DA
1328 HG
1994

STUDIES ON THE STRUCTURE AND FUNCTION OF THE MAP KINASE FAMILY IN *ARABIDOPSIS THALIANA*

Tsuyoshi MIZOGUCHI

January, 1995

Submitted in partial fulfillment of the requirements for the degree of **Doctor of Philosophy in Sciences**, in Doctoral program in Biological Sciences, University of Tsukuba.

寄贈
平成 年 月 日
Tsuyoshi MIZOGUCHI 氏

95314100

Table of Contents

		page
Contents	----	i
Abbreviations	----	v
Abstract	----	1
Introduction	----	5
Materials and Methods	----	11
Preparation of a cDNA library and DNA templates for the polymerase chain reaction (PCR)	----	12
PCR	----	12
Cloning and sequencing of cDNAs encoding putative components of MAP kinase cascades in Arabidopsis	----	13
Plant growth and stress treatments	----	14
RNA and DNA blot hybridization analyses	----	15
Construction of pGEX-MPK1 and pGEX-MPK2 and expression of the recombinant ATMPK1 and ATMPK2 proteins in Escherichia coli	----	16
Culturing of tobacco cells and preparation of cell extracts	----	17
Detection of the MBP kinase activity in extracts of 2,4-D treated tobacco BY-2 cells	----	18
Activation of recombinant ATMPK1 and ATMPK2 proteins by the purified Xenopus MAP kinase kinase	----	18
cDNA clones	----	19

Results ---- 20

Part I.

Isolation of cDNAs that encode plant protein kinases using PCR ---- 21

Amplification of plant protein kinase cDNA fragments using PCR and analysis of PCR-generated products ---- 22

Part II.

Characterization of two cDNA clones that encode MAP kinases in *Arabidopsis* ---- 24

Cloning and sequence analysis of two cDNA clones that encode MAP kinases in Arabidopsis ---- 24

Primary structure of the putative ATMPK1 and ATMPK2 proteins ---- 25

Analysis by Northern and Southern blotting of the ATMPK1 and ATMPK2 genes ---- 27

Phosphorylation and activation of the ATMPK1 and ATMPK2 proteins expressed in Escherichia coli cells by Xenopus MAP kinase kinase ---- 28

Activation by auxin of the MBP kinase activity in auxin -starved tobacco BY-2 cells ---- 29

Identification of *Arabidopsis* MAP kinase family ---- 32

Cloning and sequence analysis of seven additional cDNA clones that encode MAP kinases in Arabidopsis ---- 32

Primary structure of the putative ATMPK proteins ---- 33

Phylogenetic tree analysis ---- 35

Identification of the putative upstream factors of MAP kinases in *Arabidopsis* ---- 36

Cloning and sequence analysis of cDNAs for protein kinases ---- 37

related to MAPKKKs

Primary structure of the putative ATMEKK1 protein ---- 37

Analysis by Southern blotting of the ATMEKK1 gene ---- 38

Identification of the putative downstream factors of MAP kinases in *Arabidopsis* ---- 40

Cloning and sequence analysis of two cDNAs for protein kinases related to ribosomal S6 kinases (RSKs) ---- 40

Primary structure of the putative ATPK6 and ATPK19 proteins ---- 40

Analysis by Southern blotting of the ATPK6 and ATPK19 genes ---- 42

Part III.

Effect of a variety of stresses on the expression of genes for putative components of *Arabidopsis* MAP kinase cascades ---- 43

ATMPK gene expression is induced after a variety of stimuli ---- 43

The mRNA level of ATMEKK1 is affected by a variety of stimuli ---- 44

The mRNA levels of ATPK6 and ATPK19 are affected by a variety of stimuli ---- 44

The transcription levels of three protein kinases, ATMPK3, ATPK19, and ATMEKK1 increase simultaneouly in response to a variety of stresses ---- 46

Induction of the ATMPK3, ATPK19, and ATMEKK1 genes following touch stimulation, salinity stress, and low temperature ---- 46

Discussion ---- 49

Conclusion ---- 64

Figures and Tables ---- 67

References ---- 112

Acknowledgments ---- 133

Abbreviations

ABA	:abscisic acid
ATMPKs	:*Arabidopsis thaliana* MAP kinases
ATMPKs	:genes that encode ATMPKs
b	:nucleotide bases
bp	:base pairs
BY-2	:*Nicotiana tabacum* 'Bright Yellow 2'
cATMPKs	:cDNA clones that encode ATMPKs
CBB	:Coomassie Brilliant Blue
cAPK	:cyclic AMP-dependent protein kinase
cDNA	:complementary DNA
cGPK	:cyclic GMP-dependent protein kinase
2, 4 - D	:2, 4 - dichlorophenoxyacetic acid
DNA	:deoxyribonucleic acid
DTT	:dithiothreitol
EDTA	:ethylenediaminetetraacetic acid
g	:centrifugal force (x unit gravitational field)
GST	:glutathione *S*-transferase
IPTG	:isopropyl-β-D-thiogalactopyranoside
kb	:kilo base
kbp	:kilo base pairs
KDa	:kilo Dalton
Mw	:molecular weight
mRNA	:messenger RNA
MAPK	:MAP kinase
MAP kinase	:mitogen-activated protein kinase
MAPKK	:MAP kinase kinase
MAPKKK	:MAP kinase kinase kinase
MBP	:myelin basic protein
MPF	:M phase-promoting factor
MS medium	:Murashige and Skoog's medium
$p70^{S6K}$	:p70 ribosomal S6 kinase
PAGE	:polyacrylamide gel electrophoresis
PCR	:polymerase chain reaction
PKC	:protein kinase C
PMSF	:phenylmethylsulfonyl fluolide
$pp90^{rsk}$	:p90 ribosomal S6 kinase
RNA	:ribonucleic acid
SDS	:sodium dodecyl sulfate
Tris	:tris(hydroxymethyl) aminomethane

Abstract

I have isolated several PCR-amplified DNA fragments which are different from each other and contain partial sequences of protein kinases from *Arabidopsis thaliana*. One of these amplified DNA sequences has high homology with those of MAP kinases in animals and yeasts.

Using the PCR-amplified DNA as a probe, two cDNA clones, cATMPK1 and cATMPK2, encoding MAP kinases (mitogen-activated protein kinases) have been cloned from *Arabidopsis thaliana*, and their nucleotide sequences have been determined. Putative proteins encoded by *ATMPK1* and *ATMPK2*, designated ATMPK1 and ATMPK2, contain open reading frames of 1110 and 1128 base pairs, respectively, and are 88.7% identical at the amino acid sequence level. ATMPK1 and ATMPK2 exhibit significant similarity to rat ERK2 (49%) and *Xenopus* MAP kinase (50%). The amino acid residues corresponding to the sites of phosphorylation (Thr-Glu-Tyr) that are involved in the activation of MAP kinases are conserved in ATMPK1 and ATMPK2. Northern blot analysis indicates that the mRNAs that correspond to cATMPK1 and cATMPK2 are significantly present in all the organs except seeds. Genomic Southern blot analysis suggests that there are a few additional genes that are related to ATMPK1 and ATMPK2 in the *Arabidopsis* genome. Purified *Xenopus* MAP kinase kinase (MAPKK) phosphorylates tyrosine as well as threonine/serine residues in ATMPK1 and ATMPK2 proteins that have been expressed in *Escherichia coli*, activating these enzymes. Analysis of the activation of MAP kinase using tobacco BY-2 cells demonstrates that the recombinant ATMPK1 and ATMPK2 proteins become more active as a result of phosphorylation upon incubation with extracts prepared from tobacco BY-2 cells that have been treated with the synthetic auxin, 2,4-dichlorophenoxyacetic acid (2,4-D). These results suggest that auxin may

function as an activator of plant MAP kinases, as do various mitogens in animal system.

Next I isolated seven additional cDNAs encoding novel MAP kinases in *Arabidopsis*, cATMPK3 to 9, by screening three cDNA libraries under low-stringency conditions. The amino acid residues corresponding to the site of phosphorylation (Thr-Glu-Tyr) that are involved in the activation of MAP kinases are conserved in all the nine putative ATMPK proteins. Genes for MAP kinases in *Arabidopsis* constitute a family that contains more than nine members. Sequence analysis suggests that there are at least four subfamilies in the family of *Arabidopsis* genes for MAP kinases.

Then I isolated the putative upstream factors of *Arabidopsis* MAP kinases (ATMPK1 to 9). The putative catalytic domain of ATMEKK1 shows approximately 40% identity with the amino acid sequences of the catalytic domains of MAP kinase kinase kinases (MAPKKKs) (byr2 from *Schizosaccharomyces pombe*, STE11 from *Saccharomyces cereviciae*, BCK1 from *S. cereviciae*, MEKK from mouse, and NPK1 from tobacco). Genomic Southern blot analyses suggested that ATMEKK genes constitute a small gene family.

Concerning the putative downstream factors of *Arabidopsis* MAP kinases (ATMPK1 to 9), I isolated two closely related cDNAs, cATPK6 and cATPK19, which are most related (approximately 52% identity) to ribosomal S6 kinases, $p70^{S6K}$ and $pp90^{rsk}$, in vertebrates. ATPK6 and ATPK19 have the PEST regions in their N- and C-terminal, and several phosphorylation consensus sites recognized by casein kinase II or proline-directed protein kinases such as cdc2, MAP kinase, and p54 MAP-2 kinase (SAPK) within these regions.

Expression of the MAP kinase genes (*ATMPK1* to *9*) was demonstrated to be positively and negatively regulated in plants under a variety of stimuli or stress conditions. The mRNA levels of three protein kinases, ATMPK3, ATPK19, and ATMEKK1, which are structurally related to MAP kinases, ribosomal S6 kinases, and MAPKKKs, respectively, increased markedly and simultaneously in response to touch, cold, and salinity stresses. These results suggest that MAP kinase cascades, which are thought to respond to a variety of extracellular signals, are regulated not only at the post-translational level but also at the transcriptional level in plants. It is also suggested that MAP kinase cascades in plants may function in the transduction of signals under environmental stresses.

Introduction

Plants are not mobile and adapt themselves to environmental changes. They are thought to develop unique systems to respond a variety of environmental stimuli on the process of evolution. Therefore one of the purposes of present study is to understand how these environmental stimuli are perceived and modulated to the signals that control gene expression and enzyme activities in plants. A number of reports on isolation and characterization of genes that are transcriptionally regulated by signals including changes in light, temperature, and osmolarity has been increasing year by year (Wray *et al.*, 1992; Thomashow *et al.*, 1990; Yamaguchi-Shinozaki *et al.*, 1992; Gallie *et al.*, 1993; Chandler *et al.*, 1994). Many cDNAs encoding putative trans-acting factors have been isolated from plants mainly by the use of three different experimental approaches: mutant studies, DNA-protein interaction, and homology to known transcription factors (Katagiri *et al.*, 1992). However, we still have little knowledge on the molecular process of the regulation of these transcription factors that may play key roles in the signal transduction pathways under environmental stress conditions.

Protein kinases have been shown to play important roles in the transduction of environmental and developmental signals in animals and yeasts (Hanks *et al.*, 1988; Hanks and Quinn, 1991). Many transcription factors are shown to be regulated by protein kinases in these organisms. In plants, the transduction of hormone (Daminov *et al.*, 1992), light (Datta *et al.*, 1989; Sarokin *et al.*, 1992), low-temperature (Monroy *et al.*, 1993), and pathogen-elicitor (Felix *et al.*, 1991) stimuli has been reported to be involved protein phosphorylation, although what kind of protein kinases function in these pathways largely remain to be elucidated.

Molecular cloning of genes which encode protein kinases from animals and yeasts, and the disruption of the genes in the organisms have provided us much information about their functions *in vivo*. Therefore I think it important to identify a variety of genes for protein kinases from plants in order to understand the molecular mechanisms of signal transduction by phosphorylation in plants. Since protein kinases have highly conserved regions in their catalytic domains (Hanks *et al.*, 1988; Hanks and Quinn, 1991), PCR has been frequently used, with degenerate oligonucleotide that correspond to such conserved regions, to amplify particular target DNA sequences (Fig. 1). Many cDNA and genomic DNA clones encoding a variety of protein kinases have been isolated using this strategy from a number of different plant species (Walker and Zhang, 1990; Feiler and Jacobs, 1990; Hirt *et al.*, 1991; Hashimoto *et al.*, 1992; Anderberg and Walker-Simmons, 1992; Hayashida *et al.*, 1992, 1993; Mizoguchi *et al.*, 1992; Banno *et al.*, 1993). I have isolated several PCR-ampified fragments of DNA that contain partial sequences of different protein kinases from *Arabidopsis* (Fig. 1 and Table 1). These amplified sequences exhibit high degree of homology to those of genes for MAP kinases (Mizoguchi *et al.*, 1993b; 1994a), ribosomal S6 kinases, casein kinase II (Mizoguchi *et al.*, 1993a), calmodulin dependent protein kinase II (Mizoguchi *et al.*, 1994b), and SNF1.

In vertebrates, the functional kinase units composed of three protein kinases (MAPKKK, MAPKK, and MAP kinase) integrate mitogenic signals from a large number of growth factor receptors, cytokine receptors, several membrane-bound oncogenes, and some mitogenic peptides with G protein-linked receptors (Sturgill and Wu, 1991; Cobb *et al.*, 1991; Thomas, 1992; Nishida and Gotoh, 1992; Pelech and Sanghera, 1992; Nishida and Gotoh,

1993). MAP kinases, a family of 40 to 45 KDa serine/threonine kinases, are activated by a variety of mitogenic stimuli and play key roles not only in growth factor- and differentiating factor-stimulated signalling pathways but also in the phosphorylation cascade that is induced by MPF (M phase-promoting factor) (Sturgill and Wu, 1991; Nishida and Gotoh, 1993; Pelech and Sanghera, 1992). MAP kinases have been identified in *Xenopus* (Gotoh *et al.*, 1991; Posada *et al.*, 1991), rat (Boulton *et al.*, 1990; Boulton *et al.*, 1991), human (Gonzalez *et al.*, 1992) and mouse (Her *et al.*, 1991). MAP kinases have also been shown to directly phosphorylate several transcription factors including Myc (Alvarez *et al.*, 1991; Seth *et al.*, 1991), Elk (Gille *et al.*, 1992), Jun (Pulverer *et al.*, 1991). Thus one of their functions is thought to integrate multiple intracellular signals transmitted by various second messengers, and regulate gene expressions by phosphorylation of trancsription factors. MAP kinases are activated by phosphorylation of threonine and tyrosine residues by upstream kinases MAPKKs that are themselves regulated by serine and threonine phosphorylation by their upstream kinases MAPKKKs (Sturgill and Wu, 1991; Cobb *et al.*, 1991; Thomas, 1992; Nishida and Gotoh, 1992; Pelech and Sanghera, 1992; Nishida and Gotoh, 1993; Neiman, 1993). At least three protein kinases, Raf (Dent *et al.*, 1992; Howe *et al.*, 1992; Kyriakis*et al.*, 1992), Mos (Posada *et al.*, 1993), and MEKK (Lange-Carter, 1993), function as MAPKKKs in animals. It has also been reported that MAP kinases directly phosphorylate not only transcription factors, such as Jun, Fos, and Myc, but also two protein kinases, $pp90^{rsk}$ (Erikson, 1991) and MAPKAP kinase-2 (Stokoe *et al.*, 1992), cytoplasmic phospholipase A_2 ($cPLA_2$) (Lin *et al.*, 1993), and microtubule associated protein, p220 (Shiina *et al.*, 1992).

In *S. cerevisiae*, three MAP kinase cascades function in different signaling pathways (Neiman *et al.*, 1993). The first MAP kinase cascade to be identified functions in the mating factor response pathway. The second one functions in a signal transduction pathway mediated by protein kinase C. The third one is required for response to increase in extracellular osmolarity. The phosphorylation cascades including MAPKKK, MAPKK, and MAP kinase are conserved among various signal transduction pathways from yeast to vertebrate (Sturgill and Wu, 1991; Cobb *et al.*, 1991; Thomas, 1992; Nishida and Gotoh, 1992; Pelech and Sanghera, 1992; Nishida and Gotoh, 1993; Neiman, 1993).

Thus, if we are to understand signal transduction pathways in plants, it seems to be important to ascertain whether MAP kinase cascades exist and function in the intracellular signal pathways of plants because these enzymes appear to play key roles in a wide variety of responses to extracellular signals in animals and yeasts. In this thesis, I first demonstrated that ATMPK1 and 2 are plant MAP kinases. It has also be shown that genes for MAP kinases in *Arabidopsis* constitute a family that contains more than nine members. Then I showed the isolation of the putative upstream and downstream factors, ATMEKK1-4 and ATPK6 and 19, which are structurally most related to MAPKKKs and ribosomal S6 kinases, respectively. Finally, I demonstrated that the mRNA levels of three protein kinases, ATMPK3, ATPK19, and ATMEKK1, which are structurally related to MAP kinases, ribosomal S6 kinases, and MAPKKKs, respectively, increased markedly and simultaneously in response to touch, cold, and salinity stresses. On the basis of the characterization of the putative components of MAP kinase cascades in plants, the function of

these protein kinases in plants under a variety of environmental stresses, was discussed.

Materials and Methods

Preparation of a cDNA library and DNA templates for the polymerase chain reaction (PCR)

Total RNA was isolated by the method of Nagy *et al.* (Nagy *et al.*, 1988) from whole plants of *Arabidopsis* (Columbia ecotype) that have been grown under continuous illumination of approximately 2,500 lux at 22°C for 4 to 5 weeks and harvested prior to bolting.

Poly(A)$^+$-RNA was prepared by the method of Maniatis *et al.* (Maniatis *et al.*, 1982). Double-stranded cDNA was synthesized from poly(A)$^+$-RNA using cDNA Synthesis System Plus from Amersham. A library was prepared from this cDNA using cDNA cloning system λgt11 (4 x 10^7 pfu/mg cDNA) from Amersham. The cDNA library was amplified once for the preparation of DNA templates for PCR.

PCR

DNA was prepared from an amplified phage library by phenol extraction and CsCl gradient centrifugation (Maniatis *et al.*, 1982). The DNA inserts were amplified by PCR with two oligonucleotide primers that corresponded to the boundary sequences of the arms of λgt11 (5'-GGTGGCGACGACTCCTGGAGCCCG-3' and 5'-TTGACACCAGACCAACTGGTAATG-3'). The oligonucleotide sequences shown in Table 1 were used as primers for isolation of PCR-amplified fragments of DNA that contained the partial sequences of genes for protein kinases from *Arabidopsis.*

All the primers were phosphorylated with T4 polynucleotide kinase under standard conditions to increase the efficiency of ligation (Maniatis *et*

al., 1982). PCR was performed using a GeneAmp kit according to the manufacturer's protocol (Perkin Elmer Cetus, Norwalk, CT, USA). Thirty-five cycles of denaturation, annealing and polymerization were carried out at 93°C for 1 min, at 37°C for 2 min and at 72°C for 3 min, respectively. The amplified DNA fragments, which ranged from 400 to 600 base pairs (bp) in length, were cloned into the pBluescript vector, pSKII$^-$, (Stratagene, La Jolla, CA, USA), digested with *Sma*I and dephosphorylated.

Cloning and sequencing of cDNAs encoding putative components of MAP kinase cascades in Arabidopsis

All the cDNA clones were isolated from two *Arabidopsis* cDNA libraries prepared from rosset plants grown for 4 weeks and from dehydrated rosset plants (Yamaguchi-Shinozaki *et al.*, 1992) by plaque hybridization, as described earlier (Mizoguchi *et al.*, 1993a; 1993b; 1994a; 1994b). The inserted DNA of the E68, the 23A12T7, and the C11 were used as probes to isolate cATMPK1 to 9 (Mizoguchi *et al.*, 1993b; 1994a), cATMEKK1 to 4, and cATPK6 and 19, respectively. The probes were labeled with ^{32}P using a random primer kit in accordance with the manufacturer's instruction (Boehringer Mannheim, Mannheim, Germany).

Positive clones were plaque-purified and the DNA was isolated by phenol extraction and CsCl gradient centrifugation (Maniatis *et al.*, 1982). Total of 8.0 x 10^5 , 8.0 x 10^5 , and 6.0 x 10^5 plaques were screened under high-stringency hybridization conditions to isolate cATMPK1 and 2 (Mizoguchi *et al.*, 1994a), cATMEKK1 to 4, and cATPK6 and 19, respectively. A total of 8.0 x 10^5 plaques were screened under low-stringency

hybridization conditions to isolate the seven cDNA clones, cATMPK3 to 9 (Mizoguchi *et al.*, 1993b).

Filters were prehybridized (8 h) and hybridized (18 h) with either 50% (high-stringency conditions) or 30% (low-stringency conditions) formamide, 100 μg/ml sonicated salmon sperm testes DNA, 5x Denhardt's solution, 50 mM sodium phosphate (pH 6.5), 5x SSC, and 0.2% SDS at 42°C. The filters were washed twice in either 0.1x SSC, 0.1% SDS at 65°C (high-stringency conditions) or 0.5x SSC, 0.5% SDS at 37°C (low-stringency conditions), then dried and exposed overnight to X-ray film with an intensifying screen. Positive clones were classified into the distinct groups based on restriction endonuclease mapping and partial sequencing analyses.

The longest inserted DNA fragments of each clone were subcloned into the pBluescript vector, pSKII$^-$ (Stratagene). A DNA sequencer model 373A (Applied Biosystems, San Jose, CA, U.S.A.) was used for DNA sequencing. Nucleotide and amino acid sequences were analyzed using the GENETYX software system (Software Development Co., Tokyo, Japan).

Plant growth and stress treatments

Arabidopsis (Columbia ecotype) plants were grown on GM agar plates under continuous illumination of - 2500 lux at 22°C for 4 to 5 weeks and used in stress treatment experiments prior to bolting as reported previously (Yamaguchi-Shinozaki *et al.*, 1992; 1994).

Plants subjected to treatment with abscisic acid (ABA), 2,4-D (2,4-D) and salt stress (NaCl) were grown hydroponically in a solution of 100 μM ABA, 100 μM 2,4-D, and 250 mM NaCl, respectively, under dim light. Heat treatment (Heat) was performed under continuous light by exposure of

plants grown at 22°C to a temperature of 40°C, or grown at 22°C and then transferred from GM agar plates, grown hydroponically in 40°C water. Plants subjected to treatments with water (Water) or cold stress (Cold) were grown hydroponically under dim light in 22°C or 4°C water, respectively. Desiccation treatment (Desiccation) was conducted as previously reported (Yamaguchi-Shinozaki *et al.*, 1994). Touch treatment (Touch) was conducted as previously reported (Braam and Davis, 1990). An gloved hand was placed gently touching the rosette leaves, and then the hand was moved so as to move the leaves back and forth 20 times. In each case, the plants were subjected to the stress treatments for various periods and frozen in liquid nitrogen.

RNA and DNA blot hybridization analyses

Genomic Southern hybridization was performed as previously described (Maniatis *et al.*, 1982).

Northern hybridization was carried out by the method of Maniatis *et al.* (Maniatis *et al.*, 1982). Total RNA (40 μg) was subjected to electrophoresis on an 1% agarose gel that contained formaldehyde (Maniatis *et al.*, 1982). Nitrocellulose or nylon filters blotted with the RNA were hybridized with [^{32}P]-labeled DNA probes that included cDNAs in 50% formamide, 5x SSC, 25 mM sodium phosphate buffer (pH 6.7), 10x Denhardt's solution and 250 μg/ml denatured salmon sperm DNA (Maniatis *et al.*, 1982). Nitrocellulose or nylon filters blotted with RNA were stained with methylene blue (Maniatis *et al.*, 1982). The hybridization signals corresponding to *ATMPK1* to *9* (Mizoguchi *et al.*, 1993b; 1994a), *ATMEKK1*, *ATPK6* and *19*, *cdc2a* (Hirayama *et al.*, 1991), and *Tchγ*

mRNAs were quantified with a Bio Image Analyzer (FUJIX BAS2000, Fuji Film Co., Ltd., Tokyo, Japan).

Construction of pGEX-MPK1 and pGEX-MPK2 and expression of the recombinant ATMPK1 and ATMPK2 proteins in Escherichia coli

A 1.5-kbp *Bam*HI fragment containing *ATMPK1* cDNA and a 1.6-kbp *Eco*RI fragment containing *ATMPK2* cDNA were cloned into the *Bam*HI and *Eco*RI sites of the pGEX vector (Smith and Johnson, 1988) and then used to transform *E. coli* JM109 cells. The constructs were named pGEX-MPK1 and pGEX-MPK2, respectively (Mizoguchi *et al.*, 1994a). Expression of the fused ATMPK1 and ATMPK2 proteins in *E. coli* was achieved by the method described previously (Smith and Johnson, 1988) with some modifications.

Overnight cultures of *E. coli* cells that has been transformed with pGEX-MPK1 or pGEX-MPK2 were diluted 1:10 in 800 ml of 2YT medium (16 g Bacto-tryptone, 10 g Bacto-yeast extract, 5 g NaCl/liter) that contained ampicillin (100 μg/ml) and grown for 1 h at 37°C before addition of IPTG to 0.1 mM.

After a further 7 h of growth, cells were pelleted and resuspended in 1/100 culture volume of MTPBS (150 mM NaCl, 16 mM Na_2HPO_4, 4 mM NaH_2PO_4, pH 7.3). Cells were lysed on ice by mild sonication and, after addition of Triton X-100 to 1% (v/v), the lysate was subjected to centrifugation at 10,000 x *g* for 5 min at 4 °C.

The supernatant was mixed at room temperature in a 50-ml polypropylene tube with 2 ml of 50% glutathione-Sepharose beads (Pharmacia, Uppsala, Sweden). After absorption had been allowed to

proceed for 2 min, the beads were collected by a brief centrifugation at 500 x g and washed three times with 50 ml of MTPBS. Fusion proteins were eluted by washing of beads with free glutathione. Two 2-min washes were performed with a volume of 50 mM Tris-HCl (pH 8.0) that contained 5 mM reduced glutathione (Sigma, St. Lous, MO, USA; final pH 7.5; freshly prepared) equivalent to the volume of the beads.

Culturing of tobacco cells and preparation of cell extracts

Tobacco cells, BY-2 (*Nicotiana tabacum* 'Bright Yellow 2'), in suspension culture, were kindly provided by Dr. T. Nagata of University of Tokyo. BY-2 cells were maintained in Linsmaier and Skoog's medium that had been modified by increasing the concentration of KH_2PO_4 to a final concentration of 370 mg/l. This modified medium was supplemented with 3% sucrose and 1 μM 2,4-D. Cells were subcultured every 7 days (Nagata *et al.*, 1981).

The cells were rendered quiescent by incubation in 2,4-D free medium for 5 days (Mizoguchi *et al.*, 1994a). The cells were further incubated with or without 1 μM 2,4-D at 24°C for the indicated periods of time, then they were rapidly collected by aspiration and frozen in liquid nitrogen. The cells were homogenized in 10 ml of an ice-cold extraction buffer solution which consisted of 20 mM Hepes (pH 7.6), 100 mM KCl, 1 mM magnesium acetate, 2 mM $CaCl_2$, 1 mM DTT, 1 mM PMSF and 20 mM β-glycerophosphate. The homogenate was centrifuged first at 30,000 x *g* for 30 min, and then the supernatant was centrifuged at 100,000 x *g* for 4 h at 4°C. The final supernatant was concentrated and used immediately as the cell extract.

Detection of the MBP kinase activity in extracts of 2,4-D treated tobacco BY-2 cells

Extracts of BY-2 cells (50 μg of protein), prepared as described above, were subjected to the kinase-detection assay on gels that contained myelin basic protein (MBP) as described previously (Mizoguchi *et al.*, 1993; Geahlen *et al.*, 1986; Gotoh *et al.*, 1990). In this assay, MBP kinase activity was measured in terms of the phosphorylation of MBP in a polyacrylamide gel that contained 0.5 mg/ml MBP, after denaturation and renaturation that followed fractionation by SDS-PAGE of the sample.

Activation of recombinant ATMPK1 and ATMPK2 proteins by the purified Xenopus MAP kinase kinase

To detect the phosphorylation of recombinant ATMPK1 and ATMPK2 proteins, GST-ATMPK1 or GST-ATMPK2 was incubated for 30 min with the purified *Xenopus* MAPKK (Matsuda *et al.*, 1992; Kosako *et al.*, 1992; Kosako *et al.*, 1993) or the extracts of tobacco BY-2 cells that had been treated with 2,4-D, prepared as described above, in a solution that consisted of 20 mM Tris-HCl (pH 7.5), 20 mM $MgCl_2$ and 100 μM [γ-^{32}P]ATP (1 μCi; Amersham) in a final volume of 15 μl. Samples were subjected to electrophoresis and autoradiography.

To detect the MBP kinase activity of the activated GST-ATMPK1 and GST-ATMPK2 proteins, GST-ATMPK1 and GST-ATMPK2 were preincubated with *Xenopus* MAPKK or extracts of tobacco cells that had been treated with 2,4-D for 5 min, in the above reaction mixture, without radiolabeled ATP. MBP (0.5 mg/ml) and [γ-^{32}P]ATP (1 μCi) were added to

the reaction mixture, and then the incubation was continued for another 30 min. The reaction was stopped by the addition of Laemmli's sample buffer (Laemmli, 1970) and boiling. Labeled proteins were analyzed by SDS-PAGE and subsequent autoradiography. For the quantification of the labeled proteins, the Bio Image Analyzer (Fuji Film) was used.

cDNA clones

The clone 23A12T7 (GenBank T04696) was obtained from the expressed sequence tag (EST) clones, isolated and partially sequenced by Dr. Newman, Michigan State University. The EST clone was supplied by the Arabidopsis Biological Research Center at Ohio State. The clone 23A12T7 contains partial sequences (approximately 700 bp) for protein kinases.

The plasmid DNA clone containing the cdc2a cDNA (Hirayama *et al.*, 1991) is a kind gift from Dr. Hirayama and Dr. Oka, University of Kyoto.

The plasmid DNA clone (designated Tchγ3) containing the 3'-noncoding region of Tchγ is a kind gift from Dr. Okada, National Institute for Basic Biology. The *Arabidopsis* Tchγ encodes calmodulin (Braam and Davis, 1990) and the mRNA level increases in response to water-spray (Itoh and Okada, personal communication).

Results

Part I.

Isolation of cDNAs that encode plant protein kinases using PCR

All the eukaryotic protein kinases, of which amino acid sequences have been reported, contain a homologous region of about 250 - 300 amino acid residues (Fig. 1) (Hanks *et al.*, 1988; Hanks and Quinn, 1991). This region is responsible for the phosphotransferase activity and is commonly referred to as the catalytic domain. Amino acid sequence alignment of protein kinase catalytic domains have revealed several residues that are invariant or nearly invariant throughout the eukaryotic protein kinase family (Hanks *et al.*, 1988; Hanks and Quinn, 1991). This alignment further demonstrated the existence of 11 highly conserved subdomains (Fig. 1). Many of the subdomains contain five or more contiguous amino acid residues that are nearly invariant among the members of certain subgroups of the protein kinases. These patches of homology are excellent candidates for the generation of degenerate oligonucleotide primers for use in the polymerase chain reaction (PCR).

At first, I used two DNA sequences encoding the conserved regions VI and VIII (Hanks *et al.*, 1988; Hanks and Quinn, 1991) to carry out PCR using *Arabidopsis* genomic DNA as a template ((1) in Fig. 1). A 380-bp PCR-amplified DNA fragment was obtained and used as a probe to screen a cDNA library prepared from the rosettes of *Arabidopsis* plants which had not yet bolted. In this screening, I could isolate only one type of cDNA clone, named cATPK64 (Mizoguchi*et al.*, 1992). ATPK64 has 36% to 40% sequence identity to all the members of the cyclic nucleotide- and calcium-phospholipid-dependent subfamilies of serine/threonine kinases in animals

and yeasts (Hanks *et al.*, 1988; Hanks and Quinn, 1991). The putative catalytic domain of ATPK64 is most similar to that of PVPK-1 (73% identity) which was the first cDNA encoding a protein kinase in plants (Lawton *et al.*, 1989).

At that time, there were the reports on isolation of only two subfamilies of protein kinases, cdc2 (Hanks *et al.*, 1988; Hanks and Quinn, 1991; Hirayama *et al.*, 1991) and PVPK-1 (Lawton *et al.*, 1989) subfamilies, in plants. In order to learn more about the structure and the role of protein kinases as signal transducers in plants, I carried out PCR using the two DNA sequences encoding the conserved regions I and IX (Hanks *et al.*, 1988; Hanks and Quinn, 1991) as primers ((2) in Fig. 1). Amino acid residues surrounding the subdomains I and IX are diagnostic of individual protein kinase subfamilies (Hanks *et al.*, 1988; Hanks and Quinn, 1991). To isolate a variety of plant protein kinases, it was better to use the subdomains I and IX than the subdomains VI and VIII.

