

Isolation and characterization of novel Ser/Thr protein kinase genes from higher plants.

Takehiro IMAI

(Received September 12, 2005)
(Revised December 13, 2005)

Abstract

According to the similarity of amino acid sequences in the catalytic domains, eukaryotic protein kinases have been classified into the five main groups: 'AGC', 'CaMK', 'CMGC', 'PTK', and other group. The AGC group, represented by the cyclic nucleotide-dependent kinases (PKA and PKG), the calcium-phospholipid-dependent kinases (PKC) and the ribosomal S6 protein kinase, are poorly characterized in plants except a few cases. In order to gain a better understanding of plant protein kinases in the 'AGC' group, novel Ser/Thr protein kinases genes were isolated from radish (*R. sativus*), *Arabidopsis* (*A. thaliana*) and common wheat (*T. aestivum*), and their primary structures were characterized by molecular biological methods.

In the first chapter, five cDNAs (*RsNdr1*, *2a*, *2b*, *3* and *4*) were isolated from radish. The deduced amino acid sequences of *RsNdr1* and *RsNdr2a/b* contained all 12 conserved catalytic subdomains that are characteristic of the eukaryotic Ser/Thr protein kinases, suggesting that the translation products of the cDNAs are functional. A cell lysate from *E. coli* overexpressing *RsNdr1* fusion protein had the protein kinase activity. *RsNdrs* showed a significant homology to the fungal *cot-1* like and animal *Ndr* protein kinases, and it was suggested that *RsNdrs* consist of a small multi-gene family in radish. In the second chapter, *RsNdr* homologs were surveyed in a completely-sequenced *Arabidopsis* genome, to characterize the genes as a family. Eight homologs (*AtNdr*) were identified in *A. thaliana*, and five full-length cDNAs out of eight (*AtNdr1*, *3~6*) genes were obtained. DNA sequencing of cDNAs indicated that these protein kinases have mRNA variants produced by alternate splicing, showing a possibility that the expression and activity might be regulated at the splicing level. In the last chapter, three cDNA clones (wheat *Ndr*) encoding a Ser/Thr protein kinase were isolated from common wheat by data mining of the wheat EST database KOMUGI, and their primary structure were determined by RT-PCR and the 5'-RACE method.

Molecular phylogenetical analysis showed that *RsNdr*, *AtNdr* and wheat *Ndr* are clearly distinguished from those of other 'AGC' protein kinases. I concluded that these protein kinases are members of a new protein kinase superfamily that universally exist across the plant, animal and fungal kingdoms.

Key Words: Ser/Thr protein kinase, *Ndr*, radish, *Arabidopsis*, wheat

General introduction

Reversible protein phosphorylation plays a crucial role in the regulation of many fundamental cellular processes including cell division and differentiation, growth, and development (Hanks *et al.* 1988, Berridge 1993). In higher plants, protein kinases are used in the signal transduction cascades in response to a variety of signals such as light, water, temperature and hormones. The phosphorylation status of a specific target protein is regulated by the opposing actions of protein kinases and phosphatases. The “protein kinases” are a group of enzymes that catalyze transfer of a phosphate group from an ATP molecule to a hydroxyl group on a serine, threonine or tyrosine side chain of a protein, though protein kinases related to the prokaryotic His kinase family have been identified in eukaryotes (Chang *et al.* 1993). Since the activity of the protein is often changed by the status of phosphorylation/dephosphorylation, protein kinases in a network of the cells are compared with on-off switches in computer chips. Now it is estimated that in eukaryote 1-3 per cent of functional genes encode protein kinases.

After the pioneering work of Lawton *et al.* (1989) in which the first plant protein kinase, PVPK-1, was documented, several protein kinases have been isolated, and in some cases, their crucial roles in signal transduction cascades have been clarified (e.g. Meskiene and Hirt 2000, Innes 2001, Asai *et al.* 2002).

Eukaryotic protein kinases contain 12 highly conserved regions that are referred to as subdomains (Hanks *et al.* 1988, Hanks and Quinn 1991). Invariant or nearly invariant residues in the subdomains are important for maintaining three-dimensional structure of the protein and their catalytic activity. Using a phylogenetic analysis of amino acid sequences in the subdomains, Hanks and Hunter (1995) first classified the eukaryotic protein kinases into the five main groups: ‘AGC’, ‘CaMK’, ‘CMGC’, ‘PTK’ and ‘other’ group. After that, Stone and Walker (1995) adjusted this classification for protein kinases, and pointed out the similarities and the differences between plant protein kinases and other eukaryotic protein kinases. Among plant protein kinases, those belonging to the AGC group are poorly understood, except for a few cases (Zhang *et al.* 1994a, b). The AGC group is represented by cyclic nucleotide-dependent kinases (PKA and PKG), the calcium-phospholipid-dependent kinases (PKC) and the ribosomal S6 protein kinases. There is a common feature in these kinases; they are regulated by second messengers (i.e. cAMP, cGMP, diacylglycerol and Ca^{2+}) with the exception of the ribosomal S6 protein kinases. Since the function of cAMP in plants has been disputed for many years, and the presence of authentic PKA in plants has been doubted, the gene for PKA has never been cloned from any plant species. Moreover, the plant uses Ca^{2+} signaling in a

different way from that of mammals, i.e. the major calcium-dependent kinases cloned at present are CDPKs (calmodulin-like domain protein kinases), which are structurally different from PKC and unique to the plant. The CDPKs belong to CaMK group. Considering the importance of cAMP and Ca^{2+} as ubiquitous messengers, the absence of PKA and PKC, and the absence of signal transduction cascades in which both protein kinases are involved, is somewhat mysterious, leading me to search for protein kinases in AGC group from higher plants. In this paper, I will describe the results of experiments to isolate cDNAs encoding protein kinase in AGC group from higher plants. In chapter 1, the first isolation and characterization of cDNAs encoding novel protein kinases in AGC group from higher plant are documented using radish (*Raphanus sativus*) as the material plants. It shows that isolated protein kinases are homologous to animal *Ndr* and fungal *cot-1* like protein kinases that influenced cell division and/or morphology. In chapter 2, the nature of a multi-gene family of protein kinase is discussed using *A. thaliana*. In the last chapter, the results of isolation of protein kinase in monocots using wheat EST database are presented.

Chapter 1. Cloning and characterization of novel Ser/Thr protein kinases in AGC group from radish.

1.1 Introduction

Radish is an annual vegetable in the *Brassicaceae* and widely cultivated in Japan. Besides the long history of use in the field of plant physiology, radish has become a suitable source for isolating novel genes due to the tremendous information on the model plant *Arabidopsis thaliana*. In addition, studying radishes other than *Arabidopsis* will give an advantage for a variety of biochemical work. In this chapter, in order to gain a better understanding of plant protein kinases in the 'AGC' group, I have cloned and characterized five cDNAs encoding novel protein kinases, *RsNdrs* (*R. sativus Ndr*), and expressed two of them in *E. coli* to show their enzymatic activities. The results show that at least one *RsNdr* is functional and that *RsNdrs* are a new family of protein kinases in radish.

1.2 Materials and methods

Plant materials

A European cultivar of radish (*Raphanus sativus* L. cv. 'Comet') was used as the material plant. The plants were grown in a greenhouse under normal day-light/temperature conditions to maturity.

Reverse transcription-polymerase chain reaction (RT-PCR)

Total RNA was isolated from 100 mg of young leaves using an RNeasy plant mini kit (Qiagen). cDNA was reverse-transcribed from 5 µg of total RNA using Superscript preamplification system for first strand cDNA synthesis (GIBCO BRL). PCR amplification was carried out on tenfold dilution of RNaseH-treated cDNA mixture with the four-oligonucleotide primers (forward: # 1; 5'-AGAGGACAGGTTGAGCATGT-3', # 2; 5'-ATATGATGACCTTGCTCATGAG-3', reverse: # 3; 5'-GAGGATACCCAACGAGCATTTC-3', # 4; 5'-CATGCCATATCCTTTCTTCAGC-3'). The oligonucleotide primers were designed from the nucleotide sequences of Arabidopsis ESTs (N64988, nucleotide positions 134~153 and 270~291, and N65509, nucleotide positions 177~198 and 119~140), referring to the sequences for the catalytic domains of PKAs from budding yeast (GenBank accessions M17072-04, M17224 and Y00694) and fission yeast (U08622), and a cDNA for a putative protein kinase from tobacco (=NTPKTL7, X71057). PCR was performed in a System 2400 programmable incubator (PE Biosystems) using ExTaq DNA polymerase (Takara) with the following setting; initial denaturation (94°C, 3 min), 30 cycles of denaturation (94°C, 30 sec), annealing (55°C, 30 sec) and extension (72°C, 1 min). The PCR products were cloned into a plasmid vector, pGEM-T Easy (Promega), and their nucleotide sequences were determined as described below. Rapid amplification of cDNA ends (RACE) was performed with various radish-specific primers designed from the DNA sequences of initial PCR products. Primers specific to the 5' and 3' end of the radish cDNAs were designed from the sequences of RACE products, and then putative full-length cDNAs were amplified by RT-PCR. Complete nucleotide sequences of the putative full-length cDNAs were determined after cloning them into the pGEM-T Easy as described.

DNA sequencing

Nucleotide sequence of the selected plasmid DNA was determined by cycle sequencing procedures using Thermo Sequence fluorescent labeled primer cycle sequencing kit with 7-deaza-dGTP (AP Biotech) and CEQ2000 Dye Terminator Cycle Sequencing with Quick Start kit (BECKMAN COULTER). DNA sequencing was carried out using ALFexpress DNA sequencer (AP Biotech) and CEQ2000 DNA Analysis system (BECKMAN COULTER) following the instructions of the manufacturer. The nucleotide and amino acid sequences were analyzed using the software GeneWorks (Oxford Molecular Group), clustal W at Genomenet (<http://clustalw.genome.ad.jp/>), PAUP* version 4.0 (Swofford, 2002), and the BLAST program at NCBI (<http://www.ncbi.nlm.nih.gov/BLAST/>).

Northern blot analysis and RT-PCR

For Northern blot analysis, total RNA was prepared from the mature and young (six-week old)

leaves, stems, roots and flower buds of the material plants using Sepasol-RNAI (Nacalai Tesque). About 50 µg total RNA was electrophoresed on a 1.2% agarose gel containing 5.0% (v/v) formaldehyde and was blotted onto a nylon membrane, Biodyne Plus (Pall). The RNA was hybridized with Dig-labeled RNA probe in a Dig Easy Hyb solution (Roche) at 65°C overnight. The Dig-labeled anti-sense RNA probes were prepared by *in vitro* transcription on linearized plasmid containing a full-length cDNA for *RsNdr1* and *RsNdr2*. Post hybridization washes were performed twice at low stringency (2xSSC/0.1%SDS, at room temperature) for 5 min and twice at high stringency (0.1xSSC/0.1%SDS, at 65°C) for 15 min. After post hybridization washes, signals were detected by the Dig chemiluminescence detection procedures following the instructions of the manufacturer (Roche). An RNA probe, which was transcribed from a cDNA clone containing radish cytosolic glyceraldehydes-3-phosphate dehydrogenase (*GapC*), was used as a control in the Northern hybridization experiment.

To investigate accumulation of a small quantity of mRNAs, 25 and 30 cycles of RT-PCR were performed with the primer pairs specific to each of *RsNdrs*. The same RNA as that used in the Northern blot analysis was employed as a template for RT-PCR, except that RNA was treated with RNase-free DNase to remove residual amounts of contaminated DNA. *GapC* was used as a control for RT-PCR experiment.

Expression of RsNdrs in E. coli

A 1.7-kb *SaII* cDNA fragment of *RsNdr1* and a 1.7-kb *BamHI/SaII* cDNA fragment of *RsNdr2b* were cloned into *SaII*, *BamHI/SaII* sites, respectively, of an expression vector pMAL-c2x (NEB). These restriction fragments were obtained by PCR with KOD DNA polymerase (Toyobo) and specific primers in which appropriate restriction sites were incorporated. A fusion constructs were named pMAL-RsNdr1, pMAL-RsNdr2b, respectively. A fusion construct with a kinase-dead RsNdr1/K77A was constructed by PCR-based site-directed mutagenesis on the pMAL-RsNdr1 template using Pfu DNA polymerase (Stratagene) and primers (#375; 5'-GGCAAACCTGAAAAAACTG-3', #374; 5'-ATGGCATAAACCTCGCCTGT-3'), following the instructions. For expression analysis, an overnight culture of *E. coli* harboring fusion constructs was transferred to a flask containing 25 ml of SOB, and *E. coli* was cultured at 37°C with vigorous shaking. When an OD₆₀₀ of the culture reached 0.4~0.6, isopropyl-β-thiogalactopyranoside (IPTG) was added to a final concentration of 1 mM. Aliquots of 1 ml of the culture were taken at hourly intervals, and the cells were harvested by centrifugation. The cells were resuspended in 100 µl of 20 mM phosphate buffer (pH 7.0) and disrupted by repeated freeze/thaw procedures. After centrifugation at 14,000 x g for 20 min at 4°C, the supernatant (hereafter termed 'soluble fraction') was mixed with an equal volume of SDS sample

buffer (0.125M Tris-HCl pH 6.8, 4% SDS, 20% glycerol, 2% β -mercaptoethanol and bromophenol blue). The remaining pellets ('insoluble fraction') were also resuspended in 100 μ l of SDS sample buffer. Proteins in the soluble (20 μ l) and insoluble fraction (10 μ l) were separated by 7.5% SDS-PAGE (Laemmli, 1970), and they were stained with CBB. Proteins were transferred onto a nitrocellulose membrane, Hybond ECL (AP Biotech), using Trans-Blot SD (BioRad). Recombinant fusion proteins were detected by a 1: 10000 dilution of anti-MBP serum (NEB) and HRP-linked anti-rabbit antibodies (Wako) using ECL Western blotting detection reagents (AP Biotech).

Protein kinase assay

Protein kinase activity was measured using *E. coli* lysate as an enzymatic source, and histone, myelin basic protein and casein (Sigma) as a substrate. The 100 μ l reaction contains 1 or 2 μ g of proteins from *E. coli* cell lysate, 20 mM Tris-HCl pH 7.4, 10 mM $MgCl_2$, 1 mM DTT, 100 mM [γ - ^{32}P]ATP (10Ci/ml), and 50 μ g each of exogenous substrate proteins. The effects of activating factors on the activity were measured by adding either Ca^{2+} (1.5 mM at a final concentration) or cAMP (0.01 mM, at a final concentration), or both. After 2 hours incubation at 30°C, reaction was terminated by adding 10 μ l of glacial acetic acid, and then 40 μ l of a reaction mixture was taken and spotted onto P81 phosphocellulose papers (Whatman). The papers were washed in 75 mM phosphoric acid five times for 5 min each and once in distilled water. The papers were dried and the amount of incorporated ^{32}P was measured with a liquid scintillation counter, Packard TRI-CARB 2900TR. As a positive control for the entire assay, cell lysate prepared from mammalian COS7 cells was used.

