Volume 336, number 3, 440–444 FEBS 13457 December 1993

ATMPKs: a gene family of plant MAP kinases in *Arabidopsis thaliana*

Tsuyoshi Mizoguchi[a,b], Nobuaki Hayashida[a], Kazuko Yamaguchi-Shinozaki[a], Hiroshi Kamada[b], Kazuo Shinozaki[a,b,*]

[a]*Laboratory of Plant Molecular Biology, The Institute of Physical and Chemical Research (Riken), Tsukuba Life Science Center, 3-1-1 Koyadai, Tsukuba, Ibaraki 305, Japan*
[b]*Institute of Biological Sciences, The University of Tsukuba, Tennohdai, Tsukuba, Ibaraki 305, Japan*

Received 11 November 1993

We previously reported two cDNAs for MAP kinases (cATMPK1 and cATMPK2) from a dicot plant, *Arabidopsis thaliana*. We describe here the cloning and characterization of five additional cDNAs encoding novel MAP kinases in *Arabidopsis*, cATMPK3, cATMPK4, cATMPK5, cATMPK6, and cATMPK7. The amino acid residues corresponding to the sites of phosphorylation (Thr-Glu-Tyr) that are involved in the activation of animal MAP kinases are conserved in all the seven putative ATMPK proteins. Genes for MAP kinases in *Arabidopsis* constitute a family that contains more than seven members. Sequence analysis suggests that there are at least three subfamilies in the family of *Arabidopsis* genes for MAP kinases.

Arabidopsis thaliana; Protein kinase; MAP kinase

1. INTRODUCTION

MAP kinases (mitogen-activated protein kinases) are serine/threonine kinases that are activated by various growth factors, differentiating factors, and in M-phase phosphorylation cascade reactions (see reviews [1–7]). They are thought to play key roles in integrating multiple intracellular signals transmitted by various second messengers. MAP kinases are unique in that they become activated when both tyrosine and threonine residues are phosphorylated [1–7]. Recently, MAPK kinase (a direct upstream activator of MAP kinase) and MAPKK kinase (a direct upstream activator of MAPK kinase) have been identified and their cDNAs have been cloned. The phosphorylation cascades including MAPKK kinase, MAPK kinase, and MAP kinase are conserved among various signal transduction pathways from yeast to vertebrate [1–7]. It has also been reported that raf gene products can function as a MAPKK kinase and that ras gene products can function upstream of the phosphorylation cascade of MAP kinase in animal signal-transduction pathways [1–7].

Multiple isoforms of MAP kinases have been identified in animals and yeasts [1–7]. Three distinct cDNAs encoding different MAP kinases (ERK1, ERK2, and ERK3) have been isolated in rat, although there is no evidence that ERK3 is a bona fide MAP kinase [8–9]. 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* [10–13]. One MAP kinase homolog (spk1) has also been reported in *Schizosaccharomyces pombe* [14].

There have been four reports on the isolation of cDNAs that encode MAP kinase homologs [15–18] in plants. These four MAP kinase homologs can be classified into two groups [15] based on their sequence analysis. Genomic Southern analysis provided evidence that there are several genes for MAP kinases in *Arabidopsis*. It is possible that MAP kinases play various roles in the signaling processes of plants and yeasts. We screened different cDNA libraries to isolate cDNAs encoding MAP kinase homologs as a first step in studying the functions of MAP kinases in plants.

Here we report the isolation of five additional cDNAs encoding MAP kinase homologs in *Arabidopsis* and discuss the structural differences in MAP kinases between animals, yeasts, and plants.

2. MATERIALS AND METHODS

2.1. *Cloning of cDNAs encoding MAP kinases from* Arabidopsis *cDNA libraries*

All the cDNA clones were isolated from two *Arabidopsis* cDNA libraries prepared from rosset plants grown for 4 weeks and from dehydrated rosset plants [19,20] by plaque hybridization, as described by Maniatis et al. [21]. A total of 8.0×10^5 plaques were screened under low-stringency hybridization conditions with the same probe used to isolate the two cDNA clones, ATMPK1 and ATMPK2 [15]. Filters

*Corresponding author. Fax: (81) (298) 36-9060.

The nucleotide sequences reported in this paper have been submitted to the DNA Data Bank of Japan with accession numbers D21839, D21840, D21841, D21842 and D21843.

