博士論文

論文題目　Study on evolutionary genetics of early morphological evolution in land plants

（陸上植物形態の初期形態進化に関する進化遺伝学的解析）

氏　名　川井　絢子

Contents

Acknowledgments

I would like to express my sincere thanks to Professor Motomi Ito who has supported invaluably and guided me competently throughout my ten years.

I would like to express my cordial appreciation to Professor Mitsuyasu Hasebe for his highly efficient cooperation.

I am deeply grateful to Assistant Professor Tomomichi Fujita for giving me critical support, advice and valuable suggestions.

I am very much obliged to Dr. Yoichi Tanabe for his critical and warm support especially on *Selaginella*.

I am grateful to Professor Hiroyuki Sekimoto for giving me support and providing me with information especially on *Closterium*.

I appreciate to Dr. Seishirou Aoki, Dr. Tomoaki Nishiyama, Dr. Rumiko Kofuji and Dr. Yuji Hiwatashi for giving me essential support, valuable advice and comments.

I am grateful to Professor John Bowman and Dr. Sandra Floyd for providing me with information especially on *Marchantia*.

At last, I truly thank sincerely members of M. Ito laboratory and M. Hasebe laboratory.

And finally, I would like to express my great appreciation to my family.

Chapter 1.

General Introduction

Land plants originated from the charophyte algae about 450 million years ago (Langdale and Harrison, 2008), and began to diverge in the Silurian and Devonian periods. Land plants consist of two major lineages: Tracheophyta (vascular plants) and Bryophyta (non-vascular plants). Bryophytes, comprising hornworts, mosses, and liverworts, are the early diverging lineages of land plants. The haploid gametophyte is the most familiar part of the life cycle of bryophytes, and the diploid sporophyte is a short stalk without branching that has not been observed in other land plants. Lycophyta was presumed to be a monophyletic group, and one of the earliest diverged lineages of tracheophyta about 400 million years ago (Banks, 2009). Leaves (microphylls and megaphylls) and roots are thought to have evolved independently in the lycophyte and in other vascular plants. Judging from the evolutionary history of land plants, bryophytes and lycophytes have retained some specific features. Thus, these two lineages are worth investigating to obtain answers to the following questions. How were early land plants, bryophytes and/or lycophytes, established? Which genes participated in this evolution? What part in the gene sequence was modified during this evolution?

The evolution of morphology among each lineage in land plants is inferred from the comparison between similarities and differences in the genome structure, gene function, gene system, and so on (Floyd and Bowman, 2007). Many characteristics in plants are estimated to have originated through gene duplication (Moore et al., 2005; Banks et al., 2011). This increase in genes allows for amino acid changes and obtaining new functions, and lineage-specific gene duplication also contributes to lineage divergence and the origins of lineage-

specific features (Moore and Purugganan, 2005; Zhang, 2003). For example, the number of MADS-box genes, which are known to regulate some aspects of morphogenesis including floral organs, is 36 genes in *A. thaliana*, whereas *S. moellendorffii* and *P. patens* have only 3 and 6 genes, respectively. The number of AUX/IAA genes (a component of auxin signaling) is 29 genes in *A. thaliana*, whereas *S. moellendorffii* and *P. patens* have only 4 and 2 genes, respectively (Banks et al., 2011). It has been indicated that an increasing number of genes due to gene duplication leads to more complicated organization and function in land plans. Contrarily, early land plants that maintained a relatively uncomplicated organization and function might retain some peculiar gene functions and systems for specific features in each lineage, in addition to similar gene functions and systems. The function of orthologous genes might differ among angiosperms, bryophytes, and lycophytes as they differ so much in their development and morphology, following positive and purifying selection among paralogous genes (Moore and Purugganan, 2005; Zhang, 2003). Of course, not only different functions but also similar functions must be maintained between orthologous genes.

When we select a candidate gene for gene functional analysis from paralogous genes, the existence of many paralogous genes causes problems in the analysis of molecular biology. It is difficult to distinguish the function of each paralogous gene, and to judge which genes are essential for the study. For example, we cannot state that a target gene has the same function or no function (pseudogene) only from the information derived from the comparison of gene sequences among homologous genes (Ma et al., 1991; Kramer et al., 2004). Since the "molecular evolutionary clock" hypothesis was proposed by Zuckerkandl and Pauling (1962,

1965) and the "neutral theory of molecular evolution" hypothesis was proposed by Kimura (1968, 1983), molecular phylogenetics has progressed with advances in DNA sequencing and computational technologies, and the use of molecular phylogenetic trees has increased as a valuable tool for understanding the diversification of species and genomes (Kumar, 2005). In the study of angiosperms (Ferrario et al., 2003; Kim et al., 2005; Zimmermann and Werr, 2005), molecular phylogenetic analysis including functionally identified genes is a better method to indicate functional relationships between functionally identified genes and unknown genes among homologous genes. Ferrario et al. (2003) revealed that the orthologous gene that closely matched a functionally identified gene in *Arabidopsis* on the molecular phylogenetic tree had the same function as its orthologous gene by a complementation experiment in *Petunia*. In the study, the all genes isolated from *Petunia* were located within a clade including functionally identified genes of *Arabidopsis*, whereas the expressions of the paralogous genes belonging to other subgroups in the same clade were different. Therefore, it is possible to judge which gene would be important for functional analysis, because we could expect that genes in the same clade on a molecular phylogenetic tree will conserve a similar function. In contrast, the validity of this approach between long-distant lineages is uncertain, especially in morphogenetic genes, i.e. the genes to determine organ identity, pattern, and the form of plants. This is because the developmental and morphological differences are so great among bryophytes, lycophytes, and other vascular plants, and since the function of morphogenetic genes might be specialized in different plant lineages, even though the orthologous gene is in the same clade. We need to clarify that molecular phylogenetic analysis,

including functionally identified genes in angiosperms, is also a useful tool for the analysis of molecular biology in early land plants. For example, we will be able to predict that the genes in a clade consisting of angiosperms and bryophytes may retain a function common throughout the land plants and that there will be peculiar behavior in bryophytes if a clade consisting of only bryophytes is shown in the phylogenetic tree of homologous genes. Because we know little about the functions of genes in early diversified land plants and about whether the function of a gene identified in angiosperms is conserved between long-distant lineages (between bryophytes, lycophytes, and other vascular plants) or not (Floyd and Bowman, 2007; Singer and Ashton, 2007), the success of molecular phylogenetic analysis among long-distant lineages for gene functional analysis will be very useful in revealing functionally conserved genes in all land plants and the genes of peculiar to bryophytes and/or lycophytes from among many paralogous genes.

In this study, I will show that molecular phylogenetic analysis, including a gene whose function is known in some model angiosperms, is a better method to select candidate functional genes in early land plants. I demonstrated this approach using the *KNOX* gene family, the homeobox genes involved in shoot apical meristem (SAM) function, (Vollbrecht et al., 1991; Barton and Poethig, 1993) (Chapter 2) and the PIN gene family, i.e. the polar transporter of the plant signaling molecule (phytohormone) auxin (Okada et al., 1991; Gälweiler et al., 1998) (Chapters 3 and 4). I thought that these genes may be functionally conserved genes in all land plants, because SAM and auxin are critical to plant body establishment. At first, in order to reveal that molecular phylogenetic analysis is useful to

search for functionally conserved genes among angiosperms and lycophytes, I performed molecular phylogenetic analysis and gene expression analysis using the KNOX genes (Chapter 2). Next, in order to reveal that molecular phylogenetic analysis is also useful for the functional analysis of peculiar genes in early land plants, I performed molecular phylogenetic analysis including the bryophytes, lycophytes, gymnosperms, and angiosperms (Chapter 3) and gene expression analysis and functional analysis in bryophytes (Chapter 4) using the PIN genes.

Chapter 2.

The expression analysis of the *Selaginella* class 1 *KNOX* gene

2-1 Introduction

All land plants except bryophytes are thought to have evolved from a common dichotomously branching ancestor such as rhyniophytes identified in fossils. Regarding extant vascular plants, the lycophyte group (class Lycopsida) is the sister lineage of the euphyllophytes, which comprise seed plants and fern allies, including ferns, whisk ferns, and horsetails (Kenrick and Crane, 1997; Pryer et al., 2001). Based on the fossil record, the ancient tracheophyte group emerged approximately 400 million years ago (MYA) in the early Devonian, prior to the evolution of leaves and roots in vascular plants, and successfully dominated in the Carboniferous period (Stewart and Rothwell, 1993). The only three surviving modern-day plant orders of this ancient group, i.e., the Lycoposidales (clubmosses), Isoetales (quillworts), and Selaginellales (spikemosses), all of which are monophyletic, are placed in the lycophytes (Raubeson and Jansen, 1992; Kenrick and Crane, 1997). The remarkably common features of lycophytes are the distinguishable microphylls, which are poorly developed, single-veined leaves without a leaf gap, in contrast to the euphylls of ferns and seed plants.

The spikemosses, which are composed of the single genus *Selaginella*, are heterosporous and herbaceous lycophytes. Approximately 700 species belonging to this genus are distributed worldwide (Wochok and Sussex, 1975). In addition to microphylls, *Selaginella* have a unique rhizophore, which is defined as a root-producing, positively geotropic, leafless, and capless axis. For about a century, this axial organ has been a

histologically controversial structure, which had been interpreted as an aerial root or stem-like root (Webster and Jagels, 1977; Gifford and Foster, 1989). However, Imaichi and Kato (1989, 1991) revealed that the exogenous developmental process of the rhizophore is clearly distinguishable from the developmental process of the endogenous root. They demonstrated that the *Selaginella* rhizophore, like the lepidodendrid rhizomorph, could be fundamental axial organ coordinated with the roots and stems.

Molecular genetic and genomic studies have successfully generated much information on the genetic networks that control organ differentiation in higher plants; however, the ancient vascular plant lineage has been little studied at the molecular level, except for the construction of bacterial artificial chromosome (BAC) and expressed sequence tag (EST) libraries (Wang et al., 2005; Weng et al., 2005). The orthologous genes of key regulator genes that are involved in shoot apical meristem and/or lateral architecture differentiation in higher plants could reveal new characters concerning the unique *Selaginella* appendages.

As a first step in examining the molecular characteristics of the spikemoss pleurogeous organs, I focused on the *Knotted1*-like homeobox (*KNOX*) gene family, which was the first homeobox gene identified in plants (Hake et al., 1989; Vollbrecht et al., 1991). Members of the *KNOX* gene family are divided phylogenetically into two classes in land plants, class 1 and 2. Class 1 *KNOX* genes are typically expressed only in the shoot apical meristem, whereas class 2 *KNOX* genes have more diverse expression patterns (Bharathan et al., 1997; Reiser et al., 2000). In simple-leaved angiosperms such as maize, rice,

Arabidopsis, tobacco, and snapdragon, class 1 *KNOX* genes are expressed preferentially in shoot apical meristems and are negatively regulated by *ASYMMETRIC LEAVES1/ROUGHSHEATH2/PHANTASTICA* (*ARP*) genes in the domains in which leaves are expected to develop (Schneeberger et al., 1998; Byrne et al., 2000). In complex-leaved plants such as tomato, class 1 *KNOX* genes are expressed later in leaf development (Bharathan et al., 2002). The loss of the *Arabidopsis SHOOTMERISTEMLESS* (*STM*) gene, a well-characterized class 1 *KNOX* gene, induces a shoot meristem deficiency in the developmental process (Long et al., 1996). The fern *Ceratopteris* class 1 KNOX genes show expression patterns similar to those of their angiosperm counterparts, except in initial leaf formation (Sano et al., 2005). Thus, class 1 *KNOX* genes regulate shoot meristem and leaf formation. In a recent study of class 1 *KNOX* genes in *Selaginella kraussiana*, leaf development was regulated by the interaction of class 1 *KNOX* and *ARP* genes, similar to that in euphyllophytes (Harrison et al., 2005).

Here, I report the class 1 *KNOX* gene expression in *Selaginella uncinata*, with particular attention to the *Selaginella* rhizophore concept at the molecular level.

2-2 Materials and Methods

2-2-1 Cloning of *Selaginella KNOX* genes

Selaginella uncinata (Desv.) Spring was collected at the Koishikawa Garden of the University of Tokyo, Tokyo, Japan (Fig. 1). Total RNA extraction from various organs and 3' and 5' RACE were performed as described by Tanabe et al. (2003). The materials were ground in liquid nitrogen and dissolved completely in extraction buffer (4 M guanidine thiocyanate, 1 M ammonium thiocyanate, 1% lauryl sarcosine, 0.5% PVP, and 1% 2-mercaptoethanol). After three chloroform / isoamyl alcohol (24:1) extractions, the nucleic acids were precipitated in ethanol. The extracted RNA was purified by CTAB precipitation, followed by lithium chloride precipitation or ISOGEN-LS treatment (Nippon Gene).

Complementary DNA was synthesized from the total RNA according to the instructions of the 3' RACE system kit using SuperScript II reverse transcriptase and the universal primer (Invitrogen). Specific degenerate primers were designed to target class 1, KNd41 (5'-$\{CAU\}_4$ AAR AAR AAR GGI AAR YTN CC-3') , and all KNOX genes, KNd2 (5'-$\{CAU\}_4$ AAY AAY TGG TTY ATH AAY CAR MG-3'. The PCR conditions were an initial step at 94°C for 1 min; 35 cycles at 94°C for 1 min, 52°C for 1 min, and 72°C for 1.5 min; and a final step at 72°C for 5 min. The amplified products were cloned into the pAMP1 vector (Gibco-BRL) according to the manufacturer's instructions. The cloned products were sequenced using the DNA Analysis System (Beckman Coulter), using a Dye Terminator Cycle Sequencing kit (Beckman Coulter). The 5' regions of the cloned genes

were also isolated according to the instructions of the 5' RACE system kit (Invitrogen) and then sequenced.

2-2-2 Phylogenetic analysis

To construct a phylogenetic tree of *KNOX* genes, the amino acid sequences shown in Figure 2 were obtained from EMBL/DDBJ/GenBank DNA databases (Table 1) and aligned using the program Clustal W, version 1.8 (Thompson et al., 1994). The maximum likelihood (ML) distances were calculated using the ProtML program with the Jones, Taylor, and Thornton (JTT) model (Jones et al., 1992), and a neighbor-joining (NJ) tree was obtained using the program NJdist (Adachi and Hasegawa, 1992–1996). The trees were analyzed further with a local rearrangement search using the program ProtML to obtain the ML tree. Bootstrap values calculated using the resampling of estimated log-likelihoods (RELL) method (Hasegawa and Kishino, 1994) are indicated on nodes reconstructed using both the ML and NJ methods.

2-2-3 RT-PCR expression analysis

To perform the RT-PCR expression analysis, complementary DNA was synthesized from total RNA extracted from apical tips, including microphyll buds, internodes, microphylls, root tips, and rhizophore tips, as described above. The PCR conditions were an initial step at 95°C for 5 min; 40 cycles at 95°C for 15 s, 52°C for 15 s, and 72°C for 1 min; and a final step at 72°C for 7 min. The PCR amplification test was performed with

SuKNOX1-specific internal primers, SuKN1F1 (5'-ATCACCTGAGGTAGCCACAGTTGA-3') and SuKN1R1 (5'-AAAAGGATCAATCTCAAACTCCA-3'). The *S. uncinata* orthologous gene of the *ribosomal protein L6* gene (*SuRPL6*, DDBJ Accession No. AB521036), which was constitutively expressed in all tissues that I examined, was used as a quantifying control. *SuRPL6* was amplified by PCR with forward SuRPL6F1 (5'-CGTCAACCAGGCGTACGTGAT-3') and reverse SuRPL6R1 (5'-GGACTCGGTCTGCTCGATGA-3') primers. The partial fragments of *SuKNOX1* amplified by internal primers were cloned and used as specific probes for Southern hybridization. Southern hybridization was performed according to the instructions of the AlkPhos Direct Labelling and Detection System (GE Healthcare). The amplified products were fractionated on 1% (w/v) agarose gel and transferred to Hybond N+ nitrocellulose membranes (GE Healthcare) in an alkali transfer buffer (0.008 N NaOH). The membranes were hybridized at 55°C for 16 h in hybridization buffer (GE Healthcare).

2-2-4 *In situ* hybridization

Apical tips and parts producing rhizophores were collected in fixation buffer (4% paraformaldehyde and 50 mM sodium phosphate), dehydrated with tertiary butyl alcohol, and embedded in Paraplast Plus (Oxford Labware). Sections 8 μm thick were prepared. The procedure for in situ hybridization followed that of Jack et al. (1992) using the digoxigenin (DIG)-labeled *SuKNOX1*-specific RNA probe and a DIG RNA labeling kit (Roche), following cloning of the 565 bp of partial *SuKNOX1* cDNA outside of the well-conserved

ELK [glutamic acid (E), leucine (L), and lysine (K)] and homeo domains. The sense strand of the *SuKNOX1* mRNA was used as a negative control. The incorporation of DIG-labeled uridine triphosphate (UTP) into RNA was accompanied by the synthesis of RNA with T7 or SP6 RNA polymerase. Hybridization was performed at 46 °C for 16 h in hybridization buffer (100 mM NaCl, 10 mM Tris-HCl, 10 mM sodium phosphate buffer [pH 6.8], 5 mM EDTA, 50% formamide, 1 mM DTT, 1 mg/ml tRNA, 10% dextran sulfate, and 0.5 U/ml RNase inhibitor). After the washing steps, the signal was detected using a DIG detection kit (Roche).

2-3 Results

2-3-1 Cloning of *Selaginella KNOX* genes and phylogenetic analysis

I examined the relatively large spikemoss *Selaginella uncinata* because this species provides advantages for histological studies. The *Selaginella KNOX* cDNA, named *SuKNOX1* (DDBJ Accession No. AB288208, Fig. 3A), was isolated from *S. uncinata* cDNA using 3' and 5' RACE methods. The *KNOX* cDNA clone was amplified successfully using the described PCR conditions, using the degenerate primer that targets the class 1 *KNOX*-specific sequence KNd41. The deduced amino acid sequence of the *SuKNOX1* contained a MEINOX domain [the amalgam of the animal Myeloid ecotropic viral integration site 1 (MEIS) and plant KNOX domain], ELK domain, and TALE (three amino acid loop extension) type homeo domain with three additional amino acid residues between helix 1 and 2 (Bertolino et al., 1995), all of which are hallmarks of *KNOX* genes (Fig. 3 and Fig. 4). In addition, I successfully amplified a class 2 *KNOX* gene, named *SuKNOX2* (DDBJ Accession No. AB288209, Fig. 3B), using the degenerate primer KNd2, which corresponds to the best-conserved homeobox region. The KNd2 primer can be used to clone both class 1 and 2 *KNOX* genes, but no class 1 *KNOX* genes were amplified, except *SuKNOX1*.

To address the evolutionary relationship between *SuKNOX1* and class 1 *KNOX* genes of land plants, I constructed an ML gene tree that contained 33 *KNOX* genes from a wide range of green plants, including seed plants, ferns, spikemosses, mosses, and ulvophyceans. Metazoan genes were used as an outgroup (Burglin, 1998). The tree was

constructed using 83 amino acid residues covering the ELK and homeo domains (Fig. 2 and Fig. 4, underlined). No reliable alignment was obtained when the set included the MEINOX domain because of the low sequence similarity between metazoan and plant genes; thus, the MEINOX domain was excluded from the alignment data for the phylogenetic analysis. The gene tree indicates that *S. uncinata SuKNOX1* is clearly placed within the class 1 clade and that *SuKNOX2* was placed within the class 2 clade, with high statistical confidence (Fig. 5). Relationships among *S. uncinata SuKNOX1*, *S. kraussiana SkKNOX1* and *2* (Harrison et al., 2005), and the class 1 genes of other vascular plants were not resolved reliably. The gene tree predicts that *S. uncinata* likely possesses another class 1 *KNOX* gene(s) in its genome.

2-3-2 Expression of *SuKNOX1* mRNA

Selaginella has creeping or ascendant stems with leafy microphylls, which are produced by a dome-shaped shoot apex having an apical cell. Rhizophores are initiated exogenously at the junctions of branching stems and endogenously give rise to roots (Imaichi and Kato, 1991). The expression pattern of *SuKNOX1* was investigated using *in situ* hybridization. The longitudinal section shows that *SuKNOX1* mRNA was clearly localized around the shoot apical meristem, including the shoot apical cell (Fig. 6A). No *SuKNOX1* expression was detected in microphylls, although there was weak detection in juvenile microphyll primordia (Fig. 6A). To examine *SuKNOX1* expression in the rhizophores, I prepared longitudinal sections containing the junctions of branching stems (Fig. 6C). These showed clear *SuKNOX1* mRNA accumulation in the outer layer around the

young rhizophore apex and weaker accumulation in the associated vascular bundles. *SuKNOX1* expression in the rhizophore apex was confirmed in another section (Fig. 6E).

I also examined the *SuKNOX1* expression pattern using RT-PCR (Fig. 6F). *SuKNOX1*-specific amplification was detected in the shoot apical tips of microphylls and in rhizophore tips, whereas it was generally not detected in the internodes, microphylls, or root tips. *SuKNOX1* amplification was infrequently detected in the internodes and root tips (data not shown). These results were in accordance with those of in situ hybridization. Unfortunately, I could not examine *SuKNOX1* expression patterns in the gametophyte because I did not have sufficient samples for mRNA extraction. Further assessment of *KNOX* gene expression in the gametophyte will be necessary to infer the comprehensive functions of these genes in the *Selaginella* life cycle.

2-4 Discussion

I characterized the *SuKNOX1* cDNA from the spikemoss *Selaginella uncinata*. Based on the gene tree, *SuKNOX1* is clearly included within the class 1 *KNOX* gene cluster containing *Arabidopsis STM*, *Knotted1-like from Arabidopsis thaliana 1* (*KNAT1*), and *2* (*KNAT2/ATK1*) genes. As reported by Serikawa and Mandoli (1999), *Acetablaria acetablum AaKNOX1*, a *KNOX* gene from an ulvophycean green alga, branched out before the divergence of class 1 and 2 genes. The moss *Physcomitrella* and the fern *Ceratopteris* possess both class 1 and 2 genes (Champagne and Ashton, 2001; Sano et al., 2005), indicating the appearance of the two large *KNOX* gene groups prior to the terrestrialization of green plants approximately 470 MYA (Kenrick and Crane, 1997). Therefore, further studies of *KNOX* genes from charophycean algae, the closest relatives to land plants, will shed light on the ancient split event.

Most class 1 *KNOX* genes exhibit similar expression in shoot apical meristems despite differences in the meristem structures among the diversified euphyllophytes, which include seed plants and ferns. In the euphyllophytes, multicellular and unicellular meristems are regulated similarly by class 1 *KNOX* genes during development (Sano et al., 2005). Lycophytes are the most ancient of the modern vascular plants, branching at approximately 400 MYA, and possess unicellular meristems that are thought to have originated independently in this lineage (Pryer et al., 2001; Sano et al., 2005). With respect to leaf formation, the expression of *KNOX* genes in seed plants is negatively regulated by *ARP*

genes during leaf development, with notable exceptions observed in the compound leaves of tomato (Hareven et al., 1996; Janssen et al., 1998), whereas class 1 *KNOX* genes are expressed in the primordia and immature leaves in the ferns *Ceratopteris* and *Anagramma*, suggesting differences in the leaf developmental mechanisms between seed plants and ferns (Sano et al., 2005). According to paleobotanical evidence and molecular phylogenetic reports of early land plants, seed plant leaves, fern fronds, and spikemoss microphylls originated in parallel (Gifford and Foster, 1989; Stewart and Rothwell, 1993; Kenrick and Crane, 1997; Pryer et al., 2001). Recently, Harrison et al. (2005) reported the detailed expression patterns of the *Selaginella kraussiana* class 1 *KNOX* genes, *SkKNOX1* and *2*, in the shoot apical meristem and in leaf formation. *SkKNOX1* is specifically expressed in the shoot apical meristem, whereas *SkKNOX2* is preferentially expressed in the internode regions; neither is expressed in microphyll primordia. I found clear expression of the *S. uncinata SuKNOX1* in the shoot apex, supporting the evidence that class 1 *KNOX* genes have conserved functions in the primitive sporophytic apices. Reportedly, *SkKNOX1* and *2* originally diverged in the *Selaginella* lineage (Harrison et al., 2005). However, according to my gene trees based on ML and NJ methods, the phylogenetic relationships among *S. uncinata SuKNOX1*, *S. kraussiana SkKNOX1* and *2*, and other vascular plant class 1 genes were not resolved fully, suggesting that the common ancestor of lycophytes and euphyllophytes, a rhyniophyte, possessed at least two or more class 1 genes in the genome.

I also detected *SuKNOX1* expression in the axial apex of the unique spikemoss rhizophore, but not in the root tips. Meanwhile, the expression patterns of *SkKNOX1* and *2*

were not reported by Harrison et al (2005). My expression data suggest that the rhizophore has developmental mechanisms distinct from those of the root. This finding supports the "rhizophore concept" at the molecular level, which postulates that the rhizophore is a fundamental organ distinguishable from the root (Imaichi and Kato, 1989; 1991). Further molecular studies of other developmentally critical regulator homologous genes such as *homeodomain-leucine zipper* (*HD-Zip*) genes and *NO APICAL MERISTEM* (*NAM*)/ *ATAF*/ *CUP-SHAPED COTYLEDON* (*CUC*) (*NAC*) genes will reveal additional characters of *Selaginella* appendages. It was found that *Selaginella class III HD-Zip* gene expression predicts organ initiation site, similar to that of *Arabidopsis* orthologous genes; however, there have been no reports in rhizophores (Floyd et al., 2006; Prigge et al., 2006). A BAC library containing ten genome-equivalents and ESTs containing 1301 non-redundant clones have recently been constructed for *S. moellendorffii* (Wang et al., 2005; Weng et al., 2005). Further studies using genomic resources to construct libraries for other large-sized *Selaginella* species, as well as studies using microarray techniques, will reveal the gene expression profiles of spikemoss pleurogenous architectures, which will aid in depicting evolutionary scenarios for early land plants.

Chapter 3.

The cloning and characterization of *PIN* genes from a chalophyte alga, a liverwort, a moss, and two gymnosperms, and the phylogenetic analysis of land plant *PIN* genes

3-1 Introduction

The plant hormone (phytohormone) auxin plays a critical role in various aspects of plant growth and development, such as shoot elongation, phyllotaxy, root development, embryo patterning, cell elongation, vascular tissue differentiation, and tropic responses to light and gravity, by forming local concentration gradient by polar auxin transport (Estelle, 1992; Woodward and Bartel, 2005). In *Arabidopsis thaliana*, *pin1* mutants are characterized by pin-shaped inflorescence and it became clear that PIN proteins play a key role in facilitation of auxin efflux from cells (Okada et al., 1991; Gälweiler et al., 1998). Molecular genetic studies in the *A. thaliana* and other species have shown auxin transport systems comprising the PINFORMED (PIN) auxin efflux carriers, ATP-BINDING CASSETTE GROUP B/P-GLYCOPROTEIN/MULTIDRUG RESISTANCE (ABCB/PGP/MDR) auxin transporters, and AUXIN-RESISTANT1 (AUX1)/ LIKE-AUX1 (AUX1/LAX) auxin influx carriers (Zazímalová et al., 2010). Above all, polar auxin transport and auxin gradients are mainly mediated by PIN family that exhibits asymmetrical positioning, while MDR/ABCB/PGPs proteins are found on all sides of the membrane (Zazímalová et al., 2010), and it was suggested that AUX1/LAX proteins that exhibit asymmetrical positioning at the plasma membrane act to buffer the PIN-mediated patterning mechanism against environmental or developmental influences (Reinhardt et al., 2003; Bainbridge et al., 2008).

The *A. thaliana* genome contains eight *PIN* genes (*AtPIN1-AtPIN8*) and the AtPIN proteins have specific patterns of tissue localization and specific roles in development

(Gälweiler et al., 1998; Luschnig et al., 1998; Müller et al., 1998; Friml et al., 2002a; 2002b; 2003; Benková et al., 2003; Vieten et al., 2005; Krecek et al., 2009; Mravec et al., 2009; Keuskamp et al., 2010). The PIN family proteins were classified into PIN1-type (PIN1, 2, 3, 4, and 7) and PIN5-type (PIN5, 6, and 8). The former were localized to the plasma membrane (PM) and the latter to the endoplasmic reticulum (ER) (Mravec et al., 2009). Moreover, localization of PIN1-type protein is various and dynamic: PIN1-type protein is internalized and localized at basal, apical, and/or lateral side on PM. For example, AtPIN1 was polarly localized to the basal side of xylem parenchyma cells of inflorescence axes (Gälweiler et al., 1998), AtPIN2 was polarly localized to the apical side of root epidermal cells (Müller et al., 1998), AtPIN3 was polarly localized to the lateral side of root pericycle cells and the polar localization of AtPIN3 in columella cells was changed from symmetric to asymmetric by a change in the gravity (Friml et al., 2002b), AtPIN4 was localized to surface of the hypophysis cell, basal cell, and provascular cells of a globular embryo (Friml et al., 2002a), and the polar localization of AtPIN1 and AtPIN7 in embryogenesis were switched during embryogenesis (Friml et al., 2003). The predicted structure of all PINs contains two hydrophobic domains separated by a hydrophilic loop domain. The hydrophilic loop domain, which is the crucial region related to the function of transmembrane proteins, exhibits more variability than hydrophobic domains. The hydrophilic loop domain contains the following motifs: the internalization motif NPXXY (Zazímalová et al., 2007; Mravec et al., 2009), the cluster of glycosylation and two phosphorylation sites (Gly/2P cluster) (Zazímalová et al., 2007), three TPRXS (N/S) motifs phosphorylated by PINOID (PID) protein Ser/Thr kinase (Dhonukshe et

al., 2010; Huang et al., 2010), and other phosphorylation sites (Michniewicz et al., 2007; Zhang et al., 2010). Changes of amino acid sequences in the hydrophilic loop domain are successful in getting different function containing various localizations of PIN protein, and hence change the morphology: Mutational analysis of tyrosine-based NPNSY motif in AtPIN1 indicated that the mutated AtPIN1:GFP with NSLSL motif was localized to ER (Mravec et al., 2009); The phosphorylation at three TPRXS (N/S) motifs of the AtPIN2 was reduced or abolished when the middle serine residue (Ser) in three TPRXS (N/S) motifs were replaced with alanine residues (Ala). Thus, the localization of mutated AtPIN2 in root cells was changed and the gravitropic root growth exhibited abnormality like the *atpin2* mutant (Dhonukshe et al., 2010); Loss-of-phosphorylation AtPIN1:GFP with altered three TPRXS (N/S) motifs induced inflorescence, flower, and embryo defects correlating with their mislocalization (Huang et al., 2010); And the mutated AtPIN1:GFP caused defective phyllotaxis and floral morphology when Ser337 and Thr340 were both converted into Ala or Asp residue. Furthermore, it has been exhibited that the defects in the subcellular localization of phosphomimic PIN1:GFP(Asp) were observed in the embryo and root though the nonphosphorylation PIN1:GFP(Ala) localization was not changed, and that the AtPIN1:GFP(Asp) was functionally replaced AtPIN2 in its endogenous expression domain (Zhang et al., 2010). These revealed that PIN motifs contribute to functional diversification in PIN family and plant morphology.

Several outcomes occur on duplicated genes: subfunctionalization, neofunctionalization, and pseudogenization (Zhang, 2003; Moore and Purugganan, 2005;

Floyd and Bowman, 2007). The *AtPIN*s may have evolved from a common single ancestral sequence and *AtPIN1* gene is more closely related to *AtPIN3*, *AtPIN4*, and *AtPIN7* genes than *AtPIN2* gene (Paponov et al., 2005). The phylogenetic structure of *PIN* gene family has been broadly conserved in eudicots and more divergent in grasses. It is likely that *PIN* diversity is related to the plant characters, for example, development of root system or phyllotaxy (Paponov et al., 2005).

Determining the phylogenetic relationships in the *PIN* gene family among plants is necessary step for elucidating the evolution of PIN gene family, which has been thought to be important for morphogenesis. The precise phylogeny of *PIN* gene family including relationships among duplicated genes and their functional diversity is still not clear (Paponov et al., 2005; Zazímalová et al., 2007; Krecek et al., 2009). And, the origin of *PIN* gene in the evolutionary history of plants is also unknown (Paponov et al., 2005; Zazímalová et al., 2007; Krecek et al., 2009). To determining the phylogenetic relationships in the *PIN* gene family among plants, we need to find the *PIN* gene family in charophyte alga. *PIN* genes are observed in almost all of the land plants. However, *PIN-like* genes, which are similar in amino acids sequence and not identified as the *PIN* gene, are found in prokaryotes and eukaryotes on NCBI database (http://www.ncbi.nlm.nih.gov/), of course in *A. thaliana*, whereas the *PIN* orthologous gene was not found in the whole-genome analysis of chlorophyta (*Chlamydomonas reinhardtii*, *Ostreococcus tauri*, *Micromonas pusilla*, *Chlorella vulgaris*, and *Volvox carteri*) (Krecek et al., 2009). In a latest study, a putative *PIN* orthologous gene lacking 5' end has been found from EST libraries in *Spirogyra pratensis* (De Smet et al.,

2011), but this was not enough for phylogenetic analysis.

To elucidate the evolution of the *PIN* gene family, I performed exhaustive analysis of *PIN* genes in land plants and a charophyte alga. At the beginning, I found a new variant sequence of *PpPIND* from the moss, *Physcomitrella patens*. I isolated *PIN* genes from plants in the five major lineages: *Pinus thunbergii* (Coniferophyta), *Ginkgo biloba* (Ginkgophyta), *Physcomitrella patens* (Mosses), *Marchantia polymorpha* (Liverworts), and *Closterium peracerosum-strigosum-littorale complex* (Charophyta). For the genes from *P. patens* and *C. psl.* complex, I have determined the exon-intron structures for phylogenetic inference. Then, I describe phylogenetic relationships of the *PIN* gene family of land plants with the phylogenetic tree using *C. psl.* complex as an outgroup gene and with the comparison of the amino acid sequence motifs.

3-2 Materials and Methods

3-2-1 Plant materials

***C. psl.* complex** The strains of heterothallic *Closterium peracerosum-strigosum-littorale complex* (*C. psl.* complex), unicellular charophyte alga, were NIES-67 (mt^+) and NIES-68 (mt^-), which were obtained from the National Institute for Environmental Studies, Ibaraki, Japan. The respective vegetative cells (mt^+-V and mt^--V) were obtained from cultures grown in nitrogen-supplemented medium (C medium; http://www.nies.go.jp/biology/mcc/home.htm), as described previously (Sekimoto et al., 1990). The sexual reproduction of *C. psl.* complex was induced as described previously (Sekimoto et al., 2006). At 24 and 72 hours after the mating reaction began, the cells were harvested, frozen in liquid nitrogen, and stored at –80°C (mix-24 h and mix-72 h, respectively).

P. patens *Physcomitrella patens* (Hedw.) Bruch & Schimp subsp. *patens* originally collected in Gransden Wood, Huntingdonshire, U. K. (Ashton and Cove, 1977), was used in this study. It was cultured in plant incubator on the same condition as follows.

Physcomitrella patens (*P. patens*) was grown on medium based on BCDAT medium, which contains BCD medium [1 mM $MgSO_4$, 10 mM KNO_3, 45 μM $FeSO_4$, 1.8 mM KH_2PO_4 (adjusted pH to 6.5 with 4 M KOH), 1 mM $CaCl_2$], 5 mM ammonium tartrate, and trace elements (Knight et al., 1988). The trace elements included 0.22 μM $CuSO_4$, 0.19 μM $ZnSO_4$,

10 μM H_3BO_4, 0.10 μM Na_2MoO_4, 2 μM $MnCl_2$, 0.23 μM $CoCL_2$ and 0.17 μM KI (Nishiyama et al., 2000). These media were solidified with 0.8 % (w/v) agar (01028-85, NACALAI TESQUE, Kyoto, Japan) in 9 cm Petri dishes. The solidified medium was covered with a layer of cellophane (TAIKO, Hyogo, Japan) to facilitate collection of the moss from the medium (Grimsley et al., 1977). Polytron homogenizer (Kinematica, Littau, Switzerland) or a homogenizer (AM-2, Nippon Seiki, Niigata, JAPAN) were used to homogenize the moss with sterilized water, and approximately 5 ml of homogenized moss was applied to one solid medium. The medium was put in the plant incubator under continued light condition at 25 °C. BCDAT and BCD media were used for gametophyte development.

Gymnosperms Young male reproductive organs of *Pinus thunbergii* and 1week-old seedling plant of *Ginkgo biloba* were collected at University of Tokyo, Japan, in April. All materials were frozen in liquid nitrogen.

3-2-2 Extraction of total RNA and preparation of cDNA

***C. psl.* complex** Total RNA was isolated using the Trizol plus RNA purification kit (invitrogen), according to the manufacturer's instructions. Total RNAs derived from cells in various stages of sexual reproduction (1, 2, 4, 6, 8, 12, 16, 20, 24 h after mixing) were isolated, mingled and used as primary materials to construct full length cDNA library. Various expressed sequence tag (EST) related sexual reproduction were obtained (Sekimoto et al., unpublished data) and used for similarity search by blastn program. Furthermore, the

Complementary DNA (cDNA) was synthesized from total RNA that was isolated from respective cells at two stages of sexual reproduction (mix-24 h and mix-72 h) using the TaKaRa RNA LA PCR kit (AMV) version 1.1 (TaKaRa Bio) and oligo dT - adapter primer, according to the manufacturer's instructions.

P. patens Total RNA was extracted using the RNeasy Plant Mini Kit (QIAGEN K. K., Tokyo, Japan) from chloronemata and gametophore, and treated with the RNase-free DNase I (Roche) according to the manufacturer's instructions. Single stranded complementary DNA (cDNA) was synthesized from DNase I treated total RNA using the SuperScript II (invitrogen) reverse transcriptase and the Oligo dT-Adaptor Primer (Gibco-BRL, Rockvill, MD, USA).

Gymnosperms Total RNA extraction from gymnosperms for 3' and 5' RACE were performed as described by Shigyo and Ito (2004). Total RNAs of *Pinus thunbergii* and *Ginkgo biloba* were extracted from male cones approximately 0.5–1.0 cm long and young seedling approximately 1.5 cm, respectively. Complementary DNA was synthesized from DNase I treated total RNA (1 μg) using SuperScript III (invitrogen) reverse transcriptase and an oligo dT adapter primer (3'AP: 5'-GGCCACGCGTCGACTAGTACT-T17-3'), according to the manufacturer's instructions.

3-2-3 Cloning of the *PIN* genes

***C. psl.* complex** cDNAs containing the entire coding regions of *CpslcPIN* were cloned using the partial sequences found in the EST database as follows. A cDNA of *CpslcPIN* was amplified by PCR using the *CpslcPIN*-specific primers: CpPIN-F2 (5'-TCATCAGCGTCGATCATTAGCGTCGAGC-3'), CpPIN-F3 (5'-AGCGTCGAGCATTAGTGTTACACTG-3'), CpPIN-R1 (5'-ACCAAAGAACGTCCTTAACGGAATAGCCTCC-3') and CpPIN-R2 (5'-ACGGAATAGCCTCCTAGTTGACTG-3'), designed using sequence information from EST data (Sekimoto et al., unpublished data). The PCR products were inserted into pGEM-T Easy Vector (Promega, Madison, WI, USA).

To determine the exon–intron junctions, I searched for whole genome sequences data (Sekimoto et al., unpublished data) and compared with the cDNA of *CpslcPIN* (Fig. 7).

Bryophytes The partial sequences of *PhypaPIND* were obtained from the *Physcomitrella patens* EST database, clone MS481.seq.1 (PHYSCObase: http://moss.nibb.ac.jp/). A cDNA of *PhypaPIND* was amplified by PCR using the *PhypaPIND*-specific primers: PpPIN-f2 (5'-GGGGTTCTTCGAGCTCGACTTAAATTC-3') and PpPIN-r2 (5'-GAAGTCCTTGCAAGTGATTTATGGTTA-3') and sequenced it.

To determine the exon–intron junctions, I searched for whole-genome sequences data (NCBI: http://www.ncbi.nlm.nih.gov/) and compared with the cDNA of *PhypaPIND*. Moreover, *PhypaPIND* genomic DNA was cloned by PCR with a high fidelity enzyme using the *PhypaPIND*-specific primers: PpPIN-5'f1 (5'-

GGTCGAGAGAGATGAAGTATGTAGGGTTTTACC-3') and PpPIN-3'r1 (5'-AGTGGAGAAGCAGAGACTCAGCGGACC-3') in Chapter 4 (Fig. 8).

PIN cDNA sequences of *Marchantia polymorpha* were provided by Prof. John Bowman and Dr. Sandra Floyd at Monash University.

Gymnosperms To clone the *PIN* genes of the two gymnosperms, I proceeded as follows. The partial PIN coding sequence was amplified by PCR using cDNA template obtained and *PIN*-specific degenerate primers. I synthesized four *PIN*-specific degenerate primers based on a comparison of the amino acid sequences of previously reported *PIN* genes: PIN-Nf1 (5'-TN CCN YTN TAY GTN GC-3'), PIN-Nf2 (5'-$\{CAU\}_4$ AAY RAY CCN TAY RMN ATG AA-3'), PIN-Cr1 (5'- TC NYK NGC RAA NAC RAA-3'), and PIN-Cr2 (5'-$\{CUA\}_4$ TG NGG NAR NGC NGC YTG-3') where K, M, N, R, and Y follow the IUPAC code. The 5' and 3' terminal sequences of the PIN gene were determined by rapid amplification of cDNA ends (RACE) experiments. To obtain the full-length cDNA of the *PIN* gene, we synthesized gene-specific primers based on regions located close to 5' ends of each cDNA according to the nucleotide sequences of the RACE products. The cDNA of *PinthPIN1* and *GinbiPIN1* was amplified by PCR using the universal adaptor primer (UAP: 5'-CUACUACUACUAGGCCACGCGTCGACTAGT-3') and the *PinthPIN1* and *GinbiPIN1*-specific primers, respectively: PtPIN-F1 (5'- TGCCAAGCTATCAATAGACCAGAGG-3') and PtPIN-F2 (5'- TCCTGATTTGTATGCCGCGACTGTC-3'), and GbPIN-F1 (5'-ACTGTAGGCTAGAGTATCCCAATATCATAG-3') and GbPIN-F2 (5'-

TAGAGTATCCCAATATCATAGATTAGCTGC-3'). The PCR products were inserted into pGEM-T Easy Vector (Promega, Madison, WI, USA), and were sequenced (Fig. 9 and Fig. 10).

3-2-4 cDNA sequence determination

The cloned products were sequenced using a Dye Terminator Cycle Sequencing kit and a CEQ 8000 Genetic Analysis System (Beckman Coulter). Sequencing was conducted according to the manufacturer's instructions.

3-2-5 Protein prediction

The transmembrane domains were estimated using TMHMM Server v. 2.0: http://www.cbs.dtu.dk/services/TMHMM/.

Motifs were predicted according to the PROSITE database: http://prosite.expasy.org/, and were cited from Michniewicz et al. (2007); Zazímalová et al. (2007); Mravec et al. (2009); Dhonukshe et al. (2010); Huang et al. (2010); and Zhang et al. (2010).

3-2-6 Phylogenetic analysis

To construct a phylogenetic tree of *PIN* genes, the amino acid sequences showed in Figure 11 were obtained from the Kyoto Encyclopedia of Genes and Genomes (KEGG, http://www.genome.jp/kegg/) database and from Miyashita et al. (2010) (Table 2). Sequences

were aligned using the program MAFFT version 5, method L-INS-i (Katoh and Toh, 2008). The maximum likelihood (ML) distances were calculated using the MEGA 5 software (Tamura et al., 2011) with the Whelan and Goldman (WAG) model (Whelan and Goldman, 2001) of the amino acid replacements assuming a proportion of the invariant sites and four gamma-distributed rates (WAG + I + G4 model) of amino-acid sequence evolution with all sites containing gaps. ML heuristic method was the nearest-neighbor-interchange (NNI), and initial tree for ML was made automatically. The robustness of the trees was assessed by bootstrap analysis (100 replicates). Some sequences in database were excluded from the phylogenetic analysis because of insertions, deletions, and/or atypical length. Bayesian inference (BI) analysis was performed using MrBayes version 3.12 (Ronquist and Huelsenbeck, 2003) with Metropolis Coupled Markov chain Monte Carlo (MCMC) with all sites containing gaps used as the missing data. Trees were sampled every 100 generations for a total of 1,000,000 generation MCMC analysis, and the first 2,500 trees (250,000 generations) were discarded as burn-in. The Convergence of the MCMC procedure was assessed by calculating the effective sampling size (ESS) with the program Tracer version 1.3 (Rambaut and Drummond, 2004). An ESS above 100 was considered significant. Posterior probabilities of nodes were estimated based on 50% majority rule consensus of the remaining 7,500 trees.