1.3 Results

Primary structure of RsNdrs

RT-PCR was performed on the radish total RNA prepared from the mature leaves with the four primers (#1-#4). They were designed to amplify a segment between subdomains III and IX in catalytic domain of PKAs from various organisms. The size of RT-PCR products, which were amplified from the cDNA of leaves, was about 450-600bp with all possible primer combinations, slightly larger than that expected for yeast PKAs (data not shown). DNA sequencing of the cloned PCR products obtained by primer pairs #1/#3 and #2/#4 showed that the former product consisted of two different cDNAs of 583bp and 601bp, whereas the latter product was a single cDNA of 486bp. DNA sequencing of the 3'-RACE product with primer #2 also revealed a new cDNA fragment different from the aforementioned cDNAs. The deduced amino acid sequences of each of four cDNA clones contained highly conserved motifs that constitute a catalytic domain of eukaryotic Ser/Thr protein kinases. Here, putative protein kinase revealed by the 3'-RACE-PCR is desig-

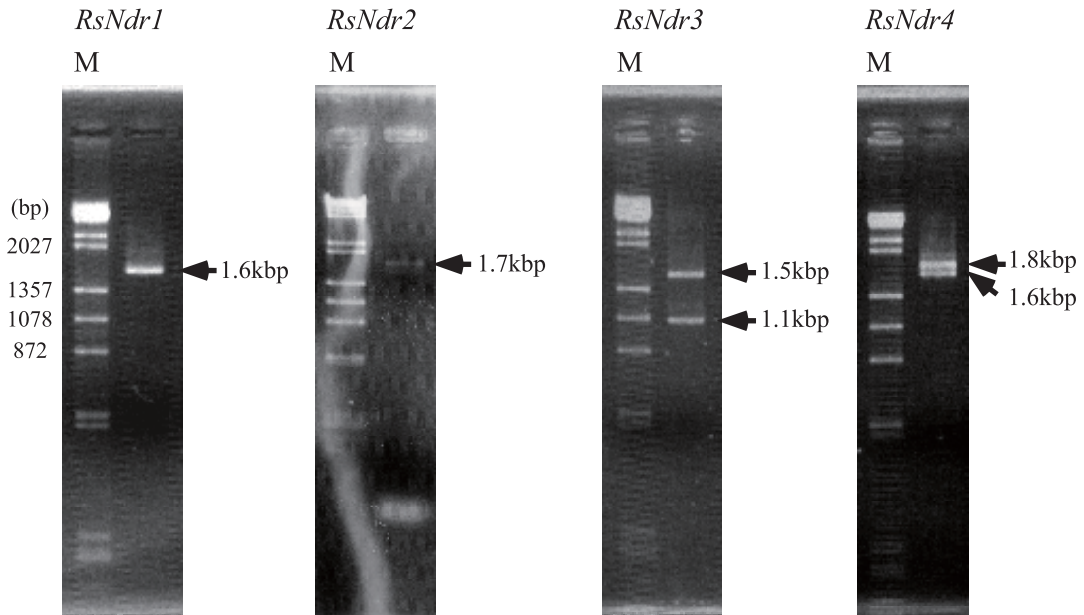


Fig. 1. A full-length *RsNdr* cDNA fragment amplified by RT-PCR. cDNA was reverse-transcribed from radish's leaves. PCR was performed using each of the 5'- and 3'- end specific primers. M: A mixture of the λ /HindIII and ϕ X174 HaeIII digests used as molecular size standards.

nated as *RsNdr1*, whereas the other three protein kinases represented by RT-PCR products of 583bp, 601bp and 486bp are named *RsNdr2*, *RsNdr3* and *RsNdr4*, respectively.

After a series of RACE experiments, primers specific to the 5'- and 3'- end of *RsNdr* cDNAs were synthesized, and a putative full-length cDNA was amplified by RT-PCR. A single RT-PCR product was amplified for each of *RsNdr1* (ca. 1.6-kb) and *RsNdr2* (ca. 1.7-kb), whereas two amplification products were observed for each of *RsNdr3* (ca. 1.5-kb and ca. 1.1-kb) and *RsNdr4* (ca. 1.8-kb and ca. 1.6-kb) (Fig. 1). The larger one of the two products for *RsNdr3* is named *RsNdr3L*, whereas the smaller *RsNdr3S*. Similarly the larger and the smaller ones for *RsNdr4* are named *RsNdr4L* and *RsNdr4S*, respectively.

DNA sequencing of the 1.6-kb product for *RsNdr1* showed that cDNA (1643bp) contained a single open reading frame (ORF) of 1386bp encoding a protein of 461 amino acids with a calculated molecular mass of 53kDa (Fig. 2). DNA sequencing of the 1.8-kb product for *RsNdr2* revealed that the product contains two cDNAs of 1730bp (named as *RsNdr2a*) and 1718bp (*RsNdr2b*). *RsNdr2a* contains a single ORF of 1626bp encoding a protein of 541 amino acids with a calculated molecular mass of 63kDa, whereas *RsNdr2b* contains an ORF of 1557bp encoding for 518 amino acids of 60kDa (Fig. 2). Although the overall DNA sequences for *RsNdr2a* and *RsNdr2b* are quite similar, they are different from each other in the number of GA repeats (ten vs. six, in *RsNdr2a* and *RsNdr2b*,

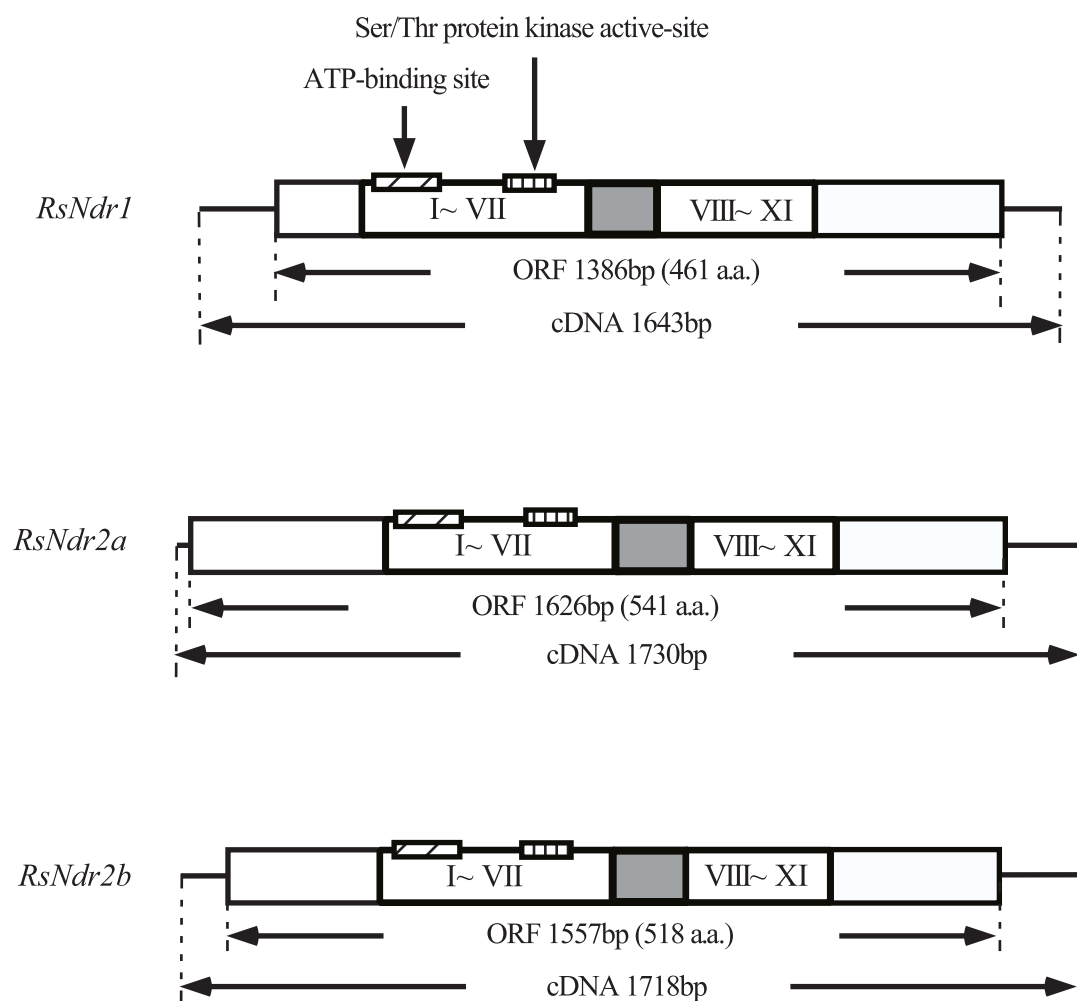


Fig. 2. The architecture of cDNAs for *RsNdr1* and *RsNdr2a/2b*. The open reading frame is indicated by a box, whereas 5'- and 3'-UTRs are shown by a horizontal line. Conserved catalytic subdomains are shown by Roman numerals. The ATP-binding site, Ser/Thr protein kinase active-site and an extra amino acid sequence between catalytic domains VII and VIII are shown by an obliquely striped, vertically striped and shaded boxes, respectively.

respectively) at the 5'-end, the twenty-eight synonymous and four nonsynonymous changes in the two ORFs, and the five base changes in the 3'-UTR (data not shown). There are three insertions/deletions in the 3'-UTR of the two cDNAs. Deduced amino acid sequences of *RsNdr1*, *RsNdr2a* and *RsNdr2b* contain all 12 highly conserved subdomains of the eukaryotic Ser/Thr protein kinase, including the ATP-binding site (in subdomains I-II) and Ser/Thr protein kinase active-site (in subdomain VI) (Fig. 2).

DNA sequencing of the 1.1-kb product for *RsNdr3S* showed that cDNA (1076bp) contained a single ORF of 834bp encoding a protein of 278 amino acid sequences (data not shown). The deduced

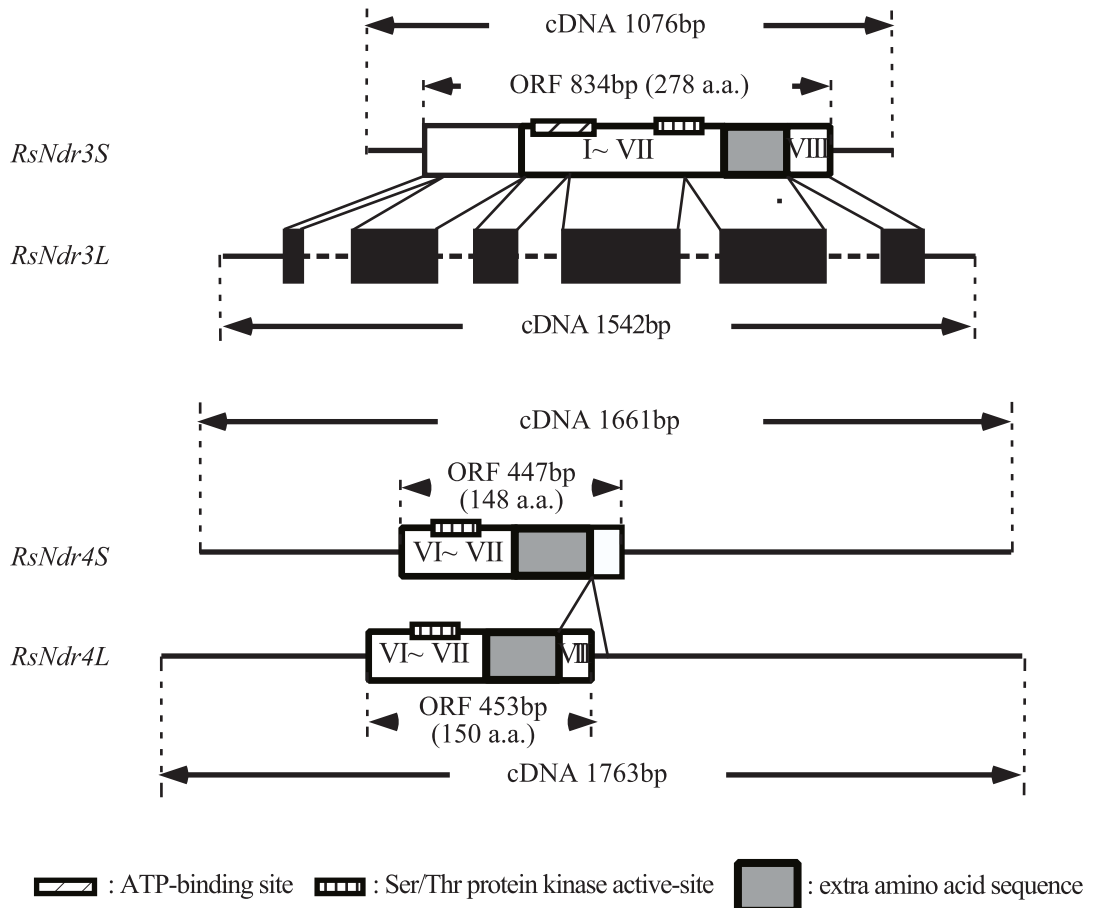


Fig. 3. The architecture of cDNAs for *RsNdr3S*, *RsNdr3L*, *RsNdr4S* and *RsNdr4L*. Notes and comments are the same as those in Fig. 2. Sequences identified with ORF sequences of *RsNdr3S* and intron-like sequences in *RsNdr3L* are shown by black boxes and dotted line, respectively.

amino acid sequence of *RsNdr3S* included ATP-binding site (in subdomain I-II), but it lacked subdomains IX~XII (Fig. 3). DNA sequencing of the 1.5-kb product for *RsNdr3L* showed that cDNA (1542bp) did not contain any long ORFs, because five intron-like sequences were present in the cDNA. If the five intron-like sequences were removed, *RsNdr3L* is identical to the *RsNdr3S* (Fig. 3). DNA sequencing of the 1.8-kb and 1.6-kb products for *RsNdr4L* (1763bp) and *RsNdr4S* (1661bp) showed that the former contains a single ORF of 453bp encoding for 150 amino acid sequences and the latter a single ORF of 447bp encoding for 148 amino acid sequences. Deduced amino acid sequences of both *RsNdr4L* and *RsNdr4S* lack the ATP-binding site, but the former contains subdomains VI-VIII and the latter has subdomains VI-VII (Fig. 3).

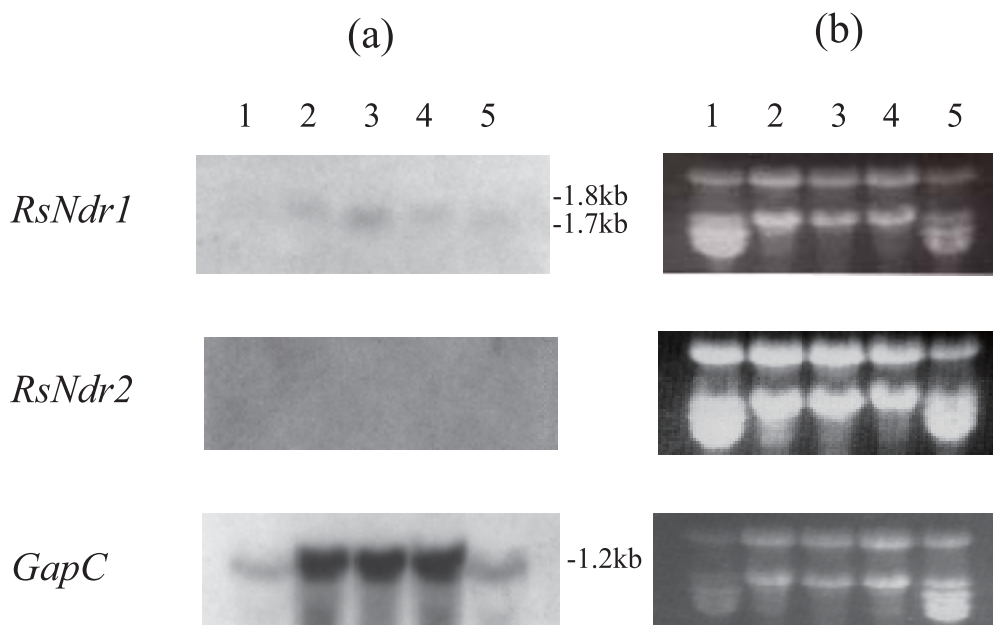


Fig. 4. RNA accumulation studies of *RsNdrs* by Northern blot. Total RNA prepared from leaves (1), stems (2), roots (3), flower buds (4) and young leaves (5) was hybridized with Dig-labeled *RsNdr1*, *RsNdr2* and *GapC* RNA probes. The left part (a) of each figure shows the hybridization signals, whereas the right (b) shows the EtBr-stained rRNAs.

RNA accumulation of *RsNdrs* in a plant

To study the steady state level of RNA accumulation of *RsNdrs*, Northern blot analysis was conducted with the *RsNdr1* and *RsNdr2b* probes on the total RNAs prepared from leaves, stems, roots and flower buds (Fig. 4). Although a clear transcript of 1.2-kb was detected with a control probe (*GapC*), very weak signals, either a 1.8-kb transcript in stems and flower buds or a 1.7-kb transcript in roots and young leaves, were observed with the *RsNdr1* probe (Fig. 4). No hybridization signal was detected with *RsNdr2b* probes under our experimental conditions (Fig. 4). For *RsNdr3* and *RsNdr4*, RNA accumulation was not tested by Northern blot analysis.