 Published by Elsevier Science Publishers B.V.

were prehybridized (8 h) and hybridized (18 h) with 30% formamide, 100 mg/ml sonicated salmon sperm testes DNA, 5 × Denhardt's solution, 50 mM sodium phosphate (pH 6.5), 5 × SSC, and 0.2% SDS at 42°C. The filters were washed twice in 0.5 × SSC, 0.5% SDS at 37°C, then dried and exposed overnight to X-ray film with an intensifying screen. Positive clones were plaque-purified and classified into 10 distinct groups based on restriction endonuclease mapping and partial sequencing analyses.

2.2. Sequence analysis of cDNAs for MAP kinases

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

3. RESULTS

3.1. *Cloning and sequence analysis of five novel cDNAs for MAP kinases in* Arabidopsis

The 560-bp PCR-amplified fragment containing the conserved sequence of MAP kinase cDNAs was used as a probe for screening two different cDNA libraries to isolate cDNA clones for homologs of MAP kinases [15]. Thirty-nine positive cDNA clones were obtained from a total of 8.0×10^5 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 homologs: two were cATMPK1 and cATMPK2 [15]. We subcloned and sequenced the largest inserts of eight cDNAs encoding putative MAP kinases and determined the nucleotide sequences of five of the cDNAs (cATMPK3 to -7). Table I summarizes the nucleic acid lengths, deduced amino acid lengths, and molecular weights of the *ATMPK* cDNA clones. The deduced molecular weights range from 42,381 to 45,056 kDa.

3.2. *Primary Structure of the putative ATMPK proteins*

Fig. 1 compares the deduced amino acid sequences of the putative proteins encoded by *ATMPK1* to seven genes (named ATMPK1 to -7) with *Xenopus* MPK1 [22]. 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 [23]. The amino acid residues (Thr-Glu-Tyr), the TEY sequence [1–7], 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 [1–7]. To analyze the similarities between ATMPK proteins and other MAP kinases from various organisms [8–18,22,24–31], the deduced amino acid sequences of known MAP kinases were compared around kinase subdomains VI to VIII, including the TEY sequence

Fig. 1. Comparison of the deduced amino acid sequences of *Arabidopsis* ATMPKs (ATMPK1 to -7) and *Xenopus* MPK1 [22]. 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. [23]. 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 and rat ERK1.

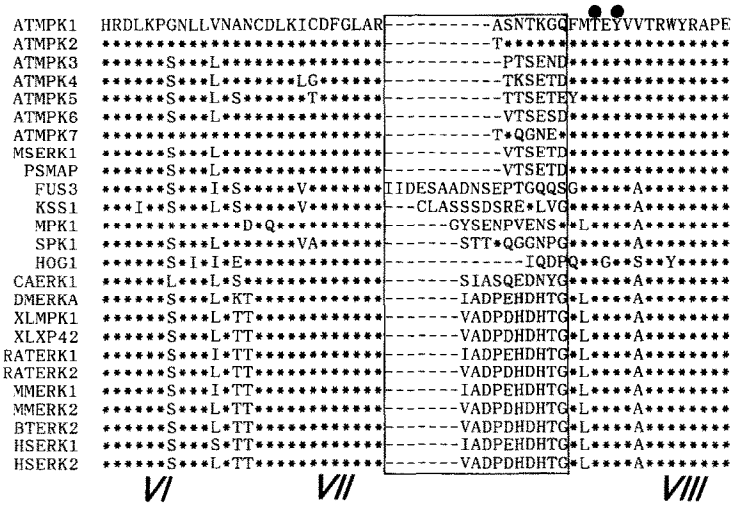

Fig. 2. Comparison of the deduced amino acid sequences around kinase subdomains VI to VIII between *Arabidopsis* ATMPKs and MAP kinases of other organisms. Asterisks represent identical amino acid residues and dashes indicate gaps introduced to maximize alignment. Roman numerals indicate the major conserved subdomains of protein kinases identified by Hanks et al [21]. Closed circles show residues that correspond to the sites of regulatory phosphorylation reported for *Xenopus* MPK1 and rat ERK1. The boxed sequences represent the divergent region between subdomains VII and VIII [14].

(Fig. 2). It has been reported that there are two regions where the amino acid sequences of MAP kinases are markedly divergent [14]. 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. 2), although the subdomains VI, VII, and VIII are highly conserved among all MAP kinases.