3-3 Results

3-3-1 Isolation of *PIN* genes from a charophyte, a liverwort, a moss, and two gymnosperms

By searching the EST database of *Closterium peracerosum-strigosum-littorale complex* (*C. psl.* complex) (Sekimoto et al., unpublished data), a unicellular charophyte alga, the *PIN* homologous gene sequences were obtained. A PIN homolog was named *C. psl.* complex *PIN-FORMED* gene (*CpslcPIN*, DDBJ accession number AB748929). The *CpslcPIN* cDNA was amplified from each sample of 24 and 72 hours after the mating reaction using the *CpslcPIN*-specific primers. The full-length *CpslcPIN* cDNA encodes a putative protein of 762 amino acids residues (Fig. 7).

PIN gene sequences of *Marchantia polymorpha* (liverwort), were provided from Professor John Bowman and Dr. Sandra Floyd at Monash University. *MarpoPIN1*, *MarpoPIN2*, *MarpoPIN3*, and *MarpoPIN4* have the open reading flame of 2160, 1230, 1377, and 1548 bp encoding 719, 409, 458, and 515 amino acids residues, respectively.

PhypaPIND cDNA (DDBJ accession number AB748930) has the open reading flame of 1290 bp encoding 429 amino acids residues, which has 71 amino acids residues longer than already known PhypaPIND (NCBI accession number XM_001765711; Rensing et al., 2008) in central hydrophilic loop region (Fig. 8). I could not confirm existence of precedent *PpPIND* (NCBI accession number XM_001765711; Rensing et al., 2008) using PCR in this study. This indicated that *PhypaPIND* (DDBJ accession number AB748930)

isolated in this study is probably the complete and correct sequence of the cDNA.

Candidate PIN homologs of *P. thunbergii* and *G. biloba* (Gymnosperms), were obtained from young male reproductive organs and one week-old seedling plant, respectively, using the 3' and 5' RACE systems with PIN-specific degenerate primers based on the alignment of *PIN* genes in *P. patens*, *A. thaliana*, and *O.sativa*. Four of these degenerate primers located in the amino and carboxy-terminal hydrophobic domain in predicted structure of PIN protein. Using these primers, I obtained partial fragments of the putative *P. thunbergii* PIN homolog and the putative *G. biloba* PIN homolog. The remaining parts of these DNA fragments were cloned using 3' and 5' RACE systems. To confirm that the 3' and 5' RACE products were from the same cDNA, I synthesized gene-specific primers based on regions located close to 3' and 5' ends of each cDNA according to the nucleotide sequences of the RACE products. The PIN homologs were named *P. thunbergii PIN-FORMED* gene (*PinthPIN*, DDBJ accession number AB748932) and *G. biloba PIN-FORMED* (*GinbiPIN*, DDBJ accession number AB748931). The full-length *PinthPIN* and *GinbiPIN* cDNA encodes a putative protein with 695 and 707 amino acid residues, respectively (Fig. 9 and Fig. 10).

3-3-2 Characterization of *PIN* genes from a charophyte, a liverwort, a moss, and two gymnosperms

PIN protein prediction The PIN proteins are known to have two hydrophobic domains that are separated by a hydrophilic loop domain (Zazímalová et al., 2007). The transmembrane domains were estimated using TMHMM2 server. The predicted

amino acid sequences of *CpslcPIN*, *PhypaPIND*, *MarpoPIN1*, *MarpoPIN2*, *MarpoPIN3*, *MarpoPIN4*, *PinthPIN* and *GinbiPIN* cDNA showed the three-domain topology typically found in other PIN proteins (Fig. 12). Hydrophilic loop domains of CpslcPIN, MarpoPIN1, PinthPIN, and GinbiPIN is longer than most of other PINs and they are more similar to the *P. patens* PINA-C (Fig. 12). And a hydrophilic loop domain of PhypaPIND, MarpoPIN2, MarpoPIN3, and MarpoPIN4 is intermediate length (middle-looped PIN) that is neither short-looped PIN (ArathPIN5, 8) nor long-looped PIN (ArathPIN1-4, 6, and 7) (Fig. 12). In the hydrophilic loop domain, the internalization motif NPXXY (Barak et al., 1994; Zazímalová et al., 2007; Mravec et al., 2009), three TPRXS (N/S) phosphorylation motifs (Dhonukshe et al., 2010; Huang et al., 2010), and a cluster of motifs Gly/2P (Zazímalová et al., 2007) were found in MarpoPIN1, PinthPIN and GinbiPIN amino acid sequences, whereas these motifs were not found in CpslcPIN and PhypaPIND, MarpoPIN2, MarpoPIN3, and MarpoPIN4 (Fig. 11). GinbiPIN contains two amino acids serine and threonine that are phosphorylated by PID in ArathPIN1 (Ser337 and Thr 340) (Michniewicz et al., 2007; Zhang et al., 2010) (Fig. 11), which suggests that the function of *GinbiPIN* gene may be similar to *ArathPIN1* gene.

Exon-intron structures analysis of the *CpslcPIN* and *PhypaPIND* genes It has been supported that exon-intron structure in orthologous genes and in paralogous genes is conserved, and is useful for phylogenetic inference (Rokas et al., 1999; Krem and Di Cera, 2001; Kolukisaoglu et al., 2002; Wada et al., 2002). The exon–intron structure of *PIN* gene

family is well conserved (Paponov et al., 2005; Wang et al., 2009): For example, the exons and intron posistions of *AtPIN1* gene are conserved in the orthologous genes consisting of *AtPIN3*, *4*, and *7*, which are *AtPIN1* paralogous genes, and in *OsPIN1a*, *1b*, *1c*, and *1d*, which are orthologous genes of *AtPIN1*. In addition to domain and motif analysis, the *CpslcPIN* and *PhypaPIND* genes were confirmed to belong the *PIN* gene family based on arrangement of exons and introns. The putative homologous exon-intron structures and transmembrane domains are shown in Figure 13. The 3^{rd}-4^{th}-5^{th}-6^{th}-7^{th} exons and introns of *PhypaPIND* gene showed highly similar distribution to the 2^{nd}-3^{rd}-4^{th}-5^{th}-6^{th} exons and introns of other *PhypaPINA-C* genes (Fig. 13). Furthermore, the 3^{rd}-4^{th}-5^{th}-6^{th}-7^{th} exons and intron positions of *PhypaPIND* gene are also highly similar to it of all *ArathPIN* genes except *ArathPIN5* gene which conserved only the 2^{nd}-3^{rd}-4^{th}-5^{th} exons and intron positions; the 2^{nd}-3^{rd}-4^{th}-5^{th}-6^{th}, the 5^{th}-6^{th}-7^{th}-8^{th}-9^{th}, and the 3^{rd}-4^{th}-5^{th}-6^{th}-7^{th} exons and intron positions of *ArathPIN1*, *3*, *4*, and *8*, *ArathPIN2*, and *ArathPIN6* genes, respectively. The 4^{th}-5^{th}-6^{th}-7^{th} exons, transmembrane domains and intron positions, although 4^{th} and 5^{th} introns are long, of *CpslcPIN* gene showed high similarity to the 4^{th}-5^{th}-6^{th}-7^{th} exons, transmembrane domains and intron positions of *PhypaPIND* gene. Thus Figure 13 shows that in the 3'-terminal half region of the gene structure, distribution of exons and C-terminal transmembrane domains is highly conserved among *PIN* genes, and it suggests that *CpslcPIN* and *PhypaPIND* genes are members of *PIN* gene family.

3-3-3 Analysis of phylogenetic relationships among *PIN* homologous genes

Gene tree of *PIN* gene family The phylogenetic tree of *PIN* genes among land plants was constructed based on 79 amino acid sequences of the full-length cording region of *PIN* genes from the angiosperms, gymnosperms, lycophyta, and bryophyta (moss and liverwort), and an amino acid sequence of charophyta to outgroup using ML and BI methods (Fig. 11, Fig. 14, and Fig. 15). I used the CpslcPIN sequence of *C. psl.* complex as outgroup for rooting the tree, because it is thought that several genes of land plants have been gained in ancestor common to the streptophytes consisting of charophytes and land plants (Gogarten et al., 1989; Iwabe et al., 1989). The alignment of 80 entire PIN sequences provided a matrix of 1089 amino acid residues containing insertions and deletions (Fig. 11).

The resulting ML tree (Fig. 14) showed ten major monophyletic groups roughly similar to BI tree (Fig. 15) with ten monophyletic groups expect tree topology: (1) The group A consisted of bryophytes was supported with 99% bootstrap value and 1.00 posterior probability in ML and BI, respectively (99/1.00). (2) The group B consisted of grasses was supported with 100/1.00. (3) The group C was supported with 99/1.00. The branch of this group C consisted of two subgroups: eudicots and grasses groups. (4) The group D was supported with 97/1.00. The branch of this group D consisted of two subgroups: eudicots and grasses groups. (5) The group E consisted of eudicots was supported with 100/1.00. This group E in ML tree formed a clade with PoptrPIN15 and vvi_100253235 group of the group E', but was low bootstrap support (24%). (6) The group F consisted of bryophytes was supported with 100/1.00. (7) The group G consisted of *S. moellendorffii* was supported with 70/1.00. (8) The group H was supported with 77/1.00. The branch of this group H

consisted of two subgroups: eudicots and grasses groups. (9) The group I was supported with 78/0.87. The branch of this group I consisted of two subgroups: eudicots and grasses groups. This group I in BI tree formed a clade with a PoptrPIN10 of the group E' in ML tree with 1.00 posterior probability. (10) The group J was supported with 72/0.98. The branch of this group J consisted of two subgroups: subgroup J-1 and subgroup J-2 was supported with 98/1.00 and 49/0.95, respectively, moreover, the both subgroups consisted of eudicots and grasses groups, respectively. Gymnosperms (PinthPIN and GinbiPIN) group has low bootstrap support (39%) and not formed a clade in BI tree. Although ten monophyletic groups were recognized by high bootstrap support, phylogenetic relationship among these ten groups is still inconclusive because there are some clades with low bootstrap probability in this ML and BI tree. A summary of ML tree was shown in Figure 17A.

Motif comparison among PIN genes To clarify phylogenetic relationships in the PIN family furthermore, I compared with signature motifs of PIN amino acid sequences. In a recent study, it has been revealed that ArathPINs have some motifs related PIN localization and/or plant morphology (Michniewicz et al., 2007; Mravec et al., 2009; Dhonukshe et al., 2010; Huang et al., 2010; Zhang et al., 2010). These motifs were found as some molecular synapomorphies in the multiple alignment of amino acid sequences of *PIN* genes which showed many molecular synapomorphies (Fig. 16): (1) All groups (group B to J) except group A have a PNXY amino acid residues in site of NPXXY motif; (2) Groups C to J have a NPNXY amino acid residues in NPXXY motif; (3) Groups E, F, G, H, I, and J

have a Gly/2P cluster (second cluster) and three TPRXS(N/S) motifs, but the group E has two TPRXS(N/S) motifs; (4) Groups F, G, H, and J have one more Gly/P cluster at site of the third cluster. The site corresponding to this third Gly/2P cluster lost the Gly (N-glycosylation site, NX[S/T]X) and gained the CK2 (casein kinase II phosphorylation site, [S/T]XX[D/E]) in the group I because Asn residue (N) in the Gly was altered to Glu residue (E) (the 401st site in Fig. 11); (5) Group G, H, I, and J have one more PKC (Protein kinase C phosphorylation site, [S/T]X[R/K]) at site of the third cluster and they have the third Gly/2P cluster except group I; (6) Group H and subgroup J-1 have one more Gly/2P cluster (first cluster). The site corresponding to this first Gly/2P cluster was altered to PKC from Gly in the subgroup J-2; (7) Group I has no first Gly/2P cluster; (8) Subgroup J-2 has no third TPRXS(N/S) motif; (9) Grasses groups of groups H, I, and subgroup J-2 have no first Gly/2P cluster, no third TPRXS(N/S) motif, and no first TPRXS(N/S) motif, respectively; and (10) The GinbiPIN has SPNT amino acid residues including two phosphorylation sites by PID which had been found in ArathPIN1 (Ser337 and Thr340) by Michniewicz et al. (2007) and Zhang et al. (2010). I constructed the phylogenetic tree using the signature motifs (Fig. 17B), it indicated that the gene tree of the PIN family were supported and reinforced with these molecular synapomorphies (Fig. 17A and Fig. 17B). Especially, the monophyletic group consisting of groups B to J was supported by PNXY amino acid residues in site of NPXXY motif, and groups C to J was supported by NPNXY amino acid residues in NPXXY motif; the monophyletic group consisting of groups E to J was supported by the second Gly/2P cluster and the first and third TPRXS(N/S) motifs; the

monophyletic group consisting of groups F to J was supported by the second TPRXS(N/S) motifs; and the monophyletic group of J-2 was supported by absence of the first Gly/2P cluster and the third TPRXS(N/S) motif (Fig. 17A and Fig. 17B).

3-4 Discussion

By characterizing the *PIN* genes from the charophyta, liverwort, moss, and gymnosperms, I have provided new insight into the evolution of this gene family of auxin efflux carriers. In particular, the presence of *PhypaPIND*, *MarpoPIN2*, *MarpoPIN3*, and *MarpoPIN4* suggested a new hypothesis about role of *PIN* genes.

3-4-1 Sequence analysis of the *PIN* family of Streptophyta

Streptophyta consist of charophyte algae and land plants that have separated from Chlorophyta (McCourt et al., 2004). Homologs of the *PIN* gene family have been found in many land plants, and *PIN*-like genes are widely present in life. But, until the present report, no *PIN* orthologous gene had been reported in species of the Chlorophyta (Paponov et al., 2005; Zazímalová et al., 2007; Krecek et al., 2009), and there was no report on these genes in green algae. *CpsIcPIN* is the first full-length *PIN* orthologous gene to be reported from charophyte algae, although a *PIN* fragmental sequence was found in the EST library in Charophyta, i.e. *S. pratensis* (De Smet et al., 2011). Furthermore, the full-length cDNAs of *PIN* genes were isolated from two of the three major lineages of Bryophyta, the liverwort *Marchantia* and the moss *Physcomitrella*. Although the *PIN* sequence has been found in the moss *P. patens* (Paponov et al., 2005; Rensing et al., 2008), I could not confirm the existence of precedent *PpPIND* cDNA (NCBI accession number XM_001765711; Rensing et al., 2008) using PCR in this study. It is likely that XM_001765711 is a splicing variant

and the expression level is very low. Thus, the sequence of *PhypaPIND* isolated in this study provided perfect information for the analyses of phylogenetic relationships and function of its gene and protein. I isolated full-length cDNAs of *PIN* genes from two of the four major lineages of Gymnosperms, Coniferophyta *Pinus* and Ginkgophyta *Ginkgo*, although a *PIN* gene from Coniferophyta *Pinus* has been reported in Palovaara et al. (2010). These sequence data could contribute to improving the phylogenetic analysis of *PIN* genes among streptophytes. The analysis of exon-intron structures suggests that the *CpslcPIN* and *PhypaPIND* genes are orthologous genes of the *PIN* gene family (Fig. 13). The gene tree suggested that the *MarpoPIN2*, *MarpoPIN3*, and *MarpoPIN4* genes are orthologous genes of the *PhypaPIND* gene (Fig. 14 and Fig. 15).

It has been suggested that the hydrophilic loop domain of the AtPIN protein is involved in regulation via protein–protein interactions (Kerr and Bennett, 2007) and also interacts with ABCB/PGP/MDR proteins to enhance the auxin efflux (Blakeslee et al., 2007; Geisler and Murphy, 2006). In this study, it became clear that two hydrophobic domains and a long-loop domain of PIN protein were well-conserved among the Streptophyta (Charophyta and land plants) (Fig. 12). The Gly/2P motif, considered to be important for posttranslational modifications (Zazímalová et al., 2007), three TPRXS (N/S) motifs phosphorylated by the PID protein (Dhonukshe et al., 2010; Huang et al., 2010), and a tyrosine-based NPXXY internalization motif (Zazímalová et al., 2007; Mravec et al., 2009) in a hydrophilic loop domain are conserved in all land plants, whereas the PhypaPIND, MarpoPIN2, MarpoPIN3, and MarpoPIN4 proteins with a middle-loop have none of these motifs (Fig. 12). In addition,

the CpslcPIN protain from charophycean algae does not have these motifs. This suggests that the function of the *CpslcPIN* gene and the *PhypaPIND*, *MarpoPIN2*, *MarpoPIN3*, and *MarpoPIN4* genes are likely to be different from other *PIN* genes. I cannot specify whether the putative function of the CpslcPIN and PhypaPIND and MarpoPIN2, MarpoPIN3, and MarpoPIN4 proteins is as an auxin carrier or whether there is an alternative or additional function only from the present sequence information. The hydrophilic loop domain of GinbiPIN contains the amino acids serine and threonine that may be target sites of PID-dependent phosphorylation like the ArathPIN1 protein (Michniewicz et al., 2007; Zhang et al., 2010) (Fig. 16), and thus it is likely that the behavior of GinbiPIN is similar to the ArathPIN1 protein.

3-4-2 Phylogenetic relationships and molecular evolution of the *PIN* gene family

I found many molecular synapomorphies in the *PIN* genes of land plants. The phylogenetic tree based on amino acid sequences of *PIN* genes was reinforced by the presence and absence of some signature motifs. The phylogenetic relationships and molecular evolution of *PIN* gene family were underpinned by the gene tree and signature motifs.

In the ML tree obtained in the present study, I could recognize some gene groups; however, the phylogenetic relationships among these groups could not be clarified due to low support of the branches (Fig. 14). Thus, I employed motif information to recognize

monophyletic groups, and constructed the maximum parsimonious tree using the signature motifs (Fig. 17B). Figure 17 shows a summary of phylogenetic relationships among the *PIN* genes in land plants. The phylogenetic trees of *PIN* genes were supported by some motifs in some aspects; however, there was some conflict between the gene tree and the motif tree (Fig. 17A and Fig. 17B). The monophyletic group consisting of groups F and G in the ML tree could not be supported by the motifs because there is a feature common to groups G, H, I, and J which is two PKC at the site corresponding to the third Gly/2P cluster (Fig. 16, Fig. 17A, and Fig. 17B). The third Gly/2P cluster shows molecular synapomorphy in the groups G, H, and J. However, group I has no Gly corresponding to this third Gly/2P cluster because the Asn amino acid residue in the Gly was altered to a Glu amino acid residue (Fig. 11 and Fig. 16). I inferred that the third Gly/2P cluster has been individually lost in groups F and I because the parallel gain of functional motif is a very low probability event in groups G, H, and J in the ML tree, respectively. Therefore, Figure 17C proposed a possible phylogenetic relationships among the *PIN* genes in land plants with the constraint of monophyly of the group F and G. It was inferred that the third Gly/2P cluster was gained in an ancestor common to groups F, G, H, I, and J; thereafter, the PKC and the Gly in the third Gly/2P cluster was lost in groups F and I, respectively (Fig. 17C). This issue will be resolved when more species are included in the phylogenetic analyses because our analyses were performed with sequences of the full-length coding region. It would be highly desirable to include more PIN sequences from the gymnosperms and ferns because the phylogenetic distance between flowering plants, lycophyta, and bryophyta is very large.

According to the ML and BI tree obtained in this study, it was interpreted that the lineages of bryophytes and lycophytes have individually lost orthologous genes corresponding to groups B, C, and D. However, it is also possible that problems with random sampling error could have occurred (Fig. 14 and Fig. 15). Considering the present understanding of land plant phylogeny, my results show that the *PIN* genes of land plants consist of major three groups, which correspond to "Core", "Basal", and "Bryophyta-specific" groups (Fig. 14, Fig. 15, and Fig. 17). Land plants consist of Bryophyta and Tracheophyta. The "Core" group includes bryophytes and tracheophytes, which have a long-loop and several motifs: a NPXXY motif, at least one Gly/2P cluster, and one TPRXS (N/S) motif. The "Basal" group includes only angiosperms, which lack the long-loop and have a PNTY or NPNXY amino acid motif. The absence of the "Basal" group gene in *P. patens* and *S. moellendorffii* has been confirmed by genome sequence analysis (Rensing et al., 2008; Banks et al., 2011). Therefore, the results suggest that at least two *PIN* genes existed in the last common ancestor between eudicots and monocots. Group E, including ArathPIN6 (Fig. 14 and Fig. 15), is likely to have separated from the "Core" group early in evolution because Mravec et al. (2009) have found that the AtPIN5, 6, 8 proteins are localized in the ER. The *PIN* homologous genes of Bryophytes, the liverwort *M. polymorpha* and the moss *P. patens* are separated into the "Core" group and "Bryophyta-specific" group. Proteins of the "Bryophyta-specific" group have a middle-loop and the NPXXY and TPRXS (N/S) motifs are absent (Fig. 16). The *PhypaPIND* gene has suggested the possibility of horizontal transfer of this gene from grasses because of the presence of

intron sequences (Krecek et al., 2009; Mravec et al., 2009). However, the *PhypaPIND* gene formed a cluster with the *MarpoPIN2*, *MarpoPIN3*, and *MarpoPIN4* genes, supported by a 99% bootstrap value and 1.00 posterior probability, which indicate orthologous genes (Fig. 14 and Fig. 15). Thus, it is suggested that at least two *PIN* genes existed in the last common ancestor between bryophytes and tracheophytes, and the orthologous *PhypaPIND* gene is absent in tracheophytes. This suggests that genes of the "Bryophyta-specific" group may have a bryophyte-specific role that is absent in tracheophytes. The lycophyte *S. moellendorffii PIN* genes formed a clade, suggesting that *SelmoPIN* genes were generated by *S. moellendorffii*-specific duplications, and *SelmoPIN* genes may have an *S. moellendorffii*-specific role. Furthermore, I found that the group J clearly diverged into subgroup J-1 and subgroup J-2 (Fig. 17). The subgroup J-1 includes *ArathPIN1*, whereas no gene of *A. thaliana* was found in subgroup J-2, and the features of the motifs different greatly between the subgroups (Fig. 16). This result suggests that the genes of subgroup J-2 probably have new role which is not required for *A. thaliana*. No *PIN* genes could be found in the Chlorophyta *Chlamydomonas*, *Ostreococcus*, *Micromonas*, *Chlorella*, and *Volvox*, whereas they were present in the filamentous charophyte algae *S. pratensis* (De Smet et al., 2011) and the unicellular *C. psl.* complex. This indicates that the *PIN* family appeared early in the evolution of Streptophyta. The cDNA of the *CpslcPIN* gene was amplified in samples 24 and 72 hours after the mating reaction, whereas I could not isolate it from vegetative cells (mt^+-V and mt^--V), suggesting that *CpslcPIN* is likely to function in sexual reproduction and/or zygote formation. A gene introduction and stable transformation system

has been created in the unicellular charophyte alga *C. psl.* complex (Abe et al., 2008; 2011). Furthermore, I found at least two *PIN* homologous genes from the EST database in *Chara braunii* (Sakayama et al., unpublished data), and am trying to isolate the cDNA now. The *Charales*, including *Chara braunii*, are often considered to be a sister lineage to land plants, although it has been suggested that either the Zygnematophyceae or a clade consisting of Zygnematophyceae and Coleochaetophyceae might be the most likely sister group to land plants (Becker and Marin, 2009; Turmel et al., 2006; Wodniok et al., 2011). The elucidation of the Charophyta *PIN* orthologous genes is expected to reveal a more comprehensive evolutionary history of the *PIN* gene family.

Chapter 4.

The analysis of the moss *Physcomitrella patens PIND* gene

4-1 Introduction

Phylogenetic analysis of PIN family using a charophytes, bryophytes, lycophytes, gymnosperms, and angiosperms suggests that bryophytes have a specific gene group, presumably exhibiting bryophytes specific function (Chapter 3). In this chapter, the moss *Physcomitrella patens* was used as a model plant for further research, and the *PpPIND* gene belonging to the "Bryophytes-specific" group has been analyzed in detail.

The moss *P. patens* has many suitable features as a model plant. The simple morphology of *P. patens* makes it a useful system for studies on plant physiology and developmental biology (Cove et al., 1997; Reski, 1998; Schumaker and Dietrich, 1998). The ability to target gene disruption in *P. patens* is attributed to a high rate of homologous recombination that has not been observed in other land plants (Kammerer and Cove, 1996; Schaefer, 2001). The life cycle of *P. patens* comprises haploid gametophyte and diploid sporophyte (Cove and Knight, 1993). The haploid gametophyte of *P. patens* has two major developmental stages: protonemata are branching system of cell filaments and gametophores are leafy shoots that are the more familiar part of most moss plants. The diploid sporophyte of *P. patens* has a short stalk without branching and lateral organ, and is a typical one in bryophytes. In the evolution of vascular plants, tracheophyte, the dramatic change in the gametophyte and sporophyte could have occurred, because many differences in organization are found between bryophytes and tracheophytes both in gametophyte and sporophyte (Kenrick and Crane, 1997). The gene tree shown in Chapter 3 (Fig. 14 and Fig.

15) suggests that *PpPIND* gene is expected to carry the role specific to bryophytes.

Here I describe expression pattern of *PpPIND* gene using semi-quantitative RT-PCR and real time PCR analysis. Furthermore, I report the results of functional studies of *PpPIND* gene, as an initial contribution towards gaining an understanding of the role of *PIN* gene family in a moss: The *pppinD* disruptant lines were generated. Additionally, to analyze relationship between PpPIND protein and auxin transport, I generated *pppinD* disruptant lines in the background of GH3:GUS1-1 and GH3:GUS36-1 lines (Fujita et al., 2008) that indicate GUS activity in response to auxin.

4-2 Materials and Methods

4-2-1 Plant materials and culture conditions

The plant materials were the moss *Physcomitrella patens* as written in Chapter 3. BCDAT, BCD, and Jiffy-7 media (Jiffy-7; Jiffy Products International AS, Kristansand, Norway) were used for the culture of *P. patens*. To induce the gametangia and sporophytes, protonemata that had been vegetatively propagated on 9 cm petri dishes with BCDAT medium were inoculated onto sterile 42-mm diameter peat pellets. Jiffy-7 medium was autoclaved in a plastic plant box (Asahi Techno Glass, Chiba, Japan) with milli-Q water (MILLIPORE, MA, USA). The mosses inoculated on the Jiffy-7 medium were cultured at 25°C under continued light conditions for 1 to 1.5 months, and then transferred to 15°C under 8-hour light/16-hour dark conditions for induction of gametangia development. Under these conditions, first antheridia and then archegonia differentiated at the shoot apices of the gametophores. Fertilization was promoted by pouring sterilized water to the surface of the gametophores colony. Sporophytes were visible under stereomicroscope 4 weeks after the transfer to 15°C and were collected after an additional month of culture.

4-2-2 total RNA extraction

Total RNA was extracted using the RNeasy Plant Mini Kit (QIAGEN K. K., Tokyo, Japan) from different stages of the moss life cycle: chloronemata, caulonemata, gametophore with rhizoids, antheridia, archegonia, and sporophytes.

4-2-3 Preparation of cDNA

Total RNA was quantified on the NanoDrop (NanoDrop Technologies, Wilmington, DE, USA), and treated with the RNase-free DNase I (Roche) according to the manufacturer's instructions. For the semi-quantitative RT-PCR analysis, single stranded complementary DNA (cDNA) was synthesized from 500 ng DNase I treated total RNA using the SuperScript II reverse transcriptase and the Oligo dT-Adaptor Primer (Gibco-BRL, Rockvill, MD, USA). For the real-time PCR analysis, single stranded cDNA was synthesized from 500 ng DNase I treated total RNA using the PrimeScript RT reagent kit (Perfect Real Time) with the random hexamers and the Oligo dT Primer (TaKaRa) according to the manufacturer's instructions.

4-2-4 Semi-quantitative RT-PCR

To examine the patterns of expression, cDNAs were synthesized from total RNA extracted from different stages. The PCR amplification test was performed with *PpPIND* gene-specific internal primer set PpPIN-F300 (5'-TGGGTGATTACCTTATTCCAGCTCTC-3') and PpPIN-R395 (5'-ATGAGCGAGTAGACGATTCCCATAAC-3'). The *P. patens* orthologous gene of the glyceraldehyde-3-phosphate dehydrogenase C subunit gene (*PpGapC*) was used as an internal control. *PpGapC* was amplified by PCR with PpGapC-F850 (5'-CGATGCCATTAAGACGGCTATCAA-3') and PpGapC-R59 (5'-

CATGAACGCTGGCGATGGAAATTA-3') primers. PCR reactions were carried out using 0.1 μl TaKaRa Ex Taq HS (TaKaRa), 2μl 10 x Ex Taq buffer, 1.6 μl 2.5 mM dNTP, 0.5 μl each of 10 μM gene-specific forward and reverse primers, and 1 μl of a cDNA in 20 μl of final reaction volume. The PCR conditions were 1 cycle at 94°C for 60 sec, followed by 36, 38, or 40 cycles of 94°C for 30 sec, 56°C for 30 sec, and 72°C for 45 sec, and a final step at 72°C for 5 min. The results of expression analysis were confirmed by conducting an additional independent experiment as a repeat of the whole procedure from the starting point isolation of total RNA, and similar results were obtained (data not shown).

4-2-5 Quantitative real-time RT-PCR

To perform the quantification expression analysis, real-time PCR analysis was carried out using the ABI PRISM 7000 Sequence Detection System (Applied Biosystems, Foster City, CA, USA) with the TaqMan Probe (Applied Biosystems) and the Premix Ex Taq (Perfect Real Time) (TaKaRa) according to the manufacturer's instructions. The primers and probe for real-time PCR were designed using the Primer Express 2.0, a software program provided with the ABI PRISM 7000 Sequence Detection System (Applied Biosystems). *PpPIND* gene was detected using the PpPINDcDNA-303F (5'-ACAAACGTCTGCGCTGAACA-3') and PpPINDcDNA-375R (5'-CCCACCGGGCAAATTTC-3') primers and PpPINDcDNA probe (5'-TGGAAGCTGGCAGAATCGTTGAGGAG-3'). The *P. patens* orthologous gene of the alpha-1 tubulin gene (*PpTUA1*, DDBJ Accession No. AB096718) was used as an internal

control to normalize the amount of templates in RT-PCR. *PpTUA1* gene was detected using the PpTUA1cDNA-1002F (5'-ACATGGCGTGTTGCCTGAT-3') and PpTUA1cDNA-1088R (5'-CCGTGCGCTTCGTCTTG-3') primers and PpTUA1 probe (5'-AAGGACGTGAACGCCGCGGT-3'). To test the accuracy of the amplification plot method on DNA standard samples of known concentration, data from standard curves were analyzed. The DNA standard samples of *PpPIND* and *PpTUA1* were obtained by PCR. PCR amplification was carried out with the plasmid of each gene that obtained in cloning and the T7 (5'-TAATACGACTCACTATAGGG-3') and SP6 (5'-ATTTAGGTGACACTATAGAA-3') primers. The duplicate measurement was carried out using each of the cDNA and standard samples as template. Statistical analysis was performed according to standard procedures. Standard errors are shown in the figure as bars on the data points. The results of expression analysis were confirmed by conducting an additional independent experiment as a repeat of the whole procedure from the starting point isolation of total RNA, and similar results were obtained (data not shown).

4-2-6 Genomic DNA extraction

Genomic DNA was extracted from protonemata of *P. patens* by the CTAB (Hexadecyltrimethylammonium bromide) method (Murray et al., 1980).

4-2-7 Genomic Southern hybridization

2 μg of *P. patens* genomic DNA were digested, separated on 0.7 % (w/v)

SeakemGTG agarose (BMG, Rockland, ME, USA), and transferred to a Hybond N^+ nylon membrane (GE Healthcare, UK). An Alkphos Direct kit (GE Healthcare, UK) was used for labeling DNA probes, hybridization and detection.

4-2-8 Construction for *PpPIND* gene-knock-out

To disrupt the *PpPIND* gene, the plasmid pGUSmutNPTII (provided by Dr. Hiwatashi), which contained the coding sequence of ß-glucuronidase gene (*uidA*), nopaline synthase polyadenylation signal (nos-ter), and neomycin phosphotransferase II (*nptII*) cassette (Nishiyama et al., 2000), was used for construction of PpPIND-dis lines.

PpPIND genomic DNA (5130 bp) amplified using the gene-specific primer set PpPIN-5'f1 (5'-GGTCGAGAGAGATGAAGTATGTAGGGTTTTACC-3') and PpPIN-3'r1 (5'-AGTGGAGAAGCAGAGACTCAGCGGACC-3') was cloned into pGEM-T Easy vector (Promega, Madison, WI, USA), thereby generating pgPpPIND (Fig. 8).

A 1053 bp fragment of the 5' region containing the estimated start codon of *PpPIND* gene was amplified using the PpPIN-KO5'f (5'-GGGGTACCGACAGGCAGAGGGGAGATCTG-3') and PpPIN-KO5'r (5'-CCATCGATTGTCAACATGACGATAACGTTGATTGTCC-3') primers and pgPpPIND plasmid. The 1053 bp fragment was digested by *Kpn*I and *Cla*I and cloned into *Kpn*I-*Cla*I site of pGUSmutNPTII plasmid. A 664 bp fragment of the 3' region containing the stop codon of *PpPIND* gene was amplified using the PpPIN-KO3'f (5'-CACTAGTCTGAGAAATACAAATTCCAGACAACGAATAC-3') and PpPIN-KO3'r (5'-

TCCCCGCGGGGTACCATTGACCTGTTTCTGAGATG-3') primers and pgPpPIND plasmid. This 664 bp fragment was digested by *Spe*I and *Sac*II and cloned into *Spe*I-*Sac*II site of pGUSmutNPTII plasmid including the previous fragment of 5' region of *PpPIND* gene. Then the recombinant plasmid was designated as pPpPIND-dis. pPpPIND-dis plasmid was linearized with *Kpn*I for gene targeting.

To disrupt the *PpPIND* gene in the GH3:GUS1-1 and GH3:GUS36-1 transgenic lines provided by Asst. Prof. Fujita, the plasmid pPpPIND-dis was modified stripping it of the coding sequence of *uidA* by *Eco*RI digestion. Then this recombinant plasmid was designated as pGH3GUS/PpPIND-dis. pGH3GUS/PpPIND-dis plasmid was linearized with *Kpn*I, that was used for construction of GH3:GUS/PpPIND-dis lines.

4-2-9 Transformation of *P. patens*

Polyethylene glycol-mediated transformation was performed as previously described (Nishiyama et al., 2000). Stable transformants were screened with PCR using appropriate primer sets to confirm the insertion of the construct in the locus. PCR-positive candidates were further analyzed by Southern analyses.

4-2-10 Detection of phenotypic abnormality by dissection

Since there is no information on the difference in intensity of connection between the sporophyte and the gametophyte in *P. patens*, shoot tips of gametophores bearing a fully mature sporophyte were collected 8 weeks after induction of gametangia development from

the wild type and *PpPIND* disruptant lines. Differences in intensity of the junction between vaginula and sporophyte foot were classified into three types, "looseness", "moderateness", and "tightness" junction types. A method of dissection and classification was as follows. Fresh plants were laid on a slide glass and were dissected immediately in 50 mM sodium phosphate buffer (pH 7.0) without fixation treatment under a stereomicroscope (LEICA MZ16, Leica Microsystems, Wetzlar, Germany). Shoot tip of gametophores bearing a fully mature sporophyte was separated from gametophores at the site of innermost perichaetial leaf, which is near to end of the sporophyte foot. Under holding the sporangium using a tweezers, a slit was put into the vaginula from the bottom end using a sharp single-edged razor, which was made from a twin bladed razor (Feather, Osaka, Japan). Then, the vaginula was slipped off or was dragged away from foot using a sharp single-edged razor. At this time, I recognized three types among them as follows. It was classified into the "looseness" junction type when the vaginula was pushed very lightly and could be separated from foot, and when the adhesion between vaginula and foot could not be perceived sensibly. It was classified into the "moderateness" junction type when the vaginula was pushed lightly and could be separated from foot, and when the adhesion between vaginula and foot could be subtly perceived sensibly. It was classified into the "tightness" junction type when the adhesion between vaginula and foot could be perceived sensibly and the vaginula was dragged away from foot. I performed blind test when phenotypic abnormality between the wild type and *PpPIND* disruptant lines was detected.

4-2-11 Statistical analysis

The corelation between the *PpPIND* disruptant lines and junction types were analyzed using a Pearson's Chi-squared test with Yates' continuity correction for evaluation that there is no variation between the *PpPIND* disruptant lines. For testing difference between the wild type and the *PpPIND* disruptants, we used a binomial generalized linear model (GLM). All statistical analyses were performed with the software R v2.11.1 (R-Development-Core-Team, 2010). For binomial, the junction type was classified into the "looseness" and the "moderateness and tightness".

4-2-12 SEM

For scanning electron microscopy (SEM), fully matured sporophytes at 8 weeks after induction of gametangia development were fixed with 2.5% glutaraldehyde in 0.1 M sodium phosphate buffer pH 7.0 for 48 hours at 4°C after 30 min under vacuum. The samples were washed with 0.1 M sodium phosphate buffer (pH 7.0) for 1 hour at 4°C, and then fixed in 2% osmium tetroxide for 2 hours at 4°C. The fixed samples were dehydration through a graded ethanol series. The freeze-dried preparations were soaked through a graded tert-butyl alcohol and ethanol mixture series. Change the tert-butyl alcohol again and only small amount enough to soak the samples was used. The samples were frozen in refrigerator for few minutes and freeze-drying using a JEOL JFD-300 Freeze Drying Device (JEOL Ltd, Tokyo, Japan). The dried samples were mounted on stubs using carbon tape and coated with gold with an ion sputter JFC-1100 (JEOL Ltd, Tokyo, Japan). The preparations

were observed using a SEM_SU1510 (Hitachi High-Technologies, Tokyo, Japan) operating at 15 kV and photographed.

4-2-13 TEM

For transmission electron microscopy (TEM), fully matured sporophytes at 8 weeks after induction of gametangia development were injured using needle and were fixed with 2.5% glutaraldehyde in 0.1 M sodium phosphate buffer pH 7.0 for 48 hours at 4°C after 30 min under vacuum. The samples were washed with 0.1 M sodium phosphate buffer (pH 7.0) for 1 hour at 4°C, and then fixed in 2% osmium tetroxide for 2 hours at 4°C. The fixed samples were dehydration through a graded ethanol series. Subsequently, embedded in spuur low viscosity resin (Polysciences, Inc. Pennsylvania, USA) through a graded spuur resin and ethanol series. Sections were prepared and stained with uranyl acetate and lead citrate. TEM micrographs were taken with a JEM-1011 (JEOL Ltd, Tokyo, Japan) operating at 75 kV.

4-2-14 Microtome section

For microtome sections (MS) using paraffin, fully matured sporophytes at 8 weeks after induction of gametangia development were injured using needle and were fixed in FAA (5% [v/v] formalin, 5% [v/v] acetic acid, and 45% [v/v] ethanol) for 24-48 hours after 30 min under vacuum. The fixed samples were dehydrated in ethyl alcohol and butyl alcohol series, embedded in paraffin through a graded paraffin and butyl alcohol series, and

cut into 7 μm thick sections using LEICA RM2145 (Leica Microsystems, Wetzlar, Germany). Sections were stained with safranine for 24 h and fast green for 20 s.

4-2-15 GUS assay

GUS staining was performed as described by Nishiyama et al. (2000). The disrupted lines were cultured on BCDAT or Jiffy-7. The tissues were not fixed before GUS staining, and infiltrated for 30 minutes under vacuum in a substrate solution (50 mM NaH_2PO_4 [pH 7.0], 0.5 mM 5-bromo-4-chloro-3-indolyl b-D-glucuronide [X-Gluc, Wako, Osaka, Japan], 0.5 mM $K_3Fe(CN)_6$, 0.5 mM $K_4Fe(CN)_6$, and 0.05 % [v/v] Triton X-100), and then incubated at 37 °C for 12-24 hours for staining. After the incubation, the tissues were dehydrated through a graded ethanol series. Images of the stained tissues were digitized with CCD camera (LEICA DFC480, Leica Microsystems, Wetzlar, Germany).

4-3 Results

4-3-1 Expression pattern of the *PpPIND* gene

Semi quantitative RT-PCR analysis To investigate the expression pattern of *PpPIND* gene, I performed gene-specific semi quantitative RT-PCR using total RNA extracted from gametophytes (chloronemata; caulonemata; shoot tips of adult gametophores without rhizoids at 25°C; base of adult gametophores with rhizoids at 25°C; shoot tips of gametophores bearing antheridia, 2 weeks after induction; and shoot tips of gametophores bearing antheridia and archegonia, 3weeks after induction) and sporophyte with gametophyte (shoot tips of gametophores bearing young sporophyte, 3.5 weeks after induction; and shoot tips of gametophores bearing mature sporophyte, 5 weeks after induction). As shown in Figure 18, *PpPIND* gene was highly expressed in chloronemata, caulonemata, and young and mature sporophytes with shoot tips of gametophores, whereas low expression was detected in reproductive organs and not detected in adult gametophores with rhizoids at 25°C. This result indicates that the *PpPIND* gene is constantly expressed in the protonemata and sporophyte phase, although slightly lower in the sporophyte, and shows very low-level expression in reproductive organs. Furthermore, Figure 18 shows that expression of *PpPIND* gene (XM_001765711; Rensing et al., 2008) was not detected in all the sampling organs because the band is 386 bps, whereas *PpPIND* gene from this study is 559 bps. This suggested that *PpPIND* gene (XM_001765711; Rensing et al., 2008) is a splicing variant and probably without a clear function.

Quantitative real-time RT-PCR analysis To analyze the expression pattern of the *PpPIND* gene in detail, I employed a real time PCR method that is capable to quantitatively detect the differences of gene expressions among samples.

The expression of the *PpPIND* gene in sporophytes phase was highest in fully matured sporophytes. In shoot tips of gametophores, the expression increased at fully matured sporophytes phase, and the expression level of the *PpPIND* gene increased as the sporothytes matured (Fig. 19). In gametophyte generation, high expression level was observed in caulonemata (Fig. 19). This result indicates that the *PpPIND* gene is gradually highly expressed in sporophytes and shoot tips of gametophores till sporophytes become fully mature, and that the expression level in caulonemata is by far the highest among gametophyte.

4-3-2 Functional analyses of the *PpPIND* gene and PpPIND protein

Generation of *PpPIND* and GH3:GUS/*PpPIND* disruptants To investigate the function of *PpPIND* gene, the disruptants were generated by homologous recombination using targeting constructs (Fig. 20). A DNA fragment containing the *uidA* and *nptII* cassette was replaced into *P. patens PIND* gene locus to obtain *PpPIND* disruptants. After screening of stable transformants of *PpPIND* gene locus by PCR analysis, eight lines (PpPIND-dis-5-33, 5-113, 5-116, 5-135, 5-154, 5-162, 24-168, 24-292) were found. In addition to the *PpPIND* disruptants, to analyze the relationship between PpPIND protein and auxin

transport, GH3:GUS/*PpPIND* disruptants were generated by disrupting *PpPIND* gene in the background of the two types GH3:GUS transgenic lines which have auxin-responsive *GH3* promoter with a GUS reporter gene in *P. patens* (Fujita et al., 2008). A DNA fragment that contained the *nptII* cassette was replaced into *PpPIND* gene locus in *P. patens* GH3:GUS lines (GH3:GUS 1-1 and GH3:GUS 36-1 lines which the GH:GUS insert is the reverse and forward direction) to obtain GH3:GUS/*PpPIND* disruptants. By PCR analysis, five and three lines (GH3:GUS 1-1/PpPIND-dis-1, 6, 9, 104, 206, and GH3:GUS 36-1/PpPIND-dis-1, 8, 9) were found to have a disrupted *PpPIND* gene locus in GH3:GUS 1-1 and 36-1 lines, respectively. Genomic southern analyses showed that all lines contained a single insertion in *PpPIND* gene locus (Fig. 21, Fig. 22, and Fig. 23).