Amplification of plant protein kinase cDNA fragments using PCR and analysis of PCR-generated products

I carried out PCR as described in Materials and Methods, using 8 sets of the mixed primers (Table 1) designed for the cAMP-dependent protein kinase (cAPK), the cGMP-dependent protein kinase (cGPK), the calcium-phospholipid-dependent protein kinase (PKC), the calcium/calmodulin-dependent protein kinase II (CaMIIK), the cdc2 (cdc2), the Src (Src), the Fyn (Fyn), and the Abl (Abl) subfamilies (Hanks *et al.*, 1988; Hanks and Quinn, 1991). The amplified DNA fragments, which ranged from 400 to 600 bp in length, were subcloned and sequenced.

Partial sequencing analysis indicated that the 5 independent clones, designated C11, C23, D12, E68, and F15, are different from each other and contain partial sequences of protein kinases. The deduced amino acid sequences of the C11, C23 (Mizoguchi *et al.*, 1994b), D12 (Mizoguchi *et al.*, 1993a), E68 (Mizoguchi *et al.*, 1994a), and F15 are most similar to the corresponding regions of the p70 S6 kinase (48% identity), the calcium/calmodulin dependent protein kinase II (40% identity), casein kinase II (73% identity), the MAP kinase (50% identity), and the SNF1 (49% identity), respectively (Hanks *et al.*, 1988; Hanks and Quinn, 1991).

Part II.

Characterization of two cDNA clones that encode MAP kinases in *Arabidopsis*

I obtained several partial sequences encoding protein kinases from a higher plant, *Arabidopsis,* with PCR technique. Obtained PCR fragments include those of casein kinase II (Mizoguchi *et al.*, 1993a) and MAP kinase (Mizoguchi *et al.*, 1994a), both of which are shown to play central roles in various signal transduction pathways. MAP kinases in animals and yeasts are thought to be activated by various extracellular signals, integrate multiple intracellular signals transmitted by various second messengers, and regulate gene expressions by phosphorylation of trancsription factors (see Introduction). The analysis of plant MAP kinases is most suited for my purpose to understand the molecular process in which plants respond to a variety of the extracellular stimuli. Therefore I decided to concentrate on the study of genes for plant MAP kinases.

Cloning and sequence analysis of two cDNA clones that encode MAP kinases in Arabidopsis

The inserted DNA of clone E68 was used as a probe for the screening of a cDNA library for cloning cDNAs that encode homologues of MAP kinases (Mizoguchi *et al.*, 1994a). Twenty positive signals were obtained from a total of 8.0 x 10^5 plaques. The cDNA inserts were subcloned into the pSKII$^-$ vector. Partial sequence analysis revealed that the cloned DNA

inserts had two distinct but closely related sequences. We subcloned and sequenced the largest inserts and named them *ATMPK1* and *ATMPK2*.

Fig. 2A and B show the nucleotide sequences of the two cDNAs and the corresponding deduced amino acid sequences. *ATMPK1* (Fig. 2A) contains a 1110-bp open reading frame, encoding a putative polypeptide of 370 amino acids with a molecular weight of 42,685. *ATMPK2* (Fig. 2B) contains a 1128-bp open reading frame, encoding a putative polypeptide of 376 amino acids with a predicted molecular weight of 43,138. The coding regions of *ATMPK1* and *ATMPK2* are 86.3% identical at the nucleotide sequence level. Putative initiation codons encoded by *ATMPK1* and *ATMPK2* are flanked by sequences that are similar to a consensus sequence that is frequently found in plants, AAAAATGGC and AAGAATGGC, respectively (Lütcke *et al.*, 1987). Direct repeat sequences, CCAAATCT(A/C)(T/A/G)C, are found in the 5'-noncoding regions of both *ATMPK1* and *ATMPK2*. No typical polyadenylation signals were found in the 3' flanking regions.

Primary structure of the putative ATMPK1 and ATMPK2 proteins

The deduced amino acid sequences of the putative proteins encoded by *ATMPK1* and *ATMPK2* (named ATMPK1 and ATMPK2) are compared with those of other MAP kinases in Fig. 3. ATMPK1 and ATMPK2 contain all the conserved residues and the eleven subdomains that are typical of protein kinases (Hanks *et al.*, 1988; Hanks and Quinn, 1991). The amino acid residues (Thr^{191}-Glu-Tyr^{193}) corresponding to the sites of the phosphorylation that is involved in the activation of MAP kinases (Payne *et al.*, 1991; Gartner *et al.*, 1992) are conserved in both ATMPK1 and ATMPK2.

The phosphorylation of these threonine and tyrosine residues is required for the activation of MAP kinases. These sites of phosphorylation are also conserved in four protein kinases (FUS3, KSS1, MPK1, HOG1, and spk1) from yeast which show extensive homology to MAP kinases from animals (Neiman *et al.*, 1993).

As shown in Table 2, ATMPK1 and ATMPK2 are more closely related to *Xenopus* and rat MAP kinases (48.9 to 50.1%) and to yeast protein kinases (spk1, MPK1, KSS1, and FUS3; 44.2 to 49.0%) than to the *Arabidopsis* cdc2 protein kinase (39.7%) (Ferreira *et al.*, 1991; Hirayama *et al.*, 1991). *Arabidopsis* cdc2 protein kinase does not contain the TEY sequence that is conserved in other MAP kinases. ATMPK1 and ATMPK2 are much more closely related to one another (88.7% identity) than to other kinases. ATMPK1 and ATMPK2 are most similar (57.2% and 56.1% identity, respectively) to MsERK1, a homologue of MAP kinase from alfalfa (Duerr *et al.*, 1993). The extent of sequence similarity between *Arabidopsis* ATMPKs and alfalfa MsERK1 seems rather low even though they are all members of the plant MAP kinase family. By contrast, greater similarity (82%) has been noted between cdc2 proteins from *Arabidopsis* and alfalfa (Imajuku *et al.*, 1992).

Analysis by Northern and Southern blotting of the ATMPK1 and ATMPK2 genes

The existence of two closely related but structurally distinct isozymes of MAP kinase in *Arabidopsis* suggests the possibility that these isozymes may serve different functions. Such functional specialization might be expected to be reflected in the tissue-specific expression of two MAP kinases. I examined the relative levels of transcripts of *ATMPK1* and *ATMPK2* in different organs of *Arabidopsis* (Mizoguchi *et al.*, 1994a). Northern blot analysis revealed the presence of transcripts of *ATMPK1* and *ATMPK2* in all the tissues examined although the levels of their expression varied among different tissues (Fig. 4). The levels of transcripts of *ATMPK1* and *ATMPK2* were high in stems, but quite low in seeds. Single transcripts were detected in *Arabidopsis* tissues when each of these cDNAs was used as probe.

The number of genes that encode MAP kinases in the *Arabidopsis* genome was estimated by Southern blot analyses (Fig. 5) (Mizoguchi *et al.*, 1994a). Nuclear DNA from *Arabidopsis* was digested with *Bam*HI, *Eco*RI and *Hin*dIII, blotted onto nylon membranes and hybridized under both high- and low-stringency conditions with the *ATMPK1* and *ATMPK2* cDNAs as probe. Under high-stringency hybridization conditions, *ATMPK1* hybridized with one *Bam*HI, two *Eco*RI and three *Hin*dIII fragments. *ATMPK2* hybridized with one *Bam*HI, one *Eco*RI and three *Hin*dIII fragments. However, under low-stringency conditions, a few extra bands of DNA were detected, suggesting that there may be a few additional genes for MAP kinases on the *Arabidopsis* genome.

Phosphorylation and activation of the ATMPK1 and ATMPK2 proteins expressed in Escherichia coli cells by Xenopus MAP kinase kinase

The proteins encoded by *ATMPK1* and *ATMPK2* cDNAs were expressed in *E. coli* cells as products of fusions with glutathione S-transferase by use of the pGEX expression vector (Mizoguchi *et al.*, 1994a). Fig. 6A shows details of the constructs, namely, pGEX-MPK1 and pGEX-MPK2 (see Materials and Methods). The recombinant ATMPK1 and ATMPK2 fusion proteins (named GST-ATMPK1 and GST-ATMPK2, respectively) that had been induced by treatment of the cells with IPTG were recovered mainly in the insoluble protein fraction, but small amounts of the fusion proteins were recovered in the soluble fraction (Fig. 6B). The fusion proteins in the soluble fraction were purified by use of glutathione-Sepharose beads and gave two bands after SDS-PAGE. The closed and open triangles in Fig. 6B indicate the electrophoretic positions of the full-length and truncated recombinant MAP kinases, respectively.

In order to examine whether GST-ATMPK1 and GST-ATMPK2 had MAP kinase activity, protein kinase activities of the fusion proteins were measured with MBP as the substrate (Mizoguchi *et al.*, 1994a). The MBP kinase activity of GST-ATMPK1 and GST-ATMPK2 was virtually undetectable in the "in-gel kinase" assay (data not shown) and very low in the "test-tube" assay (Fig. 7C, lanes 1 and 3). Recombinant MAP kinase of *Xenopus* has been shown to be activated by incubation with MAPKK (Matsuda *et al.*, 1992; Kosako *et al.*, 1992). Therefore, I examined whether the purified *Xenopus* MAPKK could phosphorylate GST-ATMPK1 and GST-ATMPK2 and activate their MBP kinase activities. The extent of total phosphorylation of both GST-ATMPK1 and GST-ATMPK2 increased after

incubation with the purified *Xenopus* MAPKK (Fig. 7A, lanes 3 and 4, upper bands) and the activation of their respective MBP kinase activities (Fig. 7C, lanes 2 and 4) in the standard test-tube assay was detected. Autophosphorylation of GST-ATMPK1 was also detected but the extent of such phosphorylation was clearly increased by incubation with *Xenopus* MAPKK (Fig. 7A, lanes 1 and 3, upper band). Phosphorylation of the lower bands was the result from autophosphorylation of *Xenopus* MAPKK (Fig. 7A, lanes 3 and 4). Autophosphorylation of GST-ATMPK1 occurred on tyrosine as well as on threonine and serine residues and total phosphorylation of GST-ATMPK1 on tyrosine and threonine residues increased after incubation with the purified *Xenopus* MAPKK (Fig. 7B, lanes 1 and 2, respectively). Strong phosphorylation of serine residues was observed. These results indicate that ATMPK1 and ATMPK2, two plant MAP kinases, can be phosphorylated and activated by *Xenopus* MAPKK *in vitro*, although there is another possibility that the increase of total phosphorylation of the GST-ATMPK1 after incubation with the *Xenopus* MAPKK may be due to autophosphorylation of activated GST-ATMPK1.

Activation by auxin of the MBP kinase activity in auxin-starved tobacco BY-2 cells

Previous reports have indicated that MAP kinase activities are stimulated by exposure of quiescent animal and yeast cells to a variety of stimuli (Sturgill and Wu, 1991; Nishida and Gotoh, 1993; Pelech and Sanghera, 1992). I examined whether MAP kinase activity can be stimulated by phytohormones, such as auxin, that play important roles in the proliferation and growth of plant cells (Mizoguchi *et al.*, 1994a). I first

analyzed the activation of an MBP-phosphorylating activity by 2,4-D, a synthetic auxin, in tobacco BY-2 cells.

Tobacco BY-2 cells were rendered quiescent by incubation in 2,4-D free medium. The addition of 2,4-D to the 2,4-D free medium induced mitosis in the quiescent BY-2 cells, as shown in Fig. 8A. Treatment of BY-2 cells with 2,4-D transiently increased the MBP kinase activity of the 46-KDa protein kinase in the cell extracts (Fig. 8B). The MBP kinase activity of the 46-KDa kinase reached a maximum level 10 min after the addition of 1 μM 2,4-D to the quiescent BY-2 cells. The MBP kinase activity then fell to nearly the basal level within 15-20 min (Fig. 8B). The MBP kinase activity of 2,4-D treated BY-2 cells was 4-fold greater than that of the untreated cells. These results suggest that MBP kinase activity with Mw 46 KDa is activated by auxin treatment in BY-2 cells.

I then examined whether 2,4-D treated BY-2 cells have MAPKK activity that activates MAP kinases, using inactive GST-ATMPK2 protein as a substrate (Mizoguchi *et al.*, 1994a). Incubation of GST-ATMPK2 with cell extracts prepared from BY-2 cells that had been treated with 1 μM 2,4-D for 5 and 10 min resulted in the increase of the extent of total phosphorylation of GST-ATMPK2. In case of the extract from the cells treated with 2,4-D for 10 min alone, no phosphorylated band at the position indicated by an arrowhead was observed. By contrast, the extent of the total phosphorylation of GST-ATMPK2 was quite low when an extract of untreated BY-2 cells was used (Fig. 8C, D). Maximal activation of the MBP kinase activity of the 46-KDa kinase was observed after 10 min treatment with 2,4-D (Fig. 8B), while maximal total phosphorylation of GST-ATMPK2 was obtained with the extract of 5-min 2,4-D treated BY-2 cells. The extent of total phosphorylation of GST-ATMPK2 by an extract of 5-min 2,4-D

treated cells was 3.5-fold greater than that in the case of an extract of untreated cells.

Identification of *Arabidopsis* MAP kinase family

I described above the isolation of two cDNAs for MAP kinases (cATMPK1 and cATMPK2) from a dicot plant, *Arabidopsis* (Mizoguchi *et al.*, 1994a). The genomic Southern analysis (Fig. 5) (Mizoguchi *et al.*, 1994a) and the isolation of a different MAP kinase homologue from alfalafa (Duerr *et al.*, 1993; Jonak *et al.*, 1993) suggested that there are several genes for MAP kinases in plants. Four protein kinases (FUS3, KSS1, MPK1, and HOG1), which show high homology with mammalian MAP kinases, have been identified by genetic studies in *Saccharomyces cereviciae* (Elion *et al.*, 1990; Courchesne *et al.*, 1989; Torres *et al.*, 1991; Brewster *et al.*, 1993). These obserbations encouraged me to speculate that MAP kinases play various roles in the signaling processes of plants. Therefore, I screened different cDNA libraries under low-stringency conditions (see Materials and Methods) to isolate cDNAs encoding MAP kinase homologues as many as possible in the first step of the studying on the functions of MAP kinases in plants (Mizoguchi *et al.*, 1993b).

Cloning and sequence analysis of seven additional cDNA clones that encode MAP kinases in Arabidopsis

The inserted DNA of the clone E68 (Mizoguchi *et al.*, 1994a) containing the conserved sequence of MAP kinase cDNAs was used as a probe for screening two different cDNA libraries to isolate cDNA clones for homologues of MAP kinases (Mizoguchi *et al.*, 1993b). Thirty-nine positive cDNA clones were obtained from a total of 8.0 x 10^5 plaques. The cDNA inserts were subcloned into the pSKII$^-$ vector. Partial sequence analysis

revealed that the cloned DNA inserts had 10 distinct but closely related sequences for MAP kinase homologues: two were cATMPK1 and cATMPK2 (Fig. 2) (Mizoguchi *et al.*, 1994a). I subcloned and sequenced the largest inserts of eight cDNAs encoding putative MAP kinases and determined the nucleotide sequences of seven of the cDNAs (cATMPK3 to 9) (Fig. 9). Table 3 summarizes the nucleic acid length, deduced amino acid length, and molecular weights of the *ATMPK* cDNA clones. The deduced molecular weight ranges from approximately 42 to 64 KDa.

Primary Structure of the putative ATMPK proteins

Fig. 10 compares the deduced amino acid sequences of the putative proteins encoded by *ATMPK1* to nine genes (named ATMPK1 to 9) (Mizoguchi *et al.*, 1993b; 1994a) with *Xenopus* MPK1 (Gotoh *et al.*, 1991). Putative ATMPK proteins show extensive homology to *Xenopus* MAP kinase. ATMPK proteins contain all the conserved residues and the 11 subdomains that are typical of protein kinases (Hanks *et al.*, 1988; Hanks and Quinn, 1991). The amino acid residues (Thr-Glu-Tyr), the TEY sequence (Sturgill and Wu, 1991; Cobb *et al.*, 1991; Thomas, 1992; Nishida and Gotoh, 1992; Pelech and Sanghera, 1992; Nishida and Gotoh, 1993; Neiman *et al.*, 1993), are conserved in all the ATMPK proteins. The phosphorylation of these threonine and tyrosine residues is required for the activation of MAP kinases, and it has been speculated that the TEY sequence represents the regulatory phosphorylation sites common to all MAP kinases (Sturgill and Wu, 1991; Cobb *et al.*, 1991; Thomas, 1992; Nishida and Gotoh, 1992; Pelech and Sanghera, 1992; Nishida and Gotoh, 1993; Neiman *et al.*, 1993).

To analyze the similarities between ATMPK proteins and other MAP kinases from various organisms (Boulton *et al.*, 1990; Boulton *et al.*, 1991; Elion *et al.*, 1990; Courchesne *et al.*, 1989; Torres *et al.*, 1991; Brewster *et al.*, 1993; Toda *et al.*, 1991; Mizoguchi *et al.*, 1993b; 1994a; Duerr *et al.*, 1993; Stafstrom *et al.*, 1993; Jonak *et al.*, 1993; Gotoh *et al.*, 1991; Whiteway *et al.*, 1992; Biggs and Zipursky, 1992; Posada *et al.*, 1991; Tanner and Mueckler, 1993; Her*et al.*, 1991; GenBank, accession number: Z14089; Charest *et al.*, 1993; Owaki *et al.*, 1992), the deduced amino acid sequences of known MAP kinases were compared around kinase subdomains VI to VIII (Hanks *et al.*, 1988; Hanks and Quinn, 1991), including the TEY sequence (Fig. 11). It has been reported that there are two regions where the amino acid sequences of MAP kinases are markedly divergent (Toda *et al.*, 1991). One is a short stretch located between subdomains VII and VIII, and another is the region of the carboxyl-terminal 50-amino acid residues. The ATMPK proteins are also quite divergent in the regions between subdomain VII and VIII (Fig. 11), although the subdomains VI, VII, and VIII are highly conserved among all MAP kinases.

ATMPK8 and 9 have putative extra regulatory domains in their C-terminal (Fig. 9F, G and Fig. 10). The N-terminal region of ATMPK8 is rich in serine (S) and asparagine (N) residues.

Phylogenetic tree analysis

Amino acid identities between ATMPK proteins and other members of MAP kinase and JNK/SAPK family (Pelech and Sanghera, 1992; Nishida and Gotoh, 1993; Neiman, 1993; Dérijard *et al.*, 1994; Kyriakis *et al.*, 1994) from animals and yeasts range from 33 to 56%. Among plant MAP kinase homologues, ATMPK1, ATMPK2 (Mizoguchi *et al.*, 1994a), ATMPK7 (Mizoguchi *et al.*, 1993b), and NTF3 (Wilson *et al.*, 1993) from tobacco are closely related (80% to 89% identity). ATMPK3, ATMPK6 (Mizoguchi *et al.*, 1993b), MsERK1 (Duerr *et al.*, 1993; Jonak *et al.*, 1993) from alfalfa and PsMAP (Stafstrom *et al.*, 1993) from pea are also closely related to one another (76% to 94% identity). ATMPK4 and ATMPK5 (75% identity) (Mizoguchi *et al.*, 1993b), ATMPK8 and ATMPK9 (82% identity) are closely related to each other, respectively.

In order to better understand the relationships between the reported proteins and known MAP kinases from other organisms, we constructed a phylogenetic tree (Fig. 12) by using the UPGMA program (Nei, 1987). The results obtained indicate that ATMPK proteins from *Arabidopsis* are located in a separate branch and constitute a subfamily which is slightly different from those of homologous proteins of animals or yeasts. Based on the phylogenetic tree, plant MAP kinases can be grouped in, at least, four separate subfamilies.

Identification of putative upstream factors of MAP kinases in *Arabidopsis*

I described above the isolation of nine cDNAs (cATMPK1 to 9) from a dicot plant *Arabidopsis* which exhibit high degree of homology to those of genes for MAP kinases (Fig. 10) (Mizoguchi *et al.*, 1993b; 1994a). The plant MAP kinases including these *Arabidopsis* genes can be classified into four subgroups based on sequence similarities (Fig. 12) (Mizoguchi *et al.*, 1993b). I also demonstrated that MAP kinase and MAPKK-like activities transiently increased in response to 2, 4-D (an auxin analogue) treatment in tobacco BY-2 cells (Fig. 7) (Mizoguchi *et al.*, 1994a). cNPK1, a cDNA encoding a protein kinase that exhibits a high degree of homology to yeast MAPKKK BCK1 (Lee and Levin, 1992) and STE11 (Rhodes *et al.*, 1990), was cloned from tobacco (Banno *et al.*, 1993). cNPK2 which are structurally related to members of MAPKK family has also been isolated from tobacco (Jonak *et al.*, 1994). These observations support the existence of similar phosphorylation cascades that include MAPKKKs, MAPKKs and MAP kinases with those of MAP kinase cascades of animals and yeasts (Mizoguchi *et al.*, 1993b; 1994a; Banno *et al.*, 1993; Jonak *et al.*, 1994).

I screened *Arabidopsis* cDNA libraries to isolate cDNAs encoding MAPKK or MAPKKK homologues, which may function as putative upstream factors of ATMPK1 to 9, using the PCR fragments or the cDNA inserts isolated by random sequencing project as probes (GenBank, accession number: T04696 and T13020, see Materials and Methods). Here I report the isolation and characterization of a cDNA (designated as cATMEKK1) encoding a protein kinase which shows high homology to the members of MAPKKK family (see Introduction) in *Arabidopsis*.

Cloning and sequence analysis of cDNAs for protein kinases related to MAPKKKs

The 700-bp fragment containing the conserved sequence of MAPKKKs (GenBank, accession number: T04696 and T13020) was used as a probe for screening two different cDNA libraries to isolate full-length cDNA clones. Thirty-nine positive cDNA clones were obtained from a total of 8.0 x 10^5 plaques. The cDNA inserts were subcloned into the pSKII$^-$ vector. Partial sequence analysis revealed that the cloned DNA inserts had at least 4 distinct but closely related sequences. I subcloned and sequenced the largest inserts of four cDNAs and determined the nucleotide sequences of a cDNA clone (cATMEKK1) completely. cATMEKK1 contains 2386-bp and the starting methionine can be predicted to occur at nucleotide 136. With this methionine at the start, the cDNA encodes a protein of 608 amino acids, corresponding to a molecular size of 65.980 KDa (Fig. 13).

Primary structure of the putative ATMEKK1 protein

A homology search with the amino acid sequence of the predicted ATMEKK1 protein was performed using the sequences in the DNA and protein data bases. As shown in Fig. 14A, the putative ATMEKK1 protein shows extensive homology to byr2 (40.1% identity) from *S. pombe* (Wang *et al.*, 1991), STE11 (42.2% identity) from *S. cereviciae* (Rhodes *et al.*, 1990), BCK1 (42.2% identity) from *S. cereviciae* (Lee and Levin, 1992), MEKK (36.4% identity) from mouse (Lange-Carter *et al.*, 1993), and NPK1 (45.9% identity) from tobacco (Banno *et al.*, 1993). ATMEKK1 contains all the conserved residues and the 11 subdomains that are typical of protein

kinases (Hanks *et al.*, 1988; Hanks and Quinn, 1991). A phylogenetic tree was constructed to measure evolutional distance among the ATMEKK1, STE11, Byr2, BCK1, NPK1, and MEKK proteins (Fig. 14B). The evolutional distance between ATMEKK1 and other MAPKKKs are almost equal. ATMEKK1 belongs to a distinct subfamily from that NPK1 belongs to.

The putative ATMEKK1 protein has non-catalytic domain in the N-terminal half (Fig. 14C). The STE11, byr2, BCK1, and MEKK proteins also have kinase-unrelated domains in their N-terminal halves, which are thought to regulate their protein kinase activity, although the NPK1 protein has a kinase-unrelated region in its C-terminal half (Lee and Levin, 1992; Rhodes *et al.*, 1990; Banno *et al.*, 1993; Wang *et al.*, 1991; Lange-Carter *et al.*, 1993). In contrast to the homologies among the catalytic domains of these protein kinases, the noncatalytic domain of the ATMEKK1 has no significant homology with those of byr2, STE11, BCK1, MEKK, and NPK1. I have compared the deduced amino acid sequence of the noncatalytic domain in the N-terminal half of the ATMEKK1 protein with the sequences present in the data banks. This analysis revealed that no proteins have significant homology with the sequence of this domain of the ATMEKK1 protein (data not shown).

Analysis by Southern blotting of the ATMEKK1 gene

The number of genes that encode ATMEKK1 related genes in the *Arabidopsis* genome was estimated by Southern blot analyses (Fig. 15). Nuclear DNA from *Arabidopsis* was digested with *Eco*RI and *Hind*III, blotted onto nylon membranes and hybridized under both high- and low-stringency conditions with the cATMEKK1 insert as a probe. Under high-

stringency hybridization conditions, the cATMEKK1 insert hybridized with one strong band and several faint bands when genomic DNA was digested with both *EcoR*I and *Hind*III. However, under low-stringency conditions, additional DNA bands were detected. This result indicates that the ATMEKK genes constitute a gene family on the *Arabidopsis* genome. Isolation of three additional cDNAs (ATMEKK2, ATMEKK3, and ATMEKK4) related to ATMEKK1 also support this result (data not shown).

Identification of putative downstream factors of MAP kinases in *Arabidopsis*

Cloning and sequence analysis of two cDNAs for protein kinases related to ribosomal S6 kinases

The inserted DNA of the clone C11 was used as a probe in the screening of a cDNA library to clone cDNAs encoding novel serine/threonine protein kinases. Eight positive signals were obtained out of 6 x 10^5 plaques. The cDNA inserts were subcloned into pSKII$^-$ vector. Partial sequence analysis revealed that the cloned DNA inserts had two distinct but closely related sequences. I subcloned and sequenced the largest insert for each of these cDNAs and named them cATPK6 and cATPK19. Fig. 16 (A) and (B) show the nucleotide sequences and the corresponding amino acid sequences of the two cDNAs. cATPK6 (Fig. 16A) contains a 1395 bp open reading frame, encoding a polypeptide of 465 amino acids and has a predicted molecular weight of 52,587. cATPK19 (Fig. 16B) contains a 1413 bp open reading frame, encoding a polypeptide of 471 amino acids and has a predicted molecular weight of 53,111. cATPK6 and cATPK19 are 82.0% identical at the nucleotide sequence level.

Primary structure of the putative ATPK6 and ATPK19 proteins

The deduced amino acid sequences of ATPK6 and ATPK19 are highly homologous to the conserved catalytic subdomains of other serine/threonine-specific protein kinases (Fig. 17A). Conserved amino acids characteristic of each of the 11 catalytic subdomains found in almost

all serine/threonine-specific protein kinases are present (Hanks *et al.*, 1988; Hanks and Quinn, 1991). ATPK6 and ATPK19 are closely related each other and 85% identical at the amino acid level.

A search of the sequence data bases (GenBank, EMBL, and Swiss-Prot, October 1993) using GCG version 7.0 software showed that the ATPK6 and ATPK19 sequences within the putative catalytic domains are most related to human $p70^{S6K}$ (Grove *et al.*, 1991) (51.7 - 52.1% identity), followed by *Xenopus* $pp90^{rsk}$ (Jones *et al.*, 1988) (49.0 - 49.4% identity), human protein kinase A (Maldonado *et al.*, 1988) (41.5 - 41.1% identity), and human protein kinase C (Finkenzeller *et al.*, 1990) (40.2 - 41.1% identity). Comparisons of the deduced amino acid sequences of ATPK6 and ATPK19 with these protein kinases showed that there was a high degree of sequence identity and similarity within 11 conserved subdomains of the catalytic domain (Fig. 17A) (Hanks *et al.*, 1988; Hanks and Quinn, 1991). To trace the phylogenetic relationship of ATPK6 and ATPK19 to other protein kinases, an unrooted phylogenetic tree was constructed using parsimony analysis (Fig. 17B). Based on this analysis, ATPK6 and ATPK19 are most closely related to ribosomal S6 kinase (RSK) family, but are more distantly related to the PKC or the PKA families (Hanks *et al.*, 1988; Hanks and Quinn, 1991).

ATPK6 and ATPK19 have putative regulatory domains in their N- and C-terminal regions (Fig. 16A and B). These regions contain the PEST regions (Rogers *et al.*, 1986; Rechsteiner *et al.*, 1987). The PEST regions are high local concentration of proline (P), glutamic acid (E), serine (S), threonine (T), and to a lesser extent aspartic acid (D) residues. Proteins that contain PEST regions are thought to be rapidly degraded. Known PEST proteins including oncogene products (myc, fos, E1A, myb, and p53), key enzymes (ODC, TAT, HMG, and CoA reductase), and components of signal

pathways (cAMP- and cGMP-dependent kinases, protein kinase C, phytochrome, cyclin, and steroid receptors), are thought to be important regulatory molecules, and their fast turnover surely reflects a metabolic requirement for rapidly changing their concentrations (Rechsteiner *et al.*, 1987). Several consensus phosphorylation sites for casein kinase II ($S/TX_2E/D$) and proline-directed protein kinases such as cdc2 (XS/TPX), MAP kinase ($PX_{1\ or\ 2}S/TP$) (Fig. 16A and B) (Pearson and Kemp, 1991).

Analysis by Southern blotting of the ATPK6 and ATPK19 genes

I estimated the number of genes related to *ATPK6* and *ATPK19* in the *Arabidopsis* genome by Southern blot analyses (Fig. 18). *Arabidopsis* nuclear DNA was digested with *Bam*HI, *Eco*RI, and *Hin*dIII, blotted onto nylon membranes and hybridized under both high- and low-stringency conditions using the ATPK6 and ATPK19 cDNAs as probes. Under high-stringency hybridization conditions ATPK6 hybridized with one strong and one weak *Bam*HI, one strong and one weak *Eco*RI, and two strong and three weak *Hin*dIII fragments. ATPK19 hybridized with one strong and one weak *Bam*HI, four *Eco*RI, and three strong and two weak *Hin*dIII fragments under the same conditions. Under low-stringency hybridization conditions, no other signals than cross-hybridizing bands which correspond to *ATPK6* or *ATPK19* genes was detected in each case. This result indicates that there is no closely related gene to *ATPK6* and *ATPK19* in *Arabidopsis* genome.

Part III.

Effect of a variety of stresses on the expression of genes for the putative components of the *Arabidopsis* MAP kinase cascades

ATMPK gene expression is induced after a variety of stimuli

I examined whether expression of MAP kinase genes of *Arabidopsis* were affected in plants after treatment of a variety of stimuli, because my preliminary experiments indicated that the level of *ATMPK3* mRNA was affected by high salt treatment (data not shown). Genomic Southern blotting analysis using *ATMPK* cDNAs as probes demonstrated that the *ATMPK* genes almost did not cross-hybridize one another (Fig. 19) under high-stringency conditions (Mizoguchi *et al.*, 1993a; 1994a).

Fig. 20 shows the accumulation of transcripts for *ATMPK*s in response to a number of stress conditions, as analyzed by Northern blott. Quantification of the accumulated mRNAs for *ATMPK*s was plotted. The level of the *ATMPK3* mRNA was low in fully hydrated plants grown on GM agar plates. However, transcripts for *ATMPK3* increased dramatically in response to low temperature (about 10 fold). The level of *ATMPK3* mRNA increased within 1 h after the initiation of cold stress, and then fell at least for the first 10 h of cold stress. High-salt treatment also affected the level of *ATMPK3* mRNA and it began to increase within 1 h. When plants were transferred from GM agar plates to water as controls, the amount of *ATMPK3* mRNA almost did not change. Those of *ATMPK2* and 6 gradually increased by extracellular 2,4-D treatment (3 fold). The level of *ATMPK5* increased in response to ABA, low-temperature, and NaCl treatments (2.5-

4 fold). In contrast, the mRNA levels of *ATMPK1*, *8*, and *9* were almost not affected (less than 2 fold) by these treatments.

