTGGTGGTGGAGGCCAAATACAGTACAATCTTCCCAATGTGAAGCTTGAGAAACTCAATCTTG 346
Ltk3 CATGAGCCAAAATAACCTTGGCGGTGGAGGTCAAATACAGTACAATCTTCCCAATATGAAGCTTGAGAAACTCAATCTTG 240
     *1*******************1********1***2********************1************************

Ltk1 CAGGAAATCAATTTGGTGGAAATTTACCCTACTCAATTTCGACGATGCCTAATCTTAAGTATTTAAACCTTAATCATAAC 300
Ltk2 CAGGAAATCAATTTGGTGGAAATTTACCCTACTCAATTTCGACGATGCCTAATCTTAAGTATTTAAACCTTAATCATAAC 426
Ltk3 CAGGAAATCAATTTGGTGGAAATTTACCCTACTCAATTTCGACGATGCCTAATCTTAAGTATTTAAACCTTAATCATAAC 320
     ********************************************************************************

Ltk1 CAACTACAAGGAAACATCAGTGATGTATTTTCCAACCTTTACAGTTTGTCAGAACTGGATCTGTCCTTTAATTCCCTTAC 380
Ltk2 CAACTACAAGGAAACATCACTGATGTATTTTCCAACCTTTATAGTTTGTCAGAGCTGGACCTCTCCTTTAATTCTCTTAC 506
Ltk3 CAACTACAAGGAAACATCACTGATGTATTTTCCAACCTTTATAGTTTGTCAGAGCTGGATCTCTCCTTTAATTCTCTTAC 400
     *******************2*********************2***********2*****1**2***********2*****

Ltk1 TGGTGATCTACCACAAAGTTTCACTGGATTGTCAAGCCTGAAAAAAGTGTATTTGCAGAACAACCAATTTACTGGTAATA 460
Ltk2 TGGTGATCTACCACAAAGTTTCACTGGATTGTCAAGCCTGAAAAGAATGCATTTGCAGAACAACCAATTTACTGGTTATA 586
Ltk3 TGGTGATCTACCACAAGGTTTCACTGGATTGTCAAGCCTGAAAAGAATGTATTTGCAGAACAACCAATTTACTAGTTATA 480
     ****************3***************************2*2**1**************************22**

Ltk1 TCAATGTCTTAGCCAATCTCCCCCTTGAAACTTTGAATGTTGCGAACAACCATTTCACTGGTTGGATTCCTAGTCAGCTT 540
Ltk2 TCAATGTCTTAGCCAACCTTCCCCTTGAAACCCTGAATGTTGGAAACAACCATTTCACTGGTTGGATTCCCAGTCAGCTT 666
Ltk3 TCAATGTCTTAGCCAACCTTCCCCTTGAAACCCTGAATGTTGGAAACAACCATTTCACTGGTTGGATTCCCAGTCAGCTT 560
     ****************2**2***********22*********22**************************2*********

Ltk1 AAGAAGATAAACAGTCTACAGACTGATGGAAATTCTTGGAGCACAGGACCAGCGCCACCTCCACCTCCATACACAGCGCC 620
Ltk2 AAGAAGATAAACAATCTACAGACCGACGGAAATTCTTGGAGCACAGGACCAGCGCCACCTCCACCTCCATATACAGCG-- 744
Ltk3 AAGAAGATAAACAATCTACAGACCGACGGAAATTCTTGGAGCACAGGACCAGCGCCACCTCCACCTCCATATACAGCG-- 638
     *************2*********2**2********************************************2******22

Ltk1 GCCTCCTCCTCCAAACCATTGGAATG--------CTGATGGTTCATCAAGTTCATCAAGTTCTGGTGGAAGATCTGGGAT 692
Ltk2 -CCTCCTCCTCCAAACCATTGGAATGGCGCGGGTCAGAATGATGAT--GGTTCATCAAGTTCTGGTGGAAGACCTGGGAT 821
Ltk3 -CCTCCTCCTCCAAACCATTGGAATGGCGCGGGTCAGAATGGTGAT--GGTTCATCAAGTTCTGGTGGAAGACCTGGGAT 715
     2*************************22222222*2**22*3*2**222***********************2*******

Ltk1 AGGTGGTGGAGGTGCAGCAGGAATCATTATATCGTTGCTGGTTGTTGGATCAGTTGTTGCATTTCTTGTGATCAAAAGAA 772
Ltk2 AGGTGGTGGAGGTGTAGCAGGAATCATTATATCATTGCTGATTGTTGGATCAGTTGTTGCATTTTTTCTAATCAAAAGAA 901
Ltk3 AGGTGGTGGAGGTGTAGCAGGAATCATTATATCATTGCTGATTGTTGGATCAGTTGTTGCATTTTTTCTAATCAAAAGAA 795
     **************2******************2******2***********************2**2*2**********

Ltk1 GAAAACGCAAAGCTGCTATGAAAGAACATTTTGAACAGCACCAGCCATTCACTTCCTTCCCTTCAAATGAAGTTAAAGAC 852
Ltk2 GAAAACACAAAGCTATTATGGAAGAACAATTTGAACAGCATCAGCCGTTCACTTCCTTCCCTTCAAATGAAGTTAACGAC 981
Ltk3 GAAAACACAAAGCTATTATGGAAGAACAATTTGAACAGCATCAGCCGTTCACTTCCTTCCCTTCAAATGAAGTTAACGAC 875
     ******2*******22****2*******2***********2*****2*****************************2***

Ltk1 ATGAAGCCTGTCTACGAGGCCACCACAGTAGACATAGAGTCTTTGGCTTCACCTGCTTCAGTTAATCTGAAACCACCTCC 932
Ltk2 ATGAAGCCTATCTATGAGTCCACCACAGTAGACATAGAGTCTTTGGCTTCACCTGCTTCGATTAATCTGAAACCACCCCC 1161
Ltk3 ATGAAGCCTATCTATGAGTCCACCACAGTAGACATAGAGTCTTTGGCTTCACCTGCTTCGATTAATCTGAAACCACCCCC 955
     *********2****2***2****************************************22****************2**

Ltk1 GAAGATAGAACGAAACAAATCATTTGATGATGACGATGATTTTTCAAACAAGCTTGTTGCAAAGAAAAGTAATATAACAC 1012
Ltk2 GAAGATAGAACAGAACAAATCATTTGATGATGATGATGATTTTTCAAACAAGACTGCTGCAAATAGAAGTAATATAACAC 1241
Ltk3 GAAGATAGAACAGAACAAATCATTTGATGATGATGATGATTTTTCAAACAAGACTGCTGCAAATAGAAGTAATATAACAC 1035
     ***********22********************2******************22**2******2*2**************

Ltk1 CTATAAATGCAACTGTTTATTCAGTTGCAGATCTACAGATGG 1054
Ltk2 CTATGAAGGCAACTGTTTATTCAGTTGCAGATCTACAGATGG 1283
Ltk3 CTATGAAGGCAACTGTTTATTCAGTTGCAGATCTACAGATGG 1077
     ****2**2**********************************
```

Figure 5-2. Comparison of the nucleotide sequence of *Ltk1*, *Ltk2* and *Ltk3*. Identical nucleotides are indicated with asterisks. Number 1 indicates *Ltk3* has identical nucleotides to *Ltk1* but not *Ltk2*; Number 2 indicates *Ltk3* has identical nucleotides to *Ltk2* but not *Ltk1*; Number 3 indicates *Ltk3* has unique nucleotides different from *Ltk1* or *Ltk2*. Sequence number reflects the number in Appendix 4.

```
LTK1 RAWRANGGDPCGQSWKGITCSGSGVTKIQLPNLSLTGNLAYNMNNLGSLV 71
LTK2 RAWRANGGDPCGQSWQGITCSGSGVTKILLPNLSLTGNLAYNMNNLGSLV 110
LTK3 RAWRANGGDPCGQSWKGITCSGSGVTKIQLPNLSLAGNLAYNMNNLGSLV 50
     ***************1************1******3**************
     1--                    -------- 2 ------ ---------
LTK1 ELDMSQNNLGGGGQVQYNLPNMKLEKLNLAGNQFGGNLPYSISTMPNLKY 121
LTK2 ELDLSQNNLGGGGQIQYNLPNVKLEKLNLAGNQFGGNLPYSISTMPNLKY 160
LTK3 ELDMSQNNLGGGGQIQYNLPNMKLEKLNLAGNQFGGNLPYSISTMPNLKY 100
     ***1**********2******1****************************
     - 3 ---------- --------- 4 ----------- ------- 5 -
LTK1 LNLNHNQLQGNISDVFSNLYSLSELDLSFNSLTGDLPQSFTGLSSLKKVY 171
LTK2 LNLNHNQLQGNITDVFSNLYSLSELDLSFNSLTGDLPQSFTGLSSLKRMH 210
LTK3 LNLNHNQLQGNITDVFSNLYSLSELDLSFNSLTGDLPQGFTGLSSLKRMY 150
     ************2*************************3********221
     ------------- ---------- 6 --------- ------- 7 ---
LTK1 LQNNQFTGNINVLANLPLETLNVANNHFTGWIPSQLKKINSLQTDGNSWS 221
LTK2 LQNNQFTGYINVLANLPLETLNVGNNHFTGWIPSQLKKINNLQTDGNSWS 260
LTK3 LQNNQFTSYINVLANLPLETLNVGNNHFTGWIPSQLKKINNLQTDGNSWS 200
     *******32**************2******************2*******
     ---------- ---------- 8 --------
LTK1 TGPAPPPPPYTAPPPPPNHWNADG--SSSSSSSGGRSGIGGGGAAGIIIS 269
LTK2 TGPAPPPPPYTAPPPP-NHWNGAGQNDDGSSSSGGRPGIGGGGVAGIIIS 309
LTK3 TGPAPPPPPYTAPPPP-NHWNGAGQNGDGSSSSGGRPGIGGGGVAGIIIS 249
     ****************2****22*22222*******2******2******
                          =================
LTK1 LLVVGSVVAFLVIKRRKRKAAMKEHFEQHQPFTSFPSNEVKDMKPVYEAT 319
LTK2 LLIVGSVVAFFLIKRRKHKAIMEEQFEQHQPFTSFPSNEVNDMKPIYEST 359
LTK3 LLIVGSVVAFFLIKRRKHKAIMEEQFEQHQPFTSFPSNEVNDMKPIYEST 299
     **2*******22*****2**2*2*2***************2****2**2*