Phenotype of *PpPIND* disruptants All *PpPIND* disruptant lines possessed phenotypically normal morphology. However, phenotypic abnormality becomes apparent through the dissection of sporophytes. I counted the number of the appearances of the loose, moderate, and tight junction types under the blind condition without information of wild types or *PpPIND* disruptant lines. The *PpPIND* disruptant lines, line 5-33, 5-135, and 5-162, were characterised by a more increase in population that loosed junction between vaginula (gametophyte generation) and sporophyte foot in fully matured phase, and conversely tightened population was decreased (Fig. 24): The frequency of the loose, moderate, and tight junction were 27.3%, 32.8%, and 39.8% in wild type; 56.5%, 26.1%, and 17.4% in PpPIND-dis-5-33; 55.1%, 24.6%, and 20.3% in PpPIND-dis-5-135; 63.3%, 23.3%, and

13.3% in PpPIND-dis-5-162, respectively. For the statistical analysis, at first, I performed a Pearson's Chi-squared test with Yates' continuity correction for evaluation. As a result, no significant differences among the *PpPIND* disruptant lines were found. Then, a GLM method was applied to test an effect of disruption of *PpPIND* gene. The equivalence among the *PpPIND* disruptant lines was presented: p-value = 0.9036 between PpPIND-dis-5-33 and PpPIND-dis-5-135, p-value = 0.8265 between PpPIND-dis-5-33 and PpPIND-dis-5-162, and p-value = 0.5871 between PpPIND-dis-5-135 and PpPIND-dis-5-162 (Table 3). As the no significant difference among the *PpPIND* disruptant lines, I used these *PpPIND* disruptant lines for functional analysis of *PpPIND*. The effect of disruption of *PpPIND* gene was clarified using a GLM data analysis (p-value = 0.00768, Table 4). The present results supported that the junction between vaginula and sporophyte foot in fully matured phase are likely to become looser in the *PpPIND* disruptant than in wild type.

To examine a reason why it become loose, I observed intracellular structure, cell structure, and surface of cell walls of sporophyte-gametophyte junction using transmission electron microscopy (TEM), microtome sections (MS), and scanning electron microscopy (SEM). A slight difference was recognized with SEM: the surface of cell walls of foot in *PpPIND* disruptant lines was slightly smooth compared with wild type (Fig. 25). Under TEM (not shown) and MS (Fig. 26), no morphological difference was detected between wild type and disraptants.

GUS assay of GH3:GUS/*PpPIND* disruptants In the *P. patens*, GH3:GUS

lines have indicated GUS activity of a GH3 auxin-responsive promoter in gametophyte and sporophyte (Fujita et al., 2008). I examined changes in GUS staining patterns between GH3:GUS lines and GH3:GUS/*PpPIND* disruptant lines. The localization of GUS signal was detected, which was changed at upper region of vaginula in GH3:GUS/*PpPIND* disruptant lines (GH3:GUS 1-1/PpPIND-dis-1, 9 and GH3:GUS 36-1/PpPIND-dis-8) as compared with GH3:GUS lines (Fig. 27): The upper region of vaginula in GH3:GUS 1-1 was stained, whereas GH3:GUS 1-1/PpPIND-dis-1 and 9 were not detected; and the upper region of vaginula in GH3:GUS 36-1 was weakly stained, whereas GH3:GUS 36-1/PpPIND-dis-8 was increased GUS staining. The two different phenotypes caused by the same gene disruption are unusual, however, there was difference in construction to insert a GH3:GUS expression cassette between GH3:GUS 1-1 and GH3:GUS 36-1 transgenic lines. In other organs, a significant difference in GUS staining could not be detected between GH3:GUS lines and GH3:GUS/*PpPIND* disruptant lines because the GUS staining signal was unstable among individuals.

4-4 Discussion

4-4-1 Functional analysis of the *PpPIND* gene

PpPIND gene expression was found in the matured sporophyte and shoot tip of the gametophore during sporophyte maturation, and in protonemata, especially caulonemata (Fig. 18 and Fig. 19). Additionally, disruptants defective in the *PpPIND* gene showed substantial increases in the population that loosened the junction between the vaginula and sporophyte foot in the fully matured phase (GLM data analysis, $p = 0.00768$) (Fig. 24 and Table 4). These results suggest that the *PpPIND* gene may function in the regulation of the connection between the sporophyte and the gametophyte in the fully matured phase. The fully matured sporophytes phase occurs just before the release of spores. The *PpPIND* gene may have a role specific to bryophytes, because tight junctions of the sporophyte–gametophyte may help to widely disperse spores by wind and to provide a survival advantage. Interestingly, the liverwort *Marchantia,* which has sporophytes suspended by a foot on the undersurface of an archegoniophore (Smith, 1955), has three *PIN*s which are in the sister group of the *PpPIND* gene, as shown in Chapter 3 (Fig. 14 and Fig. 15). Thus, it is likely that the *PIN* genes in *Marchantia* are more important for sporophyte development and the connection between the sporophyte and the gametophyte than the *PpPIND* gene. Thus, the functional analyses of *Marchantia PIN*s will give us a better understanding of their function.

In bryophytes, the sporophyte–gametophyte junction also seems to be important in nutrient transfer from the gametophyte to the sporophyte and in sporophyte penetration and

growth (Uzawa and Higuchi, 2010). The *PpPIND* gene, which encodes for a putative transporter protein and is highly expressed in sporophytes and the shoot tips of gametophores in the fully matured sporophyte phase (Fig. 19), is expected to have a functional role in substance transport and signal transduction, especially in the transport of nutrients or auxin, in the fully matured sporophyte phase. Bryophytes have a sporophyte dependent on the gametophyte throughout the life cycle, whereas tracheophytes have a reduced gametophyte and a free-living sporophyte (Uzawa and Higuchi, 2010). The sporophyte–gametophyte junction in bryophytes functions as a placenta. Placental cells generally have a dense cytoplasm rich in mitochondria, endoplasmic reticulum and ribosomes, and they present a wall-membrane apparatus typical of transfer cells characterized by wall ingrowths or a wall labyrinth (Gunning and Pate, 1969; Ligrone et al., 1993). In this study, an abnormality in the intracellular structure, cell structure, and wall ingrowths of placental cells in *PpPIND* disruptants could not be detected by preliminary TEM and careful MS observation (Fig. 26). More careful observations by TEM are needed because of the complex structure of placental cells. It is likely that *PpPIND* gene is not important for nutrient transfer through placenta cells or for the normal development of the sporophyte, but rather plays a role in signal transduction (especially auxin signal transduction) without morphogenesis, because the sporophytes of *PpPIND* disruptants looked normal without a clear change in placental cells. A reduced ratio of auxin to cytokinin has been shown to promote cell adhesion in plant tissue culture (Gamborg et al., 1968; Schiavone and Cooke, 1987), and auxin has been implicated in providing positional cues during leaf abscission (McManus et al., 1998). The present

phenotypic observation by dissection and SEM in the fully matured sporophyte phase showed that the cell adhesion between the vaginula and sporophyte foot became loose in the *PpPIND* disruptants (Fig. 24), and that the surface of cell walls in the foot of *PpPIND* disruptants was slightly smoother without peeling damage or adhesion of fragments of the vaginula, whereas the wild type seemed to be damaged and had fragments on the surface (Fig. 25). Feraru et al. (2011) reported that cell wall integrity is required for PIN polarity in *Arabidopsis*; hence, *PIN* gene defects may cause an abnormal condition in the cell wall. Because the regulation of auxin concentration participates in the control of cell adhesion (Gamborg et al., 1968; Schiavone and Cooke, 1987; McManus et al., 1998) and the *Arabidopsis PIN* gene is involved in cell wall integrity (Feraru at al., 2011), one possibility is that the *PpPIND* gene encodes an auxin efflux carrier and plays a role in cell adhesion between the sporophyte and gametophyte. The study of the sporophyte–gametophyte junction is considered to be important for understanding the relationship between the two generations in the evolution of land plants (Ligrone et al., 1993; Uzawa and Higuchi, 2010), and a detailed study of the *PpPIND* gene function would help us to understand this.

The phenotype of *PpPIND* disruptants obtained in this study showed a partial abnormality; not all the *PpPIND* disruptants showed "looseness" and the surface of the cell walls of the foot in *PpPIND* disruptants was not highly abnormal (Fig. 24 and Fig. 25). Auxin transport is mediated at the cellular level by three independent mechanisms that are characterized by the PIN-FORMED (PIN), ATP-BINDING CASSETTE GROUP B/P-GLYCOPROTEIN/MULTIDRUG RESISTANCE (ABCB/PGP/MDR), and AUXIN-

RESISTANT1 (AUX1)/LIKE-AUX1 (AUX1/LAX) transport proteins in *A. thaliana* (Zazímalová et al., 2010). The ABCB/PGP/MDR and AUX1/LAX genes have been shown in the *P. patens* genome (Rensing et al., 2008). It is possible that these genes could compensate for the function a defective *PpPIND* gene, and further study of these genes is needed.

4-4-2 Functional analysis of the PpPIND protein

The GH3:GUS/*PpPIND* disruptants were produced to test whether the PpPIND protein plays a role in auxin transport or not. As a result, different expression patterns of the auxin response GH3:GUS reporter were found only in the vaginula (Fig. 27), which almost corresponded but was not an exact match to the location of disruptant phenotypes of *PpPIND* gene (Fig. 24 and Fig. 25) .The peculiar expression patterns in other organs could not be detected because of the variance between individuals. This result suggests that the PpPIND protein correlates with auxin transport in the vaginula and/or sporophyte foot. Although a change in the GUS signal was detected only in the upper region of the vaginula (Fig. 27), the auxin concentration might have been slightly changed over the entire vaginula because the gradient of the GUS signal does not conform with the auxin concentration gradient, and we cannot detect slight differences in the auxin concentration. There is the possibility that a minor change in the auxin concentration could have an effect on the junction between the vaginula and the sporophyte foot. On the other hand, it is possible that the PpPIND protein may not be an auxin flux carrier, because a relationship between GH3:GUS activity and the phytohormone abscisic acid has been reported by Fujita et al.

(2008): The change in GH3:GUS activity can be suppressed by adding abscisic acid, which is transported by the ATP-BINDING CASSETTE (ABC) transporter, which is in a superfamily also containing the ABCB/PGP/MDR auxin transporter (Fujita et al., 2008; Seo and Koshiba, 2011). Fujita et al. (2008) indicated that an auxin transport inhibitor can disturb sporophyte development in the moss *P. patens*. I tested for a disturbance of the sporophyte–gametophyte junction caused by an auxin transport inhibitor; however, no clear effect was observed because some water controls also showed a loosened the sporophyte–gametophyte junction and an insufficient number of developed sporophytes was examined. A more systematic analysis is required to detect the function of the PpPIND protein. The *A. thaliana* PIN5 protein has been tested in an auxin efflux assay using a yeast system, indicating that ArathPIN5 has an auxin export function (Mravec et al., 2009). This strategy will clarify if the PpPIND protein plays a role in auxin transport.

Chapter 5.

General Discussion

It is unclear whether the function of morphogenetic genes identified in angiosperms is conserved between long-distant lineages (e.g., between bryophytes, lycophytes, and other vascular plants) and, especially, which genes are responsible for some traits specific to the bryophytes, lycophytes, and angiosperms (Floyd and Bowman, 2007; Singer and Ashton, 2007; Rensing et al., 2008; Okano et al., 2009; Banks et al., 2011). Because of this redundancy, pseudogenization, subfunctionalization, and neofunctionalization events occurred following gene duplication over a long period of time during evolution (Zhang, 2003; Moore and Purugganan, 2005). It is difficult to identify the gene responsible for the birth of a new trait only by functional analysis. It is also a question whether molecular phylogenetic analysis including morphogenetic genes, whose function is known in some model plants of angiosperms, is a better method to select candidate genes for the responsible genes in early land plants, or not. For example, in gymnosperms, the function of the class B and class C MADS-box genes could partially or fully substitute for their angiosperm orthologous genes in complementation and ectopic expression experiments. This revealed that these gene functions were conserved over more than 300 million years of seed plant evolution, because extant gymnosperms and angiosperms diverged more than 300 million years ago (Sundström and Engström, 2001; Winter et al., 2002; Zhang et al., 2004). This result was consistent with the results of gene phylogenetic analysis. Similarly, in lycophytes and bryophytes, if the selection of genes based on molecular phylogenetic analysis can succeed in functional analysis, such an approach makes a contribution to the study of plant evolutionary genetics. I addressed these questions at the

beginning of the present study, and the results presented in this thesis could serve as a framework for future studies in order to reveal gene function in early land plants.

In Chapter 2, the *SuKNOX1* gene was expressed in the shoot apical meristem but not in the root meristem, as estimated by the molecular phylogenetic tree of *KNOX* genes (Fig. 5 and Fig. 6). The expression of the *SuKNOX1* gene was also detected in the rhizophore meristem (Fig. 6). This suggests that the rhizophore is an organ with stem characteristics, not root characteristics. It was revealed that the *SuKNOX1* gene participates in the development of the rhizophore, which is the peculiar organ to lycophyte. In this study, I indicated that the function of the class 1 *KNOX* gene, expressed in SAM, but not in the root meristem, might be conserved between lycophytes and angiosperms. Here, molecular phylogenetic analysis including the gene identified in angiosperms was shown to be a useful method to select a candidate gene for the functional analysis of molecular biology, also in early land plants. In the present study, two *KNOX*-like genes were isolated from *S. uncinata*. The phylogenetic tree of *KNOX* genes indicated that the *SuKNOX1* and *SuKNOX2* genes belong to class 1 and class 2, respectively (Fig. 5). Most of the class 1 *KNOX* genes exhibit a similar expression pattern in SAM in vascular plants, whereas class 2 *KNOX* genes have more diverse expression patterns (Bharathan et al., 1997; Reiser et al., 2000; Harrison et al., 2005; Sano et al., 2005). The analysis of the expression pattern of the class 2 *SuKNOX2* gene will likely give us a greater understanding of the mechanism of morphogenesis in *S. uncinata* and of the validity of molecular phylogenetic analysis to select a candidate functional gene.

In the Chapter 3, the phylogenetic relationships of *PIN* genes were estimated by molecular phylogenetic tree and motifs comparison (Fig. 14, Fig. 15, and Fig. 16). I suggested relationships between duplicated genes and their functional diversity among *PIN* genes and showed that the *PIN* genes of land plants consist of major three groups ("Core", "Basal", and "Bryophyta-specific" groups) (Fig. 17). Therefore, because it was predicted that the *PpPIND* gene has a presumably bryophyte-specific function, I demonstrated the result of functional analysis of the *PpPIND* gene in Chapter 4. A high expression level of the *PpPIND* gene was especially detected in shoot tips of gametophores and sporophytes in the mature sporophyte phase (Fig. 19). In addition, I could detect a phenotypic abnormality and a change in the localization of the GUS signal at the junction between the vaginula (gametophyte generation) and the sporophyte foot in *PpPIND* disruptants, and in the upper region of the vaginula in GH3:GUS/*PpPIND* disruptants, respectively (Fig. 24, Fig. 25, and Fig. 27). These results suggest that the *PpPIND* gene probably participates in subordination, in which the sporophyte is dependent on the gametophyte, and the function of the *PpPIND* gene is probably specific to bryophytes, as predicted by the molecular phylogenetic analysis of each gene. The phenotype of *PpPIND* disruptants showed incomplete dysfunction (Fig. 24), indicating the possibility that instability of the cell wall composition and structure was caused by the *PpPIND* gene defect. However, it is also possible to consider that the redundancy of functions among *PpPINA*, *PpPINB*, *PpPINC*, and *PpPIND* caused this instability. The redundancy of functions between *P. patens KNOX* genes belonging to different groups on the phylogenetic tree has been reported by Singer and Ashton (2007).

The functional analysis of three *MKN* genes, i.e. moss *KNOX* orthologous genes, has been performed by gene knockout to produce single, double and triple mutants in *P. patens*. It has been indicated that *MKN2*, a class 1 *KNOX* gene, functions in sporophyte ontogeny and spore development; mutants of *MKN4*, another class 1 *KNOX* gene, were phenotypically normal. *MKN1-3*, a class 2 *KNOX* gene, functions in spore coat formation; thus, although the two classes differ in terms of their main function, double and triple mutants indicated functional redundancy between class 1 and class 2 *KNOX* genes. This report is the first to demonstrate a case of functional redundancy between classes of *KNOX* genes. Considering this result, although functional redundancy between *PIN* genes is only known in closely related genes, i.e. *AtPIN1*, *AtPIN2*, *AtPIN3*, *AtPIN4*, and *AtPIN7* (Friml et al., 2003; Vieten et al., 2005), it is likely that functional redundancy between the "Core" and "Bryophyte-specific" groups of *PpPIN* genes is also possible by remaining partial function for their main function before the gene duplication. This might be the reason why the abnormal phenotypes of the *PpPIND* disruptants did not exhibit complete dysfunction (Fig. 24). To support this hypothesis, further study is needed to create disruptants lacking gene expression in both groups. Although well-known for closely related genes in flowering plants (Byrne et al., 2002), functional redundancy between functionally different classes has not been described. This phenomenon may peculiarly appear as an ancestral function, because ancestral functions may not be fully diversified in bryophytes, compared with flowering plants. *PpPIN* multiple disruptants may provide the ancestral functions of the *PIN* gene.

Genes as a subject of research in functional analyses are generally selected based on an expectation of function which is known in some model plant of the angiosperms. In my thesis, the study of the *KNOX* gene family in lycophytes followed this approach. However, the *PpPIND* gene, whose function is probably unknown in some model angiosperms, was selected in expectation of a bryophyte-specific function according to the results of molecular phylogenetic analysis. As a consequence, I suggest that molecular phylogenetic analysis is broadly helpful for functional analysis, not only for functionally conserved genes, but also for functionally peculiar genes in early land plants. Thus, the molecular phylogenetic analysis including a gene whose function is known in some model angiosperm plant would be helpful in the functional analysis of both of functionally conserved and unknown genes in early land plants. In addition, it will be interesting and important if the functional redundancy between clades discriminated by clearly different function exists in bryophytes.

References

Abe J, Hiwatashi Y, Ito M, Hasebe M, Sekimoto H (2008) Expression of exogenous genes under the control of endogenous HSP70 and CAB promoters in the Closterium peracerosum–strigosum–littorale complex. Plant Cell Physiol 49: 625-632

Abe J, Hori S, Tsuchikane Y, Kitao N, Kato M, Sekimoto H (2011) Stable nuclear transformation of the *Closterium peracerosum-strigosum-littorale* complex. Plant Cell Physiol 52: 1676-1585

Adachi J, Hasegawa M (1992-1996) ProtML 2.3b3 maximum likelihood inference of protein phylogeny. The Institute of Statistical Mathematics, Tokyo

Ashton NW, Cove DJ (1977) The isolation and preliminary characterisation of auxotrophic and analogue resistant mutants in the moss *Physcomitrella patens*. Molecular and General Genetics 154: 87-95

Banks JA (2009) Selaginella and 400 million years of separation. Annu Rev Plant Biol 60: 223-238

Banks JA, Nishiyama T, Hasebe M, Bowman JL, Gribskov M, et al. (2011) The Selaginella genome identifies genetic changes associated with the evolution of vascular plants. Science 332: 960-963

Bainbridge K, Guyomarc'h S, Bayer E, Swarup R, Bennett M, Mandel T, Kuhlemeier C (2008) Auxin influx carriers stabilize phyllotactic patterning. Genes Dev 22: 810-823

Barak LS, Tiberi M, Freedman NJ, Kwatra MM, Lefkowitz RJ, Caron MG (1994) A highly conserved tyrosine residue inGprotein-coupled receptors is required for agonist-mediated beta 2-adrenergic receptor sequestration. J. Biol. Chem. 269: 2790-2795

Barton MK, RS Poethig (1993) Formation of the shoot apical meristem in Arabidopsis thaliana: an analysis of development in the wild type and in the shoot meristemless mutant Development 119: 823-831

Becker B, Marin B (2009) Streptophyte algae and the evolution of the embryophytes. Ann. Bot. 103: 999-1004

Benková LE, Michniewicz M, Sauer M, Teichmann T, Seifertová LD, Jürgens G, Friml J (2003) Local, efflux-dependent auxin gradients as a common module for plant organ formation. Cell 115: 591-602

Bertolino E, Reimund B, WildtPerinic D, Clerc RG (1995) A novel homeobox protein which recognizes a TGT core and functionally interferes with a retinoid-responsive motif. J Biol Chem 270: 31178-31188

Bharathan G, Janssen BJ, Kellogg EA, Sinha NR (1997) Did homeodomain proteins duplicate before the origin of angiosperms, fungi, and metazoa? Proc Natl Acad Sci USA 94: 13749-13753

Bharathan G, Goliber TE, Moore C, Kessler S, Pham T, Sinha NR (2002) Homologies in leaf form inferred from *KNOXI* gene expression during development. Science 296: 1858-1860

Blakeslee JJ, Bandyopadhyay A, Lee OR, et al. (2007) Interactions among PIN-FORMED and P-glycoprotein auxin transporters in Arabidopsis. Plant Cell 19: 131-147

Burglin TR (1998) The PBC domain contains a MEINOX domain: Coevolution of Hox and TALE homeobox genes? Dev Genes Evol 208: 113-116

Byrne ME, Barley R, Curtis M, Arroyo JM, Dunham M, Hudson A, Martienssen RA (2000) *Asymmetric leaves1* mediates leaf patterning and stem cell function in *Arabidopsis*. Nature 408: 967-971

Byrne ME, Simorowski J, Martienssen RA (2002) ASYMMETRIC LEAVES1 reveals knox gene redundancy in Arabidopsis. Development 129:1957–1965

Champagne CEM, Ashton NW (2001) Ancestry of *KNOX* genes revealed by bryophyte (*Physcomitrella patens*) homologs. New Phytol 150: 23-36

Cove DJ, Knight CD (1993) The moss Physcomitrella patens, a model system with potential for the study of plant reproduction. Plant Cell 5: 1483-1488

Cove DJ, Knight CD, Lamparter T (1997) Mosses as model systems. Trends Plant Sci. 2: 99-105

De Smet I, Voss U, Lau S, Wilson M, Shao N, Timme R, Swarup R, Kerr I, Hodgman

C, Bock R, Bennett M, Ju¨rgens G, Beeckman T (2011) Unraveling the evolution of auxin signaling. Plant Physiol 155: 209-221

Dhonukshe P, Huang F, Galvan-Ampudia CS, Mähönen AP, Kleine-Vehn J, Xu J, Quint A, Prasad K, Friml J, Scheres B, et al. (2010) Plasma membrane-bound AGC3 kinases phosphorylate PIN auxin carriers at TPRXS(N/S) motifs to direct apical PIN recycling. Development 137: 3245-3255

Estelle M (1992) The plant hormone auxin: insight in sight. Bioessays 14: 439-444

Feraru E, Feraru MI, Kleine-Vehn J, Martiniere A, Mouille G, Vanneste S, Vernhettes S, Runions J, Friml J (2011) PIN polarity maintenance by the cell wall in Arabidopsis. Curr Biol 21: 338-343

Ferrario S, Immink RG, Shchennikova A, Busscher-Lange J, Angenent GC (2003) The MADS box gene FBP2 is required for SEPALLATA function in petunia. Plant Cell 15: 914-925

Floyd SK, Bowman JL (2007) The ancestral patterning tool kit of land plants. Int. J. Plant Sci. 168: 1-35

Floyd SK, Zalewski CS, Bowman JL (2006) Evolution of class III homeodomain-leucine zipper genes in streptophytes. Genetics 173: 373-388

Friml J, Benková L E, Blilou I, Wísniewska J, Hamann T, Ljung K, Woody S, Sandberg G, Scheres B, Jürgens G, Palme K (2002a) AtPIN4 mediates sink-driven auxin gradients and root patterning in *Arabidopsis*. Cell 108: 661-673

Friml J, Vieten A, Sauer M, Weijers D, Schwarz H, Hamann T, Offringa R, Jürgens G (2003) Efflux-dependent auxin gradients establish the apical-basal axis of *Arabidopsis*. Nature 426: 147-153

Friml J, Wísniewska J, Benková E, Mendgen K, Palme K (2002b) Lateral relocation of auxin efflux regulator PIN3 mediates tropism in *Arabidopsis*. Nature 415: 806-809

Fujita T, Sakaguchi H, Hiwatashi Y, Wagstaff SJ, Ito M, Deguchi H, Sato T, Hasebe M (2008) Convergent evolution of shoots in land plants: lack of auxin polar transport in moss shoots. Evol. Dev. 10: 176-86

Gälweiler L, Guan C, Müller A, Wisman E, Mendgen K, Yephremov A, Palme K (1998) Regulation of polar auxin transport by AtPIN1 in *Arabidopsis* vascular tissue. *Science* 282: 2226-2230

Gamborg OL, Miller RA, Ojima K (1968) Nutrient requirements of suspension cultures of soybean root cells. Exp Cell Res 50: 151-158

Geisler M, Murphy AS (2006) The ABC of auxin transport: the role of p-glycoproteins in plant development. FEBS Lett. 580: 1094-1102

Gifford EM, Foster AS (1989) Morphology and Evolution of Vascular Plants. Freeman, New York

Gogarten JP, Kibak H, Dittrich P, Taiz L, Bowman BJ, Manolson MF, Poole RJ, Date T, Oshima T, Konishi J, Denda K, Oshida M (1989) Evolution of the vacuolar H+-ATPase, implications for the origin of eukaryotes. Proc. Natl. Acad. Sci. USA 68: 6661-6685

Grimsley NH, Ashton NW, Cove DJ (1977) The production of somatic hybrids by protoplast fusion in the moss, *Physcomitrella patens*. Molecular and General Genetics 154: 97-100

Gunning BES, Pate JS (1969) "Transfer cells": Plant cells with wall ingrowths, specialized in relation to short distance transport of solutes—Their occurrence, structure, and development. Protoplasma 68: 107-133

Hake S, Vollbrecht E, Freeling M (1989) Cloning Knotted, the dominant morphological mutant in maize using *Ds2* as a transposon tag. EMBO J 8: 15-22

Hareven D, Gutfinger T, Parnis A, Eshed Y, Lifschitz E (1996) The making of a compound leaf: genetic manipulation of leaf architecture in tomato. Cell 84: 735-744

Harrison CJ, Corley SB, Moylan EC, Alexander DL, Scotland RW, Langdale JA (2005) Independent recruitment of a conserved developmental mechanism during leaf evolution. Nature 434: 509-514

Hasegawa M, Kishino H (1994) Accuracies of the simple methods for estimating the bootstrap probability of a maximum-likelihood tree. Mol Biol Evol 11: 142-145

Huang F, Kemel Zago M, Abas L, van Marion A, Galvan Ampudia CS, Offringa R

(2010) Phosphorylation of conserved PIN motifs directs Arabidopsis PIN1 polarity and auxin transport. Plant Cell 22: 1129-1142

Imaichi R, Kato M (1989) Developmental anatomy of the shoot apical cell, rhizophore and root of *Selaginella-uncinata*. Bot Mag Tokyo 102: 369-380

Imaichi R, Kato M (1991) Developmental-study of branched rhizophores in 3 *Selaginella* species. Am J Bot 78: 1694-1703

Iwabe N, Kuma K, Hasegawa M, Osawa S, Miyata T (1989) Evolutionary relationship of archaebacteria, eubacteria and eukaryotes inferred from phylogenetic trees of duplicated genes. Proc. Natl. Acad. Sci. USA 86: 9355-9359

Jack T, Brockman LL, Meyerowitz EM (1992) The homeotic gene *APETALA3* of *Arabidopsis thaliana* encodes a MADS-box and is expressed in petal and stamens. Cell 68: 683-897

Janssen BJ, Williams A, Chen JJ, Mathern J, Hake S, Sinha N (1998) Isolation and characterization of two knotted-like homeobox genes from tomato. Plant Mol Biol 36: 417-425

Jones DT, Taylor WR, Thornton JM (1992) The rapid generation of mutation data matrices from protein sequences. Comput Appl Biosci 8: 275-282

Kammerer W, Cove DJ (1996) Genetic analysis of the result of retransformation of transgenic lines of the moss, *Physcomitrella patens*. Mol Gen Genet 250: 380-382

Katoh K, Toh H (2008) Recent developments in the MAFFT multiple sequence alignment program. Brief Bioinform 9: 286-298

Kenrick P, Crane PR (1997) The origin and early evolution of plants on land. Nature 389: 33-39

Kerr ID, Bennett MJ (2007) New insight into biochemical mechanisms regulating auxin transport in plants. Biochem. J. 401: 613-622

Keuskamp DH, Pollmann S, Voesenek LACJ, Peeters AJM, Pierik R (2010) Auxin transport through PIN-FORMED 3 (PIN3) controls shade avoidance and fitness during competition. Proc. Natl Acad. Sci. USA, 107: 22740-22744

Kim S, Koh J, Yoo MJ, Kong H, Hu Y, Ma H, Soltis PS, Soltis DE (2005) Expression of floral MADS-box genes in basal angiosperms: Implications on evolution of floral regulators and the perianth. Plant Journal 43: 724-744

Kimura M (1968) Evolutionary rate at the molecular level. Nature 217: 624-626

Kimura M (1983) The Neutral Theory of Molecular Evolution. Cambridge Univ. Press, Cambridge, U.K.

Knight CD, Cove DJ, Boyd PJ, Ashton NW (1988) The isolation of biochemical and developmental mutants in *Physcomitrella patens*. Methods in Bryology. Glime, J. M. Miyazaki, Japan, Hattori Botanical Laboratory: 47-58

Kolukisaoglu HU, Bovet L, Klein M, Eggmann T, Geisler M, Wanke D, Martinoia E, Schulz B (2002) Family business: the multidrugresistance related protein (MRP) ABC transporter genes in Arabidopsis thaliana. Planta 216: 107-119

Kramer EM, Jaramillo MA, Di Stilio VS (2004) Patterns of gene duplication and functional evolution during the diversification of the AGAMOUS subfamily of MADS box genes in angiosperms. Genetics 166: 1011-1023

Krecek P, Skupa P, Libus J, Naramoto S, Tejos R, Friml J, Zazímalová E (2009) The PIN-FORMED (PIN) protein family of auxin transporters. Genome Biol 10: 249

Krem MM, Di Cera E (2001) Molecular markers of serine protease evolution. EMBO J 20: 3036-3045

Kumar S (2005) Molecular clocks: four decades of evolution. Nature Rev. Genet. 6: 654-662

Langdale JA, Harrison CJ (2008) Developmental transitions during the evolution of plant form. Cambridge University Press. pp. 297-314

Ligrone R, Duckett JG, Renzaglia KS (1993) The gametophyte^sporophyte junction in land plants. Adv. Bot. Res. 19: 231-317

Long JA, Moan EI, Medford JI, Barton MK (1996) A member of the KNOTTED class of homeodomain proteins encoded by the *STM* gene of *Arabidopsis*. Nature 379: 66-69

Luschnig C, Gaxiola RA, Grisafi P, Fink GR (1998) EIR1, a root-specific protein involved in auxin transport, is required for gravitropism in *Arabidopsis thaliana*. Genes

Dev 12: 2175-2187

Ma H, Yanofsky MF, Meyerowitz EM (1991) AGL1-AGL6, an Arabidopsis gene family with similarity to floral homeotic and transcription factor genes. Genes Dev. 5: 484-495

McManus MT, Thompson DS, Merriman, Cledwyn LL, Osbourne DJ (1998) Transdifferentiation of mature cortical cells to functional abscission cells in bean. Plant Physiol 116: 891-899

McCourt RM, Delwiche CF, Karol KG (2004) Charophyte algae and land plant origins. Trends in Ecology & Evolution 19: 661-666

Michniewicz M, Zago MK, Abas L,Weijers D, Schweighofer A, et al. (2007) Antagonistic Regulation of PIN Phosphorylation by PP2A and PINOID Directs Auxin Flux. Cell 130: 1044-1056

Miyashita Y, Takasugi T, Ito Y (2010) Identification and expression analysis of PIN genes in rice. Plant Sci. 178: 424-428

Moore RC, Grant SR, Purugganan MD (2005) Molecular population genetics of redundant floral regulatory genes in Arabidopsis thaliana. Mol Biol Evol 22: 91-103

Moore RC, Purugganan MD (2005) The evolutionary dynamics of plant duplicate genes. Curr. Opin. Plant Biol. 8: 122-128

Müller A, Guan C, Gälweiler L, Tanzler P, Huijser P, Marchant A, Parry G, Bennett M, Wisman E, Palme K (1998) AtPIN2 defines a locus of *Arabidopsis* for root gravitropism control. EMBO J 17: 6903-6911

Murray MG, Thompson WF (1980) Rapid isolation of high molecular weight plant DNA. Nucleic Acids Research 8: 4321-4325

Mravec J, Skupa P, Bailly A, Hoyerová K, Krecek P, Bielach A, Petrásek J, Zhang J, Gaykova V, Stierhof YD, Dobrev PI, Schwarzerová K, Rolcík J, Seifertová D, Luschnig C, Benková E, Zazímalová E, Geisler M, Friml J (2009) Subcellular homeostasis of phytohormone auxin is mediated by the ER-localized PIN5 transporter. Nature 459: 1136-1140

Nishiyama T, Hiwatashi Y, Sakakibara K, Kato M, Hasebe M (2000) Tagged mutagenesis

and gene-trap in the moss, *Physcomitrella patens* by Shuttle mutagenesis. DNA Research 7: 9-17

Okada K, Ueda J, Komaki MK, Bell CJ, Shimura Y (1991) Requirement of the auxin polar transport system in early stages of *Arabidopsis* floral bud formation. Plant Cell 3: 677-684

Okano Y, Aono N, Hiwatashi Y, Murata T, Nishiyama T, Ishikawa T, Kubo M, Hasebe M (2009) A polycomb repressive complex 2 gene regulates apogamy and gives evolutionary insights into early land plant evolution. Proc. Natl. Acad. Sci. USA 106: 16321-16326

Palovaara J, Hallberg H, Stasolla C, Luit B, Hakman I (2010) Expression of a gymnosperm PIN homologous gene correlates with auxin immunolocalization pattern at cotyledon formation and in demarcation of the procambium during *Picea abies* somatic embryo development and in seedling tissues. Tree Physiology 30: 479-489

Paponov IA, Teale WD, Trebar M, Blilou K, Palme K (2005) The PIN auxin efflux facilitators: evolutionary and functional perspectives. Trends Plant Sci. 10: 170-177

Prigge MJ, Clark SE (2006) Evolution of the class III HD-Zip gene family in land plants. Evol Dev 8: 350-361

Pryer KM, Schneider H, Smith AR, Cranfill R, Wolf PG, Hunt JS, Sipes SD (2001) Horsetails and ferns are a monophyletic group and the closest living relatives to seed plants. Nature 409: 618-622

Rambaut A, Drummond AJ (2004) Tracer. Version 1.3: MCMC Trace Analysis Tool. University of Oxford, Oxford, UK.

Raubeson LA, Jansen RK (1992) Chloroplast DNA evidence on the ancient evolutionary split in vascular land plants. Science 255: 1697-1699

R Development Core Team (2010) R: A Language and Environment for Statistical Computing (Vienna: R Foundation for Statistical Computing).

Reinhardt D, Pesce ER, Stieger P, Mandel T, Baltensperger K, Bennett M, Traas J, Friml J, Kuhlemeier C (2003) Regulation of phyllotaxis by polar auxin transport.

Nature 426: 255-260

Reiser L, Sanchez-Baracaldo P, Hake S (2000) Knots in the family tree: evolutionary relationships and functions of *knox* homeobox genes. Plant Mol Biol 42: 151-166

Rensing SA, Lang D, Zimmer AD, Terry A, Salamov A, et al. (2008) The *Physcomitrella* genome reveals evolutionary insights into the conquest of land by plants. Science 319: 64-69

Reski R (1998) *Physcomitrella* and *Arabidopsis*: the David and Goliath of reverse genetics. Trends Plant Sci 3: 209-210

Rokas A, Kathirithamby J, Holland PW (1999) Intron insertion as a phylogenetic character: the engrailed homeobox of Strepsiptera does not indicate affinity with Diptera. Insect Mol Biol 8: 527-530

Ronquist F, Huelsenbeck JP (2003) MrBayes 3: Bayesian phylogenetic inference under mixed models Bioinformatics. 19: 1572-1574

Sano R, Juarez CM, Hass B, Sakakibara K, Ito M, Banks JA, Hasebe M (2005) KNOX homeobox genes potentially have similar function in both diploid unicellular and multicellular meristems, but not in haploid meristems. Evol Dev 7: 69-78

Schaefer DG (2001) Gene targeting in Physcomitrella patens. Curr Opin Plant Biol 4: 143-150

Schiavone FM, Cooke TJ (1987) Unusual patterns of somatic embryogenesis in the domesticated carrot: developmental effects of exogenous auxins and auxin transport inhibitors. Cell Diff 2: 53-62

Schneeberger R, Tsiantis M, Freeling M, Langdale JA (1998) The *rough sheath2* gene negatively regulates homeobox gene expression during maize leaf development. Development 125: 2857-2865

Schumaker KD, Dietrich MA (1998) Hormone-induced signaling during moss development. Annu. Rev. Plant Physiol. Plant Mol. Biol. 49: 501-523

Sekimoto H, Satoh S, Fujii T (1990) Biochemical and physiological properties of a protein inducing protoplast release during conjugation in the Closterium peracerosum-

strigosum-littorale complex. Planta 182: 348-354

Sekimoto H, Tanabe Y, Tsuchikane Y, Shirosaki H, Fukuda H, Demura T, Ito M (2006) Gene expression profiling using cDNA microarray analysis of the sexual reproduction stage of the unicellular charophycean alga Closterium peracerosum–strigosum–littorale complex. Plant Physiol. 141: 271-279

Seo M, Koshiba T (2011) Transport of ABA from the site of biosynthesis to the site of action. J Plant Res. 124(4): 501-7

Serikawa KA, Mandoli DF (1999) *Aaknox1*, a *kn1*-like homeobox gene in *Acetabularia acetabulum*, undergoes developmentally regulated subcellular localization. Plant Mol Biol 41: 785-793

Shigyo M, Ito M (2004) Analysis of gymnosperm two-AP2-domain-containing genes. Dev. Genes Evol. 214: 105-114

Singer SD, Ashton NW (2007) Revelation of ancestral roles of KNOX genes by a functional analysis of Physcomitrella homologues. Plant Cell Rep 26: 2039-2054

Singer SD, Krogan NT, Ashton NW (2007) Clues about the ancestral roles of plant MADS-box genes from a functional analysis of moss homologues. Plant Cell Rep 26: 1155-1169

Smith GM (1955) Cryptogamic Botany: Volume II, Bryophytes and Pteridophytes. New York: McGraw-Hill.