The mRNA level of ATMEKK1 is affected by a variety of stimuli

I examined whether the expression of ATMEKK1 was affected in plants after treatment with a variety of stimuli. My preliminary results indicated that the level of *ATMEKK1* mRNA increased (about 2.3 fold) in response to high-salt treatment for 10 h (data not shown). I then analyzed the accumulation of the *ATMEKK1* mRNA in *Arabidopsis* in response to a number of stress conditions for various lengths of time up to 24 h (Fig. 21). The level of the *ATMEKK1* mRNA was low in fully hydrated plants grown on GM agar plates. However, when plants were transferred from GM agar plates to water (4 °C), the level of the *ATMEKK1* mRNA began to increase within 1 h and peaked (4 fold) after 2 h. The elevated level then fell gradually at least for the first 24 h of cold stress. Under high-salt condition, the level of mRNA increased within 1 h. The transcript remained at an elevated level for at least 24 h. When plants were transferred from GM agar to water (room temperature) as controls, the level of the *ATMEKK1* mRNA slightly increased (about 1.5 fold) within 1 h and then went back immediately. Similar results were obtained when plants were treated with ABA, heat, or 2,4-D.

The mRNA levels of ATPK6 and ATPK19 are affected by a variety of stimuli

Recently, the transcription levels of protein kinases (Lin *et al.*, 1991; Anderberg and Walker-Simmons, 1992; Urao *et al.*, 1994), transcription

factors (Aguan *et al.*, 1993; Urao *et al.*, 1993), G proteins (Yoshida *et al.*, 1993), and calmodulins (Braam and Davis, 1990) in plants have been reported to be affected by environmental stresses (*e.g.* light, temperature, mechanical stress, and desiccation). The transcripts for two protein kinases, ATMPK3 and ATMEKK1, structurally related to MAP kinase and MAPKKK, respectively, increased in response to cold or high salt stresses within 1 h, as described above. Therefore, I analyzed the accumulation of *ATPK6* and *ATPK19* mRNAs in *Arabidopsis* plants in response to a variety of stress conditions for various lengths of time up to 24 h (Fig. 22). The levels of *ATPK6* and *ATPK19* mRNAs was low in fully hydrated plants grown on GM agar plates. However, when plants were transferred from GM agar plates to water (4 °C), the level of *ATPK6* mRNA began to increase within 2 h and peaked (6.8 fold) after 10 h. The *ATPK19* mRNA also began to increase in response to cold shock within 1 h, peaked (8.8 fold) after 2 h, and then fell gradually at least for the first 24 h of cold stress. Under high-salt condition, the level of *ATPK6* mRNA increased within 2 h, peaked (4.8 fold) after 10 h, and remained at an elevated level for at least 24 h. The *ATPK19* mRNA also began to accumulate within 1 h, peaked (6.3 fold) after 10 h, and remained at an elevated level for at least 24 h. When plants were transferred from GM agar to water (room temperature) as controls, the levels of *ATPK6* and *ATPK19* mRNAs slightly increased within 1 h and then went back immediately. Although the *ATPK6* mRNA increased in response to the application of exogenous ABA or 2,4-D and *ATPK19* mRNA increased in response to 2,4-D, the increase of mRNAs for *ATPK6* and *ATPK19* were relatively low compared with those in cold or high-salt stresses.

The transcription levels of three protein kinases, ATMPK3, ATPK19, and ATMEKK1 increase simultaneouly in response to a variety of stresses

As shown in Fig. 20, the expression level of *ATMPK3* increases dramatically when plants are under cold or high-salt stress conditions. Interestingly, the mRNA levels of two protein kinases, *ATPK19* and *ATMEKK1*, increase similarly in response to these stresses (Fig. 23). *ATPK19* and *ATMEKK1* are most structurally related to ribosomal S6 kinases ($p70^{S6K}$ and $pp90^{rsk}$) (Blenis, 1993; Downward, 1994) (Fig. 17) and *MAPKKK*s (Neiman, 1993; Lange-Carter *et al.*, 1993; Banno *et al.*, 1993) (Fig. 14), respectively. The levels of the *ATMPK3*, *ATPK19*, and *ATMEKK1* mRNAs did not change when plants were transferred from GM agar plates to water as controls (Fig. 23). The mRNA levels of *ATMPK1* (Mizoguchi *et al.*, 1994a) and cdc2a (Hirayama *et al.*, 1991) were not so affected by these treatments (less than 2 fold) (Fig. 23).

Induction of the ATMPK3, ATPK19, and ATMEKK1 genes following touch stimulation, salinity stress, and low temperature

I examined the levels of the *ATMPK3*, *ATPK19*, *ATMEKK1* mRNAs within 1 hr after a variety of stresses, because their changes in response to cold shock were more dramatic than those of any other *ATMPK*s (Fig. 20). As shown in Fig. 24A, the levels of the *ATMPK3*, *ATPK19*, and *ATMEKK1* mRNAs in response to low-temperature, high-salt stress began to increase within 5 min. Unexpectedly, when plants were transferred from GM agar plates to water and MS medium as controls, the mRNA levels were also increased, although the extent of increase in response to water was

relatively small (Fig. 24B). Therefore, it is possible for me to assume that the mRNA levels of these three protein kinases increase in response to mechanical (or touch) stimuli (Braam and Davis, 1990; Jaffe, 1973) or wounding stress (Gallie, 1993), when they are transferred from GM agar plates to other conditions.

Thus I first examined the effect of wounding stress (Gallie, 1993) on their expression. Under the same conditions the transcription level of *cdc2a* (Hirayama *et al.*, 1991) did not change (Fig. 24A and B). It has been reported that the mRNA level of *cdc2a* rapidly and markedly increases in response to wounding stress within 30 min (Hamerly *et al.*, 1993). These results indicate that the accumulation of the transcripts for *ATMPK3*, *ATPK19*, and *ATMEKK1* are not due to wounding stress.

Next I examined the effect of mechanical/touch stimuli (Braam and Davis, 1990; Jaffe, 1973) on the expression of three protein kinases. As shown in Fig. 24C, the mRNA levels of *ATMPK3*, *ATPK19*, and *ATMEKK1* increased simultaneously in response to touch stimuli. The expression pattern of one of touch-induced genes, TCHγ (see Materials and Methods) encoding *Arabidopsis* calmodulin were similar with those of these genes (Braam and Davis, 1990), which is a positive control. The mRNA levels of *ATMPK1* and *cdc2a* did not change under the same conditions.

Then I examined whether other stresses such as cold or high-salt condition could affect the mRNA levels of *ATMPK3*, *ATPK19*, and *ATMEKK1*. When plants were transferred from GM agar plates and grown hydroponically in water, the mRNA level of *ATMPK3* increased transiently and went back within 1 h (Fig. 24B). Treatments of plants with water or NaCl solution after 1 - 2 h of pretreatments of plants with water did not affect the expression levels of *ATMPK3* and *ATPK19*. However, after 10 - 24

h of pretreatments with water, the mRNA levels increased again within 20 min only by high-salt treatment (data not shown). Therefore, I examined whether the mRNA levels increased by cold or high-salt stresses after 22 h of pretreatment with water (Fig. 25A). The expression levels of *ATMPK3* (Fig. 25A) began to increase after 5 min in response to salinity stress and cold shock. The mRNA level of *ATPK19* also increased similarly (data not shown). In contrast, the mRNA levels of these kinases did not change in response to water. The expression levels of *cdc2a* (Fig. 25A) and *ATMPK1* (data not shown) were almost not affected by same conditions. As shown in Fig. 25B, similar pattern of responsiveness to water, heat, cold, and salinity stresses after 22h of pretreatment with water were obtained with *ATMPK3*, *ATPK19*, and *Tchγ*. The expression levels of these three genes which increased in response to touch (Fig. 24C) were markedly affected by salinity stress (Fig. 25B). In contrast, treatment with water did not affect the expression levels of these three genes. Heat stress also slightly affected their induction.

Discussion

Structures of protein kinases are highly conserved in yeasts, plants, and animals

Using the PCR method, I isolated a variety of cDNAs that contain conserved sequences of eukaryotic protein kinases, G proteins, and calmodulins (Fig. 26). They include casein kinase II catalytic subunits (ATCKA1 and 2) (Mizoguchi *et al.*, 1993a), MAP kinases (ATMPK1 - 9) (Mizoguchi *et al.*, 1993b; 1994a), MAPKK (pMAPKK), MAPKKKs (ATMEKK1 - 4 and pNPK1), and ribosomal S6 kinases (ATPK6 and 19), and these protein kinases in animals and yeasts are reported to play important roles in signal transduction pathways of various growth factors (see Introduction). ATCKA1 and ATCKA2 show more than 70% identity to the casein kinase II in the evolutionary distant organisms, yeasts and vertebrates (Mizoguchi *et al.*, 1993a). It has also demonstrated that ATCKA1 and ATCKA2 proteins expressed in *E. coli* had several similar biochemical properties to those of yeasts and vertebrates (Mizoguchi *et al.*, 1993a). ATMPK1 and ATMPK2 are not only structurally related to MAP kinases in yeasts and vertebrates, but also have similar biochemical properties (Mizoguchi *et al.*, 1994a), as discussed fully in the following parts. pMAK and pSNF1 also show approximately 60% identity to the counterparts of yeast SNF1 and rat MAK, respectively. The product of the *SNF1* of yeast is a protein kinase that functions in the regulation of carbon metabolism (Celenza and Carlson, 1986). The *MAK* gene product is suggested to participate in sensory signal transduction (Matsushime *et al.*, 1990; Bladt and Birchmeier, 1993).

The existence of the highly conserved protein kinases in evolutionary distant organisms, yeasts, plants, and animals, provides us an insight that

fundamental mechanisms of signal transduction are highly conserved among all the eukaryotes and these conserved protein kinases play important and similar roles in all the eukaryotic cells. Therefore, the comparison of these highly conserved protein kinases in different three systems, yeasts, plants, and animals, may give us further information concerning common or specific features of signal transduction pathways modulated by phosphorylation in eukaryotic cells.

In contrast, ATPK64 (Mizoguchi *et al.*, 1992) and ATPK27 are involved in the PVPK-1 and the SRK (S locus receptor kinase) family, respectively (Lawton *et al.*, 1989; Stein *et al.*, 1991). In the *Brassica* family, fertilization can be controlled by a self-incompatibility system that is inherited as a dominant genetic locus called the *S* locus (Nasrallah *et al.*, 1991). The SRK gene has been shown to cosegregate with this self-incompatibility (Dzelzkalns *et al.*, 1992). The *Pto*-encoded protein, which responds to and interacts with the *Pseudomonas syringae avrPto* gene, encodes a serine/threonine kinase involved in the SRK family (Martin *et al.*, 1993). There have been no reports of the protein kinases in yeasts and animals which are related to the members of these two families. Therefore the PVPK-1 and the SRK families may be specific in plants, and studies on these putative plant-specific protein kinases may help us to understand unique mechanisms of signal transduction in plants.

ATMPK1 and ATMPK2 from Arabidopsis encode MAP kinases

To understand signal transduction pathways in plants, it seems to be important to ascertain whether MAP kinases exist and function in the intracellular signal pathways of plants since these enzymes appear to play key roles in a wide variety of responses to extracellular signals in animals and yeasts (Sturgill and Wu, 1991; Nishida and Gotoh, 1993; Pelech and Sanghera, 1992). Two lines of evidence described herein indicate that cloned *ATMPK1* and *ATMPK2* encode homologues of MAP kinase in *Arabidopsis* (Mizoguchi *et al.*, 1994a). First, the putative ATMPK1 and ATMPK2 proteins show extensive sequence homology to *Xenopus* MPK1 (Gotoh *et al.*, 1991), rat ERK2 (Boulton *et al.*, 1991) and yeast KSS1 (Courchesne *et al.*, 1989) (Fig. 3 and Table 1), all of which are members of the MAP kinase family. Furthermore, ATMPK1 and ATMPK2 contain the conserved subdomain VIII (Hanks *et al.*, 1988) that includes the phosphorylation sites (TEY) that are typical of MAP kinases (see Introduction). Phosphorylation of tyrosine residues, as well as of threonine residues, is necessary for the activation of MAP kinases (Payne *et al.*, 1991; Gartner *et al.*, 1992). Second, the recombinant proteins, GST-ATMPK1 and GST-ATMPK2 expressed in *E. coli*, were phosphorylated and activated by the purified *Xenopus* MAPKK (Matsuda *et al.*, 1992; Kosako *et al.*, 1992; Kosako *et al.*, 1993), which functions as a direct upstream activator that is sufficient for full activation of inactive *Xenopus* MAP kinase (Fig. 6A and C). Analysis of phosphoamino acids revealed that the activation of the GST-ATMPK2 protein was accompanied by an increase in the extent of phosphorylation of tyrosine as well as of threonine and serine residues (Fig. 6B). These observations indicate that *ATMPK1* and *ATMPK2* encode

homologues of MAP kinases in *Arabidopsis*. However, the extent of activation and phosphorylation of the GST-ATMPK1 and GST-ATMPK2 by *Xenopus* MAPKK was quite small, perhaps because of the heterologous combination of a plant MAP kinase with a MAPKK from an amphibian.

MAP kinases play key roles in the growth factor- and differentiating factor-stimulated signaling pathways in animals and yeasts. Various growth factors and mitogens stimulate the activities of MAP kinases in animals (see Introduction). In plants, auxins are thought to play a fundamental role in the control of plant growth and development (Krikorian *et al.*, 1987). Therefore, I examined whether or not MAP kinases could be activated by 2,4-D (an auxin analogue) in tobacco BY-2 cells (Mizoguchi *et al.*, 1994a). I detected a rapid and transient activation of the MBP-phosphorylating activity of the 46-KDa kinase in BY-2 cells that had been treated with 2,4-D for 5 min using an in-gel phosphorylation assay (Fig. 7B). However, protein kinase activities were not detected when casein and histone were used as substrates under the same conditions (data not shown). MAP kinases have been reported to be rapidly and transiently activated by mitogens and to phosphorylate MBP but not casein or histone *in vitro* (Sturgill and Wu, 1991; Nishida and Gotoh, 1993; Pelech and Sanghera, 1992). These results suggest that both MAP kinase and MAPKK may be activated by auxin in tobacco BY-2 cells, even though the 46-KDa kinase remains to be fully characterized. The extent of total phosphorylation of GST-ATMPK2 increased after incubation with extracts prepared from 5-min 2,4-D-treated BY-2 cells (Fig. 7C, D). I examined whether the extract of BY-2 cells that had been treated with 2,4-D for 5 min could activate GST-ATMPK2 *in vitro* (data not shown). The incubation of GST-ATMPK2 with Mg^{2+} ion, ATP and the extract of the 2,4-D treated BY-2

cells increased the MBP kinase activity of GST-ATMPK2. The MBP kinase activity was increased about 3- and 7-fold than its basal level after incubation in the extract of the cells alone and with native GST-ATMPK2, respectively (data not shown). Matsuda *et al.* reported that increases and decreases in the activity of the *Xenopus* MAPKK are associated with the activation and inactivation of MAP kinase, respectively (Matsuda *et al.*, 1992). These observations lead me to speculate that MAPKK, that is activated by treatment with auxin, in turn activates MAP kinase in tobacco BY-2 cells. In plant as well as in animal system, MAP kinase may play important roles in mitogen activated gene expression. A number of auxin-inducible genes that have been reported in various plants can be possible downstream target genes that are controlled through MAP kinase-signal pathways.

Arabidopsis MAP kinase constitutes a gene family

Genomic Southern blot analysis indicated that the *Arabidopsis* genome has several related genes in addition to ATMPK1 and ATMPK2 (Mizoguchi *et al.*, 1994a). Duerr *et al.* reported the cloning of a cDNA, MsERK1, that encodes a homologue of MAP kinase from alfalfa (Duerr *et al.*, 1993). However, a low level of sequence homology was observed between MsERK1 and ATMPKs (56%) at amino acid level (Duerr *et al.*, 1993; Mizoguchi *et al.*, 1994a). The region outside of the conserved catalytic domain of MsERK1 exhibits no homology to those of ATMPK1 and ATMPK2 even though the putative catalytic domains of these kinases are highly conserved (Fig. 2). The amino acid sequences of two homologues of MAP kinases in budding yeast, KSS1 and MPK1, show evidence of conservation of

amino acids only in the catalytic domains and these enzymes have been shown to function in different signal transduction pathways (Courchesne *et al.*, 1989; Torres *et al.*, 1991). I screened cDNA libraries under low-stringency hybridization conditions with the same probe used to isolate two cDNA clones, cATMPK1 and cATMPK2, and found that *Arabidopsis* has at least nine different genes for MAP kinases and that these genes can be classified into subgroups based on sequence similarities (Mizoguchi *et al.*, 1993b; Mizoguchi *et al.*, manuscript in preparation).

Putative ATMPK proteins have extensive homology to MAP kinases isolated from animals and yeasts, and contain the TEY sequence characteristic of the MAP kinase family. Based on the comparison of amino acid identities among all the related MAP kinases, plant MAP kinases (ATMPK1 to 9 from *Arabidopsis* (Mizoguchi *et al.*, 1993b; 1994a), MsERK1 from alfalfa (Duerr *et al.*, 1993; Jonak *et al.*, 1993), PsMAP from pea (Stafstrom *et al.*, 1993), and NTF3 from tobacco (Wilson *et al.*, 1993)) apparently form a large group different from MAP kinases obtained from animals and yeasts (Table 2). The reported plant MAP kinases seem to be classified into the following four subgroups: subgroup 1: ATMPK1, ATMPK2, ATMPK7, and NTF3; subgroup 2: ATMPK3, ATMPK6, MsERK1, and PsMAP; subgroup 3: ATMPK4 and ATMPK5; subgroup 4: ATMPK8 and ATMPK9. Amino acid identities among members of subgroups 1, 2, and 3 are between 79.7 and 88.7%, 75.6 and 94.4%, or is 74.5%, respectively. By contrast, amino acid similarities among the three subfamilies range from 52.7 to 72.2%. ATMPK6 shows high homology to MsERK1 (88.4%) and PsMAP (87.1%) not only in the catalytic domains but also in the divergent regions between subdomains VII and VIII (Mizoguchi *et al.*, 1994a). This suggests that ATMPK6 is a counterpart of MsERK1 (Duerr *et al.*, 1993;

Jonak *et al.*, 1993) and PsMAP (Stafstrom *et al.*, 1993) in *Arabidopsis*. An *S. pombe* MAP kinase, Spk1 (Toda *et al.*, 1991), is essential for mating. Spk1-disrupted mutant, which is defective in sporulation, has evidently restored its ability to sporulate by complementation with genes for *Xenopus* MAP kinase (Gotoh *et al.*, 1991) that show 55.3% identity with Spk1. ATMPK5 shows a significantly higher homology (55.8% identity) with animal and yeast MAP kinases, while other plant MAP kinases have less than 52.3% identity (Mizoguchi *et al.*, 1993b). ATMPK5 may restore the sporulation ability of Spk1-disrupted *S. pombe* cells. Genomic Southern analysis of alfalfa indicated the existence of a number of genes for MAP kinase, and PCR has produced at least 4 MAP kinases in alfalfa (Jonak *et al.*, 1993), which provides further evidence for multiple MAP kinases in plants.

I identified at least 9 independent MAP kinases, ATMPKs, in *Arabidopsis* and classified them into four subgroups based on comparisons between their amino acid sequences (Fig. 2) (Mizoguchi *et al.*, 1993b; 1994a). It was reported that in *S. cereviciae* three different MAP kinases (FUS3, MPK1, and HOG1) function in different signaling pathways (Neiman, 1993), while in vertebrate cells at least one MAP kinase (MAPK or ERK) are thought to function in various signal transduction pathways (Sturgill and Wu, 1991; Cobb *et al.*, 1991; Thomas, 1992; Nishida and Gotoh, 1992; Pelech and Sanghera, 1992; Nishida and Gotoh, 1993). Recently, however, two additional protein kinases, JNK1 (Dérijard *et al.*, 1994) and p38 (Han *et al.*, 1994), related to MAP kinases have been identified in vertebrates. These results suggest that multiple protein kinases related to MAP kinases may function in multiple signal transduction pathways in eukaryotic cells.

MAP kinases have a divergent region between subdomains VII and VIII (Fig. 2) (Toda *et al.*, 1991; Mizoguchi *et al.*, 1993b). Since this region is

located just before the TEY sequence, which is the phosphorylation site for the activation by MAPKKs, it may play an important role in the interaction of MAP kinases with its direct upstream activator MAPKKs (Zhang *et al.*, 1994). This also suggests that there may be multiple MAPKKKs, MAPKKs, and MAP kinases in *Arabidopsis*.

Similar phosphorylation cascades composed of MAPKKK, MAPKK, and MAPK may functon in plants

cNPK1, a cDNA encoding a protein kinase that exhibits a high degree of homology to yeast BCK1 (Lee and Levin, 1992) and STE11 (Rhodes *et al.*, 1990), was cloned from tobacco (Banno *et al.*, 1993). Overexpression of the putative catalytic domain of cNPK1 in yeast cells was shown to be able to substitute for defective growth-control functions due to mutations in BCK1 and PKC1 (Levin *et al.*, 1990; Banno *et al.*, 1993). BCK1 and STE11 are protein kinases that are upstream of MKK1/MKK2 (Irie *et al.*, 1993) and STE7 (Teague *et al.*, 1986), respectively, which are homologues of MAPKKs. cNPK2, a cDNA encoding a protein kinase that exhibits a high degree of homology to MAPKKs was also cloned from tobacco (Machida *et al.*, unpublished data).

I found that ATMEKK1 from *Arabidopsis* has extensive homology to byr2 from *S. pombe* (Wang et al., 1991), STE11 from *S. cereviciae* (Rhodes *et al.*, 1990), BCK1 from *S. cereviciae* (Lee and Levin, 1992), MEKK from mouse (Lange-Carter et al., 1993), and NPK1 from tobacco (Banno *et al.*, 1993), all of which are members of the MAPKKK family (see Introduction). Genomic Southern analysis indicated that genes for ATMEKKs constitute a small gene family. Corroborative evidence was obtained through the

isolation of cDNAs corresponding to at least three additional ATMEKK1 related genes in *Arabidopsis*. To isolate cDNAs encoding protein kinases which are structurally related to the reported MAPKKKs in other organisms, I obtained a PCR-amplified DNA fragment which has a high degree of homology with the corresponding region of tobacco NPK1 (Banno *et al.*, 1993) using *Arabidopsis* cDNA as a template (data not shown). Therefore, *Arabidopsis* has at least two types of protein kinases, ATMEKK1-type and NPK1-type, which are structurally related to MAPKKKs. I also obtained a PCR-amplified DNA fragment which has a high degree of homology with MAPKKs (see Introduction) using *Arabidopsis* cDNA as a template (Fig. 26 and 27).

These observations support the probable existence of similar phosphorylation cascades, that include a MAPKKK, a MAPKK and a MAP kinase, in higher plants (Fig. 27).

ATPK6 and ATPK19 from Arabidopsis have high homology with animal ribosomal S6 kinases, and may be the putative downstream factors of the Arabidopsis MAP kinases

I described that ATPK6 and ATPK19 from *Arabidopsis* are more closely related to animal $p70^{S6K}$, $pp90^{rsk}$, protein kinase C, and protein kinase A than to reported plant protein kinases (Fig. 2) (Hanks *et al.*, 1988; Hanks and Quinn, 1991). When this work had been completed, a cDNA clone, *atpk1*, encoding apparently the same protein kinase as ATPK6 was isolated from *Arabidopsis* by Zhang *et al.* (Zhang *et al.*, 1994a; 1994b). They also reported that there is only a closely related gene, *atpk2*, in *Arabidopsis* genome. My genomic Southern blot analysis indicated that there is no

closely related gene to ATPK6 and ATPK19. Therefore, ATPK19 may be *atpk2*.

There have been reported two ribosomal S6 kinases, p90 ribosomal S6 kinase (now referred to as $pp90^{rsk}$ or RSK) (Jones *et al.*, 1988; Alcorta *et al.*, 1989) and p70 ribosomal S6 kinases ($p70^{S6K}$) (Grove *et al.*, 1991; Kozma *et al.*, 1990) in animals. Both kinases are activated by serine/threonine phosphorylation (Downward, 1994). MAP kinase was demonstrated to phosphorylate and partially reactivate dephosphorylated $pp90^{rsk}$ (Sturgill *et al.*, 1988). ATPK6 and ATPK19 have putative PEST sequences in their N- and C-terminal (Rogers *et al.*, 1986; Rechsteiner *et al.*, 1987). $p70^{S6K}$ is also shown to have PEST sequences in their N- and C-terminal, and its C-terminal PEST region is thought to function as autoinhibitory domains (Mukhopadhyay *et al.*, 1992). This region of $p70^{S6K}$ has been demonstrated to be phosphorylated by three proline-directed protein kinases, cdc2, MAP kinase, and p54 MAP-2 kinase, but not to be activated by these kinases (Mukhopadhyay *et al.*, 1992). It has been also reported that phosphorylation by other unidentified protein kinases may function in the activation of the $p70^{S6K}$ (Mukhopadhyay *et al.*, 1992). Thus it can be speculated that ATPK6 and ATPK19 may function as the putative downstream factors of *Arabidopsis* MAP kinases (Fig. 27). This idea is supported by the observation that the mRNA levels of *ATPK19* and *ATMPK3* increase markedly and simultaneously in response to several environmental stresses, as is fully described in the following part.

In animals $p70^{S6K}$ phosphorylates 40S ribosomal protein S6 and activates protein synthesis (Grove *et al.*, 1991; Kozma *et al.*, 1990; Jefferies *et al.*, 1994). Recently, Zhang *et al.* has demonstrated that Atpk1, which is apparently identical with ATPK6, can phosphorylate two plant ribosomal

proteins (Zhang *et al.*, 1994a; 1994b). Thus it is possible to speculate that ATPK6 and ATPK19 may play important roles in increasing the efficiency of protein synthesis through phosphorylation of plant ribosomal proteins.

The expression levels of three protein kinases, ATMPK3, ATPK19, and ATMEKK1, increased markedly and simultaneously in response to cold, touch, and salinity stimuli

I examined the expression of nine genes for MAP kinases (ATMPKs) (Mizoguchi *et al.*, 1993b; 1994a), at the mRNA level in *Arabidopsis* under a variety of stress conditions (Fig. 20). I examined stresses that affect the expression of *ATMPK*s and timing of the induction of *ATMPK*s, and found different pattern of induction of the *ATMPK*s. The different pattern of *ATMPK* induction by different stimuli may reflect different roles of *ATMPK*s in signal transduction under a variety of environmental or developmental stimuli. Furthermore, I found that the pattern of mRNA accumulation of three protein kinases, *ATMPK3*, *ATPK19*, and *ATMEKK1*, by touch, cold, and high-salt treatments resemble each other. By contrast the mRNA levels of *ATMPK1* (Mizoguchi *et al.*, 1994a) and cdc2a (Hirayama *et al.*, 1991) did not change under the same conditions (Fig. 23). In *S. cereviciae,* HOG1 and PBS2, which encode a MAP kinase and a MAPKK, respectively, have been shown to constitute an osmosensing signal transduction pathway (Brewster *et al.*, 1993). Recently, two mammalian protein kinases, JNK1 and p38, have also been demonstrated to be activated in response to osmotic shock, and complement a HOG1 mutation (Dérijard *et al.*, 1994; Han *et al.*, 1994; Brewster *et al.*, 1993; Galcheva-Gargova *et al.*, 1994). In plants, the transduction of low-temperature stimuli has been shown to be mediated by protein phosphorylation (Monroy *et al.*, 1993). Therefore, it is possible to be speculated that three protein kinases, ATMPK3, ATPK19, and ATMEKK1

may play key roles in the signal transduction pathways under low-temperature and osmotic stress conditions.

It has been reported that proteasome constitutes a large polysubunit complex with multi-protease activities and that the positive and negative regulation of the mRNAs encoding multiple proteasomal subunits occur simultaneously (Shimbara *et al.*, 1992). Recent genetic and biochemical evidence have indicated that Ste5 binds simultaneously to Ste11 (a MAPKKK), Ste7 (a MAPKK), and either Fus3 or Kss1 (MAP kinases) and forms a multikinase complex (Choi *et al.*, 1994; Marcus *et al.*, 1994). In *Xenopus* similar possibility that three protein kinases, MAP kinase, MAPKK, and MAPKKK, may exist as a kinase complex, has been demonstrated (Matsuda *et al.*, 1993). Thus these observations give us an idea that three protein kinases, ATMPK3, ATPK19, and ATMEKK1, whose mRNA levels are positively regulated by touch, cold, and high-salt stresses simultaneously, may form a multikinase complex. Protein-protein interactions among these three protein kinases are now investigated.

I described above that mRNA levels of *ATMPK3*, *ATPK19*, and *ATMEKK1* rapidly increase in response to touch stimuli. A variety of plant species including *Arabidopsis* stimulated by touch display altered development (Braam and Davis, 1990; Jaffe, 1973). This morphological change is one of the alterations in plant growth patterns in response to a variety of environmental stimuli. The molecular mechanism by which plants respond to touch stimuli had been largely unknown. However, five cDNAs (TCH1 to five) whose expression levels increase by touch have been isolated from *Arabidopsis*. TCH1, 2, and 3 encode calmodulin or calmodulin-like proteins (Braam and Davis, 1990). It has also reported that touch and cold shock cause immediate increases in cytosolic calcium in

tobacco seedlings (Knight *et al.*, 1992). These results suggest that calcium ions, calmodulin, calmodulin-like proteins, and three protein kinases (ATMPK3, ATPK19, and ATMEKK1) may function in signaling pathways which transduce cold and touch stimuli from the environment.

Conclusion

In this thesis I revealed the presence of genes for the putative components of MAP kinase cascades in a higher plant, *Arabidopsis* (Fig. 27). Nine cDNA clones, designated cATMPK1 to 9, encode protein kinases related to MAP kinases, and they can be classified into four subgroups based on their amino acid sequences. Four cDNA clones (cATMEKK1-4) and one PCR-amplified DNA fragment have high homology to the MAPKKKs. One PCR-amplified DNA fragment, designated pMAPKK, contains partial sequences which shows high sequence homology to MAPKKs (Fig. 26). As the putative downstream factors, I described the isolation of two cDNAs that encode closely related protein kinases, designated ATPK6 and ATPK19, which are related to animal ribosomal S6 protein kinases, $p70^{S6K}$ and $pp90^{rsk}$. The existence of several protein kinases which are related to MAP kinases, MAPKKs, MAPKKKs, and ribosomal S6 protein kinases strongly suggests that there may be multiple similar phosphorylation cascades which function in the transduction of signals elicited by a variety of extracellular stimuli in plants.

Compared with animals and yeasts, that contain two and four members of MAP kinase family, respectively, plants apparently have larger gene family for MAP kinase. It had been reported that only one type of two closely related MAP kinases, ERK1 and ERK2, integrates mitogenic signals from a large number of growth factor receptors, cytokine receptors, several membrane-bound oncogenes, and some mitogenic peptides with G protein-linked receptors. Recently, however, two additional MAP kinase homologues, JNK1/SAPK and p38, have been reported in mammals. They have been shown to function in distinct signal transduction pathways from those of ERK1 and ERK2. Therefore, MAP kinases in eukaryotic cells may contain more members and function in a variety of signaling pathways.

I demonstrated that the mRNA levels of nine MAP kinases (ATMPK1 - 9) are regulated by different environmental stresses. Furthermore, the expression levels of genes for three protein kinases, designated *ATMPK3*, *ATPK19*, and *ATMEKK1*, which are most related to MAP kinases, ribosomal S6 protein kinases, and MAPKKKs, increased markedly and simultaneously in response to touch, cold, and salinity stresses. The increase of the transcripts by these stresses indicates that three protein kinases may function together in signal transduction pathways in plants under these stress conditions. These results suggest that ATMEKK1 and ATPK19 may function as an upstream- and a downstream-factor of ATMPK3, respectively.

The following hypothesis can be deduced from the above observations on the mechanism of the activation of plant MAP kinase cascades in response to a variety of environmental stresses (Fig. 28). In animals, MAP kinase cascades are activated by posttranslational modification such as phosphorylation. In contrast, there are at least two mechanisms of the activation of MAP kinase cascades in plants. One is the posttranslational modification such as phosphorylation, which is the same mechanism as that of animals. The other is the transcriptional regulation of genes involved in MAP kinase cascades. The mRNA levels of certain components of the cascades markedly increase after treatments with various stimuli, which probably results in the increase of the components in response to such stimuli. The higher level of activities is likely to amplify the efficiency of signal transductin by phosphorylation. This type of regulation may include three protein kinases, ATMPK3, ATPK19, and ATMEKK1, in *Arabidopsis*.