Table II summarizes sequence similarities among all the reported MAP kinases from various organisms. Amino acid identities between ATMPK proteins and other members of MAP kinases range from 43.8% to 55.8%. In contrast, amino acid identities among ATMPK proteins are between 53.2% and 88.7%. Among plant MAP kinase homologs, ATMPK1, ATMPK2, and ATMPK7 are closely related (79.7% to 88.7% identity). ATMPK3, ATMPK6, MsERK1 from alfalfa [16] and PsMAP from pea [17] are also closely related to one another (75.6% to 94.4% identity).

Table I

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 (kDa)
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

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

4. DISCUSSION

Genomic Southern blot analyses have indicated that the *Arabidopsis* genome has several related genes in addition to *ATMPK1* and *ATMPK2* [15]. We 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 seven different genes for MAP kinases and that these genes can be classified into subgroups based on sequence similarities.

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 -7 from *Arabidopsis*, MsERK1 from alfalfa [16,18], and PsMAP from pea [17]) apparently form a large group different from MAP kinases obtained from animals and yeasts (Table I). The reported plant MAP kinases seem to be classified into the following three subgroups: subgroup 1 = ATMPK1, ATMPK2, and ATMPK7; subgroup 2 = ATMPK3, ATMPK6, MsERK1, and PsMAP; subgroup 3 = ATMPK4 and ATMPK5. Amino acid identities among members of subgroups 1, 2, 3 may between 79.7% and 88.7%, 75.6% and 94.4%, or may be 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%)

Table II

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

	PK[b]	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25
1.	ATMPK1	*	88.7	58.0	55.0	54.4	55.9	79.9	56.1	55.6	46.7	46.2	46.2	49.3	45.2	46.2	48.3	49.4	49.4	49.0	49.3	49.0	49.3	49.3	48.8	49.3
2.	ATMPK2		*	55.7	53.2	54.5	54.6	79.7	55.2	54.6	44.2	45.3	46.2	49.3	43.8	46.8	48.5	50.1	50.1	48.9	50.0	49.4	50.0	50.0	49.3	50.0
3.	ATMPK3			*	65.1	63.8	75.6	56.8	76.9	76.7	49.3	48.2	46.8	50.1	48.1	48.3	49.6	49.7	49.1	50.0	49.7	50.0	49.7	49.4	49.7	49.7
4.	ATMPK4				*	74.5	71.4	53.6	72.2	72.2	49.9	47.3	48.3	51.9	47.6	50.1	48.0	51.2	50.9	49.3	51.2	49.3	51.2	51.5	48.8	51.2
5.	ATMPK5					*	67.6	53.4	67.5	67.3	49.7	49.8	48.1	55.8	47.7	49.6	49.9	50.3	49.7	50.6	50.3	50.6	50.3	50.6	50.3	50.3
6.	ATMPK6						*	52.7	88.4	87.1	50.4	50.1	48.4	51.3	47.7	50.3	51.6	51.4	51.2	50.6	51.3	50.6	51.3	51.3	50.0	51.3
7.	ATMPK7							*	53.9	53.1	46.2	44.5	45.3	49.3	45.2	47.2	47.2	47.1	47.1	47.1	46.7	47.1	46.7	46.7	46.8	46.7
8.	MsERK1								*	94.4	50.1	49.1	49.1	52.8	48.3	51.1	51.3	52.3	52.0	51.7	51.6	51.7	51.6	51.9	51.2	51.6
9.	PsMAP									*	49.3	48.2	48.6	51.3	47.2	46.3	50.4	51.7	51.4	51.2	51.0	51.2	51.0	51.3	50.6	51.0
10.	FUS3										*	57.5	45.3	55.4	44.3	58.1	49.4	49.9	49.9	50.6	50.4	50.6	50.4	50.1	49.7	50.4
11.	KSS1											*	50.2	61.2	49.4	56.7	53.3	52.0	52.3	49.7	52.0	49.7	52.0	52.0	52.0	52.0
12.	MPK1												*	48.5	40.6	50.3	48.3	49.1	48.8	48.5	49.4	49.9	49.4	49.7	49.4	49.4
13.	SPK1													*	47.7	58.0	52.0	55.3	55.3	52.8	56.3	51.5	56.3	56.3	54.8	56.3
14.	HOG1														*	46.6	48.2	48.1	47.5	46.9	48.1	46.6	48.1	48.1	46.6	47.8
15.	Ysaerk1															*	50.4	52.9	52.9	49.7	53.8	49.7	53.8	53.6	51.8	53.9
16.	DmERKA																*	80.4	79.1	76.5	80.4	76.8	80.4	79.9	77.4	79.7
17.	XLMPK1																	*	98.1	84.7	95.5	85.0	95.5	96.1	84.9	95.8
18.	XLp42																		*	83.8	95.2	84.1	95.2	95.8	84.1	95.5
19.	RatERK1																			*	85.6	99.5	86.0	86.0	96.8	85.7
20.	RatERK2																				*	85.9	100	99.2	85.6	99.4
21.	MmERK1																					*	85.9	86.2	96.8	86.0
22.	MmERK2																						*	99.2	85.6	99.4
23.	BtERK2																							*	86.0	99.7
24.	HsERK1																								*	85.7
25.	HsERK2																									*