Stewart WN, Rothwell GW (1993) Paleobotany and the Evolution of Plants. Cambridge University Press, Cambridge

Sundström J, Engström P (2002) Conifer reproductive development involves B-type MADS-box genes with distinct and different activities in male organ primordia. The Plant Journal 31: 161-169

Tamura K, Peterson D, Peterson N, Stecher G, Nei M, Kumar S (2011) MEGA5: molecular evolutionary genetics analysis using maximum likelihood, evolutionary distance, and maximum parsimony methods. Mol Biol Evol. 28: 2731-2739

Tanabe Y, Uchida M, Hasebe M, Ito M (2003) Characterization of the *Selaginella*

remotifolia MADS-box gene. J Plant Res 116: 71-75

Thompson J, Higgins DG, Gibson TJ (1994) CLUSTAL W: improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position specific gap penalties and weight matrix choice. Nucleic Acids Res 22: 4673-4680

Turmel M, Otis C, Lemieux C (2006) The chloroplast genome sequence of Chara vulgaris sheds new light into the closest green algal relatives of land plants. Molecular Biology and Evolution 23: 1324-1338

Uzawa M, Higuchi M (2010) Comparative development of the sporophyte–gametophyte junction in six moss species. J. Plant Res. 123(6): 777-787

Vieten A, Vanneste S, Wísniewska J, Benková E, Benjamins R, Beeckman T, Luschnig C, Friml J (2005) Functional redundancy of PIN proteins is accompanied by auxin-dependent crossregulation of PIN expression. Development 132: 4521-4531

Vollbrecht E, Veit B, Sinha N, Hake S (1991) The developmental gene *Knotted-1* is a member of a maize homeobox gene family. Nature 350: 241-243

Wada H, Kobayashi M, Sato R, Satoh N, Miyasaka H, Shirayama Y (2002) Dynamic insertion-deletion of introns in deuterostome EF-1 alpha genes. J. Mol. Evol 54: 1181-1128

Wang J, Hu H, Wang G, Li J, Chen J, Wu P (2009) Expression of PIN genes in rice (*Oryza sativa* L.): Tissue specificity and regulation by hormones. Mol Plant. 2: 823-831

Wang W, Tanurdzic M, Luo M, Sisneros N, Kim HR, Weng JK, Kudrna D, Mueller C, Arumuganathan K, Carlson J, Chapple C, Pamphilis C, Mandoli D, Tomkins J, Wing RA, Banks JA (2005) Construction of a bacterial artificial chromosome library from the spikemoss *Selaginella moellendorffii*: A new resource for plant comparative genomics. BMC Plant Biol 5: 10

Webster TR, Jagels R (1977) Morphology and development of aerial roots of *Selaginella-martensii* grown in moist containers. Can J Bot/Rev Can Bot 55: 2149-2158

Weng JK, Tanurdzic M, Chapple C (2005) Functional analysis and comparative genomics of expressed sequence tags from the lycophyte *Selaginella moellendorffii*. BMC

Genomics 6: 85

Whelan S, Goldman N (2001) A general empirical model of protein evolution derived from multiple protein families using a maximum-likelihood approach. Mol Biol Evol 18: 691-699

Winter KU, Saedler H, Theissen G (2002) On the origin of class B floral homeotic genes: functional substitution and dominant inhibition in Arabidopsis by expression of an orthologue from the gymnosperm Gnetum. The Plant Journal 31: 457-475

Wochok ZS, Sussex IM (1975) Morphogenesis in *Selaginella*: III. Meristem determination and cell differentiation. Dev Biol 47: 376-383

Wodniok S, Brinkmann H, Glockner G, Heidel A, Philippe H, Melkonian M, Becker B (2011) Origin of land plants: Do conjugating green algae hold the key? BMC Evol. Biol. 11: 104

Woodward AW, Bartel B (2005) Auxin: regulation, action, and interaction. Ann. Bot. 95: 707-35

Zazímalová E, Krerek P, Skupa P, Hoyerová K, Petrásek J (2007) Polar transport of the plant hormone auxin- the role of PIN-FORMED (PIN) proteins. Cell. Mol. Life Sci. 64: 1621-1637

Zazímalová E, Murphy AS, Yang H, Hoyerova K, Hosek P (2010) Auxin transporters-why so many? Cold Spring Harbor Perspect. Biol. 2: a001552

Zhang J (2003) Evolution by gene duplication—an update. Trends Ecol. Evol. 18: 292-298

Zhang J, Nodzynski T, Pencík A, Rolcík J, Friml J (2010) PIN phosphorylation is sufficient to mediate PIN polarity and direct auxin transport. Proc. Natl. Acad. Sci. USA 107: 918-922

Zhang PY, Tan HTW, Pwee KH, Kumar PP (2004) Conservation of class C function of floral organ development during 300 million years of evolution from gymnosperms to angiosperms. The Plant Journal 37: 566-577

Zimmermann R, Werr W (2005) Pattern formation in the monocot embryo as revealed by NAM and CUC3 orthologues from *Zea mays* L. Plant Mol. Biol. 58: 669-685

Zuckerkandl E, Pauling L (1965) Evolutionary Divergence and Convergence in Proteins. in Bryson V and Vogel HJ, eds., Evolving Genes and Proteins. Academic Press, New York, pp. 97-166

Zuckerkandl E, Pauling L (1962) Molecular disease, evolution, and genic heterogeneity. in Kasha M and Pullman B, eds, Horizons in Biochemistry. Academic Press, New York, pp. 189-225

Table 1. *KNOX* genes

Accession no.	Gene name	Species	Fmily	class	Subdivision	Division
U32344	STM	*Arabidopsis thaliana*	Brassicaceae (Mustard)	Dicotyledoneae (Eudicot)	Angiospermae (Angiosperm)	Spermatophyta (Seed plant)
AF482995	KNAT1	*Arabidopsis thaliana*	Brassicaceae (Mustard)	Dicotyledoneae (Eudicot)	Angiospermae (Angiosperm)	Spermatophyta (Seed plant)
X81353	KNAT2/ATK1	*Arabidopsis thaliana*	Brassicaceae (Mustard)	Dicotyledoneae (Eudicot)	Angiospermae (Angiosperm)	Spermatophyta (Seed plant)
AY086091	KNAT3	*Arabidopsis thaliana*	Brassicaceae (Mustard)	Dicotyledoneae (Eudicot)	Angiospermae (Angiosperm)	Spermatophyta (Seed plant)
X92393	KNAT4	*Arabidopsis thaliana*	Brassicaceae (Mustard)	Dicotyledoneae (Eudicot)	Angiospermae (Angiosperm)	Spermatophyta (Seed plant)
AF306661	KNAT5	*Arabidopsis thaliana*	Brassicaceae (Mustard)	Dicotyledoneae (Eudicot)	Angiospermae (Angiosperm)	Spermatophyta (Seed plant)
AB072361	KNAT6	*Arabidopsis thaliana*	Brassicaceae (Mustard)	Dicotyledoneae (Eudicot)	Angiospermae (Angiosperm)	Spermatophyta (Seed plant)
AF308451	KNAT7	*Arabidopsis thaliana*	Brassicaceae (Mustard)	Dicotyledoneae (Eudicot)	Angiospermae (Angiosperm)	Spermatophyta (Seed plant)
X61308	KN1	*Zea mays*	Poaceae (Grass)	Monocotyledoneae (Monocot)	Angiospermae (Angiosperm)	Spermatophyta (Seed plant)
AAA86287	RS1	*Zea mays*	Poaceae (Grass)	Monocotyledoneae (Monocot)	Angiospermae (Angiosperm)	Spermatophyta (Seed plant)
AF100455	LG3	*Zea mays*	Poaceae (Grass)	Monocotyledoneae (Monocot)	Angiospermae (Angiosperm)	Spermatophyta (Seed plant)
AF457118	LG4a	*Zea mays*	Poaceae (Grass)	Monocotyledoneae (Monocot)	Angiospermae (Angiosperm)	Spermatophyta (Seed plant)
P56659	KNOX1	*Zea mays*	Poaceae (Grass)	Monocotyledoneae (Monocot)	Angiospermae (Angiosperm)	Spermatophyta (Seed plant)
AAB33488	KNOX4	*Zea mays*	Poaceae (Grass)	Monocotyledoneae (Monocot)	Angiospermae (Angiosperm)	Spermatophyta (Seed plant)
P56663	KNOX5	*Zea mays*	Poaceae (Grass)	Monocotyledoneae (Monocot)	Angiospermae (Angiosperm)	Spermatophyta (Seed plant)
AAB33490	KNOX6	*Zea mays*	Poaceae (Grass)	Monocotyledoneae (Monocot)	Angiospermae (Angiosperm)	Spermatophyta (Seed plant)
P56666	KNOX8	*Zea mays*	Poaceae (Grass)	Monocotyledoneae (Monocot)	Angiospermae (Angiosperm)	Spermatophyta (Seed plant)
AAB33489	KNOX10	*Zea mays*	Poaceae (Grass)	Monocotyledoneae (Monocot)	Angiospermae (Angiosperm)	Spermatophyta (Seed plant)
U90091	SKN1	*Picea mariana*	Pinaceae (Pine)	Coniferopsida	Gimnospermae (Gymnosperm)	Spermatophyta (Seed plant)
U90092	SKN2	*Picea mariana*	Pinaceae (Pine)	Coniferopsida	Gimnospermae (Gymnosperm)	Spermatophyta (Seed plant)
AAV63997	PmKN3	*Picea mariana*	Pinaceae (Pine)	Coniferopsida	Gimnospermae (Gymnosperm)	Spermatophyta (Seed plant)
AB043954	CRKNOX1	*Ceratopteris richardii*	Adiantaceae	Pteridopsida		Pteridophyta (fern)
AB043956	CRKNOX2	*Ceratopteris richardii*	Adiantaceae	Pteridopsida		Pteridophyta (fern)
AB043957	CRKNOX3	*Ceratopteris richardii*	Adiantaceae	Pteridopsida		Pteridophyta (fern)
AB288208*	SuKNOX1	*Selaginella uncinata*	Selaginellaceae (Spike moss)	Lycopodiopsida (Club moss)		Lycopodiophyta
AB288209*	SuKNOX2	*Selaginella uncinata*	Selaginellaceae (Spike moss)	Lycopodiopsida (Club moss)		Lycopodiophyta
AY667449	SkKNOX1	*Selaginella kraussiana*	Selaginellaceae (Spike moss)	Lycopodiopsida (Club moss)		Lycopodiophyta
AY667450	SkKNOX2	*Selaginella kraussiana*	Selaginellaceae (Spike moss)	Lycopodiopsida (Club moss)		Lycopodiophyta
AY667451	SkKNOX3	*Selaginella kraussiana*	Selaginellaceae (Spike moss)	Lycopodiopsida (Club moss)		Lycopodiophyta
AF285148	MKN1-3	*Physcomitrella patens* subsp. *patens*	Funariaceae	Bryopsida (moss)		Bryophyta
AF285147	MKN2	*Physcomitrella patens* subsp. *patens*	Funariaceae	Bryopsida (moss)		Bryophyta
AF284817	MKN4	*Physcomitrella patens* subsp. *patens*	Funariaceae	Bryopsida (moss)		Bryophyta
AF170172	AaKNOX1	*Acetabularia acetabulum*	Polyphysaceae	Ulvophyceae		Chlorophyta
CCD63401	CEH-20	*Caenorhabditis elegans*	Caenorhabditis	Secernentea		Nematoda
CAA92154	CEH-40	*Caenorhabditis elegans*	Caenorhabditis	Secernentea		Nematoda
BAA05957	PBX2	*Homo sapiens*	Hominidae	Mammalia	Vertebrata	Chordata
AAH94883	PBX3	*Homo sapiens*	Hominidae	Mammalia	Vertebrata	Chordata
CAC28212	PBX4	*Homo sapiens*	Hominidae	Mammalia	Vertebrata	Chordata

* is isolated from this study.

Table 2. *PIN* genes

KEGG Entry name	Gene name	Species	Fmily	Class	Subdivision	Division
AT1G23080	ArathPIN7	*Arabidopsis thaliana*	Brassicaceae (Mustard)	Dicotyledoneae (Eudicot)	Angiospermae (Angiosperm)	Spermatophyta (Seed plant)
AT1G70940	ArathPIN3	*Arabidopsis thaliana*	Brassicaceae (Mustard)	Dicotyledoneae (Eudicot)	Angiospermae (Angiosperm)	Spermatophyta (Seed plant)
AT1G73590	ArathPIN1	*Arabidopsis thaliana*	Brassicaceae (Mustard)	Dicotyledoneae (Eudicot)	Angiospermae (Angiosperm)	Spermatophyta (Seed plant)
AT1G77110	ArathPIN6	*Arabidopsis thaliana*	Brassicaceae (Mustard)	Dicotyledoneae (Eudicot)	Angiospermae (Angiosperm)	Spermatophyta (Seed plant)
AT2G01420	ArathPIN4	*Arabidopsis thaliana*	Brassicaceae (Mustard)	Dicotyledoneae (Eudicot)	Angiospermae (Angiosperm)	Spermatophyta (Seed plant)
AT5G15100	ArathPIN8	*Arabidopsis thaliana*	Brassicaceae (Mustard)	Dicotyledoneae (Eudicot)	Angiospermae (Angiosperm)	Spermatophyta (Seed plant)
AT5G16530	ArathPIN5	*Arabidopsis thaliana*	Brassicaceae (Mustard)	Dicotyledoneae (Eudicot)	Angiospermae (Angiosperm)	Spermatophyta (Seed plant)
AT5G57090	ArathEIR1/ PIN2	*Arabidopsis thaliana*	Brassicaceae (Mustard)	Dicotyledoneae (Eudicot)	Angiospermae (Angiosperm)	Spermatophyta (Seed plant)
POPTR 231887	PoptrPIN3	*Populus trichocarpa*	Salicaceae (Willow)	Dicotyledoneae (Eudicot)	Angiospermae (Angiosperm)	Spermatophyta (Seed plant)
POPTR 560374	PoptrPIN8	*Populus trichocarpa*	Salicaceae (Willow)	Dicotyledoneae (Eudicot)	Angiospermae (Angiosperm)	Spermatophyta (Seed plant)
POPTR 728847	PoptrPIN7	*Populus trichocarpa*	Salicaceae (Willow)	Dicotyledoneae (Eudicot)	Angiospermae (Angiosperm)	Spermatophyta (Seed plant)
POPTR 734743	PoptrPIN2	*Populus trichocarpa*	Salicaceae (Willow)	Dicotyledoneae (Eudicot)	Angiospermae (Angiosperm)	Spermatophyta (Seed plant)
POPTR 759514	PoptrPIN13	*Populus trichocarpa*	Salicaceae (Willow)	Dicotyledoneae (Eudicot)	Angiospermae (Angiosperm)	Spermatophyta (Seed plant)
POPTR 774965	PoptrPIN15	*Populus trichocarpa*	Salicaceae (Willow)	Dicotyledoneae (Eudicot)	Angiospermae (Angiosperm)	Spermatophyta (Seed plant)
POPTR 780408	PoptrPIN12	*Populus trichocarpa*	Salicaceae (Willow)	Dicotyledoneae (Eudicot)	Angiospermae (Angiosperm)	Spermatophyta (Seed plant)
POPTR 797521	PoptrPIN10	*Populus trichocarpa*	Salicaceae (Willow)	Dicotyledoneae (Eudicot)	Angiospermae (Angiosperm)	Spermatophyta (Seed plant)
POPTR 798645	PoptrPIN5	*Populus trichocarpa*	Salicaceae (Willow)	Dicotyledoneae (Eudicot)	Angiospermae (Angiosperm)	Spermatophyta (Seed plant)
POPTR 803601	PoptrPIN6	*Populus trichocarpa*	Salicaceae (Willow)	Dicotyledoneae (Eudicot)	Angiospermae (Angiosperm)	Spermatophyta (Seed plant)
POPTR 809416	PoptrPIN9	*Populus trichocarpa*	Salicaceae (Willow)	Dicotyledoneae (Eudicot)	Angiospermae (Angiosperm)	Spermatophyta (Seed plant)
POPTR 824601	PoptrPIN1	*Populus trichocarpa*	Salicaceae (Willow)	Dicotyledoneae (Eudicot)	Angiospermae (Angiosperm)	Spermatophyta (Seed plant)
POPTR 827532	PoptrPIN11	*Populus trichocarpa*	Salicaceae (Willow)	Dicotyledoneae (Eudicot)	Angiospermae (Angiosperm)	Spermatophyta (Seed plant)
POPTR 831533	PoptrPIN4	*Populus trichocarpa*	Salicaceae (Willow)	Dicotyledoneae (Eudicot)	Angiospermae (Angiosperm)	Spermatophyta (Seed plant)
RCOM 0583560		*Ricinus communis*	Euphorbiaceae (Spurge)	Dicotyledoneae (Eudicot)	Angiospermae (Angiosperm)	Spermatophyta (Seed plant)
RCOM 0843030		*Ricinus communis*	Euphorbiaceae (Spurge)	Dicotyledoneae (Eudicot)	Angiospermae (Angiosperm)	Spermatophyta (Seed plant)
RCOM 1437510		*Ricinus communis*	Euphorbiaceae (Spurge)	Dicotyledoneae (Eudicot)	Angiospermae (Angiosperm)	Spermatophyta (Seed plant)
vvi:100242778		*Vitis vinifera*	Vitaceae (Grape)	Dicotyledoneae (Eudicot)	Angiospermae (Angiosperm)	Spermatophyta (Seed plant)
vvi:100244520		*Vitis vinifera*	Vitaceae (Grape)	Dicotyledoneae (Eudicot)	Angiospermae (Angiosperm)	Spermatophyta (Seed plant)
vvi:100249181		*Vitis vinifera*	Vitaceae (Grape)	Dicotyledoneae (Eudicot)	Angiospermae (Angiosperm)	Spermatophyta (Seed plant)
vvi:100250503		*Vitis vinifera*	Vitaceae (Grape)	Dicotyledoneae (Eudicot)	Angiospermae (Angiosperm)	Spermatophyta (Seed plant)
vvi:100253234		*Vitis vinifera*	Vitaceae (Grape)	Dicotyledoneae (Eudicot)	Angiospermae (Angiosperm)	Spermatophyta (Seed plant)
vvi:100256460		*Vitis vinifera*	Vitaceae (Grape)	Dicotyledoneae (Eudicot)	Angiospermae (Angiosperm)	Spermatophyta (Seed plant)
vvi:100258578		*Vitis vinifera*	Vitaceae (Grape)	Dicotyledoneae (Eudicot)	Angiospermae (Angiosperm)	Spermatophyta (Seed plant)
vvi:100259491		*Vitis vinifera*	Vitaceae (Grape)	Dicotyledoneae (Eudicot)	Angiospermae (Angiosperm)	Spermatophyta (Seed plant)
vvi:00263725		*Vitis vinifera*	Vitaceae (Grape)	Dicotyledoneae (Eudicot)	Angiospermae (Angiosperm)	Spermatophyta (Seed plant)
vvi:100268124		*Vitis vinifera*	Vitaceae (Grape)	Dicotyledoneae (Eudicot)	Angiospermae (Angiosperm)	Spermatophyta (Seed plant)

Table 2 (continuation)

KEGG Entry name	Gene name	Species	Fmily	class	Subdivision	Division
BR000827*	OrysaPIN1a	*Oryza sativa*	Poaceae (Grass)	Monocotyledoneae (Monocot)	Angiospermae (Angiosperm)	Spermatophyta (Seed plant)
BR000828*	OrysaPIN1b	*Oryza sativa*	Poaceae (Grass)	Monocotyledoneae (Monocot)	Angiospermae (Angiosperm)	Spermatophyta (Seed plant)
BR000829*	OrysaPIN1c	*Oryza sativa*	Poaceae (Grass)	Monocotyledoneae (Monocot)	Angiospermae (Angiosperm)	Spermatophyta (Seed plant)
BR000830*	OrysaPIN1d	*Oryza sativa*	Poaceae (Grass)	Monocotyledoneae (Monocot)	Angiospermae (Angiosperm)	Spermatophyta (Seed plant)
BR000831*	OrysaPIN2	*Oryza sativa*	Poaceae (Grass)	Monocotyledoneae (Monocot)	Angiospermae (Angiosperm)	Spermatophyta (Seed plant)
BR000832*	OrysaPIN3a	*Oryza sativa*	Poaceae (Grass)	Monocotyledoneae (Monocot)	Angiospermae (Angiosperm)	Spermatophyta (Seed plant)
BR000833*	OrysaPIN3b	*Oryza sativa*	Poaceae (Grass)	Monocotyledoneae (Monocot)	Angiospermae (Angiosperm)	Spermatophyta (Seed plant)
BR000834*	OrysaPIN5a	*Oryza sativa*	Poaceae (Grass)	Monocotyledoneae (Monocot)	Angiospermae (Angiosperm)	Spermatophyta (Seed plant)
BR000835*	OrysaPIN5b	*Oryza sativa*	Poaceae (Grass)	Monocotyledoneae (Monocot)	Angiospermae (Angiosperm)	Spermatophyta (Seed plant)
BR000836*	OrysaPIN5c	*Oryza sativa*	Poaceae (Grass)	Monocotyledoneae (Monocot)	Angiospermae (Angiosperm)	Spermatophyta (Seed plant)
BR000837*	OrysaPIN8	*Oryza sativa*	Poaceae (Grass)	Monocotyledoneae (Monocot)	Angiospermae (Angiosperm)	Spermatophyta (Seed plant)
BR000838*	OrysaPIN9	*Oryza sativa*	Poaceae (Grass)	Monocotyledoneae (Monocot)	Angiospermae (Angiosperm)	Spermatophyta (Seed plant)
SORBI 02g029210	SorbiPIN1	*Sorghum bicolor*	Poaceae (Grass)	Monocotyledoneae (Monocot)	Angiospermae (Angiosperm)	Spermatophyta (Seed plant)
SORBI 03g029320	SorbiPIN2	*Sorghum bicolor*	Poaceae (Grass)	Monocotyledoneae (Monocot)	Angiospermae (Angiosperm)	Spermatophyta (Seed plant)
SORBI 03g032850	SorbiPIN3	*Sorghum bicolor*	Poaceae (Grass)	Monocotyledoneae (Monocot)	Angiospermae (Angiosperm)	Spermatophyta (Seed plant)
SORBI 03g037350	SorbiPIN4	*Sorghum bicolor*	Poaceae (Grass)	Monocotyledoneae (Monocot)	Angiospermae (Angiosperm)	Spermatophyta (Seed plant)
SORBI 05g002150	SorbiPIN7	*Sorghum bicolor*	Poaceae (Grass)	Monocotyledoneae (Monocot)	Angiospermae (Angiosperm)	Spermatophyta (Seed plant)
SORBI 07g026370	SorbiPIN8	*Sorghum bicolor*	Poaceae (Grass)	Monocotyledoneae (Monocot)	Angiospermae (Angiosperm)	Spermatophyta (Seed plant)
SORBI 10g004430	SorbiPIN9	*Sorghum bicolor*	Poaceae (Grass)	Monocotyledoneae (Monocot)	Angiospermae (Angiosperm)	Spermatophyta (Seed plant)
SORBI 10g008290	SorbiPIN10	*Sorghum bicolor*	Poaceae (Grass)	Monocotyledoneae (Monocot)	Angiospermae (Angiosperm)	Spermatophyta (Seed plant)
SORBI 10g026300	SorbiPIN11	*Sorghum bicolor*	Poaceae (Grass)	Monocotyledoneae (Monocot)	Angiospermae (Angiosperm)	Spermatophyta (Seed plant)
zma:100191787		*Zea mays*	Poaceae (Grass)	Monocotyledoneae (Monocot)	Angiospermae (Angiosperm)	Spermatophyta (Seed plant)
zma:100273056		*Zea mays*	Poaceae (Grass)	Monocotyledoneae (Monocot)	Angiospermae (Angiosperm)	Spermatophyta (Seed plant)
zma:100281763		*Zea mays*	Poaceae (Grass)	Monocotyledoneae (Monocot)	Angiospermae (Angiosperm)	Spermatophyta (Seed plant)
zma:100285745		*Zea mays*	Poaceae (Grass)	Monocotyledoneae (Monocot)	Angiospermae (Angiosperm)	Spermatophyta (Seed plant)
zma:100381964		*Zea mays*	Poaceae (Grass)	Monocotyledoneae (Monocot)	Angiospermae (Angiosperm)	Spermatophyta (Seed plant)
zma:100383548		*Zea mays*	Poaceae (Grass)	Monocotyledoneae (Monocot)	Angiospermae (Angiosperm)	Spermatophyta (Seed plant)
AB748932**	PinthPIN	*Pinus thunbergii*	Pinaceae (Pine)	Coniferopsida	Gimnospermae (Gymnosperm)	Spermatophyta (Seed plant)
AB748931**	GinbiPIN	*Ginkgo biloba*	Ginkgoaceae (Ginkgo)	Cycadopsida	Gimnospermae (Gymnosperm)	Spermatophyta (Seed plant)
SELMODRAFT 102666	SelmoPIN2-1	*Selaginella moellendorffii*	Selaginellaceae (Spike moss)	Lycopodiopsida (Club moss)		Lycopodiophyta
SELMODRAFT 105586	SelmoPIN3-2	*Selaginella moellendorffii*	Selaginellaceae (Spike moss)	Lycopodiopsida (Club moss)		Lycopodiophyta
SELMODRAFT 131936	SelmoPIN1-2	*Selaginella moellendorffii*	Selaginellaceae (Spike moss)	Lycopodiopsida (Club moss)		Lycopodiophyta
SELMODRAFT 234325	SelmoPIN1-1	*Selaginella moellendorffii*	Selaginellaceae (Spike moss)	Lycopodiopsida (Club moss)		Lycopodiophyta
SELMODRAFT 422990	SelmoPIN4-1	*Selaginella moellendorffii*	Selaginellaceae (Spike moss)	Lycopodiopsida (Club moss)		Lycopodiophyta
SELMODRAFT 88887	SelmoPIN5-1	*Selaginella moellendorffii*	Selaginellaceae (Spike moss)	Lycopodiopsida (Club moss)		Lycopodiophyta
SELMODRAFT 98910	SelmoPIN3-1	*Selaginella moellendorffii*	Selaginellaceae (Spike moss)	Lycopodiopsida (Club moss)		Lycopodiophyta

Table 2 (continuation)

KEGG Entry name	Gene name	Species	Fmily	Class	Subdivision	Division
PHYPADRAFT 117036	PhypaPINB	*Physcomitrella patens* subsp. *patens*	Funariaceae	Bryopsida (moss)		Bryophyta
PHYPADRAFT 202504	PhypaPINA	*Physcomitrella patens* subsp. *patens*	Funariaceae	Bryopsida (moss)		Bryophyta
PHYPADRAFT 71313	PhypaPINC	*Physcomitrella patens* subsp. *patens*	Funariaceae	Bryopsida (moss)		Bryophyta
AB748930**	PhypaPpPIND	*Physcomitrella patens* subsp. *patens*	Funariaceae	Bryopsida (moss)		Bryophyta
***	MarpoPIN1	*Marchantia polymorpha L.*	Marchantiaceae	Hepaticopsida (liverwort)		Bryophyta (Marchantiophyta)
***	MarpoPIN2	*Marchantia polymorpha L.*	Marchantiaceae	Hepaticopsida (liverwort)		Bryophyta (Marchantiophyta)
***	MarpoPIN3	*Marchantia polymorpha L.*	Marchantiaceae	Hepaticopsida (liverwort)		Bryophyta (Marchantiophyta)
***	MarpoPIN4	*Marchantia polymorpha L.*	Marchantiaceae	Hepaticopsida (liverwort)		Bryophyta (Marchantiophyta)
AB748929**	CpsIcPIN	*Closterium peracerosum-strigosum-littorale complex*	Closteriaceae	Charophyceae		Charophyte(Green alga)

* is GenBank Accession number (Miyashita et al. 2010).

** is DDBJ Accession number (isolated from this study).

*** is provided by Dr. Sandra Floyd and Professor John Bowman.

Table 3. Effects of disruption of *PpPIND* gene between the *PpPIND* disruptants.

The number of samples		looseness	moderateness + tightness
	WT	35	93
disruptants	PpPIND-dis-5-33	13	10
	PpPIND-dis-5-135	38	31
	PpPIND-dis-5-162	19	11

Effects of disruption of *PpPIND* gene was equivalence between the *PpPIND* disruptants (2 x 2 Pearson's chi-square test with Yates's continuity correction, $\chi 2 = 0.0147$, p-value = 0.9036 between PpPIND-dis-5-33 and PpPIND-dis-5-135, $\chi 2 = 0.048$, p-value = 0.8265 between PpPIND-dis-5-33 and PpPIND-dis-5-162, and $\chi 2 = 0.2949$, p-value = 0.5871 between PpPIND-dis-5-135 and PpPIND-dis-5-162).

Table 4. Effects of disruption of *PpPIND* gene on junction between vaginula and sporophyte foot in fully matured phase using a binomial GLM.

	Estimate	Std. Error	z value	Pr(>|z|)
(Intercept)	-0.97725	0.19830	-4.928	8.3e-07 ***
disruption	1.23962	0.46502	2.666	0.00768 **
PpPIND-dis-5-135	-0.05877	0.48528	-0.121	0.90362
PpPIND-dis-5-162	0.28418	0.56610	0.502	0.61567
PpPIND-dis-5-33	NA	NA	NA	NA

Asterisk indicates a significant (*** $p < 0.001$, and ** $p < 0.01$). NA indicates not analyzed.

Table 5-1. Primers used in this study.

primer name	sequence (5' to 3')
T7	TAATACGACTCACTATAGGG
SP6	ATTTAGGTGACACTATAGAA
3'AP	GGCCACGCGTCGACTAGTACTTTTTTTTTTTTTTTT
5'AP	CUACUACUACUAGGCCACGCGTCGACTAGTACGGGIIGGGIIGGGIIG
UAP	CUACUACUACUAGGCCACGCGTCGACTAGT
CpPIN-F2	TCATCAGCGTCGATCATTAGCGTCGAGC
CpPIN-F3	AGCGTCGAGCATTAGTGTTACACTG
CpPIN-R1	ACCAAAGAACGTCCTTAACGGAATAGCCTCC
CpPIN-R2	ACGGAATAGCCTCCTAGTTGACTG
PIN-Nf1	TN CCN YTN TAY GTN GC
PIN-Nf2	CAUCAUCAUCAU AAY RAY CCN TAY RMN ATG AA
PIN-Cr1	TC NYK NGC RAA NAC RAA
PIN-Cr2	CAUCAUCAUCAU TG NGG NAR NGC NGC YTG
PtPIN-F1	TGCCAAGCTATCAATAGACCAGAGG
PtPIN-F2	TCCTGATTTGTATGCCGCGACTGTC
GbPIN-F1	ACTGTAGGCTAGAGTATCCCAATATCATAG
GbPIN-F2	TAGAGTATCCCAATATCATAGATTAGCTGC
PpPIN-f2	GGGGTTCTTCGAGCTCGACTTAAATTC
PpPIN-r2	GAAGTCCTTGCAAGTGATTTATGGTTA
PpPIN-5'f1	GGTCGAGAGAGATGAAGTATGTAGGGTTTTACC
PpPIN-3'r1	AGTGGAGAAGCAGAGACTCAGCGGACC
PpPIN-F300	TGGGTGATTACCTTATTCCAGCTCTC
PpPIN-R395	ATGAGCGAGTAGACGATTCCCATAAC
PpGapC-F850	CGATGCCATTAAGACGGCTATCAA
PpGapC-R59	CATGAACGCTGGCGATGGAAATTA
PpPIN-KO5'f	GGGGTACCGACAGGCAGAGGGGAGATCTG
PpPIN-KO5'r	CCATCGATTGTCAACATGACGATAACGTTGATTGTCC
PpPIN-KO3'f	CACTAGTCTGAGAAATACAAATTCCAGACAACGAATAC
PpPIN-KO3'r	TCCCCGCGGGGTACCATTGACCTGTTTCTGAGATG

Table 5-2. *PpPIND* specific TaqMan probe and primers used in real time PCR.

primer or probe name	sequence (5' to 3')
PpPINDcDNA-303F	ACAAACGTCTGCGCTGAACA
PpPINDcDNA-375R	CCCACCGGGCAAATTTC
PpPINDcDNA-328T (Taqmqn probe)	TGGAAGCTGGCAGAATCGTTGAGGAG
PpTUA1cDNA-1002F	ACATGGCGTGTTGCCTGAT
PpTUA1cDNA-1088R	CCGTGCGCTTCGTCTTG
PpTUA1cDNA-1043T (Taqmqn probe)	AAGGACGTGAACGCCGCGGT

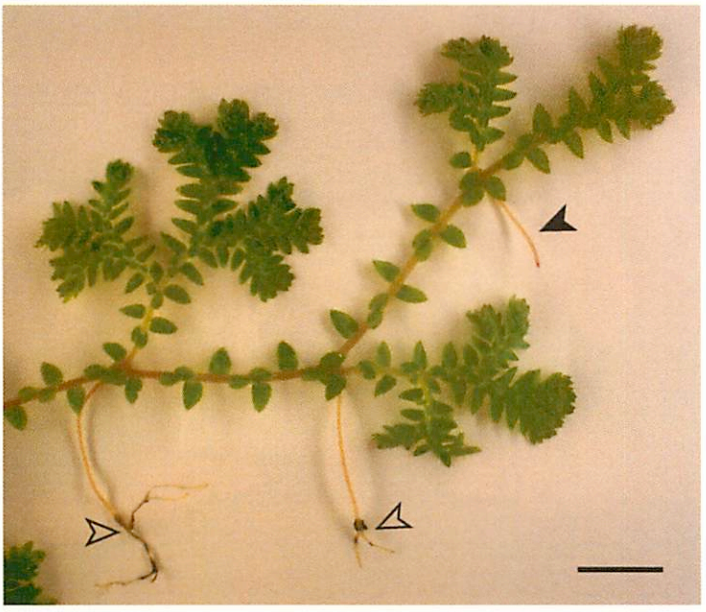
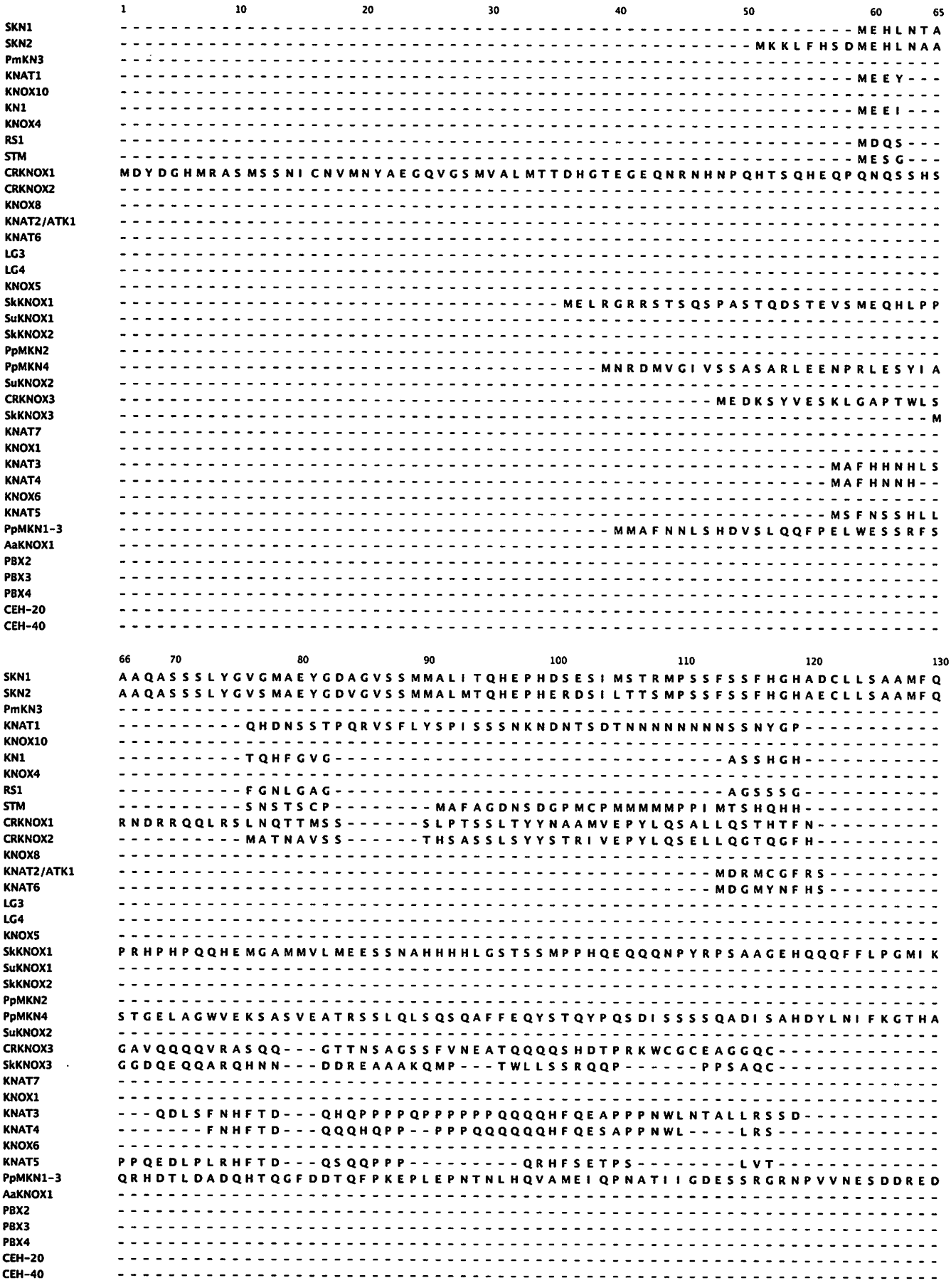
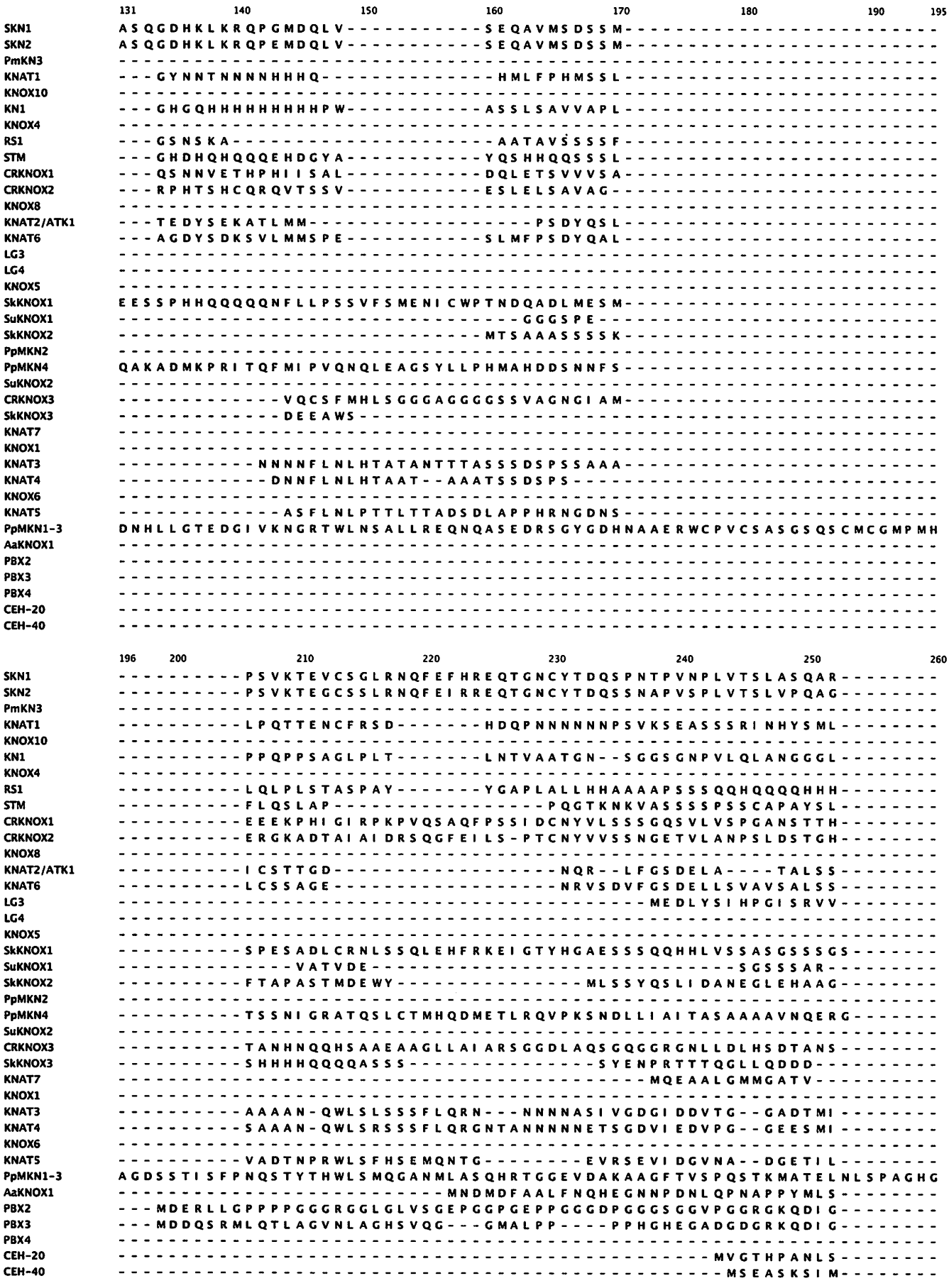


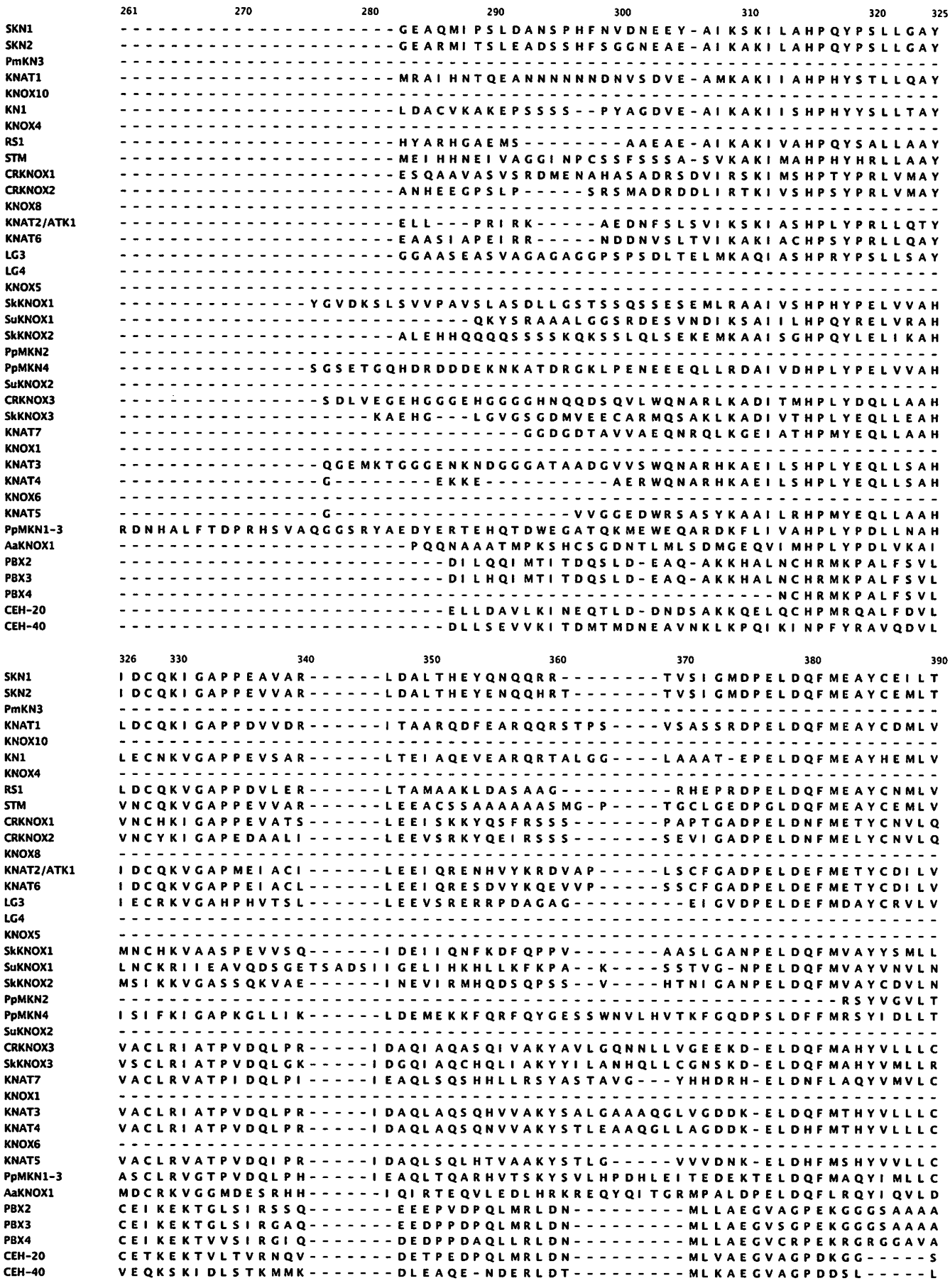
Fig. 1. A photograph of *Selaginella uncinata*.

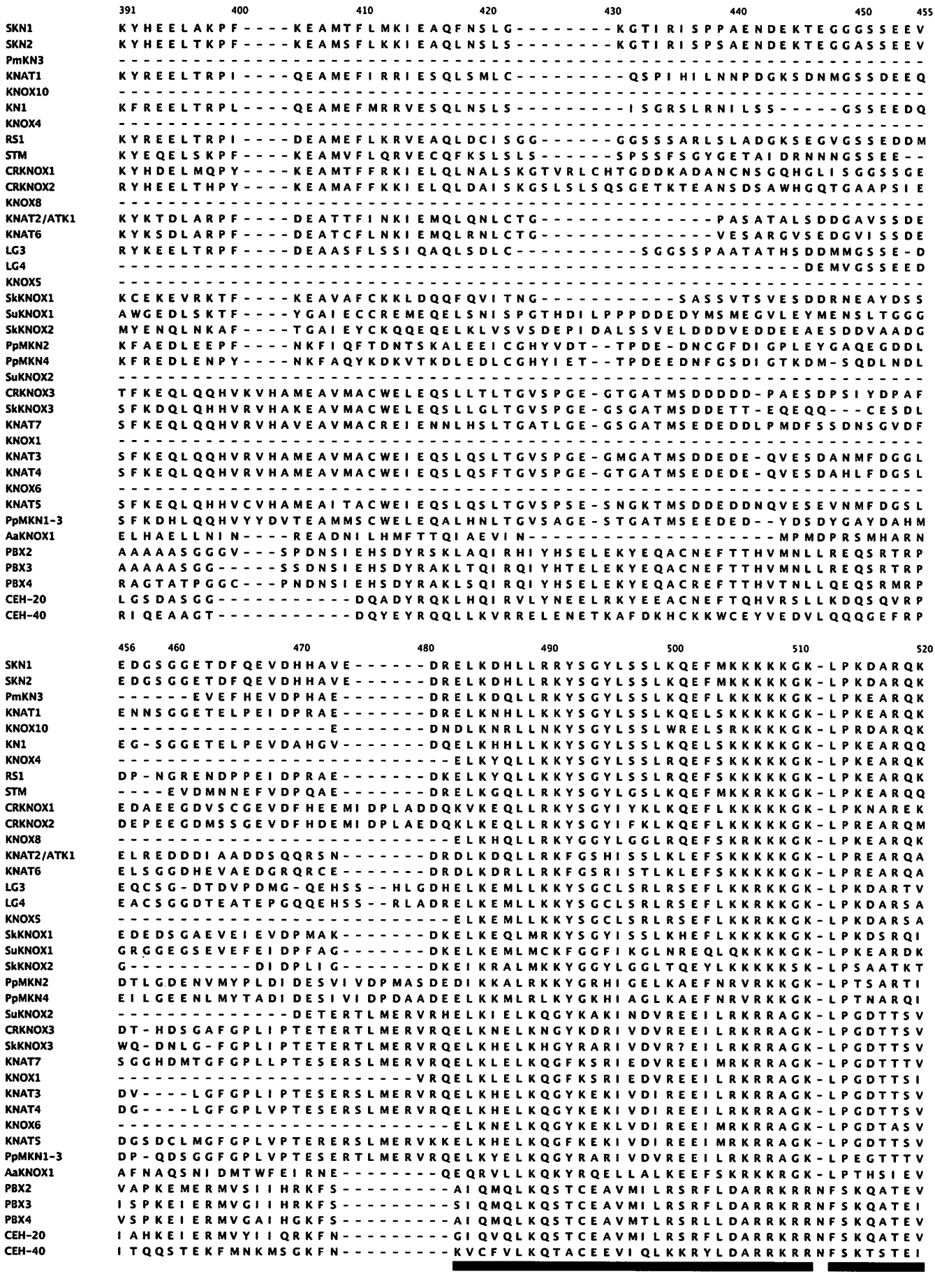
The spikemoss *Selaginella uncinata* has creeping stems with leafy microphylls and rhizophores (closed arrowhead), which produce roots (opened arrowhead). Scale bar = 10 mm.