Figures and Tables

Figure 1. Catalytic domains of eukaryotic protein kinases (Hanks *et al.*, 1988; Hanks and Quinn, 1991). Primary structure of seven diverse members of the protein kinase family are shown by the horizontal bars. The green portions of bars indicate the homologous catalytic domains. cAPK, cGPK, PKC, CaMIIK, cdc2, Src, and EGFR represent cAMP-dependent protein kinase catalytic subunit, cGMP-dependent protein kinase, calcium-phospholipid-dependent protein kinase, calcium/calmodulin-dependent protein kinase II, cell division control protein kinase 2, cellular homologue of Rous sarcoma virus transforming protein, and epidermal growth factor receptor, respectively. Conserved 11 catalytic subdomains are also represented by the roman numerals (I - XI) within the larger green bar. Amino acid residues or motifs that are highly conserved throughout the entire family are shown in single letter code with nonconserved positions represented by 'x'. Two pairs of arrowheads represented by (1) and (2) indicate the regions used to generate the primers for PCR.

(A)

```
GTCCGAGCGTTGGCCAAATCTCTCTCACTTCCACAGGTTTCTCTCTCCGGCCAAATCTAAGCTCCGGGGAACGTCGTTGGTCACTTATCACCGAGGGAAA   100
ACAAAAAATGGCGACTTTGGTTGATCCTCCTAATGGGATAAGGAATGAAGGGAAGCATTACTTCTCAATGTGGCAAACTCTGTTCGAGATCGACACTAAG   200
        M  A  T  L  V  D  P  P  N  G  I  R  N  E  G  K  H  Y  F  S  M  W  Q  T  L  F  E  I  D  T  K

TACATGCCTATCAAGCCTATTGGTCGTGGAGCTTACGGTGTTGTCTGCTCCTCTGTTAACAGTGACACCAACGAGAAAGTTGCTATCAAGAAGATTCACA   300
Y  M  P  I  K  P  I  G  R  G  A  Y  G  V  V  C  S  S  V  N  S  D  T  N  E  K  V  A  I  K  K  I  H  N

ATGTTTATGAGAATAGGATCGATGCGTTGAGGACTCTTCGGGAGCTCAAGCTTCTACGCCATCTTCGACATGAGAATGTCATTGCTTTGAAAGATGTCAT   400
 V  Y  E  N  R  I  D  A  L  R  T  L  R  E  L  K  L  L  R  H  L  R  H  E  N  V  I  A  L  K  D  V  M

GATGCCAATTCATAAGATGAGCTTCAAGGATGTTTATCTTGTTTATGAGCTCATGGACACTGATCTCCACCAGATTATCAAGTCTTCTCAGCGTCTTAGT   500
M  P  I  H  K  M  S  F  K  D  V  Y  L  V  Y  E  L  M  D  T  D  L  H  Q  I  I  K  S  S  Q  R  L  S

AACGATCATTGCCAATACTTCTTGTTCCAGTTGCTTCGAGGGCTCAAGTATATTCATTCAGCCAATATCCTGCACCGAGATTTGAAACCTGGTAACCTTC   600
N  D  H  C  Q  Y  F  L  F  Q  L  L  R  G  L  K  Y  I  H  S  A  N  I  L  H  R  D  L  K  P  G  N  L  L

TTGTCAACGCAAACTGCGATTTAAAGATATGCGATTTTGGACTAGCGCGTGCGAGCAACACCAAGGGTCAGTTCATGACTGAATATGTTGTGACTCGTTG   700
 V  N  A  N  C  D  L  K  I  C  D  F  G  L  A  R  A  S  N  T  K  G  Q  F  M  T  E  Y  V  V  T  R  W

GTACCGAGCCCCAGAGCTTCTCCTCTGTTGTGACAACTATGGAACATCCATTGATGTTTGGTCTGTTGGTTGCATTTTCGCCGAGCTTCTTGGTAGGAAA   800
 Y  R  A  P  E  L  L  L  C  C  D  N  Y  G  T  S  I  D  V  W  S  V  G  C  I  F  A  E  L  L  G  R  K

CCGATATTCCAAGGAACGGAATGTCTTAACCAGCTTAAGCTCATTGTCAACATTATCGGAAGCCAAAGAGAAGAAGATCTTGAGTTCATAGTTAACCCGA   900
P  I  F  Q  G  T  E  C  L  N  Q  L  K  L  I  V  N  I  I  G  S  Q  R  E  E  D  L  E  F  I  V  N  P  K

AAGCTAAAAGATACATTAGATCACTTCCGTACTCACCTGGGATGTCTTTATCCAGACTTTACCCGTGCGCTCATGTATTGGCCATCGACCTTCTGCAGAA  1000
 A  K  R  Y  I  R  S  L  P  Y  S  P  G  M  S  L  S  R  L  Y  P  C  A  H  V  L  A  I  D  L  L  Q  K

AATGCTTGTTTTTGATCCGTCAAAGAGGATTAGTGCCTCTGAAGCACTCCAGCATCCATACATGGCGCCACTATATGACCCGAATGCAAACCCTCCTGCT  1100
M  L  V  F  D  P  S  K  R  I  S  A  S  E  A  L  Q  H  P  Y  M  A  P  L  Y  D  P  N  A  N  P  P  A

CAAGTTCCTATCGATCTCGATGTAGATGAGGATTTGAGAGAGGAGATGATAAGAGAAATGATATGGAATGAGATGCTTCACTACCATCCACAAGCTTCAA  1200
Q  V  P  I  D  L  D  V  D  E  D  L  R  E  E  M  I  R  E  M  I  W  N  E  M  L  H  Y  H  P  Q  A  S  T

CCTTAAACACTGAGCTCTGAGCTCAAGTCTTGTTTGTACGGGTAATTTACAGAAAACTTCTTCTTCTTATGTCTGATTGTCATCATAGACTCATAGTGTA  1300
 L  N  T  E  L  *

TATAGTCTTGAAAAATAAGATGAAGACTAACTTATAGTTTAAGCGAATAGTGATGCCATGGAAGCTCTGTTTTATTTAATTACAAGCTTGATGTGTGTCT  1400
GTAACATATGTACATAGAGAGAGCTGTTTTTTTTTTTTAATTACAAGTTTGATGTGTGTCTGTAACATATGTACATAGAAAGAGCTGTGTTTTTTTTTTA  1500
AAAAAAAAAAAAAAAAAAAA                                                                                  1520
```

(B)

```
CCAAATCTCTCACTCTCACCAGCCAAATCTCTCCAACTTCCTCTTTCTCTCCACCCAAATCTCGCCGGCGATCGTCACTGACGAATCTGATTCGTCCCCA   100
ACACTCAACGGGAGGTAGAAGAATGGCGACTCCTGTTGATCCACCTAATGGAATTAGGAATCAAGGGAAGCATTACTTCTCAATGTGGCAAACACTTTTC   200
                M  A  T  P  V  D  P  P  N  G  I  R  N  Q  G  K  H  Y  F  S  M  W  Q  T  L  F

GAGATCGATACCAAATACATGCCTATCAAACCGATAGGCCGAGGCGCGTACGGTGTGGTTTGCTCTTCGGTTAACAGAGAGAGTAATGAGAGAGTGGCGA   300
E  I  D  T  K  Y  M  P  I  K  P  I  G  R  G  A  Y  G  V  V  C  S  S  V  N  R  E  S  N  E  R  V  A  I

TCAAGAAGATCCACAATGTGTTTGAGAATAGGATTGATGCGTTGAGGACTCTTAGGGAGCTCAAGCTTCTACGTCATCTTCGACATGAGAATGTGGTTGC   400
 K  K  I  H  N  V  F  E  N  R  I  D  A  L  R  T  L  R  E  L  K  L  L  R  H  L  R  H  E  N  V  V  A

TCTTAAAGATGTAATGATGGCTAATCATAAGAGAAGCTTTAAGGATGTTTATCTTGTTTATGAGCTTATGGATACTGATCTTCATCAGATTATTAAGTCT   500
 L  K  D  V  M  M  A  N  H  K  R  S  F  K  D  V  Y  L  V  Y  E  L  M  D  T  D  L  H  Q  I  I  K  S

TCTCAAGTTCTAAGTAATGACCATTGCCAATACTTCTTGTTCCAGTTGCTTCGAGGGCTCAAGTATATTCATTCAGCAAACATTCTCCATCGGGATCTGA   600
S  Q  V  L  S  N  D  H  C  Q  Y  F  L  F  Q  L  L  R  G  L  K  Y  I  H  S  A  N  I  L  H  R  D  L  K

AACCCGGTAACCTCCTTGTGAATGCAAACTGCGACTTAAAGATATGTGACTTTGGGCTAGCGAGGACGAGCAACACCAAAGGTCAGTTCATGACTGAATA   700
 P  G  N  L  L  V  N  A  N  C  D  L  K  I  C  D  F  G  L  A  R  T  S  N  T  K  G  Q  F  M  T  E  Y

TGTTGTGACTAGATGGTACCGAGCACCAGAGCTACTCCTCTGTTGTGACAACTATGGAACCTCCATTGATGTCTGGTCAGTCGGTTGCATATTCGCCGAG   800
 V  V  T  R  W  Y  R  A  P  E  L  L  L  C  C  D  N  Y  G  T  S  I  D  V  W  S  V  G  C  I  F  A  E

CTTCTTGGAAGAAAACCAGTATTCCCGGGAACAGAATGCCTAAACCAGATTAAACTCATCATTAACATTTTGGGTAGCCAGAGAGAGGAAGATCTCGAGT   900
L  L  G  R  K  P  V  F  P  G  T  E  C  L  N  Q  I  K  L  I  I  N  I  L  G  S  Q  R  E  E  D  L  E  F

TTATAGATAACCCAAAAGCCAAAAGATACATAGAATCACTCCCTTACTCACCAGGGATATCATTCTCTCGTCTTTACCCGGGTGCAAATGTTTTAGCCAT  1000
 I  D  N  P  K  A  K  R  Y  I  E  S  L  P  Y  S  P  G  I  S  F  S  R  L  Y  P  G  A  N  V  L  A  I

TGATCTGCTTCAGAAAATACTCGTTCTTGACCCTTCGAAAAGGATTAGTGTCACGGAAGCGCTTCAACATCCATACATGGCGCCTTTATATGACCCGAGT  1100
 D  L  L  Q  K  I  L  V  L  D  P  S  K  R  I  S  V  T  E  A  L  Q  H  P  Y  M  A  P  L  Y  D  P  S

GCAAATCCTCCTGCTCAAGTTCCTATTGATCTCGATGTAGATGAAGACGAGGATTTGGGAGCAGAGATGATAAGAGAATTAATGTGGAAGGAAATGATTC  1200
A  N  P  P  A  Q  V  P  I  D  L  D  V  D  E  D  E  D  L  G  A  E  M  I  R  E  L  M  W  K  E  M  I  H

ATTATCATCCAGAAGCTGCTACCATAAACAACAATGAGGTCTCTGAGTTTTGAGTTCAAGTCTTTTGCAGGTACTGTTCAGAGAGATCTTTCAACAACTT  1300
 Y  H  P  E  A  A  T  I  N  N  N  E  V  S  E  F  *

AAACTTATTTATTTTCATGTTTGCTTCTTGAAATGTTGTTAATAATAGTGTCGAAGAAGAAGAACAATAACTTTAAATTATGTGATATGTTTGTGTCATG  1400
GGGCTTTGTTTTGTTTTTGGTTACAAGCTTAAGTGTCTGTAACGTTTGTACATAAGTGTATGTGTCCTTAAAACTTCATACTCTGTCTCTGATCTTATTG  1500
TTGGTGATTTTTCTTATTGAAGATTTTCGAATGTTCAAAGAGCAAACCATTGGATTAAAAAA                                         1562
```

Figure 2. Nucleotide and deduced amino acid sequences of two cDNAs, *ATMPK1* (A) and *ATMPK2* (B). The DNA sequences include the putative coding regions and the 5'- and 3'-noncoding regions. The amino acid sequences of the putative coding regions are shown beneath the DNA sequences. The sequences corresponding to the probe (a 560-bp PCR-amplified sequence) that was used to screen the library are underlined. The repeated sequences, CCAAATCT(C/A)(T/A/G)C, in the putative 5'-noncoding regions are boxed.

```
                                                                  I
ATMPK1          MATLVDPPNGIRNEGKHYFSMWQTLFEIDTKYMPIKPIGRGAY   43
ATMPK2          ***P*********Q***************************   43
MsERK1    MEGGGAPP*DT*MSDAAPAPPQMGIENIPAV*SHGGRFKP**M***K***   50
  MPK1              MAAAGAAS*P*GGPEMVRGQA*DVGPR*INLAY**E***   39
  ERK2              MAAAAAA-----GPEMVRGQV*DVGPR*TNLSY**E***   34
  SPK1    MASATSTPTIADGNS*KESVATSRSPHTHDLN**LPEE*EM*NL**Q***   50
                          II                  III       IV
ATMPK1    GVVCSSVNSDTNEKVAIKKIHNVYENRIDALRTLRELK-LLRHLRHENVI   92
ATMPK2    ********RES**R*********F**************-**********V   92
MsERK1    *I***AH**E***H**V***A*AFD*K***K*****I*L*--**MD****V   99
  MPK1    *M***AHDNVNKVR******SP-F*HQTYCQ*****I*I***FK-***I*   87
  ERK2    *M***AYDNLNKVR******SP-F*HQTYCQ*****I*I***FR-***I*   82
  SPK1    ****AALHKPSGL***V****P-FNHPVFC******I*-****F****I*   98
                                  V
ATMPK1    ALKDVMMPIHKMSFKDVYLVYELMDTDLHQIIKSSQRLSNDHCQYFLFQL  142
ATMPK2    *******AN**R************************V*************  142
MsERK1    *IR*IVP*PQREV*N***IA************R*N*A**EE******Y*I  149
  MPK1    GIN*IIRAPTIEQM****I*QD**E***YKLL*-T*H*****IC***Y*I  136
  ERK2    GIN*IIRAPTIEQM****I*QD**E***YKLL*-T*H*****IC***Y*I  131
  SPK1    SIL*ILP*PSYQELE***I*Q***E***YRV*R-**P**D******TY*I  147
                     VI                   VII
ATMPK1    LRGLKYIHSANILHRDLKPGNLLVNANCDLKICDFGLAR-ASNTKGQ--F  189
ATMPK2    ***************************************-T******--*  189
MsERK1    ***********V*******S***L***************VT*E*D----*  195
  MPK1    ***********V*******S***L*TT************V*DPDHDHTG*  186
  ERK2    ***********V*******S***L*TT************V*DPDHDHTG*  181
  SPK1    **A**AM***GVV******S***L*******VA******STTAQG*NPG*  197
                VIII                      IX
ATMPK1    MTEYVVTRWYRAPELLLCCDNYGTSIDVWSVGCIFAELLGRKPIFQGTEC  239
ATMPK2    ******************************************V*P****  239
MsERK1    ****************NSSD*TAA**********M**MD***L*P*RDH  245
  MPK1    L****A********IM*NSKG*TK***I******L**M*SNR***P*KHY  236
  ERK2    L****A********IM*NSKG*TK***I******L**M*SNR***P*KHY  231
  SPK1    *****A********IM*SFRE*SKA**L**T***L**M*SAR*L*P*KDY  247
                     X
ATMPK1    LNQLKLIVNIIGSQREEDLEFIVNPKAKRYIRSLPYSPGMSLSRLYPCAH  289
ATMPK2    ***I***I**L***********D********E*******I*F*****G*N  289
MsERK1    VH**R*LMEL**TPS*D**G*L-*EN******Q**PYRRQ*FQEKF*HV*  294
  MPK1    *D**NH*LG*L**PSQ***NC*I*L**RN*LL***HKNKVPWN**F*N*D  286
  ERK2    *D**NH*LG*L**PSQ***NC*I*L**RN*LL***HKNKVPWN**F*N*D  281
  SPK1    HS*IT**L**L*TPTMD*FSR*KSAR*RK**K***FT*KV*FKA*F*Q*S  297
                XI
ATMPK1    VLAIDLLQKMLVFDPSKRISASEALQHPYMAPLYDPNANPPAQVPIDLDV  339
ATMPK2    *********I**L*******VT**************S*************  339
MsERK1    PE****VE***T***R***TVED**A***LTS*H*ISDE*VCMT*FSF*F  344
  MPK1    PK*L***D***T*N*H***EVEA**A***LEQY***SDE*V*EA*FKFEM  336
  ERK2    SK*L***D***T*N*H***EVEQ**A***LEQY***SDE*I*EA*FKF*M  331
  SPK1    PD*****E*L*T*N*D***T*E***K***V*AYH*ASDE*T*SPMPPNL*  347

ATMPK1    DE--DLREEMIREMIWNEMLHYHPQASTLNTEL                  370
ATMPK2    **DE**GA*****LM*K**I****E*A*I*NNEVSEF              376
MsERK1    EQ-HA*T**QMK*L*YR*A*AFN*EYQQ                       371
  MPK1    EL-D**PK*TLK*L*FE*TARFQ*GY                         361
  ERK2    EL-D**PK*KLK*L*FE*TARFQ*GYRS                       358
  SPK1    *L--YCNK*DLEIPVLKALIFREVNFR                        372
```

Figure 3. Comparison of the deduced amino acid sequences of ATMPK1, ATMPK2 (Mizoguchi *et al.*, 1994a), alfalfa MsERK1 (Duerr *et al.*, 1993), *Xenopus* MPK1 (Gotoh *et al.*, 1991), rat ERK2 (Boulton *et al.*, 1991) and yeast spk1 (Toda *et al.*, 1991). Asterisks represent identical amino acid residues and dashes indicate gaps introduced to maximize alignment. Roman numerals indicate the eleven major conserved subdomains of protein kinases identified previously (Hanks *et al.*, 1988; Hanks and Quinn, 1991). Amino acid residues that are invariant or nearly invariant among known protein kinases are indicated by closed triangles. Closed circles show residues that correspond to the sites of regulatory phosphorylation reported for *Xenopus* MPK1 (Gotoh *et al.*, 1991) and rat ERK1 (Payne *et al.*, 1991; Gartner *et al.*, 1992).

Figure 4. Northern analysis of the expression of the *ATMPK1* and *ATMPK2* genes in a variety of organs of *Arabidopsis*. For detection of the respective mRNAs, total RNA (40 μg) prepared from leaves (lanes 1, 6 and 11), seeds (lanes 2, 7 and 12), stems (lanes 3, 8 and 13), roots (lanes 4, 9 and 14) and flowers (lanes 5, 10 and 15) was subjected to electrophoresis on 1% agarose gels that contained formaldehyde (Maniatis *et al.*, 1982). Nitrocellulose filters blotted with RNA were hybridized with [^{32}P]-labeled *ATMPK1* or *ATMPK2* cDNAs. The sizes of RNA markers are indicated in kb. A nitrocellulose filter blotted with RNA was stained with methylene blue to show that similar amounts of RNA were loaded per lane. The two prominent bands are rRNA.

Figure 5. Southern blot analysis of genomic sequences that correspond to *ATMPK1* and *ATMPK2*. Genomic DNA was digested with *Bam*HI (B), *Eco*RI (E), and *Hin*dIII (H), fractionated on 0.7% agarose gels and transferred to nylon membranes. Filters were hybridized with [^{32}P]-labeled fragments of *ATMPK1* and *ATMPK2* cDNA at 42°C and washed in either 0.5x SSC/0.5% SDS at 50°C (low-stringency) or 0.1x SSC/0.1% SDS at 65°C (high-stringency). ATMPK1 and ATMPK2 indicate the probe DNA used and High and Low represent high- and low-stringency hybridization conditions, respectively. The sizes of DNA markers are indicated in kbp.

Figure 6. (A) Structures of the ATMPK1 and ATMPK2 fusion proteins expressed in *E. coli* cells. GST, A, ATMPK1 and ATMPK2 indicate the partial sequence of the gene for glutathione-S-transferase derived from the pGEX vector, the *EcoRI* adapter (5'-AATTCGAGGATCCGGGTACCATGG-3'), *ATMPK1* cDNA and *ATMPK2* cDNA, respectively. Numbers (1, 1215, 1250) above the boxes show the nucleotide numbers of *ATMPK1* and *ATMPK2* cDNAs, respectively (Fig. 1). Arrows indicate the sites of the predicted first ATG triplets of *ATMPK1* and *ATMPK2* cDNAs. **(B)** Expression and purification of recombinant ATMPK1 and ATMPK2 proteins from *E. coli* cells containing pGEX-MPK1 and pGEX-MPK2, respectively. The insoluble pellet fraction (P) and the soluble fraction (S) of crude extracts prepared from IPTG-treated *E. coli* cells containing pGEX-MPK1 (ATMPK1) or pGEX-MPK2 (ATMPK2) were analyzed by SDS-PAGE. Fusion proteins (G) purified by use of glutathione-Sepharose beads from the soluble fraction were also analyzed by SDS-PAGE. The closed and open triangles indicate the electrophoretic positions of the full-length and truncated recombinant MAP kinases, respectively. The positions and sizes (KDa) of molecular-weight markers (M) are also indicated.

Figure 7. Activation of the recombinant ATMPK1 and ATMPK2 proteins by *Xenopus* MAPKK. Partially purified GST-ATMPK1 and GST-ATMPK2 were incubated for 30 min at 30 °C with (+) or without the purified *Xenopus* MAPKK (5 μg/ml) and [γ-^{32}P]ATP (50 μM) as described in Materials and Methids. The samples were then subjected to electrophoresis in a 12% acrylamide gel and then the gel was autoradiographed. **(A)** Phosphorylation of ATMPK1 and ATMPK2. Closed and open triangles denote the electrophoretic positions of ATMPK1 and ATMPK2, and *Xenopus* MAPKK, respectively. **(B)** Phosphoamino acid analysis of phosphorylated bands in **(A)**. The bands corresponding to GST-ATMPK1 of lanes 1 and 3 in **(A)** were excised and the labeled phosphoamino acids were analyzed in lanes 1 and 2 in **(B)**, respectively, as described elsewhere (Boyle *et al.*, 1991). The positions of the phosphoamino acid standards, Ser, Thr and Tyr, are also indicated as pS, pT and pY, respectively. **(C)** Results of "test-tube" assays of MBP-phosphorylating activity (in arbitrary units, see Materials and Methods). The MBP-phosphorylating activity of purified *Xenopus* MAPKK alone was negligible (Matsuda *et al.*, 1992; Kosako *et al.*, 1992), and is not shown here.

Figure 8. Activation of the recombinant ATMPK2 protein by a crude extract of tobacco BY-2 cells that had been treated with 2,4-D. **(A)** Changes in mitotic index of BY-2 cells during treatment with 2,4-D (see Materials and Methods). **(B)** Time course of the activation of the 46-KDa MBP kinase. The quiescent BY-2 cells were treated with 1 μM 2,4-D for the indicated times. MBP kinase activities of cell extracts were assayed on a gel that contained MBP (0.5 mg/ml) as described in Materials and Methods. Closed and open triangles show the electrophoretic positions of the 46-KDa MBP kinase and of a 30-KDa MBP kinase whose activity did not respond to treatment with 2,4-D, respectively. The position and sizes (KDa) of molecular-weight markers (M) are also indicated. **(C)** Phosphorylation of the purified GST-ATMPK2 by extracts of 2,4-D treated BY-2 cells. GST-ATMPK2 was incubated for 20 min at 30 °C in the absence (lane 1) or presence of extracts of

BY-2 cells that had been treated with 2,4-D for 0 min (lane 2), 5 min (lane 3) and 10 min (lane 4). The extract of BY-2 cells treated with 2,4-D for 10 min alone was also incubated (lane 5). The samples were then subjected to electrophoresis in a 12% acrylamide gel and then the gel was autoradiographed. (D) Relative levels of phosphorylation of the purified GST-ATMPK2 by extracts of 2,4-D treated BY-2 cells. Relative levels of phosphorylation were calculated from the intensity of the bands (in arbitrary units, see Materials and Methods). Lanes 1, 2, 3 and 4 correspond to the bands indicated by the arrowhead in Fig. 8 (C).

(A) ATMPK3

```
        10        20        30        40        50        60        70        80        90       100
CGACTTTGATCAATTGAGAGAGAAATGAACACCGGCGGTGGCCAATACACGGATTTTCCGGCGGTGGACACTCACGGAGGACAGTTCATAAGTTACGATA
                          M  N  T  G  G  G  Q  Y  T  D  F  P  A  V  D  T  H  G  G  Q  F  I  S  Y  D  I

       110       120       130       140       150       160       170       180       190       200
TCTTCGGTAGTTTATTCGAGATCACATCTAAGTATCGTCCTCCGATTATACCAATTGGTCGTGGAGCTTATGGAATCGTTTGCTCTGTGTTGGATACGGA
 F  G  S  L  F  E  I  T  S  K  Y  R  P  P  I  I  P  I  G  R  G  A  Y  G  I  V  C  S  V  L  D  T  E

       210       220       230       240       250       260       270       280       290       300
GACGAACGAGCTAGTAGCGATGAAGAAGATAGCTAATGCTTTTGATAATCATATGGATGCTAAGCGTACGCTTCGTGAGATCAAGCTTCTTCGTCATCTT
 T  N  E  L  V  A  M  K  K  I  A  N  A  F  D  N  H  M  D  A  K  R  T  L  R  E  I  K  L  L  R  H  L

       310       320       330       340       350       360       370       380       390       400
GATCATGAAAACATTATAGCTATAAGAGATGTTGTTCCACCACCACTAAGAAGACAGTTCAGTGATGTTTATATCTCTACTGAATTAATGGATACTGATC
D  H  E  N  I  I  A  I  R  D  V  V  P  P  P  L  R  R  Q  F  S  D  V  Y  I  S  T  E  L  M  D  T  D  L

       410       420       430       440       450       460       470       480       490       500
TTCATCAAATCATCAGATCTAACCAGAGTTTATCAGAAGAACACTGTCAGTACTTCTTGTACCAGCTACTTCGAGGACTGAAGTATATCCACTCAGCTAA
 H  Q  I  I  R  S  N  Q  S  L  S  E  E  H  C  Q  Y  F  L  Y  Q  L  L  R  G  L  K  Y  I  H  S  A  N

       510       520       530       540       550       560       570       580       590       600
CATTATTCATAGGGATTTAAAGCCGAGCAATCTTCTGTTGAACGCGAATTGCGATTTAAAGATTTGTGATTTCGGTCTTGCTAGACCTACTTCAGAGAAT
 I  I  H  R  D  L  K  P  S  N  L  L  L  N  A  N  C  D  L  K  I  C  D  F  G  L  A  R  P  T  S  E  N

       610       620       630       640       650       660       670       680       690       700
GATTTTATGACTGAGTATGTTGTTACGAGATGGTATAGAGCACCTGAGCTTCTGTTGAACTCTTCAGATTACACAGCTGCTATTGATGTTTGGTCTGTTG
D  F  M  T  E  Y  V  V  T  R  W  Y  R  A  P  E  L  L  L  N  S  S  D  Y  T  A  A  I  D  V  W  S  V  G

       710       720       730       740       750       760       770       780       790       800
GTTGTATCTTTATGGAGCTTATGAATAGAAAGCCTTTGTTCCCTGGTAAAGACCATGTTCATCAAATGCGCTTATTGACAGAGTTGCTTGGCACACCGAC
 C  I  F  M  E  L  M  N  R  K  P  L  F  P  G  K  D  H  V  H  Q  M  R  L  L  T  E  L  L  G  T  P  T

       810       820       830       840       850       860       870       880       890       900
AGAATCTGATCTCGGTTTTACTCACAATGAGGATGCGAAAAGATACATCCGGCAACTTCCCAACTTCCCACGTCAGCCCTTAGCTAAACTTTTCTCTCAT
 E  S  D  L  G  F  T  H  N  E  D  A  K  R  Y  I  R  Q  L  P  N  F  P  R  Q  P  L  A  K  L  F  S  H

       910       920       930       940       950       960       970       980       990      1000
GTTAACCCAATGGCCATTGATCTTGTTGACAGAATGTTGACGTTTGACCCCAACAGAAGAATCACTGTTGAACAAGCTCTGAATCACCAGTACCTTGCTA
V  N  P  M  A  I  D  L  V  D  R  M  L  T  F  D  P  N  R  R  I  T  V  E  Q  A  L  N  H  Q  Y  L  A  K

      1010      1020      1030      1040      1050      1060      1070      1080      1090      1100
AATTGCACGACCCGAATGATGAGCCAATCTGTCAAAAGCCATTCTCTTTTGAGTTCGAACAACAGCCTCTGGATGAGGAACAGATAAAAGAGATGATCTA
 L  H  D  P  N  D  E  P  I  C  Q  K  P  F  S  F  E  F  E  Q  Q  P  L  D  E  E  Q  I  K  E  M  I  Y

      1110      1120      1130      1140      1150      1160      1170      1180      1190      1200
CCAAGAAGCCATAGCACTCAATCCAACATACGGTTAGAAGTGCAGCAGCCCCGTGAATGCCTGGTATTACCCAATAACCATCCGAATGGCTACTTAGTAT
 Q  E  A  I  A  L  N  P  T  Y  G  *

      1210      1220      1230      1240      1250      1260      1270      1280      1290      1300
CTTTGCCTGTTCTTTTATTCATGTACTTCTTCCACATATGTAATCTTGTAACTTTTATTTGTTTGTTCATGTTATTTACTGCTAGTGATTAAGTGTAGCT

      1310      1320      1330      1340      1350      1360      1370
CCAATGTAAAGTACTCCAGTTATTAAAGAATATCTATAGACTTTTTTACTTCACATAAAAAAAAAAAAAA
```

Figure 9. Nucleotide and deduced amino acid sequences of seven cDNAs, *ATMPK3*.(A), *ATMPK4*.(B), *ATMPK5*.(C), *ATMPK6*.(D), *ATMPK7*.(E), *ATMPK8*.(F), and *ATMPK9* (G). The DNA sequences include the putative coding regions and the 5'- and 3'-noncoding regions. The amino acid sequences of the putative coding regions are shown beneath the DNA sequences.