RNA accumulation of *RsNdrs* was also checked by RT-PCR on the same RNA as that used in the Northern blot analysis (Fig. 5). Using a *GapC*-specific primer pair, a large amount of RT-PCR product with the expected size was amplified, even after 25 cycles of the reaction. With a primer pair specific to the *RsNdr1*, a single PCR product with the expected size was amplified from all RNA preparations after 25 cycles of reactions, though unspecific minor fragments began to appear after 30 cycles of reactions in all but flower bud RNA. With a primer pair common to *RsNdr2a* and *RsNdr2b*, a single PCR product of the expected size was detected clearly after 30 cycles of the reactions. With a primer pair specific to *RsNdr3*, a single PCR product for *RsNdr3S* was detected

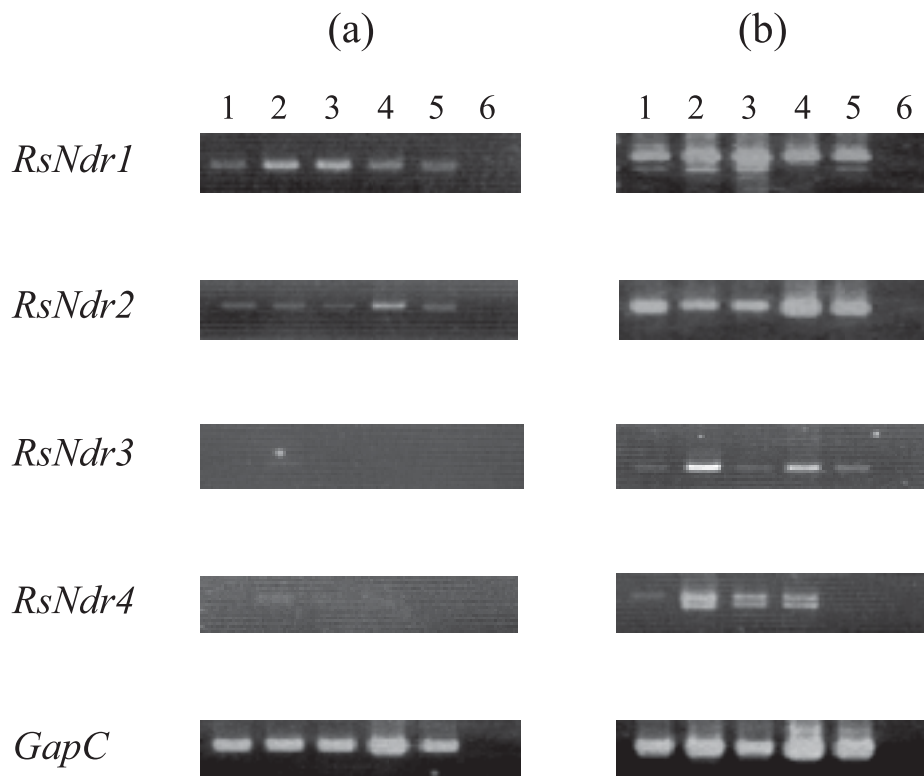


Fig. 5. RNA accumulation studies of *RsNdrs* by RT-PCR. The RNA from leaves (1), stems (2), roots (3), flower buds (4) and young leaves (5) was employed as a template for RT-PCR. PCR was performed using *RsNdrs* and *GapC* specific primers. RT-PCR products are shown, together with the number of cycles (25 (a) or 30 (b)) employed. Lane 6 indicates a negative control of the RT-PCR (without a template).

after 30 cycles of the reactions in all RNA preparations, though the amount of amplification product was low. With a primer pair specific to *RsNdr4*, two RT-PCR products, corresponding to the both *RsNdr4S* and *RsNdr4L*, were clearly detected in RNA from stems, roots and buds after 30 cycles of PCR. In the RNA from leaves, only a product for *RsNdr4L* was amplified.

Detection of the protein kinase activity of RsNdrs expressed in E. coli

To examine whether *RsNdr* proteins have protein kinase activity *in vitro*, each putative full-length cDNA for *RsNdr1* and *RsNdr2b* was cloned into an expression vector pMAL-c2x and expressed in *E. coli* as a fusion protein with a tag, maltose binding protein (MBP). A protein of about 100kDa was overproduced in *E. coli* harboring pMAL-RsNdr1, and mainly accumulated in an insoluble fraction. Using anti-MBP serum, the 100kDa protein was immunologically detected in an insoluble fraction together with several smaller proteins. In a soluble fraction, in contrast, only the 100kDa protein was detected by the same anti-MBP serum. In the total protein of *E. coli* harboring

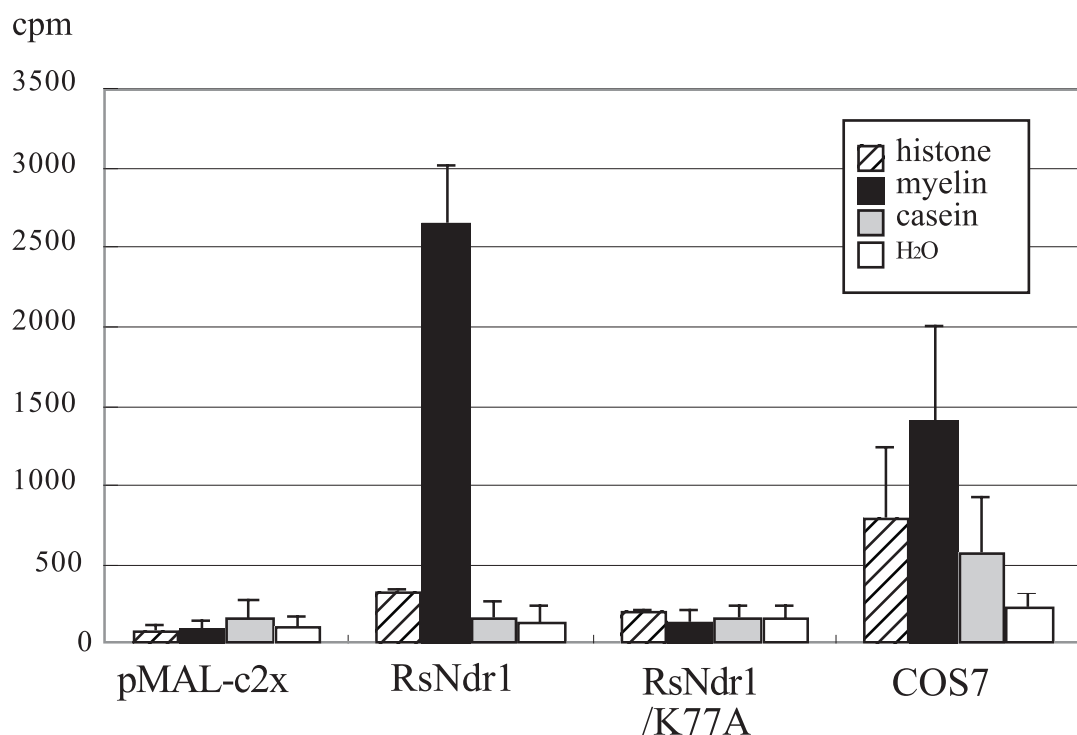


Fig. 6. Phosphorylation activity of the *RsNdr1* fusion protein. The lysate of *E. coli* cells, which overexpress MBP (pMAL-c2x), *RsNdr1* or *RsNdr1/K77A* fusion protein ($2\mu\text{g}$), was incubated in the kinase buffer containing $[\gamma\text{-}^{32}\text{P}]$ ATP, 0.01 mM cAMP, 1.5 mM CaCl_2 , and each protein substrate. The lysate from mammalian COS7 cells was used as a positive control of the experiment. Bars represent mean \pm standard deviation of duplicate determinations.

pMAL-*RsNdr2b*, the overexpressed 110kDa protein was detected only in an insoluble fraction. Anti-MBP serum detected the 110kDa protein in both insoluble and soluble fractions, but the signals were weak, especially in a soluble fraction. In a soluble fraction, several smaller proteins (less than 50kDa) than MBP were detected.

When the cell lysate expressing *RsNdr1* fusion protein was incubated with the conventional protein substrates, the highest activity was observed on the myelin basic protein, and a weak activity was detected on the histone (Fig. 6). When the activity was measured with the kinase-dead *RsNdr1/K77A*, the activity was lost on both the myelin basic protein and the histone (Fig. 6). Protein kinase activity was not detected for *RsNdr2* fusion protein in any of three substrates tested (data not shown).

1.4 Discussion

Isolation of cDNAs encoding novel protein kinases in AGC group from radish

In this chapter, I have cloned and sequenced five putative full-length cDNAs (*RsNdr1*, *RsNdr2a*, *RsNdr2b*, *RsNdr3* and *RsNdr4*) encoding novel protein kinases from radish. Deduced amino acid sequences of three cDNAs (*RsNdr1*, *RsNdr2a* and *RsNdr2b*) contain all 12 conserved subdomains necessary for the eukaryotic Ser/Thr protein kinases, suggesting that the translation products of these cDNAs are functional. Although the difference between *RsNdr2a* and *RsNdr2b* is so small that I can not assign these two as separate genes (i.e. their relationship might be allelic), the sequencing and Southern blot analysis showed that at least two different but related kinases, *RsNdr1* and *RsNdr2* are present in radish. BLAST homology search using deduced amino acid sequences of *RsNdrs* hit several protein kinases belonging to the AGC group, as expected from my cloning strategy. The AGC group of protein kinases was first defined by Hanks and Hunter (1995), and is represented by PKA, PKG, PKC and ribosomal S6 kinase. Later, the group was divided into nine subgroups based on sequence similarities (Smith *et al.* 1997). To reveal the relative relationship of *RsNdrs* among protein kinases in the AGC group, molecular phylogenetical analysis was carried out. Fig. 7 shows a phylogenetic tree for selected 17 protein kinases from various organisms. The tree was reconstructed by NJ method after aligning amino acid sequences for catalytic domains of each protein kinase with the computer program Clustal W. The tree clearly shows that the *RsNdrs* are more closely related to the protein kinases in a particular subgroup of the 'AGC' than to the authentic 'AGC' protein kinases, such as PKA, PKC or ribosomal S6 kinase. This subgroup consists of the protein kinases from fungi, i.e. *cot-1* (Yarden *et al.* 1992), *TB3* (Buhr *et al.* 1996), *orb6* (Verde *et al.* 1998), *cbk1* (Racki *et al.* 2000, Bidlingmaier *et al.* 2001) and *ukc1* (Durrenberger and Kronstad 1999) of *Neurospora crassa*, *Colletotrichum trifolii*, *S. pombe*, *S. cerevisiae* and *Ustilago maydis*, respectively, and those referred to as *Ndr* (named after nuclear dbf2 related) identified in human, *D. melanogaster* and *C. elegans* (Millward *et al.* 1995, Geng *et al.* 2000, Zallen *et al.* 2000). For the catalytic domain, the deduced amino acid sequences of *RsNdr1* and *RsNdr2* show about 50% identity to those of the aforementioned protein kinases, though the organisms from which the protein kinases were isolated diverge at the kingdom level (Fig. 7). In addition to the sequence similarity, *RsNdrs* and these protein kinases share an unusual structural feature; they have an insert between subdomains VII and VIII of the kinase domain. The significance of an insert is largely unknown except for animal *Ndr1*, where the insert is shown to be used as a nuclear localization signal (NLS). The transient expression of an *RsNdr1* (amino acid 1-440) construct fused with GFP shows that the insert of *RsNdr1* may function as NLS (data not shown). A similar insert is found in some other protein kinases (e.g. plant PVPK-1 and yeast DBF2) belonging to a different subgroup of AGC protein kinases (Smith *et al.*

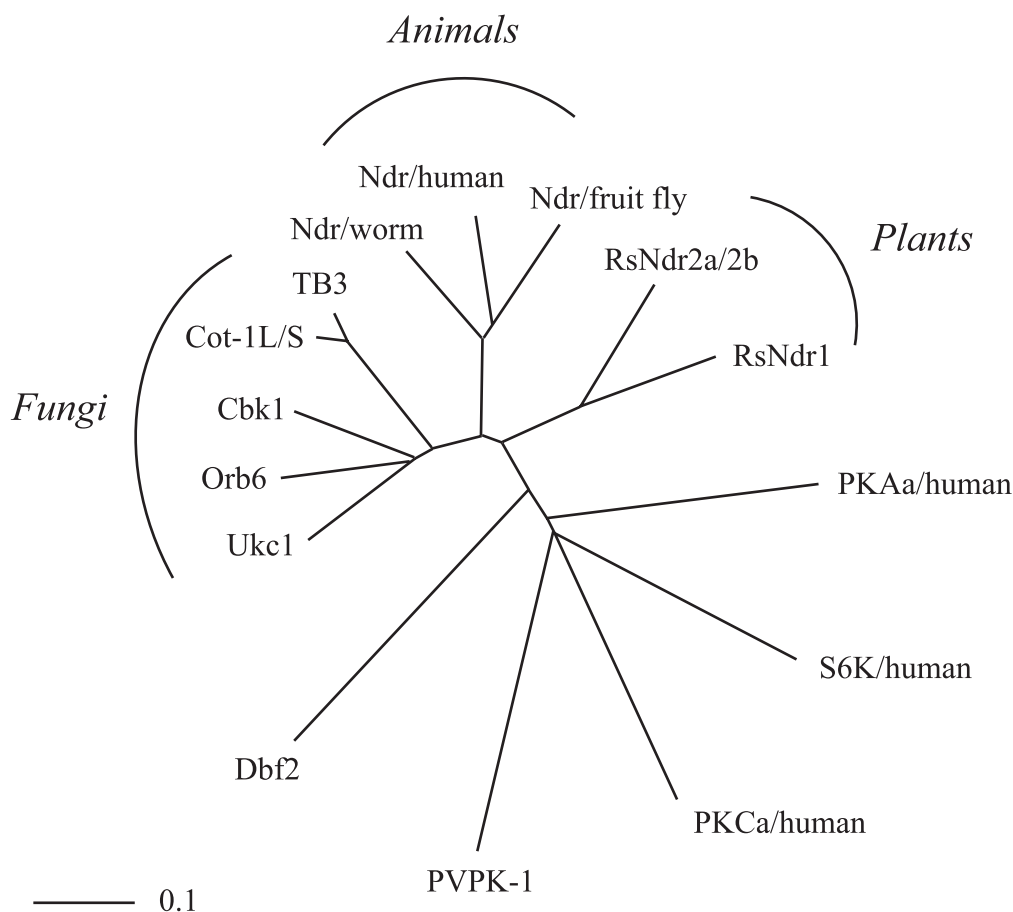


Fig. 7. A phylogenetic tree showing the relative relationship of *RsNdrs* among selected protein kinases in the AGC group. For the abbreviation of the name of a protein kinase, refer to the text.

1997). The insert is never found in the authentic AGC protein kinases. From these data, I consider *RsNdrs* as plant orthologues of animal *Ndr* and fungal *cot-1* like protein kinases. It is also reported that *Ndr* and *cot-1*, belonging to NDR kinase family in AGC group, have three conserved Ser/Thr residues that activation of Ndr protein kinase involves phosphorylation of. The fact that *RsNdrs* also have three Ser/Thr residues supports the conclusion (Fig. 8).