LTK1 TVDIESLASPASVNLKPPPKIERNKSFDDDDDFSNKLVAKKSNITPINAT 369
LTK2 TVDIESLASPASINLKPPPKIEQNKSFDDDDDFSNKTAANRSNITPMKAT 419
LTK3 TVDIESLASPASINLKPPPKIEQNKSFDDDDDFSNKTAANRSNITPMKAT 349
     ************2*********2*************22*22*****22**

LTK1 VYSVADLQM   378
LTK2 VYSVADLQM   328
LTK3 VYSVADLQM   358
     *********
```

Figure 5-3. Comparison of the protein sequence of LTK1, LTK2 and LTK3. Identical amino acid residues are marked with asterisks. Number 1, 2 and 3 indicates LTK3 has identical aa to LTK1, LTK2 and unique residue respectively. The leucine-rich repeats (LRRs) are underlined and numbered 1 to 8 as in Figure 4-3. Residues comprising the transmembrane domain are bold. Residues within the PEST motif are italicized, while those in the PRR are italicized and bold. A double underline indicates residues in the SRR. Sequence number reflects the number in Figure 4-3.

Appendix

Appendix 1: List of plasmid clones.

Clone Name	Brief Description	Insert Size (kb)	Chapter Number	GenBank Accession Number
pACCRT-EB	*Erwinia CrtE* and *CrtB*, accumulates phytoene		2, 3	
pMPDS3	Maize *Pds* cDNA clone 3	2.0	2, 3	U37285
pMPDSd3-33	Maize *Pds* cDNA clone 3 in frame with lacZ	2.0	2	
pMPDS 1	Maize *Pds* cDNA clone 1	1.1	3	
pMPDS 2	Maize *Pds* cDNA clone 2, identical to clone 10, a chimeric cDNA containing *Ltk1* and *Pds*	2.0	3	
pMPDS 4	Maize *Pds* cDNA clone 4	0.7	3	
pMPDS 6	Maize *Pds* cDNA clone 6	1.0	3	
pMPDS 7	Maize *Pds* cDNA clone 7, encoding PDS andTCTP-like protein.	3.1	3	
pMPDS 10	Maize *Pds* cDNA clone 10, identical to clone 2, a chimeric cDNA containing *Ltk1* and *Pds*	2.0	3, 4	
p624	Maize genomic DNA clone with *Pds* promoter and *Ltk1* 3' end.	12.2	3,	AF039585
pH58	a Hind III fragment of p624 containing *Pds* promoter and *Ltk1* 3' end.	5.8	3, 4	AF039585
p5R900	Maize cDNA of *Ltk1*, 5' RACE product	0.99	4	
p3R41	Maize cDNA of *Ltk1*, 3' RACE product	2.2	4	
p10F(E)	Maize cDNA with inserts of pMPDS10 and p5R900	2.8	4	
pLTK1	Maize cDNA of *Ltk1*	2.3	4	AF023164
pLTK2	Maize cDNA of *Ltk2*	2.4	4	AF023165
pLTK3	Maize cDNA of *Ltk3*	1.07	4	AF023166
p531	Maize genomic DNA clone containing *Ltk2* promoter	16.6	4	AF023267
p531E2	an EcoR I fragment from p531 containing *Ltk2* promoter	5.07	4	AF023267

Appendix 2: List of primers.

Oligo #	Sequence (5'→ 3')	Chapter	Gene
94	CCTGATGAAATC(T)TCG(A,T)GCG(A,T)GAC(T)CA	2	*Pds*
95	ACAGCA(G)CCTTCCATG(T)GAAGCC(T)AA	2	*Pds*
96	AGGCACATGGCTGAATATCG	2	λgt11 vector
97	CGGCAGTACAATGGATTTCC	2	λgt11 vector
106	TGTGAGGAGTATGCCAAGACG	3	*Psy*
109	CAGGTACGCTCATTAACCCG	3	*Psy*
110	GGAACTGTGAAACACTTCGC	2,3	*Pds*
111	GAAACCTTCGATAGGTGACC	2,3	*Pds*
141	ATCCCTGAGAAAGGCAGAGG	2,3	*Sh1*
142	AGTGACTCCCAACTTGTGCG	2.3	*Sh1*
189	CGTCCAAAAGTACCCTCT	4	*Ltk1*
207	CCATCTGTAGATCTGCAACTGAA	4	*Ltk1, Ltk2, Ltk3*
208	AAGACTCTATGTCTACTGTGG	4	*Ltk1*
219	GGGCATTACTTGCTCAGGAT	4	*Ltk1*
226	CCCTGATCCTGACAAGTAAT	4	*Ltk1*
227	GGGAAGATTGTACTGTACTTG	4	*Ltk1*
229	CAAGGTCGACGAAGCTAAAGC	4	*Ltk1, Ltk2*
231	CGCAATAACCCTCACTAAAGG	4	*Ltk1, Ltk2*
232	GGGAAGATTGTACTGTACTTG	4	*Ltk2*
240	ACGATCGTTGGGGTCAA	4	*Ltk1, Ltk2*
243	TGCACTTCACTTGTCAACAG	4	*Ltk2* promoter
246	GCTCTTCCTTATCGCCAT	4	*Ltk2* promoter
247	CGGATTTGGAGGAGTCGAT	4	*Ltk2*
276	TCACAGGTTGGCAGGCGAAT	4	*Ltk1, Ltk2, Ltk3*

Appendix 3: Partial sequence of pH58 aligned with the 3' ends of *Lkt1* and *Ltk2*.*

```
 H58    1  TTAATTCTTTTCTTGTTATTGTCCCCGTCTCAATTGTAACGGTATCGGGGCCCTTCCTTA
 H58   61  TAGTATAAATTTGGTATTGATGATCAGATATGCCACTGTTTCCATTACAGCTCTAGACCC
 H58  121  AGGTCAGACAGTCCTTGTGCGGTGGGCAACACCCCACTGCATGACATCGACGCATTGGAC
 H58  181  AGGATGGTCGATCCTGCACTCAAGGGTCTATACCCAGCCAAATCTCTTATCCCGGATTTG
 H58  241  CTGATGTCCTTGCCCTGTGTGTCCAGGTAGCAGTTCCATTCATCATCCATGCAGGATAGC

 H58  301  CATAGCATATCCATCTAAATATAACCTGAACCTTGTGAACCAAATTCTTATTGCATCCTG
LTK1 1913                                                          CCTG
LTK2 2132                                                          CCTG

 H58  361  AACCAGAATTCAGGCCACCAATGTCAGAGGTGGTGCAAGCATTGGTTCGACTTGTGCAGA
LTK1 1917  AACCAGAATTCAGGCCACCAATGTCAGAGGTGGTGCAAGCATTGGTTCGACTTGTGCAGA
LTK2 2136  AACCAGAATTCAGGCCACCAATGTCAGAGGTGGTGCAAGCATTGGTTCGACTTGTGCAGA

 H58  421  GGGCCAACATGACGAAGAGGAATGCTTGATGGGGA---TACTTCTCGGCGACCAGATGAC
LTK1 1977  GGGCCAACATGACGAAGAGG-ATGCTTGATGGGGA---TACTTCTCGGCGACCAGATGAC
LTK2 2196  GGGCCAACATGACGAAGAGG-ATGCTTGATGGAGGAGATACTTCTCGTGGACCAGATGAT

 H58  478  CTGGACCAAGATTTCATATGACAAGCTGAACTCT----TGTCAGA-----------TTGT
LTK1 2033  CTGGACCAAGATTTCATATGACAAGCTGAACTCT----TGTCAGA-----------TTGT
LTK2 2255  CAGGACCAATATTTCATATGACAAGCTGAACTCGAGCTTTTCAGCTACCTCTCAGATTGT

 H58  523  ATTTATTTCGTT-CGA----CTAGCGGCCCA-GCA-GAGACATNAATCACACANAGAGGA
LTK1 2078  ATTTATTTCGTT-CGA----CTAGCGGCCCAAGCAAGAGACATAATCACACA--AGAGGA
LTK2 2315  ATTTATCTCGTCTCGTTCCACTAGCGCCC--GGCA-GAGACATAATCACACA--AGAGGA

 H58  576  AGAAACAGAACTCTGTCAGCTTCGCCAGTACATAATCACAAGTTTTTCTTGGTAGCAAAG
LTK1 2131  AGAAACAGAACTCTGTCAGCTTCGCCAGTACATAATCACAAGTTTTTCCTGGTAGCAAAG
LTK2 2370  AGAAACAGAACTCTGTC

 H58  636  TTGTGCTATATGGTTGTACTGTGTCTACAAAATAGTACCTTGATCTTACGNCTAATGGCA
LTK1 2191  TTGTGCTATATGGTTGTACTGTGTCTACAAAATAGTACCTTGATCTTAC-GCTAATGGCA

 H58  696  TATTGTTATAATTCTTTGGTAACTTATATGCAAATGCCCCAGATAAACANANCTATGGTG
LTK1 2190  TATTGTTATAATTCTTTGGTAACTTATATGCAAAT-CCCCAGATAAACA