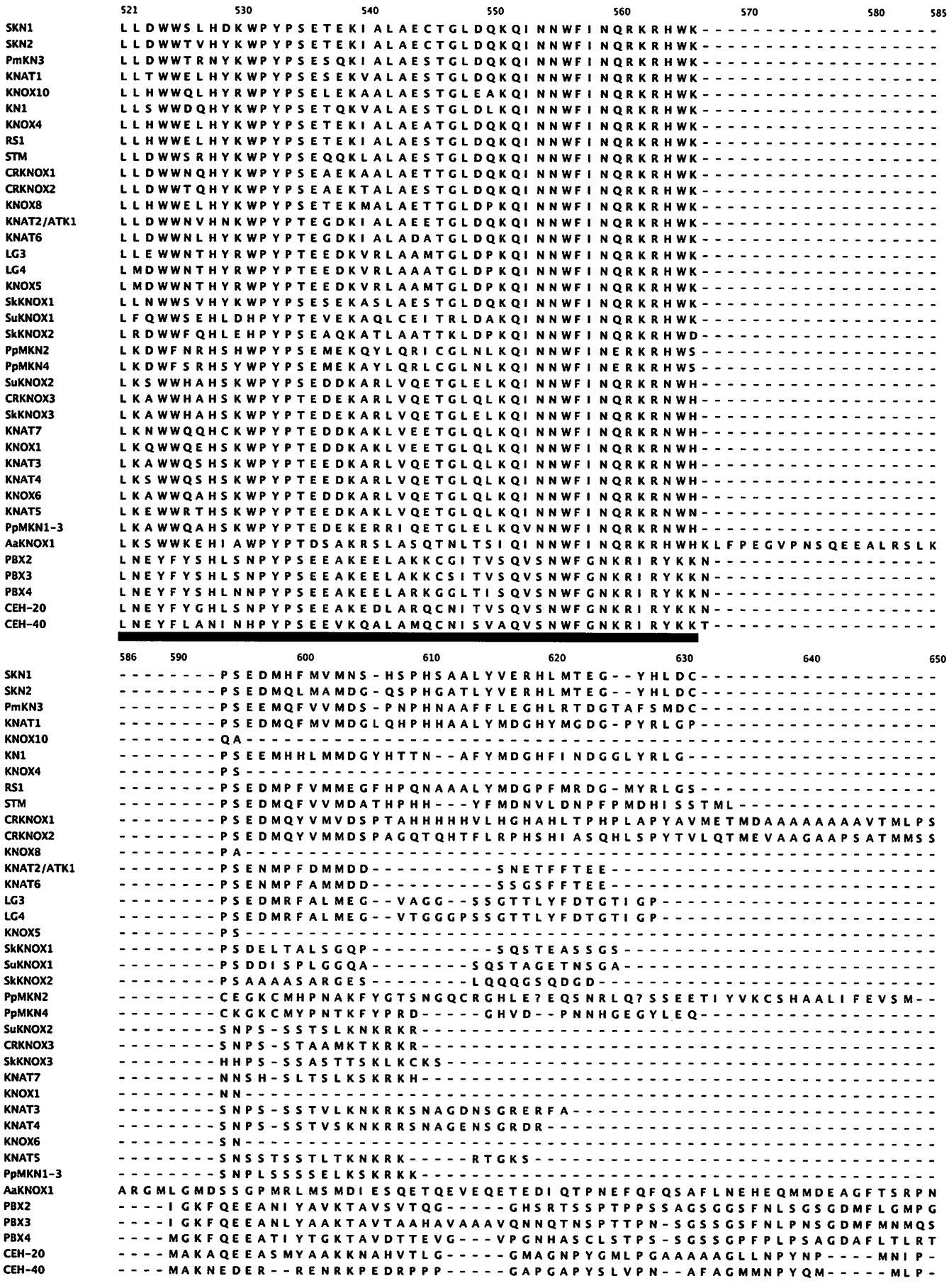
Figure 2


1 10 20 30 40 50 60 65
SKN1 MEHLNTA
SKN2 MKKLFHSDMEHLNAA
PmKN3
KNAT1 MEEY
KNOX10
KN1 MEEI
KNOX4
RS1 MDQS
STM MESG
CRKNOX1 MDYDGHMRASMSSNICNVMNYAEGQVGSMVALMTTDHGTEGEQNRNHNPQHTSQHEQPQNQSSHS
CRKNOX2
KNOX8
KNAT2/ATK1
KNAT6
LG3
LG4
KNOX5
SkKNOX1 MELRGRRSTSQSPASTQDSTEVSMEQHLPP
SuKNOX1
SkKNOX2
PpMKN2
PpMKN4 MNRDMVGIVSSASARLEENPRLESYIA
SuKNOX2
CRKNOX3 MEDKSYVESKLGAPTWLS
SkKNOX3 M
KNAT7
KNOX1
KNAT3 MAFHHNHLS
KNAT4 MAFHNNH
KNOX6
KNAT5 MSFNSSHLL
PpMKN1-3 MMAFNNLSHDVSLQQFPELWESSRFS
AaKNOX1
PBX2
PBX3
PBX4
CEH-20
CEH-40
66 70 80 90 100 110 120 130
SKN1 AAQASSSLYGVGMAEYGDAGVSSMMALITQHEPHDSESIMSTRMPSSFSSFHGHADCLLSAAMFQ
SKN2 AAQASSSLYGVSMAEYGDVGVSSMMALMTQHEPHERDSILTTSMPSSFSSFHGHAECLLSAAMFQ
PmKN3
KNAT1 QHDNSSTPQRVSFLYSPISSSNKNDNTSDTNNNNNNNNSSNYGP
KNOX10
KN1 TQHFGVG ASSHGH
KNOX4
RS1 FGNLGAG AGSSSG
STM SNSTSCP MAFAGDNSDGPMCPMMMMMPPIMTSHQHH
CRKNOX1 RNDRRQQLRSLNQTTMSS SLPTSSLTYYNAAMVEPYLQSALLQSTHTFN
CRKNOX2 MATNAVSS THSASSLSYYSTRIVEPYLQSELLQGTQGFH
KNOX8
KNAT2/ATK1 MDRMCGFRS
KNAT6 MDGMYNFHS
LG3
LG4
KNOX5
SkKNOX1 PRHPHPQQHEMGAMMVLMEESSNAHHHHLGSTSSMPPHQEQQQNPYRPSAAGEHQQQFFLPGMIK
SuKNOX1
SkKNOX2
PpMKN2
PpMKN4 STGELAGWVEKSASVEATRSSLQLSQSQAFFEQYSTQYPQSDISSSSQADISAHDYLNIFKGTHA
SuKNOX2
CRKNOX3 GAVQQQQVRASQQ GTTNSAGSSFVNEATQQQQSHDTPRKWCGCEAGGQC
SkKNOX3 GGDQEQQARQHNN DDREAAAKQMP TWLLSSRQQP PPSAQC
KNAT7
KNOX1
KNAT3 QDLSFNHFTD QHQPPPPQPPPPPPQQQQHFQEAPPPNWLNTALLRSSD
KNAT4 FNHFTD QQQHQPP PPPQQQQQQHFQESAPPNWL LRS
KNOX6
KNAT5 PPQEDLPLRHFTD QSQQPPP QRHFSETPS LVT
PpMKN1-3 QRHDTLDADQHTQGFDDTQFPKEPLEPNTNLHQVAMEIQPNATIIGDESSRGRNPVVNESDDRED
AaKNOX1
PBX2
PBX3
PBX4
CEH-20
CEH-40


131 140 150 160 170 180 190 195
SKN1 ASQGDHKLKRQPGMDQLV----------SEQAVMSDSSM----------
SKN2 ASQGDHKLKRQPEMDQLV----------SEQAVMSDSSM----------
PmKN3 ----------
KNAT1 ---GYNNTNNNNHHHQ----------HMLFPHMSSL----------
KNOX10 ----------
KN1 ---GHGQHHHHHHHHHPW----------ASSLSAVVAPL----------
KNOX4 ----------
RS1 ---GSNSKA----------AATAVSSSSF----------
STM ---GHDHQHQQQEHDGYA----------YQSHHQQSSSL----------
CRKNOX1 ---QSNNVETHPHIISAL----------DQLETSVVVSA----------
CRKNOX2 ---RPHTSHCQRQVTSSV----------ESLELSAVAG----------
KNOX8 ----------
KNAT2/ATK1 ---TEDYSEKATLMM----------PSDYQSL----------
KNAT6 ---AGDYSDKSVLMMSPE----------SLMFPSDYQAL----------
LG3 ----------
LG4 ----------
KNOX5 ----------
SkKNOX1 EESSPHHQQQQQNFLLPSSVFSMENICWPTNDQADLMESM----------
SuKNOX1 ----------GGGSPE----------
SkKNOX2 ----------MTSAAASSSSK----------
PpMKN2 ----------
PpMKN4 QAKADMKPRITQFMIPVQNQLEAGSYLLPHMAHDDSNNFS----------
SuKNOX2 ----------
CRKNOX3 ----------VQCSFMHLSGGGAGGGGSSVAGNGIAM----------
SkKNOX3 ----------DEEAWS----------
KNAT7 ----------
KNOX1 ----------
KNAT3 ----------NNNNFLNLHTATANTTTASSSDSPSSAAA----------
KNAT4 ----------DNNFLNLHTAAT--AAATSSDSPS----------
KNOX6 ----------
KNAT5 ----------ASFLNLPTTLTTADSDLAPPHRNGDNS----------
PpMKN1-3 DNHLLGTEDGIVKNGRTWLNSALLREQNQASEDRSGYGDHNAAERWCPVCSASGSQSCMCGMPMH
AaKNOX1 ----------
PBX2 ----------
PBX3 ----------
PBX4 ----------
CEH-20 ----------
CEH-40 ----------
196 200 210 220 230 240 250 260
SKN1 ----------PSVKTEVCSGLRNQFEFHREQTGNCYTDQSPNTPVNPLVTSLASQAR---------
SKN2 ----------PSVKTEGCSSLRNQFEIRREQTGNCYTDQSSNAPVSPLVTSLVPQAG---------
PmKN3 ----------
KNAT1 ----------LPQTTENCFRSD--------HDQPNNNNNNPSVKSEASSSRINHYSML---------
KNOX10 ----------
KN1 ----------PPQPPSAGLPLT--------LNTVAATGN--SGGSGNPVLQLANGGGL---------
KNOX4 ----------
RS1 ----------LQLPLSTASPAY--------YGAPLALLHHAAAAPSSSQQHQQQQHHH---------
STM ----------FLQSLAP----------PQGTKNKVASSSSPSSCAPAYSL---------
CRKNOX1 ----------EEEKPHIGIRPKPVQSAQFPSSIDCNYVLSSSGQSVLVSPGANSTTH---------
CRKNOX2 ----------ERGKADTAIAIDRSQGFEILS-PTCNYVVSSNGETVLANPSLDSTGH---------
KNOX8 ----------
KNAT2/ATK1 ----------ICSTTGD----------NQR--LFGSDELA-----TALSS---------
KNAT6 ----------LCSSAGE----------NRVSDVFGSDELLSVAVSALSS---------
LG3 ----------MEDLYSIHPGISRVV---------
LG4 ----------
KNOX5 ----------
SkKNOX1 ----------SPESADLCRNLSSQLEHFRKEIGTYHGAESSSQQHHLVSSASGSSSGS-------
SuKNOX1 ----------VATVDE----------SGSSSAR---------
SkKNOX2 ----------FTAPASTMDEWY----------MLSSYQSLIDANEGLEHAAG---------
PpMKN2 ----------
PpMKN4 ----------TSSNIGRATQSLCTMHQDMETLRQVPKSNDLLIAITASAAAAVNQERG-------
SuKNOX2 ----------
CRKNOX3 ----------TANHNQQHSAAEAAGLLAIARSGGDLAQSGQGGRGNLLDLHSDTANS---------
SkKNOX3 ----------SHHHHQQQQASSS----------SYENPRTTTQGLLQDDD----------
KNAT7 ----------MQEAALGMMGATV----------
KNOX1 ----------
KNAT3 ----------AAAAN-QWLSLSSSFLQRN---NNNNASIVGDGIDDVTG--GADTMI---------
KNAT4 ----------SAAAN-QWLSRSSSFLQRGNTANNNNNETSGDVIEDVPG--GEESMI---------
KNOX6 ----------
KNAT5 ----------VADTNPRWLSFHSEMQNTG--------EVRSEVIDGVNA--DGETIL---------
PpMKN1-3 AGDSSTISFPNQSTYTHWLSMQGANMLASQHRTGGEVDAKAAGFTVSPQSTKMATELNLSPAGHG
AaKNOX1 ----------MNDMDFAALFNQHEGNNPDNLQPNAPPYMLS---------
PBX2 ---MDERLLGPPPPGGGRGGLGLVSGEPGGPGEPPGGGDPGGGSGGVPGGRGKQDIG---------
PBX3 ---MDDQSRMLQTLAGVNLAGHSVQG---GMALPP----PPHGHEGADGDGRKQDIG---------
PBX4 ----------
CEH-20 ----------MVGTHPANLS---------
CEH-40 ----------MSEASKSIM---------


261 270 280 290 300 310 320 325
SKN1 - G E A Q M I P S L D A N S P H F N V D N E E Y - A I K S K I L A H P Q Y P S L L G A Y
SKN2 - G E A R M I T S L E A D S S H F S G G N E A E - A I K A K I L A H P Q Y P S L L G A Y
PmKN3 -
KNAT1 - M R A I H N T Q E A N N N N N N D N V S D V E - A M K A K I I A H P H Y S T L L Q A Y
KNOX10 -
KN1 - L D A C V K A K E P S S S S - - P Y A G D V E - A I K A K I I S H P H Y Y S L L T A Y
KNOX4 -
RS1 - H Y A R H G A E M S - - - - - - - - A A E A E - A I K A K I V A H P Q Y S A L L A A Y
STM - M E I H H N E I V A G G I N P C S S F S S S A - S V K A K I M A H P H Y H R L L A A Y
CRKNOX1 - E S Q A A V A S V S R D M E N A H A S A D R S D V I R S K I M S H P T Y P R L V M A Y
CRKNOX2 - A N H E E G P S L P - - - - - S R S M A D R D D L I R T K I V S H P S Y P R L V M A Y
KNOX8 -
KNAT2/ATK1 - E L L - - - P R I R K - - - - - A E D N F S L S V I K S K I A S H P L Y P R L L Q T Y
KNAT6 - E A A S I A P E I R R - - - - - N D D N V S L T V I K A K I A C H P S Y P R L L Q A Y
LG3 - G G A A S E A S V A G A G A G G P S P S D L T E L M K A Q I A S H P R Y P S L L S A Y
LG4 -
KNOX5 -
SkKNOX1 - - - - - - - - - - - - - - - - Y G V D K S L S V V P A V S L A S D L L G S T S S Q S S E S E M L R A A I V S H P H Y P E L V V A H
SuKNOX1 - Q K Y S R A A A L G G S R D E S V N D I K S A I I L H P Q Y R E L V R A H
SkKNOX2 - A L E H H Q Q Q Q S S S S K Q K S S L Q L S E K E M K A A I S G H P Q Y L E L I K A H
PpMKN2 -
PpMKN4 - S G S E T G Q H D R D D D E K N K A T D R G K L P E N E E E Q L L R D A I V D H P L Y P E L V V A H
SuKNOX2 -
CRKNOX3 - S D L V E G E H G G G E H G G G G H N Q Q D S Q V L W Q N A R L K A D I T M H P L Y D Q L L A A H
SkKNOX3 - K A E H G - - - L G V G S G D M V E E C A R M Q S A K L K A D I V T H P L Y E Q L L E A H
KNAT7 - G G D G D T A V V A E Q N R Q L K G E I A T H P M Y E Q L L A A H
KNOX1 -
KNAT3 - Q G E M K T G G G E N K N D G G G A T A A D G V V S W Q N A R H K A E I L S H P L Y E Q L L S A H
KNAT4 - G - - - - - - - - E K K E - - - - - - - - - - A E R W Q N A R H K A E I L S H P L Y E Q L L S A H
KNOX6 -
KNAT5 - G - V V G G E D W R S A S Y K A A I L R H P M Y E Q L L A A H
PpMKN1-3 R D N H A L F T D P R H S V A Q G G S R Y A E D Y E R T E H Q T D W E G A T Q K M E W E Q A R D K F L I V A H P L Y P D L L N A H
AaKNOX1 - P Q Q N A A A T M P K S H C S G D N T L M L S D M G E Q V I M H P L Y P D L V K A I
PBX2 - D I L Q Q I M T I T D Q S L D - E A Q - A K K H A L N C H R M K P A L F S V L
PBX3 - D I L H Q I M T I T D Q S L D - E A Q - A K K H A L N C H R M K P A L F S V L
PBX4 - N C H R M K P A L F S V L
CEH-20 - E L L D A V L K I N E Q T L D - D N D S A K K Q E L Q C H P M R Q A L F D V L
CEH-40 - D L L S E V V K I T D M T M D N E A V N K L K P Q I K I N P F Y R A V Q D V L
326 330 340 350 360 370 380 390
SKN1 I D C Q K I G A P P E A V A R - - - - - - L D A L T H E Y Q N Q Q R R - - - - - - - - T V S I G M D P E L D Q F M E A Y C E I L T
SKN2 I D C Q K I G A P P E V V A R - - - - - - L D A L T H E Y E N Q Q H R T - - - - - - - T V S I G M D P E L D Q F M E A Y C E M L T
PmKN3 -
KNAT1 L D C Q K I G A P P D V V D R - - - - - - I T A A R Q D F E A R Q Q R S T P S - - - - V S A S S R D P E L D Q F M E A Y C D M L V
KNOX10 -
KN1 L E C N K V G A P P E V S A R - - - - - - L T E I A Q E V E A R Q R T A L G G - - - - L A A A T - E P E L D Q F M E A Y H E M L V
KNOX4 -
RS1 L D C Q K V G A P P D V L E R - - - - - - L T A M A A K L D A S A A G - - - - - - - - - R H E P R D P E L D Q F M E A Y C N M L V
STM V N C Q K V G A P P E V V A R - - - - - - L E E A C S S A A A A A A S M G - P - - - - T G C L G E D P G L D Q F M E A Y C E M L V
CRKNOX1 V N C H K I G A P P E V A T S - - - - - - L E E I S K K Y Q S F R S S S - - - - - - - P A P T G A D P E L D N F M E T Y C N V L Q
CRKNOX2 V N C Y K I G A P E D A A L I - - - - - - L E E V S R K Y Q E I R S S S - - - - - - - S E V I G A D P E L D N F M E L Y C N V L Q
KNOX8 -
KNAT2/ATK1 I D C Q K V G A P M E I A C I - - - - - - L E E I Q R E N H V Y K R D V A P - - - - - L S C F G A D P E L D E F M E T Y C D I L V
KNAT6 I D C Q K V G A P P E I A C L - - - - - - L E E I Q R E S D V Y K Q E V V P - - - - - S S C F G A D P E L D E F M E T Y C D I L V
LG3 I E C R K V G A H P H V T S L - - - - - - L E E V S R E R R P D A G A G - - - - - - - - - E I G V D P E L D E F M D A Y C R V L V
LG4 -
KNOX5 -
SkKNOX1 M N C H K V A A S P E V V S Q - - - - - - I D E I I Q N F K D F Q P P V - - - - - - - A A S L G A N P E L D Q F M V A Y Y S M L L
SuKNOX1 L N C K R I I E A V Q D S G E T S A D S I I G E L I H K H L L K F K P A - - K - - - - S S T V G - N P E L D Q F M V A Y V N V L N
SkKNOX2 M S I K K V G A S S Q K V A E - - - - - - I N E V I R M H Q D S Q P S S - - V - - - - H T N I G A N P E L D Q F M V A Y C D V L N
PpMKN2 - R S Y V G V L T
PpMKN4 I S I F K I G A P K G L L I K - - - - - - L D E M E K K F Q R F Q Y G E S S W N V L H V T K F G Q D P S L D F F M R S Y I D L L T
SuKNOX2 -
CRKNOX3 V A C L R I A T P V D Q L P R - - - - - I D A Q I A Q A S Q I V A K Y A V L G Q N N L L V G E E K D - E L D Q F M A H Y V L L L C
SkKNOX3 V S C L R I A T P V D Q L G K - - - - - I D G Q I A Q C H Q L I A K Y Y I L A N H Q L L C G N S K D - E L D Q F M A H Y V M L L R
KNAT7 V A C L R V A T P I D Q L P I - - - - - I E A Q L S Q S H H L L R S Y A S T A V G - - - Y H H D R H - E L D N F L A Q Y V M V L C
KNOX1 -
KNAT3 V A C L R I A T P V D Q L P R - - - - - I D A Q L A Q S Q H V V A K Y S A L G A A A Q G L V G D D K - E L D Q F M T H Y V L L L C
KNAT4 V A C L R I A T P V D Q L P R - - - - - I D A Q L A Q S Q N V V A K Y S T L E A A Q G L L A G D D K - E L D H F M T H Y V L L L C
KNOX6 -
KNAT5 V A C L R V A T P V D Q I P R - - - - - I D A Q L S Q L H T V A A K Y S T L G - - - - - V V V D N K - E L D H F M S H Y V V L L C
PpMKN1-3 A S C L R V G T P V D Q L P H - - - - - I E A Q L T Q A R H V T S K Y S V L H P D H L E I T E D E K T E L D Q F M A Q Y I M L L C
AaKNOX1 M D C R K V G G M D E S R H H - - - - - - I Q I R T E Q V L E D L H R K R E Q Y Q I T G R M P A L D P E L D Q F L R Q Y I Q V L D
PBX2 C E I K E K T G L S I R S S Q - - - - - - - E E E P V D P Q L M R L D N - - - - - - - - - M L L A E G V A G P E K G G G S A A A A
PBX3 C E I K E K T G L S I R G A Q - - - - - - - E E D P P D P Q L M R L D N - - - - - - - - - M L L A E G V S G P E K G G G S A A A A
PBX4 C E I K E K T V V S I R G I Q - - - - - - - D E D P P D A Q L L R L D N - - - - - - - - - M L L A E G V C R P E K R G R G G A V A
CEH-20 C E T K E K T V L T V R N Q V - - - - - - - D E T P E D P Q L M R L D N - - - - - - - - - M L V A E G V A G P D K G G - - - - - S
CEH-40 V E Q K S K I D L S T K M M K - - - - - - - D L E A Q E - N D E R L D T - - - - - - - - - M L K A E G V A G P D D S L - - - - - L


391 400 410 420 430 440 450 455
SKN1 KYHEELAKPF----KEAMTFLMKIEAQFNSLG--------KGTIRISPPAENDEKTEGGGSSEEV
SKN2 KYHEELTKPF----KEAMSFLKKIEAQLNSLS--------KGTIRISPSAENDEKTEGGASSEEV
PmKN3 ----------
KNAT1 KYREELTRPI----QEAMEFIRRIESQLSMLC--------QSPIHILNNPDGKSDNMGSSDEEQ
KNOX10 ----------
KN1 KFREELTRPL----QEAMEFMRRVESQLNSLS--------ISGRSLRNILSS-----GSSEEDQ
KNOX4 ----------
RS1 KYREELTRPI----DEAMEFLKRVEAQLDCISGG------GGSSSARLSLADGKSEGVGSSEDDM
STM KYEQELSKPF----KEAMVFLQRVECQFKSLSLS------SPSSFSGYGETAIDRNNNGSSEE--
CRKNOX1 KYHDELMQPY----KEAMTFFRKIELQLNALSKGTVRLCHTGDDKADANCNSGQHGLISGGSSGE
CRKNOX2 RYHEELTHPY----KEAMAFFKKIELQLDAISKGSLSLSQSGETKTEANSDSAWHGQTGAAPSIE
KNOX8 ----------
KNAT2/ATK1 KYKTDLARPF----DEATTFINKIEMQLQNLCTG----------PASATALSDDGAVSSDE
KNAT6 KYKSDLARPF----DEATCFLNKIEMQLRNLCTG----------VESARGVSEDGVISSDE
LG3 RYKEELTRPF----DEAASFLSSIQAQLSDLC--------SGGSSPAATATHSDDMMGSSE-D
LG4 ----------DEMVGSSEED
KNOX5 ----------
SkKNOX1 KCEKEVRKTF----KEAVAFCKKLDQQFQVITNG--------SASSVTSVESDDRNEAYDSS
SuKNOX1 AWGEDLSKTF----YGAIECCREMEQELSNISPGTHDILPPPDDEDYMSMEGVLEYMENSLTGGG
SkKNOX2 MYENQLNKAF----TGAIEYCKQQEQELKLVSVSDEPIDALSSVELDDDVEDDEEAESDDVAADG
PpMKN2 KFAEDLEEPF----NKFIQFTDNTSKALEEICGHYVDT--TPDE-DNCGFDIGPLEYGAQEGDDL
PpMKN4 KFREDLENPY----NKFAQYKDKVTKDLEDLCGHYIET--TPDEEDNFGSDIGTKDM-SQDLNDL
SuKNOX2 ----------
CRKNOX3 TFKEQLQQHVKVHAMEAVMACWELEQSLLTLTGVSPGE-GTGATMSDDDDD-PAESDPSIYDPAF
SkKNOX3 SFKDQLQHHVRVHAKEAVMACWELEQSLLGLTGVSPGE-GSGATMSDDETT-EQEQQ---CESDL
KNAT7 SFKEQLQQHVRVHAVEAVMACREIENNLHSLTGATLGE-GSGATMSEDEDDLPMDFSSDNSGVDF
KNOX1 ----------
KNAT3 SFKEQLQQHVRVHAMEAVMACWEIEQSLQSLTGVSPGE-GMGATMSDDEDE-QVESDANMFDGGL
KNAT4 SFKEQLQQHVRVHAMEAVMACWEIEQSLQSFTGVSPGE-GTGATMSEDEDE-QVESDAHLFDGSL
KNOX6 ----------
KNAT5 SFKEQLQHHVCVHAMEAITACWEIEQSLQSLTGVSPSE-SNGKTMSDDEDDNQVESEVNMFDGSL
PpMKN1-3 SFKDHLQQHVYYDVTEAMMSCWELEQALHNLTGVSAGE-STGATMSEEDED--YDSDYGAYDAHM
AaKNOX1 ELHAELLNIN----READNILHMFTTQIAEVIN----------MPMDPRSMHARN
PBX2 AAAAASGGGV---SPDNSIEHSDYRSKLAQIRHIYHSELEKYEQACNEFTTHVMNLLREQSRTRP
PBX3 AAAAASGG------SSDNSIEHSDYRAKLTQIRQIYHTELEKYEQACNEFTTHVMNLLREQSRTRP
PBX4 RAGTATPGGC---PNDNSIEHSDYRAKLSQIRQIYHSELEKYEQACREFTTHVTNLLQEQSRMRP
CEH-20 LGSDASGG----------DQADYRQKLHQIRVLYNEELRKYEEACNEFTQHVRSLLKDQSQVRP
CEH-40 RIQEAAGT----------DQYEYRQQLLKVRRELENETKAFDKHCKKWCEYVEDVLQQQGEFRP
456 460 470 480 490 500 510 520
SKN1 EDGSGGETDFQEVDHHAVE------DRELKDHLLRRYSGYLSSLKQEFMKKKKKGK-LPKDARQK
SKN2 EDGSGGETDFQEVDHHAVE------DRELKDHLLRKYSGYLSSLKQEFMKKKKKGK-LPKDARQK
PmKN3 -------EVEFHEVDPHAE-------DRELKDQLLRKYSGYLSSLKQEFLKKKKKGK-LPKEARQK
KNAT1 ENNSGGETELPEIDPRAE-------DRELKNHLLKKYSGYLSSLKQELSKKKKKGK-LPKEARQK
KNOX10 ----------E-------DNDLKNRLLNKYSGYLSSLWRELSRKKKKGK-LPRDARQK
KN1 EG-SGGETELPEVDAHGV-------DQELKHHLLKKYSGYLSSLKQELSKKKKKGK-LPKEARQQ
KNOX4 ----------ELKYQLLKKYSGYLSSLRQEFSKKKKKGK-LPKEARQK
RS1 DP-NGRENDPPEIDPRAE-------DKELKYQLLKKYSGYLSSLRQEFSKKKKKGK-LPKEARQK
STM ----EVDMNNEFVDPQAE-------DRELKGQLLRKYSGYLGSLKQEFMKKRKKGK-LPKEARQQ
CRKNOX1 EDAEEGDVSCGEVDFHEEMIDPLADDQKVKEQLLRKYSGYIYKLKQEFLKKKKKGK-LPKNAREK
CRKNOX2 DEPEEGDMSSGEVDFHDEMIDPLAEDQKLKEQLLRKYSGYIFKLKQEFLKKKKKGK-LPREARQM
KNOX8 ----------ELKHQLLRKYGGYLGGLRQEFSKRKKKGK-LPKEARQK
KNAT2/ATK1 ELREDDDIAADDSQQRSN-------DRDLKDQLLRKFGSHISSLKLEFSKKKKKGK-LPREARQA
KNAT6 ELSGGDHEVAEDGRQRCE-------DRDLKDRLLRKFGSRISTLKLEFSKKKKKGK-LPREARQA
LG3 EQCSG-DTDVPDMG-QEHSS--HLGDHELKEMLLKKYSGCLSRLRSEFLKKRKKGK-LPKDARTV
LG4 EACSGGDTEATEPGQQEHSS--RLADRELKEMLLKKYSGCLSRLRSEFLKKRKKGK-LPKDARSA
KNOX5 ----------ELKEMLLKKYSGCLSRLRSEFLKKRKKGK-LPKDARSA
SkKNOX1 EDEDSGAEVEIEVDPMAK-------DKELKEQLMRKYSGYISSLKHEFLKKKKKGK-LPKDSRQI
SuKNOX1 GRGGEGSEVEFEIDPFAG-------DKELKEMLMCKFGGFIKGLNREQLQKKKKGK-LPKEARDK
SkKNOX2 G----------DIDPLIG-------DKEIKRALMKKYGGYLGGLTQEYLKKKKKSK-LPSAATKT
PpMKN2 DTLGDENVMYPLDIDESVIVDPMASDEDIKKALRKKYGRHIGELKAEFNRVRKKGK-LPTSARTI
PpMKN4 EILGEENLMYTADIDESIVIDPDAADEELKKMLRLKYGKHIAGLKAEFNRVRKKGK-LPTNARQI
SuKNOX2 ----------DETERTLMERVRHELKIELKQGYKAKINDVREEILRKRRAGK-LPGDTTSV
CRKNOX3 DT-HDSGAFGPLIPTETERTLMERVRQELKNELKNGYKDRIVDVREEILRKRRAGK-LPGDTTSV
SkKNOX3 WQ-DNLG-FGPLIPTETERTLMERVRQELKHELKHGYRARIVDVR?EILRKRRAGK-LPGDTTSV
KNAT7 SGGHDMTGFGPLLPTESERSLMERVRQELKLELKQGFKSRIEDVREEIMRKRRAGK-LPGDTTTV
KNOX1 ----------VRQELKLELKQGFKSRIEDVREEILRKRRAGK-LPGDTTSI
KNAT3 DV----LGFGPLIPTESERSLMERVRQELKHELKQGYKEKIVDIREEILRKRRAGK-LPGDTTSV
KNAT4 DG----LGFGPLVPTESERSLMERVRQELKHELKQGYKEKIVDIREEILRKRRAGK-LPGDTTSV
KNOX6 ----------ELKNELKQGYKEKLVDIREEIMRKRRAGK-LPGDTASV
KNAT5 DGSDCLMGFGPLVPTERERSLMERVKKELKHELKQGFKEKIVDIREEIMRKRRAGK-LPGDTTSV
PpMKN1-3 DP-QDSGGFGPLVPTESERTLMERVRQELKYELKQGYRARIVDVREEILRKRRAGK-LPEGTTTV
AaKNOX1 AFNAQSNIDMTWFEIRNE--------QEQRVLLKQKYRQELLALKEEFSKRKKRGK-LPTHSIEV
PBX2 VAPKEMERMVSIIHRKFS----------AIQMQLKQSTCEAVMILRSRFLDARRKRRNFSKQATEV
PBX3 ISPKEIERMVGIIHRKFS----------SIQMQLKQSTCEAVMILRSRFLDARRKRRNFSKQATEI
PBX4 VSPKEIERMVGAIHGKFS----------AIQMQLKQSTCEAVMTLRSRLLDARRKRRNFSKQATEV
CEH-20 IAHKEIERMVYIIQRKFN----------GIQVQLKQSTCEAVMILRSRFLDARRKRRNFSKQATEV
CEH-40 ITQQSTEKFMNKMSGKFN----------KVCFVLKQTACEEVIQLKKRYLDARRKRRNFSKTSTEI


521 530 540 550 560 570 580 585
SKN1 LLDWWSLHDKWPYPSETEKIALAECTGLDQKQINNWFINQRKRHWK
SKN2 LLDWWTVHYKWPYPSETEKIALAECTGLDQKQINNWFINQRKRHWK
PmKN3 LLDWWTRNYKWPYPSESQKIALAESTGLDQKQINNWFINQRKRHWK
KNAT1 LLTWWELHYKWPYPSESEKVALAESTGLDQKQINNWFINQRKRHWK
KNOX10 LLHWWQLHYRWPYPSELEKAALAESTGLEAKQINNWFINQRKRHWK
KN1 LLSWWDQHYKWPYPSETQKVALAESTGLDLKQINNWFINQRKRHWK
KNOX4 LLHWWELHYKWPYPSETEKIALAEATGLDQKQINNWFINQRKRHWK
RS1 LLHWWELHYKWPYPSETEKIALAESTGLDQKQINNWFINQRKRHWK
STM LLDWWSRHYKWPYPSEQQKLALAESTGLDQKQINNWFINQRKRHWK
CRKNOX1 LLDWWNQHYKWPYPSEAEKAALAETTGLDQKQINNWFINQRKRHWK
CRKNOX2 LLDWWTQHYKWPYPSEAEKTALAESTGLDQKQINNWFINQRKRHWK
KNOX8 LLHWWELHYKWPYPSETEKMALAETTGLDPKQINNWFINQRKRHWK
KNAT2/ATK1 LLDWWNVHNKWPYPTEGDKIALAEETGLDQKQINNWFINQRKRHWK
KNAT6 LLDWWNLHYKWPYPTEGDKIALADATGLDQKQINNWFINQRKRHWK
LG3 LLEWWNTHYRWPYPTEEDKVRLAAMTGLDPKQINNWFINQRKRHWK
LG4 LMDWWNTHYRWPYPTEEDKVRLAAATGLDPKQINNWFINQRKRHWK
KNOX5 LMDWWNTHYRWPYPTEEDKVRLAAMTGLDPKQINNWFINQRKRHWK
SkKNOX1 LLNWWSVHYKWPYPSESEKASLAESTGLDQKQINNWFINQRKRHWK
SuKNOX1 LFQWWSEHLDHPYPTEVEKAQLCEITRLDAKQINNWFINQRKRHWK
SkKNOX2 LRDWWFQHLEHPYPSEAQKATLAATTKLDPKQINNWFINQRKRHWD
PpMKN2 LKDWFNRHSHWPYPSEMEKQYLQRICGLNLKQINNWFINERKRHWS
PpMKN4 LKDWFSRHSYWPYPSEMEKAYLQRLCGLNLKQINNWFINERKRHWS
SuKNOX2 LKSWWHAHSKWPYPSEDDKARLVQETGLELKQINNWFINQRKRNWH
CRKNOX3 LKAWWHAHSKWPYPTEDEKARLVQETGLQLKQINNWFINQRKRNWH
SkKNOX3 LKAWWHAHSKWPYPTEDEKARLVQETGLELKQINNWFINQRKRNWH
KNAT7 LKNWWQQHCKWPYPTEDDKAKLVEETGLQLKQINNWFINQRKRNWH
KNOX1 LKQWWQEHSKWPYPTEDDKAKLVEETGLQLKQINNWFINQRKRNWH
KNAT3 LKAWWQSHSKWPYPTEEDKARLVQETGLQLKQINNWFINQRKRNWH
KNAT4 LKSWWQSHSKWPYPTEEDKARLVQETGLQLKQINNWFINQRKRNWH
KNOX6 LKAWWQAHSKWPYPTEDDKARLVQETGLQLKQINNWFINQRKRNWH
KNAT5 LKEWWRTHSKWPYPTEEDKAKLVQETGLQLKQINNWFINQRKRNWN
PpMKN1-3 LKAWWQAHSKWPYPTEDEKERRIQETGLELKQVNNWFINQRKRNWH
AaKNOX1 LKSWWKEHIAWPYPTDSAKRSLASQTNLTSIQINNWFINQRKRHWHKLFPEGVPNSQEEALRSLK
PBX2 LNEYFYSHLSNPYPSEEAKEELAKKCGITVSQVSNWFGNKRIRYKKN
PBX3 LNEYFYSHLSNPYPSEEAKEELAKKCSITVSQVSNWFGNKRIRYKKN
PBX4 LNEYFYSHLNNPYPSEEAKEELARKGGLTISQVSNWFGNKRIRYKKN
CEH-20 LNEYFYGHLSNPYPSEEAKEDLARQCNITVSQVSNWFGNKRIRYKKN
CEH-40 LNEYFLANINHPYPSEEVKQALAMQCNISVAQVSNWFGNKRIRYKKT
586 590 600 610 620 630 640 650
SKN1 PSEDMHFMVMNS-HSPHSAALYVERHLMTEG--YHLDC
SKN2 PSEDMQLMAMDG-QSPHGATLYVERHLMTEG--YHLDC
PmKN3 PSEEMQFVVMDS-PNPHNAAFFLEGHLRTDGTAFSMDC
KNAT1 PSEDMQFMVMDGLQHPHHAALYMDGHYMGDG-PYRLGP
KNOX10 QA
KN1 PSEEMHHLMMDGYHTTN--AFYMDGHFINDGGLYRLG
KNOX4 PS
RS1 PSEDMPFVMMEGFHPQNAAALYMDGPFMRDG-MYRLGS
STM PSEDMQFVVMDATHPHH---YFMDNVLDNPFPMDHISSTML
CRKNOX1 PSEDMQYVMVDSPTAHHHHHVLHGHAHLTPHPLAPYAVMETMDAAAAAAAAAVTMLPS
CRKNOX2 PSEDMQYVMMDSPAGQTQHTFLRPHSHIASQHLSPYTVLQTMEVAAGAAPSATMMSS
KNOX8 PA
KNAT2/ATK1 PSENMPFDMMDD----------SNETFFTEE
KNAT6 PSENMPFAMMDD----------SSGSFFTEE
LG3 PSEDMRFALMEG--VAGG--SSGTTLYFDTGTIGP
LG4 PSEDMRFALMEG--VTGGGPSSGTTLYFDTGTIGP
KNOX5 PS
SkKNOX1 PSDELTALSGQP----------SQSTEASSGS
SuKNOX1 PSDDISPLGGQA---------SQSTAGETNSGA
SkKNOX2 PSAAAASARGES---------LQQQGSQDGD
PpMKN2 CEGKCMHPNAKFYGTSNGQCRGHLE?EQSNRLQ?SSEETIYVKCSHAALIFEVSM--
PpMKN4 CKGKCMYPNTKFYPRD-----GHVD--PNNHGEGYLEQ
SuKNOX2 SNPS-SSTSLKNKRKR
CRKNOX3 SNPS-STAAMKTKRKR
SkKNOX3 HHPS-SSASTTSKLKCKS
KNAT7 NNSH-SLTSLKSKRKH
KNOX1 NN
KNAT3 SNPS-SSTVLKNKRKSNAGDNSGRERFA
KNAT4 SNPS-SSTVSKNKRRSNAGENSGRDR
KNOX6 SN
KNAT5 SNSSTSSTLTKNKRK-----RTGKS
PpMKN1-3 SNPLSSSSELKSKRKK
AaKNOX1 ARGMLGMDSSGPMRLMSMDIESQETQEVEQETEDIQTPNEFQFQSAFLNEHEQMMDEAGFTSRPN
PBX2 ----IGKFQEEANIYAVKTAVSVTQG-----GHSRTSSPTPPSSAGSGGSFNLSGSGDMFLGMPG
PBX3 ----IGKFQEEANLYAAKTAVTAAHAVAAAVQNNQTNSPTTPN-SGSSGSFNLPNSGDMFMNMQS
PBX4 ----MGKFQEEATIYTGKTAVDTTEVG---VPGNHASCLSTPS-SGSSGPFPLPSAGDAFLTLRT
CEH-20 ----MAKAQEEASMYAAKKNAHVTLG-----GMAGNPYGMLPGAAAAAGLLNPYNP----MNIP-
CEH-40 ----MAKNEDER--RENRKPEDRPPP-----GAPGAPYSLVPN--AFAGMMNPYQM-----MLP-

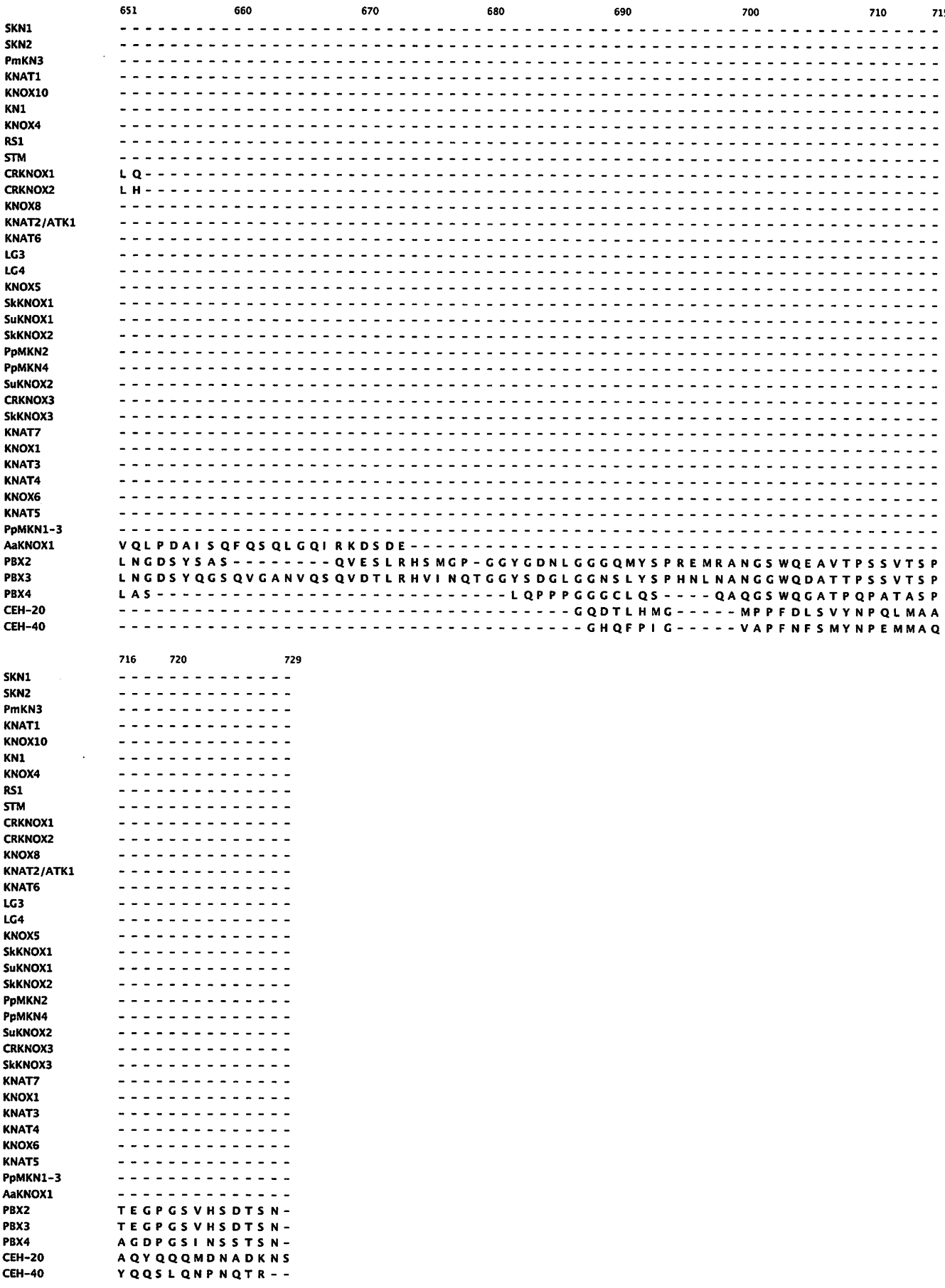


Fig. 2. Alignment of deduced amino acid sequences of *KNOX* genes.

Dashes indicate gaps. The underlined amino acids were used in the phylogenetic analysis shown in Fig. 5.