Although the functions of *RsNdrs* have not yet been clarified, it is interesting to note that mutants deficient in the aforementioned protein kinases in the other organisms exhibit interesting phenotypes. For example, temperature-sensitive mutations in *cot-1* result in hyperbranched hyphae (Yarden *et al.* 1992). Mutations in *orb6* cause a loss of polarized growth and delayed entry into mitosis (Verde *et al.* 1998) and mutations in *cbk1* result in altered cell morphology (Bidlingmaier *et al.* 2001). It is also reported that *D. melanogaster Ndr (trc)* is required for the integrity of epidermal

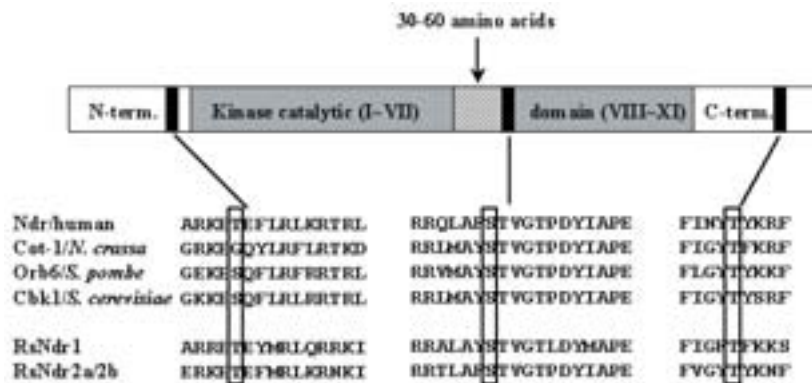


Fig. 8. Structure of Ndr protein kinase. The insert is shown by an obliquely striped. Three conserved Ser/Thr residues are boxed.

cell extensions (Geng *et al.* 2000). Moreover, it is interesting to note that both the fungal *cot-1* like protein kinases and *Ndr* influence cell division and/or cell morphology through the control of cell polarity (Tamaskovic *et al.* 2003). Since a similar role for *RsNdrs* is expected in plant cells, it will be interesting to see the phenotype of the plant in which the expression of *RsNdrs* is altered. *RsNdr*-like sequences are also found in other plant species such as *A. thaliana*, *N. tabacum*, *S. oleracea* and *M. crystallinum*, though none of their protein products are characterized. When I started the experiment, only two *Arabidopsis* ESTs (N65509 and N64988) relevant to *Ndrs* were deposited in the databases, but now it turned out that several (probably eight) genes, including '*AtNdr1*' (DDBJ accession no. AB047278, in their nomenclature), may encode a putative Ndr protein kinase in *Arabidopsis* (Wang *et al.* 2003). Further characterization of *RsNdr* homologs in *Arabidopsis* will be presented in the next chapter. In summary, I concluded that *RsNdrs* are members of a new protein kinase superfamily that universally exist across the plant, animal and fungal kingdoms.

Expression of *RsNdrs*

To examine the expression pattern of *RsNdrs*, I performed Northern blot and RT-PCR analyses. Because the results of these two experiments were not consistent with each other, the expression pattern of *RsNdrs* in different organs could not be clarified. The present results, however, showed that the transcript levels of *RsNdrs* are quite low. Although the signals were very weak, Northern blot analysis using an *RsNdr1* probe showed an interesting result; mRNAs that hybridized with the probe show differences in molecular size among the organs (Fig. 4). The result suggests that *RsNdr1* might be regulated by a processing or splicing event. In *N. crassa*, a transcript of *cot-1* is subjected to photoregulation, resulting in the change of a transcript of size and a translation initiation codon (Lauter *et al.* 1998). Further work is needed to reveal the significance of this observa-

tion.

It should be emphasized that aberrant cDNAs were isolated for *RsNdr3* and *RsNdr4*, i.e. *RsNdr3L* included intron-like sequences in the potential ORF, whereas *RsNdr3S*, *RsNdr4S* and *RsNdr4L* all encoded proteins lacking some of catalytic domains. Although the nucleotide sequence of the 5' half of *RsNdr4* was almost identical to the *RsNdr2*, genome sequences for the *RsNdr2* and *RsNdr4* showed that the two cDNAs are derived from different loci (data not shown). At present, the reason why such aberrant mRNAs are produced in some *RsNdrs* is not known. If aberrant mRNAs are translated into a protein, such molecules may compete with fully functional protein kinases toward substrates or regulatory factors that might bind to some domains shared with the proteins. It is not known whether kinds of regulatory mechanism exist behind this observation or not.

Protein kinase activity of RsNdr proteins expressed in E. coli

In order to prove protein kinase activity of the products, fusion proteins of *RsNdrs* were expressed in *E. coli*. Initially, two other expression vectors, pRSET (Invitrogen) and pGEX (Pharmacia), were tested for *RsNdr1* cDNA expression. In the former case, however, introduction of a fusion plasmids containing *RsNdr1* into *E. coli* strain BL21 (DE3)pLysS made the cells lethal even before IPTG induction. In the latter case, the expressed fusion protein of *RsNdr1* with a glutathione S-transferase (GST) tag was insoluble, suggesting that the protein was localized in the inclusion bodies (data not shown). Thus, the expression vector, pMAL-c2x, was employed for constructing a fusion plasmid pMAL-*RsNdr1*. In the construct, a moderate amount of the fusion protein was detected in the soluble fraction of total *E. coli* proteins.

When the protein kinase activity of the *E. coli* cell lysate expressing pMAL-*RsNdr1* was measured, activity was observed on myelin basic protein, indicating that the *RsNdr1* is a functional protein kinase, as suggested by the amino acid sequence. The fact that expression of a kinase-dead version of *RsNdr1* abolished its activity supports our conclusion (Fig. 6). Purification of the fusion protein, however, led to the loss of enzymatic activity for unknown reason(s). The fusion protein may be unstable and denatured during the process of purification, or additional factor(s) required for the full activity of the enzyme might be present in the lysate of *E. coli* cells. In this assay, the protein kinase activity of *RsNdr1* was not changed by adding cAMP or Ca^{2+} , well known regulatory factors for PKA and PKC, respectively, to the reaction mixture (data not shown). As for *RsNdr2*, protein kinase activity was not detected in *E. coli* lysate harboring a fusion plasmid, although *RsNdr2*, like *RsNdr1*, contains all 12 catalytic domains in its ORF (Fig. 2). The amount of the *RsNdr2* fusion protein in cell lysate may be so low that the activity cannot be detected in our assay condi-

tions. Several smaller proteins than the expected one were detected by anti-MBP serum in the soluble fraction of *E. coli* lysate with *RsNdr2*. This may indicate significant degradation of the protein in *E. coli*. Alternatively, the three substrates used here may be inappropriate for *RsNdr2* protein. In some cases, it is difficult to demonstrate protein kinase activity in bacterial expression systems, because of the lack of appropriate post-translational modification system that is, in some cases, very important for enzymatic activity. Therefore, other expression systems than *E. coli*, such as yeast, mammalian cells and insect cells, should be tested to prove the protein kinase activity of *RsNdr2*.

Chapter 2. Genome-wide analysis of protein kinase genes using a model plant *A. thaliana*.

2.1 Introduction

In the first chapter, I have isolated and characterized cDNAs encoding novel Ser/Thr protein kinases, *RsNdrs*, from radish. Based on the sequence and structural similarity, I concluded that these protein kinases are homologs of animal protein kinases called *Ndr* (Millward *et al.* 1995). The results also showed that *RsNdrs* consist of a small multi-gene family in radish. It is not clear, however, that how many *RsNdrs* exist in the radish genome, and whether the genes homologous to *RsNdr* consist of a multi-gene family in other plant species. I, therefore, performed genome-wide analysis using a model plant *Arabidopsis thaliana*, which belongs to the same family (*Brassicaceae*) as radish. For the genome-wide analysis, *A. thaliana* has many advantages over radish, including information on the almost completely-sequenced genome, availability of a large number of EST clones and T-DNA insertion or transposon-tagged lines. These useful information are deposited in the databases such as TAIR (The *Arabidopsis* information resource), and one can easily access to the resources that are necessary for an exhaustive study on the new genes of interest.

Since the sequence information of two ESTs from *A. thaliana* (N64988 and N65509) have been used for isolating *RsNdrs*, it is not surprising that the genes homologous to *RsNdrs* are present in the genome of *A. thaliana*. In this chapter, in order to gain a better understanding of novel protein kinase genes isolated as a family, I have made genome-wide survey for *RsNdr*-like genes in *Arabidopsis*. The study revealed that at least eight genes encoding for the *RsNdr*-like protein kinase exist in *Arabidopsis*, and seven of eight genes are transcribed to RNA. I have cloned full-length cDNAs into a plasmid vector by RT-PCR, and determined their primary structure. I have also cloned and characterized new two full-length cDNAs from radish using the information obtained from *Arabidopsis*.

2.2 Materials and methods

Plant materials

Arabidopsis thaliana ecotype Columbia was used as the material plant. The plants were grown in a greenhouse under normal day-light/temperature conditions to maturity.

Data mining

BLAST homology search was performed in the database of NCBI with the sequences of two *Arabidopsis* ESTs (N64988 and N65509) and *RsNdrs*. Eight candidate genes for *RsNdr*-related genes (hereafter *AtNdr*) were identified. The various PCR primers were designed from the nucleotide sequences of *AtNdrs*.

RT-PCR

Total RNA was isolated from 100 mg of leaves, stems and roots of *Arabidopsis* plants using an RNeasy plant mini kit (Qiagen). cDNA was reverse-transcribed from 5 µg of total RNA using Superscript preamplification system for first-strand cDNA synthesis (GIBCO BRL). PCR amplification of cDNA was performed as described. The PCR products were cloned into a plasmid vector, pCR 2.1 (Invitrogen), and their nucleotide sequences were determined as described. Based on the sequences of RACE products, primers specific to the 5'- and 3'- ends of the *A. thaliana* cDNAs were synthesized, and the primers were used to obtain putative full-length cDNAs. Complete nucleotide sequences of putative full-length cDNAs were determined after cloning them into a plasmid vector as described.

Isolation of new RsNdrs from radish

RT-PCR was performed on radish RNA with the primer pairs designed for *AtNdr3* sequence (forward: AtNdr3/F2; 5'-ATGCAAATGCAGAGTCTGCAGG-3', reverse: AtNdr3/R2; 5'-TCACCCGGATGAGCTCGACTTAC-3'). The nucleotide positions of the primers were +84 ~ +105 and +1531 ~ +1553 for AtNdr3/F2 and AtNdr3/R2, respectively. The PCR products were cloned into a plasmid vector, pGEM-T Easy (Promega), and their nucleotide sequences were determined as described. RACE was performed with various radish-specific primers. The primers designed for *AtNdr5* (AtNdr5/5RACE/1; 5'-TAGCGAGAACAGATTCGGCG-3', AtNdr5/5RACE/2; 5'-CTTCATCTTCACTTAAAGTC-3', AtNdr5/3RACE/1; 5'-TG GTAAGCTCTTGTGTAGCG-3', AtNdr5/3RACE/2; 5'-TGAGGTCAACGACGATTGG-3') were also used. Nucleotide positions of AtNdr5/5RACE/1, AtNdr5/5RACE/2, AtNdr5/3RACE/1, and AtNdr5/3RACE/2 were +768 ~ +787, +735 ~ +754, +1275 ~ +1294 and +1401 ~ +1420 in *AtNdr5a*, respectively. Putative full-length cDNAs were amplified by

RT-PCR with the primer pairs specific to the 5'- and 3'- ends of the radish cDNAs. Complete nucleotide sequences of the putative full-length cDNAs for *RsNdr5* and *RsNdr6* were determined after cloning them into a plasmid vector as described.

2.3 Results

Identification of AtNdrs

The first BLAST homology search using the two EST and *RsNdr2* sequences hit an *Arabidopsis* gene located on the chromosome 4 (GenBank accession At4g33080). BLASTN and BLASTP homology searches using nucleotide and amino acid sequences of At4g33080 and *RsNdrs* found the following seven *AtNdr* genes: AAD10677 and At1g30640 located on the chromosome 1, AAC16470 and At2g20470 on the chromosome 2, At3g23310 on the chromosome 3, At4g14350 on the chromosome 4 and ABO16893 on the chromosome 5.

In order to check the expression of these *AtNdr* genes, RT-PCR was performed with a primer pairs specific to each *AtNdrs*. Except for At2g20470, a PCR product with an expected size was amplified, though unspecific, minor fragments were observed in some cases (Table 1).

Table 1. Expression of *AtNdrs*.

<i>A. thaliana</i>	Expression			gene name
	leaf	stem	root	
Chr. 1/AAD10677	—	+	+	<i>AtNdr5</i>
Chr. 1/At1 g30640	+	+	+	<i>AtNdr7</i>
Chr. 2/AAC16470	+	+	+	<i>AtNdr6</i>
Chr. 2/At2 g20470	—	—	—	
Chr. 3/At3 g23310	+	+	+	<i>AtNdr2</i>
Chr. 4/At4 g14350	+	+	+	<i>AtNdr3</i>
Chr. 4/At4 g33080	+	+	+	<i>AtNdr1</i>
Chr. 5/ABO16893	+	+	+	<i>AtNdr4</i>

- a) *AtNdrs* are indicated by chromosome number/Genbank accession.
b) Expression + indicates that product was amplified by RT-PCR using cDNA from leaves, stems or roots, whereas expression — indicates that product was not present. When the expression is confirmed, genome sequences are named as *AtNdr* (in indicated in gene name).

In AAD10677, no PCR product was amplified in the RNA from leaves, but the product was observed in the RNAs from stems and roots. Here, the expressed *AtNdrs* were named *AtNdr1* (= At4g14350), *AtNdr2* (= At3g23310), *AtNdr3* (= At4g14350), *AtNdr4* (= ABO16893), *AtNdr5* (= AAD10677), *AtNdr6* (= AAC16470) and *AtNdr7* (= At1g30640) (because *AtNdr* was expediently numbered, gene number was not correspond to homolog of *RsNdr*). To obtain a full-length cDNA for each

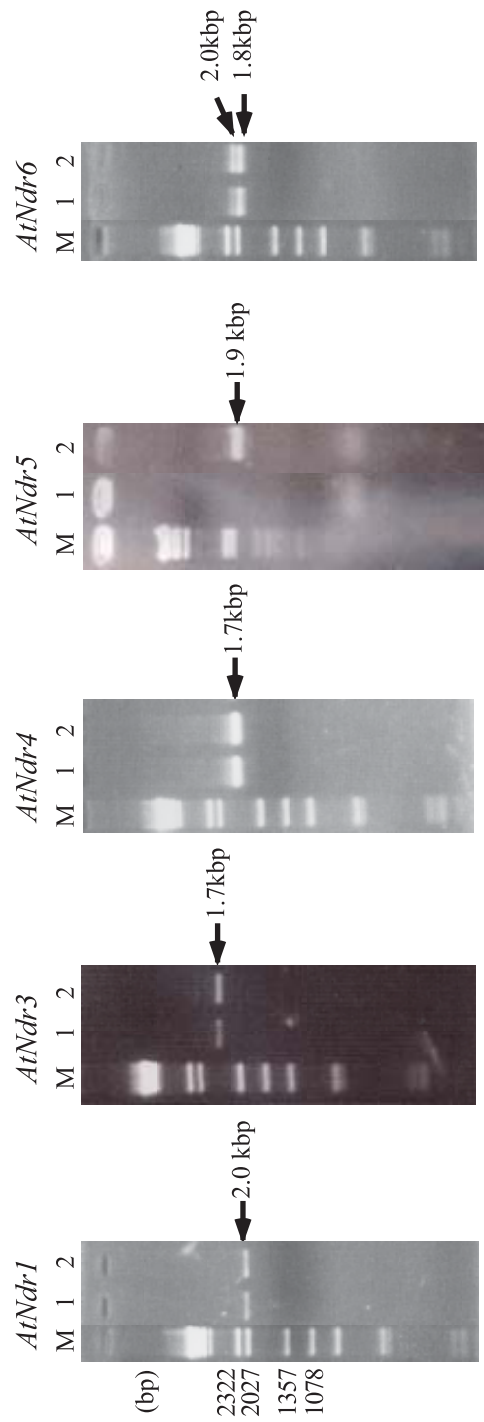


Fig. 9. A full-length *AtNdr* cDNA fragment amplified by RT-PCR. cDNA was reverse-transcribed from arabidopsis's leaves (1) and stems (2). PCR was performed using each of the 5'- and 3'- end specific primers designed from the sequences of RACE products. M: A mixture of the λ /HindIII and ϕ X174 HaeIII digests used as molecular size standards.