 H58  756  AAGTGTGNNCCGAATGTTAAAATAACTGATTAGTCCAGTACTGCTGCATCAACTGAAACA
 H58  816  AGTTGCACAGTGGAAATATTTATCTTGGTCCCTCAACTATCCCCTTGAGGCTTATTTTCG
 H58  876  ACCTTGAACTTTCAACATGACTTTCCAGACTTCATCGGTTCTTAAACTTTACTCTTGGGT
 H58  936  CAGTTTTATCCTAAAACTATCAAAATAACTGTTTTGGTCCTCGAACTTTGCATTGTTAGT
 H58  996  TCAATTTTATATGAACTATNCAAAGTAAA
 H58 1025  TTTCAGACCTCAAACTTTTTTTTATTCAACGAGCCGCTGAAAAAAGGTTTTATGTA
 H58 1081  TGAAAACAACAACTTCACTGCTTGTTAGGTAAATGGATAAACAAAGCGCTGATGAATCCC
 H58 1141  CTAATGATTTTTATCAAAATCATTAAGTTAAGGTAGATACACATCTTGTCATATGATCAA
 H58 1201  ATGGTTTCGCCAAAAATCAATAATCAGACAACAAAATGTGCGAACTCGATATTTTACACG
 H58 1261  ACTCTCTTTACCAATTCTGCCCCGAATTACACTTAAAACGACTCAACAGCTTAANCGTTG
 H58 1321  GCTTGCCACGCCTTACTTGACTGTAAAACTCTCACTCTTACCGAACTTGGCCGTAACCTG
 H58 1381  CCAACCAAAGCGAGAANCAAAACATAACATCAAACGAATCGACCGANTTGTTAGGTAATC
 H58 1441  GTCACCNTCCNACAAAGAGCGACTCGCTGTATACCGTTGGCATGCTAGCTTTATCTGTTC
 H58 1501  GGGCAATACGATGCCCATTGTACTTGTTGACTGGTCTGATATCCGTGAGCAAAAACGGCT
 H58 1561  TATGGTATTGCGAGCTTCAGTCGCACTACACGGTCGTTCTGTTACTCTTTATGAGAAAGC
 H58 1621  GTTCCCGCTTTCAGAGCAATGTTCAAAGAAAGCTCATGACCAATTTCTAGCCGACCTTGC
 H58 1681  GAGCATTCTACCGAGTAACACCACACCGCTCATTGTCAGTGATGCTGGCTTTAAAGTGCC
 H58 1741  ATGGTATAAATCCGTTGAGAAGCTGGGTTGGTACTGGTTAAGTCGAGTAAGAGGAAAAGT
```

H58	1801	**ACAATATGCAGACCTAGGAGCGGAAAACTGGAAACCTATCAGCAACTTACATGATATGTC**
H58	1861	**ATCTAGTCACTCAAAGACTTTAGGCTATAAGAGGCTGACTAAAAGCAATCCAATCTCATG**
H58	1921	**CCAAATTCTATTGTATAAATCGGGCTCCTCAAGGCCGAAAAAATCATTCGTCGACACGGA**
H58	1981	**CTCATTGTCACCACCCGTCACCTAAAATCTACTCAGCGTCGGCAAAGGAGCCATGGATTC**
H58	2041	**TAGCAACTAACTTACCTGTTGAAATTCGAACACCCAAACAACTTGTTAATATCTATTCGA**
H58	2101	**AGCGAATGCAGATTGAAGAAACCTTCCGAGACTTGAAAAGTCCTGCCTACGGACTAGCCT**
H58	2161	**ACGCCATAGCCGAACGACAGCTCAGAGCGTTTTGATATCATGCTGCTAATCGCCCTGATG**
H58	2221	**CTTCAACTAACATGTTGGCTTGCGGGCGTTCATGCTCAGAAACAAGGTTGGGACAAGCAC**
H58	2281	**TTCCAGGCTAACACAGTCAGAAATCGAAACGTACTCTCAACAGTTCGCTTAGGCATGGAA**
H58	2341	**GTTTTGCGGCATTCTGGCTACACAATAACAGGGAAGACTTACTCGTGGCTGCAACCCTAC**
H58	2401	**TAGCTCAAAATTTATTCACACATGGTTACGCTTTGGGGAAATTATGAGGGGATCTCTCAG**
H58	2461	AACAAAGCGTCGCTGCACTCCTACGGCAAATCTCCGCTCCGTCTATTCCTGCACACGGGG
H58	2521	AACACTGCGTTCGGCTGAAACGATTGAGCGAGCACGACAGTGGTATTGGGCCGTTGGGCT
H58	2581	CAAAGAACGAGCCAAAAAACAGGGTGGGGTATTTTGACTCTTGATACTAAAGTATAGTGA
H58	2641	TGACAATGGATCGGATTCAGATTTGGTATTACAAATATCCACTCGCAATTATATAAATAG
H58	2701	TTGCAACTATTATCCGTTATCGATCATATCCACAGGTAGAAATTTATATTTATGTTTGTG
H58	2761	TCCATTGAATTTTGGACGGGGTTCGGACATTCATCGGATATGACAAACATAATCAATTTT
H58	2821	CAACAATTCAATAGCATAATTAATCAATTTTTTTCTCAATTCACTATCATATATAATTTA
H58	2881	AAATTTAATAAGAGAATTATTATGGCCCGCGGGCTACGGCTAAAGGTACTAGAGCTTGA
H58	2941	AAAGGGCAGGCCACGGCACTTACTGTTGCGGCCTATAAGACCTTAGCAAATGGCACAAGG
H58	3001	CATAGCGCAGGAAGCCAACGAGGCTAAAAGGGGCGACCAACCGTGGAGAATTGGAAGAAG
H58	3061	TTAGGGTTTACTCTTTGCTATGTAATATGGATTAGACCAAAAAATAAACTAAAGTCAATT
H58	3121	TTGGATATCTAACAGATAACCTGTGGGTGAGAAGGTAATATTCTAATTCATGCTCGTCCG
H58	3181	ACTTTGACTCAAATATGGGTTTAATACACGGGTCTAAAAATGTATTATACTCAGCTTCAT
H58	3241	CGGATCAGATATACAACGAATATCCAAATCCACGTTTTAAATTGTCATCTCTAAGTATAC
H58	3301	ATGGCATGATTGGTTAGCAGCCTAATTCGCAATAATTTACTTAATATTCTTTACTAAAGT

*See text (Chapter 3) for detailed information. The Tn10 sequence is bold.

Appendix 4: The alignment of *Ltk1* and *Ltk2* cDNA.

```
Ltk2 GATTTGGAGGAGTCGAATAAACAAGGAAACTCCTCCCGGCCCAGAAAGCA -50
Ltk2 CTTCACTTGCAACAGTCCCCGTCAGGGAGGGAGCATTGCTGTGCGGGATG -100
Ltk2 GCGATAAGGAAGAGCGGAGGCGGTGGCGGGGTGGGGGGTGTGGCGCGGCG -150
Ltk2 GCTGCTGCTTCTGGCCTGCTGCTGCATTTGGCCTGGTGGACGGATCTTCA -200
Ltk1                 CCCCAACGATCTTAACGTCCTCAACACGCTCTTC -34
Ltk2 CTGCTGCGGACACTGACCCCAACGATCTTAATGTCCTCAATACGCTCTTC -250
Ltk1 ACCAGTCTGAATTCTCCTGGGCAGCTCACAGGTTGGCAGGCGAATGGGGG -84
Ltk2 ACCAGTCTGAATTCTCCTGGGCAGCTCAGAGCTTGGCGGGCGAATGGCGG -300
Ltk1 TGATCCTTGTGGCCAGTCATGGAAGGGCATTACTTGCTCAGGATCAGGGG -134
Ltk2 TGATCCTTGTGGCCAGTCATGGCAGGGCATTACCTGCTCAGGATCAGGGG -350
Ltk1 TCACTAAAATCCAATTACCTAACTTGTCACTCACTGGAAATTTGGCCTAC -184
Ltk2 TCACTAAAATCCTATTACCAAACTTGTCACTCACCGGAAATCTGGCCTAC -400
Ltk1 AACATGAATAACTTGGGTTCATTAGTTGAGCTTGACATGAGCCAAAATAA -234
Ltk2 AACATGAATAACTTGGGTTCATTAGTTGAGCTTGACTTGAGCCAAAATAA -450
Ltk1 CCTTGGCGGTGGAGGTCAAGTACAGTACAATCTTCCCAATATGAAGCTTG -284
Ltk2 CCTTGGTGGTGGAGGCCAAATACAGTACAATCTTCCCAATGTGAAGCTTG -500
Ltk1 AGAAACTCAATCTTGCAGGAAATCAATTTGGTGGAAATTTACCCTACTCA -334
Ltk2 AGAAACTCAATCTTGCAGGAAATCAATTTGGTGGAAATTTACCCTACTCA -550
Ltk1 ATTTCGACGATGCCTAATCTTAAGTATTTAAACCTTAATCATAACCAACT -384
Ltk2 ATTTCGACGATGCCTAATCTTAAGTATTTAAACCTTAATCATAACCAACT -600
Ltk1 ACAAGGAAACATCAGTGATGTATTTTCCAACCTTTACAGTTTGTCAGAAC -434
Ltk2 ACAAGGAAACATCACTGATGTATTTTCCAACCTTTATAGTTTGTCAGAGC -650
Ltk1 TGGATCTGTCCTTTAATTCCCTTACTGGTGATCTACCACAAAGTTTCACT -484
Ltk2 TGGATCTCTCCTTTAATTCTCTTACTGGTGATCTACCACAAAGTTTCACT -700
Ltk1 GGATTGTCAAGCCTGAAAAAAGTGTATTTGCAGAACAACCAATTTACTGG -534
Ltk2 GGATTGTCAAGCCTGAAAAGAATGTATTTGCAGAACAACCAATTTACTGG -750
Ltk1 TAATATCAATGTCTTAGCCAATCTCCCCCTTGAAACTTTGAATGTTGCGA -584
Ltk2 TTATATCAATGTCTTAGCCAACCTTCCCCTTGAAACCCTGAATGTTGGAA -800
Ltk1 ACAACCATTTCACTGGTTGGATTCCTAGTCAGCTTAAGAAGATAAACAGT -634
Ltk2 ACAACCATTTCACTGGTTGGATTCCCAGTCAGCTTAAGGAGATAAACAAT -850
Ltk1 CTACAGACTGATGGAAATTCTTGGAGCACAGGACCAGCGCCACCTCCACC -684
Ltk2 CTACAGACCGACGGAAATTCTTGGAGCACAGGACCAGCGCCACCTCCACC -900
Ltk1 TCCATACACAGCGCCGCCTCCTCCTCCAAACCATTGGAATGCTGATGGTT -734
Ltk2 TCCATATACAGCGCC---TCCTCCTCCAAACCATTGGAATGGCGCGGGTC -947
Ltk1 -----CATCAAGT-TCATCAAGTTCTGGTGGAAGATCTGGGATAGGTGGT -778
Ltk2 AGAATGATGATGGTTCATCAAGTTCTGGTGGAAGACCTGGGATAGGTGGT -997
Ltk1 GGAGGTGTAGCAGGAATCATTATATCGTTGCTGGTTGTTGGATCAGTTGT -828
Ltk2 GGAGGTGTAGCAGGAATCATTATATCATTGCTGATTGTTGGATCAGTTGT -1047
Ltk1 TGCATTTCTTGTGATCAAAAGAAGAAAACGCAAAGCTGCTATGAAAGAAC -878
Ltk2 TGCATTTTTTCTAATCAAGAGAAGAAAACACAAAGCTATTATGGAAGAAC -1097
Ltk1 ATTTTGAACAGCACCAGCCATTCACTTCCTTCCCTTCAAATGAAGTTAAA -928
Ltk2 AATTTGAACAGCATCAGCCGTTCACTTCCTTCCCTTCAAATGAAGTTAAC -1147
Ltk1 GACATGAAGCCTGTCTGCGAGGCCACCACAGTAGACATAGAGTCTTTGGC -978
Ltk2 GACATGAAGCCTATCTATGAGTCCACCACAGTAGACATAGAGTCTTTGGC -1197
Ltk1 TTCACCTGCTTCAGTTAATCTGAAACCACCTCCGAAGATAGAACGAAACA -1028
Ltk2 TTCACCTGCTTCGATTAATCTGAAACCACCCCCGAAGATAGAACAGAACA -1247
Ltk1 AATCATTTGATGATGACGATGATTTTTCAAACAAGCTTGTTGCAAAGAAA -1078
Ltk2 AATCATTTGATGACGATGATGATTTTTCAAACAAGACTGCTGCAAATAGA -1297
```