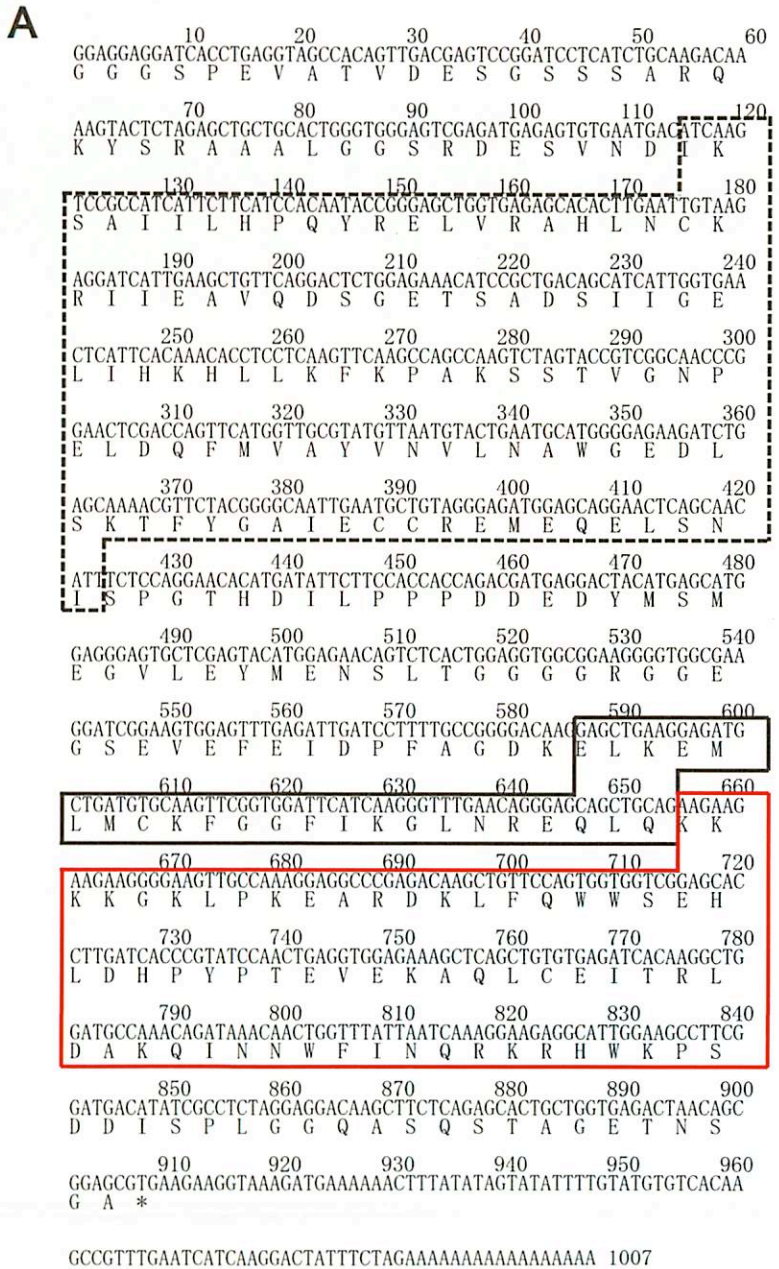


Fig. 3. The nucleotide and deduced amino acid sequences of the *SuKNOX1* (A) and *SuKNOX2* (B) cDNA isolated in this study.

The numberings are nucleotide position when the end of the isolated sequence is 1. The MEINOX domain is broken-line boxed. The ELK domain is solid-line boxed. The homeo domain is red-line boxed.

Figure 3

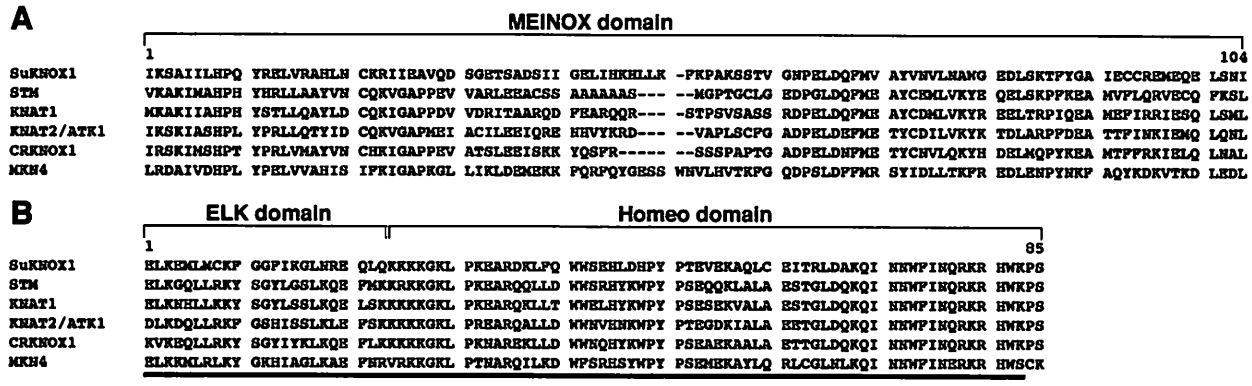


Fig. 4. Alignment of deduced amino acid sequences of *SuKNOX1* and representative class 1 *KNOX* genes from *Arabidopsis thaliana* (*STM*, *KNAT1*, and *KNAT2/ATK*), the fern *Ceratopteris richardii* (*CRKNOX1*), and the moss *Physcomitrella patens* (*MKN4*), including (A) the MEINOX domain and (B) the ELK and homeo domains. The underlined amino acids were used in the phylogenetic analysis shown in Fig. 5.

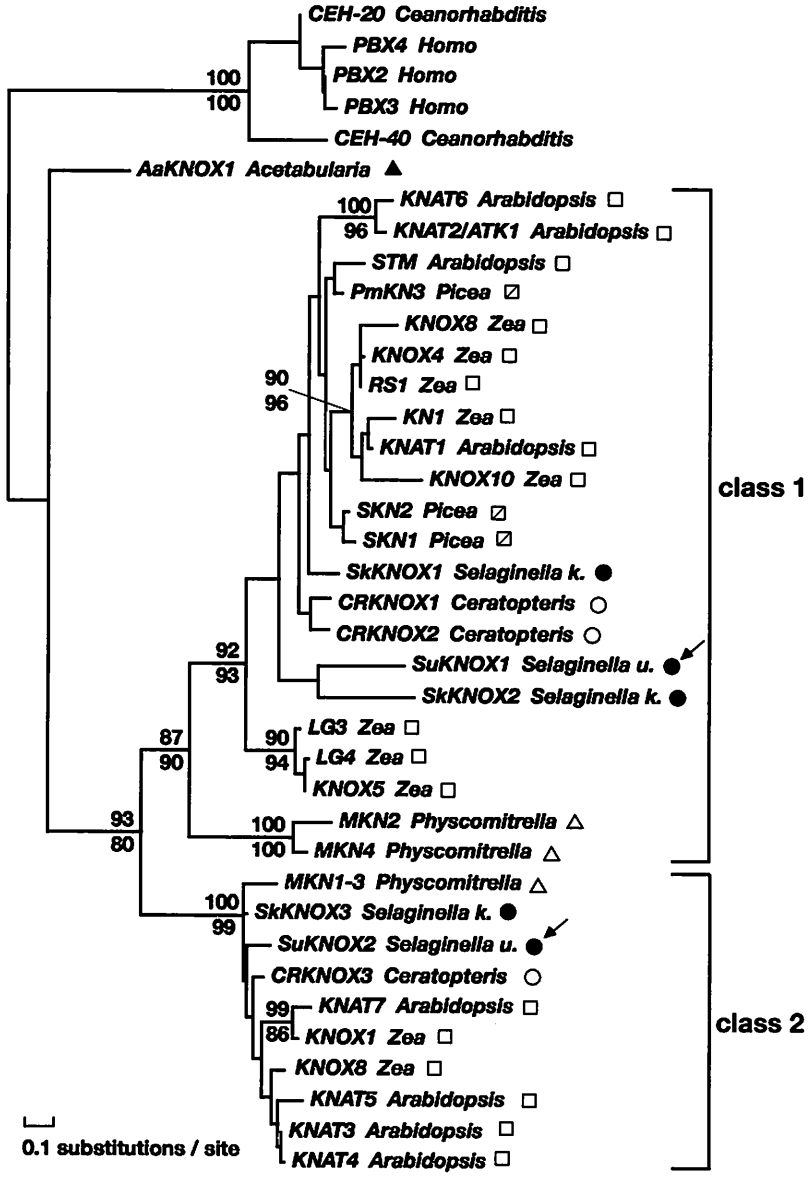


Fig. 5. Maximum likelihood (ML) tree.

ML tree showing the phylogenetic relationships among *KNOX* genes of the spikemoss and representatives from other land plants, a green alga, and metazoans (outgroup). Symbols following the genus names represent plant classifications: open square, angiosperms; square filled with an angled bar, gymnosperms; open circle, ferns; filled circle, spikemosses; open triangle, mosses; filled triangle, green alga. The *Selaginella uncinata* genes are indicated by arrows. Bootstrap values calculated using the ML and neighbor-joining (NJ) methods are indicated above and below the nodes, respectively; only values exceeding 80% and supported by both the ML and NJ topologies are indicated reliably. The length of the bar represents 0.1 amino acid substitutions per residue.

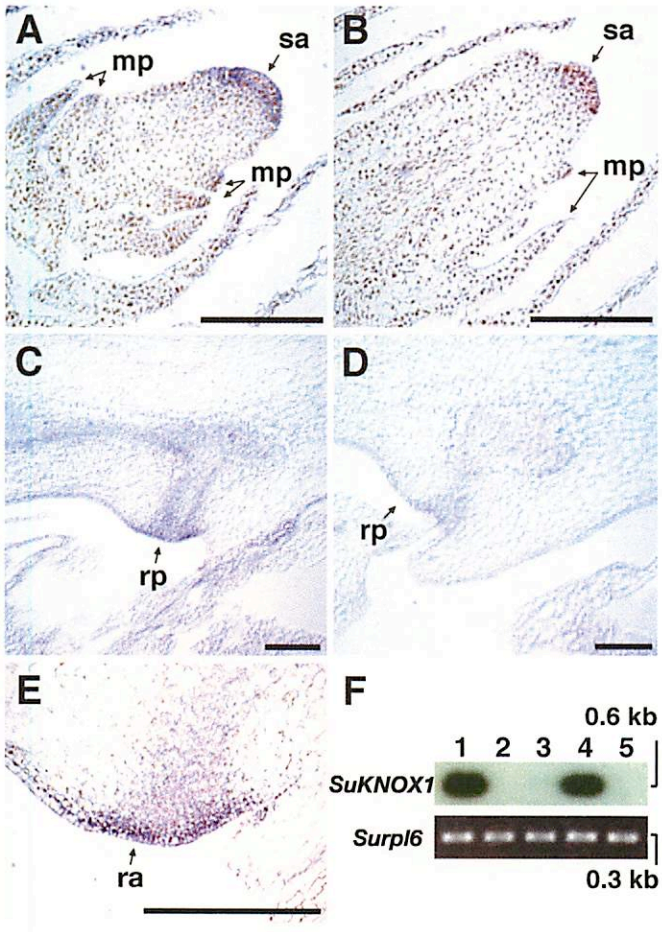


Fig. 6. Expression pattern of the *SuKNOX1*.

(A-E) Longitudinal sections show the shoot apex (sa), microphyll primordium (mp), rhizophore primordium (rp), and rhizophore apical tip (ra). Scale bars = 100 μm. (A, C, E) The locations of *SuKNOX1* mRNA expression were detected using *in situ* hybridization. (B, D) The sense probe was used as a negative control. (F) Amplification of *SuKNOX1* RT-PCR products. Complementary DNA was synthesized from RNA extracted from the apical tips of microphylls (lane 1), internodes (lane 2), microphylls (lane 3), rhizophore tips (lane 4), and root tips (lane 5). *SuKNOX1* PCR products were hybridized with *SuKNOX1*-specific probes. The *SuRPL6* was used as a quantifying control.

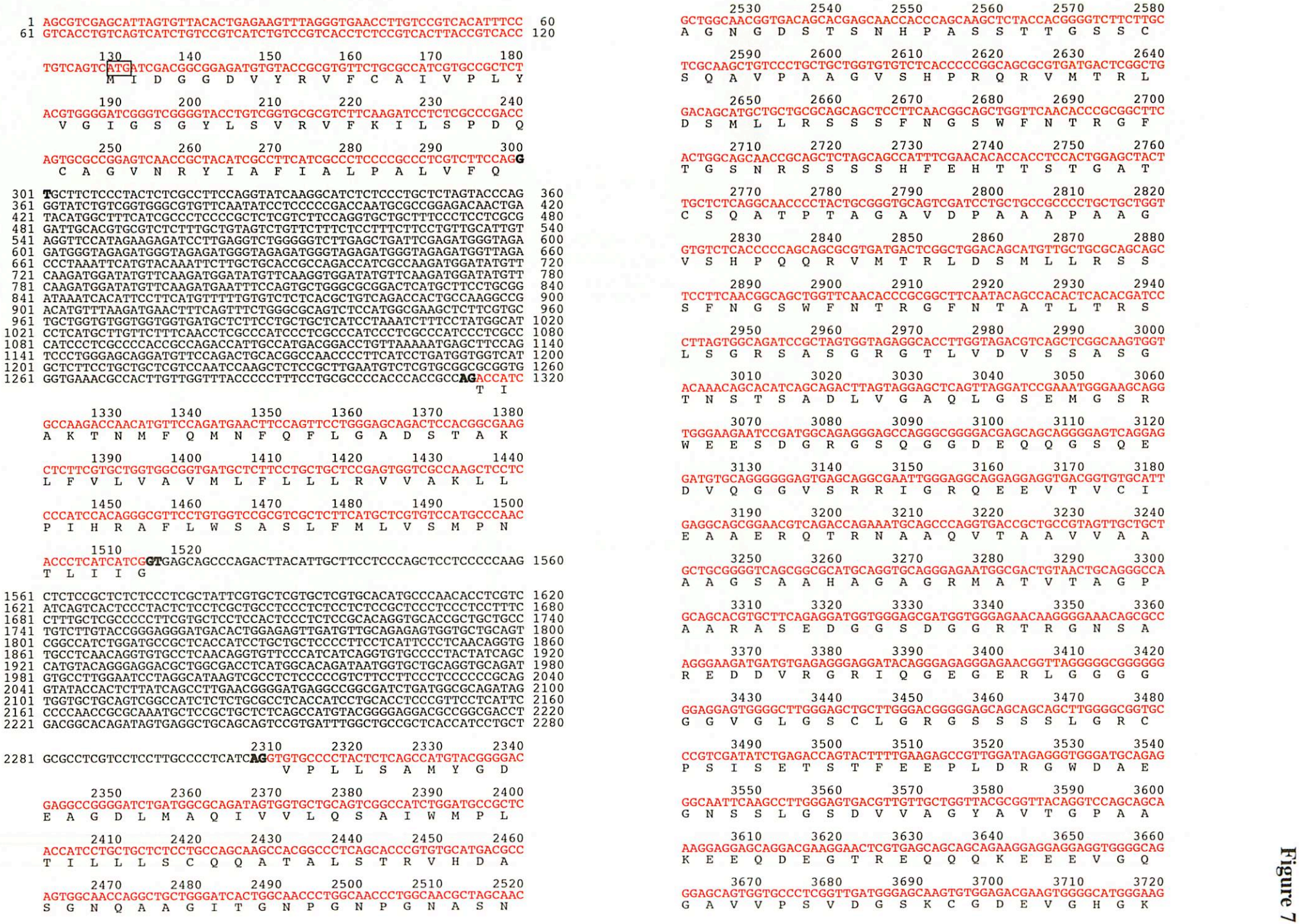


Figure 7

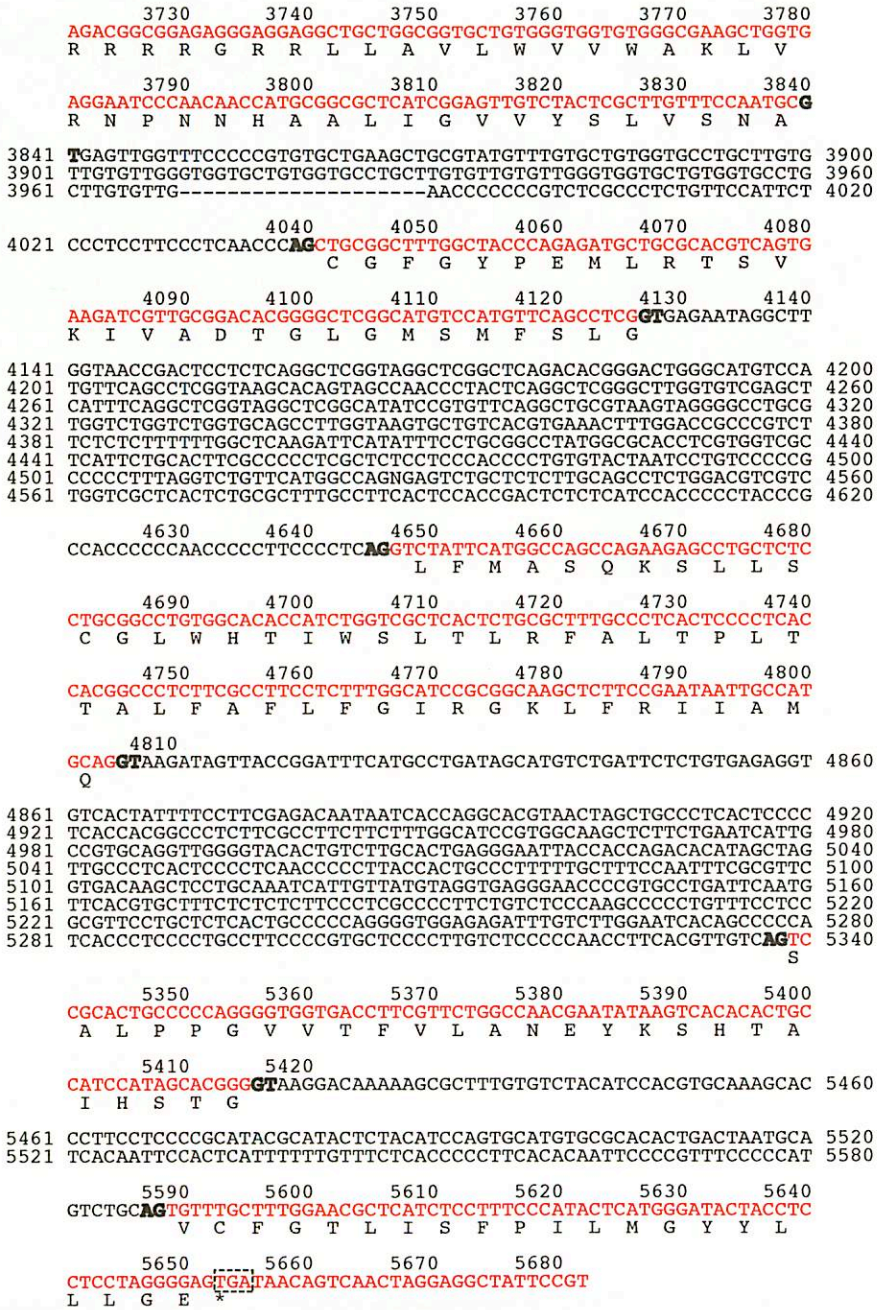


Fig. 7. *CpslcPIN* genomic DNA, cDNA, and deduced amino acid sequences isolated in this study.

The numberings are nucleotide position when the end of the isolated sequence is 1. The putative start codon is boxed. Broken-lined box indicates stop codon. Bold "GT" and "AG" show the exon-intron splice junctions according to the GT-AG rule (Mount, 1982). cDNA sequence is indicated by red. An unclear domain in intron 3 is indicated by gap (-).

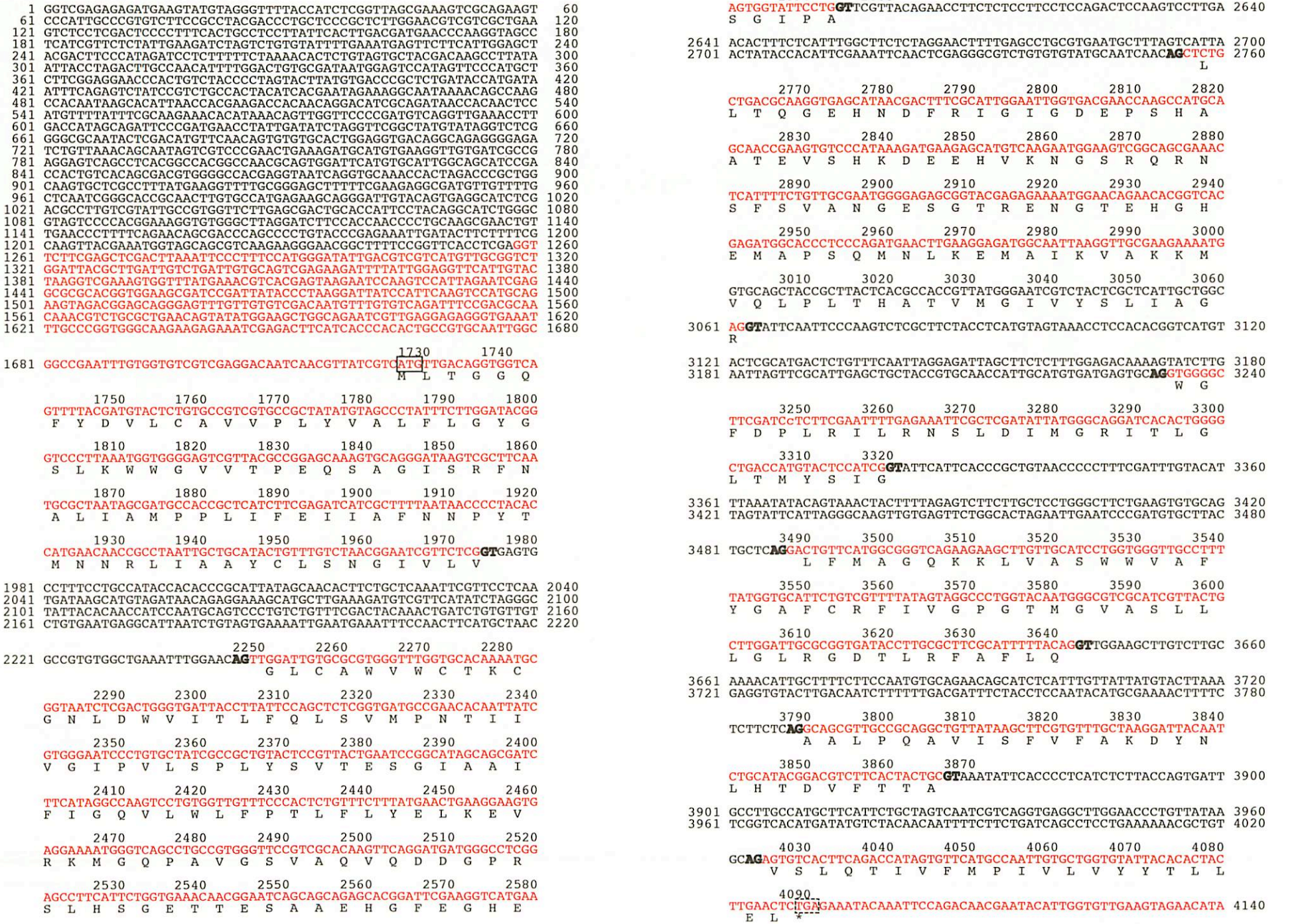


Figure 8

```
4141 AACGTGGCGAATCTTGGTCGCCAACAACATCAAAAGTGGGGAACGATTGTAGAACTCTGA 4200
4201 TGCTGCGTCGGCCCCAAGGTGACTGATTGGTCACATTCATACTTATGTTGATTAATAGTT 4260
4261 TTACAGGAAATTATAACACTCACTGGACTCGGATTTTTAACCATAAATCACTTGCAAGGA 4320
4321 CTTCCTTCGGAAAAAATGGTTATTTTGTTGCAAAGCTAACCATGTGGGGGTAAACAATTT 4380
4381 AGTATCGAAGGGGTGACACGCCGAGCAAAGTGAACAGACTACGTGTGCAGAGGAAGATTC 4440
4441 AGGACTGAAACTGTAACGAAGACTGGAGATATACCACGTATTTTATATTATCAACTGATA 4500
4501 AGTGGGGATCAGAGAACCCAAGTGAAGACATGCCTCATTTGAAGGCAAGCTCCAGCACAT 4560
4561 AAGTTTTGTTACATGGGTTACGATGTATATGCACGTGTAGGTTGAGTTTAAGCACGTACA 4620
4621 ACAATTTAAACACCAACAAGTGATGTGAAAGTTGGAAGATGAGATTAGGGAAGAAGTAGA 4680
4681 AGTTATTAGTAAATTATTCACTCATTCCATCATCTCAGAAACAGGTCAATGGTACCACAG 4740
4741 CAGGTGCAGAGCACTATATACTGGATCGGATATATACCTCTAAAGATTCACTCTTGCTGT 4800
4801 CCGAGGATGGTATATTTCTCAAGTTGAAAGTAAAGTGAGTTAACTTCAATTAATGAAGTT 4860
4861 CAAAAGCTGTGTTTTCACACTTGAGTTGGGCCTGAGGAATAACGATGTTACAAAAAGTAC 4920
4921 AATGCTAGTTACAGCCTTTTTGGGCGAAAACAATATACATGTGAACTGAAGTTAGGACTA 4980
4981 CATGCTTCAACTAAGAGGTTATGTGACGCTGTTGAGAGAATTTTAATGCGCAATAATACA 5040
5041 CAATATAGTTACACAATGGCAAAGATTACAATATATCACAGTATGGTCCGAGGCATTCTG 5100
5101 TTGGTCCGCTGAGTCTCTGCTTCTCCACT   5129
```

Fig. 8. *PhypaPIND* genomic DNA, cDNA, and deduced amino acid sequences isolated in this study.

The numberings are nucleotide position when the end of the isolated sequence is 1. The putative start codon is boxed. Broken-lined box indicates stop codon. Bold "GT" and "AG" show the exon-intron splice junctions according to the GT-AG rule (Mount, 1982). cDNA sequence is indicated by red.

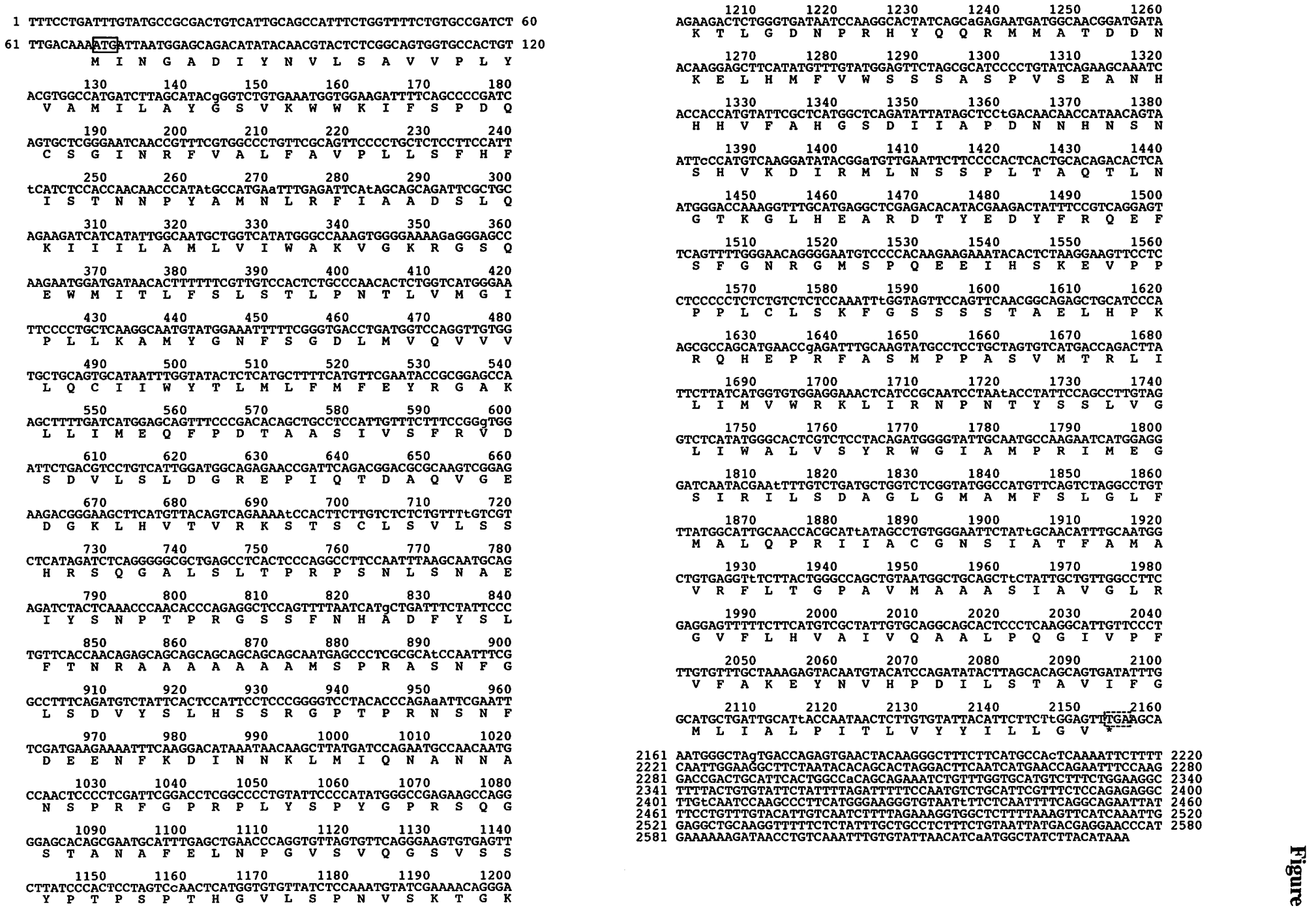

1 TTTCCTGATTTGTATGCCGCGACTGTCATTGCAGCCATTTCTGGTTTTCTGTGCCGATCT 60
61 TTGACAAAATGATTAATGGAGCAGACATATACAACGTACTCTCGGCAGTGGTGCCACTGT 120
M I N G A D I Y N V L S A V V P L Y
130 140 150 160 170 180
ACGTGGCCATGATCTTAGCATACgGGTCTGTGAAATGGTGGAAGATTTTCAGCCCCGATC
V A M I L A Y G S V K W W K I F S P D Q
190 200 210 220 230 240
AGTGCTCGGGAATCAACCGTTTCGTGGCCCTGTTCGCAGTTCCCCTGCTCTCCTTCCATT
C S G I N R F V A L F A V P L L S F H F
250 260 270 280 290 300
tCATCTCCACCAACAACCCATAtGCCATGAaTTTGAGATTCAtAGCAGCAGATTCGCTGC
I S T N N P Y A M N L R F I A A D S L Q
310 320 330 340 350 360
AGAAGATCATCATATTGGCAATGCTGGTCATATGGGCCAAAGTGGGGAAAAGaGGGAGCC
K I I I L A M L V I W A K V G K R G S Q
370 380 390 400 410 420
AAGAATGGATGATAACACTTTTTTCGTTGTCCACTCTGCCCAACACTCTGGTCATGGGAA
E W M I T L F S L S T L P N T L V M G I
430 440 450 460 470 480
TTCCCCTGCTCAAGGCAATGTATGGAAATTTTTCGGGTGACCTGATGGTCCAGGTTGTGG
P L L K A M Y G N F S G D L M V Q V V V
490 500 510 520 530 540
TGCTGCAGTGCATAATTTGGTATACTCTCATGCTTTCATGTTCGAATACCGCGGAGCCA
L Q C I I W Y T L M L F M F E Y R G A K
550 560 570 580 590 600
AGCTTTTGATCATGGAGCAGTTTCCCGACACAGCTGCCTCCATTGTTTCTTTCCGGgTGG
L L I M E Q F P D T A A S I V S F R V D
610 620 630 640 650 660
ATTCTGACGTCCTGTCATTGGATGGCAGAGAACCGATTCAGACGGACGCGCAAGTCGGAG
S D V L S L D G R E P I Q T D A Q V G E
670 680 690 700 710 720
AAGACGGGAAGCTTCATGTTACAGTCAGAAAAtCCACTTCTTGTCTCTCTGTTTtGTCGT
D G K L H V T V R K S T S C L S V L S S
730 740 750 760 770 780
CTCATAGATCTCAGGGGGCGCTGAGCCTCACTCCCAGGCCTTCCAATTTAAGCAATGCAG
H R S Q G A L S L T P R P S N L S N A E
790 800 810 820 830 840
AGATCTACTCAAACCCAACACCCAGAGGCTCCAGTTTTAATCATgCTGATTTCTATTCCC
I Y S N P T P R G S S F N H A D F Y S L
850 860 870 880 890 900
TGTTCACCAACAGAGCAGCAGCAGCAGCAGCAATGAGCCCTCGCGCAtCCAATTTCG
F T N R A A A A A A A M S P R A S N F G
910 920 930 940 950 960
GCCTTTCAGATGTCTATTCACTCCATTCCTCCCGGGGTCCTACACCCAGAaATTCGAATT
L S D V Y S L H S S R G P T P R N S N F
970 980 990 1000 1010 1020
TCGATGAAGAAAATTTCAAGGACATAAATAACAAGCTTATGATCCAGAATGCCAACAATG
D E E N F K D I N N K L M I Q N A N N A
1030 1040 1050 1060 1070 1080
CCAACTCCCCTCGATTCGGACCTCGGCCCCTGTATTCCCCATATGGGCCGAGAAGCCAGG
N S P R F G P R P L Y S P Y G P R S Q G
1090 1100 1110 1120 1130 1140
GGAGCACAGCGAATGCATTTGAGCTGAACCCAGGTGTTAGTGTTCAGGGAAGTGTGAGTT
S T A N A F E L N P G V S V Q G S V S S
1150 1160 1170 1180 1190 1200
CTTATCCCACTCCTAGTCcAACTCATGGTGTGTTATCTCCAAATGTATCGAAAACAGGGA
Y P T P S P T H G V L S P N V S K T G K
1210 1220 1230 1240 1250 1260
AGAAGACTCTGGGTGATAATCCAAGGCACTATCAGCaGAGAATGATGGCAACGGATGATA
K T L G D N P R H Y Q Q R M M A T D D N
1270 1280 1290 1300 1310 1320
ACAAGGAGCTTCATATGTTTGTATGGAGTTCTAGCGCATCCCCTGTATCAGAAGCAAATC
K E L H M F V W S S S A S P V S E A N H
1330 1340 1350 1360 1370 1380
ACCACCATGTATTCGCTCATGGCTCAGATATTATAGCTCCtGACAACAACCATAACAGTA
H H V F A H G S D I I A P D N N H N S N
1390 1400 1410 1420 1430 1440
ATTcCCATGTCAAGGATATACGGaTGTTGAATTCTTCCCCACTCACTGCACAGACACTCA
S H V K D I R M L N S S P L T A Q T L N
1450 1460 1470 1480 1490 1500
ATGGGACCAAAGGTTTGCATGAGGCTCGAGACACATACGAAGACTATTTCCGTCAGGAGT
G T K G L H E A R D T Y E D Y F R Q E F
1510 1520 1530 1540 1550 1560
TCAGTTTTGGGAACAGGGGAATGTCCCCACAAGAAGAAATACACTCTAAGGAAGTTCCTC
S F G N R G M S P Q E E I H S K E V P P
1570 1580 1590 1600 1610 1620
CTCCCCCTCTCTGTCTCTCCAAATTtGGTAGTTCCAGTTCAACGGCAGAGCTGCATCCCA
P P L C L S K F G S S S S T A E L H P K
1630 1640 1650 1660 1670 1680
AGCGCCAGCATGAACCgAGATTTGCAAGTATGCCTCCTGCTAGTGTCATGACCAGACTTA
R Q H E P R F A S M P P A S V M T R L I
1690 1700 1710 1720 1730 1740
TTCTTATCATGGTGTGGAGGAAACTCATCCGCAATCCTAAtACCTATTCCAGCCTTGTAG
L I M V W R K L I R N P N T Y S S L V G
1750 1760 1770 1780 1790 1800
GTCTCATATGGGCACTCGTCTCCTACAGATGGGGTATTGCAATGCCAAGAATCATGGAGG
L I W A L V S Y R W G I A M P R I M E G
1810 1820 1830 1840 1850 1860
GATCAATACGAAtTTTGTCTGATGCTGGTCTCGGTATGGCCATGTTCAGTCTAGGCCTGT
S I R I L S D A G L G M A M F S L G L F
1870 1880 1890 1900 1910 1920
TTATGGCATTGCAACCACGCATtATAGCCTGTGGGAATTCTATtGCAACATTTGCAATGG
M A L Q P R I I A C G N S I A T F A M A
1930 1940 1950 1960 1970 1980
CTGTGAGGTtTCTTACTGGGCCAGCTGTAATGGCTGCAGCTtCTATTGCTGTTGGCCTTC
V R F L T G P A V M A A A S I A V G L R
1990 2000 2010 2020 2030 2040
GAGGAGTTTTTCTTCATGTCGCTATTGTGCAGGCAGCACTCCCTCAAGGCATTGTTCCCT
G V F L H V A I V Q A A L P Q G I V P F
2050 2060 2070 2080 2090 2100
TTGTGTTTGCTAAAGAGTACAATGTACATCCAGATATACTTAGCACAGCAGTGATATTTG
V F A K E Y N V H P D I L S T A V I F G
2110 2120 2130 2140 2150 2160
GCATGCTGATTGCATtACCAATAACTCTTGTGTATTACATTCTTCTtGGAGTTTGAAGCA
M L I A L P I T L V Y Y I L L G V *
2161 AATGGGCTAgTGACCAGAGTGAACTACAAGGGCTTTCTTCATGCCAcTCAAAATTCTTTT 2220
2221 CAATTGGAAGGCTTCTAATACACAGCACTAGGACTTCAATCATGAACCAGAATTTCCAAG 2280
2281 GACCGACTGCATTCACTGGCCaCAGCAGAAATCTGTTTGGTGCATGTCTTTCTGGAAGGC 2340
2341 TTTTACTGTGTATTCTATTTTAGATTTTTCCAATGTCTGCATTCGTTTCTCCAGAGAGGC 2400
2401 TTGtCAATCCAAGCCCTTCATGGGAAGGGTGTAATtTTCTCAATTTTCAGGCAGAATTAT 2460
2461 TTCCTGTTTGTACATTGTCAATCTTTTAGAAAGGTGGCTCTTTTAAAGTTCATCAAATTG 2520
2521 GAGGCTGCAAGGTTTTTCTCTATTTGCTGCCTCTTTCTGTAATTATGACGAGGAACCCAT 2580
2581 GAAAAAAGATAACCTGTCAAATTTGTGTATTAACATCaATGGCTATCTTACATAAA

Figure 9

Fig. 9. *PinthPIN* cDNA and deduced amino acid sequences isolated in this study.

The numberings are nucleotide position when the end of the isolated sequence is 1. The putative start codon is boxed. Broken-lined box indicates stop codon. Bold "GT" and "AG" show the exon-intron splice junctions according to the GT-AG rule (Mount, 1982).