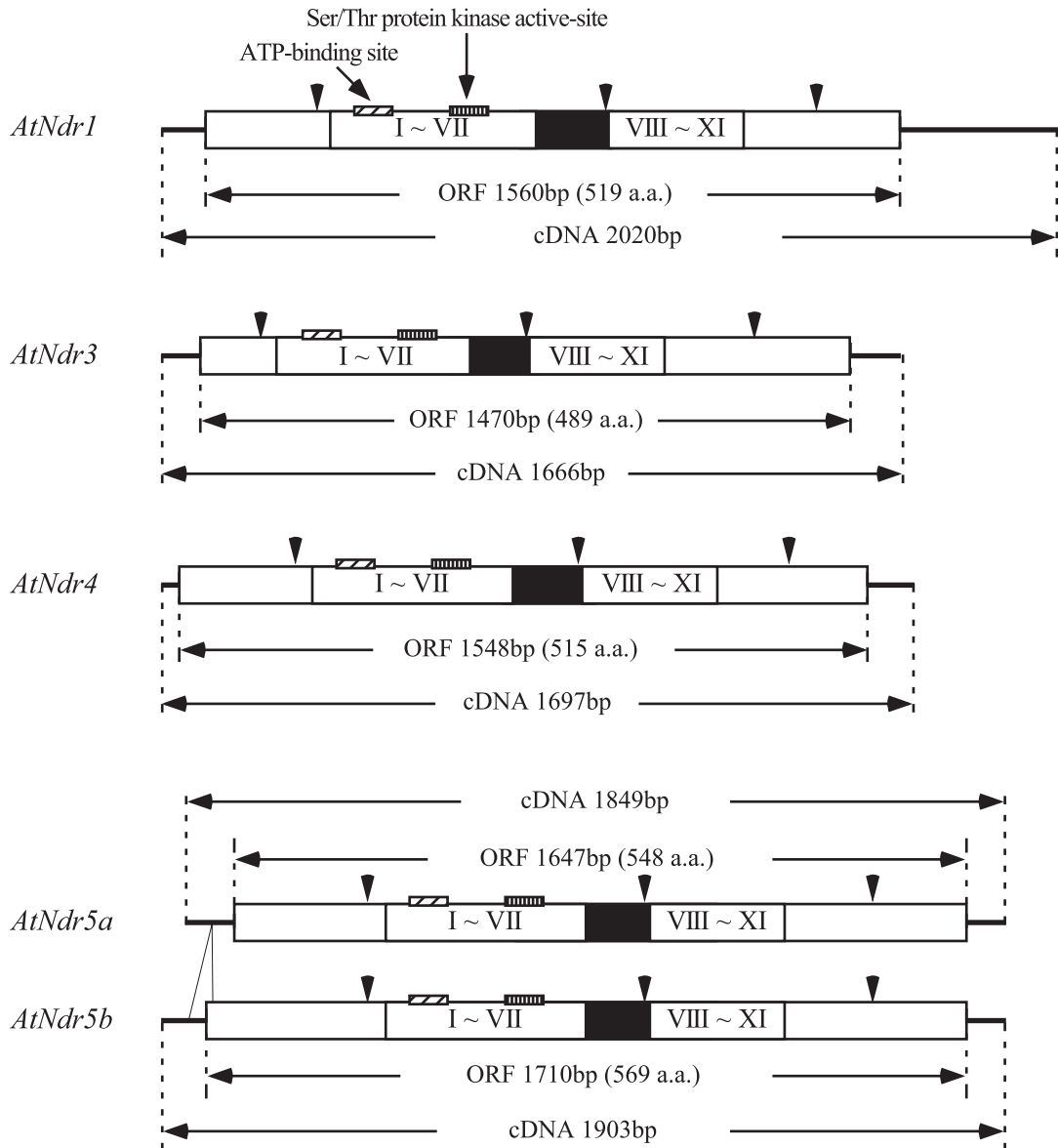


Fig. 10. The architecture of cDNAs for *AtNdr1*, *AtNdr3*, *AtNdr4* and *AtNdr5a/5b*. An open reading frame is indicated by a box, whereas 5'- and 3'- UTRs are shown by a horizontal line. Conserved catalytic sub-domains are shown by Roman numerals. The ATP-binding site, Ser/Thr protein kinase active-site and an extra amino acid sequence between catalytic domains VII and VIII are shown by an obliquely striped, vertically striped and black boxes, respectively. The three conserved Ser/Thr residues are marked by horizontal arrowheads.

AtNdr, a series of RACE experiments was performed. With a pair of specific primers to each *AtNdr*, putative full-length cDNAs for five of the seven *AtNdrs* were obtained by RT-PCR (Fig. 9). The size of RT-PCR products was 2.0kb, 1.7kb, 1.7kb and 1.9kb for *AtNdr1*, *AtNdr3*,

AtNdr4 and *AtNdr5*, respectively. For *AtNdr6*, two cDNA fragments of 2.0kb and 1.8kb were amplified. Full-length cDNAs for *AtNdr2* and *AtNdr7* have not been obtained.

Primary structures of *AtNdrs*

DNA sequencing of the 2.0-kb product for *AtNdr1* showed that cDNA (2020bp) contained a single ORF of 1560bp encoding a protein of 519 amino acids with a calculated molecular mass of 60kDa (Fig. 10), whereas that of the 1.7-kb product for *AtNdr3* revealed that cDNA (1666bp) contained a single ORF of 1460bp encoding a protein of 489 amino acids with 56.4kDa (Fig. 10). DNA sequencing of a 1.7-kb fragment for *AtNdr4* showed that cDNA (1697bp) contained a single ORF of 1548bp encoding a protein of 515 amino acids of 59kDa (Fig. 10), and that of a 1.9-kb product for *AtNdr5* revealed that the product actually contained two cDNAs with similar size (1849bp, named *AtNdr5a*, and 1903bp, *AtNdr5b*). *AtNdr5a* contained a single ORF of 1647bp encoding a protein of 548 amino acids with a calculated molecular mass of 62kDa, whereas *AtNdr5b* contained a single ORF of 1710bp encoding for 569 amino acids of 65kDa (Fig. 10). The DNA sequences for *AtNdr5a* and *AtNdr5b* are quite similar to each other; only the difference is that *AtNdr5a* lacks the sequence from +61 ~ +114 in *AtNdr5b*. Deduced amino acid sequences of *AtNdr1*, *AtNdr3*, *AtNdr4*, *AtNdr5a* and *AtNdr5b* contain all of 12 highly conserved subdomains of the eukaryotic Ser/Thr protein kinases, including the ATP-binding site (in subdomains I-II) and Ser/Thr protein kinase active-site (in subdomain VI) (Fig. 10). The sequences also contain an insert between subdomains VII and VIII, and three conserved Ser/Thr residues, being characteristics of the NDR kinase family of eukaryotic protein kinases (Fig. 10) (Tamaskovic *et al.* 2003).

DNA sequencing of the amplification products for *AtNdr6* revealed that a 2.0-kb fragment contained two different cDNAs of 2046bp (*AtNdr6L*) and 1973bp (*AtNdr6M*) and a 1.8-kb fragment contained the cDNA of 1868bp (*AtNdr6S*). *AtNdr6M* contains a single ORF of 1584bp encoding for 527 amino acids of 60kDa (Fig. 11). A deduced amino acid sequence of *AtNdr6M* showed all characteristics for NDR kinase family (Fig. 11). Deduced amino acid sequences of *AtNdr6L* and *AtNdr6S*, on the contrary, lack the conserved subdomains VII ~ XI in their ORFs, although they contain a single ORF of 750bp encoding 249 amino acids (Fig. 11).

Gene structures for *AtNdrs*

By comparing full-length cDNAs of *AtNdrs* with the genomic sequences of *Arabidopsis*, gene structures for the *AtNdr1*, *AtNdr3* and *AtNdr4* were determined (Fig. 12). Genes for *AtNdr1* and *AtNdr3* consist of 13 exons and 12 introns, whereas that for *AtNdr4* consists of 12 exons and 11 introns. A single gene consisting of 13 exons and 12 introns encodes both *AtNdr5a* and *AtNdr5b*

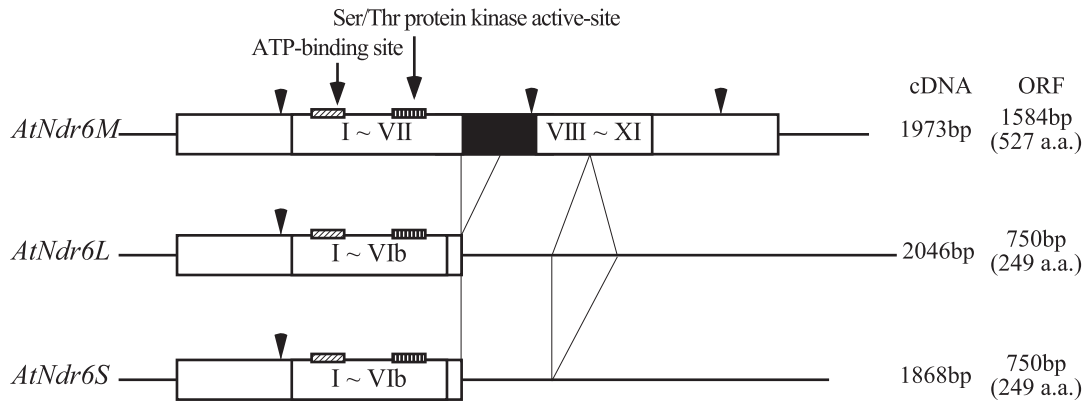


Fig. 11. The architecture of cDNAs for *AtNdr6S*, *AtNdr6M* and *AtNdr6L*. Notes and comments are the same as those in Fig. 10.

cDNAs. Two cDNAs were different from each other in the 5' sequence of exon 2, suggesting that a difference in the two cDNAs was produced by alternate splicing of the first intron in a gene. This alternate splicing event resulted in the difference in the position of an initiation codon (ATG) of *AtNdr5*; putative amino acid sequences of *AtNdr5a* are 21 amino acids longer than those of *AtNdr5b*. A single gene consisting of 14 exons and 13 introns encodes the three related cDNAs, *AtNdr6S*, *AtNdr6M*, and *AtNdr6L*. In *AtNdr6M*, the 5th intron was not spliced out, whereas in *AtNdr6L* the 8th intron was not spliced out (Fig. 12).

Primary structure of *RsNdr5* and *RsNdr6*

Based on the similarity of the nucleotide and amino acid sequences, new *RsNdrs* were isolated from radish. In order to isolate a radish homolog of *AtNdr3*, RT-PCR was performed on the RNA from radish leaves with the primers designed for *AtNdr3* (data not shown). The size of the RT-PCR product was 1.5-kb. The deduced amino acids sequences of the product were highly homologous to the *AtNdr3* and contained catalytic domains of Ser/Thr protein kinase. Here, a radish homolog of *AtNdr3* is named *RsNdr5*.

To check the presence of a homolog of *AtNdr5* in radish, genomic Southern blot analysis was performed with the *AtNdr5* gene probe. Hybridization signal was detected in radish genome (data not shown), suggesting that a homolog of *AtNdr5* is present in radish genome. A 3'RACE product was obtained in the RNA from stems, and nucleotide sequence was determined. The deduced amino acid sequence of 3'RACE product was highly homologous to the *AtNdr5*, and it contained catalytic domains of Ser/Thr protein kinase. Here, a radish homolog of *AtNdr5* is named *RsNdr6*.

The size of an RT-PCR product was 1.9kb for *RsNdr5* and 1.5kb for *RsNdr6* (data not

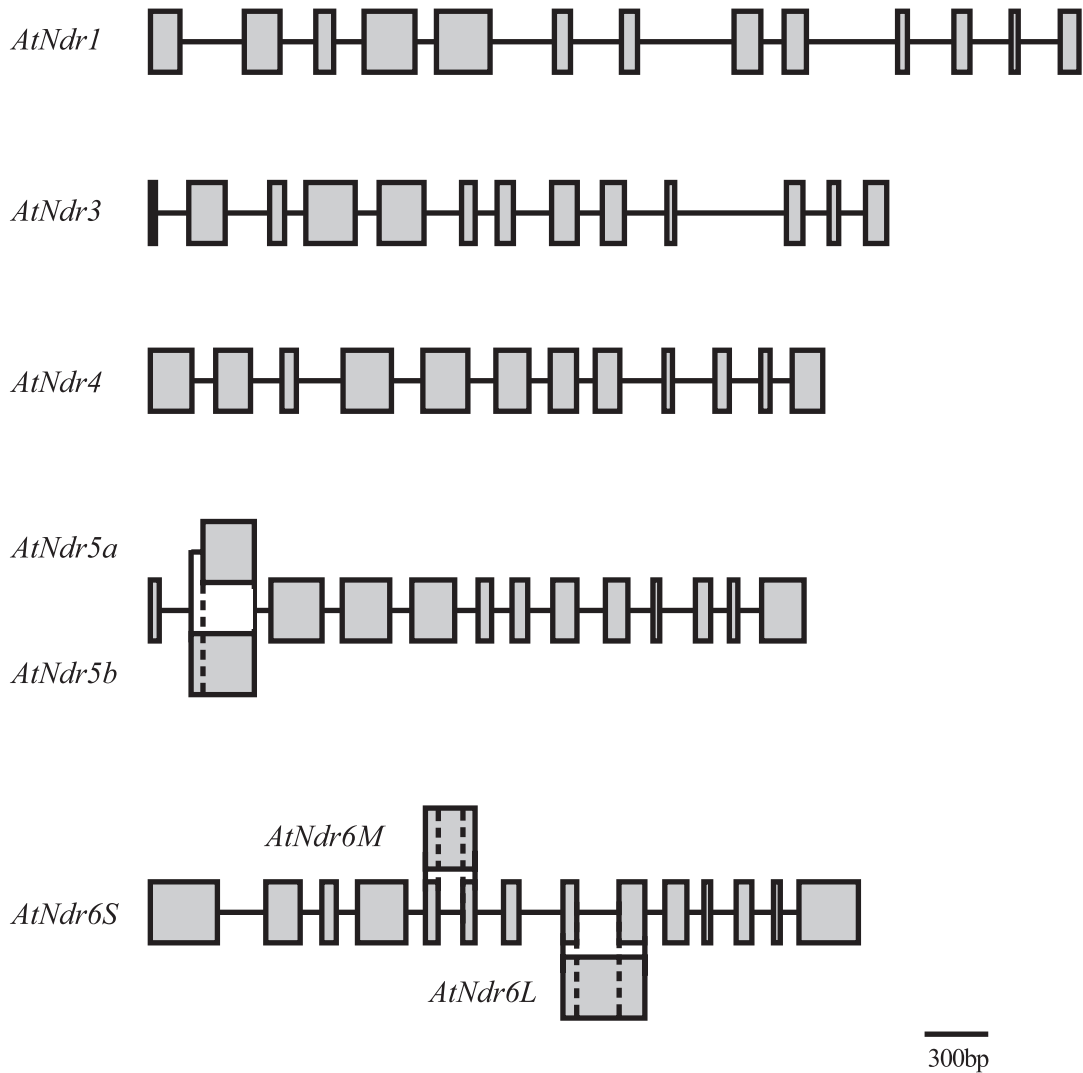


Fig. 12. The genome structure of *AtNdr1*, *AtNdr3*, *AtNdr4*, *AtNdr5* and *AtNdr6*. The exon is indicated by a box, whereas the intron is shown by a horizontal line.

shown). A 1.5-kb product was not amplified in the RNA from leaves but stems. DNA sequencing of a 1.9-kb product for *RsNdr5* showed that the cDNA (1918bp) contained a single ORF of 1470bp encoding a protein of 489 amino acids with a calculated molecular mass of 56.4kDa (Fig. 13). DNA sequencing of the 1.5-kb product for *RsNdr6* showed that cDNA (1543bp) contained a single ORF of 1461bp encoding a protein of 486 amino acids with a calculated molecular mass of 56.5kDa (Fig. 13). Deduced amino acid sequences of *RsNdr5* and *RsNdr6* contain all of 12 highly conserved subdomains of the eukaryotic Ser/Thr protein kinase, including the ATP-binding site (in subdomains I-

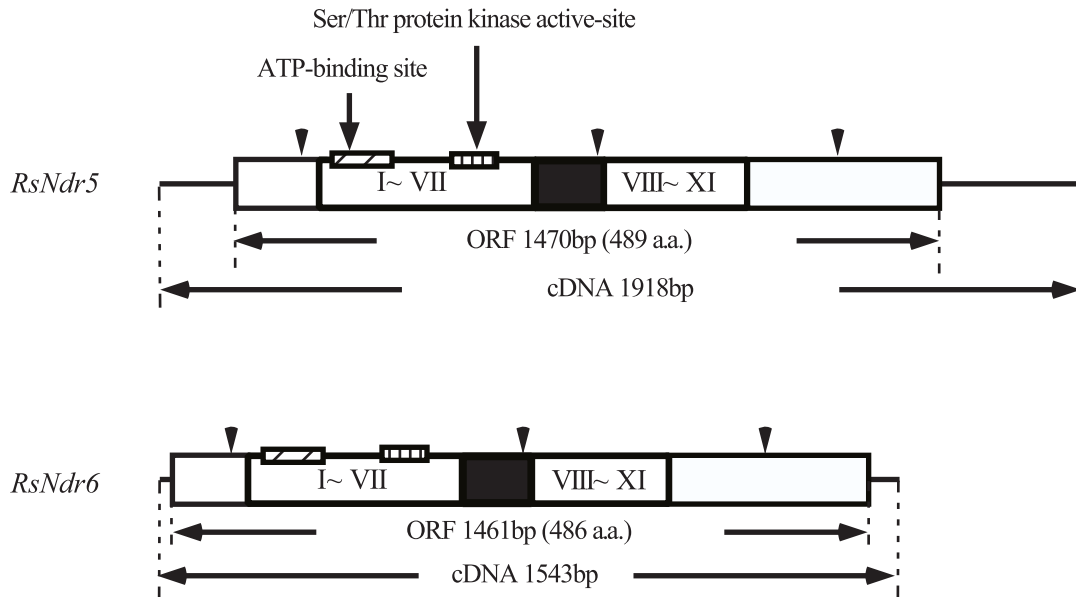


Fig. 13. The architecture of cDNAs for *RsNdr5* and *RsNdr6*. Notes and comments are the same as those in Fig. 10.