```
Ltk1 AGTAATATAACACCTATAAATGCAACTGTTTATTCAGTTGCAGATCTACA -1128
     ||||||||||||||||| || |||||||||||||||||||||||||||||
Ltk2 AGTAATATAACACCTATGAAGGCAACTGTTTATTCAGTTGCAGATCTACA -1347
Ltk1 GATGGCAACAGATAGCTTTAGCTTCGACAACCTTGTTGGAGAGGGTACTT -1178
     |||||||||||||||||| |||||||| ||||||||||||||||||||||
Ltk2 GATGGCAACAGATAGCTTCAGCTTCGATAACCTTGTTGGAGAGGGTACTT -1397
Ltk1 TTGGACGTGTTTACAGGGCACAATTCAATGATGGAAAGGTTCTGGCGATC -1228
     |||||||||||||||||||||| ||||||| ||||||||||||||| |||
Ltk2 TTGGACGTGTTTACAGGGCACAGTTCAATGGTGGAAAGGTTCTGGCCATC -1447
Ltk1 AAGAAATTAGATAGTACTGTGATGCCATTTCAATCATCTGATGACTTTGC -1278
     ||||||||||| | ||||||||||||||||||||| ||||||||||||||
Ltk2 AAGAAATTAGACATTACTGTGATGCCATTTCAATCGTCTGATGACTTTGC -1497
Ltk1 TGAACTGGTCTCGAACATTTCAAAATTGCACCATCCGAATCTGAATGAGC -1328
     ||||||||||||||||||||||||||||||||||||||||||||||||||
Ltk2 TGAACTGGTCTCGAACATTTCAAAATTGCACCATCCGAATCTGAATGAGC -1547
Ltk1 TTGTGGGCTATTGCATGGAACATGGCCAGCACTTGCTTGTGTATGATTTC -1378
     |||| |||||||||||||||||||||||||||||||||||||||||||||
Ltk2 TTGTAGGCTATTGCATGGAACATGGCCAGCACTTGCTTGTGTATGATTTC -1597
Ltk1 CACAGGAATGGATCACTTCATGATCTACTCCACCTTTCAGATGAGTACAG -1428
     || ||||||||||| ||||||||||||| |||||||||||||||||||||
Ltk2 CATAGGAATGGATCGCTTCATGATCTACGCCACCTTTCAGATGAGTACAG -1647
Ltk1 CAAGCCACTTAGCTGGAACACTCGTATCAAGATCGCACTCGGCTCTGCCC -1478
     |||| |||||||||||||| |||| |||||| ||||||||||||||||||
Ltk2 CAAGGCACTTAGCTGGAACTCTCGCATCAAGTTCGCACTCGGCTCTGCCC -1697
Ltk1 GTGCACTGGAGTACCTTCATGAAATATGTTCTCCATCCATCATCCATAAG -1528
     ||||||||||||||||||||||||||||||||||||||||||||||||||
Ltk2 GTGCACTGGAGTACCTTCATGAAATATGTTCTCCATCCATCATCCATAAG -1747
Ltk1 AATTTCAAGTCGTCAAACATTTTGCTGGACTCAGAATTCAATCCACACCT -1578
     ||||||||| | ||||| ||||||||||||||||||||||||||||||||
Ltk2 AATTTCAAGGCATCAAATATTTTGCTGGACTCAGAATTCAATCCACACCT -1797
Ltk1 TTCAGATGCTGGACTTGCAAGCTTTATTCCTGATGCTGAATTCCAGGCAG -1628
     ||||||| |||||||||||||||| ||||||| |||||||||||||||||
Ltk2 TTCAGATACTGGACTTGCAAGCTTCATTCCTGGTGCTGAATTCCAGGCAG -1847
Ltk1 CAGAACAAAGTGCCGGATACACTGCCCCAGAAGTGGACATGACCGGTCAG -1678
     ||||||||||||||||||||||||||||||||||||||||||||||||||
Ltk2 CAGAACAAAGTGCCGGATACACTGCCCCAGAAGTGGACATGACCGGTCAG -1897
Ltk1 TACACCTTCAAGAGTGATGTCTACAGCTTTGGAGTTGTCATGCTTGAACT -1728
     |||||| ||||||||||||| |||||||||||||||||||||||||| ||
Ltk2 TACACCCTCAAGAGTGATGTTTACAGCTTTGGAGTTGTCATGCTTGAGCT -1947
Ltk1 ATTGACAGGACGTAGACCATTTGACAGCTCTAGACCCAGGTCAGAGCAGT -1778
     | ||||||||||||||||||||||||||||||||||||||||||||||||
Ltk2 ACTGACAGGACGTAGACCATTTGACAGCTCTAGACCCAGGTCAGAGCAGT -1997
Ltk1 CACTTGTGCGGTGGGCAACACCCCAGCTGCATGACATCGACGCATTGGAC -1828
     ||||||||||||||||| |||| ||||||||||||||||| |||||||||
Ltk2 CACTTGTGCGGTGGGCAGCACCTCAGCTGCATGACATCGATGCATTGGAC -2047
Ltk1 AGGATGGTCGATCCTGCACTCAAGGGTCTATACCCAGCCAAATCTCTATC -1878
     || |||||||||||||||||||||||||||||||||||||||||||||||
Ltk2 AGAATGGTCGATCCTGCACTCAAGGGTCTATACCCAGCCAAATCTCTATC -2097
Ltk1 CCGATTTGCTGATGTCCTTGCCCTGTGTGTCCAGCCTGAACCAGAATTCA -1928
     |||||||||||||||||| |||||||||||||||||||||||||||||||
Ltk2 CCGATTTGCTGATGTCCTCGCCCTGTGTGTCCAGCCTGAACCAGAATTCA -2147
Ltk1 GGCCACCAATGTCAGAGGTGGTGCAAGCATTGGTTCGACTTGTGCAGAGG -1978
     ||||||||||||||||||||||||||||||||||||||||||||||||||
Ltk2 GGCCACCAATGTCAGAGGTGGTGCAAGCATTGGTTCGACTTGTGCAGAGG -2197
Ltk1 GCCAACATGACGAAGAGGATGCTTGATGGGGA---TACTTCTCGGCGACC -2025
     ||||||||||||||||||||||||||||| |    |||||||||  ||||
Ltk2 GCCAACATGACGAAGAGGATGCTTGATGGAGGAGATACTTCTCGTGGACC -2247
Ltk1 AGATGACCTGGACCAAGATTTCATATGACAAGCTGAACTCT----TGTCA -2071
     |||||| | ||||||| |||||||||||||||||||||||     | |||
Ltk2 AGATGATCAGGACCAATATTTCATATGACAAGCTGAACTCGAGCTTTTCA -2297
Ltk1 GA-----------TTGTATTTATTTCGTT-CGA----CTAGCGGCCCAAG -2105
     |            |||||||||| ||||  ||     |||||| ||   |
Ltk2 GCTACCTCTCAGATTGTATTTATCTCGTCTCGTTCCACTAGCGCCC--GG -2345
Ltk1 CAAGAGACATAATCACACAAGAGGAAGAAACAGAACTCTGTCAGCTTCGC -2155
     || |||||||||||||||||||||||||||||||||||||||
Ltk2 CA-GAGACATAATCACACAAGAGGAAGAAACAGAACTCTGTC         -2386
Ltk1 CAGTACATAATCACAAGTTTTTCCTGGTAGCAAAGTTGTGCTATATGGTT -2205
Ltk1 GTACTGTGTCTACAAAATAGTACCTTGATCTTACGCTAATGGCATATTGT -2255
Ltk1 TATAATTCTTTGGTAACTTATATGCAAATCCCCAGATAAACAAAAAAAAA -2305
Ltk1 A -2306
```

Bibliography

Al-Babili, S., von Lintig, J., Haubruck, H., and Beyer, P. (1996). A novel, soluble form of phytoene desaturase from *Narcissus pseudonarcissus* chromoplasts is Hsp70-complexed and competent for flavinylation, membrane association and enzymatic activation. Plant J *9*, 601-612.

Albrecht, M., Klein, A., Hugueney, P., Sandmann, G., and Kuntz, M. (1995). Molecular cloning and functional expression in *E. coli* of a novel plant enzyme mediating ζ-carotene desaturation. FEBS Lett *372*, 199-202.

Albrecht, M., and Sandmann, G. (1994). Light-stimulated carotenoid biosynthesis during transformation of maize etioplasts is regulated by increased activity of isopentenyl pyrophosphate isomerase. Plant Physiol *105*, 529-534.

Armstrong, G. A. (1994). Eubacteria show their true colors: genetics of carotenoid pigment biosynthesis from microbes to plants. J Bacteriol *176*, 4795-4802.

Armstrong, G. A., Alberti, M., Leach, F., and Hearst, J. E. (1989). Nucleotide sequence, organization, and nature of the protein products of the carotenoid biosynthesis gene cluster of Rhodobacter capsulatus. Mol Gen Genet *216*, 254-268.

Badillo, A., Steppuhn, J., Deruere, J., Camara, B., and Kuntz, M. (1994). Structure of a functional geranylgeranyl pyrophosphate synthase gene from *Capsicum annuum*. Plant Mol Biol *27*, 425-428.

Bartley, G. E., and Scolnik, P. A. (1993). cDNA cloning, expression during development, and genome mapping of Psy2, a second tomato gene encoding phytoene synthase. J Biol Chem *268*, 25718-25721.

Bartley, G. E., and Scolnik, P. A. (1995). Plant carotenoids: pigments for photoprotection, visual attraction, and human health. Plant Cell *7*, 1027-1038.

Bartley, G. E., Scolnik, P. A., and Giuliano, G. (1994). Molecular biology of carotenoid biosynthesis in plants. Annu Rev Plant Physiol Plant Mol Biol *45*, 287-301.

Bartley, G. E., Viitanen, P. V., Bacot, K. O., and Scolnik, P. A. (1992). A tomato gene expressed during fruit ripening encodes an enzyme of the carotenoid biosynthesis pathway. J Biol Chem *267*, 5036-5039.

Bartley, G. E., Viitanen, P. V., Pecker, I., Chamovitz, D., Hirschberg, J., and Scolnik, P. A. (1991). Molecular cloning and expression in photosynthetic bacteria of a soybean cDNA coding for phytoene desaturase, an enzyme of the carotenoid biosynthesis pathway. Proc Natl Acad Sci USA *88*, 6532-6536.

Becraft, P. W., Stinard, P. S., and McCarty, D. R. (1996). CRINKLY4: A TNFR-like receptor kinase involved in maize epidermal differentiation. Science *273*, 1406-1409.

Berman, S. (1991). Epidemiology of acute respiratory infections in children of developing countries. Rev Infect Dis *13 Suppl 6*, S454-62.

Bertram, J. S., and Bortkiewicz, H. (1995). Dietary carotenoids inhibit neoplastic transformation and modulate gene expression in mouse and human cells. Am J Clin Nutr *62*, 1327S-1336S.