[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: 1. and 2. *Arabidopsis* ATMPK1 and ATMPK2 [15]; 3, 4, 5, 6 and 7. *Arabidopsis* ATMPK3, ATMPK4, ATMPK5, ATMPK6 and ATMPK7 [this manuscript]; 8. alfalfa MsERK1 [16]; 9. pea PsMAP [17]; 10. budding yeast FUS3 [10]; 11. budding yeast KSS1 [11]; 12. budding yeast MPK1 [12]; 13. fission yeast SPK1 [14]; 14. budding yeast HOG1 [13]; 15. candida Yserk1 [24]; 16. *Drosophila* DmERKA [25]; 17. *Xenopus* MPK1 [22]; 18. *Xenopus* Xp42 [26]; 19. rat ERK1 [8]; 20. rat ERK2 [9]; 21. mouse MmERK1 [27]; 22. mouse MmERK2 [28]; 23. bovine BtERK2 [29]; 24. human HsERK1 [30]; 25. human HsERK2 [31].

and PsMAP (87.1%) not only in the catalytic domains but also in both the N- and C-terminal regions and in the divergent regions between subdomains VII and VIII [14]. This suggests that ATMPK6 is a counterpart of MsERK1 [16] and PsMAP [17]. An *S. pombe* MAP kinase, Spk1, 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 that show 55.3% identity with Spk1 [32]. 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. ATMPK5 may restore the sporulation ability of Spk1-disrupted *S. pombe* cells. Genomic Southern analysis of alfalfa indicated the existence of a number of MAP kinases, and PCR has produced at least 4 MAP kinases in alfalfa [18], which provides further evidence for multiple MAP kinases in plants.

It has been pointed out that in yeast cells different MAP kinases function in different signaling pathways, while in vertebrate cells at least one MAP kinase may function in various signal transduction pathways [1–7]. We identified at least 7 independent MAP kinases, ATMPKs, in plants and classified them into three subgroups based on comparisons between their amino acid sequences. Thus, ATMPKs may function in multiple signal transduction pathways.

MAP kinases have a divergent region between subdomains VII and VIII (Fig. 2). Since this region is located just before the TEY sequence, which is the phosphorylation site for the activation by MAPK kinases, it may play an important role in the interaction of MAP kinases with its direct upstream activator MAPK kinases. This suggests that there may be multiple MAPKK kinases, MAPK kinases, and MAP kinases in *Arabidopsis*. Recently, we identified multiple MAPKK kinases in *Arabidopsis* (Mizoguchi et al., unpublished data). Thus, plants may possess a multiple MAP kinase cascade (MAPKK kinase–MAPK kinase–MAP kinase), whereas animals do not.

Acknowledgements: This work was supported in part by the Special Coordination Fund of the Science and Technology Agency of the Japanese Government and by a Grant-In-Aid from the Ministry of Education, Science and Culture of Japan to K.S.. It was also supported by a Grant for 'Biodesign Research Programs' from Riken to N.H. K.Y.-S. was supported by a fellowship from the Science and Technology Agency of Japan. We appreciate the helpful discussions and encouragement provided by Profs. H. Harada and T. Fujii of the University of Tsukuba.