II) and Ser/Thr protein kinase active-site (in subdomain VI) (Fig. 13). The sequences also contain an insert between subdomains VII and VIII, and three conserved Ser/Thr residues, being characteristic of the NDR kinase family of eukaryotic protein kinase (Fig. 13) (Tamaskovic *et al.* 2003).

2.4 Discussion

Eight candidate genes encoding for RsNdr-like genes are present in the genome of Arabidopsis

In this chapter, I have explored candidate genes for *RsNdr*-like protein kinases in *Arabidopsis* using information on the genome and EST sequences of this model plant. With this genome-wide study, it appeared that a total of eight *AtNdr* genes present in the *Arabidopsis* genome. The study also showed that *AtNdr* genes are not clustered in the genome of *Arabidopsis*; one or two *AtNdrs* are located on each of the five *Arabidopsis* chromosomes. RT-PCR experiments confirmed that at least seven out of eight *AtNdrs* genes are expressed. Presence of several EST sequences related to *AtNdr* also supports my conclusion that all *AtNdrs* but *At2g20470* are transcribed. DNA sequencing of full-length cDNAs indicates the most of cDNAs contain a long ORF encoding a protein with all characteristics for Ndr protein kinases. This suggests that transcribed genes identified here are functional ones. Since a PCR product for *At2g20470* was not obtainable, and ESTs corresponding to the *At2g20470* are not found in the database, I wonder whether *At2g20470* is a functional gene or not. A deduced amino acid sequence of *At2g20470* in the database, however, contains all 12 con-

served subdomains necessary for the eukaryotic Ser/Thr protein kinases, and it exhibits all structural characteristics for the Ndr protein kinases. Therefore it cannot be ruled out the possibility that At2g20470 is a functional gene. It may be expressed under very specific conditions, e.g. a certain developmental stage or under a particular biotic or abiotic stress. In the present RT-PCR analysis, organ-specific expression was observed only for the *AtNdr5* (AAD10677) (Fig. 9 and Table 1). The available data on a microarray in the public database showed that the expression of *AtNdr3* (At4g14350) is also regulated in an organ-specific manner, i.e. *AtNdr3* is strongly expressed in roots (data not shown). Organ-specific expression of *AtNdr3*, however, was not clear in my RT-PCR data. Further study focused on the promoter sequences of *AtNdrs*, including At2g20470, will reveal the unique and overlapping expression patterns of the members of *AtNdr* gene family. The present genome-wide survey using *A. thaliana* clearly shows that the gene for Ndr protein kinase consists of a small multi-gene family in the genome. In the chapters 1 and 2, I have identified at least six cDNAs encoding Ndr protein kinases from radish. Therefore, a small multi-gene family for Ndr protein kinases is expected in other plant species, at least in the plants of *Brassicaceae*. Although the functions of plant Ndr protein kinases have not yet been clarified, it is interesting to see whether each member of a multi-gene family has a separate function or not.

Alternate splicing as a possible mechanism to produce cDNA variants of AtNdrs

DNA sequencing shows that both *AtNdr6S* and *AtNdr6L* do not encode a functional protein kinase, because their ORFs are interrupted by an intron-like sequence. A stop codon present *in frame* in the sequence abolishes catalytic subdomains VII~XI in both cDNAs. Although aberrant splicing is a possible mechanism to generate such cDNA variants, it should be noted that *AtNdr6S* is produced by an additional splicing of an intron-like sequence from the 5th exon of *AtNdr6M* and that *AtNdr6L* is done by no splicing of 8th intron of *AtNdr6S* (Fig. 12). The cDNA for *AtNdr6M* encodes a protein with all catalytic subdomains, indicating that this splicing abolishes a functional mRNA species. Similarly, a gene for *AtNdr5* expressed two mRNA variants, *AtNdr5a* and *AtNdr5b*, by an alternate splicing of the 1st intron of the *AtNdr5* gene. In this case, each mRNA variant may encode a functional protein with a different N-terminal sequence. In the first chapter, I have reported that cDNAs containing incomplete ORFs are present in radish (i.e. *RsNdr3* and *RsNdr4*). Taken together, these results indicate that some *RsNdrs* and *AtNdrs* have mRNA variants produced by alternate splicing. The results further suggest that the expression and activity of *RsNdrs* and *AtNdrs* might be regulated at the splicing level. The importance of this observation remains to be studied.

Table 2. Homology among *RsNdrs*.

	<i>RsNdr1</i>	<i>RsNdr2a</i>	<i>RsNdr2b</i>	<i>RsNdr5</i>
<i>RsNdr2a</i>	52/69			
<i>RsNdr2b</i>	54/69	95/100		
<i>RsNdr5</i>	58/78	55/74	58/74	
<i>RsNdr6</i>	55/67	54/73	57/73	74/79

a) Homology among *RsNdrs* was calculated using amino acid sequences.

Homology using entire amino acid sequences (%) / Homology using catalytic domains (%)

Orthology among *RsNdrs* and *AtNdrs*

In this chapter, I have cloned and characterized two cDNAs (*RsNdr5* and *RsNdr6*) encoding protein kinases from radish using the information on *AtNdrs*. Deduced amino acid sequence of *RsNdr5* and *RsNdr6* cDNAs showed all characteristic of the NDR kinase family of eukaryotic protein kinases (Tamaskovic *et al.* 2003), making the number of full-length cDNAs characterized in radish four (i.e. *RsNdr1*, *RsNdr2a/b*, *RsNdr5* and *RsNdr6*). Southern blot analysis (data not shown) showed that the four cDNAs are not cross-hybridized with each other. DNA sequencing study revealed that deduced amino acid sequences of *RsNdrs* show about 55% and 75% identity in their entire and catalytic domains, respectively (Table 2). Currently, it can be said that at least four different but related *RsNdrs* are present in radish.

In order to infer the relationship among *RsNdrs* and *AtNdrs*, a phylogenetic analysis was carried out using their deduced amino acid sequences. Fig. 14 shows a NJ tree inferring phylogenetic relationships among *Ndr* sequences from radish and *Arabidopsis*. The tree clearly shows the orthology among *RsNdrs* and *AtNdrs*; *RsNdr1*, *RsNdr2a/2b*, *RsNdr5*, *RsNdr6* are the ortholog of *AtNdr4*, *AtNdr1*, *AtNdr3* and *AtNdr5a/b*, respectively. *AtNdr2* is more similar to a *RsNdr5/AtNdr3* group than to any other *Ndrs*, and radish orthologs of *AtNdr6M*, *AtNdr7* and *At2g20470* have not been identified. In the protein kinase assay using four synthetic peptides, *RsNdr1* and *AtNdr4* showed the same substrate specificity (data not shown), indicating the two protein kinases are an ortholog in the two species. As mentioned earlier, mRNA for *RsNdr6* was expressed in all RNAs tested but leaves. This pattern of expression is similar to that of *AtNdr5* in *Arabidopsis*, supporting their orthology inferred from the phylogenetic analysis. The study on the expression pattern and substrate specificity of each *Ndr* protein kinases will reveal the orthology among *Ndrs* from radish and *Arabidopsis*.

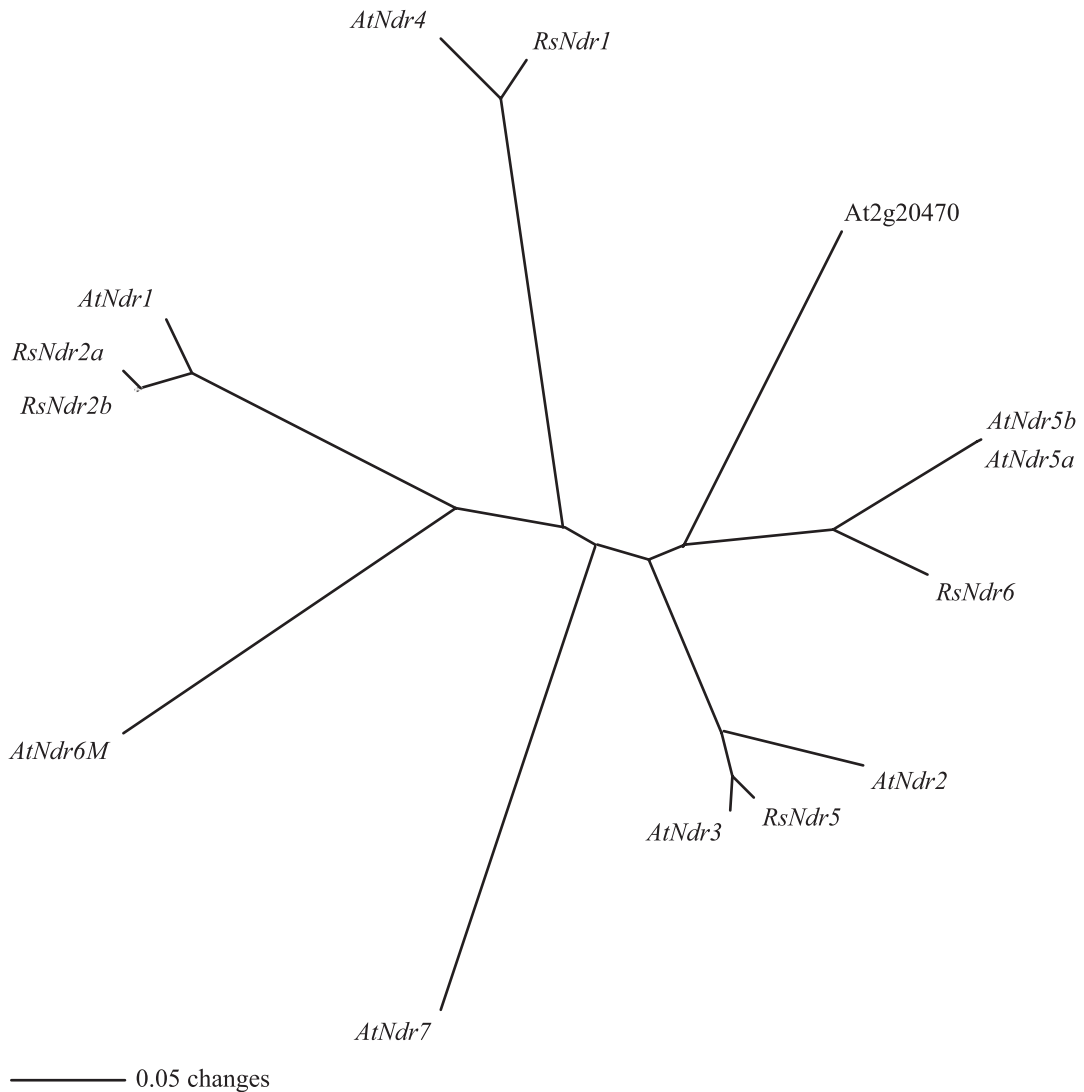


Fig. 14. A phylogenetic tree showing the relative relationship between *RsNdrs* and *AtNdrs*.

Chapter 3. Cloning and structural analysis of wheat cDNAs encoding *RsNdr* homolog by data mining of the EST database KOMUGI.

3.1 Introduction

Although common wheat (*Triticum aestivum*) is one of the world's most important crops, molecular cloning and functional analysis of wheat genes have been hampered by a large genome size, polyploidy, and a lack of molecular tools such as ESTs, ordered genomic clones, and gene-tag-

ged lines. In addition, difficulty in obtaining transformants has made wheat an inconvenient material for molecular study. However, the recent establishment of wheat EST databases such as KOMUGI (<http://shigen.lab.nig.ac.jp/wheat/komugi/top/top.jsp>) has dramatically changed the situation; at present, the number of EST sequences deposited in databases has grown to a point where it is comparable to that of ESTs for the model plant *Arabidopsis thaliana*, and one can now consider wheat a primary source for gene cloning.

In the chapters 1 and 2, I cloned and characterized cDNAs encoding novel protein kinases in the 'AGC' group from radish and *Arabidopsis*, to gain a better understanding of plant protein kinases in the group. Based on the sequence and structural similarity, I concluded that the protein kinases are a plant homolog of the fungal *cot-1* like protein kinases (Yarden *et al.* 1992) and animal *Ndr* (Millward *et al.* 1995). The results also showed that *RsNdrs* and *AtNdrs* consist of a small multi-gene family in radish and *Arabidopsis*, respectively. However, no information exists on the *RsNdr* like protein kinases in monocots. The main purpose of this study was to examine for the first time the gene encoding *RsNdr* like protein kinase in monocots. In this chapter, three cDNA clones encoding a Ser/Thr protein kinase which is homologous to *RsNdr*, were isolated by data mining of the wheat EST database KOMUGI, and the primary structure of these cDNAs was determined by RT-PCR and the 5'-RACE method. The results show that at least three homeologous genes are present in common wheat, each of which is expressed in young seedlings.

3.2 Materials and methods

Data mining and primary RT-PCR

The database KOMUGI, established as part of National BioResource Project in Japan, and GenBank were searched using the sequence of *RsNdr1* cDNA as an electronic probe, and 19 homologous EST sequences were selected. Fig. 15 shows the selected EST sequences and their alignment into a contig to make a hypothetical cDNA sequence. Although some heterogeneity was present among the sequences, the existence of several conserved regions allowed us to synthesize a primer pair, F01 (=5'-GGTGCTATCATGTATGAAATGC-3') and R01 (=5'-TTGTAATACACATGCTCGTCAG-3'), to be used in the initial RT-PCR to amplify the 3' half (~1kb) of a hypothetical cDNA (Fig. 15). Total RNA was isolated from young seedlings of the wheat cv. Chinese Spring, using an RNeasy plant mini kit (Qiagen), and RT-PCR was performed. Complete nucleotide sequences of RT-PCR products of an expected size (ca. 0.95kb) were determined after cloning them into a plasmid vector as described.