Beyer, P., Mayer, M., and Kleinig, H. (1989). Molecular oxygen and the state of geometric isomerism of intermediates are essential in the carotene desaturation and cyclization reaction in daffodil chromoplasts. Eur J Biochem *184*, 141-150.

Beyer, P., Weiss, G., and kleinig, H. (1985). Solubilization and reconstitution of the membrane-bound carotenogenic enzymes from daffodil chromoplasts. Eur J Biochem *153*, 341-346.

Bhaskaram, P. (1995). Measles & malnutrition. Indian J Med Res *102*, 195-199.

Bonk, M., Hoffmann, B., Von Lintig, J., Schledz, M., Al-Babili, S., Hobeika, E., Kleinig, H., and Beyer, P. (1997). Chloroplast import of four carotenoid biosynthetic enzymes in vitro reveals differential fates prior to membrane binding and oligomeric assembly. Eur J Biochem *247*, 942-950.

Bonk, M., Tadros, M., Vandekerckhove, J., Al-Babili, S., and Beyer, P. (1996). Purification and characterization of chaperonin 60 and heat-shock protein 70 from chromoplast of *Narcissus pseudonarcissus*. Involvement of heat-shock protein 70 in a soluble protein complex containing phytoene desaturase. Plant Physiol *111*, 931-939.

Bramley, P., Teulieres, C., Blain, I., Bird, C., and Schuch, W. (1992). Biochemical characterization of transgenic tomato plants in which carotenoid synthesis has been inhibited through the expression of antisense RNA to pTOM5. Plant J *2*, 343-349.

Bramley, P. M. (1994). Carotenoid biosynthesis: a target site for bleaching herbicides. Biochem Soc Trans *22*, 625-629.

Bramley, P. M. (1985). The in vitro biosynthesis of carotenoids. In Advances in lipid research., R. Paoletti and D. Kritchevsky, eds. (New York: Academic Press, Inc.), pp.

243-279.

Britton, G., and Goodwin, T. W. (1982). Carotenoid Chemistry and Biochemistry (New York: Pergamon Press).

Britton, G., Liaaen-Jensen, S., and Pfander, H. (1995). Carotenoids today and challenges for the future. In Carotenoids, G. Britton, S. Liaaen-Jensen and H. Pfander, eds. (Basel: Birkhauser Verlag).

Bucher, P. (1990). Weight matrix descriptions of four eukaryotic RNA polymerase II promoter elements derived from 502 unrelated promoter sequences. J Mol Biol *212*, 563-578.

Buckner, B., Kelson, T. L., and Robertson, D. S. (1990). Cloning of the *y1* locus of maize, a gene involved in the biosynthesis of carotenoids. Plant Cell *2*, 867-876.

Buckner, B., Miguel, P. S., Janick-Buckner, D., and Bennetzen, J. L. (1996). The *y1* gene of maize codes for phytoene synthase. Genetics *143*, 479-488.

Burkhardt, P. K., Beyer, P., Wunn, J., Kloti, A., Armstrong, G. A., Schledz, M., von Lintig, J., and Potrykus, I. (1997). Transgenic rice (*Oryza sativa*) endosperm expressing daffodil (*Narcissus pseudonarcissus*) phytoene synthase accumulates phytoene, a key intermediate of provitamin A biosynthesis. Plant J *11*, 1071-1078.

Burr, B., Burr, F. A., Thompson, K. H., Albertson, M. C., and Stuber, C. W. (1988). Gene mapping with recombinant inbreds in maize. Genetics *118*, 519-526.

Camara, B., Bardat, F., and Moneger, R. (1982). Sites of biosynthesis of carotenoids in *Capsicum* chromoplasts. Eur J Biochem *127*, 255-258.

Chamovitz, D., Pecker, I., and Hirschberg, J. (1991). The molecular basis of resistance to the herbicide norflurazon. Plant Mol Biol *16*, 967-974.

Chang, C., Schaller, G. E., Patterson, S. E., Kwok, S. F., Meyerowitz, E. M., and Bleecker, A. B. (1992). The TMK1 gene from *Arabidopsis* codes for a protein with structural and biochemical characteristics of a receptor protein kinase. Plant Cell *4*, 1263-1271.

Chappell, J. (1995). Biochemistry and molecular biology of the isoprenoid biosynthetic pathway in plants. Annu Rev Plant Physiol Plant Mol Biol *46*, 521-547.

Clark, S. E., Williams, R. W., and Meyerowitz, E. M. (1997). The *CLAVATA1* gene encodes a putative receptor kinase that controls shoot and floral meristem size in *Arabidopsis*. Cell *89*, 575-585.

Corona, V., Aracri, B., Kosturkova, G., Bartley, G. E., Pitto, L., Giorgetti, L., Scolnik, P. A., and Giuliano, G. (1996). Regulation of a carotenoid biosynthesis gene promoter during plant development. Plant J *9*, 505-512.

Cunningham, F. X., Jr., Pogson, B., Sun, Z., McDonald, K. A., DellaPenna, D., and Gantt, E. (1996). Functional analysis of the beta and epsilon lycopene cyclase enzymes of *Arabidopsis* reveals a mechanism for control of cyclic carotenoid formation. Plant Cell *8*, 1613-1626.

Demming, B., Adams, W. W., Heber, U., Neimanis, S., Winter, K., Kruger, A., Czygan, F.-C., Bilger, W., and Bjorkman, O. (1990). Inhibition of zeaxanthin formation and of rapid changes in radiationless energy dissipation by dithiothreitol in spinach leaves and chloroplasts. Plant Physiol *92*, 293-301.

Demming, B., Winter, K., Kruger, A., and Czygan, F.-C. (1989). Light response of CO_2 assimilation, dissipation of excess excitation energy, and zeaxanthin content of sun and shade leaves. Plant Physiol *90*, 881-886.

Dixon, M. S., Jones, D. A., Keddie, J. S., Thomas, C. M., Harrison, K., and Jones, J. D. (1996). The tomato *Cf-2* disease resistance locus comprises two functional genes encoding leucine-rich repeat proteins. Cell *84*, 451-459.

Dogbo, O., Laferriere, A., D'Harlingue, A., and Camara, B. (1988). Carotenoid biosynthesis: isolation and characterization of a bifunctional enzyme catalyzing the synthesis of phytoene. Proc Natl Acad Sci USA *85*, 7054-7058.

Elmazar, M. M., Reichert, U., Shroot, B., and Nau, H. (1996). Pattern of retinoid-induced teratogenic effects: possible relationship with relative selectivity for nuclear retinoid receptors RAR alpha, RAR beta, and RAR gamma. Teratology *53*, 158-167.

Ester, W. H. (1931). Vivipary in maize. Genetics *16*, 574-590.

Eugester, C. H. (1995). History: 175 years of carotenoid chemistry. In Carotenoids, G. Britton, S. Liaaen-Jensen and H. Pfander, eds. (Basel: Birkhauser Verlag).

Fontes, E. B. P., Shank, B. B., Wrobel, R. L., Moose, S. P., OBrian, G. R., Wurtzel, E. T., and Boston, R. B. (1991). Characterization of an immunoglobulin binding protein homolog in the maize *floury-2* endosperm mutant. Plant Cell *3*, 483-496.

Fraser, P. D., Hedden, P., Cooke, D. T., Bird, C. R., Schuch, W., and Bramley, P. M. (1995). The effect of reduced activity of phytoene synthase on isoprenoid levels in tomato pericarp during fruit development and ripening. Planta *196*, 321-326.

Fraser, P. D., Truesdale, M. R., Bird, C. R., Schuch, W., and Bramley , P. M. (1994). Carotenoid biosynthesis during tomato fruit development. Evidence for tissue specific gene expression. Plant Physiol *105*, 405-413.

Fray, R. G., Wallace, A., Fraser, P. D., Valero, D., Hedden, P., Bramley, P. M., and Grierson, D. (1995). Constitutive expression of a fruit phytoene synthase gene in transgenic tomatoes causes dwarfism by redirecting metabolites from gibberellin pathway. Plant J *8*, 693-701.