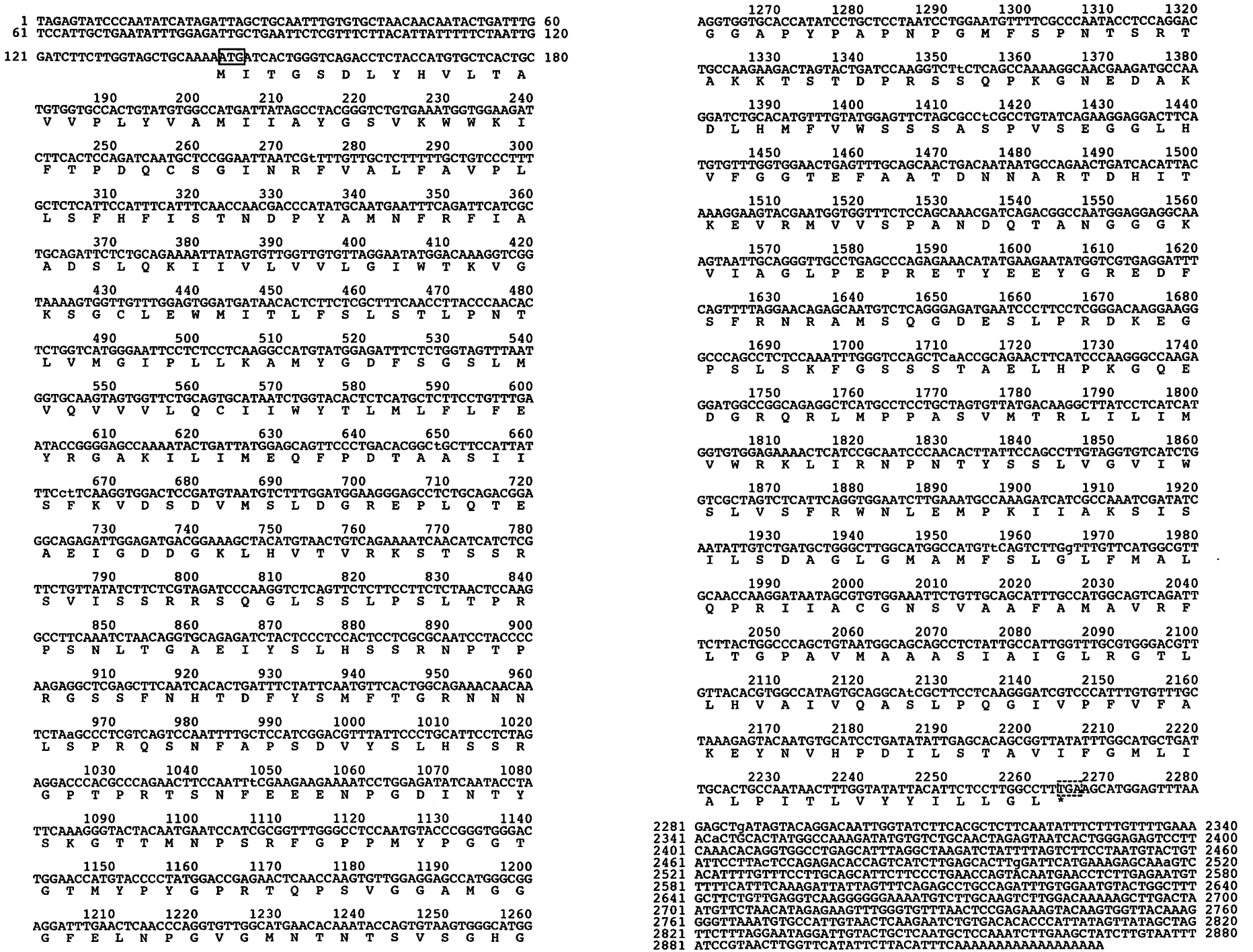

1 TAGAGTATCCCAATATCATAGATTAGCTGCAATTTGTGTGCTAACAACAATACTGATTTG 60
61 TCCATTGCTGAATATTTGGAGATTGCTGAATTCTCGTTTCTTACATTATTTTTCTAATTG 120
121 GATCTTCTTGGTAGCTGCAAAAATGATCACTGGGTCAGACCTCTACCATGTGCTCACTGC 180
M I T G S D L Y H V L T A
190 200 210 220 230 240
TGTGGTGCCACTGTATGTGGCCATGATTATAGCCTACGGGTCTGTGAAATGGTGGAAGAT
V V P L Y V A M I I A Y G S V K W W K I
250 260 270 280 290 300
CTTCACTCCAGATCAATGCTCCGGAATTAATCGtTTTGTTGCTCTTTTTGCTGTCCCTTT
F T P D Q C S G I N R F V A L F A V P L
310 320 330 340 350 360
GCTCTCATTCCATTTCATTTCAACCAACGACCCATATGCAATGAATTTCAGATTCATCGC
L S F H F I S T N D P Y A M N F R F I A
370 380 390 400 410 420
TGCAGATTCTCTGCAGAAAATTATAGTGTTGGTTGTGTTAGGAATATGGACAAAGGTCGG
A D S L Q K I I V L V V L G I W T K V G
430 440 450 460 470 480
TAAAAGTGGTTGTTTGGAGTGGATGATAACACTCTTCTCGCTTTCAACCTTACCCAACAC
K S G C L E W M I T L F S L S T L P N T
490 500 510 520 530 540
TCTGGTCATGGGAATTCCTCTCCTCAAGGCCATGTATGGAGATTTCTCTGGTAGTTTAAT
L V M G I P L L K A M Y G D F S G S L M
550 560 570 580 590 600
GGTGCAAGTAGTGGTTCTGCAGTGCATAATCTGGTACACTCTCATGCTCTTCCTGTTTGA
V Q V V V L Q C I I W Y T L M L F L F E
610 620 630 640 650 660
ATACCGGGGAGCCAAAATACTGATTATGGAGCAGTTCCCTGACACGGCtGCTTCCATTAT
Y R G A K I L I M E Q F P D T A A S I I
670 680 690 700 710 720
TTCctTCAAGGTGGACTCCGATGTAATGTCTTTGGATGGAAGGGAGCCTCTGCAGACGGA
S F K V D S D V M S L D G R E P L Q T E
730 740 750 760 770 780
GGCAGAGATTGGAGATGACGGAAAGCTACATGTAACTGTCAGAAAATCAACATCATCTCG
A E I G D D G K L H V T V R K S T S S R
790 800 810 820 830 840
TTCTGTTATATCTTCTCGTAGATCCCAAGGTCTCAGTTCTCTTCCTTCTCTAACTCCAAG
S V I S S R R S Q G L S S L P S L T P R
850 860 870 880 890 900
GCCTTCAAATCTAACAGGTGCAGAGATCTACTCCCTCCACTCCTCGCGCAATCCTACCCC
P S N L T G A E I Y S L H S S R N P T P
910 920 930 940 950 960
AAGAGGCTCGAGCTTCAATCACACTGATTTCTATTCAATGTTCACTGGCAGAAACAACAA
R G S S F N H T D F Y S M F T G R N N N
970 980 990 1000 1010 1020
TCTAaGCCCTCGTCAGTCCAATTTTGCTCCATCGGACGTTTATTCCCTGCATTCCTCTAG
L S P R Q S N F A P S D V Y S L H S S R
1030 1040 1050 1060 1070 1080
AGGACCCACGCCCAGAACTTCCAATTtCGAAGAAGAAAATCCTGGAGATATCAATACCTA
G P T P R T S N F E E E N P G D I N T Y
1090 1100 1110 1120 1130 1140
TTCAAAGGGTACTACAATGAATCCATCGCGGTTTGGGCCTCCAATGTACCCGGGTGGGAC
S K G T T M N P S R F G P P M Y P G G T
1150 1160 1170 1180 1190 1200
TGGAACCATGTACCCCTATGGACCGAGAACTCAACCAAGTGTTGGAGGAGCCATGGGCGG
G T M Y P Y G P R T Q P S V G G A M G G
1210 1220 1230 1240 1250 1260
AGGATTTGAACTCAACCCAGGTGTTGGCATGAACACAAATACCAGTGTAAGTGGGCATGG
G F E L N P G V G M N T N T S V S G H G
1270 1280 1290 1300 1310 1320
AGGTGGTGCACCATATCCTGCTCCTAATCCTGGAATGTTTTCGCCCAATACCTCCAGGAC
G G A P Y P A P N P G M F S P N T S R T
1330 1340 1350 1360 1370 1380
TGCCAAGAAGACTAGTACTGATCCAAGGTCTtCTCAGCCAAAAGGCAACGAAGATGCCAA
A K K T S T D P R S S Q P K G N E D A K
1390 1400 1410 1420 1430 1440
GGATCTGCACATGTTTGTATGGAGTTCTAGCGCCtCGCCTGTATCAGAAGGAGGACTTCA
D L H M F V W S S S A S P V S E G G L H
1450 1460 1470 1480 1490 1500
TGTGTTTGGTGGAACTGAGTTTGCAGCAACTGACAATAATGCCAGAACTGATCACATTAC
V F G G T E F A A T D N N A R T D H I T
1510 1520 1530 1540 1550 1560
AAAGGAAGTACGAATGGTGGTTTCTCCAGCAAACGATCAGACGGCCAATGGAGGAGGCAA
K E V R M V V S P A N D Q T A N G G G K
1570 1580 1590 1600 1610 1620
AGTAATTGCAGGGTTGCCTGAGCCCAGAGAAACATATGAAGAATATGGTCGTGAGGATTT
V I A G L P E P R E T Y E E Y G R E D F
1630 1640 1650 1660 1670 1680
CAGTTTTAGGAACAGAGCAATGTCTCAGGGAGATGAATCCCTTCCTCGGGACAAGGAAGG
S F R N R A M S Q G D E S L P R D K E G
1690 1700 1710 1720 1730 1740
GCCCAGCCTCTCCAAATTTGGGTCCAGCTCaACCGCAGAACTTCATCCCAAGGGCCAAGA
P S L S K F G S S S T A E L H P K G Q E
1750 1760 1770 1780 1790 1800
GGATGGCCGGCAGAGGCTCATGCCTCCTGCTAGTGTTATGACAAGGCTTATCCTCATCAT
D G R Q R L M P P A S V M T R L I L I M
1810 1820 1830 1840 1850 1860
GGTGTGGAGAAAACTCATCCGCAATCCCAACACTTATTCCAGCCTTGTAGGTGTCATCTG
V W R K L I R N P N T Y S S L V G V I W
1870 1880 1890 1900 1910 1920
GTCGCTAGTCTCATTCAGGTGGAATCTTGAAATGCCAAAGATCATCGCCAAATCGATATC
S L V S F R W N L E M P K I I A K S I S
1930 1940 1950 1960 1970 1980
AATATTGTCTGATGCTGGGCTTGGCATGGCCATGTtCAGTCTTGgTTTGTTCATGGCGTT
I L S D A G L G M A M F S L G L F M A L
1990 2000 2010 2020 2030 2040
GCAACCAAGGATAATAGCGTGTGGAAATTCTGTTGCAGCATTTGCCATGGCAGTCAGATT
Q P R I I A C G N S V A A F A M A V R F
2050 2060 2070 2080 2090 2100
TCTTACTGGCCCAGCTGTAATGGCAGCAGCCTCTATTGCCATTGGTTTGCGTGGGACGTT
L T G P A V M A A A S I A I G L R G T L
2110 2120 2130 2140 2150 2160
GTTACACGTGGCCATAGTGCAGGCAtCGCTTCCTCAAGGGATCGTCCCATTTGTGTTTGC
L H V A I V Q A S L P Q G I V P F V F A
2170 2180 2190 2200 2210 2220
TAAAGAGTACAATGTGCATCCTGATATATTGAGCACAGCGGTTATATTTGGCATGCTGAT
K E Y N V H P D I L S T A V I F G M L I
2230 2240 2250 2260 2270 2280
TGCACTGCCAATAACTTTGGTATATTACATTCTCCTTGGCCTTTGAAGCATGGAGTTTAA
A L P I T L V Y Y I L L G L *
2281 GAGCTgATAGTACAGGACAATTGGTATCTTCACGCTCTTCAATATTTCTTTGTTTTGAAA 2340
2341 ACaCTGCACTATGGCCAAAGATATGTGTCTGCAACTAGAGTAATCACTGGGAGAGTCCTT 2400
2401 CAAACACAGGTGGCCTGAGCATTTAGGCTAAGATCTATTTTAGTCTTCCTAATGTACTGT 2460
2461 ATTCCTTACTCCAGAGACACCAGTCATCTTGAGCACTTgGATTCATGAAAGAGCAAaGTC 2520
2521 ACATTTTGTTTCCTTGCAGCATTCTTCCCTGAACCAGTACAATGAACCTCTTGAGAATGT 2580
2581 TTTTCATTTCAAAGATTATTAGTTTCAGAGCCTGCCAGATTTGTGGAATGTACTGGCTTT 2640
2641 GCTTCTGTTGAGGTCAAGGGGGGAAAATGTCTTGCAAGTCTTGGACAAAAAGCTTGACTA 2700
2701 ATGTTCTAACATAGAGAAGTTTGGGTGTTTAACTCCGAGAAAGTACAAGTGGTTACAAAG 2760
2761 GGGTTAAATGTGCCATTGTAACTCAAGAATCTGTGACACACCCATTATAGTTATAGCTAG 2820
2821 TTCTTTAGGAATAGGATTGTACTGCTCAATGCTCCAAATCTTGAAGCTATCTTGTAATTT 2880
2881 ATCCGTAACTTGGTTCATATTCTTACATTTCAAAAAAAAAAAAAAAAAA

Figure 10

Fig. 10. *GinbiPIN* cDNA and deduced amino acid sequences isolated in this study.

The numberings are nucleotide position when the end of the isolated sequence is 1. The putative start codon is boxed. Broken-lined box indicates stop codon. Bold "GT" and "AG" show the exon-intron splice junctions according to the GT-AG rule (Mount, 1982).

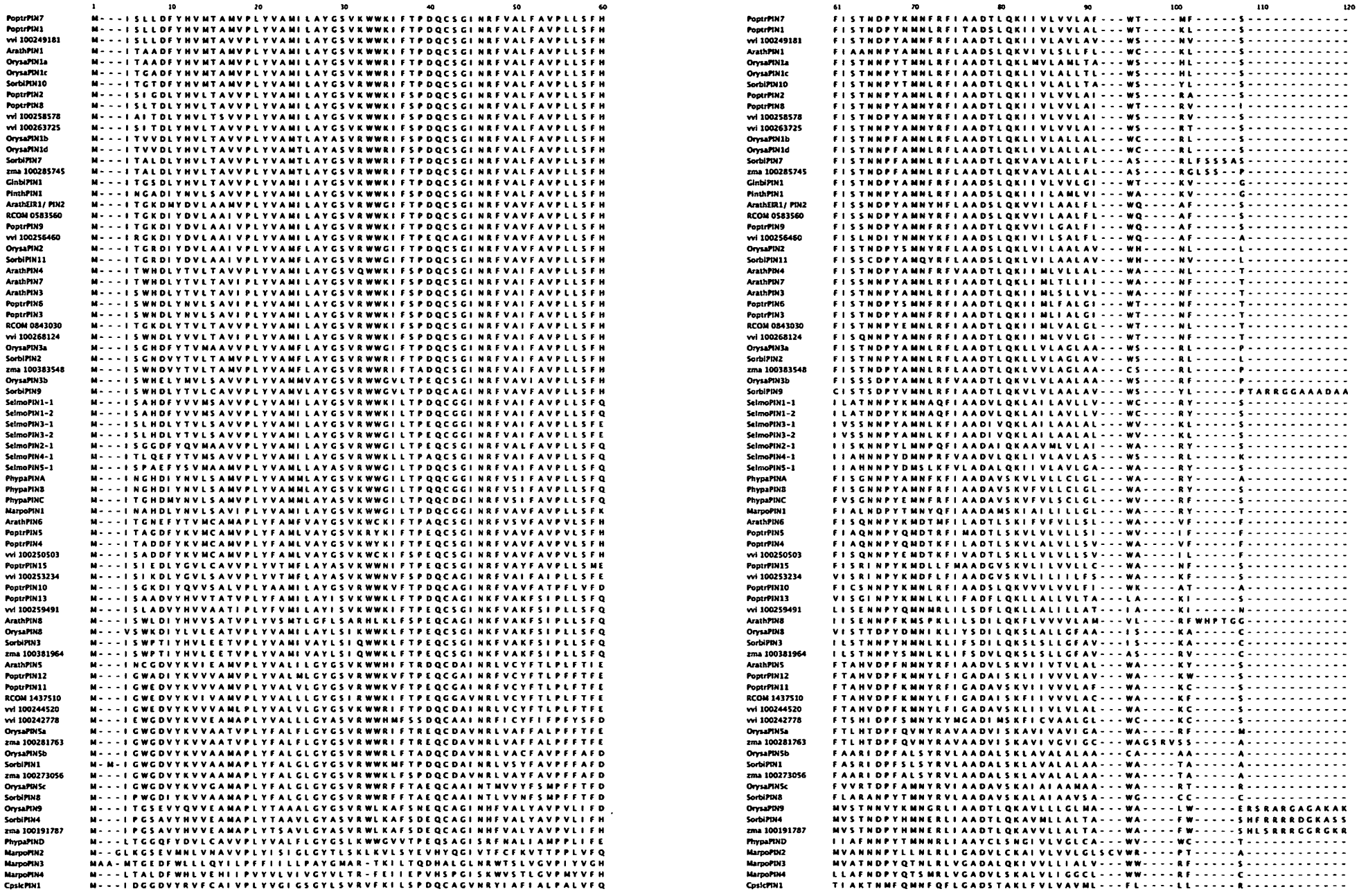
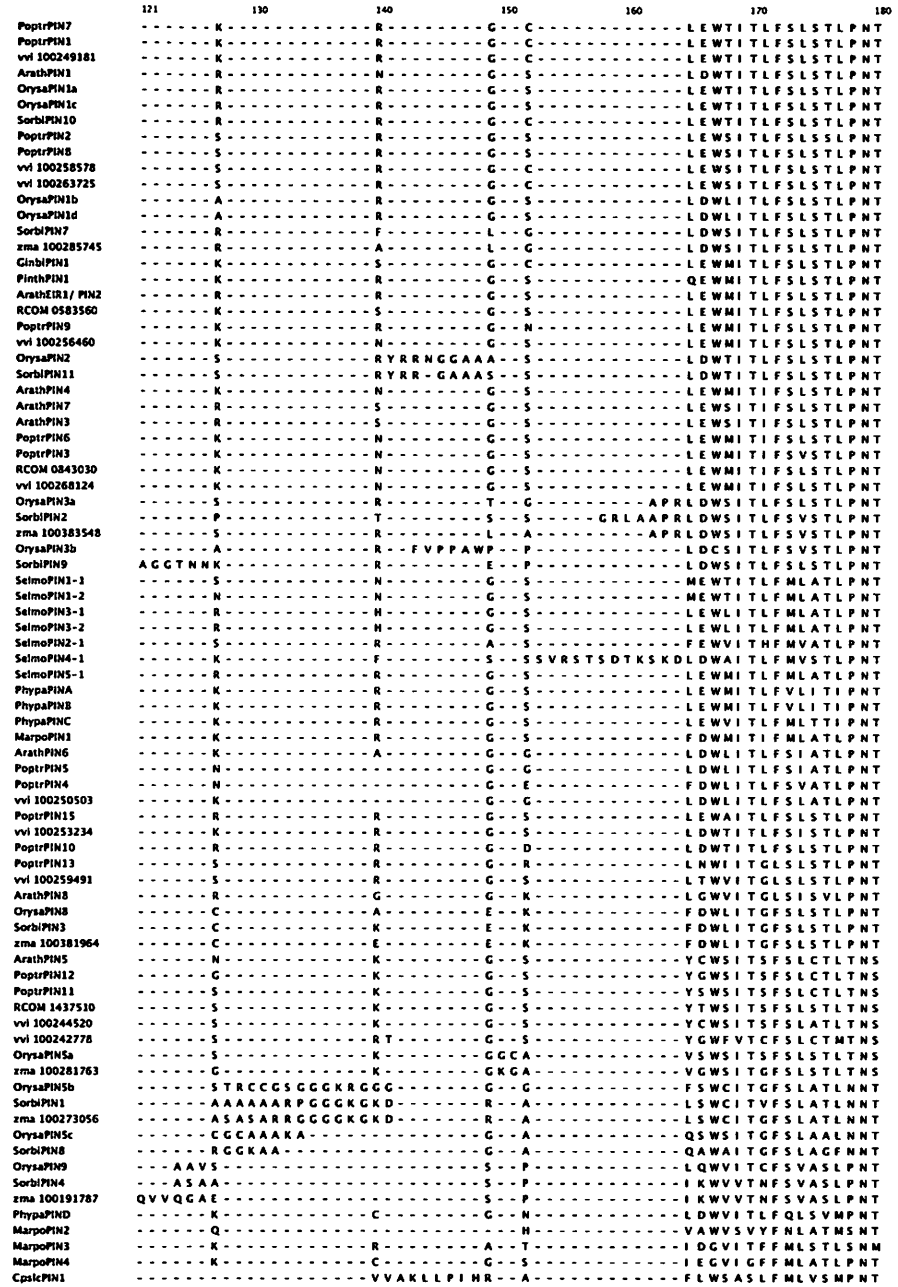
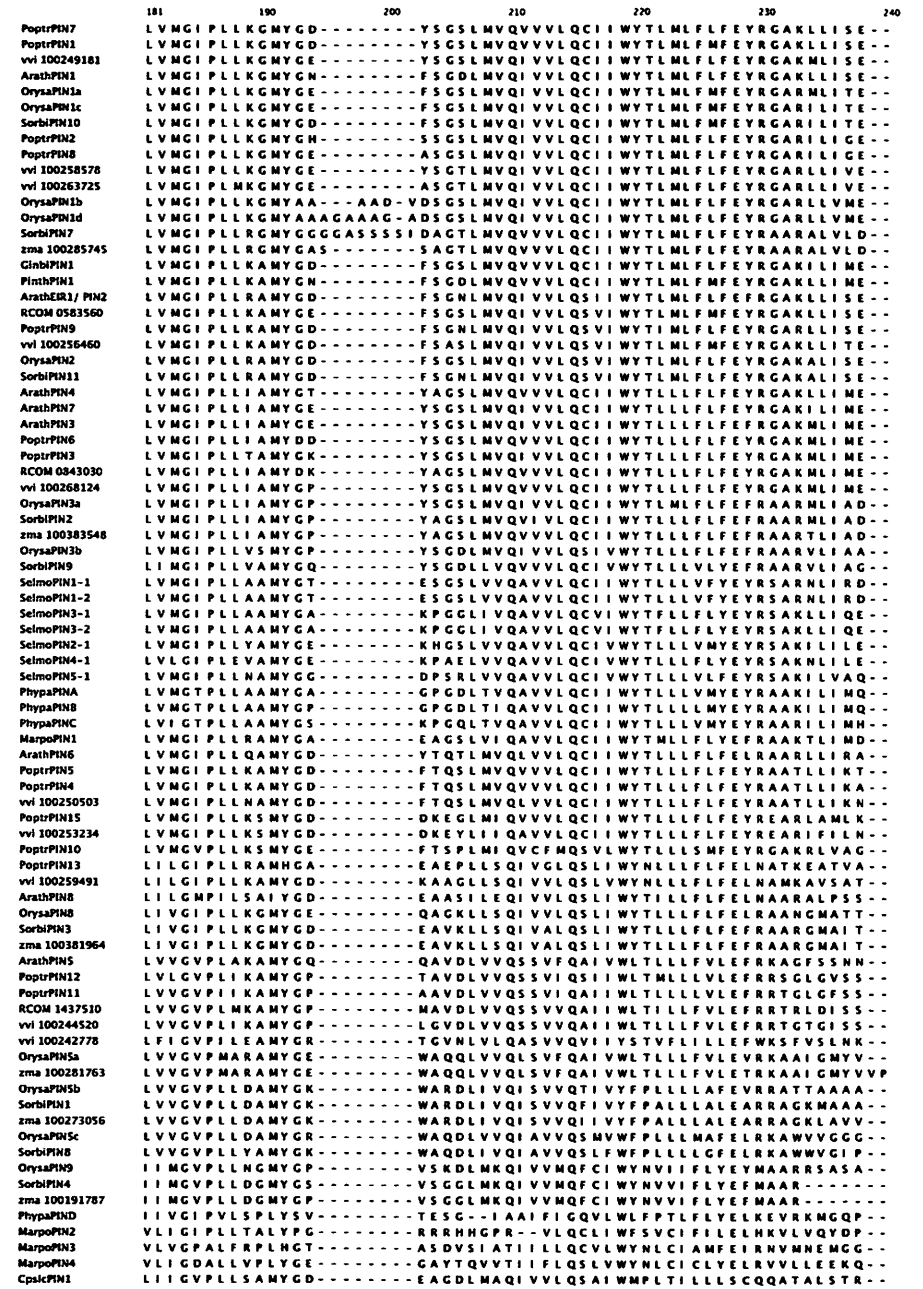
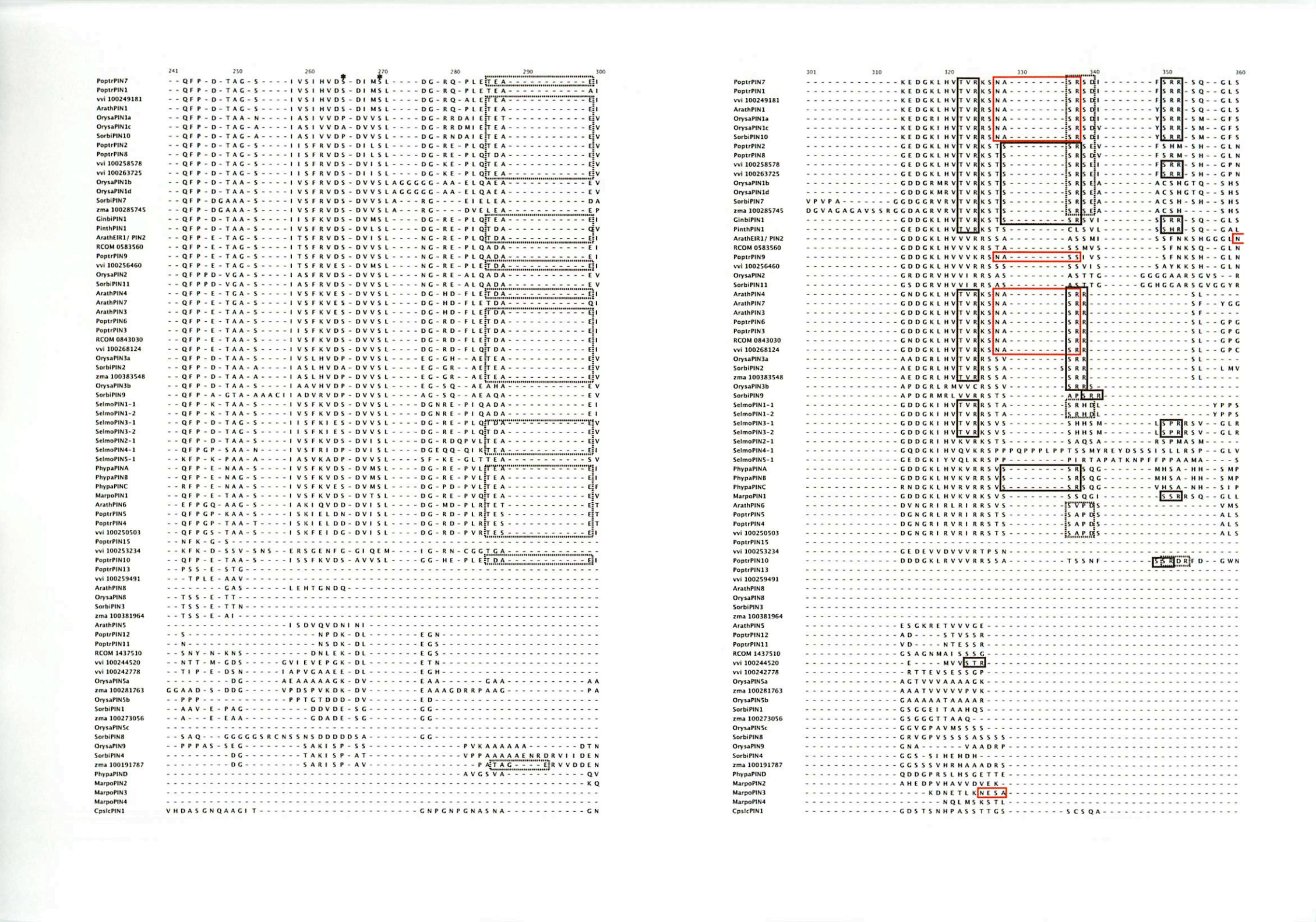
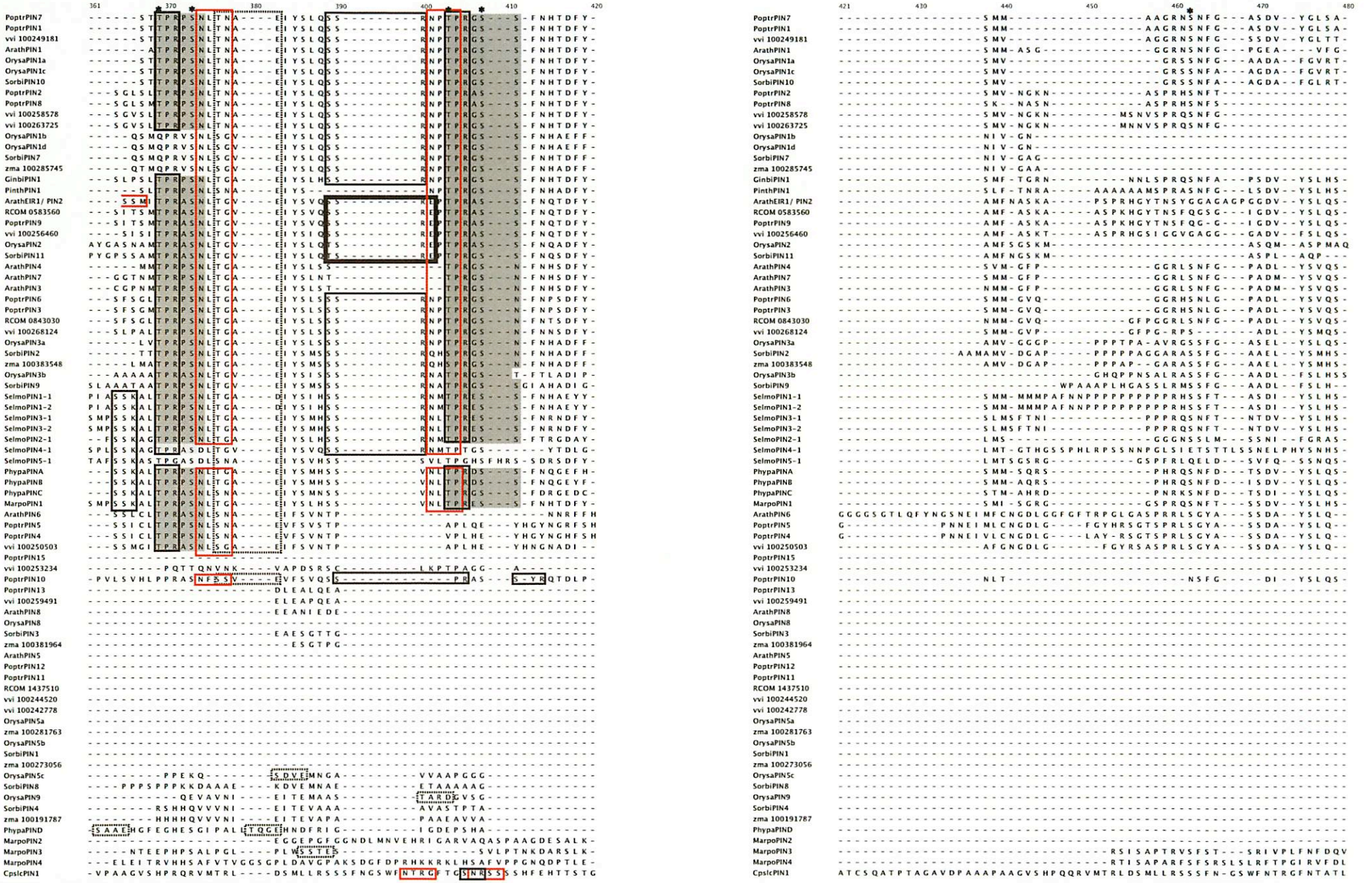

1 10 20 30 40 50 60
PoptrPIN7 M---ISLLDFYHVMTAMVPLYVAMILAYGSVKWWKIFTPDQCSGINRFVALFAVPLLSFH
PoptrPIN1 M---ISLLDFYHVMTAMVPLYVAMILAYGSVKWWKIFTPDQCSGINRFVALFAVPLLSFH
vvi 100249181 M---ISLLDFYHVMTAVVPLYVAMILAYGSVKWWKIFTPDQCSGINRFVALFAVPLLSFH
ArathPIN1 M---ITAADFYHVMTAMVPLYVAMILAYGSVKWWKIFTPDQCSGINRFVALFAVPLLSFH
OrysaPIN1a M---ITAADFYHVMTAMVPLYVAMILAYGSVKWWRIFTPDQCSGINRFVALFAVPLLSFH
OrysaPIN1c M---ITGADFYHVMTAMVPLYVAMILAYGSVKWWRIFTPDQCSGINRFVALFAVPLLSFH
SorbiPIN10 M---ITGTDFYHVMTAMVPLYVAMILAYGSVRWWRIFTPDQCSGINRFVALFAVPLLSFH
PoptrPIN2 M---ISIGDLYHVLTAVVPLYVAMILAYGSVKWWKIFSPDQCSGINRFVALFAVPLLSFH
PoptrPIN8 M---ISLTDLYHVLTAVVPLYVAMILAYGSVKWWKIFSPDQCSGINRFVALFAVPLLSFH
vvi 100258578 M---IAITDLYHVLTSVVPLYVAMILAYGSVKWWKIFSPDQCSGINRFVALFAVPLLSFH
vvi 100263725 M---ISITDLYHVLTAVVPLYVAMILAYGSVKWWKIFSPDQCSGINRFVALFAVPLLSFH
OrysaPIN1b M---ITVVDLYHVLTAVVPLYVAMTLAYASVRWWRIFSPDQCSGINRFVALFAVPLLSFH
OrysaPIN1d M---ITVVDLYHVLTAVVPLYVAMTLAYASVRWWRIFSPDQCSGINRFVALFAVPLLSFH
SorbiPIN7 M---ITALDLYHVLTAVVPLYVAMTLAYGSVRWWRIFTPDQCSGINRFVALFAVPLLSFH
zma 100285745 M---ITALDLYHVLTAVVPLYVAMTLAYGSVRWWRIFTPDQCSGINRFVALFAVPLLSFH
GinbiPIN1 M---ITGSDLYHVLTAVVPLYVAMIIAYGSVKWWKIFTPDQCSGINRFVALFAVPLLSFH
PinthPIN1 M---INGADIYNVLSAVVPLYVAMILAYGSVKWWKIFSPDQCSGINRFVALFAVPLLSFH
ArathEIR1/ PIN2 M---ITGKDMYDVLAAMVPLYVAMILAYGSVRWWGIFTPDQCSGINRFVAVFAVPLLSFH
RCOM 0583560 M---ITGKDIYDVLAAIVPLYVAMILAYGSVRWWKIFTPDQCSGINRFVAVFAVPLLSFH
PoptrPIN9 M---ITGKDIYDVLAAIVPLYVAMILAYGSVRWWKIFTPDQCSGINRFVAVFAVPLLSFH
vvi 100256460 M---IRGKDIYDVLAAIVPLYVAMILAYGSVRWWKIFTPEQCAGINRFVAVFAVPLLSFH
OrysaPIN2 M---ITGRDIYDVLAAIVPLYVAMFLAYGSVRWWGIFTPDQCSGINRFVAVFAVPLLSFH
SorbiPIN11 M---ITGRDIYDVLAAIVPLYVAMFLAYGSVRWWGIFTPDQCSGINRFVAVFAVPLLSFH
ArathPIN4 M---ITWHDLYTVLTAVVPLYVAMILAYGSVQWWKIFSPDQCSGINRFVAIFAVPLLSFH
ArathPIN7 M---ITWHDLYTVLTAVIPLYVAMILAYGSVRWWKIFSPDQCSGINRFVAIFAVPLLSFH
ArathPIN3 M---ISWHDLYTVLTAVIPLYVAMILAYGSVRWWKIFSPDQCSGINRFVAIFAVPLLSFH
PoptrPIN6 M---ISWNDLYNVLSAVIPLYVAMILAYGSVRWWKIFSPDQCSGINRFVAIFAVPLLSFH
PoptrPIN3 M---ISWNDLYNVLSAVIPLYVAMILAYGSVRWWKIFSPDQCSGINRFVAIFAVPLLSFH
RCOM 0843030 M---ITGKDLYTVLTAVVPLYVAMILAYGSVRWWKIFSPDQCSGINRFVAIFAVPLLSFH
vvi 100268124 M---ISWNDLYVVLTAVIPLYVAMILAYGSVRWWKIFSPDQCSGINRFVAIFAVPLLSFH
OrysaPIN3a M---ISGHDFYTVMAAVVPLYVAMFLAYGSVRWWGIFTPDQCSGINRFVAIFAVPLLSFH
SorbiPIN2 M---ISGNDVYTVLTAMVPLYVAMFLAYGSVRWWRIFTPDQCSGINRFVAIFAVPLLSFH
zma 100383548 M---ISWNDVYTVLTAMVPLYVAMFLAYGSVRWWRIFTADQCSGINRFVAIFAVPLLSFH
OrysaPIN3b M---ISWHELYMVLSAVVPLYVAMMVAYGSVRWWGVLTPEQCSGINRFVAVIAVPLLSFH
SorbiPIN9 M---ISWHDLYTVLCAVVPLYVAMVLAYGSVRWWGVLTPDQCAGINRFVAVFAVPLLSFH
SelmoPIN1-1 M---ISAHDFYVVMSAVVPLYVAMILAYGSVRWWKILTPDQCGGINRFVAIFAVPLLSFQ
SelmoPIN1-2 M---ISAHDFYVVMSAVVPLYVAMILAYGSVRWWKILTPDQCGGINRFVAIFAVPLLSFQ
SelmoPIN3-1 M---ISLHDLYTVLSAVVPLYVAMILAYGSVRWWGILTPEQCGGINRFVAIFAVPLLSFE
SelmoPIN3-2 M---ISLHDLYTVLSAVVPLYVAMILAYGSVRWWGILTPEQCGGINRFVAIFAVPLLSFE
SelmoPIN2-1 M---ISGGDFYQVMAAVVPLYVAMILAYGSVRWWGILTPEQCSGINRFVALFAVPLLSFQ
SelmoPIN4-1 M---ITLQEFYTVMSAVVPLYVAMILAYGSVRWWKLLTPAQCSGINRFVAIFAVPLLSFQ
SelmoPIN5-1 M---ISPAEFYSVMAAMVPLYVAMLLAYASVRWWGILTPDQCSGINRFVAIFAVPLLSFQ
PhypaPINA M---INGHDIYNVLSAMVPLYVAMMLAYGSVKWWGILTPQQCGGINRFVSIFAVPLLSFQ
PhypaPINB M---INGHDIYNVLSAMVPLYVAMMLAYGSVKWWGILTPQQCGGINRFVSIFAVPLLSFQ
PhypaPINC M---ITGHDMYNVLSAMVPLYVAMMLAYASVKWWGILTPQQCDGINRFVSIFAVPLLSFQ
MarpoPIN1 M---INAHDLYNVLSAVIPLYVAMILAYGSVKWWGILTPDQCGGINRFVAVFAVPLLSFK
ArathPIN6 M---ITGNEFYTVMCAMAPLYFAMFVAYGSVKWCKIFTPAQCSGINRFVSVFAVPVLSFH
PoptrPIN5 M---ITAGDFYKVMCAMVPLYFAMLVAYGSVKRYKIFTPEQCSGINRFVAVFAVPVLSFH
PoptrPIN4 M---ITADDFYKVMCAMVPLYFAMLVAYGSVKWYKIFTPEQCSGINRFVAVFAVPVLSFH
vvi 100250503 M---ISADDFYKVMCAMVPLYFAMLVAYGSVKWCKIFSPEQCSGINRFVAVFAVPVLSFH
PoptrPIN15 M---ISIEDLYGVLCAVVPLYVTMFLAYASVKWWNIFTPEQCSGINRFVAYFAVPLLSME
vvi 100253234 M---ISIKDLYGVLSAVVPLYVTMFLAYASVKWWNVFSPDQCAGINRFVAIFAIPLLSFE
PoptrPIN10 M---ISGKDIYQVVSALVPLYAAMILAYGSVRWWKVFTPDQCAGINRFVAVFATPFLVFD
PoptrPIN13 M---ISAADVYHVVTATVPLYFAMILAYISVKWWKLFTPDQCAGINKFVAKFSIPLLSFQ
vvi 100259491 M---ISLADVYHVVAATIPLYFVMILAYISVKWWKIFTPEQCSGINKFVAKFSIPLLSFQ
ArathPIN8 M---ISWLDIYHVVSATVPLYVSMTLGFLSARHLKLFSPEQCAGINKFVAKFSIPLLSFQ
OrysaPIN8 M---VSWKDIYLVLEATVPLYVAMILAYLSIKWWKLFTPEQCSGINKFVAKFSIPLLSFQ
SorbiPIN3 M---ISWPTIYHVLEETVPLYVAMIVAYLSIQWWKLFTPEQCSGINKFVAKFSIPLLSFQ
zma 100381964 M---ISWPTIYHVLEETVPLYVAMIVAYLSIQWWKLFTPEQCSGINKFVAKFSIPLLSFQ
ArathPIN5 M---INCGDVYKVIEAMVPLYVALILGYGSVKWWHIFTRDQCDAINRLVCYFTLPLFTIE
PoptrPIN12 M---IGWADIYKVVVAMVPLYVALMLGYGSVRWWKVFTPEQCGAINRFVCYFTLPFFTFE
PoptrPIN11 M---IGWEDVYKVVVAMVPLYVALVLGYGSVRWWKVFTPEQCGAINRFVCYFTLPLFTFE
RCOM 1437510 M---IGWEDVYKVIVAMVPLYVALLLGYGSIRWWKIFTPEQGGAVNRLVCYFTLPLFTFE
vvi 100244520 M---IGWEDVYKVVVAMLPLYVALVLGYGSVKWWRIFTPDQCDAINRLVCYFTLPLFTFE
vvi 100242778 M---IEWGDVYKVVEAMAPLYVALLLGYASVRWWHMFSSDQCAAINRFICYFIFPFYSFD
OrysaPIN5a M---IGWGDVYKVVAATVPLYFALFLGYGSVRWWRIFTREQCDAVNRLVAFFALPFFTFE
zma 100281763 M---IGWGDVYKVVAATVPLYFALFLGYGSVRWWRIFTREQCDAVNRLVAFFALPFFTFE
OrysaPIN5b M---IGWGDVYKVVAAMAPLYFALGLGYGSVRWWRLFTADQCDAVNRLVACFAVPFFAFD
SorbiPIN1 M-M-IGWGDVYKVVAAMAPLYFALGLGYGSVRWWKMFTPDQCDAINRLVSYFAVPFFAFD
zma 100273056 M---IGWGDVYKVVAAMAPLYFALGLGYGSVRWWKLFSPDQCDAVNRLVAYFAVPFFAFD
OrysaPIN5c M---IGWGDVYKVVGAMAPLYFALGLGYGSVRWWRFFTAEQCAAINTMVVYFSMPFFTFD
SorbiPIN8 M---IPWGDIYKVVAAMAPLYFALGLGYGSVRWWRFFTAEQCAAINTLVVNFSMPFFTFD
OrysaPIN9 M---ITGSEVYQVVEAMAPLYTAAALGYGSVRWLKAFSNEQCAGINHFVALYAVPVLIFD
SorbiPIN4 M---IPGSAVYHVVEAMAPLYTAAVLGYASVRWLKAFSDEQCAGINHFVALYAVPVLIFH
zma 100191787 M---IPGSAVYHVVEAMAPLYTSAVLGYASVRWLKAFSDEQCAGINHFVALYAVPVLIFH
PhypaPIND M---LTGGQFYDVLCAVVPLYVALFLGYGSLKWWGVVTPEQSAGISRFNALIAMPPLIFE
MarpoPIN2 M--GLKGSEVMNLVNAVVPLYISIGLGYTLSKLKVLSYEVHYQGIVTFCFKVTTPPLVFQ
MarpoPIN3 MAA-MTGEDFWLLLQYILPFFIILLPAYGMAR-TKILTQDHALGLNRWTSLVGVPIYVGH
MarpoPIN4 M---LTALDFWHLVEHIIPVYVLVIVGYVLTR-FEIIEPVHSPGISKWVSTLGVPMYVFH
CpsIcPIN1 M---IDGGDVYRVFCAIVPLYVGIGSGYLSVRVFKILSPDQCAGVNRYIAFIALPALVFQ
61 70 80 90 100 110 120
PoptrPIN7 FISTNDPYKMNFRFIAADTLQKIIVLVVLAF---WT----MF------S-----------
PoptrPIN1 FISTNDPYNMNLRFITADSLQKIIVLVVLAL---WT----KL------S-----------
vvi 100249181 FISTNDPYAMNFRFIAADTLQKIIVLAVLAV---WS----NV------S-----------
ArathPIN1 FIAANNPYAMNLRFLAADSLQKVIVLSLLFL---WC----KL------S-----------
OrysaPIN1a FISTNNPYTMNLRFIAADTLQKLMVLAMLTA---WS----HL------S-----------
OrysaPIN1c FISTNNPYTMNLRFIAADTLQKLIVLALLTL---WS----HL------S-----------
SorbiPIN10 FISTNNPYTMNLRFIAADTLQKLIVLALLTA---WS----YL------S-----------
PoptrPIN2 FISTNNPYAMNLRFIAADTLQKIIVLVVLAI---WS----RA------S-----------
PoptrPIN8 FISTNNPYAMNYRFIAADTLQKIIVLVVLAI---WT----RV------I-----------
vvi 100258578 FISTNDPYAMNYRFIAADTLQKIIVLVVLAI---WS----RV------S-----------
vvi 100263725 FISTNNPYAMNYRFIAADTLQKIIVLVVLAI---WS----RT------S-----------
OrysaPIN1b FISTNNPFAMNLRFLAADTLQKLIVLALLAL---WC----RL------S-----------
OrysaPIN1d FISTNNPFAMNLRFLAADTLQKLIVLALLAL---WC----RL------S-----------
SorbiPIN7 FISTNNPFAMNLRFLAADTLQKVAVLALLFL---AS----RLFSSSAS-----------
zma 100285745 FISTNDPFAMNLRFLAADTLQKVAVLALLAL---AS----RGLSS--P-----------
GinbiPIN1 FISTNDPYAMNFRFIAADSLQKIIVLVVLGI---WT----KV------G-----------
PinthPIN1 FISTNNPYAMNLRFIAADSLQKIIILAMLVI---WA----KV------G-----------
ArathEIR1/ PIN2 FISSNDPYAMNYHFLAADSLQKVVILAALFL---WQ----AF------S-----------
RCOM 0583560 FISSNDPYAMNFRFIAADSLQKVVILAALFL---WQ----AF------S-----------
PoptrPIN9 FISSNDPYAMNFRFIAADTLQKVVILGALFI---WQ----AF------S-----------
vvi 100256460 FISLNDIYNMNYKFIAADSLQKIVILSALFL---WQ----AF------A-----------
OrysaPIN2 FISTNDPYSMNYRFLAADSLQKLVILAALAV---WH----NL------L-----------
SorbiPIN11 FISSCDPYAMQYRFLAADSLQKLVILAALAV---WH----NV------L-----------
ArathPIN4 FISTNDPYAMNFRFVAADTLQKIIMLVLLAL---WA----NL------T-----------
ArathPIN7 FISSNNPYAMNLRFIAADTLQKLIMLTLLII---WA----NF------T-----------
ArathPIN3 FISTNNPYAMNLRFIAADTLQKIIMLSLLVL---WA----NF------T-----------
PoptrPIN6 FISTNDPYSMNFRFIAADTLQKIIMLFALGI---WT----NF------T-----------
PoptrPIN3 FISTNDPYAMNFRFIAADTLQKIIMLIALGI---WT----NF------T-----------
RCOM 0843030 FISTNNPYEMNLRFIAADTLQKIIMLVALGL---WT----NL------T-----------
vvi 100268124 FISQNNPYAMNFRFIAADTLQKIIMLVVLGI---WT----NF------T-----------
OrysaPIN3a FISTNDPYAMNLRFLAADTLQKLLVLAGLAA---WS----RL------P-----------
SorbiPIN2 FISTNNPYAMNLRFLAADTLQKLLVLAGLAV---WS----RL------L-----------
zma 100383548 FISTNDPYAMNLRFLAADTLQKLVVLAGLAA---CS----RL------P-----------
OrysaPIN3b FISSSDPYAMNLRFVAADTLQKVLVLAALAA---WS----RF------P-----------
SorbiPIN9 CISTSDPYVMNLRFIAADTLQKVVLVLAALAV---WS----YL------PTARRGGAAADAA
SelmoPIN1-1 ILATNNPYKMNAQFIAADVLQKLAILAVLLV---WC----RY------S-----------
SelmoPIN1-2 ILATNDPYKMNAQFIAADVLQKLAILAVLLV---WC----RY------S-----------
SelmoPIN3-1 IVSSNNPYAMNLKFIAADIVQKLAILAALAL---WV----KL------S-----------
SelmoPIN3-2 IVSSNNPYAMNLKFIAADIVQKLAILAALAL---WV----KL------S-----------
SelmoPIN2-1 IISKNNPYLMNPQFIAADAIQKAAVMLVLAI---WA----RY------S-----------
SelmoPIN4-1 IIAHNNPYDMNPRFVAADVLQKLIVLAVLAS---WS----RL------K-----------
SelmoPIN5-1 IIAHNNPYDMSLKFVLADALQKIIVLAVLGA---WA----RY------S-----------
PhypaPINA FISGNNPYAMNFKFIAADAVSKVLVLLCLGL---WA----RY------A-----------
PhypaPINB FISGNNPYAMNFRFIAADAVSKVFVLLCLGL---WA----RY------S-----------
PhypaPINC FVSGNNPYEMNFRFIAADAVSKVFVLSCLGL---WV----RF------S-----------
MarpoPIN1 FIALNDPYTMNYQFIAADAMSKIAILILLGL---WA----RY------T-----------
ArathPIN6 FISQNNPYKMDTMFILADTLSKIFVFVLLSL---WA----VF------F-----------
PoptrPIN5 FIAQNNPYQMDTRFIMADTLSKVLVLVLLSI---WV----IF------F-----------
PoptrPIN4 FIAQNNPYQMDTKFILADTLSKVLALVLLSV---WA----VF------F-----------
vvi 100250503 FISQNNPYEMDTKFIVADTLSKLLVLVLLSV---WA----IL------F-----------
PoptrPIN15 FISRINPYKMDLLFMAADGVSKVLILVVLLC---WA----NF------S-----------
vvi 100253234 VISRINPYKMDFLFIAADGVSKVLILIILFS---WA----KF------S-----------
PoptrPIN10 FICSNNPYKMNLRFIAADSLQKVVVLVVLFI---WK----AT------A-----------
PoptrPIN13 VISGINPYKMNLKLIFADFLQKLLALLVLTA---LA----KI------S-----------
vvi 100259491 LISENNPYQMNMRLILSDFLQKLLALILLAT---IA----KI------N-----------
ArathPIN8 IISENNPFKMSPKLILSDILQKFLVVVVLAM---VL----RFWHPTGG-----------
OrysaPIN8 VISTTDPYDMNIKLIYSDILQKSLALLGFAA---IS----KA------C-----------
SorbiPIN3 ILSTNNPYNMNLKLIFSDILQKSLSLLGFAV---IS----KA------C-----------
zma 100381964 ILSTNNPYSMNLKLIFSDVLQKSLSLLGFAV---AS----RV------C-----------
ArathPIN5 FTAHVDPFNMNYRFIAADVLSKVIIVTVLAL---WA----KY------S-----------
PoptrPIN12 FTAHVDPFKMNYLFIGADAISKLIIVVVLAV---WA----KW------S-----------
PoptrPIN11 FTAHVDPFKMNYRFIGADAVSKVIIVVVLAF---WA----KC------S-----------
RCOM 1437510 FTAHVDPFKMNYLFIGADAISKFIIVVVLAC---WA----KF------S-----------
vvi 100244520 FTAHVDPFKMNYLFIGADAVSKLIIVLVLAL---WA----KC------S-----------
vvi 100242778 FTSHIDPFSMNYKYMGADIMSKFICVAALGL---WC----KC------S-----------
OrysaPIN5a FTLHTDPFQVNYRAVAADVISKAVIVAVIGA---WA----RF------M-----------
zma 100281763 FTLHTDPFQVNYRAVAADVISKAVIVGVIGC---WAGSRVSS------A-----------
OrysaPIN5b FAARIDPFALSYRVLAADALSKLAVALALAA---CA----AA------A-----------
SorbiPIN1 FASRIDPFSLSYRVLAADALSKLAVALALAA---WA----TA------A-----------
zma 100273056 FAARIDPFALSYRVLAADALSKLAVALALAA---WA----TA------A-----------
OrysaPIN5c FVVRTDPFAMNYRVIAADAVSKAIAIAAMAA---WA----RT------R-----------
SorbiPIN8 FLARANPYTMNYRVLAADAVSKALAIAAVSA---WG----CC------C-----------
OrysaPIN9 MVSTNNVYKMNGRLIAADTLQKAVLLLGLMA---WA----LW------ERSRARGAGAKAK
SorbiPIN4 MVSTNDPYHMNERLIAADTLQKAVMLLALTA---WA----FW------SHFRRRRDGKASS
zma 100191787 MVSTNDPYHMNERLIAADTLQKAVVLLALTA---WA----FW------SHLSRRRGGRGKR
PhypaPIND IIAFNNPYTMNNRLIAAYCLSNGIVLVGLCA---WV----WC------T-----------
MarpoPIN2 MVANNNPYLLNLRLIGADVLCKAIVLVVLGLSCVWR----PT------A-----------
MarpoPIN3 MVATNDPYQTNLRLVGADALQKIVVLLIALV---WW----RF------S-----------
MarpoPIN4 LLAFNDPYQTSMRLVGADVLSKALVLIGGCL---WV----RF------C-----------
CpsIcPIN1 TIAKTNMFQMNFQFLGADSTAKLFVLVAVML---FL----LL------R-----------