(B) **ATMPK4**

```
        10        20        30        40        50        60        70        80        90       100
CGTTTTCGTTGTTGTGAGGAATTTTGCTCCGATGTCGGCGGAGAGTTGTTTCGGAAGCTCGGGTGATCAAAGCAGCAGCAAAGGAGTGGCTACTCATGGT
                                  M  S  A  E  S  C  F  G  S  S  G  D  Q  S  S  S  K  G  V  A  T  H  G

       110       120       130       140       150       160       170       180       190       200
GGTAGCTATGTTCAGTATAATGTCTATGGCAATCTCTTCGAAGTTTCTAGAAAGTATGTTCCTCCTCTTCGTCCTATCGGTAGAGGCGCTTATGGAATTG
G  S  Y  V  Q  Y  N  V  Y  G  N  L  F  E  V  S  R  K  Y  V  P  P  L  R  P  I  G  R  G  A  Y  G  I  V

       210       220       230       240       250       260       270       280       290       300
TCTGTGCTGCTACAAACTCAGAGACTGGAGAGGAGGTAGCTATCAAGAAGATTGGTAATGCTTTTGACAATATTATCGATGCTAAAAGGACTTTGAGAGA
 C  A  A  T  N  S  E  T  G  E  E  V  A  I  K  K  I  G  N  A  F  D  N  I  I  D  A  K  R  T  L  R  E

       310       320       330       340       350       360       370       380       390       400
GATTAAGCTTCTCAAGCATATGGATCATGAAAATGTGATTGCTGTAAAGGATATAATAAAACCACCGCAGAGAGAGAACTTCAATGATGTTTACATTGTT
 I  K  L  L  K  H  M  D  H  E  N  V  I  A  V  K  D  I  I  K  P  P  Q  R  E  N  F  N  D  V  Y  I  V

       410       420       430       440       450       460       470       480       490       500
TATGAGCTTATGGACACTGATCTCCACCAAATTATTCGCTCTAACCAACCCTTAACTGATGATCATTGTCGGTTTTTTCTGTATCAGTTGTTGCGTGGGC
Y  E  L  M  D  T  D  L  H  Q  I  I  R  S  N  Q  P  L  T  D  D  H  C  R  F  F  L  Y  Q  L  L  R  G  L

       510       520       530       540       550       560       570       580       590       600
TCAAATATGTTCATTCAGCTAATGTATTGCATCGAGATTTGAAACCTAGCAATTTGCTCCTGAATGCAAATTGTGATCTAAAGCTTGGGGATTTCGGGCT
 K  Y  V  H  S  A  N  V  L  H  R  D  L  K  P  S  N  L  L  L  N  A  N  C  D  L  K  L  G  D  F  G  L

       610       620       630       640       650       660       670       680       690       700
TGCGAGGACCAAATCCGAGACTGACTTTATGACTGAATATGTTGTTACACGTTGGTATCGAGCTCCAGAGCTGCTACTTAATTGCTCTGAATACACAGCA
 A  R  T  K  S  E  T  D  F  M  T  E  Y  V  V  T  R  W  Y  R  A  P  E  L  L  L  N  C  S  E  Y  T  A

       710       720       730       740       750       760       770       780       790       800
GCAATTGATATTTGGTCTGTCGGTTGTATACTCGGTGAAACAATGACACGAGAGCCCTTGTTTCCCGGTAAAGATTATGTTCATCAGCTTAGACTCATCA
A  I  D  I  W  S  V  G  C  I  L  G  E  T  M  T  R  E  P  L  F  P  G  K  D  Y  V  H  Q  L  R  L  I  T

       810       820       830       840       850       860       870       880       890       900
CTGAGCTTATAGGATCTCCTGATGACTCAAGCTTGGGATTCTTAAGAAGTGACAATGCAAGAAGATACGTTAGACAGCTTCCACAGTACCCTAGACAGAA
 E  L  I  G  S  P  D  D  S  S  L  G  F  L  R  S  D  N  A  R  R  Y  V  R  Q  L  P  Q  Y  P  R  Q  N

       910       920       930       940       950       960       970       980       990      1000
CTTTGCTGCTAGATTCCCAAACATGTCGGCTGGTGCAGTCGATTTGCTGGAGAAAATGCTTGTTTTTGAACCAAGCAGACGCATCACAGTTGATGAGGCG
 F  A  A  R  F  P  N  M  S  A  G  A  V  D  L  L  E  K  M  L  V  F  E  P  S  R  R  I  T  V  D  E  A

      1010      1020      1030      1040      1050      1060      1070      1080      1090      1100
TTGTGCCACCCATATTTGGCGCCGCTGCATGATATCAACGAAGAACCGGTATGTGTGAGGCCGTTCAATTTTGATTTCGAGCAACCTACTTTGACAGAAG
L  C  H  P  Y  L  A  P  L  H  D  I  N  E  E  P  V  C  V  R  P  F  N  F  D  F  E  Q  P  T  L  T  E  E

      1110      1120      1130      1140      1150      1160      1170      1180      1190      1200
AGAACATTAAGGAGCTTATATACCGTGAAACAGTCAAGTTCAATCCTCAAGACTCAGTGTGAGAGAAAGAGAGAGAGATATATATCCTTAAAAGGAAACT
 N  I  K  E  L  I  Y  R  E  T  V  K  F  N  P  Q  D  S  V  *

      1210      1220      1230      1240      1250      1260      1270      1280      1290      1300
CATCAAAATGCTATGTTTTGTTTTGCTCTGCTTTTGTTATCAAATCTCTGATAAGTTTGATAATATGTCTGTAATTTGAATAACATCCTAAAAAATTGAA

AAAAAAA
```

(C) **ATMPK5**

```
        10        20        30        40        50        60        70        80        90       100
CGAATCGAATCGATCTCATCTCCAAAAATCTTGTATTCAAAGTTTGATTTTCTGGGTTTCTTCTGTATTATCTCGATATTTGAAAAAAATGGCGAAGGAA
                                                                                          M  A  K  E

       110       120       130       140       150       160       170       180       190       200
ATTGAATCAGCGACAGATTTAGGTGATACTAACATCAAAGGAGTTCTTGTTCATGGAGGAAGATATTTTCAATACAATGTTTATGGAAACTTGTTTGAAG
I  E  S  A  T  D  L  G  D  T  N  I  K  G  V  L  V  H  G  G  R  Y  F  Q  Y  N  V  Y  G  N  L  F  E  V

       210       220       230       240       250       260       270       280       290       300
TTTCTAATAAGTATGTTCCTCCTATTCGACCTATTGGTAGAGGAGCTTACGGTTTTGTCTGTCCTGCAGTAGATTCGGAGACACACGAAGAGATTGCGAT
 S  N  K  Y  V  P  P  I  R  P  I  G  R  G  A  Y  G  F  V  C  P  A  V  D  S  E  T  H  E  E  I  A  I

       310       320       330       340       350       360       370       380       390       400
AAAGAAGATAGGTAAAGCCTTTGATAATAAAGTAGATGCAAAGAGAACATTGAGAGAGATTAAGCTACTTCGTCATTTAGAACACGAGAACGTTGTTGTT
 K  K  I  G  K  A  F  D  N  K  V  D  A  K  R  T  L  R  E  I  K  L  L  R  H  L  E  H  E  N  V  V  V

       410       420       430       440       450       460       470       480       490       500
ATAAAAGATATCATTAGACCACCAAAGAAGGAAGATTTCGTTGATGTTTATATAGTTTTTGAGTTAATGGATACTGATCTTCATCAGATTATTCGATCAA
I  K  D  I  I  R  P  P  K  K  E  D  F  V  D  V  Y  I  V  F  E  L  M  D  T  D  L  H  Q  I  I  R  S  N

       510       520       530       540       550       560       570       580       590       600
ACCAATCTTTGAATGATGATCATTGTCAGTACTTCTTGTATCAGATTCTACGAGGACTCAAATATATTCATTCAGCAAATGTTTTACATCGTGACCTTAA
 Q  S  L  N  D  D  H  C  Q  Y  F  L  Y  Q  I  L  R  G  L  K  Y  I  H  S  A  N  V  L  H  R  D  L  K

       610       620       630       640       650       660       670       680       690       700
GCCGAGTAATTTGCTTTTGAATTCGAATTGCGATCTTAAGATTACTGACTTTGGTTTGGCAAGGACTACATCAGAAACAGAGTATATGACAGAGTATGTT
 P  S  N  L  L  L  N  S  N  C  D  L  K  I  T  D  F  G  L  A  R  T  T  S  E  T  E  Y  M  T  E  Y  V

       710       720       730       740       750       760       770       780       790       800
GTAACACGATGGTACAGAGCACCAGAGCTGCTTCTTAATAGTTCTGAGTATACATCTGCTATTGATGTATGGTCTGTTGGTTGCATATTTGCTGAGATTA
V  T  R  W  Y  R  A  P  E  L  L  L  N  S  S  E  Y  T  S  A  I  D  V  W  S  V  G  C  I  F  A  E  I  M

       810       820       830       840       850       860       870       880       890       900
TGACAAGGGAACCGCTTTTCCCGGGGAAGGATTACGTTCATCAGCTTAAACTTATCACTGAGTTGATAGGATCACCTGATGGAGCGAGTCTAGAGTTTCT
 T  R  E  P  L  F  P  G  K  D  Y  V  H  Q  L  K  L  I  T  E  L  I  G  S  P  D  G  A  S  L  E  F  L

       910       920       930       940       950       960       970       980       990      1000
CAGAAGCGCGAACGGAGGAAAATATGTTAAGGAGCTACCTAAGTTCCCAAGACAAAACTTTTCAGCTAGGTTTCCTAGTATGAATTCTACAGCTATCGAT
 R  S  A  N  G  G  K  Y  V  K  E  L  P  K  F  P  R  Q  N  F  S  A  R  F  P  S  M  N  S  T  A  I  D

      1010      1020      1030      1040      1050      1060      1070      1080      1090      1100
CTACTAGAGAAAATGCTCGTGTTTGATCCAGTAAAACGTATCACGGTTGAGGAAGCACTGTGTTATCCTTACTTGTCAGCTCTTCACGACCTTAACGACG
L  L  E  K  M  L  V  F  D  P  V  K  R  I  T  V  E  E  A  L  C  Y  P  Y  L  S  A  L  H  D  L  N  D  E

      1110      1120      1130      1140      1150      1160      1170      1180      1190      1200
AGCCTGTTTGTTCCAACCATTTCAGTTTCCATTTTGAGGATCCGTCTTCCACTGAAGAAGAGATAAAGGAGCTTGTGTGGCTCGAATCTGTGAAGTTCAA
 P  V  C  S  N  H  F  S  F  H  F  E  D  P  S  S  T  E  E  E  I  K  E  L  V  W  L  E  S  V  K  F  N

      1210      1220      1230      1240      1250      1260      1270      1280      1290      1300
TCCTCTGCCGAGCATTTAAAACATTATTTAAGAAGGGGAAAGCAAGTTTTATACTCTGACATTGAGCAAAACAAATTAGTAGTGTTTAAGCTCTGTTTCT
 P  L  P  S  I  *

      1310      1320      1330      1340      1350      1360      1370      1380      1390      1400
CCTTCATCTCATGTTGATCGTTGTTTAAGCTATTTCACTAGTTAAAAAAGCTGTCTTGCAACTCAAGAAATTGTTCTATAAGGATCCTTGTAACCCAAGA

      1410      1420      1430      1440      1450      1460      1470      1480      1490      1500
CATATCTTTAGTAGTTATGTCCATTCTTATTATCAGGATGTCTTAGATAGTGTTCTGTTTCTTCTACATATCATGTTGTTATCTTGTAACTCAAGAAATT

      1510      1520      1530
AAATTAAAAAGCTATCTTTGTTTAAAAAAAAAA
```

(D) **ATMPK6**

```
        10        20        30        40        50        60        70        80        90       100
AAAAGATTTTCCATTTCCCTCTTCTCGATCCACCACCACCACCAACCTCAAAAGCTTTCATCTTTACCGGTCATGGACGGTGGTTCAGGTCAACCGGCGG
                                                                          M  D  G  G  S  G  Q  P  A  A

       110       120       130       140       150       160       170       180       190       200
CTGATACAGAGATGACAGAAGCTCCTGGAGGTTTTCCCGCGGCGGCTCCATCGCCTCAGATGCCTGGGATTGAGAATATTCCGGCGACTCTTAGCCATGG
 D  T  E  M  T  E  A  P  G  G  F  P  A  A  A  P  S  P  Q  M  P  G  I  E  N  I  P  A  T  L  S  H  G

       210       220       230       240       250       260       270       280       290       300
TGGTAGGTTTATTCAGTATAACATATTTGGAAACATCTTCGAGGTCACCGCTAAGTATAAGCCTCCGATCATGCCTATTGGCAAGGGTGCTTATGGCATC
 G  R  F  I  Q  Y  N  I  F  G  N  I  F  E  V  T  A  K  Y  K  P  P  I  M  P  I  G  K  G  A  Y  G  I

       310       320       330       340       350       360       370       380       390       400
GTTTGTTCGGCTATGAATTCTGAAACTAACGAGAGCGTTGCGATTAAGAAAATTGCTAACGCTTTTGACAATAAGATTGATGCTAAGAGGACTCTCCGTG
V  C  S  A  M  N  S  E  T  N  E  S  V  A  I  K  K  I  A  N  A  F  D  N  K  I  D  A  K  R  T  L  R  E

       410       420       430       440       450       460       470       480       490       500
AGATCAAGCTGCTTCGTCACATGGATCATGAAAATATTGTTGCAATCAGAGATATTATCCCGCCACCATTAAGAAACGCTTTCAACGATGTTTACATCGC
 I  K  L  L  R  H  M  D  H  E  N  I  V  A  I  R  D  I  I  P  P  P  L  R  N  A  F  N  D  V  Y  I  A

       510       520       530       540       550       560       570       580       590       600
GTATGAGTTAATGGACACTGATCTCCATCAAATCATTCGGTCAAATCAAGCATTATCCGAAGAACATTGCCAGTATTTTCTTTACCAGATCCTCCGTGGA
 Y  E  L  M  D  T  D  L  H  Q  I  I  R  S  N  Q  A  L  S  E  E  H  C  Q  Y  F  L  Y  Q  I  L  R  G

       610       620       630       640       650       660       670       680       690       700
TTGAAATACATTCACTCTGCAAATGTGCTTCACAGGGATTTGAAACCAAGTAATCTCCTCCTGAACGCAAACTGCGACCTAAAAATCTGCGATTTTGGGC
L  K  Y  I  H  S  A  N  V  L  H  R  D  L  K  P  S  N  L  L  L  N  A  N  C  D  L  K  I  C  D  F  G  L

       710       720       730       740       750       760       770       780       790       800
TAGCTCGAGTCACTTCTGAGAGTGATTTCATGACTGAATATGTTGTCACGAGATGGTACCGTGCACCAGAGCTTCTCTTAAACTCTTCTGATTATACTGC
 A  R  V  T  S  E  S  D  F  M  T  E  Y  V  V  T  R  W  Y  R  A  P  E  L  L  L  N  S  S  D  Y  T  A

       810       820       830       840       850       860       870       880       890       900
AGCTATCGATGTTTGGTCTGTAGGCTGTATTTTCATGGAGTTAATGGACCGTAAGCCACTCTTCCCTGGACGAGATCATGTCCATCAGCTTCGCTTGCTC
 A  I  D  V  W  S  V  G  C  I  F  M  E  L  M  D  R  K  P  L  F  P  G  R  D  H  V  H  Q  L  R  L  L

       910       920       930       940       950       960       970       980       990      1000
ATGGAGCTCATAGGAACTCCATCAGAAGAAGAGCTCGAGTTCTTGAACGAAAACGCAAAGCGATACATAAGACAGCTTCCACCTTATCCTCGCCAATCCA
M  E  L  I  G  T  P  S  E  E  E  L  E  F  L  N  E  N  A  K  R  Y  I  R  Q  L  P  P  Y  P  R  Q  S  I

      1010      1020      1030      1040      1050      1060      1070      1080      1090      1100
TCACTGATAAGTTCCCGACAGTGCATCCTTTAGCTATAGACCTTATCGAGAAGATGTTAACATTTGATCCTAGACGGAGAATCACAGTTTTAGACGCATT
 T  D  K  F  P  T  V  H  P  L  A  I  D  L  I  E  K  M  L  T  F  D  P  R  R  R  I  T  V  L  D  A  L

      1110      1120      1130      1140      1150      1160      1170      1180      1190      1200
GGCCCATCCATACCTGAACTCGTTGCACGACATAAGCGATGAGCCAGAGTGTACAATACCTTTCAACTTTGATTTTGAAAACCATGCACTCTCAGAGGAG
 A  H  P  Y  L  N  S  L  H  D  I  S  D  E  P  E  C  T  I  P  F  N  F  D  F  E  N  H  A  L  S  E  E

      1210      1220      1230      1240      1250      1260      1270      1280      1290      1300
CAGATGAAGGAACTAATCTACCGCGAGGCGCTTGCTTTCAATCCAGAATATCAGCAATAGTTTGAAGAACAGCATTTGGGAGATGAAGATGAGATAAGGG
Q  M  K  E  L  I  Y  R  E  A  L  A  F  N  P  E  Y  Q  Q  *

      1310      1320      1330      1340      1350      1360      1370      1380      1390      1400
TTTGGCGTTATTTGGTTTGGACAACATTCTTAAAAGGAGGAAAGATCATTGTTTGTCTTTTGTGTATACTTGTGAGAGATGTTGAAGGGGATGGGTCTCA

      1410      1420      1430      1440      1450      1460      1470      1480      1490      1500
AACTTTATATTTCAGTCTTTTGGATAATTTGTTCTCTCTGTTTTTTTTTTTCCTATTCGTATTTGTAATTTGCCCTCATTTTGTTGTTGCAGTTATTTAA

      1510      1520      1530      1540      1550      1560
GACTTAAATTGTCTCAGAAATAGTTATGTTATGATAGCCCACCTTCAATTCAAAAAAAAAAA
```

```
        10        20        30        40        50        60        70        80        90       100
GTCTCACCACTTCCTAACTCTTCTCTCTCTTATCTCTTCTTGTTTTAGCTATGAAGAAGCAATCTCAATTCAGATCCAATCATACCCTTCTCTTCTTTCA

       110       120       130       140       150       160       170       180       190       200
TCTCTGATCATTAACAACAACAACAACGACTCGGGTTCTGTAAATTCAGTTATTGGAGAGAAAATGGCGATGTTAGTTGAGCCACCAAATGGGATTAAGC
                                                                M  A  M  L  V  E  P  P  N  G  I  K  Q

       210       220       230       240       250       260       270       280       290       300
AACAAGGGAAACATTACTACTCCATGTGGCAAACATTGTTCGAAATCGATACAAAGTATGTTCCGATTAAACCCATTGGAAGAGGAGCTTATGGTGTGGT
 Q  G  K  H  Y  Y  S  M  W  Q  T  L  F  E  I  D  T  K  Y  V  P  I  K  P  I  G  R  G  A  Y  G  V  V

       310       320       330       340       350       360       370       380       390       400
TTGTTCTTCTATCAATCGTGAGACTAATGAGAGAGTTGCTATTAAAAAGATTCACAATGTGTTTGAGAATCGTGTTGATGCTTTGAGAACGTTGAGGGAA
 C  S  S  I  N  R  E  T  N  E  R  V  A  I  K  K  I  H  N  V  F  E  N  R  V  D  A  L  R  T  L  R  E

       410       420       430       440       450       460       470       480       490       500
CTTAAACTTCTTAGGCATGTGAGGCATGAGAATGTGATCGCGCTTAAAGATGTTATGTTGCCTGCTAATAGAACCAGTTTCAAGGATGTTTATTTGGTTT
L  K  L  L  R  H  V  R  H  E  N  V  I  A  L  K  D  V  M  L  P  A  N  R  T  S  F  K  D  V  Y  L  V  Y

       510       520       530       540       550       560       570       580       590       600
ATGAGCTAATGGATACTGATCTTCATCAGATTATTAAATCCTCTCAGTCTCTTTCTGATGATCACTGCAAGTACTTCTTGTTTCAGTTGCTTAGAGGGTT
 E  L  M  D  T  D  L  H  Q  I  I  K  S  S  Q  S  L  S  D  D  H  C  K  Y  F  L  F  Q  L  L  R  G  L

       610       620       630       640       650       660       670       680       690       700
GAAGTATCTTCATTCTGCAAACATACTTCACCGGGATTTGAAGCCCGGGAACCTCTTAGTTAATGCGAACTGCGATCTCAAGATTTGTGATTTCGGGTTG
 K  Y  L  H  S  A  N  I  L  H  R  D  L  K  P  G  N  L  L  V  N  A  N  C  D  L  K  I  C  D  F  G  L

       710       720       730       740       750       760       770       780       790       800
GCGAGAACGAGCCAAGGTAATGAACAGTTCATGACTGAGTATGTGGTTACACGTTGGTACAGAGCACCTGAGCTTCTTCTTTGTTGTGATAACTATGGAA
A  R  T  S  Q  G  N  E  Q  F  M  T  E  Y  V  V  T  R  W  Y  R  A  P  E  L  L  L  C  C  D  N  Y  G  T

       810       820       830       840       850       860       870       880       890       900
CCTCCATTGATGTATGGTCGGTTGGTTGCATATTCGCGGAGATTCTTGGTAGGAAACCGATTTTTCCAGGGACAGAATGCTTGAATCAGCTTAAGCTAAT
 S  I  D  V  W  S  V  G  C  I  F  A  E  I  L  G  R  K  P  I  F  P  G  T  E  C  L  N  Q  L  K  L  I

       910       920       930       940       950       960       970       980       990      1000
CATCAACGTTGTTGGTAGCCAACAAGAATCTGATATTCGGTTCATAGACAACCCGAAAGCTCGAAGGTTTATAAAGTCTCTTCCGTACTCAAGAGGAACT
 I  N  V  V  G  S  Q  Q  E  S  D  I  R  F  I  D  N  P  K  A  R  R  F  I  K  S  L  P  Y  S  R  G  T

      1010      1020      1030      1040      1050      1060      1070      1080      1090      1100
CATCTCTCCAATCTTTATCCACAAGCCAATCCTCTAGCTATAGATTTGTTACAGAGGATGCTTGTGTTTGAACCAACCAAGAGAATCTCTGTAACCGATG
H  L  S  N  L  Y  P  Q  A  N  P  L  A  I  D  L  L  Q  R  M  L  V  F  E  P  T  K  R  I  S  V  T  D  A

      1110      1120      1130      1140      1150      1160      1170      1180      1190      1200
CGCTCTTACACCCGTATATGGCGGGGTTGTTTGAACCTGGAACCAATCCGCCTGCACATGTCCCAATCTCTCTGGACATAGATGAAAACATGGAGGAACC
 L  L  H  P  Y  M  A  G  L  F  E  P  G  T  N  P  P  A  H  V  P  I  S  L  D  I  D  E  N  M  E  E  P

      1210      1220      1230      1240      1250      1260      1270      1280      1290      1300
AGTGATAAGAGAGATGATGTGGAATGAGATGCTTTATTACCACCCTGAAGCTGAAATCTTAAATGCCTAAATCAACATCAATGATCATTAAGAGTAAGCT
 V  I  R  E  M  M  W  N  E  M  L  Y  Y  H  P  E  A  E  I  L  N  A  *

      1310      1320      1330      1340      1350      1360      1370      1380      1390      1400
TTTTCCAGTTTTTTTTTTTTTTGCTTTCTAAAATGCAGCTTTGATGTTATGTTATATATAGCGAGTCCACAAACGCGGGAACAGATCGATAAAAAGCCCG

      1410      1420      1430      1440      1450      1460      1470      1480      1490      1500
GTGCTTTGATCTCGTCTGGTTTGATTTAGTTGCAGTGAGTGAGTAGTAGTAACAGAGTTTGAAAAGTGGCTCTCAGTATCTTTACCACTTTGTAAATGTG

      1510      1520      1530      1540      1550      1560      1570      1580      1590      1600
ACTGTTTGTTACTAAGTAAGTTAAGGCTTGTTTTTGTTACATTATCTTTGAGTTGATGAGTTTTACTACTGATTTTCAGGCCAACTCTTGTATTTTGTAT

      1610      1620      1630      1640      1650
AGCTTTGCCATTGCCCCACTTGCTAGCCACACGAAATAATTTTTTAAAAAAAAAAAA
```

```
        10        20        30        40        50        60        70        80        90       100
CGAAGGACAAAACCGAGAGACGCGGCGAGAAAGACGGTGACTTGACGTGGTCTATGGGTGGTGGTGGGAATCTCGTCGACGGTGTTCGTCGATGGCTTTT
                                                                M  G  G  G  G  N  L  V  D  G  V  R  R  W  L  F

       110       120       130       140       150       160       170       180       190       200
TCAGCGACCTTCTTCTTCTTCTTCTTCTTCTTCCAACAATAATAATAATAATCACGAACAACCCATTTTCAATTCTTCTTCTTTCTCTTCTTCCTCT
 Q  R  P  S  S  S  S  S  S  S  S  N  N  N  N  N  N  H  E  Q  P  I  F  N  S  S  S  F  S  S  S  S

       210       220       230       240       250       260       270       280       290       300
AACCCCAATCATTCCGCTAACTCTGGAGAGCTTATTATTGAAGAAGATCTCGATTTCTCTGGTCTAACGCTTATTAACGTTCCTAAACGTAATCATTTAC
N  P  N  H  S  A  N  S  G  E  L  I  I  E  E  D  L  D  F  S  G  L  T  L  I  N  V  P  K  R  N  H  L  P

       310       320       330       340       350       360       370       380       390       400
CTATGGATCCTCACAAGAAGGGTGAAACTGAATTCTTCACGGAGTATGGAGAAGCTAACAGGTATCAGATTCAAGAGGTTGTTGGCAAAGGAAGTTATGG
 M  D  P  H  K  K  G  E  T  E  F  F  T  E  Y  G  E  A  N  R  Y  Q  I  Q  E  V  V  G  K  G  S  Y  G

       410       420       430       440       450       460       470       480       490       500
TGTTGTGGCTTCTGCTGTTGACTCACACACTGGAGAGAGAGTTGCGATCAAGAAGATCAACGATGTGTTTGAGCATGTCTCCGATGCTACTAGGATTCTC
 V  V  A  S  A  V  D  S  H  T  G  E  R  V  A  I  K  K  I  N  D  V  F  E  H  V  S  D  A  T  R  I  L

       510       520       530       540       550       560       570       580       590       600
AGAGAGATCAAATTGCTTCGGTTGCTTCGTCATCCTGATGTTGTGGAGATTAAACACATCATGTTACCTCCTTCTCGCAGAGAATTCCGCGATATTTATG
R  E  I  K  L  L  R  L  L  R  H  P  D  V  V  E  I  K  H  I  M  L  P  P  S  R  R  E  F  R  D  I  Y  V

       610       620       630       640       650       660       670       680       690       700
TTGTCTTTGAGTTGATGGAATCTGATCTTCATCAAGTGATAAAGGCGAATGATGATTTGACTCCTGAGCATTATCAGTTTTTCTTGTACCAGCTTCTTCG
 V  F  E  L  M  E  S  D  L  H  Q  V  I  K  A  N  D  D  L  T  P  E  H  Y  Q  F  F  L  Y  Q  L  L  R

       710       720       730       740       750       760       770       780       790       800
TGGTCTGAAATATGTTCACGCAGCTAATGTGTTTCATCGGGATTTGAAACCGAAGAACATTCTAGCTAATGCTGATTGCAAATTGAAGATCTGTGATTTT
 G  L  K  Y  V  H  A  A  N  V  F  H  R  D  L  K  P  K  N  I  L  A  N  A  D  C  K  L  K  I  C  D  F

       810       820       830       840       850       860       870       880       890       900
GGGCTGGCTCGCGTTTCTTTTAACGATGCCCCAACCGCTATATTTTGGACTGATTATGTAGCTACTCGGTGGTATCGTGCCCCTGAACTCTGTGGATCGT
G  L  A  R  V  S  F  N  D  A  P  T  A  I  F  W  T  D  Y  V  A  T  R  W  Y  R  A  P  E  L  C  G  S  F

       910       920       930       940       950       960       970       980       990      1000
TTTTCTCCAAATATACCCCTGCGATTGATATCTGGAGTGTCGGTTGCATTTTTGCGGAAATGCTTTTGGGTAAGCCTTTGTTTCCCGGGAAAAACGTTGT
 F  S  K  Y  T  P  A  I  D  I  W  S  V  G  C  I  F  A  E  M  L  L  G  K  P  L  F  P  G  K  N  V  V

      1010      1020      1030      1040      1050      1060      1070      1080      1090      1100
GCACCAATTGGATCTAATGACTGACTTTCTTGGTACTCCACCTCCTGAGTCAATATCAAGGATTAGAAATGAAAAGGCGAGGAGATATCTTAGCAGCATG
 H  Q  L  D  L  M  T  D  F  L  G  T  P  P  P  E  S  I  S  R  I  R  N  E  K  A  R  R  Y  L  S  S  M

      1110      1120      1130      1140      1150      1160      1170      1180      1190      1200
AGGAAAAAACAACCAGTCCCTTTCTCTCATAAATTCCCTAAAGCTGAACCTTTGGCTCTCCGCCTTCTGGAACGCCTGCTTGCCTTTGATCCGAAAGATC
R  K  K  Q  P  V  P  F  S  H  K  F  P  K  A  E  P  L  A  L  R  L  L  E  R  L  L  A  F  D  P  K  D  R

      1210      1220      1230      1240      1250      1260      1270      1280      1290      1300
GTGCATCAGCTGAAGATGCACTAGCTGATCCATATTTCAGTGGTTTGTCAAATTCAGAGCGTGAACCAACAACACAGCCAATTTCAAAGCTTGAATTTGA
 A  S  A  E  D  A  L  A  D  P  Y  F  S  G  L  S  N  S  E  R  E  P  T  T  Q  P  I  S  K  L  E  F  D

      1310      1320      1330      1340      1350      1360      1370      1380      1390      1400
CTTTGAGAGAAAGAAGTTAGTTAAAGATGATGTCAGAGAATTACTTTACCGAGAGATATTGGAATATCATCCTCAGATGTTGGAGGAATACCTTCGTGGT
 F  E  R  K  K  L  V  K  D  D  V  R  E  L  L  Y  R  E  I  L  E  Y  H  P  Q  M  L  E  E  Y  L  R  G

      1410      1420      1430      1440      1450      1460      1470      1480      1490      1500
GGTGATCAGCTTAGCTTTATGTACCCAAGTGGGGTAGACCGATTTAAAAGGCAATTTGCTCATCTTGAAGAAAATCAAGGTAAACCAGGAGCAGCAGGAG
G  D  Q  L  S  F  M  Y  P  S  G  V  D  R  F  K  R  Q  F  A  H  L  E  E  N  Q  G  K  P  G  A  A  G  G

      1510      1520      1530      1540      1550      1560      1570      1580      1590      1600
GAGGAAGAAGTACTGCACTTCACAGACATCATGCTTCCTTGCCAAGAGAGAGAGTTCCTGCTCCAAATGGTGAAACTGCAGAAGAAAGCAGTGATGTTGA
 G  R  S  T  A  L  H  R  H  H  A  S  L  P  R  E  R  V  P  A  P  N  G  E  T  A  E  E  S  S  D  V  E

      1610      1620      1630      1640      1650      1660      1670      1680      1690      1700
GAGAAGAGCAGCAGCTGCTGTGGCTTCAACCTTGGAATCCGAGGAGCAGACAATGGAGGAGGTTACAGTGCTCGTAACCTCATGAGGAGTGCAAGCATCA
 R  R  A  A  A  A  V  A  S  T  L  E  S  E  E  Q  T  M  E  E  V  T  V  L  V  T  S  *

      1710      1720      1730      1740      1750      1760      1770      1780      1790      1800
GCGGTTCTAAATGCATTGGTGTCCATCTAAACTGCAACAAAGAGGATACCATAGCTGAGGAAGAAGATAATGAAACAGTTGCAGAGCTTACTGATAAAGT

      1810      1820      1830      1840      1850      1860      1870      1880      1890      1900
TGCTTCTCTTCACAATTCTTAAAAATGATTTAGTTATTTTTTTTCTTTTCTTTTCTTTTTTCACGAAAGTTTTCTGGTGAAAAGAAATTGGTTTCTCCAT

      1910      1920      1930      1940      1950      1960      1970      1980      1990      2000
TCATCCACATCTATCTTCGTGGTGGGTTTGATTGTTGTATTTAATAGTTTTGGGGGTTTGATGAAACTTTAAATAAAAACATCGATCTCCATGTAATT

      2010      2020      2030      2040      2050      2060      2070      2080
TTTTTGTTTCTCTCTTTTCTTTAAACACTCATATTTTGAAATTATTTATATACTATTCAGTATTCACTTTGCTTCAGTAAAAAAAA
```

```
        10        20        30        40        50        60        70        80        90       100
CGTACCCATGGGGAAAAGTTTCTCCCTTTGTTAGAGATTTTAATTTGTCTGGTTTGAACTGCATTAGGGTTTCTAGAAGGAATTATATTCTCATGGATCC
                                                                                          M  D  P

       110       120       130       140       150       160       170       180       190       200
TCATAAAAAGGTTGCACTGGAGACTGAGTTCTTTACAGAATACGGTGAAGCTAGCCGATACCAAATCCAAGAAGTTATAGGGAAAGGAAGTTACGGTGTT
 H  K  K  V  A  L  E  T  E  F  F  T  E  Y  G  E  A  S  R  Y  Q  I  Q  E  V  I  G  K  G  S  Y  G  V

       210       220       230       240       250       260       270       280       290       300
GTGGCATCTGCAATCGACACACATTCATGCGAAAAGGTTGCCATTAAGAAAATCAATGATGTGTTTGAGCACGTCTCTGATGCAACGCGGATTCTAAGAG
V  A  S  A  I  D  T  H  S  C  E  K  V  A  I  K  K  I  N  D  V  F  E  H  V  S  D  A  T  R  I  L  R  E

       310       320       330       340       350       360       370       380       390       400
AGATCAAGCTTCTCAGGTTGCTCCGCCATCCTGATATAGTGGAGATAAAGCATGTAATGCTTCCTCCTTCACGCCGTGAGTTCAGAGACATTTATGTTGT
 I  K  L  L  R  L  L  R  H  P  D  I  V  E  I  K  H  V  M  L  P  P  S  R  R  E  F  R  D  I  Y  V  V

       410       420       430       440       450       460       470       480       490       500
TTTCGAGTTGATGGAATCAGATCTTCACCAAGTTATCAAGGCGAACGATGATCTTACACCTGAGCACTACCAGTTTTTCCTCTACCAGCTTCTTCGTGGT
 F  E  L  M  E  S  D  L  H  Q  V  I  K  A  N  D  D  L  T  P  E  H  Y  Q  F  F  L  Y  Q  L  L  R  G

       510       520       530       540       550       560       570       580       590       600
CTGAAGTTCATACACACAGCTAATGTGTTTCATCGTGATTTGAAGCCTAAAAATATTCTAGCCAACTCTGATTGCAAGTTGAAGATTTGCGACTTTGGGC
L  K  F  I  H  T  A  N  V  F  H  R  D  L  K  P  K  N  I  L  A  N  S  D  C  K  L  K  I  C  D  F  G  L

       610       620       630       640       650       660       670       680       690       700
TTGCTCGAGTGTCCTTTAATGATGCACCTTCGGCTATATTCTGGACTGACTATGTTGCTACAAGATGGTACCGTGCGCCAGAACTTTGTGGCTCTTTCTT
 A  R  V  S  F  N  D  A  P  S  A  I  F  W  T  D  Y  V  A  T  R  W  Y  R  A  P  E  L  C  G  S  F  F

       710       720       730       740       750       760       770       780       790       800
CTCCAAATATACTCCTGCCATTGATATATGGAGCATAGGATGCATATTTGCAGAAATGCTCACTGGGAAGCCATTGTTTCCTGGGAAGAATGTGGTGCAC
 S  K  Y  T  P  A  I  D  I  W  S  I  G  C  I  F  A  E  M  L  T  G  K  P  L  F  P  G  K  N  V  V  H

       810       820       830       840       850       860       870       880       890       900
CAATTGGATATTATGACTGATTTGCTTGGTACTCCTCCTCCCGAAGCTATTGCAAGGATTCGTAATGAAAAGGCGAGGAGATATCTCGGCAATATGAGGA
Q  L  D  I  M  T  D  L  L  G  T  P  P  P  E  A  I  A  R  I  R  N  E  K  A  R  R  Y  L  G  N  M  R  R

       910       920       930       940       950       960       970       980       990      1000
GAAAACCGCCAGTTCCATTCACTCACAAATTTCCTCATGTAGACCCATTGGCCCTTCGCTTACTGCACCGTCTTCTTGCATTTGATCCCAAAGACCGTCC
 K  P  P  V  P  F  T  H  K  F  P  H  V  D  P  L  A  L  R  L  L  H  R  L  L  A  F  D  P  K  D  R  P

      1010      1020      1030      1040      1050      1060      1070      1080      1090      1100
CTCAGCTGAAGAGGCACTGGCAGATCCTTACTTCTATGGTCTGGCAAACGTTGATCGGGAACCATCTACTCAACCGATTCCGAAACTCGAGTTTGAGTTT
 S  A  E  E  A  L  A  D  P  Y  F  Y  G  L  A  N  V  D  R  E  P  S  T  Q  P  I  P  K  L  E  F  E  F

      1110      1120      1130      1140      1150      1160      1170      1180      1190      1200
GAAAGAAGGAAGTTAACAAAAGAAGATGTGAGAGAATTGATCTACAGAGAGATATTAGAGTATCATCCTCAGATGCTGCAAGAATATCTTCGAGGTGGAG
E  R  R  K  L  T  K  E  D  V  R  E  L  I  Y  R  E  I  L  E  Y  H  P  Q  M  L  Q  E  Y  L  R  G  G  E

      1210      1220      1230      1240      1250      1260      1270      1280      1290      1300
AACAGACTAGCTTCATGTACCCAAGTGGTGTTGATCGCTTTAAGAGACAATTTGCTCATCTTGAGGAAAATTACGGTAAGGGTGAGAAAGGCTCACCTCT
 Q  T  S  F  M  Y  P  S  G  V  D  R  F  K  R  Q  F  A  H  L  E  E  N  Y  G  K  G  E  K  G  S  P  L

      1310      1320      1330      1340      1350      1360      1370      1380      1390      1400
TCAACGCCAACACGCTTCCTTGCCTAGAGAGCGAGTCCCTGCTCCTAAGAAAGAGAATGGATCACATAATCACGATATCGAGAACAGATCCATAGCGTCT
 Q  R  Q  H  A  S  L  P  R  E  R  V  P  A  P  K  K  E  N  G  S  H  N  H  D  I  E  N  R  S  I  A  S

      1410      1420      1430      1440      1450      1460      1470      1480      1490      1500
TTAGTCACGACCCTCGAAAGCCCTCCAACGTCACAACACGAGGGGTCAGACTACAGGAATGGGACGAGCCAGACGGGTTACAGCGCTCGCAGCTTGCTGA
L  V  T  T  L  E  S  P  P  T  S  Q  H  E  G  S  D  Y  R  N  G  T  S  Q  T  G  Y  S  A  R  S  L  L  K

      1510      1520      1530      1540      1550      1560      1570      1580      1590      1600
AAAGTGCCAGCATCAGTGCATCTAAATGCATTGGCATGAAACCAAGAAACAAGTCCGAGTACGGGGAATCAAACAATGATACGGTTGATGCTCTGTCTCA
 S  A  S  I  S  A  S  K  C  I  G  M  K  P  R  N  K  S  E  Y  G  E  S  N  N  D  T  V  D  A  L  S  Q

      1610      1620      1630      1640      1650      1660      1670      1680      1690      1700
AAAGGTCGCGGCTCTCCACACTTGACTTGTTTCAGAGATTGAGGATTGAAACTTGTTGCTAACAGTAGACGGCTGATTCTACAAGTCTTGTACACATCTC
 K  V  A  A  L  H  T  *

      1710      1720      1730      1740      1750      1760      1770      1780      1790      1800
TCAGAGAAGACGAAGAAGAAGTCTGATCTTGATAGAGATATTGAGAAACATGACTCATGAGCTATGAACAAAACAACTTAACCTCTAATTGGTCTTCAAT

      1810      1820      1830      1840      1850      1860      1870      1880      1890      1900
TTGTCCATCAAGTCGAGTTGTATTAAAAATGGATGGAGAATATAATTAAAGGAGAATTGTGTTAACCCCTGGATTTATGCCTTACAAGTATTTGACTTTA

      1910      1920      1930      1940
TCAGGCATTATAAATTATAGTTAAAGTGTGTGTTTCCTTCAAAAAAAA
```

```
ATMPK1                                                                                                  MATLVDPPNGIRNEGKHYFSMWQTLFE   27
ATMPK2                                                                                                  ***P********Q************   27
ATMPK3                                                                                        MNTGGGQYTDF*AVDTHGGQFISYDIFGS***   32
ATMPK4                                                                                  MSAESMFGSSGDQSSSKGVATHGGSYVQYNVYGN***   37
ATMPK5                                                                                  MAKEIESATDLGDTNIKGVLVHGGRYFQYNVYGN***   37
ATMPK6                                                                  MDGGSGQPAADTEMTEAPGGFPAAAPSPQMPGIENI*ATLSHGGRFIQYNIFGNI**   57
ATMPK7                                                                                          **M**E*****KQQ****Y********   27
ATMPK8   MGGGGNLVDGVRRWLFQRPSSSSSSSSSNNNNNNHEQPIFNSSSFSSSSNPNHSANSGELIIEEDLDFSGLTLINVPKRNHLPMDPHK*--GETEFFTEY   98
ATMPK9                                                                                                  MDPHK*VALETEFFTEY   17
XLMPK1                                                                                        MAAAGAAS*P*GGPEMVRGQA*D   23
                                 I                   II                  III          IV                 V
ATMPK1   ID-TKYM-PIKPIGRGAYGVVCSSVNSDTNEKVAIKKIHNVYENRIDALRTLRELKLLRHLRHENVIALKDVMMPIHKMSFKDVYLVYELMDTDLHQIIK  125
ATMPK2   **-****-*****************RES**R*********F**********************V*******AN**R*******************  125
ATMPK3   *T-S**RP**I********I***VLDTE***L**M***A*AFD*HM**K*****I******D***I**IR**VP*PLRRQ*S***IST***********R  130
ATMPK4   VS-R**VP*LR********I**AAT**E*G*E*****G*AFD*I***K*****I***K*MD******V**IIK*PQREN*N***I************R  136
ATMPK5   VS-N**VP**R********F**PA*D*E*H*EI*****GKAFD*KV**K*****I******E****VVI**IIR*PK*ED*V***I*F***********R  136
ATMPK6   VT-A**KP**M***K****I***AM**E***S******A*AFD*K***K*****I******MD***IV*IR*IIP*PLRNA*N***IA************R  156
ATMPK7   **-***V-****************I*RE***R*********F***V**************V************L*ANRT*******************  125
ATMPK8   GEANR*Q-IQEVV*K*S****A*A*D*H*G*R******ND*F*HVS**T*I***I****L***PD*VEI*HI*L*PSRRE*R*I*V*F***ES****V**  197
ATMPK9   GEASR*Q-IQEV**K*S****A*AIDTHSC********ND*F*HVS**T*I***I****L***PDIVEI*H**L*PSRRE*R*I*V*F***ES****V**  116
XLMPK1   VG-PR*I-NLAY**E****M***AHDNVNKVR******SPFEHQTYCQ-*****I*I*LRFK***I*GIN*IIRAPTIEQM****I*QD**E***YKLL*  120
                                     VI                  VII                    VIII                   IX
ATMPK1   SSQRLSNDHCQYFLFQLLRGLKYIHSANILHRDLKPGNLLVNANCDLKICDFGLAR---ASNTKGQFMTEYVVTRWYRAPELLLCC-DNYGTSIDVWSVG  221
ATMPK2   ***V*****************************************************---T*************************-************  221
ATMPK3   *N*S**EE******Y**************I******S***L***************----PTSEND*****************NS-SD*TAA*******  226
ATMPK4   *N*P*TD***RF**Y********V****V*******S***L*******LG******----TKSETD*****************N*-SE*TAA**I****  231
ATMPK5   *N*S*ND*******Y*I***********V*******S***L*S******T******----TTSETEY*****************NS-SE*TSA*******  231
ATMPK6   *N*A**EE******Y*I***********V*******S***L***************----VTSESD*****************NS-SD*TAA********  251
ATMPK7   ***S**D***K************L*****************************---T*QGNE*******************-************  221
ATMPK8   ANDD*TPE*Y*F**Y********V*A**VF******K*I*A**D*K**********VSFNDAPTAI*W*D**A********CGSFFSK*TPA**I****  297
ATMPK9   ANDD*TPE*Y*F**Y*******F**T**VF******K*I*A*SD*K**********VSFNDAPSAI*W*D**A********CGSFFSK*TPA**I**I*  216
XLMPK1   T-*H*****IC***Y*I***********V*******S***L*TT************VADPDHDHTG*L****A********IM*NS-KG*TK***I****  218
                                 X                                                                  XI
ATMPK1   CIFAELLGRKPIFQGTECLNQLKLIVNIIGSQREEDLEFIVNPKAKRYIRSLPYSPGMSLSRLYPCAHVLAIDLLQKMLVFDPSKRISASEALQHPYMAP  321
ATMPK2   ***********V*P*******I***I**L**********D********E*******I*F*****G*N********I**L*******VT**********  321
ATMPK3   ***M**MN***L*P*KDHVH*MR*LTELL*TPT*S**G*TH*ED******Q**NF*RQP*AK*FSHVNPM****VDR**T***NR**TVEQ**N*Q*L*K  326
ATMPK4   **LG*TMT*E*L*P*KDYVH**R**TEL***PDDSS*G*LRSDN*R**V*Q**QY*RQNFAARF*NMSAG*V***E*****E**R**TVD***C***L**  331
ATMPK5   *****IMT*E*L*P*KDYVH*****TEL***PDGAS***LRSANGGK*VKE**KF*RQNF*ARF*SMNST*****E*******V***TVE***CY**LSA  331
ATMPK6   ***M**MD***L*P*RDHVH**R*LMEL**TPS**E***-L*EN******Q**PY*RQ*ITDKF*TV*P*****IE***T***RR**TVLD**A***LNS  350
ATMPK7   *****I*******P***********I*VV***Q*S*IR**D****R*F*K*****R*TH**N***Q*NP*******R****E*T****VTD**L*****G  321
ATMPK8   *****M*LG**L*P*KNVVH**D*MTDFL*TPPP*SISR*R*E**R**LS*MRKKQPVPF*HKF*K*EP**LR**ERL*A***KD*A**ED**AD**FSG  397
ATMPK9   *****M*TG**L*P*KNVVH**DIMTDLL*TPPP*AIAR*R*E**R**LGNMRRK*PVPFTHKF*HVDP**LR**HRL*A***KD*P*E***AD**FYG  316
XLMPK1   **L**M*SNR***P*KHY*D**NH*LG*L**PSQ***NC*I*L**RN*LL***HKNKVPWN**F*N*DPK*L***D***T*N*H***EVEA**A***LEQ  317

ATMPK1   LYDPNANPPAQVPID---LDVDEDLRE--EMIREMIWNEMLHYHPQASTLNTEL                                              370
ATMPK2   ****S**********---******EDLGA*****LM*K**I****E*A*I*NNEVSEF                                          376
ATMPK3   *H***DE*IC*K*FS---FEFEQQPLDE-*Q*K***YQ*AIALN*TYG                                                    370
ATMPK4   *H*I*EE*VCVR*FN---F*FEQPTLTE-*N*K*L*YR*TVKFN**D*V                                                   376
ATMPK5   *H*L*DE*VCSNHFS---FHFEDPSSTE-*E*K*LV*L*SVKFN*LP*I                                                   376
ATMPK6   *H*ISDE*ECTI*FN---F*FENHALSE-*QMK*L*YR*A*AFN*EYQQ                                                   395
ATMPK7   *FE*GT****H***S---**I**NME**--PV****M*****Y***E*EI**A                                               368
ATMPK8   *SNSERE*TT*-**SKLEF*FERKKLVK-DDV**LLYR*I*E****MLEEYLRGGDQLSFMYPSGVDRFKRQFAHLEENQGKPGAAGGGRSTALHRHHAS  495
ATMPK9   *ANVDRE*ST*-**PKLEFEFERRKLTK-*DV**L*YR*I*E****MLQEYLRGGEQTSFMYPSGVDRFKRQFAHLEENYGK-----GEKGSPLQRQHAS  409
XLMPK1   Y***SDE*V*EA*FK---FEMEL*DLPK-*TLK*L*FE*TARFQ*GY                                                     361

ATMPK8   LPRERVPAPNGETAEESSDVERRAAAAVASTLESEEQTMEEVTVLVTS                                                    543
ATMPK9   LPRERVPAPKKENGSHNHDIENRSIASLVTTLESPPTSQHEGSDYRNGTSQTGYSARSLLKSASISASKCIGMKPRNKSEYGESNNDTVDALSQKVAALH  509

ATMPK9   T                                                                                                   510
```