Giuliano, G., Bartley, G. E., and Scolnik, P. A. (1993). Regulation of carotenoid biosynthesis during tomato development. Plant Cell *5*, 379-387.

Goodwin, T. W. (1976). Chemistry and biochemistry of plant pigments, 2nd Edition, Volume 1 (New York: Academic Press).

Guerini, D. (1997). Calcineurin: not just a simple protein phosphatase. Biochem Biophys Res Commun *235*, 271-275.

Hable, W. E., and Oishi, K. K. (1995). Maize phytoene desaturase maps near the *viviparous5* locus. Plant Physiol *108*, 1329-1330.

Hanks, S. K., and Quinn, A. M. (1991). Protein kinase catalytic domain sequence database: identification of conserved features of primary structure and classification of family members. Methods Enzymol *200*, 38-62.

Herve, C., Dabos, P., Galaud, J. P., Rouge, P., and Lescure, B. (1996). Characterization of an *Arabidopsis thaliana* gene that defines a new class of putative plant receptor kinases with an extracellular lectin-like domain. J Mol Biol *258*, 778-788.

Higgins, D. G., and Sharp, P. M. (1988). CLUSTAL: a package for performing multiple sequence alignment on a microcomputer. Gene *73*, 237-244.

Hong, S. W., Jon, J. H., Kwak, J. M., and Nam, H. G. (1997). Identification of a receptor-like protein kinase gene rapidly induced by abscisic acid, dehydration, high salt, and cold treatments in *Arabidopsis thaliana*. Plant Physiol *113*, 1203-1212.

Hugueney, P., Badillo, A., Chen, H. C., Klein, A., Hirschberg, J., Camara, B., and Kuntz, M. (1995). Metabolism of cyclic carotenoids: a model for the alteration of this biosynthetic pathway in *Capsicum annuum* chromoplasts. Plant J *8*, 417-424.

Hugueney, P., Romer, S., Kuntz, M., and Camara, B. (1992). Characterization and molecular cloning of a flavoprotein catalyzing the synthesis of phytofluene and zeta-carotene in *Capsicum* chromoplasts. Eur J Biochem *209*, 399-407.

IUPAC (1972). IUPAC Commission on the Nomenclature of Organic Chemistry (CNOC) and IUPAC-IUB Commission on Biochemical Nomenclature (CBN). Tentative rules for the nomenclature of carotenoids. Eur J Biochem *25*, 397-408.

Jiang, X., Dreano, M., Buckler, D. R., Cheng, S., Ythier, A., Wu, H., Hendrickson, W. A., and el Tayar, N. (1995). Structural predictions for the ligand-binding region of glycoprotein hormone receptors and the nature of hormone-receptor interactions. Structure *3*, 1341-1353.

Karvouni, Z., John, I., Taylor, J. E., Watson, C. F., Turner, A. J., and Grierson, D. (1995). Isolation and characterisation of a melon cDNA clone encoding phytoene synthase. Plant Mol Biol *27*, 1153-1162.

Keith, F. J., and Gay, N. J. (1990). The *Drosophila* membrane receptor Toll can function to promote cellular adhesion. EMBO J *9*, 4299-4306.

Kiesselbach, T. A. (1949). The Structure and Reproduction of Corn (Lincoln, USA and London, UK: Agricultural Experiment Station, University of Nebraska).

Kirk, J. T. O., and Tilney-Bassett, R. A. E. (1978). The Plastids: their chemistry, structure, growth and inheritance, Second edition Edition (Amsterdam, The Netherlands: Elsevier/North Holland Biomedical Press).

Kobe, B., and Deisenhofer, J. (1995). A structural basis of the interactions between leucine-rich repeats and protein ligands. Nature *374*, 183-186.

Kohorn, B. D., Lane, S., and Smith, T. A. (1992). An *Arabidopsis* serine/threonine kinase homologue with an epidermal growth factor repeat selected in yeast for its specificity for a thylakoid membrane protein. Proc Natl Acad Sci USA *89*, 10989-10992.

Krantz, D. E., and Zipursky, S. L. (1990). Drosophila chaoptin, a member of the leucine-rich repeat family, is a photoreceptor cell-specific adhesion molecule. Embo J *9*, 1969-77.

Kreuz, K., Beyer, P., and Kleinig, H. (1982). The site of carotenogenic enzymes in chromoplasts from *Narcissus pseudonarcissus* L. Planta, 66-69.

Kulkarni, C. G. (1927). Inheritance studies of white capping in yellow dent maize. Pap Mich Acad Sci Arts Lett *6*, 253-273.

Kumagai, M. H., Donson, J., della-Cioppa, G., Harvey, D., Hanley, K., and Grill, L. K. (1995). Cytoplasmic inhibition of carotenoid biosynthesis with virus-derived RNA. Proc Natl Acad Sci U S A *92*, 1679-1683.

Kuntz, M., Römer , S., Suire, C., Hugueney, P., Weil, J. H., Schantz, R., and Camara, B. (1992). Identification of a cDNA for the plastid-located geranylgeranyl pyrophosphate synthase from *Capsicum annum*: correlative increase in enzyme activity and transcript level during fruit ripening. Plant J. *2*, 25-34.

Kyte, J., and Doolittle, R. F. (1982). A simple method for displaying the hydropathic character of a protein. J Mol Biol *157*, 105-132.

Li, Z. H., Matthews, P. D., Burr, B., and Wurtzel, E. T. (1996). Cloning and characterization of a maize cDNA encoding phytoene desaturase, an enzyme of the carotenoid biosynthetic pathway. Plant Mol Biol *30*, 269-79.

Li, Z-h., Wurtzel, E. T. (In press). The *Ltk* gene family encodes novel receptor-like kinases with temporal expression in developing mziae endosperm. Plant Mol Biol

Linden, H., Lucas, M. M., de Felipe, M. R., and Sandmann, G. (1993). Immunogold localization of phytoene desaturase in higher plant chloroplasts. Physiol Plant *88*, 229-231.

Linden, H., Misawa, N., Chamovitz, D., Pecker, P., Hirschberg, J., and Sandmann, G. (1991). Functional complementation in *Escherichia coli* of different phytoene desaturase genes and analysis of accumulated carotenes. Z Naturforsch *46c*, 1045-1051.

Linden, H., Misawa, N., Saito, T., and Sandmann, G. (1994). A novel carotenoid biosynthesis gene coding for zeta-carotene desaturase: functional expression, sequence and phylogenetic origin. Plant Mol Biol *24*, 369-379.

Logemann, J., Schell, J., and Willmitzer, L. (1987). Improved method for the isolation of RNA from plant tissues. Anal Biochem *163*, 16-20.

Lutke-Brinkhaus, F., Liedvogel, B., Kreuz, K., and Kleinig, H. (1982). Phytoene synthase and phytoene dehydrogenase associated with envelope membranes from spinach chloroplasts. Planta *156*, 176-180.

Maluf, M. P., Saab, I. N., Wurtzel, E. T., and Sachs, M. M. (1997). The *viviparous12* maize mutant is deficient in abscisic acid, carotenoid and chlorophyll synthesis. J Experim Botany *48*, 1259-1268.

Mangelsdorf, P. C., and Fraps, G. S. (1931). A direct quantitative relationship between vitamin A in corn and number of genes for yellow pigmentaion. Science *73*, 241-242.

Marin, E., Nussaume, L., Quesada, A., Gonneau, M., Sotta, B., Hugueney, P., Frey, A., and Marion-Poll, A. (1996). Molecular identification of zeaxanthin epoxidase of *Nicotiana plumbaginifolia*, a gene involved in abscisic acid biosynthesis and corresponding to the ABA locus of *Arabidopsis thaliana*. EMBO J. *15*, 2331-2342.

Mayer, M. P., Nievelstein, V., and Beyer, P. (1992). Purification and characterization of a NADPH dependent oxidoreductase from chromoplasts of *Narcissus pseudonarcissus*: a redox-mediator possibly involved in carotene desaturation. Plant Physiol Biochem *30*, 389-398.