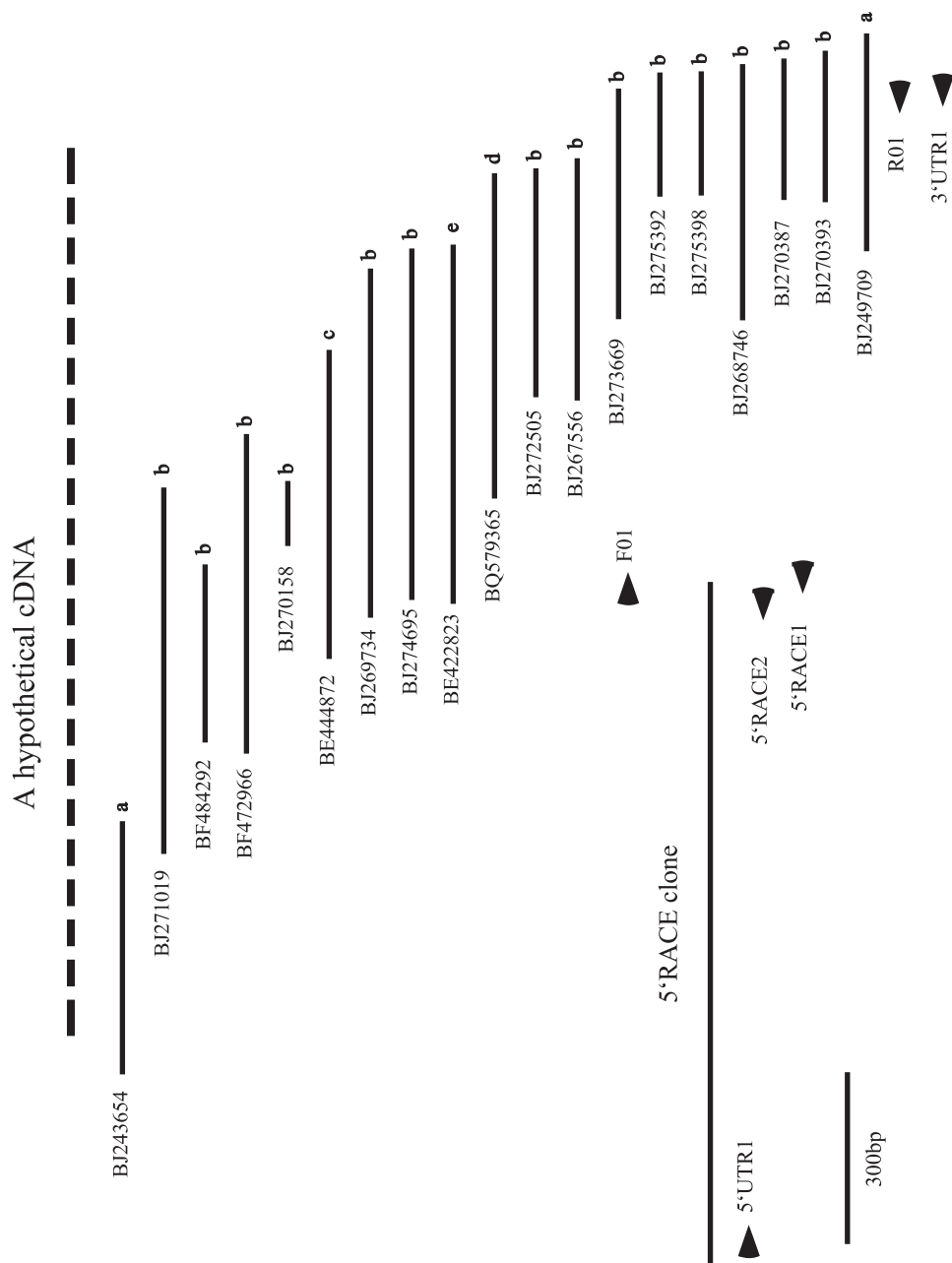


Fig. 15. Alignment of EST sequences selected from the databases KOMUGI (lines with the prefix BJ, followed by a six-digit number) and GenBank (others). cDNA libraries from which each EST clone was isolated are indicated as follows; a, spike at flowering; b, pistil at heading; c, etiolated seedling root; d, dormant embryo; e, endosperm. A hypothetical cDNA and the region covered by a 5'RACE clone are shown by a dotted and a bold line, respectively. Arrowheads indicate the position and the direction of primers used for RT-PCR and RACE.

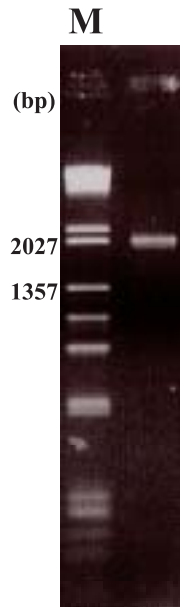


Fig. 16. A full-length cDNA fragment amplified by RT-PCR. Total RNA was prepared from wheat's leaves. PCR was performed using primer pair 5'UTR1 and 3'UTR1. M: A mixture of the λ /*Hind*III and ϕ X174 *Hae*III digests used as molecular size standards.

5'-RACE and the sequence analysis of full-length cDNAs

Since few EST clones were located in the predicted 5'-region of the cDNA (Fig. 15), RACE was performed with the primers 5'RACE1 (=5'-TGCAGGTCGACATTGAACC-3') and 5'RACE2 (=5'-GCATTTTCATACATGATAGCACC-3') designed from the DNA sequences of the initial RT-PCR products. A primer, 5'UTR1 (=5'-GGTATTGACCGCTTCAGAGC-3'), which is specific to the 5'-end of cDNA, was designed from the consensus sequence of the RACE products. Putative full-length cDNAs were then amplified by RT-PCR using 5'UTR1 and another primer, 3'UTR1 (=5'-TTCAATTGTAATACACATGCTCG-3'), whose sequence was taken from a consensus sequence of distal EST clones. Complete nucleotide sequences of the putative full-length cDNAs were determined after the cloning of RT-PCR products into a plasmid vector as described.

3.3 Results and discussion

Primary structure of cDNAs encoding putative wheat Ndr

A single fragment of expected size (~ 2 kb) was amplified by RT-PCR with the primer pair 5'UTR1 and 3'UTR1 (Fig. 16). The size of this fragment agrees with that of a hypothetical cDNA made with the assembly of EST sequences. Three clones with a putative full-length cDNA were selected and their complete nucleotide sequences were determined.

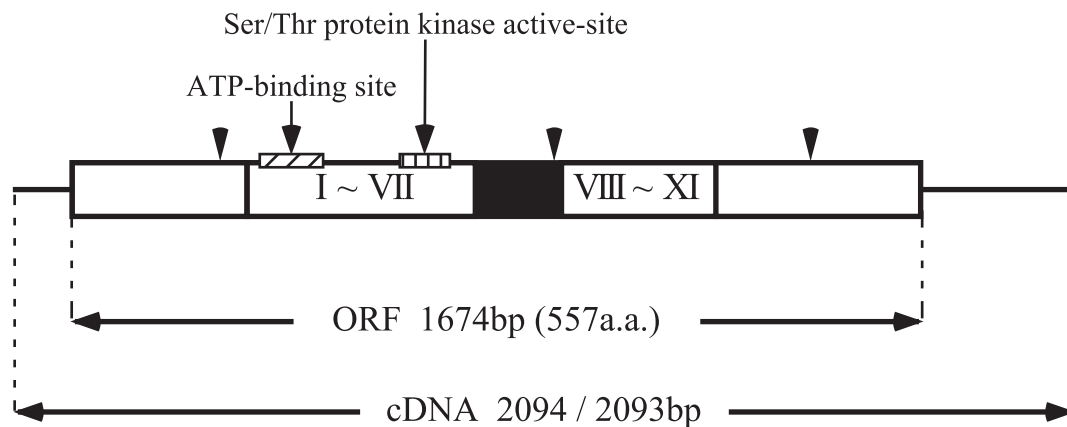


Fig. 17. The architecture of cDNAs for the wheat *Ndr*. Notes and comments are the same as those in Fig. 10.

Table 3. Position of the nucleotide differences among three cDNA clones of wheat.

Clone	5'UTR			Coding region										
	-40	-32	-26	102	144	420	501	648	663	708	819	828	1104	1155
1-7	c	g	g	C	T	T	C	C	T	T	C	C	A	A
1-9	g	a	a	G	C	C	T	A	A	A	T	T	A	G
2-3	c	g	g	C	T	T	C	C	A	T	T	T	G	G

Clone	Coding region							3'UTR						
	1443	1449	1515	1585	1689	1728	1729	1806	1817	1823	1892	1916	1949	1952
1-7	A	C	T	A	a	a	—	c	a	c	t	c	t	t
1-9	G	G	C	A	a	a	a	t	a	t	t	c	—	t
2-3	A	G	T	G	g	—	—	t	c	t	g	t	t	c

a) A nonsynonymous substitution is indicated by a boldface letter, whereas a deletion is shown by a hyphen.

Two of the clones (1-7 and 1-9) contain an RT-PCR product of 2094bp. The remaining clone (2-3) has a 2093bp product (Fig. 17). All three clones contained a single ORF of 1674bp encoding a protein of 557 amino acids with a calculated molecular mass of 64kDa (Fig. 17). Although the overall DNA sequences for the three clones are quite similar, three, 15 (14 synonymous and one nonsynonymous) and seven nucleotide differences were observed in the 5'-UTR, coding region and 3'-UTR, respectively (Table 3). Three insertions/deletions were also found in the 3'-UTR of the cDNAs. The deduced amino acid sequences of the three cDNA clones are nearly identical; only one amino acid difference (Gly vs. Ser, at position 539) was present between clone 2-3 and the others (Table 3). The deduced amino acid sequences of the three clones contain all 12 highly conserved subdomains of the eukaryotic Ser/Thr protein kinase (Hanks and Hunter 1995), including the ATP-

binding site (in subdomains I-II) and Ser/Thr protein kinase active-site (in subdomain VI) (Fig. 17). The sequences also contain an insert of 56 amino acids between subdomains VII and VIII, and three conserved Ser/Thr residues (Fig. 17), being characteristic of the NDR kinase family of eukaryotic protein kinases (Tamaskovic *et al.* 2003).

Relationship of wheat Ndr, RsNdr and AtNdr among plant protein kinases

In order to infer the relationship of the wheat *Ndrs*, *RsNdrs* and *AtNdrs* among plant protein kinases, a phylogenetic analysis was carried out using the deduced amino acid sequences. Since more than 1000 genes may encode eukaryotic protein kinases in *Arabidopsis* (Wang *et al.* 2003), and the 'protein kinases' include quite diverse families of proteins that differ in structure and function, a limited comparison was made for this purpose. Using the BLAST program, closely related sequences, putative protein kinases in tobacco (NTPKTL7), ice plant (MCPRTKINA) and spinach (SOPRTKINA) (GenBank accession nos. X71057, Z30329 and Z30330, respectively) were selected from the databases. Some plant protein kinases which had been classified into the 'AGC group' were also included in the analysis, as well as AK1 (GenBank accession no. NM 120569) (Harper *et al.* 1993) which belongs to CDPK. The latter was used as an outgroup. Fig. 18 shows a tree inferring phylogenetic relationships among the selected plant protein kinases. The tree indicates that the deduced amino acid sequences of wheat *Ndrs*, *RsNdrs* and *AtNdrs* make a cluster with the aforementioned putative protein kinases, and they are clearly distinguished from those of other 'AGC' protein kinases such as PVPK1 (Lawton *et al.* 1989). It is interesting to note that wheat *Ndr* sequences are more similar to some *AtNdrs* (e.g., At2g20470) than *RsNdr1* and its *Arabidopsis* homolog *AtNdr4*, though the *RsNdr1* sequence was used for selecting wheat ESTs from the database. Four *RsNdrs* and eight *AtNdrs* are present in radish and *Arabidopsis*, respectively (chapter 2). The three different cDNA clones characterized should be considered the products of homeologous genes in hexaploid wheat. Therefore, other members will be found in wheat with the accumulation of more data.

General conclusion

In this paper, I have cloned and characterized cDNAs encoding a novel Ser/Thr protein kinase belonging to AGC group, from radish (chapter 1), *Arabidopsis* (chapter 2) and wheat (chapter 3). Protein kinase activities of three cDNA products from radish and *Arabidopsis* were checked by using *E. coli* expression system, showing that at least two of cDNAs (*RsNdr1* and *AtNdr4*) encode a functional polypeptide. Based on the sequence and structural similarity, I concluded that these protein kinases are not the member of authentic protein kinases in the AGC group (e.g., PKA) but they are a homolog of the protein kinases in the NDR kinase family in the AGC group. Although the

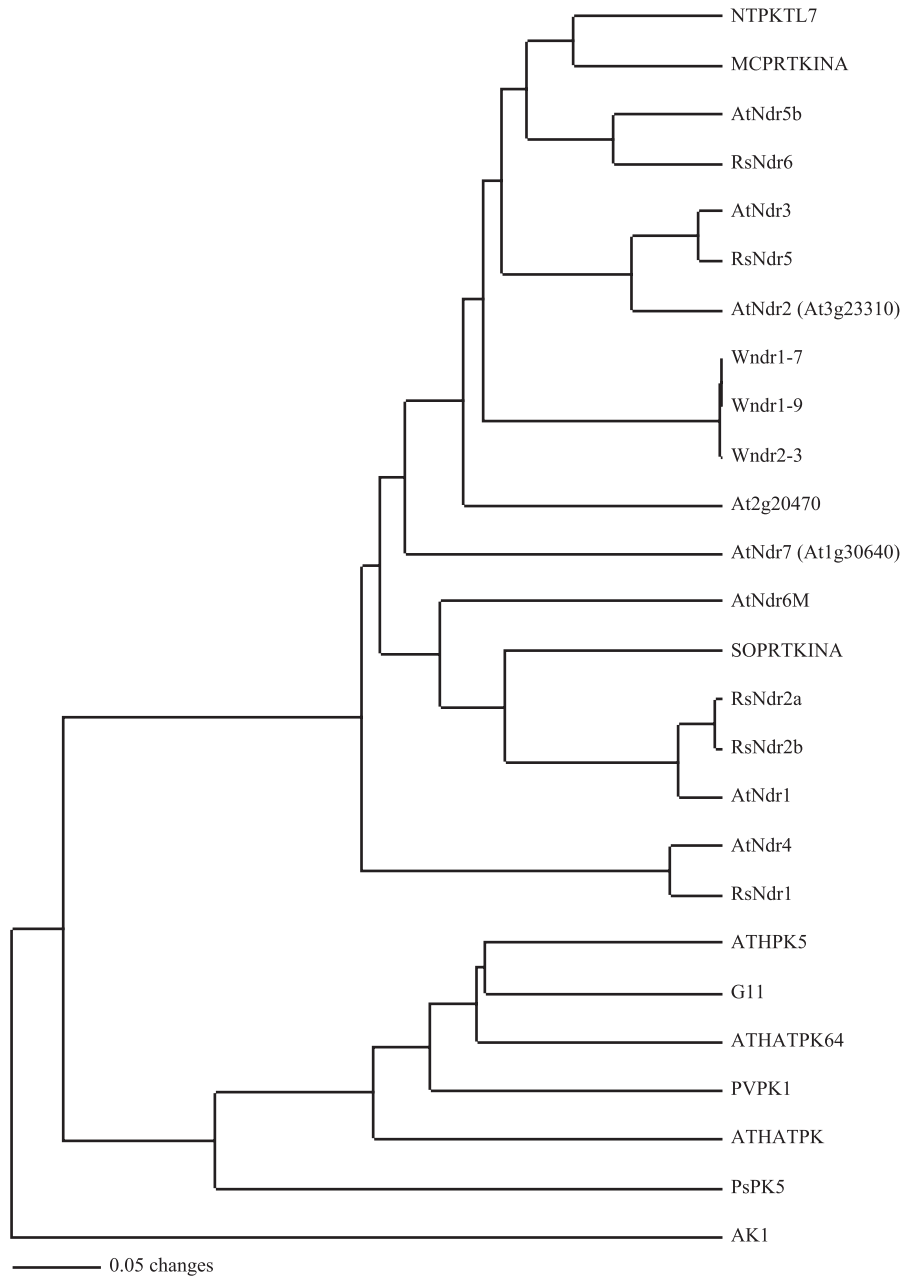


Fig. 18. A tree showing the relative relationship of wheat *Ndrs*, *RsNdrs* and *AtNdrs* among selected plant protein kinases in AGC group. The source of the protein kinases not referred to in the text is as follows; ATPK5 (Hayashida *et al.* 1993) (DDBJ accession no. D10909), ATHATPK64 (Mizoguchi *et al.* 1992) (DDBJ accession no. D10937), G11A (Lawton *et al.* 1989) (GenBank accession no. J04556), ATHATPK (Hayashida *et al.* 1992) (DDBJ accession no. D10910), PsPK5 (Lin *et al.* 1991) (GenBank accession no. M92989).

function of the plant Ndr protein kinases has not been clarified, it is interesting to note that the most of kinases in the NDR kinase family more or less influence cell division and/or cell morphology through the control of cell polarity. Thus, it is likely that plant Ndr protein kinases also have the same function as that of Ndr protein kinases in other organisms.