Figure 10. Comparison of the deduced amino acid sequences of *Arabidopsis* ATMPKs (ATMPK1 to 9) and *Xenopus* MPK1 (Gotoh *et al.*, 1991). Asterisks represent identical amino acid residues and dashes indicate gaps introduced to maximize alignment. Roman numerals indicate the 11 major conserved subdomains of protein kinases identified by Hanks *et al.* (Hanks *et al.*, 1988; Hanks and Quinn, 1991). Amino acid residues that are invariant or nearly invariant among known protein kinases are indicated by closed triangles. Closed circles show residues that correspond to the sites of regulatory phosphorylation (Sturgill and Wu, 1991; Cobb *et al.*, 1991; Thomas, 1992; Nishida and Gotoh, 1992; Pelech and Sanghera, 1992; Nishida and Gotoh, 1993; Neiman, 1993).

```
          VII                                  VIII
                                          ▽ ▽
ATMPK1    DFGLAR----------ASNTKGQFMTEYVVTRWYRAPE
ATMPK2    ******----------T*********************
ATMPK3    ******-----------PTSEND**************
ATMPK4    ******-----------TKSETD**************
ATMPK5    ******-----------TTSETEY*************
ATMPK6    ******-----------VTSESD**************
ATMPK7    ******----------T*QGNE***************
ATMPK8    ******--------VSFNDAPTAI*W*D**A********
ATMPK9    ******--------VSFNDAPSAI*W*D**A********
MSERK1    ******-----------VTSETD**************
PSMAP     ******-----------VTSETD**************
NTF3      ******----------T*SG*D***************
FUS3      ******IIDESAADNSEPTGQQSG*****A********
KSS1      ******---CLASSSDSRE*LVG******A********
MPK1      ******-------GYSENPVENS**L****A********
SPK1      ******-------STT*QGGNPG******A********
HOG1      ******-------------IQDPQ**G**S**Y*****
CAERK1    ******-------SIASQEDNYG******A********
CEMPK1    ******-------VTDPQTDHTG*L****A********
DMERKA    ******-------IADPEHDHTG*L****A********
XLMPK1    ******-------VADPDHDHTG*L****A********
XLXP42    ******-------VADPDHDHTG*L****A********
RATERK1   ******-------IADPEHDHTG*L****A********
RATERK2   ******-------VADPDHDHTG*L****A********
MMERK1    ******-------IADPEHDHTG*L****A********
MMERK2    ******-------VADPDHDHTG*L****A********
BTERK2    ******-------VADPDHDHTG*L****A********
HSERK1    ******-------IADPEHDHTG*L****A********
HSERK2    ******-------VADPDHDHTG*L****A********
```

Figure 11. Comparison of the deduced amino acid sequences around kinase subdomains VI to VIII between *Arabidopsis* ATMPKs and MAP kinases of other organisms. Asterisks represent identical amino acid residues and dashes indicate gaps introduced to maximize alignment. Roman numerals indicate the major conserved subdomains of protein kinases identified by Hanks *et al.* (Hanks *et al.*, 1988; Hanks and Quinn, 1991). Open triangles show residues that correspond to the sites of regulatory phosphorylation reported for *Xenopus* MPK1 and rat ERK1. The sequences boxed by blue represent the divergent region between subdomains VII and VIII (Toda *et al.*, 1991; Mizoguchi *et al.*, 1993b).

Figure 12. A phylogenetic tree showing evolutional relationship among the MAP kinases was constructed from the matrix of sequence similarities calculated with the UPGMA program (Nei, 1987). The evolutional distance between one protein and another is shown at the bottom. The sequences of the following proteins were included: *Arabidopsis* ATMPK1 to nine, alfalfa MsERK1, pea PsMAP, tobacco NTF3, yeast HOG1, murine p38, yeast CAERK1, yeast SPK1, yeast KSS1, yeast FUS3, *Xenopus* XLp42, *Xenopus* XLMPK1, mouse MmERK2, rat ERK2, human HsERK2, mouse MmERK1, rat ERK1, human HsERK1, *Drosophila* DmERKA, *C. elegans* CEMPK1, yeast MPK1, rat SAPKγ, human JNK1, rat SAPKβ, and rat SAPKα. All sequences used for these analyses are from the SwissProt and GenBank data bases.

```
GTCTTCTTCACTCCCGCGAGTTTTTTTTCTCGGGAAAATTTTCCAATTCCAAGCTCCATTGAAGCCAAGTAAGAGGTGCGATTCTCTTCAATTTTGACTT    100
GGAACTTGTAAATAGCGTATGGATTTGAGCTGGTTATGGACAGAATTCTAGCTCGTATGAAGAAATCAACTGGACGAAGAGGAGGAGATAAGAATATTAC    200
                                        M  D  R  I  L  A  R  M  K  K  S  T  G  R  R  G  G  D  K  N  I  T

TCCGGTACGGCGGTTAGAGCGTCGCGATGCGGCGAGGAATATCAATTACGCCGCAGCTTCATGTTCTAGTTCGTCAGCTGAAGATCTTTCCGTTTCGACT    300
 P  V  R  R  L  E  R  R  D  A  A  R  N  I  N  Y  A  A  A  S  C  S  S  S  S  A  E  D  L  S  V  S  T

TCTTCGTTGATGACTCGCTCTTTGGAGTTTCCGGAGCCTACTAGTTTCCGAATCGGTGGTGGTGTTGGAGAGATGGATCGGATTTATCGGTCTCTTGGTG    400
S  S  L  M  T  R  S  L  E  F  P  E  P  T  S  F  R  I  G  G  G  V  G  E  M  D  R  I  Y  R  S  L  G  V

TTTCTGGTCCTGATGATTTGGCTATTTCTTTTGATGCTTGGGAAGCTTGTAAGAAACGTTCTTCTTCAGATGTTGTTAATAGGTTTAAGTCTTTTGATCT    500
  S  G  P  D  D  L  A  I  S  F  D  A  W  E  A  C  K  K  R  S  S  S  D  V  V  N  R  F  K  S  F  D  L

TGATAAGGTTCGTGATCAGGATTTGAGTGAAGAAGGTCCTAGTGGTGTTGTTGTTGGTTCTGATTCAATGAATCATAAGGTTCAGGGTCAGGATTTGAGT    600
 D  K  V  R  D  Q  D  L  S  E  E  G  P  S  G  V  V  V  G  S  D  S  M  N  H  K  V  Q  G  Q  D  L  S

GAAGCAGGTCCTAGTGGTGGAATTGTTACTGAGTTGAGTGAGATTGGGAATTTAATCACTCCTGTGGATAGGTTAGTAGCTGATGGTGTGGTAGAGAATA    700
E  A  G  P  S  G  G  I  V  T  E  L  S  E  I  G  N  L  I  T  P  V  D  R  L  V  A  D  G  V  V  E  N  R

GACGTGTTATGGAAAGAACACCAACTATTGTGAAGTCGAAAGGGTATCTTGTACCAAATAATGTAGTGGCTGTTGGTGTTGGTGTTGGTGGTGGTATTAA    800
  R  V  M  E  R  T  P  T  I  V  K  S  K  G  Y  L  V  P  N  N  V  V  A  V  G  V  G  V  G  G  G  I  K

GGGGCTAAGACCACCAGTACTTAAGCCTCCTCCGGCTATGAAACGACCCCCTATTGATCATCGGGGATCGTCTTGGGATTTCCTGACGCATTTCGCTCCA    900
 G  L  R  P  P  V  L  K  P  P  P  A  M  K  R  P  P  I  D  H  R  G  S  S  W  D  F  L  T  H  F  A  P

AGTGAAACAGTTAAGCGGCCGAGTTCCTCTTCTTCTTCTTCCGAGGATGGATGCGATGAAGAGGAAGGCAAGGAAGAGGAAGCGGAAGCGGAAGAGATGG   1000
S  E  T  V  K  R  P  S  S  S  S  S  S  S  E  D  G  C  D  E  E  E  G  K  E  E  E  A  E  A  E  E  M  G

GAGCTAGGTTTATCCAGTTGGGGGATACGGCTGACGAGACGTGCTCATTCACTACAAATGAGGGTGACTCCTCAAGCACAGTATCCAATACTTCGCCTAT   1100
  A  R  F  I  Q  L  G  D  T  A  D  E  T  C  S  F  T  T  N  E  G  D  S  S  S  T  V  S  N  T  S  P  I

CTATCCAGATGGAGGAGCTATCATAACGTCTTGGCAAAAGGGTCAACTTTTGGGACGAGGATCATTTGGTTCTGTGTATGAAGGCATTTCTGGAGATGGG   1200
 Y  P  D  G  G  A  I  I  T  S  W  Q  K  G  Q  L  L  G  R  G  S  F  G  S  V  Y  E  G  I  S  G  D  G

GACTTCTTTGCTGTCAAGGAAGTTTCACTTCTTGATCAGGGAAGTCAGGCACAAGAATGCATACAACAACTTGAGGGGGAGATTAAACTACTTAGTCAGC   1300
D  F  F  A  V  K  E  V  S  L  L  D  Q  G  S  Q  A  Q  E  C  I  Q  Q  L  E  G  E  I  K  L  L  S  Q  L

TTCAGCATCAGAATATTGTGAGATATCGTGGCACAGCCAAGGATGGGTCAAACTTGTACATTTTTCTTGAGCTTGTAACCCAAGGGTCCCTTCTAAAACT   1400
  Q  H  Q  N  I  V  R  Y  R  G  T  A  K  D  G  S  N  L  Y  I  F  L  E  L  V  T  Q  G  S  L  L  K  L

CTACCAAAGATACCAGCTTCGGGACTCTGTAGTCTCCTTGTACACTAGACAGATTCTTGACGGTCTGAAATATCTCCACGATAAAGGTTTTATTCACAGG   1500
 Y  Q  R  Y  Q  L  R  D  S  V  V  S  L  Y  T  R  Q  I  L  D  G  L  K  Y  L  H  D  K  G  F  I  H  R

GACATCAAATGTGCAAATATATTGGTGGACGCTAATGGCGCCGTCAAACTTGCAGATTTTGGATTGGCAAAGGTTTCAAAGTTTAACGACATTAAGTCCT   1600
D  I  K  C  A  N  I  L  V  D  A  N  G  A  V  K  L  A  D  F  G  L  A  K  V  S  K  F  N  D  I  K  S  C

GCAAGGGAACTCCATTTTGGATGGCTCCAGAGGTTATTAACCGAAAGGATAGTGATGGCTATGGAAGTCCAGCTGATATATGGAGCCTCGGGTGCACTGT   1700
  K  G  T  P  F  W  M  A  P  E  V  I  N  R  K  D  S  D  G  Y  G  S  P  A  D  I  W  S  L  G  C  T  V

GCTGGAAATGTGTACTGGTCAGATCCCCTACTCTGATCTCGAACCCGTTCAAGCCCTGTTTAGGATCGGAAGGGGTACTCTTCCGGAAGTACCTGATACG   1800
 L  E  M  C  T  G  Q  I  P  Y  S  D  L  E  P  V  Q  A  L  F  R  I  G  R  G  T  L  P  E  V  P  D  T

TTATCACTAGATGCTCGGCTTTTCATACTTAAGTGTCTTAAAGTGAACCCGGAAGAGCGGCCAACTGCAGCTGAACTGCTGAACCATCCATTTGTGAGAA   1900
L  S  L  D  A  R  L  F  I  L  K  C  L  K  V  N  P  E  E  R  P  T  A  A  E  L  L  N  H  P  F  V  R  R

GACCCTTACCATCCGTGGGTTCAGGAGGCTCGGGATCAGCATCTCCCCTTCTCCGTAGATGAGGCTAATTTTAGTAATTTCCTCAATAGGCATAACCCGA   2000
  P  L  P  S  V  G  S  G  G  S  G  S  A  S  P  L  L  R  R  *

AGATTTATGAGAGATTAGATCCCAGAGAAACTATTGGGAGACACGAGACTTGATTCAACGACACACACACACGTGCGGTCGGGTTGCGCAAGCAAAGAGC   2100
CAAACTTACTTAGTTTCCTGTGTCATTCGTTGGTTATTTGTTGCAAGTTATGTGTTGGGCTTGATTTTAAAACTGAAGAGCCTTATTTTTCTAGGAATTA   2200
GAAATGTCGACCTTGGGGTTTTTGGGTGTATAGGTTTTGTTTAGTTTTTTTTGTTAAATTATAGTGTATGTATGTCTGTATCTATATCATGTAAATGTAT   2300
AATTATAGTTGGTTCTTATAGTTCCTGGCTTGTTGTAACTGTGTACACACATCGATGGGTTTAAGTTCTATTTGTCTCATATTTTA                  2386
```

Figure 13. Nucleotide and deduced amino acid sequences of cATMEKK1 cDNA. The DNA sequence includes the putative coding region and the 5'- and 3'-noncoding regions. The amino acid sequence of the putative coding region is shown beneath the DNA sequence. The sequence corresponding to the probe that was used to screen the library is underlined.

Figure 14. (A) Comparison of the deduced amino acid sequences of putative catalytic domains of ATMEKK1, byr2 (Wang *et al.*, 1991), BCK1 (Lee and Levin, 1992), STE11 (Rhodes *et al.*, 1990), MEKK (Lange-Carter *et al.*, 1993), and NPK1 (Banno *et al.*, 1993). Dots represent identical amino acid residues and dashes indicate gaps introduced to maximize alignment. Asterisks indicate the conserved amino acid residues among these six protein kinases. Roman numerals indicate the eleven major conserved subdomains of protein kinases identified previously (Hanks *et al.*, 1988; Hanks and Quinn, 1991). **(B)** A phylogenetic tree showing evolutional relationship among ATMEKK1, byr2, BCK1, STE11, MEKK, and NPK1 proteins was constructed from the matrix of sequence similarities calculated with the UPGMA program (Nei, 1987). Numbers above the horizontal lines indicate the evolutional distance between one protein and another. **(C)** Organization of domains of predicted ATMEKK1, byr2, BCK1, STE11, MEKK, and NPK1 proteins. Red bars represent the putative catalytic domains. Bars with purple diagonals show the non-catalytic domains which are thought to be regulatory domains.

Figure 15. Genomic DNA gel blot analysis. Genomic DNA was digested with *Eco*RI (E) and *Hin*dIII (H), fractionated on 1% agarose gels and transferred to nylon membranes. Filters were hybridized with [^{32}P]-labeled fragment of *ATMEKK1* cDNA at 42°C and washed in either 0.5x SSC/0.5% SDS at 37°C (low-stringency) or 0.1x SSC/0.1% SDS at 65°C (high-stringency). High and Low represent high- and low-stringency hybridization conditions, respectively. The sizes of DNA markers are indicated in kbp.