Mayfield, S. P., Nelson, T., Taylor, W. C., and Malkin, R. (1986). Carotenoid synthesis and pleiotropic effects in carotenoid-deficient seedlings of maize. Planta *169*, 23-32.

McCarty, D. R., Hattori, T., Carson, C. B., Vasil, V., Lazar, M., and Vasil, I. K. (1991). The *Viviparous-1* developmental gene of maize encodes a novel transcriptional activator. Cell *66*, 895-905.

McCullough, A. J., Gengenbach, B. G., and Jones, R. L. (1989). Plastid genome amplification and expression in developing endosperm. In Physiology, biochemistry, and genetics of nongreen plastids, C. D. Boyer, J. C. Shannon and R. C. Hardison, eds. (Maryland: The American Society of Plant Physiologists), pp. 203-214.

McGarvey, D. J., and Croteau, R. (1995). Terpenoid Metabolism. Plant Cell *7*, 1015-1026.

Misawa, N., Nakagawa, M., Kobayashi, K., Yamano, S., Izawa, Y., Nakamura, K., and Harashima, K. (1990). Elucidation of the *Erwinia uredovora* carotenoid biosynthetic pathway by functional analysis of gene products expressed in *Escherichia coli*. J Bacteriol *172*, 6704-6712.

Morriss-Kay, G. M., and Sokolova, N. (1996). Embryonic development and pattern formation. FASEB J *10*, 961-968.

Mu, J. H., Lee, H. S., and Kao, T. H. (1994). Characterization of a pollen-expressed receptor-like kinase gene of *Petunia inflata* and the activity of its encoded kinase. Plant Cell *6*, 709-721.

Neill, S. J., Horgan, R., and Parry, A. D. (1986). The carotenoid and abscisic acid content of viviparous kernels and seedlings of *Zea mays* L. Planta *169*, 87-96.

Noctor, G., Rees, D., Young, A. J., and Horton, P. (1991). The relationship between zeaxanthin, energy-dependent quenching of chlorophyll fluorescence and trans-thylakoid pH gradient in isolated chloroplasts. Biochim Biophys Acta *1057*, 320-330.

Norris, S. R., Barrette, T. R., and DellaPenna, D. (1995). Genetic dissection of carotenoid synthesis in *Arabidopsis* defines plastoquinone as an essential component of phytoene desaturation. Plant Cell *7*, 2139-2149.

Parry, A. D., and Hogan, R. (1991). Carotenoid metabolism and the biosynthesis of abscisic acid. Phytochem *30*, 815-821.

Pay, A., Heberle-Bors, E., and Hirt, H. (1992). An *Alfalfa* cDNA encodes a protein with homology to translationally controlled human tumor protein. Plant Mol Biol *19*, 501-503.

Pecker, I., Chamovitz, D., Linden, H., Sandmann, G., and Hirschberg, J. (1992). A single polypeptide catalyzing the conversion of phytoene to zeta-carotene is transcriptionally regulated during tomato fruit ripening. Proc Natl Acad Sci USA *89*, 4962-4966.

Pecker, I., Gabbay, R., Cunningham, F. X., Jr., and Hirschberg, J. (1996). Cloning and characterization of the cDNA for lycopene beta-cyclase from tomato reveals decrease in its expression during fruit ripening. Plant Mol Biol *30*, 807-819.

Phillip, D., Ruban, A. V., Horton, P., Asato, A., and Young, A. J. (1996). Quenching of chlorophyll fluorescence in the major light-harvesting complex of photosystem II: A systematic study of the effect of carotenoid structure. Proc Natl Acad Sci USA *93*, 1492-1497.

Purves, W. K., Orians, G. H., and Heller, H. C. (1995). Life: The science of biology, 4th Edition (USA: Sinauer Associates, Inc.).

Randolph, L. F., and Hand, D. B. (1940). Relation between carotenoid content and number of genes per cell in diploid and tetraploid corn. J Agr Res *60*, 51-64.

Rechsteiner, M. (1990). PEST sequences are signals for rapid intracellular proteolysis. Semin Cell Biol *1*, 433-440.

Rechsteiner, M., and Rogers, S. W. (1996). PEST sequences and regulation by proteolysis. Trends Biochem Sci *21*, 267-271.

Rees, D., Young, A. J., Noctor, G., Britton, G., and Horton, P. (1989). Enhancement of the pH-dependent dissipation of excitation energy in spinach chloroplasts by light activation: correlation with the synthesis of zeaxanthin. FEBS lett *256*, 85-95.

Ren, R., Mayer, B. J., Cicchetti, P., and Baltimore, D. (1993). Identification of a ten-amino acid proline-rich SH3 binding site. Science *259*, 1157-1161.

Robertson, D., Anderson, I., and Bachmann, M. (1978). Pigment-deficient mutants: Genetic, biochemical and developmental studies. In Maize Breeding and Genetics, D. Walden, ed. (New York: John Wiley & Sons), pp. 461-494.

Robertson, D. S. (1975). Survey of the albino and white-endosperm mutants of maize. J Hered *66*, 67-74.

Robertson, D. S., and Anderson, I. C. (1961). Temperature sensitive allsles of the *yl* locus in maize. J Hered *52*, 53-60.

Robertson, D. S., Bachmann, M. D., and Anderson, I. C. (1966). Role of carotenoids in protecting chlorophyll from photodestruction-II. Studies on the effect of four modifiers of the albino cl_l mutant of maize. Photochem Photobiol *5*, 797-805.

Rogers, S., Wells, R., and Rechsteiner, M. (1986). Amino acid sequences common to rapidly degraded proteins: the PEST hypothesis. Science *234*, 364-368.

Romer, S., Hugueney, P., Bouvier, F., Camara, B., and Kuntz, M. (1993). Expression of the genes encoding the early carotenoid biosynthetic enzymes in *Capsicum annuum.* Biochem Biophys Res Commun *196*, 1414-1421.

Ross, A. C. (1992). Vitamin A status: relationship to immunity and the antibody response. J Soc Exp Biol Med *200*, 303-320.

Sambrook, J., Frisch, E. F., and Maniatis, T. (1989). Molecular Cloning: A Laboratory Manual, 2 Edition (Cold Spring Harbor, New York: Cold Spring Harbor Laboratory Press).

Sandmann, G. (1991). Biosynthesis of cyclic carotenoids: Biochemistry and molecular genetics of the reaction sequence. Physiol Plant *83*, 186-193.

Sandmann, G. (1993). Carotenoid analysis in mutants from *Escherichia coli* transformed with carotenogenic gene cluster and *Scenedesmus obliquus* mutant C-6D. Methods Enzymol *214*, 341-347.

Sandmann, G. (1994). Carotenoid biosynthesis in microorganisms and plants. Eur J Biochem *223*, 7-24.

SanMiguel, P., Tikhonov, A., Jin, Y. K., Motchoulskaia, N., Zakharov, D., Melake-Berhan, A., Springer, P. S., Edwards, K. J., Lee, M., Avramova, Z., and Bennetzen, J. L. (1996). Nested retrotransposons in the intergenic regions of the maize genome. Science *274*, 765-768.

Schaller, A., and Ryan, C. A. (1996). Systemin--a polypeptide defense signal in plants. Bioessays *18*, 27-33.

Schledz, M., al-Babili, S., von Lintig, J., Haubruck, H., Rabbani, S., Kleinig, H., and Beyer, P. (1996). Phytoene synthase from Narcissus pseudonarcissus: functional expression, galactolipid requirement, topological distribution in chromoplasts and induction during flowering. Plant J *10*, 781-92.

Schmidt, E. D., Guzzo, F., Toonen, M. A., and de Vries, S. C. (1997). A leucine-rich repeat containing receptor-like kinase marks somatic plant cells competent to form embryos. Development *124*, 2049-2062.

Schneider, R., and Schweiger, M. (1991). A novel modular mosaic of cell adhesion motifs in the extracellular domains of the neurogenic *trk* and *trkB* tyrosine kinase receptors. Oncogene *6*, 1807-1811.