REFERENCES

[1] Sturgill, T.W. and Wu, J. (1991) Biochem. Biophys. Acta 1092, 350–357.
[2] Cobb, M.H., Boulton, T.G. and Robbins, D.J. (1991) Cell Regul. 2, 965–978.
[3] Thomas, G. (1992) Cell 68, 1–4.
[4] Pelech, S.L. and Sanghera, J.S. (1992) Science 257, 1355–1356.
[5] Nishida, E. and Gotoh, Y. (1992) Int. Rev. Cytol. 138, 211–238.
[6] Pelech, S.L. and Sanghera, J.S. (1992) Trends Biochem. Sci. 17, 233–238.
[7] Nishida, E. and Gotoh, Y. (1993) Trends Biochem. Sci. 18, 128–131.
[8] Boulton, T.G., Yancopoulos, G.D., Gregory, J.S., Slaughter, C., Moomaw, C., Hsu, J. and Cobb, M.H. (1990) Science 249, 64–67.
[9] 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) Cell 65, 663–675.
[10] Elion, E.A., Grisafi, P.L. and Fink, G.R. (1990) Cell 60, 649–664.
[11] Courchesne, W.E., Kunisawa, R. and Thorner, J. (1989) Cell 58, 1107–1119.
[12] Torres, L., Martín, H., García-Saez, M.I., Arroyo, J., Molina, M., Sánchez, M. and Nombela, C. (1991) Mol. Microbiol. 5, 2845–2854.
[13] Brewster, J.L., Valoir, T.D., Dwyer, N.D., Winter, E. and Gustin, M.C. (1993) Science 259, 1760–1763.
[14] Toda, T., Shimanuki, M. and Yanagida, M. (1991) Genes Dev. 5, 60–73.
[15] Mizoguchi, T., Gotoh, Y., Nishida, E., Yamaguchi-Shinozaki, K., Hayashida, N., Iwasaki, T., Kamada, H. and Shinozaki, K. (1994) Plant J. (in press).
[16] Duerr, B., Gawienowski, M., Ropp, T. and Jacobs, T. (1993) Plant Cell 5, 87–96.
[17] Stafstrom, J.P., Altschuler, M. and Anderson, D.H. (1993) Plant. Mol. Biol. 22, 83–90.
[18] Jonak, C., Páy, A., Bögre, L., Hirt, H. and Heberle-Bors, E. (1993) Plant J. 3(4), 611–617.
[19] Yamaguchi-Shinozaki, K., Koizumi, M., Urao, S. and Shinozaki, K. (1992) Plant Cell Physiol. 33(3), 217–224.
[20] Kiyosue, T., Yamaguchi-Shinozaki, K. and Shinozaki, K. (1993) XV International Botanical Congress, 6356.
[21] Maniatis, T., Fitsch, E.F. and Sambrook, J. (1982) Molecular Cloning: a Laboratory Manual, Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY.
[22] Gotoh, Y., Moriyama, K., Matsuda, S., Okumura, E., Kishimoto, T., Kawasaki, H., Suzuki, K., Yahara, I., Sakai, H. and Nishida, E. (1991) EMBO J. 10, 2661–2668.
[23] Hanks, S.K., Quinn, A.M. and Hunter, T. (1988) Science 241, 42–52.
[24] Whiteway, M., Dignard, D. and Thomas, D.Y. (1992) Proc. Natl. Acad. Sci. USA 89, 9410–9414.
[25] Biggs III, W.H. and Zipursky, S.L. (1992) Proc. Natl. Acad. Sci. USA 89, 6295–6299.
[26] Posada, J., Sanghera, J., Pelech, S., Aebersold, R. and Cooper, J.A. (1991) Mol. Cell. Biol. 11, 2517–2528.
[27] Tanner, B. and Mueckler, M. (1993) Biochim. Biophys. Acta 1171, 319–320.
[28] Her, J.-H., Wu, J., Rall, T.B., Sturgill, T.W. and Weber, M.J. (1991) Nucleic Acids Res. 19, 3743.
[29] GenBank, accession number: Z14089.
[30] Charest, D.L., Jilik, F., Harder, K., Pelech, S.L. and Mordret, G. (1993) Mol. Cell. Biol. 13, 4679–4690.
[31] Owaki, H., Marker, R., Boulton, T.G., Cobb, M.H. and Geppert, T.D. (1992) Biochem. Biophys. Res. Commun. 182, 1416–1422.
[32] Gotoh, Y., Nishida, E., Shimanuki, M., Toda, T., Imai, Y. and Yamamoto, M. (1993) Mol. Cell. Biol. 13, 6427–6434.