(A) **ATPK6**

```
CTCTGAGCTTATAAAACCTTCATGGCATTTGCGCACTTCTGCAGCATCCACTGTTTCCTT    60
AAAATCGGGAAACAACAAGAGCTTCTTCTCCTTGTTATTGGATTGCCTTTTGCCAGCCTT   120
TGTTTTAGTCAGTTTTCTCCTCATTAACTTCTGTCGGTACAGAATCTCTCTCCAGGCTGG   180
TGAACTCTTGCAGCTTTCTCAGTGATTCTGTATCGATTTTTCCCGTGTAAATTATGGTTT   240
                                                   M  V  S     3

CCTCTCAGCGTCCTGTCCCCAACAAAATCCAGAAACAGCAATATCTATCCATCAGCCCTT   300
 S  Q  R  P  V  P  N  K  I  Q  K  Q  Q  Y  L  S  I [S  P] S    23

CAAATTCTGTCTTGAAAGATGATGTTGAACTTGAGTTCTCTGATGTGTTTGGTCCACTCC   360
 N  S  V  L  K  D  D  V  E  L  E  F  S  D  V  F  G  P  L  P    43

CCGAAGAGGCCAATGATATTGCTTACGACGAGCCTGCTGTTGTCTACAGTCGATCCCACT   420
 E  E  A  N  D  I  A  Y  D  E  P  A  V  V  Y  S  R  S  H  S    63

CCTTGGTTGGCCCGTGTTCGCTCGATAGTCATTCTTTGAAGCTGACCAAGCTCACCTTAC   480
 L  V  G  P  C  S  L  D  S  H  S  L  K  L  T  K  L  T  L  L    83

TAGAGACTGAGGACTCAATTGACTTGGTGGAGTGTCTCGAAGGTGAATCTTTAAAAGAAA   540
 E  T  E  D  S  I  D  L  V  E  C  L  E  G  E  S  L  K  E  N   103

ATGATGATTTCTCTGGTAATGACGACAGTGACAATGAGAAAGCTTTAGAGGGAGATCTGG   600
 D  D  F  S  G  N  D  D  S  D  N  E  K  A  L  E  G  D  L  V   123

TGAAAGTCTCAGGTGTAGTAGGCATAGATGATTTTGAGGTTATGAAGGTTGTTGGGAAAG   660
 K  V  S  G  V  V  G  I  D  D  F  E  V  M  K  V  V  G  K  G   143

GTGCGTTTGGGAAAGTCTACCAGGTGAGGAAAAAGGAGACCTCTGAGATATACGCTATGA   720
 A  F  G  K  V  Y  Q  V  R  K  K  E  T  S  E  I  Y  A  M  K   163

AAGTCATGAGGAAAGATCACATTATGGAGAAGAACCATGCTGAATACATGAAAGCTGAGC   780
 V  M  R  K  D  H  I  M  E  K  N  H  A  E  Y  M  K  A  E  R   183

GCGATATTCTTACAAAAATTGATCACCCATTCATTGTTCAACTTAAATACTCTTTTCAGA   840
 D  I  L  T  K  I  D  H  P  F  I  V  Q  L  K  Y  S  F  Q  T   203

CCAAGTACAGGCTGTATCTTGTGCTCGACTTTATAAACGGAGGTCATCTTTTCTTCCAAC   900
 K  Y  R  L  Y  L  V  L  D  F  I  N  G  G  H  L  F  F  Q  L   223

TCTATCACCAAGGGCTTTTCAGAGAGGACTTGGCTCGAGTGTACACTGCAGAAATCGTCT   960
 Y  H  Q  G  L  F  R  E  D  L  A  R  V  Y  T  A  E  I  V  S   243

CTGCAGTTTCCCATCTCCATGAGAAAGGCATAATGCACAGAGATCTGAAACCCGAAAACA  1020
 A  V  S  H  L  H  E  K  G  I  M  H  R  D  L  K  P  E  N  I   263

TACTCATGGATACAGATGGCCACGTGATGCTAACAGATTTTGGTTTAGCAAAGGAATTTG  1080
 L  M  D  T  D  G  H  V  M  L  T  D  F  G  L  A  K  E  F  E   283

AAGAAAACACAAGATCAAACTCCATGTGCGGAACTACGGAGTATATGGCACCTGAAATTG  1140
 E  N  T  R  S  N  S  M  C  G  T  T  E  Y  M  A  P  E  I  V   303

TTCGAGGAAAAGGACATGATAAAGCAGCTGACTGGTGGAGCGTAGGGATTCTTCTGTATG  1200
 R  G  K  G  H  D  K  A  A  D  W  W  S  V  G  I  L  L  Y  E   323

AGATGCTCACAGGAAAGCCACCGTTTCTCGGGAGCAAAGGAAAGATACAGCAGAAAATTG  1260
 M  L  T  G  K  P  P  F  L  G  S  K  G  K  I  Q  Q  K  I  V   343

TGAAGGACAAGATCAAGCTTCCGCAGTTTCTGTCTAATGAAGCTCATGCAATACTGAAAG  1320
 K  D  K  I  K  L  P  Q  F  L  S  N  E  A  H  A  I  L  K  G   363

GGCTGCTGCAAAAAGAGCCAGAAAGGCGACTTGGAAGTGGACTGAGCGGAGCAGAGGAGA  1380
 L  L  Q  K  E  P  E  R  R  L  G  S  G  L  S  G  A  E  E  I   383

TAAAACAGCACAAATGGTTCAAGGGAATAAACTGGAAGAAGCTGGAGGCTAGAGAAGTGA  1440
 K  Q  H  K  W  F  K  G  I  N  W  K  K  L  E  A  R  E  V  M   403

TGCCAAGTTTCAAGCCGGAAGTATCGGGAAGGCAATGCATAGCAAATTTTGACAAGTGTT  1500
 P  S  F  K  P  E  V  S  G  R  Q  C  I  A  N  F  D  K  C  W   423

GGACTGACATGTCTGTTTTGGATTCTCCGGCGAGCAGTCCCAGTTCGGATCCTAAGGCCA  1560
 T  D  M  S  V  L  D [S  P] A  S [S  P] S  S  D  P  K  A  N   443

ACCCTTTTACCAATTTCACTTACGTCAGGCCTCCTCCTTCATTCCTTCACCAGTCCACAA  1620
 P  F  T  N  F  T  Y  V  R  P  P  P  S  F  L  H  Q  S  T  T   463

CTACTTTGTAGATTTACTATATTTTTATGTAGTTCCTACTTCTTTGGTTGTTGTTTCCTT  1680
 T  L  *                                                      465

CTGACACATGAACATTAAAGAAAACTTTGGTTACCAAGACGCTTAATTTATAATTGATTT  1740
TATGAGATTGTTATTGATTGAGATTAAAAAAAAAAAAAAAAAAAAAA               1786
```

Figure 16 Nucleotide and deduced amino acid sequences of cATPK6 (A) and cATPK19 (B) cDNAs. The DNA sequences include the putative coding regions and the 5'- and 3'-noncoding regions. The amino acid sequences of the putative coding regions are shown beneath the DNA sequences. The sequences corresponding to the probe (the insert DNA of clone C11) that was used to screen the library is underlined. Proline (P), glutamic acid (E), serine (S), threonine (T), and aspartic acid (D) residues which are involved in the PEST region (Rogers *et al.*, 1986; Rechsteiner *et al.*, 1987) are underlined. Consensus phosphorylation sites (S/TP) for proline-directed protein kinases such as cdc2 and MAP kinase are boxed (Pearson and Kemp, 1991).

```
GTTTTACTCTATATCTGGATTTGTTTGAAAATTGCTTGTGTTATGCGCCAGCTTTTTCCC    60
TAATATATAGACTAGGTCTTAATTACAGATATTAGAGTTCGTTTCCTCGAGGGTTTCTTT   120
CGACCTCGTGTCCTGGTGAAATTTTGGAATTCTAAGGCTTTGGGGTTTTATTTGGGAGGG   180
GATGATCATCTAAAGGATATTCGGGCGGTGTCGTGTTTTCTTAATCTCAATTCTGTTGAC   240
TTTAGCTTTGCGCTTAAAGTACTTTGATGGCATTGGGACACTTGTGCAGCATCCACTGTT   300
TCCTTAAAATAGGGAAACAAGAAGAGCTTCTTGTCCTAGATATTGGTTTGCCTATTCCCA   360
GTTTCTTCCTCATTAACTTTATTGGTAGCAGAGCCTCATCCATCTGGTGAACTTGTGAAG   420
CTTCCTCTGTGATTCTTTTGGCGATTGGTTAGAAATATGGTTTCTTCTCAGTGTTCTGTT   480
                                        M  V  S  S  Q  C  S  V     8

GCAAACAAAAACCAGACTGGGAAGCCGTTTCAGAAACATCTATCTTTATCCATCAGCCCT   540
A  N  K  N  Q  T  G  K  P  F  Q  K  H  L  S  L  S  I  S  P      28

CCAAAATCTGTCTTGGGTGATAACCTTGAGCTTCAATTCTCTGATGTCTTTGGACCAATG   600
P  K  S  V  L  G  D  N  L  E  L  Q  F  S  D  V  F  G  P  M      48

CCTGAAGCTAACTCTGAAGAGGCCTGTGATGTTGCTTACGACGAGCCTGCTGTTGTCTAC   660
P  E  A  N  S  E  E  A  C  D  V  A  Y  D  E  P  A  V  V  Y      68

AGTCGATCCCACTCATTGGTTGGCCCGTCTTTGGTTGTTAGCCATTCTTTGAAGATGAAC   720
S  R  S  H  S  L  V  G  P  S  L  V  V  S  H  S  L  K  M  N      88

AAGCTCACTTTACGAGAAACTGAGGACTCAGTTGACTTGGTGGAGTGTGTTGAAGGTGAA   780
K  L  T  L  R  E  T  E  D  S  V  D  L  V  E  C  V  E  G  E     108

TCGATAAAAGAGAACGATGAATTCTCTGGTAATGATGACACTGACAGTGAGAAATCTCCA   840
S  I  K  E  N  D  E  F  S  G  N  D  D  T  D  S  E  K  S  P     128

GAGGAGGTCTCAGGTGTAGTAGGCATTGAGGACTTTGAGGTATTGAAGGTTGTGGGACAA   900
E  E  V  S  G  V  V  G  I  E  D  F  E  V  L  K  V  V  G  Q     148

GGTGCATTTGGAAAAGTGTACCAGGTGAGGAAAAAAGACACGTCTGAGATATACGCGATG   960
G  A  F  G  K  V  Y  Q  V  R  K  K  D  T  S  E  I  Y  A  M     168

AAGGTCATGAGAAAAGATAAAATTGTTGAGAAGAATCATGCTGAATACATGAAAGCCGAG  1020
K  V  M  R  K  D  K  I  V  E  K  N  H  A  E  Y  M  K  A  E     188

CGCGATATTCTAACCAAAATCGATCATCCTTTCATTGTGCAACTTAAATACTCTTTTCAG  1080
R  D  I  L  T  K  I  D  H  P  F  I  V  Q  L  K  Y  S  F  Q     208

ACCAAATACAGATTGTATCTTGTTCTTGACTTTATAAACGGAGGTCATCTTTTCTTCCAG  1140
T  K  Y  R  L  Y  L  V  L  D  F  I  N  G  G  H  L  F  F  Q     228

CTCTATCACCAAGGGCTTTTCAGGGAGGACTTGGCTCGTGTGTACACTGCAGAAATCGTC  1200
L  Y  H  Q  G  L  F  R  E  D  L  A  R  V  Y  T  A  E  I  V     248

TCTGTAGTTTCCCATCTCCATGAGAAAGGCATAATGCATAGAGATCTTAAACCTGAAAAC  1260
S  V  V  S  H  L  H  E  K  G  I  M  H  R  D  L  K  P  E  N     268

ATACTCATGGACGTAGATGGCCATGTGATGTTAACTGATTTTGGTTTAGCAAAAGAATTT  1320
I  L  M  D  V  D  G  H  V  M  L  T  D  F  G  L  A  K  E  F     288

GAAGAAAACACAAGATCAAACTCCATGTGTGGGACTACGGAGTATATGGCCCCTGAGATT  1380
E  E  N  T  R  S  N  S  M  C  G  T  T  E  Y  M  A  P  E  I     308

GTTCGTGGAAAAGGACATGATAAGGCTGCCGACTGGTGGAGTGTAGGAATCCTTCTGTAT  1440
V  R  G  K  G  H  D  K  A  A  D  W  W  S  V  G  I  L  L  Y     328

GAGATGCTCACTGGAAAGCCGCCGTTTCTCGGGAGCAAAGGAAAGATACAGCAGAAAATC  1500
E  M  L  T  G  K  P  P  F  L  G  S  K  G  K  I  Q  Q  K  I     348

GTTAAGGACAAGATCAAGCTTCCACAGTTTGTGTTTAATGAAGCTCATGCCTTGCTGAAA  1560
V  K  D  K  I  K  L  P  Q  F  V  F  N  E  A  H  A  L  L  K     368

GGGCTGCTGCAAAAAGAGCCAGAAAGGAGACTGGGAAGTGGACCGAGCGGAGCAGAGGAG  1620
G  L  L  Q  K  E  P  E  R  R  L  G  S  G  P  S  G  A  E  E     388

ATAAAAAAACACAAATGGTTCAAGGCGATAAACTGGAAGAAGCTGGAGGCTAGAGAAGTA  1680
I  K  K  H  K  W  F  K  A  I  N  W  K  K  L  E  A  R  E  V     408

CAACCAAGTTTCAAGCCGGCGGTTTCGGGAAGACAATGCATAGCTAATTTTGACAAGTGT  1740
Q  P  S  F  K  P  A  V  S  G  R  Q  C  I  A  N  F  D  K  C     428

TGGACTGACATGTCTGTTTTGGATTCTCCAGCAAGTAGTCCCAACTCTGATGCTAAGGCG  1800
W  T  D  M  S  V  L  D  S  P  A  S  S  P  N  S  D  A  K  A     448

AACCCTTTTACAAACTTCACATACGTCAGGCCTCCTCATTCATTCCTTCATCGGACCACA  1860
N  P  F  T  N  F  T  Y  V  R  P  P  H  S  F  L  H  R  T  T     468

TCCAACTTGTAGATGAACTATAATATGTTATCCCCCGCTTCTGTTTTTGTAAACCGAAAC  1920
S  N  L  *                                                     471

ATATGAATATGCACAAAATCAACGGTTGTTTGAAGAAAAAAAAAAAAAAAAAAAA       1975
```

Figure 17. (A) Comparison of the deduced amino acid sequences of putative catalytic domains of ATPK6, ATPK19, human p70 ribosomal S6 kinase (Hsp70S6K) (Grove *et al.*, 1991), *Xenopus* p90 S6 kinase II a (Xlp90S6Ka) (Jones *et al.*, 1988), human PKCa (HsPKCa) (Finkenzeller *et al.*, 1990), and human PKAa (HsPKAa) (Maldonado and Hanks, 1988). Dots represent identical amino acid residues with ATPK6 and dashes indicate gaps introduced to maximize alignment. Closed triangles indicate the conserved amino acid residues among protein kinases. Roman numerals indicate the eleven major conserved subdomains of protein kinases identified previously (Hanks *et al.*, 1988, Hanks and Quinn, 1991). **(B)** A phylogenetic tree showing evolutional relationship among ATPK6, ATPK19, ribosomal S6 kinase (RSK) family, protein kinase C (PKC) family, and protein kinase A (PKA) family was constructed from the matrix of sequence similarities calculated with the UPGMA program. The evolutional distance between one protein and another is shown at the bottom. Species abbreviations: Hs, *Homo sapiens*; Rn, *Rattus norvegicus*; Gg, *Gallus gallus*; Mm, *Mus musculus*; Xl, *Xenopus laevis*; Dm, *Drosophila melanogaster*; Sc, *Saccharomyces cerevisiae*. All sequences used for these analyses are from the SwissProt and GenBank data bases.

Figure 18. Genomic DNA gel blot analysis. Genomic DNA was digested with *Bam*HI (B), *Eco*RI (E) and *Hin*dIII (H), fractionated on 1% agarose gels and transferred to nylon membranes. Filters were hybridized with [^{32}P]-labeled fragment of *ATPK6* or *ATPK19* cDNAs at 42°C and washed in either 0.5x SSC/0.5% SDS at 37°C (low-stringency) or 0.1x SSC/0.1% SDS at 65°C (high-stringency). High and Low represent high- and low-stringency hybridization conditions, respectively. The sizes of DNA markers are indicated in kbp.

Figure 19. Genomic DNA gel blot analysis. Genomic DNA was digested with *Eco*RI (E) and *Hin*dIII (H), fractionated on 1% agarose gels and transferred to nylon membranes. Filters were hybridized with [^{32}P]-labeled fragment of *ATMPK* cDNAs at 42°C and washed in 0.1x SSC/0.1% SDS at 65°C (high-stringency conditions, see Materials and Methods). The sizes of DNA markers are indicated in kbp.

Figure 20. Regulation of *ATMPK* gene expression by a variety of stimuli. The accumulation of transcripts for ATMPKs in response to a number of stress conditions was analyzed by Northern blotting analysis and quantification of the autoradiograms was plotted. Each lane was loaded with 40 μg of total RNA prepared from unbolted *Arabidopsis* plants that had been transferred from agar plates to 22 °C water and grown at 22 °C (H_2O), transferred to 4 °C water and grown at 4 °C (Cold), transferred from agar plates for hydroponic growth in 250 mM NaCl (NaCl), transferred to and grown at 40 °C (Heat), transferred from agar plates for hydroponic growth in 100 μM ABA (ABA) or 100 μM 2,4-D (2,4-D), as described in Materials and Methods. RNA was subjected to electrophoresis on 1% agarose gels that contained formaldehyde. A nylon membrane blotted with RNA were hybridized with [^{32}P]-labeled *ATMPK* cDNAs. The number of hours after the initiation of treatment prior to isolation of RNA was indicated. The ratio of hybridization signals for the *ATMPK* mRNAs were measured by Bio Image Analyzer (FUJIX BAS2000, Fuji Film Co., Ltd., Tokyo, Japan).

Figure 21. Effects of a variety of stresses on the expression of the gene that corresponds to ATMEKK1 cDNA. Northern blotting analysis was performed as described in Figure 20. A nylon membrane blotted with RNA were hybridized with [^{32}P]-labeled *ATMEKK1* cDNA. A nylon membrane blotted with RNA was stained with methylene blue to show that similar amounts of RNA were loaded per lane (rRNA). The two prominent bands are rRNAs.

Figure 22. Effects of a variety of stresses on the expression of the genes that correspond to ATPK6 and ATPK19 cDNAs. Northern blotting analysis was performed as described in Figure 20. A nylon membrane blotted with RNA were hybridized with [^{32}P]-labeled *ATPK6* or *ATPK19* cDNAs. A nylon membrane blotted with RNA was stained with methylene blue to show that similar amounts of RNA were loaded per lane (rRNA). The two prominent bands are rRNAs.

Figure 23. Similar expression profiles of three protein kinases, ATMPK3, ATPK19, and ATMEKK1, in response to a variety of stresses. The mRNA levels were examined by Northern blotting analysis as described in Figure. 20. ATMPK3, ATPK19, ATMEKK1, ATMPK1, and cdc2a cDNAs were used as probes. The number above each lane indicates the number of hours after the initiation of treatment prior to isolation of RNA. A nylon membrane blotted with RNA was stained with methylene blue to show that similar amounts of RNA were loaded per lane. The two prominent bands are rRNAs.

(A)

Figure 24. The rapid and simultaneous accumulation of the mRNA levels of ATMPK3, ATPK19, and ATMEKK1, elicited by various stimuli. **(A)** Expression of ATMPK3, ATPK19, and ATMEKK1 in response to cold, NaCl, and desiccation treatments.

(B) Expression of ATMPK3, ATPK19, and ATMEKK1 in response to water and MS media.(C) Expression of ATMPK3, ATPK19, and ATMEKK1 in response to touch stimuli. The mRNA levels were examined by Northern blotting analysis as described in Figure 20. ATMPK3, ATPK19, ATMEKK1, Tchγ3, ATMPK1, and cdc2a cDNAs were used as probes. The number above each lane indicates the number of minutes after the initiation of treatment prior to isolation of RNA. A nylon membrane blotted with RNA was stained with methylene blue to show that similar amounts of RNA were loaded per lane (rRNA). The two prominent bands are rRNAs.

(A)

Figure 25. **(A)** Expression of ATMPK3 in response to cold, NaCl, and water treatments, after pretreatment with water for 22 h. Each lane contains 40 μg of total RNA prepared from rosette plants that had been transferred from GM agar plates and grown hydroponically in water at 22 °C for 22 h, and then grown hydroponacally in water (Water) at 22 °C, 250 mM NaCl solution (NaCl) at 22 °C, or 4 °C water (Cold) at 4 °C for the indicated periods of time (min). RNA in lane C was isolated from control plants that had been exposed only to standard growth conditions. ATMPK3 and cdc2a cDNAs were used as probes.

(B)

(B) Expression of ATMPK3, ATPK19, and Tchγ in response to various stresses. Each lane contains 40 μg of total RNA prepared from rossete plants that had been transferred from GM agar plates and grown hydroponically in water for 22 h (lane 2), or grown hydroponically in water for 22 h and then grown hydroponacally in water (lane 3) at 22 °C, 40 °C water (lane 4) at 40 °C, 4 °C water (lane 5) at 4 °C, and 250 mM NaCl solution (lane 6) at 22 °C for 20 min. RNA in lane 1 was isolated from control plants that had been exposed only to standard growth conditions. ATMPK1, ATMPK3, ATPK19, and Tchγ3 cDNAs were used as probes. The mRNA levels were examined by Northern blotting analysis as described in Figure. 20. Stress treatments were performed as described in Materials and Methods. A nylon membrane blotted with RNA was stained with methylene blue to show that similar amounts of RNA were loaded per lane (rRNA). The two prominent bands are rRNAs.

clone name	related to	references
cDNA clones		
1. ATPK64	PVPK-1 (Plant type PKase)	Mizoguchi *et al.*, *Plant Mol. Biol.* (1992)
2. ATCKA1, 2	Casein kinase II α subunit	Mizoguchi *et al.*, *Plant Mol. Biol.* (1993)
3. ATMPK1, 2	MAP kinase	Mizoguchi *et al.*, *Plant Journal* (1994)
ATMPK3 - 7		Mizoguchi *et al.*, *FEBS Lett.* (1993)
ATMPK8, 9		
4. ATMEKK1 - 4	MAPKKK	
5. ATPK6C, 19B	p70 RSK, p90 RSK	Mizoguchi *et al.*, *FEBS Lett.* in press
6. ATPK10C	SNF1, CaM kinase II	Mizoguchi *et al.*, *Plant Physiol.* in press
7. ATPK27	SRK	
8. ATRAC1 - 4	rac (p21 G protein)	
PCR fragment		
9. pSNF1	SNF1	
10. pMAPKK	MAPKK	
11. pMAK	MAK	
12. pNPK1	MAPKKK	
13. pCaM	Calmodulin	

Figure 26. Summary of the isolated cDNA clones which encode putative protein kinases, G proteins, calmodulins. For details see **Discussion**.

Figure 27. A hypothetical model for kinase cascades in various signal transduction pathways in animals, yeasts, and plants. In animals, MAP kinase (MAPK) is activated by a serine/threonine/tyrosine kinase, MAP kinase kinase (MAPKK), which is in turn activated by a serine/threonine kinase, MAP kinase kinase kinase (MAPKKK). This kinase cascade plays a key role in the growth factor- or differentiation factor-stimulated signaling pathways and in the MPF-induced signaling pathway. In yeasts, there are several pathways hypothetically analogous to the MAPKKK/MAPKK/MAPK cascade in animals. MAPKKK/MAPKK/MAPK homologues in plants are also shown. Other details are given in **Discussion** and **Conclusion**.

Figure 28. A hypothetical model for the differential activation mechanisms of the MAP kinase cascades between animals and plants. MAPKKK, MAPKK, MAPK, RSK, S, Sa, and Sb represent MAP kinase kinase kinase, MAP kinase kinase, MAP kinase, ribosomal S6 kinase, substrate, substrate a, and substrate b, respectively. Closed triangles indicate the directions signals transduce in. The number of these triangles indicates the efficiency to transduce the signals from the upstream factor to the downstream factor. For details see **Conclusion**.

Table 1. Design of degenerated oligonucleotide mixtures for protein kinase recognition.

---

cAPK

(Subdomain) I

(Consensus) LGTGSFGRV

(Primers)
PKOIA1 5'-T/CTIGGIACIGGITCITTT/CGGIC/AGIGT-3'
PKOIA2 5'-T/CTIGGIACIGGIAGT/CTTT/CGGICAGIGT-3'

(Subdomain) IX

(Consensus) DWWSLGVLI

(Primers)
PKIXA1 5'-ATIAG/AIACICCIAG/AIGACCACCAG/ATC-3'
PKIXA2 5'-ATIAG/AIACICCIAG/AG/ACTCCACCAG/ATC-3'

---

cGPK

(Subdomain) I

(Consensus) LGVGGFGRV

(Primer)
PKOIB 5'-T/CTIGGIGTIGGIGGITTT/CGGIC/AGIGT-3'

(Subdomain) IX

(Consensus) DYWSLGILM

(Primers)
PKIXB1 5'-CATIAG/AIATICCIAG/AIGACCAGTAG/ATC-3'
PKIXB2 5'-CATIAG/AIATICCIAG/AGACTCCAGTAGATC-3'

PKC

(Subdomain) I

(Consensus) LGKGSFGKV

(Primers)
PKOIC1 5'-T/CTIGGIAAG/AGGITCITTT/CGGIAAG/AGT-3'
PKOIC2 5'-T/CTIGGIAAG/AGGIAGT/CTTT/CGGIAAG/AGT-3'

(Subdomain) IX

(Consensus) DWWAY/FGVLL

(Primer)
PKIXC 5'-AG/AIAG/AIACICCG/AA/TAIGCCCACCAG/ATC-3'

CaMIIK

(Subdomain) I

(Consensus) LGKGAFSVV

(Primers)
PKOID1 5'-T/CTIGGIAAG/AGGIGCITTT/CTCIGTIGT-3'
PKOID2 5'-T/CTIGGIAAG/AGGIGCITTT/CAGT/CGTIGT-3'

(Subdomain) IX

(Consensus) DIWACGVIL

(Primer)
PKIXD 5'-AG/AIATIACICCG/ACAIGCCCAIATG/ATC-3'

cdc2

(Subdomain) I

(Consensus) V/IGEGTYGVV

(Primer)
PKOIE 5'-G/ATIGGIGAG/AGGIACITAT/CGGIGTIGT-3'

(Subdomain) IX

(Consensus) DIWSVGCIF

(Primers)
PKIXE1 5'-AAIATG/ACAICCIACIGACCAIATG/ATC-3'
PKIXE2 5'-AAIATG/ACAICCIACG/ACTCCAIATG/ATC-3'

Src

(Subdomain) I

(Consensus) LGQGCFGEV

(Primer)
PKOIF 5'-T/CTIGGICAG/AGGITGT/CTTT/CGGIGAG/AGT-3'

(Subdomain) IX

(Consensus) DVWSFGILL

(Primers)
PKIXF1 5'-AG/AIAG/AIATICCGAAIGACCAIACG/ATC-3'
PKIXF2 5'-AG/AIAG/AIATICCGAAG/ACTCCAIACG/ATC-3'

Fyn

(Subdomain) I

(Consensus) LGAGQFGEV

(Primer)
PKOIG 5'-T/CTIGGIGCIGGICAG/ATTT/CGGIGAG/AGT-3'

(Subdomain) IX

(Consensus) DVWSFGILL

(Primers)
PKIXG1 5'-AG/AIAG/AIATICCGAAIGACCAIACG/ATC-3'
PKIXG2 5'-AG/AIAG/AIATICCGAAG/ACTCCAIACG/ATC-3'

Abl

(Subdomain) I

(Consensus) LGGGQYGE/DV

(Primer)
PKOIH 5'-T/CTIGGIGGIGGICAG/ATAT/CGGIGAIGT-3'

(Subdomain) IX

(Consensus) DVWAFGVLL

(Primer)
PKIXH 5'-AG/AIAG/AIACICCG/AAAIGCCCAIACG/ATC-3'

Table 2

The extent of the amino acid homology between members of the MAP kinase family and the cdc2 family[a]

PK[b]	ATMPK1	ATMPK2	MsERK1	XLMPK1	XLXp42	ERK2	ERK1	spk1	MPK1	KSS1	FUS3	ATcdc2	HsCDK2
ATMPK1[c]	*	88.7	56.1	49.4	49.4	49.3	49.0	49.3	46.2	46.2	46.7	39.7	37.7
ATMPK2[c]		*	55.2	50.1	50.1	50.0	48.9	49.3	46.2	45.3	44.2	39.0	37.3
MsERK1[d]			*	52.3	52.0	51.6	51.7	52.8	49.1	49.1	50.1	39.8	37.0
XLMPK1[e]				*	98.1	95.5	84.7	55.3	49.1	52.0	49.9	41.4	37.5
Xp42[f]					*	95.2	83.8	55.3	48.8	52.3	49.9	40.7	37.5
ERK2[g]						*	85.6	56.3	49.4	52.0	50.4	40.4	37.5
ERK1[h]							*	52.8	48.5	49.7	50.1	39.4	37.8
spk1[i]								*	48.5	61.2	55.4	38.6	37.1
MPK1[j]									*	50.2	45.3	36.7	36.5
KSS1[k]										*	57.5	41.1	38.3
FUS3[l]											*	37.5	37.3
ATcdc2[m]												*	69.0
CDK2[n]													*

[a]The extent of the identity (%) between sequences was calculated, using the GENETYX software system (Software Development Co., Tokyo, Japan) and an NEC PC9801 computer.

[b]The superscripts indicate sources, as follows: [c]*Arabidopsis* ATMPK1 and ATMPK2; [d]alfalfa MsERK1 (Duerr *et al.*, 1993); [e]*Xenopus* MPK1 (Gotoh *et al.*, 1991); [f]*Xenopus* Xp42 (Posada *et al.*, 1991); [g]rat ERK2 (Boulton *et al.*, 1991); [h]rat ERK1 (Boulton *et al.*, 1990); [i]fission yeast spk1 (Toda *et al.*, 1991); [j]budding yeast MPK1 (Torres *et al.*, 1991); [k]budding yeast KSS1 (Courchesne *et al.*, 1989); [l]budding yeast FUS3 (Elion *et al.*, 1990); [m]*Arabidopsis* cdc2 (Ferreira *et al.*, 1991); [n]human CDK2 (Ninomiya-Tsuji *et al.*, 1991)

Table 3

The nucleic acid lengths, deduced amino acid lengths, and deduced molecular weights of ATMPKs.

clone name	nucleic acid lengths (bp)	deduced amino acid lengths (aa)	deduced molecular weights (Da)
ATMPK1	1520	370	42,685
ATMPK2	1562	376	43,138
ATMPK3	1370	370	42,702
ATMPK4	1307	376	42,865
ATMPK5	1533	376	43,119
ATMPK6	1562	395	45,056
ATMPK7	1657	368	42,381
ATMPK8	2086	562	63,713
ATMPK9	1948	510	58,440

The deduced molecular weights of ATMPKs were calculated, using the GENETYX software system (Software Development Co., Tokyo, Japan) and an NEC PC9801 computer.

References

Aguan, K., Sugawara, K., Suzuku, N. and Kusano, T. (1993) Low-temperature-dependent expression of a rice gene encoding a protein with a leucine-zipper motif. *Mol. Gen. Genet.* **240,** 1-8.

Alcorta, D. A., Crews, C. M., Sweet, L. J., Bankston, L., Jones, S. W. and Erikson, R. L. (1989) Sequence and expression of chicken and mouse *rsk*: homologs of *Xenopus laevis* ribosomal S6 kinase. *Mol. Cell. Biol.* **9,** 3850-3859.

Alvarez, E., Northwook, I. C., Gonzalez, F. A., Latour, D. A., Seth, A., Abate, C., Curran, T. and Davis, J. R. (1991) Pro-Leu-Ser/Thr-Pro is a consensus primary sequence for substrate protein phosphorylation. *J. Biol. Chem.* **266,** 15277-15285.

Anderberg, R. J. and Walker-Simmons, M. K. (1992) Isolation of a wheat cDNA clone for an abscisic acid-inducible transcript with homology to protein kinases. *Proc. Natl. Acad. Sci. USA.* **89,** 10183-10187.

Anderson, N., Maller, J., Tonks, N. and Sturgill, T. W. (1990) Requirement for integration of signals from two distinct phosphorylation pathways for activation of MAP kinase. *Nature* **343**, 651-653

Ballou, L. M., Luther, H. and Thomas, G. (1991) MAP2 kinase and 70K S6 kinase lie on distinct signalling pathways. *Nature* **349,** 348-350.

Banno, H., Hirano, K., Nakamura, T., Irie, K., Nomoto, S., Matsumoto, K. and Machida, Y. (1993) NPK1, a tobacco gene that encodes a protein with a domain homologous to yeast BCK1, STE11 and Byr2 protein kinases. *Mol. Cell. Biol.* **13,** 4745-4752.

Biggs III, W. H. and Zipursky, S. L. (1992) Primary structure, expression, and signal-dependent tyrosine phosphorylation of a *Drosophila rolled* locus of extracellular signal-regulated kinase. *Proc. Natl. Acad. Sci. USA.* **89,** 6295-6299.

Bladt, F. and Birchmeier, C. (1993) Characterization and expression analysis of the murine *rck* gene: a protein kinase with a potential function in sensory cells. *Differentiation* **53,** 115-122.

Blenis, J. (1993) Signal transduction via the MAP kinases: Proceed at your own RSK. *Proc. Natl. Acad. Sci. USA.* **90,** 5889-5892.

Boulton, T. G., Yancopoulos, G. D., Gregory, J. S., Slaughter, C., Moomaw, C., Hsu, J. and Cobb, M. H. (1990) An insulin stimulated kinase similar to yeast kinases involved in cell cycle control. *Science* **249,** 64-67.

Boulton, T. G., Nye, S. H., Robbins, D. J., Ip, N. Y., Radziejewska, E., Morgenbesser, S. D., DePinho, R. A., Panayotatos, N., Cobb, M. H. and Yancopoulos, G. D. (1991) ERKs: A family of protein serine/threonine kinases that are activated and tyrosine phosphorylated in response to insulin and NGF. *Cell* **65,** 663-675.

Boyle, W. J., van der Geer, P. and Hunter, T. (1991) Phosphopeptide mapping and phosphoamino acid analysis by two-dimentional separation on thin layer cellulose plates. *Methods Enzymol.* **201,** 110-149.

Braam, J. and Davis, R. W. (1990) Rain-, wind- and touch-induced expression of calmodulin and calmodulin-related genes in *Arabidopsis*. *Cell* **60,** 357-364.

Brewster, J. L., Valoir, T. D., Dwyer, N. D., Winter, E. and Gustin, M. C. (1993) An osmosensing signal transduction pathway in yeast. *Science* **259,** 1760-1763.

Celenza, J. L. and Carlson, M. (1986) A yeast gene that is essential for release from glucose repression encodes a protein kinase. *Science* **233,** 1175-1180.

Chandler, P. M. and Robertson, M. (1994) Gene expression regulated by abscisic acid and its relation to stress tolerance. *Annu. Rev. Plant Physiol. Plant Mol. Biol.* **45,** 113-141.

Charest, D. L., Jilik, F., Harder, K., Pelech, S. L. and Mordret, G. (1993) Molecular cloning and characterization of human p43ERK1, a tyrosyl-phosphorylated and activated protein-seryl/threonyl kinase. *Mol. Cell. Biol.* **13,** 4679-4690.

Choi, K.-Y., Satterberg, B., Lyons, D. M. and Elion, E. A. (1994) Ste5 tethers multiple protein kinases in the MAP kinase cascade required for mating in *S. cerevisiae*. *Cell* **78,** 499-512.

Cobb, M. H., Boulton, T. G. and Robbins, D. J. (1991) Extracellular signal-related kinases: ERKs in progress. *Cell Regulation* **2,** 965-978.

Courchesne, W. E., Kunisawa, R. and Thorner, J. (1989) A putative protein kinase overcomes pheromone-induced arrest of cell cycling in *S. cerevisiae*. *Cell* **58,** 1107-1119.

Crews, C. M. and Erikson, R. L. (1992a) Purification of a murine protein-tyrosine/threonine kinase that phosphorylates and activates the *Erk-1* gene product: Relationship to the fission yeast *byr1* gene product. *Proc. Natl. Acad. Sci. USA.* **89,** 8205-8209.

Crews, C. M., Alessandro, A. and Erikson, R. L. (1992b) The primary structure of MEK, a protein kinase that phosphorylates the *ERK* gene product. *Science* **258,** 478-480.