Schulze-Muth, P., Irmler, S., Schroder, G., and Schroder, J. (1996). Novel type of receptor-like protein kinase from a higher plant (*Catharanthus roseus*). cDNA, gene, intramolecular autophosphorylation, and identification of a threonine important for auto- and substrate phosphorylation. J Biol Chem *271*, 26684-26689.

Schwartz, S. H., Tan, B. C., Gage, D. A., Zeevaart, J. A., and McCarty, D. R. (1997). Specific oxidative cleavage of carotenoids by VP14 of maize. Science *276*, 1872-1874.

Schwender, J., Seemann, M., Lichtenthaler, H. K., and Rohmer, M. (1996). Biosynthesis of isoprenoids (carotenoids, sterols, prenyl side-chains of chlorophylls and plastoquinone) via a novel pyruvate/glyceraldehyde 3-phosphate non-mevalonate pathway in the green alga *Scenedesmus obliquus*. Biochem J *316*, 73-80.

Scolnik, P. A., and Bartley, G. E. (1993). Phytoene desaturase from *Arabidopsis*. Plan Physiol *103*, 1745.

Serrano, A., Gimenez, P., Schmidt, A., and Sandmann, G. (1990). Immunocytochemical localization and functional determination of phytoene desaturase in photoautotrophic prokaryotes. J Gen Microbiol *136*, 2465-2469.

Siefermann-Harms, D., Fritz, B., and Ninnemann, H. (1985). Evidence for a pterin-derivative associated with the molybdenum cofactor of *Neurospora crassa* nitrate reductase. Photochem Photobiol *42*, 771-778.

Simon, M. N., and Studier, F. W. (1973). Physical mapping of the early region of bacteriophage T7 DNA. J Mol Biol *79*, 249-265.

Sommer, A. (1988). Avoidable blindness. Aust N Z J Ophthalmol *16*, 31-35.

Sommer, A. (1989). Large dose vitamin A to control vitamin A deficiency. Int J Vitam Nutr Res Suppl *30*, 37-41.

Song, W. Y., Wang, G. L., Chen, L. L., Kim, H. S., Pi, L. Y., Holsten, T., Gardner, J., Wang, B., Zhai, W. X., Zhu, L. H., Fauquet, C., and Ronald, P. (1995). A receptor kinase-like protein encoded by the rice disease resistance gene, *Xa21*. Science *270*, 1804-1806.

Springer, B., Werr, W., Starlinger, P., Bennett, D. C., Zokolica, M., and Freeling, M. (1986). The *Shrunken* gene on chromosome 9 of *Zea mays* L is expressed in various plant tissues and encodes an anaerobic protein. Mol Gen Genet *205*, 461-468.

Sun, Z., Gantt, E., and Cunningham, F. X. J. (1996). Cloning and functional analysis of the β-carotene hydroxylase of *Arabidopsis thaliana*. J Biol Chem *271*, 24349-24352.

Tomkins, A. (1991). Recent developments in the nutritional management of diarrhoea. 1. Nutritional strategies to prevent diarrhoea among children in developing countries. Trans R Soc Trop Med Hyg *85*, 4-7.

Torii, K. U., Mitsukawa, N., Oosumi, T., Matsuura, Y., Yokoyama, R., Whittier, R. F., and Komeda, Y. (1996). The *Arabidopsis ERECTA* gene encodes a putative receptor protein kinase with extracellular leucine-rich repeats. Plant Cell *8*, 735-746.

Treharne, K. J., Mercer, E. I., and Goodwin, T. W. (1966). Carotenoid biosynthesis in some maize mutants. Phytochem *5*, 581-587.

Valon, C., Smalle, J., Goodman, H. M., and Giraudat, J. (1993). Characterization of an *Arabidopsis thaliana* gene (*TMKL1*) encoding a putative transmembrane protein with an unusual kinase-like domain. Plant Mol Biol *23*, 415-21.

van de Sande, K., Pawlowski, K., Czaja, I., Wieneke, U., Schell, J., Schmidt, J., Walden, R., Matvienko, M., Wellink, J., van Kammen, A., Franssen, H., and Bisseling, T. (1996). Modification of phytohormone response by a peptide encoded by ENOD40 of legumes and a nonlegume. Science *273*, 370-373.

van der Knaap, E., Sauter, M., Wilford, R., and Kende, H. (1996). Identification of a gibberellin-induced receptor-like kinase in deepwater rice. Plant Physiol *112*, 1397.

von Lintig, J., Welsch, R., Bonk, M., Giuliano, G., Batschauer, A., and Kleinig, H. (1997). Light-dependent regulation of carotenoid biosynthesis occurs at the level of phytoene synthase expression and is mediated by phytochrome in *Sinapis alba* and *Arabidopsis thaliana* seedlings. Plant J *12*, 625-634.

Waegemann, K., and Soll, J. (1996). Phosphorylation of the transit sequence of chloroplast precursor proteins. J Biol Chem *271*, 6545-6554.

Walker, J. C. (1993). Receptor-like protein kinase genes of *Arabidopsis thaliana*. Plant J *3*, 451-456.

Walker, J. C. (1994). Structure and function of the receptor-like protein kinases of higher plants. Plant Mol Biol *26*, 1599-1609.

Walker, J. C., and Zhang, R. (1990). Relationship of a putative receptor protein kinase from maize to the S- locus glycoproteins of *Brassica*. Nature *345*, 743-746.

Wang, X., Zafian, P., Choudhary, M., and Lawton, M. (1996). The PR5K receptor protein kinase from *Arabidopsis thaliana* is structurally related to a family of plant defense proteins. Proc Natl Acad Sci USA *93*, 2598-2602.

Weedon, B. C. L., and Moss, G. P. (1995). Structure and nomenclature. In Carotenoids, G. Britton, S. Liaaen-Jensen and H. Pfander, eds. (Basel: Birkhauser Verlag).

Weinstein, J. N., Blumenthal, R., van Renswoude, J., Kempf, C., and Klausner, R. D. (1982). Charge clusters and the orientation of membrane proteins. J Membr Biol *66*, 203-212.

Weis, E., and Berry, J. A. (1988). Plants and high temperature stress. Symp Soc Exp Biol *42*, 329-346.

Williamson, M. P. (1994). The structure and function of proline-rich regions in proteins. Biochem J *297*, 249-260.

Wurtzel, E. T. (1992). Use of a *Ds* chromosome breaking element to examine maize *Vp5* expression. J Hered *83*, 109-113.

Wurtzel, E. T., Burr, F. A., and Burr, B. (1987). DNase I hypersensitivity and expression of the *Shrunken-1* gene of maize. Plant Mol Biol *8*, 251-264.

Wurtzel, E. T., Li, Z.-h., Luo, R., Matias, D., Mozoub, D., Matthews, P. D., Upasani, V. N., Valdez, G., Yoganathan, A., and Yu, J. (1996). Research towards improvement of the pro-vitamin A (carotenoid) content of rice endosperm. Intern Rice Res Notes *21*, 43-44.

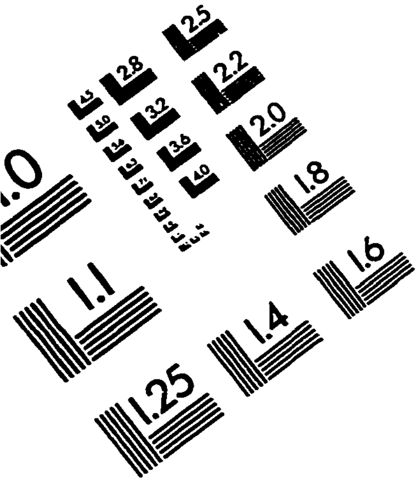
Young, A. J. (1991). The photoprotective role of carotenoids in higher plants. Physiol Plant *83*, 702-708.

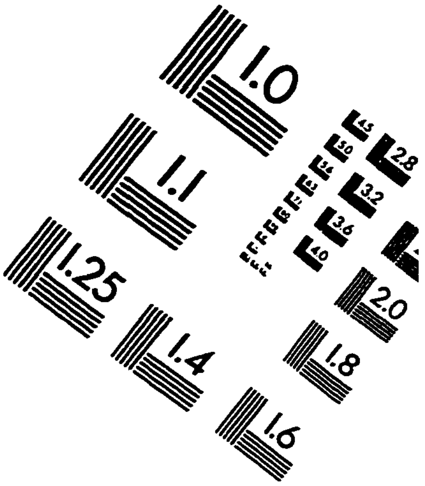
Young, A. J., and Frank, H. A. (1996). Energy transfer reactions involving carotenoids: quenching of chlorophyll fluorescence. J Photochem Photobiol B *36*, 3-15.

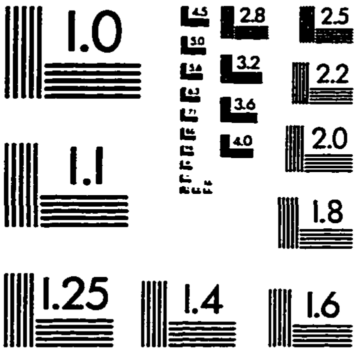
Zeevaart, J. A. D., and Creelman, R. A. (1988). Metabolism and physiology of abscisic acid. Annu Rev Plant Physiol Plant Mol Biol *39*, 439-473.

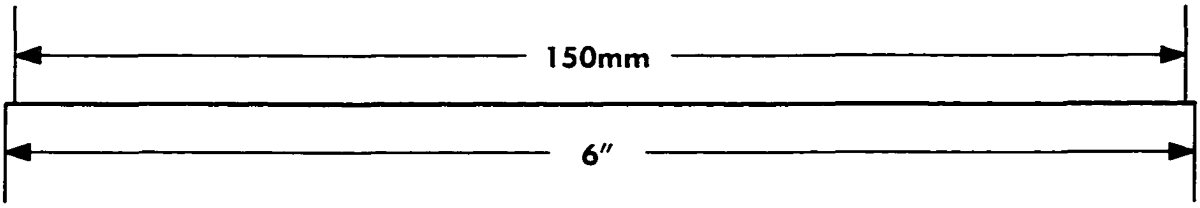
Zhou, J., Loh, Y. T., Bressan, R. A., and Martin, G. B. (1995). The tomato gene *Pti1* encodes a serine/threonine kinase that is phosphorylated by Pto and is involved in the hypersensitive response. Cell *83*, 925-935.

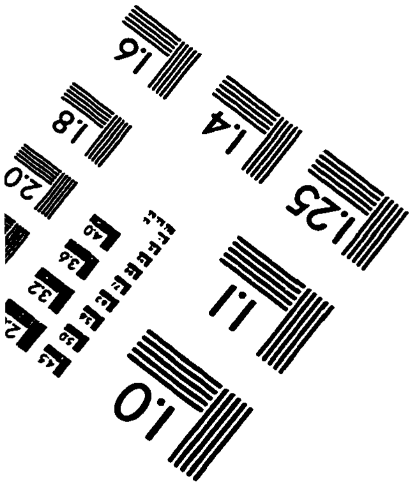
IMAGE EVALUATION TEST TARGET (QA-3)

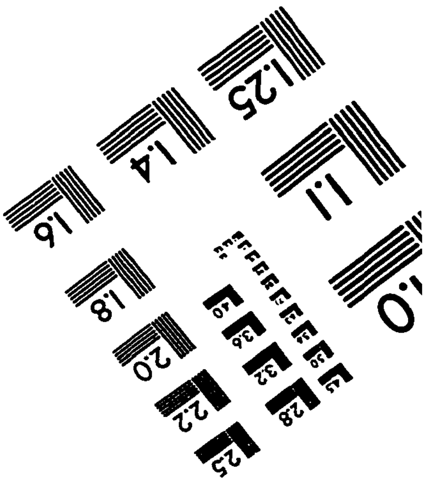












APPLIED IMAGE, Inc
1653 East Main Street
Rochester, NY 14609 USA
Phone: 716/482-0300
Fax: 716/288-5989