In order to reveal the function of plant Ndr protein kinases, I have introduced a full-length cDNA of *RsNdr1* and *AtNdr4* into tobacco plants in the sense or antisense orientations by *Agrobacterium*-mediated transformation method, and observed the phenotype of T₀ regenerants containing a transgene (data not shown). Southern, Northern and RT-PCR analyses performed on the several T₀ regenerants showed that the transgene(s) is integrated into the tobacco genome as expected and is expressed correctly under the control of a CaMV35S promoter, though the expression level was somewhat varied depending on the lines investigated. Although the transgene was expressed, the phenotypes of transgenic plants were not distinguishable from those of a control, non-transgenic tobacco plants (data not shown). All T₀ regenerants grew to the mature plants, flowered and set seeds normally in a green house. It will be interesting to see the phenotype of T₁ plants more carefully in the future.

I also tried to characterize the function of *AtNdr4*, by using Salk insertion mutants (<http://signal.salk.edu/cgi-bin/tdnaexpress>). I obtained four mutants (SALK-068700, 068791, 075540, 078846) harboring T-DNA insertion within the gene *AtNdr4*, and compared their phenotype with that of wild-type by observing the plants under a stereo microscope. Although I could not show conclusive evidence, the number of lateral roots in SALK-75540 seems to be larger than that of wild-type (Fig. 19). As the plants grow, however, the difference between a mutant line and a wild-type disappeared. These mutants are now growing in a green house, and waiting for setting the seeds. The offspring will be an indispensable tool for further study.

In this paper, it is found that the *RsNdr* and *AtNdr* consisted of a small multi-gene family in the two plant genomes. This suggests that each member may serve a distinct function in the cell. Elucidation of the signaling pathway where each gene is involved will be necessary to know a role of Ndr protein kinases that play in the plant cells. Recently, molecules which interact with protein kinases in the NDR kinase family have been reported. For example, in *S. cerevisiae*, it was reported that the association of Pag1p with Cbk1p (Ndr protein kinase in yeast) is essential for cell morphogenesis (Du *et al.* 2002). It was also reported that a *Mob2* product interacts with Cbk1p to promote polarized cell growth and to induce asymmetric cell fate by activating daughter-specific gene expression (Colman-Lerner *et al.* 2001, Weiss *et al.* 2002). In human, it was shown that Ndr1 and Ndr2 form stable complexes with MOB2, and this association dramatically stimulates Ndr1 and Ndr2 catalytic activity (Devroe *et al.* 2004). It is quite interesting to note that the sequences encoding Pag1-

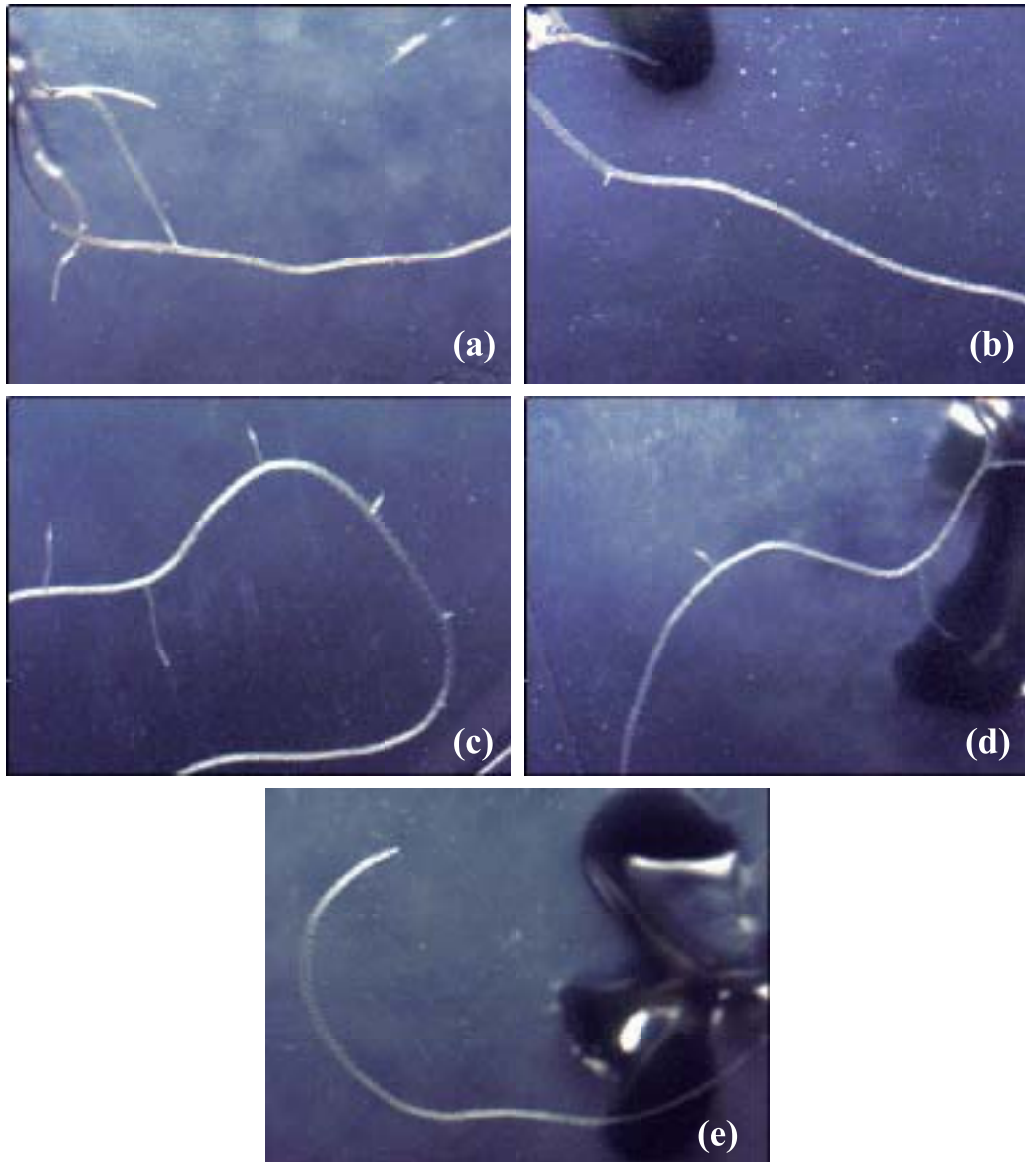


Fig. 19. Morphology of a wild-type and T-DNA insertion mutant plants. (a) SALK-068700, (b) SALK-068791, (c) SALK-075540, (d) SALK-078846 and (e) a wild-type plant. The photographs were taken at 9 days after sowing the seeds.

and Mob-like proteins exist in *A. thaliana* genome (data not shown). The finding that the related sequences to the Pag1- and Mob-like proteins exist in the *Arabidopsis* will open the way to reveal signaling pathway for plant Ndr protein kinases.

Acknowledgements

I wish to express deepest thanks to Dr. T. Terachi, Professor, Department of Biotechnology, Faculty of Engineering, Kyoto Sangyo University, for helpful advice during the course of the study. I thank Professor H. Yamagishi, Department of Biotechnology, Faculty of Engineering, Kyoto Sangyo University, for valuable suggestions, encouragement and supplying plant materials. I am also indebted Professor A. Kurosaka, Department of Biotechnology, Faculty of Engineering, Kyoto Sangyo University, for helpful advice on experiments on protein kinase assay. I am grateful to Mr. Satoshi Shimamura and Ms. Ayuka Fukai. I also thank all the members of 'the 31st Laboratory', Department of Biotechnology, Faculty of Engineering, Kyoto Sangyo University, for encouragement and various help during the course of the study.

References

- Asai, T., Tena, G., Plotnikova, J., Willmann, M. R., Chiu, W. L., Gomez-Gomez, L., Boller, T., Ausubel, F. M. and Sheen, J. (2002) *Nature* 415, 977-983.
- Berridge, M. J. (1993) *Nature* 361, 315-325
- Bhattacharya, S., Large, E., Heizmann, C. W., Hemmings, B. and Chazin, W. J. (2003) *Biochemistry* 42, 14416-14426.
- Bidlingmaier, S., Weiss, E. L., Seidel, C., Drubin, D. G. and Snyder, M. (2001) *Mol. Cell. Biol.* 21, 2449-2462.
- Buhr, T. L., Oved, S., Truesdell, G. M., Huang, C., Yarden, O. and Dickman, M. B. (1996) *Mol. Gen. Genet.* 251, 565-572.
- Chang, C., Kwok, S. F., Bleecker, A. B. and Meyerowitz, E. M. (1993) *Science* 262, 539-544.
- Colman-Lerner, A., Chin, T. E. and Brent, R. (2001) *Cell* 107, 739-750.
- Devroe, E., Erdjument-Bromage, H., Tempst, P. and Silver, P. A. (2004) *J. Biol. Chem.* 279, 24444-24451.
- Draper, J. and Scott, R. (1988) *Plant Genetic Transformation and Gene Expression, A Laboratory Manual*, 199-236, Blackwell, Oxford.
- Du, L. L. and Novick, P. (2002) *Mol. Biol. Cell.* 13, 503-514.
- Durrenberger, F. and Kronstad, J. (1999) *Mol. Gen. Genet.* 261, 281-289.
- Geng, W., He, B., Wang, M. and Adler, P. N. (2000) *Genetics* 156, 1817-1828.
- Gill, B. (2003). *Annual Wheat Newsletter* 49, 14.
- Hanks, S. K. and Hunter, T. (1995) In G. Hardie and S. K. Hanks, SK ed, *Protein Kinase Factsbook*. Academic press, London.
- Hanks, S. K., Quinn, A. M. and Hunter, T. (1988) *Science* 241, 42-52.
- Hanks, S. K. and Quinn, A. M. (1991) *Methods. Enzymol.* 200, 38-61.
- Harper, J. F., Binder, B. M. and Sussman, M. R. (1993) *Biochemistry* 32, 3282-3290.
- Hayashida, N., Mizoguchi, T., Yamaguchi-Shinozaki, K. and Shinozaki, K. (1992) *Gene* 121, 325-330.
- Hayashida, N., Mizoguchi, T. and Shinozaki, K. (1993) *Gene* 124, 251-255.
- Imai, T., Kurosaka, K. and Terachi, T. (2003) *Proc 7th Int Congress Plant Mol Biol*, 246.

- Innes, R. W. (2001) *Trends Plant Sci.* 6, 392-394.
- Laemmli, U. K. (1970) *Nature* 227, 680-685.
- Lauter, F. R., Marchfelder, U., Russo, V. E. A., Yamashiro, C. T., Yatzkan, E. and Yarden, O. (1998) *Fungal Genet. Biol.* 23, 300-310.
- Lawton, M. A., Yamamoto, R. T., Hanks, S. K. and Lamb, C. J. (1989) *Proc. Natl. Acad. Sci. USA.* 86, 3140-3144.
- Lin, X., Feng, X. H. and Watson, J. C. (1991) *Proc. Natl. Acad. Sci. USA.* 88, 6951-6955.
- Meskiene, I. and Hirt, H. (2000) *Plant Mol. Biol.* 42, 791-806.
- Millward, T., Cron, P. and Hemmings, B. A. (1995) *Proc. Natl. Acad. Sci. USA.* 92, 5022-5026.
- Millward, T. A., Heizmann, C. W., Schafer, B. W. and Hemmings, B. A. (1998) *Embo J.* 17, 5913-5922.
- Millward, T. A., Hess, D. and Hemmings, B. A. (1999) *J. Biol. Chem.* 274, 33847-33850.
- Mizoguchi, T., Hayashida, N., Yamaguchi-Shinozaki, K., Harada, H. and Shinozaki, K. (1992) *Plant Mol. Biol.* 18, 809-812.
- Racki, W. J., Becam, A. M., Nasr, F. and Herbert, C. J. (2000) *Embo J.* 19, 4524-4532.
- Smith, C. M., Shindyalov, I. N., Veretnik, S., Gribskov, M., Taylor, S. S., TenEyck, L. F. and Bourne, P. E. (1997) *Trends Biochem. Sci.* 22, 444-446.
- Stegert, M. R., Tamaskovic, R., Bichel, S. J., Hergovich, A. and Hemmings, B. A. (2004) *J. Biol. Chem.* 279, 23806-23812.
- Stone, J. M. and Walker, J. C. (1995) *Plant Physiol.* 108, 451-457.
- Swofford, D. L. (2002) PAUP*. Phylogenetic analysis using parsimony (* and other methods). Version 4. Sinauer Associates, Sunderland, Massachusetts.
- Tamaskovic, R., Bichsel, S. J. and Hemmings, B. A. (2003) *FEBS Lett.* 546, 73-80.
- Verde, F., Wiley, D. J. and Nurse, P. (1998) *Proc. Natl. Acad. Sci. USA.* 95, 7526-7531.
- Wang, D., Harper, J. F. and Gribskov, M. (2003) *Plant Physiol.* 132, 2152-2165.
- Weiss, E. L., Kurischko, C., Zhang, C., Shokat, K., Drubin, D. G. and Luca, F. C. (2002) *J. Cell Biol.* 158, 885-900.
- Yarden, O., Plamann, M., Ebbale, D. J. and Yanofsky, C. (1992) *Embo J.* 11, 2159-2166.
- Zallen, J. A., Peckol, E. L., Tobin, D. M. and Bargman, C. I. (2000) *Mol. Biol. Cell.* 11, 3177-3190.
- Zhang, S. H., Lawton, M. A., Hunter, T. and Lamb, C. J. (1994) *J. Biol. Chem.* 269, 17586-17592.
- Zhang, S. H., Broome, M. A., Lawton, M. A., Hunter, T. and Lamb, C. J. (1994) *J. Biol. Chem.* 269, 17593-17597.

高等植物の新規 Ser/Thr プロテインキナーゼ 遺伝子の単離と特徴付け

今 井 雄 大

要 旨

触媒ドメインのアミノ酸配列の相同性にもとづき、真核生物のプロテインキナーゼは‘AGC’、‘CMGC’、‘CaMK’、‘PTK’、及び‘その他’の大きく5つのファミリーに分類されている。このうち、環状ヌクレオチド依存性プロテインキナーゼ（PKA や PKG）、カルシウム - リン脂質依存性プロテインキナーゼ（PKC）やリボソーム S6 キナーゼで代表される‘AGC’グループのプロテインキナーゼは、数例の報告を除いて植物ではほとんど研究されていない。そこで本論文では、植物の‘AGC’グループの理解を深めるため、新規 Ser/Thr プロテインキナーゼ遺伝子を、ダイコン、シロイヌナズナ、及びパンコムギから単離し、それらの1次構造を分子生物学的方法により解析した。

本論文ではまず、ダイコンから5種類の cDNA クローン (*RsNdr1*, 2*a*, 2*b*, 3 及び 4) を得た。そのうち *RsNdr1*, 2*a*/2*b* の推定翻訳産物は、真核生物の Ser/Thr プロテインキナーゼの特徴である、保存された12個の触媒サブドメインをすべて含んでおり、これらの翻訳産物は機能を持つことが示唆された。*RsNdr1* の融合タンパク質を強発現する大腸菌ライセートはリン酸基転移活性を有していた。これらの *RsNdr* は、菌類の *cot-1* や動物の *Ndr* プロテインキナーゼと高い相同性を示した。また *RsNdr* はダイコンゲノム中で小さなマルチジーンファミリーを構成していることも示唆された。次に本論文では、植物 *Ndr* 遺伝子のマルチジーンファミリーとしての全容を明らかにするために、全ゲノム配列が決定されているシロイヌナズナから、*RsNdr* のホモログを探索した。8つのホモログが同定され、うち5つの遺伝子について全長 cDNA を得た。これらの cDNA の塩基配列を決定した結果、選択的スプライシングにより多様な mRNA が生み出されていることが明らかになり、植物 *Ndr* 遺伝子は、スプライシングを介した転写および活性制御を受けていることが推測された。本論文の最後の章では、EST データベース Komugi を利用し、パンコムギから Ser/Thr プロテインキナーゼをコードする3つの cDNA を単離し、それらの1次構造を RT-PCR と 5'RACE 法により決定した。

分子系統解析の結果、*RsNdr*, *AtNdr* 及び wheat *Ndr* は、AGC グループに属する他のプロテインキナーゼと明確に区別された。これらのプロテインキナーゼは、植物界、動物界、菌界にわたる真核生物に普遍的に存在する、プロテインキナーゼのスーパーファミリーの一員であることが示唆された。