Daminov, J. A., Stentzler, L., Lee, S., Schwartz, J. J., Leisner, S. and Howell, S. H. (1992) Cytokinins and auxins control the expression of a gene in *Nicotiana plumbaginifolia* cells by feedback regulation. *Plant Cell* **4,** 451-461.

Datta, N. and Cashmore, A. R. (1989) Binding of a pea nuclear protein to promoters of certain photoregulated genes is modulated by phosphorylation. *Plant Cell* **1,** 1069-1077.

Dent, P., Haser, W., Haystead, T. A. J., Vincent, L. A., Roberts, T. M. and Sturgill, T. W. (1992) Activation of mitogen-activated protein kinase kinase by v-raf in NIH 3T3 cells and *in vitro*. *Science* **257,** 1404-1407.

Dérijard, B., Hibi, M., Wu, I-H., Barrett, T., Su, B., Deng, T., Karin, M., and Davis, R. J. (1994) JNK1: a protein kinase stimulated by UV light and Ha-Ras that binds and phosphorylates the c-Jun activation domain. *Cell* **76,** 1025-1037.

Downward, J. (1994) Regulating S6 kinase. *Nature* **371,** 378-379.

Duerr, B., Gawienowski, M., Ropp, T. and Jacobs, T. (1993) MsERK1: a mitogen-activated protein kinase from a flowering plant. *Plant Cell* **5,** 87-96.

Dzelzkalns, V. A., Nasrallah, J. B. and Nasrallah, M. E. (1992) Cell-cell communication in plants: Self-incompatibility in flower development. *Dev. Biol.* **153,** 70-82.

Elion, E. A., Grisafi, P. L. and Fink, G. R. (1990) FUS3 encodes a cdc2/CDC28-related kinase required for the transition from mitosis to conjugation. *Cell* **60,** 649-664.

Erikson, R. L. (1991) Structure, expression and regulation of protein kinases involved in the phosphorylation of ribosomal protein S6. *J. Biol. Chem.* **266,** 6007-6010.

Feiler, H. S. and Jacobs, T. W. (1990) Cell division in higher plants: A *cdc2* gene, its 34-KDa product, and histone H1 kinase activity in pea. *Proc. Natl. Acad. Sci. USA.* **87,** 5397-5401.

Felix, G., Grosskopf, D. G., Regenass, M. and Boller, T. (1991) Rapid changes of protein phosphorylation are involved in transduction of the elicitor signal in plant cells. *Proc. Natl. Acad. Sci. USA.* **88,** 8831-8834.

Ferreira, P. C. G., Hemerly, A. S., Villarroel, R., Montagu, M. V. and Inzé, D. (1991) The *Arabidopsis* functional homolog of the $p34^{cdc2}$ protein kinase. *Plant Cell* **3,** 531-540.

Finkenzeller, G., Marme, D. and Hug, H. (1990) Sequence of human protein kinase C alpha. *Nucleic Acids Res.* **18,** 2183-2183.

Galau, G. A., Hughes, D. W. and Dure, L., III. (1986) Abscisic acid induction of cloned late embryogenesis abundant (*Lea*) mRNAs. *Plant Mol. Biol.* **7,** 155-170.

Galcheva-Gargova, Z., Dérijard, B., Wu, I-H. and Davis, R. J. (1994) An osmosensing signal transduction pathway in mammalian cells. *Science* **265,** 806-808.

Gallie, D. R. (1993) Posttranscriptional regulation of gene expression in plants. *Annu. Rev. Plant Physiol. Plant Mol. Biol.* **44,** 77-105.

Gartner, A., Nasmyth, K. and Ammerer, G. (1992) Signal transduction in *Saccharomyces cerevisiae* requires tyrosine and threonine phosphorylation of FUS3 and KSS1. *Genes Dev.* **6,** 1280-1292.

Geahlen, R. L., Anostario, M., Low, P. S. and Harrison, M. L. (1986) Detection of protein kinase activity in sodium dodecyl sulfate polyacrylamide gels. *Anal. Biochem.* **153,** 151-158.

GenBank, accession number: T04696 and T13020.

GenBank, accession number: Z14089.

Gille, H., Sharrocks, A. D. and Shaw, P. E. (1992) Phosphorylation of transcription factor $p62^{TCF}$ by MAP kinase stimulates ternary complex formation of c-fos promoter. *Nature* **358,** 414-417.

Gonzalez, F. A., Raden, D. L., Rigby, M. R. and Davis, R. J. (1992) Heterogeneous expression of four MAP kinase isoforms in human tissues. *FEBS Lett.* **304,** 170-178.

Gotoh, Y., Nishida, E., Yamashita, T., Hoshi, M., Kawakami, M. and Sakai, H. (1990) Microtubule-associated protein (MAP) kinase activated by NGF and EGF in PC12 cells. Identity with the mitogen-activated MAP kinase of fibroblastic cells. *Eur. J. Biochem.* **193,** 661-669.

Gotoh, Y., Moriyama, K., Matsuda, S., Okumura, E., Kishimoto, T., Kawasaki, H., Suzuki, K., Yahara, I., Sakai, H. and Nishida, E. (1991) *Xenopus* M-phase MAP-kinase: isolation of its cDNA and activation by MPF. *EMBO J.* **10,** 2661-2668.

Gotoh, Y., Nishida, E., Shimanuki, M., Toda, T., Imai, Y. and Yamamoto, M. (1993) *Schizosaccharomyces pombe* Spk1 is a tyrosine-phosphorylated protein functionally related to *Xenopus* mitogen-activated protein kinase. *Mol. Cell. Biol.* **13,** 6427-6434.

Grove, J., Bonerjee, P., Balasubramangan, A., Coffer, P., Price, P. J., Avruch, J. and Woodgett, J. R. (1991) Cloning and expression of two human p70 S6 kinase polypeptides differing only at their amino termini. *Mol. Cell. Biol.* **11,** 5541-5550.

Han, J., Lee, J.-D., Bibbs, L. and Ulevitch, R. J. (1994) A MAP kinase targeted by endotoxin and hyperosmolarity in mammalian cells. *Science* **265,** 808-811.

Hanks, S. K., Quinn, A. M. and Hunter, T. (1988) The protein kinase family: conserved features and deduced phylogeny of the catalytic domains. *Science* **241,** 42-52.

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,** 38-61.

Hashimoto, J., Hirabayashi, T., Hayano, Y., Hata, S., Ohashi, Y., Suzuka, I., Utsuji, T., Toh-E, A. and Kikuchi, Y. (1992) Isolation and characterization of cDNA clones encoding cdc2 homologues from *Oryza sativa*: a functional homologue and cognate variants. *Mol. Gen. Genet.* **233,** 10-16.

Hattori, S., Fukuda, M., Yamashita, T., Nakamura, S., Gotoh, Y. and Nishida, E. (1992) Activation of mitogen-activated protein kinase and its activator by *ras* in intact cells and in a cell-free system. *J. Biol. Chem.* **267,** 20346-20351.

Hayashida, N., Mizoguchi, T., Yamaguchi-Shinozaki, K. and Shinozaki, K. (1992) Characterization of a gene that encodes a homologue of protein kinase in *Arabidopsis thaliana*. *Gene* **121,** 325-330.

Hayashida, N., Mizoguchi, T. and Shinazaki, K. (1993) Cloning and characterization of a plant gene encoding a protein kinase. *Gene* **124,** 251-255.

Hemerly, A. S., Ferreira, P., Engler, J. de A., Van Montagu, M., Engler, G. and Inzé, D. (1993) *cdc2a* expression in *Arabidopsis* is linked with competence for cell division. *Plant Cell* **5,** 1711-1723.

Her, J.-H., Wu, J., Rall, T. B., Sturgill, T. W. and Weber, M. J. (1991) Sequence of pp42/MAP kinase, a serine/threonine kinase regulated by tyrosine phosphorylation. *Nucleic Acids Res.* **19,** 37-43.

Hirayama, T., Imajuku, Y., Anai, T., Matsui, M. and Oka, A. (1991) Identification of two cell-cycle-controlling cdc2 gene homologs in *Arabidopsis thaliana*. *Gene* **105,** 159-165.

Hirt, H., Páy, A., Györgyey, J., Bakó, L., Németh, K., Bögre, L., Schweyen, R. J., Heberle-Bors, E. and Dudits, D. (1991) Complementation of a yeast cell cycle mutant by an alfalfa cDNA encoding a protein kinase homologous to $p34^{cdc2}$. *Proc. Natl. Acad. Sci. USA.* **88,** 1636-1640.

Howe, L. R., Leevers, S. J., Gómez, N., Nakielny, S., Cohen, P. and Marshall, C. J. (1992) Activation of MAP kinase pathway by protein kinase raf. *Cell* **71,** 335-342.

Imajuku, Y., Hirayama, T., Endoh, H. and Oka, A. (1992) Exon-intron organization of the *Arabidopsis thaliana* protein kinase genes *CDC2a* and *CDC2b. FEBS Lett.* **304,** 73-77.

Irie, K., Takase, M., Lee, K. S., Levin, D. E., Araki, H., Matsumoto, K. and Oshima, Y. (1993) *MKK* and *MKK2*, which encode *Saccharomyces cerevisiae* mitogen-activated protein kinase-kinase homologs, function in the pathway mediated by protein kinase C. *Mol. Cell. Biol.* **13,** 3076-3083.

Jaffe, M. J. (1973) Thigmomorphogenesis: the response of plant growth and development to mechanical stimulation. *Planta* **114,** 143-157.

Jefferies, H. B. J., Reinhard, C., Kozma, S. C. and Thomas, G. (1994) Rapamycin selectively represses translation of the "polypyrimidine tract" mRNA family. *Proc. Natl. Acad. Sci. USA.* **91,** 4441-4445.

Jonak, C., Páy, A., Bögre, L., Hirt, H. and Heberle-Bors, E. (1993) The plant homologue of MAP kinase is expressed in a cell cycle-dependent and organ-specific manner. *Plant Journal* **3** , 611-617.

Jonak, C., Heberle-Bors, E. and Hirt, H. (1994) MAP kinases: universal multi-purpose signaling tools. *Plant Mol. Biol.* **24,** 407-416.

Jones, P. G., VanBogelen, R. A. and Neidhardt, F. C. (1987) Induction of proteins in response to low temperature in *Escherichia coli. J. Bacteriol.* **169,** 2092-2095.

Jones, S. W., Erikson, E., Blenis, J., Maller, J. L. and Erikson, R. L. (1988) A *Xenopus* ribosomal protein S6 kinase has two apparent kinase domains that are each similar to distinct protein kinases. *Proc. Natl. Acad. Sci. USA.* **85,** 3377-3381.

Katagiri, F. and Chua, N.-H. (1992) Plant trnscription factors: present knowledge and future challenges. *Trends Genet.* **8,** 22-27.

Kieber, J. J., Rothenberg, M., Roman, G., Feldmann, K. A. and Ecker, J. R. (1993) *CTR1*: a negative regulator of the ethylene response pathway in *Arabidopsis*, encodes a member of the Raf family of protein kinases. *Cell* **72,** 427-441.

Kiyosue, T., Yamaguchi-Shinozaki, K. and Shinozaki, K. (1993) In Abstracts of *XV International Botanical Congress*, p.6356.

Knight, M. R., Smith, S. M. and Trewavas, A. J. (1992) Wind-induced plant motion immediately increases cytosolic calcium. *Proc. Natl. Acad. Sci. USA.* **89,** 4967-4971.

Kondo, K. and Inoue, M. (1992) Yeast NSR1 protein that has structural similarity to mammalian nucleolin is involved in pre-RNA processing. *J. Biol. Chem.* **267,** 16252-16258.

Kosako, H., Gotoh, Y., Matsuda, S., Ishikawa, M. and Nishida, E. (1992) *Xenopus* MAP kinase activator is a serine/threonine/tyrosine kinase activated by threonine phosphorylation. *EMBO J.* **11,** 2903-2908.

Kosako, H., Nishida, E. and Gotoh, Y. (1993) cDNA cloning of MAP kinase kinase reveals kinase cascade pathways in yeasts to vertebrates. *EMBO J.* **12,** 787-794.

Kozma, S. C., Ferrari, S., Bassand, P., Siegmann, M., Totty, N. and Thomas, G. (1990) Cloning of the mitogen-activated S6 kinase from rat

liver reveals an enzyme of the second messenger subfamily. *Proc. Natl. Acad. Sci. USA.* **87,** 7365-7369.

Krikorian, A. D., Kelly, K. and Smith, D. (1987) Hormones in tissue culture and micropropagation. In *Plant Hormones and Their Role in Plant Growth and Development.* (Davies, P. J., ed.) Dordrecht, Martinus Nijhoff Publishers, pp. 593-613.

Kyriakis, J. M., App, H., Zhang, X.-F., Banerjee, P., Brautigan, D. L., Rapp, U. R. and Avuruch, J. (1992) Raf-1 activates MAP kinase-kinase. *Nature* **358,** 417-421.

Kyriakis, J. M., Banerjee, P., Nikolakaki, E., Dai, T., Rubie, E. A., Ahmad, M. F., Avuruch, J. and Woodgett, J. R. (1994) The stress-activated protein kinase subfamily of c-Jun kinases. *Nature* **369,** 156-160.

Laemmli, U. K. (1970) Cleavage of structual proteins during assembly of the head of bacteriophage T4. *Nature* **227,** 680-685.

Lange-Carter, C. A., Pleiman, C. M., Gardner, A. M., Blumer, K. J. and Johnson, G. L. (1993) A divergence in the MAP kinase regulatory network difined by MEK kinase and Raf. *Science* **260,** 315-319.

Lawton, M. A., Yamamoto, R. T., Hanks, S. K. and Lamb, C. J. (1989) Molecular cloning of plant transcripts encoding protein kinase homologs. *Proc. Natl. Acad. Sci. USA.* **86,** 3140-3144.

Lee, K. S. and Levin, D. E. (1992) Dominant mutations in a gene encoding a putative protein kinase (*BCK1*) bypass the requirement for a *Saccharomyces cerevisiae* protein kinase C homolog. *Mol. Cell. Biol.* **12,** 172-182.

Leevers, S. J. and Marshall, C. J. (1992) Activation of extracellular signal-regulated kinase, ERK2, by p21 *ras* oncoprotein. *EMBO J.* **11,** 569-574.

Levin, D. E., Field, F. O., Kunisawa, R., Bishop, J. M. and Thorner, J. (1990) A candidate protein kinase C gene, *PKC1*, is required for the *S. cerevisiae* cell cycle. *Cell* **62,** 213-224.

Lin, L., Wartmann, M., Lin, A., Knopf, J. L., Seth, A. and Davis, R. J. (1993) $cPLA_2$ is phosphorylated and activated by MAP kinase. *Cell* **72,** 269-278.

Lin, X., Feng, X. and Watson, J. C. (1991) Differential accumulation of transcripts encoding protein kinase homologs in greening pea seedlings. *Proc. Natl. Acad. Sci. USA.* **88,** 6951-6955.

Lütcke, H. A., Chow, K. C., Mickel, F. S., Moss, K. A., Kern, H. F. and Scheele, G. A. (1987) Selection of AUG initiation codons differs in plants and animals. *EMBO J.* **6,** 43-48.

Maldonado, F. and Hanks, S. K. (1988) A cDNA clone encoding human cAMP-dependent protein kinase catalytic subunit type alpha. *Nucleic Acids Res.* **16,** 8189-8190.

Maniatis, T., Fitsch, E. F. and Sambrook, J. (1982) Molecular Cloning: a Laboratory Manual. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY.

Marcus, S., Polverino, A., Barr, M. and Wigler, M. (1994) Complexes between STE5 and components of the pheromone-responsive mitogen-activated protein kinase module. *Proc. Natl. Acad. Sci. USA.* **91,** 7762-7766.

Martin, G. B., Brommonschenkel, S. H., Chunwongse, J., Fray, A., Ganal, M. W., Spivey, R., Wu, T., Earle, E. D. and Tanksley, S. D. (1993) Map-based cloning of a protein kinase gene conferring disease resistance in tomato. *Science* **262,** 1432-1436.

Matsuda, S., Kosako, H., Takenaka, K., Moriyama, K., Sakai, H., Akiyama, T. Gotoh, Y. and Nishida, E. (1992) *Xenopus* MAP kinase activator: identification and function as a key intermediate in the phosphorylation cascade. *EMBO J.* **11,** 973-982.

Matsuda, S., Gotoh, Y. and Nishida, E. (1993) Phosphorylation of *Xenopus* mitogen-activated protein (MAP) kinase kinase by MAP kinase kinase kinase and MAP kinase. *J. Biol. Chem.* **268,** 3277-3281.

Matsushime, H., Jinno, A., Takagi, N. and Shibuya, M. (1994) A novel mammalian protein kinase gene (*mak*) is highly expressed in testicular germ cells at and after meiosis. *Mol. Cell. Biol.* **10,** 2261-2268.

Mizoguchi, T., Hayashida, N., Yamaguchi-Shinozaki, K., Harada, H. and Shinozaki, K. (1992) Nucleotide sequence of a cDNA encoding a protein kinase homologue in *Arabidopsis thaliana*. *Plant Mol. Biol.* **18,** 809-812.

Mizoguchi, T., Yamaguchi-Shinozaki, K., Hayashida, N., Kamada, H. and Shinozaki, K. (1993a) Cloning and characterization of two cDNAs encoding casein kinase II catarytic subunits in *Arabidopsis thaliana*. *Plant Mol. Biol.* **21,** 279-289.

Mizoguchi, T., Hayashida, N., Yamaguchi-Shinozaki, K., Kamada, H. and Shinozaki, K. (1993b) ATMPKs: a gene family of plant MAP kinases in *Arabidopsis thaliana*. *FEBS Lett.* **336,** 440-444.

Mizoguchi, T., Gotoh, Y., Nishida, E., Yamaguchi-Shinozaki, K., Hayashida, N., Iwasaki, T., Kamada, H. and Shinozaki, K. (1994a) Characterization of two cDNAs that encode MAP kinase homologues in *Arabidopsis thaliana* and analysis of the possible role of auxin in activating such kinase activities in cultured cells. *Plant Journal* **5,** 111-122.

Mizoguchi, T., Hayashida, N., Yamaguchi-Shinozaki, K., Kamada, H. and Shinozaki, K. (1994b) Cloning and sequencing of a novel serine/threonine protein kinase in *Arabidopsis thaliana*. *Plant Physiol.* in press.

Monroy, A. F., Sarhan, F. and Dhindsa, R. S. (1993) Cold-induced changes in freezing tolerance, protein phosphorylation, and gene expression. *Plant Physiol.* **102,** 1227-1235.

Mukhopadhyay, N. K., Price, D. J., Kyriakis, J. M., Pelech, S., Sanghera, J. and Avruch, J. (1992) An array of insulin-activated, proline-directed serine/threonine protein kinases phosphorylate the p70 S6 kinase. *J. Biol. Chem.* **267,** 3325-3335.

Nagata, T., Okada, K., Takebe, I. and Mitsui, C. (1981) Delivery of tobacco mosaic virus RNA into plant protoplasts mediated by reverse-phase evaporation vesicles (liposomes). *Mol. Gen. Genet.* **184,** 161-165.

Nagy, F., Kay, S. A. and Chua, N-H. (1988) *In* SV Gelvin, RA Schilperoort eds. Plant Molecular Biology Manual B4: 1-4, Kluwer Academic Publishers, Dordrecht.

Nasrallah, J. B., Nishio, T. and Nasrallah, M. E. (1991) The self-incompatibility genes of *Brassica*. *Annu. Rev. Plant Physiol. Plant Mol. Biol.* **42,** 393-422.

Nei, M. (1987) *Molecular Evolutionary Genetics.* (New York: Columbia University Press), p.293-298.

Ninomiya-Tsuji, J., Nomoto, S., Yasuda, H., Reed, S. I. and Matsumoto, K. (1991) Cloning of a human cDNA encoding a CDC2-related kinase by complementation of a budding yeast *cdc28* mutation. *Proc. Natl. Acad. Sci. USA.* **88,** 9006-9010.

Nishida, E. and Gotoh, Y. (1992) Mitogen-activated protein kinase and cytoskeleton in mitogenic signal transduction. *Int. Rev. Cytol.* **138,** 211-238.

Nishida, E. and Gotoh, Y. (1993) The MAP kinase cascade is essential for diverse signal transduction pathways. *Trends Genet.* **18,** 128-131.

Owaki, H., Marker, R., Boulton, T. G., Cobb, M. H. and Geppert, T. D. (1992) Extracellular signal-regulated kinases in T cells: Characterization of human ERK1 and ERK2 cDNAs. *Biochem. Biophys. Res. Commun.* **182,** 1416-1422.

Payne, D., Rossomando, A., Martino, P., Erickson, A., Her, J., Shabanowitz, J., Hunt, D., Weber, M. and Sturgill, T. W. (1991) Identification of the regulatory phosphorylation sites in pp42/mitogen-activated protein kinase (MAP kinase). *EMBO J.* **10,** 885-892.

Pearson, R. B. and Kemp, B. E. (1991) Protein kinase phosphorylation site sequences and consensus specificity motifs: tabulations. *Methods Enzymol.* **200,** 62-81.

Pelech, S. L. and Sanghera, J. S. (1992a) MAP Kinases: Charting the Regulatory Pathways. *Science* **257,** 1355-1356.

Pelech, S. L. and Sanghera, J. S. (1992b) Mitogen-activated protein kinases: versatile transducers for cell signaling. *TIBS.* **17,** 233-238.

Posada, J., Sanghera, J., Pelech, S., Aebersold, R. and Cooper, J. A. (1991) Tyrosine phosphorylation and activation of homologous protein kinases during oocyte maturation and mitotic activation of fibroblasts. *Mol. Cell. Biol.* **11,** 2517-2528.

Posada, J., Yew, N., Ahn, N. G., Vande Woude, G. F. and Cooper, J. A. (1993) Mos stimulates MAP kinase in *Xenopus oocytes* and activates a MAP kinase kinase *in vitro*. *Mol. Cell. Biol.* **13,** 2546-2553.

Pulverer, B. J., Kyriakis, J. M., Avuruch, J., Nikolakaki, E. and Woodgett, J. R. (1991) Phosphorylation of c-jun mediated by MAP kinases. *Nature* **353,** 670-674.

Rechsteiner, M., Rogers, S. and Rote, K. (1987) Protein structure and intracellular stability. *Trends Biochem. Sci.* **12,** 390-394.

Rhodes, N., Connell, L. and Errede, B. (1990) STE11 is a protein kinase required for cell-type-specific transcription and signal transduction in yeast. *Genes Dev.* **4,** 1862-1874.

Roberts, D. M. and Harmon, A. C. (1992) Calcium-modulated proteins: targets of intracellular calcium signals in higher plants. *Annu. Rev. Plant Physiol Plant Mol. Biol.* **43,** 375-414.

Roe, J. L., Rivin, C. J., Sessions, R. A., Feldmann, K. A. and Zambryski, P. C. (1993) The *tousled* gene in *A. thaliana* encodes a protein kinase homolog that is required for leaf and flower development. *Cell* **75,** 939-950.

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

Sarokin, L. P. and Chua, N.-H. (1992) Binding sites for two novel phosphoproteins, 3AF5 and 3AF3, are required for *rbcS-3A* expression. *Plant Cell* **4,** 473-483.

Seger, R., Ahn, N. G., Posada, J., Munar, E. S., Jensen, A. M., Cooper, J. A., Cobb, M. H. and Krebs, E. G. (1992) Purification and characterization of mitogen-activated protein kinase activator(s) from epidermal growth factor-stimulated A431 cells. *J. Biol. Chem.* **267,** 14373-14381.

Seth, A., Alvarez, E., Gupta, S. and Davis, R. J. (1991) A phosphorylation site located in the NH_2-terminal domain of c-myc increases transactivation of gene expression. *J. Biol. Chem.* **266,** 23521-23524.

Shiina, N., Moriguchi, T., Ohta, K., Gotoh, Y. and Nishida, E. (1992) Regulation of a major microtubule-associated protein by MPF and MAP kinase. *EMBO J.* **11,** 3977-3984.

Shimbara, N., Orino, E., Sone, S., Ogura, T., Takashina, M., Shono, M., Tamura, T., Yasuda, H., Tanaka, K. and Ichihara, A. (1992) Regulation of gene expression of proteasomes (multi-protease complexes) during growth and differentiation of human hematopoietic cells. *J. Biol. Chem.* **267,** 18100-18109.

Smith, D. B. and Johnson, K. S. (1988) Single-step purification of polypeptides expressed in *Escherichia coli* as fusions with glutathione *S*-transferase. *Gene* **67,** 31-40.

Stafstrom, J. P., Altschuler, M. and Anderson, D. H. (1993) Molecular cloning and expression of a MAP kinase homologue from pea. *Plant Mol. Biol.* **22,** 83-90.

Stein, J. C., Howlett, B., Boyers, D. C., Nasrallah, M. E. and Nasrallah, J. B. (1991) Molecular cloning of a putative receptor protein kinase gene encoded at the self-incompatibility locus of *Brassica oleracea*. *Proc. Natl. Acad. Sci. USA.* **88,** 8816-8820.

Stokoe, D., Campbell, D. G., Nakielny, S., Hidaka, H., Leevers, S. J., Marshall, C. and Cohen, P. (1992) MAPKAP kinase-2; a novel protein kinase activated by mitogen-activated protein kinase. *EMBO J.* **11,** 3985-3994.

Sturgill, T. W., Ray, L. B., Erikson, E. and Maller J. L. (1988) Insulin-stimulated MAP-2 kinase phosphorylates and activates ribosomal protein S6 kinase II. *Nature* **334,** 715-718.

Sturgill, T. W. and Wu, J. (1991) Recent progress in characterization of protein cascades for phosphorylation of ribosomal protein S6. *Biochem. Biophys. Acta.* **1092,** 350-357.

Tanner, B. and Mueckler, M. (1993) Molecular cloning of a mouse extracellular signal regulated kinase (erk-1). *Biochem. Biophys. Acta.* **1171,** 319-320.

Teague, M. A., Chaleff, D. T. and Errede, B. (1986) Nucleotide sequence of the yeast regulatory gene *STE7* predicts a protein homologous to protein kinase. *Proc. Natl. Acad. Sci. USA.* **83,** 7371-7375.

Thomas, G. (1992) MAP kinase by any other name smells just as sweet. *Cell* **68,** 1-4.

Thomas, S. M., DeMarco, M., D'Arcangelo, G., Halegoua, S. and Brugge, J. S. (1992) Ras is essential for nerve growth factor- and phorbol ester-induced tyrosine phosphorylation of MAP kinases. *Cell* **68,** 1031-1040.

Tilly, B. C., Van Den Berghe, N., Tertoolen, L. G. J., Edixhoven, M. J. and De Jonge, H. R. (1993) Protein tyrosine phosphorylation is involved in osmoregulation of ionic conductances. *J. Biol. Chem.* **268,** 19919-19922.

Toda, T., Shimanuki, M. and Yanagida, M. (1991) Fission yeast genes that confer resistance to staurosporine encode an AP-1-like transcription factor and a protein kinase related to the mammalian ERK1/MAP2 and budding yeast *FUS3* and *KSS1* kinases. *Genes Dev.* **5,** 60-73.

Thomashow, M. F. (1990) Molecular genetics of cold acclimation in higher plants. In *Advances in Genetics.: Academic Press.* **28,** pp.99-131.

Thomashow, M. F., Gilmour, S. J. and Lin, C. (1993) Cold-regulated genes of *Arabidopsis thaliana*. In *Advances in Plant Cold Hardiness.: CRC Press.* pp.31-44.

Torres, L., Martín, H., García-Saez, M. I., Arroyo, J., Molina, M., Sánchez, M. and Nombela, C. (1991) A protein kinase gene complements the lytic phenotype of *Saccharomyces cerevisiae lyt2* mutants. *Mol. Microbiol.* **5,** 2845-2854.

Tsuda, L., Inoue, H. Y., Yoo, M-A., Mizuno, M., Hata, M., Lim, Y-M., Adachi-Yamada, T., Ryo, H., Masamune, Y. and Nishida, Y. (1993) A protein kinase similar to MAP kinase activator acts downstream of the Raf kinase in *Drosophila*. *Cell* **72,** 407-414.

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

Urao, T., Katagiri, T., Mizoguchi, T., Yamaguchi-Shinozaki, K., Hayashida, N. and Shinozaki, K. (1994) Two genes that encode Ca^{2+}-dependent protein kinases are induced by drought and high-salt stresses in *Arabidopsis thaliana*. *Mol. Gen. Genet.* **244,** 331-340.

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, Y., Xu, H. -P., Riggs, M., Rodgers, L. and Wigler, M. (1991) *byr2*, a *Schizosaccharomyces pombe* gene encoding a protein kinase capable of partial suppression of the *ras1* mutant phenotype. *Mol. Cell. Biol.* **11,** 3554-3563.

Whiteway, M., Dignard, D. and Thomas, D. Y. (1992) Dominant negative selection of heterologous genes: isolation of *Candida albicans* genes that interfere with *Saccharomyces cerevisiae* mating factor-induced cell cycle arrest. *Proc. Natl. Acad. Sci. USA.* **89,** 9410-9414.

Wood, K. W., Sarnecki, C., Roberts, T. M. and Blenis, J. (1992) *ras* mediates nerve growth factor receptor modulation of three signal-transducing protein kinases: MAP kinase, Raf-1, and RSK. *Cell* **68,** 1041-1050.

Wray, J. L. (1992) Inducible plant proteins:Their biochemistry and molecular biology. Society for experimental biology seminar series 49 Cambridge University Press.

Wu, J., Harrison, J. K., Vincent, L. A., Haystead, C., Haystead, T. A. J., Michel, H., Hunt, D. F., Lynch, k. R. and Sturgill, T. W. (1993) Molecular structure of a protein-tyrosine/threonine kinase activating p42 mitogen-activated protein (MAP) kinase: MAP kinase kinase. *Proc. Natl. Acad. Sci. USA.* **90,** 173-177.

Yamaguchi-Shinozaki, K., Koizumi, M., Urao, S. and Shinozaki, K. (1992) Molecular cloning and characterization of 9 cDNAs for genes that are responsive to desiccation in *Arabidopsis thaliana*: sequence analysis of one cDNA clone that encodes a putative transmembrane channel protein. *Plant Cell Physiol.* **33,** 217-224.

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

Yoshida, K., Nagano, Y., Murai, N. and Sasaki, Y. (1993) Phytochrome-regulated expression of the genes encoding the small GTP-binding proteins in peas. *Proc. Natl. Acad. Sci. USA.* **90,** 6636-6640.

Zhang, F., Strand, A., Robbins, D., Cobb, M. H. and Goldsmith, E. J. (1994) Atomic structure of the MAP kinase ERK2 at 2.3 Å resolution. *Nature* **367,** 704-711.

Zhang, S. - H., Lawton, M. A., Hunter, T. and Lamb, C. J. (1994a) *atpk1*, a novel ribosomal protein kinase gene from *Arabidopsis*: I. isolation, characterization, and expression. *J. Biol. Chem.* **269,** 17586-17592.

Zhang, S. - H., Broome, M. A., Lawton, M. A., Hunter, T. and Lamb, C. J. (1994b) *atpk1*, a novel ribosomal protein kinase gene from *Arabidopsis*: II. function and biochemical analysis of the encoded protein. *J. Biol. Chem.* **269,** 17593-17599.

Acknowledgments

I wish to express my sincere appreciation to Dr. K. Shinozaki of RIKEN, Tsukuba Life Science Center, Professor H. Kamada, Professor T. Fujii, and Professor H. Harada of the University of Tsukuba for their invaluable suggestions and encouragement throughout the course of my research and in the preparation of this manuscript.

I am also indebted to Dr. N. Hayashida and Dr. K. Yamaguchi-Shinozaki of RIKEN, Tsukuba Life Science Center, for their invaluable suggestions, advices, and constructive criticisms of my research. My sincere thanks are also due to Dr. T. Hirayama, Dr. T. Urao, Dr. T. Kiyosue, Dr. T. Iwasaki of RIKEN, Tsukuba Life Science Center, Dr. T. Saitoh, Assistant Professor S. Satoh, Assistant Professor M. Wada, and Associate Professor S. Sakai of the University of Tsukuba, for their valuable suggestions and criticisms. I wish to thank Mrs. Y. Takagi for her assistance during the preparation of several manuscripts. Special thanks must be accorded to all members of the laboratories for useful discussions of the results.

I am also grateful to Professor E. Nishida and Dr. Y. Gotoh of University of Kyoto, Associate Professor S. Matsumoto of University of Gifu, Associate Professor K. Okada of National Institute for Basic Biology, for their valuable suggestions for several experimental techniques and the analyses of the results of this research.