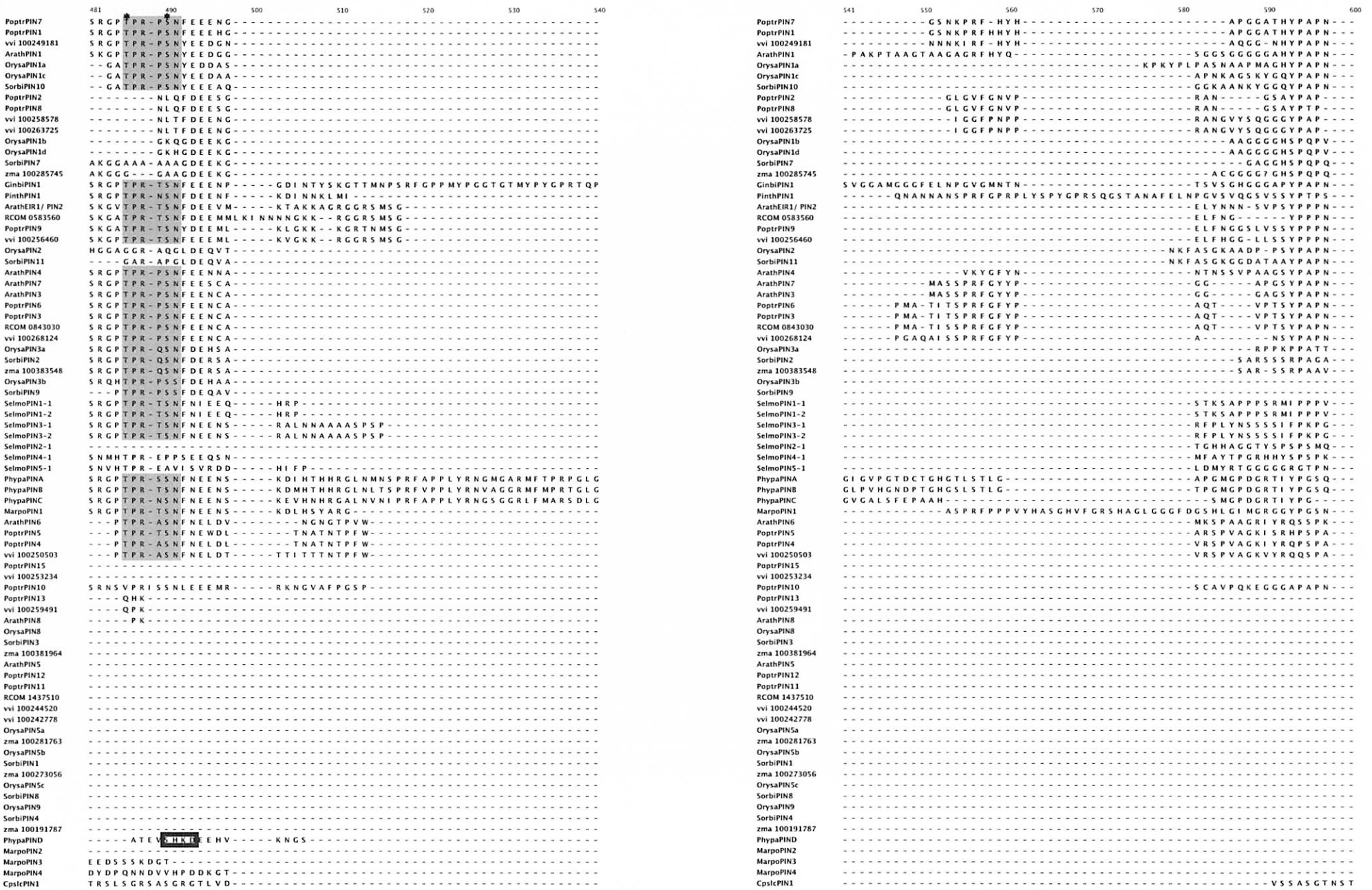
Figure 11

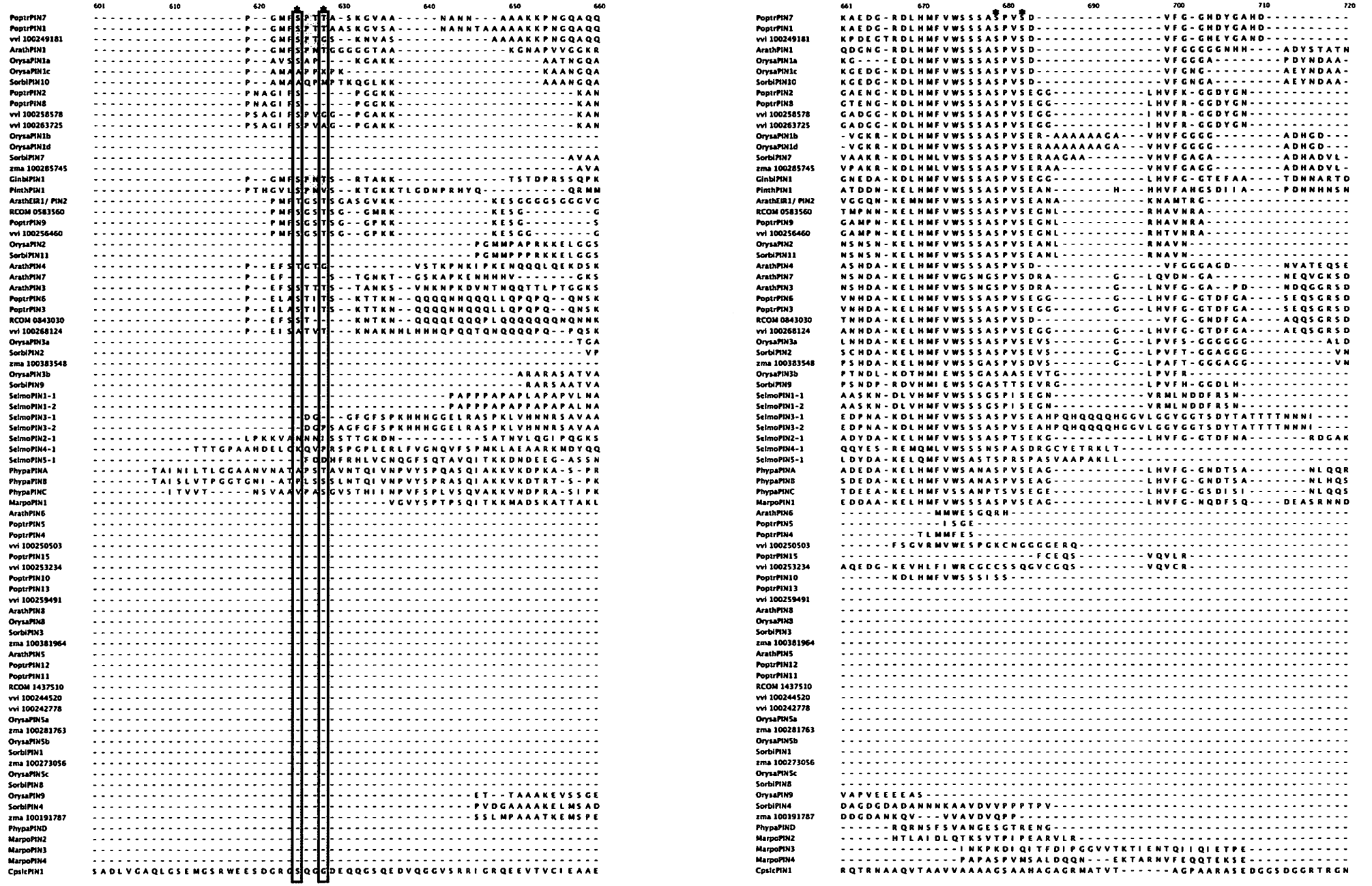

121 130 140 150 160 170 180
PoptrPIN7
PoptrPIN1
vvi 100249181
ArathPIN1
OrysaPIN1a
OrysaPIN1c
SorbiPIN10
PoptrPIN2
PoptrPIN8
vvi 100258578
vvi 100263725
OrysaPIN1b
OrysaPIN1d
SorbiPIN7
zma 100285745
GinbiPIN1
PinthPIN1
ArathEIR1/ PIN2
RCOM 0583560
PoptrPIN9
vvi 100256460
OrysaPIN2
SorbiPIN11
ArathPIN4
ArathPIN7
ArathPIN3
PoptrPIN6
PoptrPIN3
RCOM 0843030
vvi 100268124
OrysaPIN3a
SorbiPIN2
zma 100383548
OrysaPIN3b
SorbiPIN9
SelmoPIN1-1
SelmoPIN1-2
SelmoPIN3-1
SelmoPIN3-2
SelmoPIN2-1
SelmoPIN4-1
SelmoPIN5-1
PhypaPINA
PhypaPINB
PhypaPINC
MarpoPIN1
ArathPIN6
PoptrPIN5
PoptrPIN4
vvi 100250503
PoptrPIN15
vvi 100253234
PoptrPIN10
PoptrPIN13
vvi 100259491
ArathPIN8
OrysaPIN8
SorbiPIN3
zma 100381964
ArathPIN5
PoptrPIN12
PoptrPIN11
RCOM 1437510
vvi 100244520
vvi 100242778
OrysaPIN5a
zma 100281763
OrysaPIN5b
SorbiPIN1
zma 100273056
OrysaPIN5c
SorbiPIN8
OrysaPIN9
SorbiPIN4
zma 100191787
PhypaPIND
MarpoPIN2
MarpoPIN3
MarpoPIN4
CpsIcPIN1

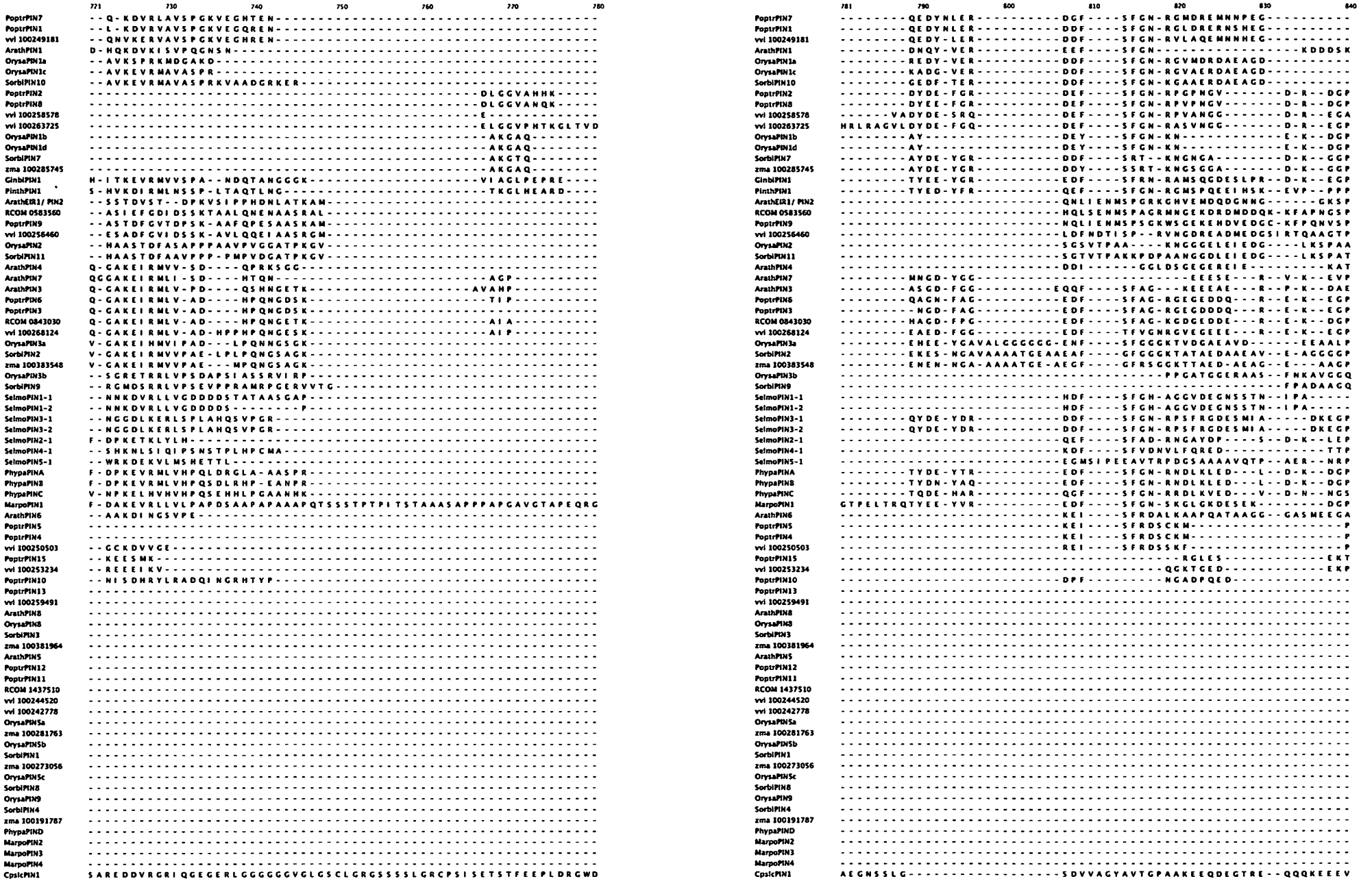

181 190 200 210 220 230 240
PoptrPIN7
PoptrPIN1
vvi 100249181
ArathPIN1
OrysaPIN1a
OrysaPIN1c
SorbiPIN10
PoptrPIN2
PoptrPIN8
vvi 100258578
vvi 100263725
OrysaPIN1b
OrysaPIN1d
SorbiPIN7
zma 100285745
GinbiPIN1
PinthPIN1
ArathEIR1/ PIN2
RCOM 0583560
PoptrPIN9
vvi 100256460
OrysaPIN2
SorbiPIN11
ArathPIN4
ArathPIN7
ArathPIN3
PoptrPIN6
PoptrPIN3
RCOM 0843030
vvi 100268124
OrysaPIN3a
SorbiPIN2
zma 100383548
OrysaPIN3b
SorbiPIN9
SelmoPIN1-1
SelmoPIN1-2
SelmoPIN3-1
SelmoPIN3-2
SelmoPIN2-1
SelmoPIN4-1
SelmoPIN5-1
PhypaPINA
PhypaPINB
PhypaPINC
MarpoPIN1
ArathPIN6
PoptrPIN5
PoptrPIN4
vvi 100250503
PoptrPIN15
vvi 100253234
PoptrPIN10
PoptrPIN13
vvi 100259491
ArathPIN8
OrysaPIN8
SorbiPIN3
zma 100381964
ArathPIN5
PoptrPIN12
PoptrPIN11
RCOM 1437510
vvi 100244520
vvi 100242778
OrysaPIN5a
zma 100281763
OrysaPIN5b
SorbiPIN1
zma 100273056
OrysaPIN5c
SorbiPIN8
OrysaPIN9
SorbiPIN4
zma 100191787
PhypaPIND
MarpoPIN2
MarpoPIN3
MarpoPIN4
CpsIcPIN1

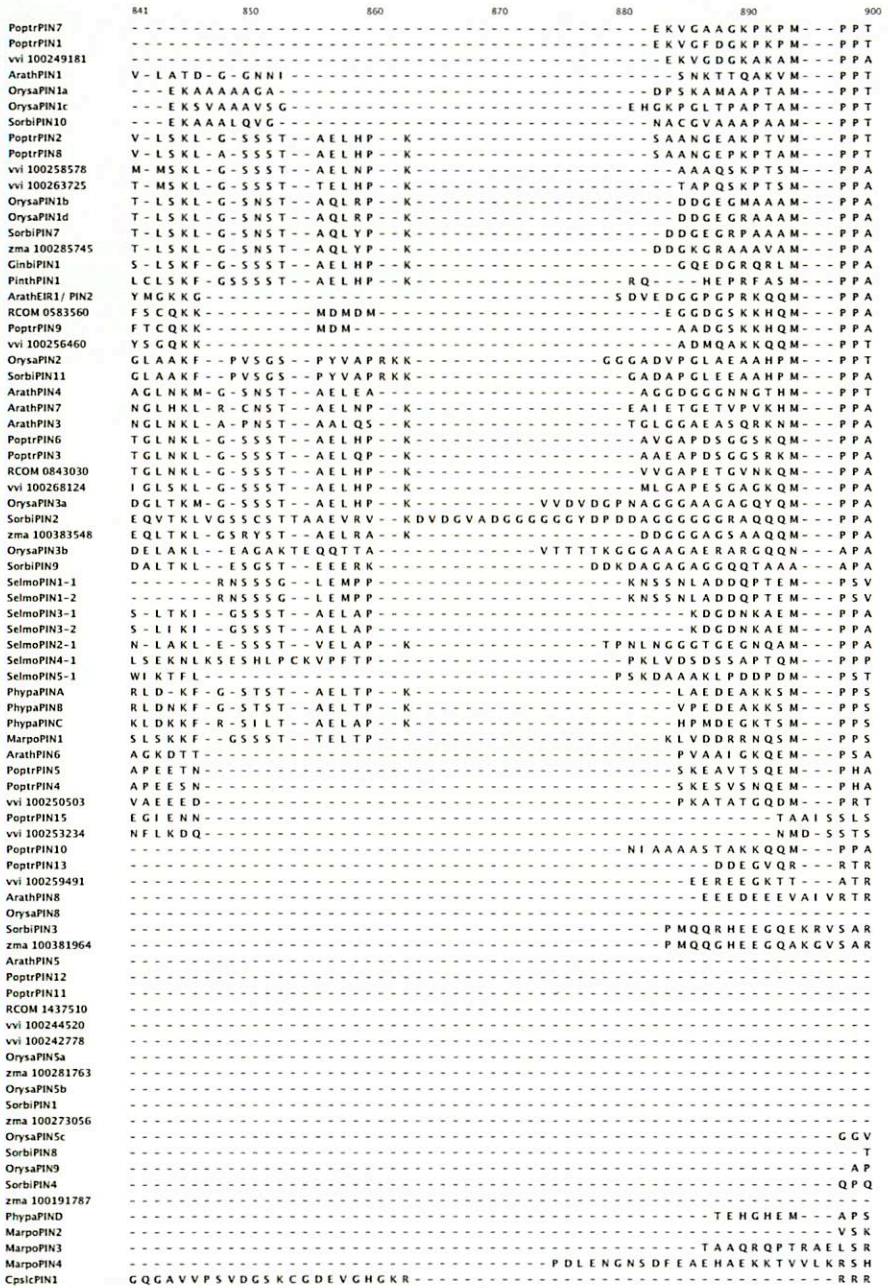

241 250 260 270 280 290 300
PoptrPIN7
PoptrPIN1
vvi 100249181
ArathPIN1
OrysaPIN1a
OrysaPIN1c
SorbiPIN10
PoptrPIN2
PoptrPIN8
vvi 100258578
vvi 100263725
OrysaPIN1b
OrysaPIN1d
SorbiPIN7
zma 100285745
GinbiPIN1
PinthPIN1
ArathEIR1/ PIN2
RCOM 0583560
PoptrPIN9
vvi 100256460
OrysaPIN2
SorbiPIN11
ArathPIN4
ArathPIN7
ArathPIN3
PoptrPIN6
PoptrPIN3
RCOM 0843030
vvi 100268124
OrysaPIN3a
SorbiPIN2
zma 100383548
OrysaPIN3b
SorbiPIN9
SelmoPIN1-1
SelmoPIN1-2
SelmoPIN3-1
SelmoPIN3-2
SelmoPIN2-1
SelmoPIN4-1
SelmoPIN5-1
PhypaPINA
PhypaPINB
PhypaPINC
MarpoPIN1
ArathPIN6
PoptrPIN5
PoptrPIN4
vvi 100250503
PoptrPIN15
vvi 100253234
PoptrPIN10
PoptrPIN13
vvi 100259491
ArathPIN8
OrysaPIN8
SorbiPIN3
zma 100381964
ArathPIN5
PoptrPIN12
PoptrPIN11
RCOM 1437510
vvi 100244520
vvi 100242778
OrysaPIN5a
zma 100281763
OrysaPIN5b
SorbiPIN1
zma 100273056
OrysaPIN5c
SorbiPIN8
OrysaPIN9
SorbiPIN4
zma 100191787
PhypaPIND
MarpoPIN2
MarpoPIN3
MarpoPIN4
CpslcPIN1
301 310 320 330 340 350 360
PoptrPIN7
PoptrPIN1
vvi 100249181
ArathPIN1
OrysaPIN1a
OrysaPIN1c
SorbiPIN10
PoptrPIN2
PoptrPIN8
vvi 100258578
vvi 100263725
OrysaPIN1b
OrysaPIN1d
SorbiPIN7
zma 100285745
GinbiPIN1
PinthPIN1
ArathEIR1/ PIN2
RCOM 0583560
PoptrPIN9
vvi 100256460
OrysaPIN2
SorbiPIN11
ArathPIN4
ArathPIN7
ArathPIN3
PoptrPIN6
PoptrPIN3
RCOM 0843030
vvi 100268124
OrysaPIN3a
SorbiPIN2
zma 100383548
OrysaPIN3b
SorbiPIN9
SelmoPIN1-1
SelmoPIN1-2
SelmoPIN3-1
SelmoPIN3-2
SelmoPIN2-1
SelmoPIN4-1
SelmoPIN5-1
PhypaPINA
PhypaPINB
PhypaPINC
MarpoPIN1
ArathPIN6
PoptrPIN5
PoptrPIN4
vvi 100250503
PoptrPIN15
vvi 100253234
PoptrPIN10
PoptrPIN13
vvi 100259491
ArathPIN8
OrysaPIN8
SorbiPIN3
zma 100381964
ArathPIN5
PoptrPIN12
PoptrPIN11
RCOM 1437510
vvi 100244520
vvi 100242778
OrysaPIN5a
zma 100281763
OrysaPIN5b
SorbiPIN1
zma 100273056
OrysaPIN5c
SorbiPIN8
OrysaPIN9
SorbiPIN4
zma 100191787
PhypaPIND
MarpoPIN2
MarpoPIN3
MarpoPIN4
CpslcPIN1

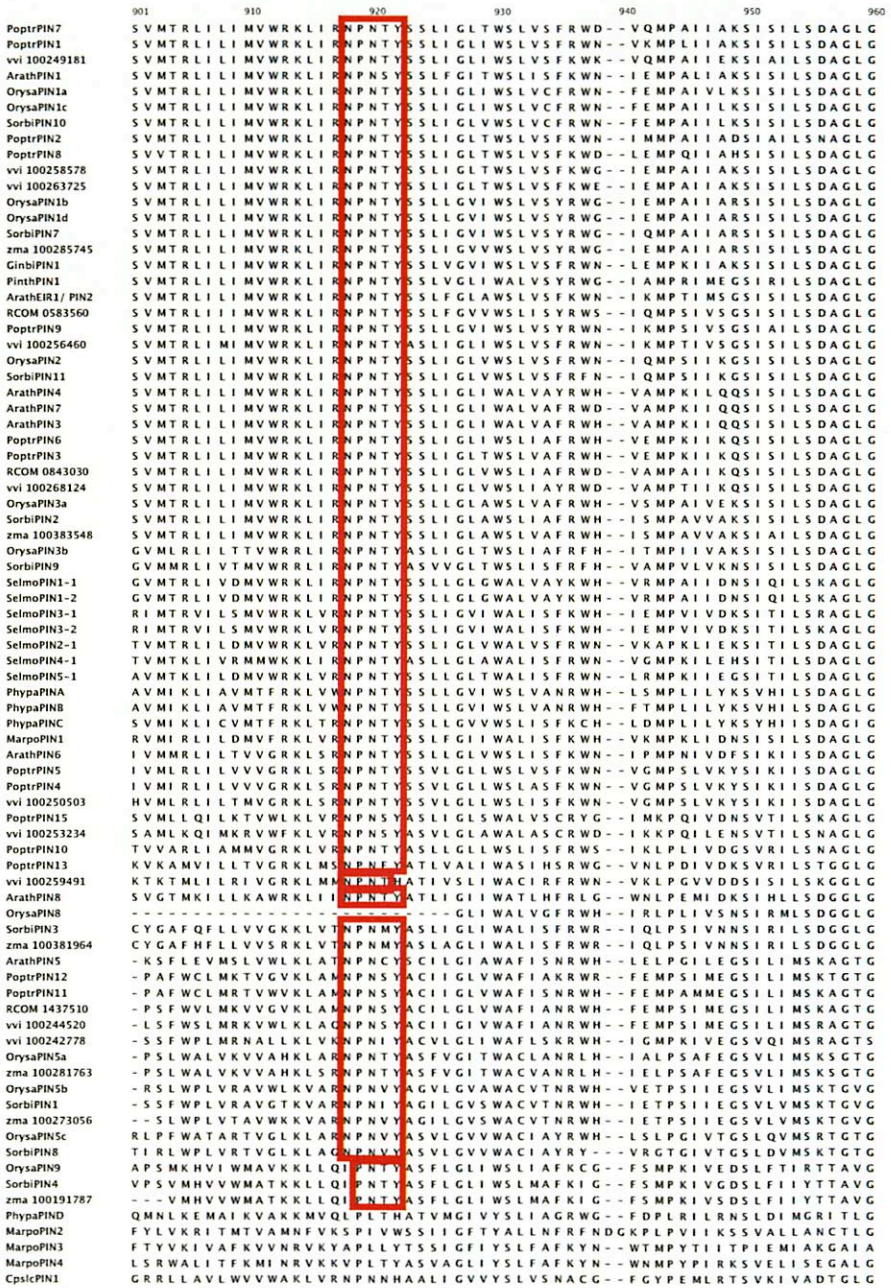


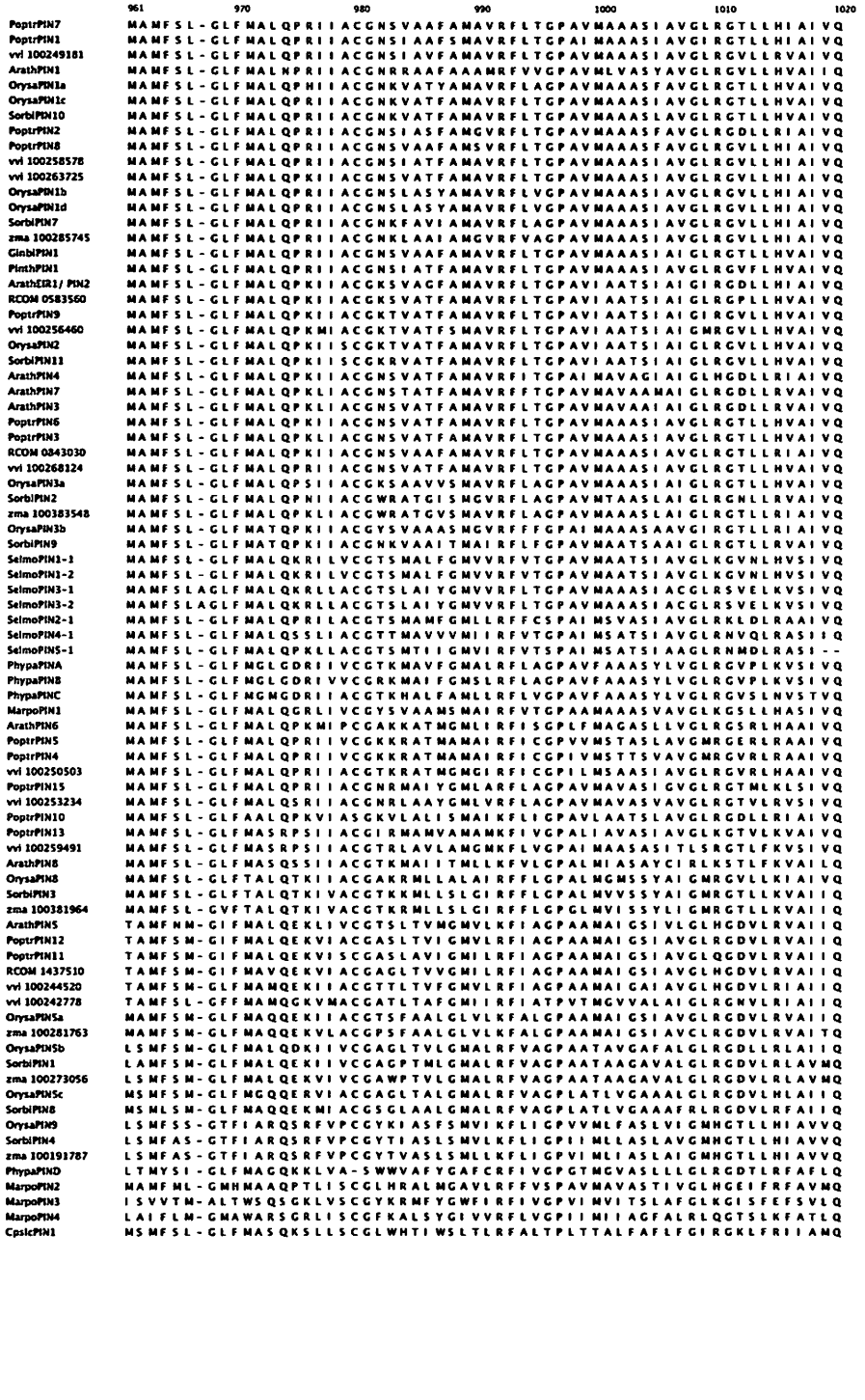

481
490
500
510
520
530
540
541
550
560
570
580
590
600
PoptrPIN7
PoptrPIN1
vvi 100249181
ArathPIN1
OrysaPIN1a
OrysaPIN1c
SorbiPIN10
PoptrPIN2
PoptrPIN8
vvi 100258578
vvi 100263725
OrysaPIN1b
OrysaPIN1d
SorbiPIN7
zma 100285745
GinbiPIN1
PinthPIN1
ArathEIR1/ PIN2
RCOM 0583560
PoptrPIN9
vvi 100256460
OrysaPIN2
SorbiPIN11
ArathPIN4
ArathPIN7
ArathPIN3
PoptrPIN6
PoptrPIN3
RCOM 0843030
vvi 100268124
OrysaPIN3a
SorbiPIN2
zma 100383548
OrysaPIN3b
SorbiPIN9
SelmoPIN1-1
SelmoPIN1-2
SelmoPIN3-1
SelmoPIN3-2
SelmoPIN2-1
SelmoPIN4-1
SelmoPIN5-1
PhypaPINA
PhypaPINB
PhypaPINC
MarpoPIN1
ArathPIN6
PoptrPIN5
PoptrPIN4
vvi 100250503
PoptrPIN15
vvi 100253234
PoptrPIN10
PoptrPIN13
vvi 100259491
ArathPIN8
OrysaPIN8
SorbiPIN3
zma 100381964
ArathPIN5
PoptrPIN12
PoptrPIN11
RCOM 1437510
vvi 100244520
vvi 100242778
OrysaPIN5a
zma 100281763
OrysaPIN5b
SorbiPIN1
zma 100273056
OrysaPIN5c
SorbiPIN8
OrysaPIN9
SorbiPIN4
zma 100191787
PhypaPIND
MarpoPIN2
MarpoPIN3
MarpoPIN4
CpsIcPIN1


PoptrPIN7
PoptrPIN1
vvi 100249181
ArathPIN1
OrysaPIN1a
OrysaPIN1c
SorbiPIN10
PoptrPIN2
PoptrPIN8
vvi 100258578
vvi 100263725
OrysaPIN1b
OrysaPIN1d
SorbiPIN7
zma 100285745
GinbiPIN1
PinthPIN1
ArathEIR1/ PIN2
RCOM 0583560
PoptrPIN9
vvi 100256460
OrysaPIN2
SorbiPIN11
ArathPIN4
ArathPIN7
ArathPIN3
PoptrPIN6
PoptrPIN3
RCOM 0843030
vvi 100268124
OrysaPIN3a
SorbiPIN2
zma 100383548
OrysaPIN3b
SorbiPIN9
SelmoPIN1-1
SelmoPIN1-2
SelmoPIN3-1
SelmoPIN3-2
SelmoPIN2-1
SelmoPIN4-1
SelmoPIN5-1
PhypaPINA
PhypaPINB
PhypaPINC
MarpoPIN1
ArathPIN6
PoptrPIN5
PoptrPIN4
vvi 100250503
PoptrPIN15
vvi 100253234
PoptrPIN10
PoptrPIN13
vvi 100259491
ArathPIN8
OrysaPIN8
SorbiPIN3
zma 100381964
ArathPIN5
PoptrPIN12
PoptrPIN11
RCOM 1437510
vvi 100244520
vvi 100242778
OrysaPIN5a
zma 100281763
OrysaPIN5b
SorbiPIN1
zma 100273056
OrysaPIN5c
SorbiPIN8
OrysaPIN9
SorbiPIN4
zma 100191787
PhypaPIND
MarpoPIN2
MarpoPIN3
MarpoPIN4
CpsiePIN1


721 730 740 750 760 770 780
781 790 800 810 820 830 840
PoptrPIN7
PoptrPIN1
vvi 100249181
ArathPIN1
OrysaPIN1a
OrysaPIN1c
SorbiPIN10
PoptrPIN2
PoptrPIN8
vvi 100258578
vvi 100263725
OrysaPIN1b
OrysaPIN1d
SorbiPIN7
zma 100285745
GinbPIN1
PinthPIN1
ArathEIR1/ PIN2
RCOM 0583560
PoptrPIN9
vvi 100256460
OrysaPIN2
SorbiPIN11
ArathPIN4
ArathPIN7
ArathPIN3
PoptrPIN6
PoptrPIN3
RCOM 0843030
vvi 100268124
OrysaPIN3a
SorbiPIN2
zma 100383548
OrysaPIN3b
SorbiPIN9
SelmoPIN1-1
SelmoPIN1-2
SelmoPIN3-1
SelmoPIN3-2
SelmoPIN2-1
SelmoPIN4-1
SelmoPIN5-1
PhypaPINA
PhypaPINB
PhypaPINC
MarpoPIN1
ArathPIN6
PoptrPIN5
PoptrPIN4
vvi 100250503
PoptrPIN15
vvi 100253234
PoptrPIN10
PoptrPIN13
vvi 100259491
ArathPIN8
OrysaPIN8
SorbiPIN3
zma 100381964
ArathPIN5
PoptrPIN12
PoptrPIN11
RCOM 1437510
vvi 100244520
vvi 100242778
OrysaPIN5a
zma 100281763
OrysaPIN5b
SorbiPIN1
zma 100273056
OrysaPIN5c
SorbiPIN8
OrysaPIN9
SorbiPIN4
zma 100191787
PhypaPIND
MarpoPIN2
MarpoPIN3
MarpoPIN4
CpsiPIN1


841 850 860 870 880 890 900
PoptrPIN7
PoptrPIN1
vvi 100249181
ArathPIN1
OrysaPIN1a
OrysaPIN1c
SorbiPIN10
PoptrPIN2
PoptrPIN8
vvi 100258578
vvi 100263725
OrysaPIN1b
OrysaPIN1d
SorbiPIN7
zma 100285745
GinbiPIN1
PinthPIN1
ArathEIR1/ PIN2
RCOM 0583560
PoptrPIN9
vvi 100256460
OrysaPIN2
SorbiPIN11
ArathPIN4
ArathPIN7
ArathPIN3
PoptrPIN6
PoptrPIN3
RCOM 0843030
vvi 100268124
OrysaPIN3a
SorbiPIN2
zma 100383548
OrysaPIN3b
SorbiPIN9
SelmoPIN1-1
SelmoPIN1-2
SelmoPIN3-1
SelmoPIN3-2
SelmoPIN2-1
SelmoPIN4-1
SelmoPIN5-1
PhypaPINA
PhypaPINB
PhypaPINC
MarpoPIN1
ArathPIN6
PoptrPIN5
PoptrPIN4
vvi 100250503
PoptrPIN15
vvi 100253234
PoptrPIN10
PoptrPIN13
vvi 100259491
ArathPIN8
OrysaPIN8
SorbiPIN3
zma 100381964
ArathPIN5
PoptrPIN12
PoptrPIN11
RCOM 1437510
vvi 100244520
vvi 100242778
OrysaPIN5a
zma 100281763
OrysaPIN5b
SorbiPIN1
zma 100273056
OrysaPIN5c
SorbiPIN8
OrysaPIN9
SorbiPIN4
zma 100191787
PhypaPIND
MarpoPIN2
MarpoPIN3
MarpoPIN4
CpslcPIN1


901 910 920 930 940 950 960
PoptrPIN7
PoptrPIN1
vvi 100249181
ArathPIN1
OrysaPIN1a
OrysaPIN1c
SorbiPIN10
PoptrPIN2
PoptrPIN8
vvi 100258578
vvi 100263725
OrysaPIN1b
OrysaPIN1d
SorbiPIN7
zma 100285745
GinbiPIN1
PinthPIN1
ArathEIR1/ PIN2
RCOM 0583560
PoptrPIN9
vvi 100256460
OrysaPIN2
SorbiPIN11
ArathPIN4
ArathPIN7
ArathPIN3
PoptrPIN6
PoptrPIN3
RCOM 0843030
vvi 100268124
OrysaPIN3a
SorbiPIN2
zma 100383548
OrysaPIN3b
SorbiPIN9
SelmoPIN1-1
SelmoPIN1-2
SelmoPIN3-1
SelmoPIN3-2
SelmoPIN2-1
SelmoPIN4-1
SelmoPIN5-1
PhypaPINA
PhypaPINB
PhypaPINC
MarpoPIN1
ArathPIN6
PoptrPIN5
PoptrPIN4
vvi 100250503
PoptrPIN15
vvi 100253234
PoptrPIN10
PoptrPIN13
vvi 100259491
ArathPIN8
OrysaPIN8
SorbiPIN3
zma 100381964
ArathPIN5
PoptrPIN12
PoptrPIN11
RCOM 1437510
vvi 100244520
vvi 100242778
OrysaPIN5a
zma 100281763
OrysaPIN5b
SorbiPIN1
zma 100273056
OrysaPIN5c
SorbiPIN8
OrysaPIN9
SorbiPIN4
zma 100191787
PhypaPIND
MarpoPIN2
MarpoPIN3
MarpoPIN4
CpslcPIN1


961 970 980 990 1000 1010 1020
PoptrPIN7
PoptrPIN1
vvi 100249181
ArathPIN1
OrysaPIN1a
OrysaPIN1c
SorbiPIN10
PoptrPIN2
PoptrPIN8
vvi 100258578
vvi 100263725
OrysaPIN1b
OrysaPIN1d
SorbiPIN7
zma 100285745
GinbiPIN1
PinthPIN1
ArathEIR1/ PIN2
RCOM 0583560
PoptrPIN9
vvi 100256460
OrysaPIN2
SorbiPIN11
ArathPIN4
ArathPIN7
ArathPIN3
PoptrPIN6
PoptrPIN3
RCOM 0843030
vvi 100268124
OrysaPIN3a
SorbiPIN2
zma 100383548
OrysaPIN3b
SorbiPIN9
SelmoPIN1-1
SelmoPIN1-2
SelmoPIN3-1
SelmoPIN3-2
SelmoPIN2-1
SelmoPIN4-1
SelmoPIN5-1
PhypaPINA
PhypaPINB
PhypaPINC
MarpoPIN1
ArathPIN6
PoptrPIN5
PoptrPIN4
vvi 100250503
PoptrPIN15
vvi 100253234
PoptrPIN10
PoptrPIN13
vvi 100259491
ArathPIN8
OrysaPIN8
SorbiPIN3
zma 100381964
ArathPIN5
PoptrPIN12
PoptrPIN11
RCOM 1437510
vvi 100244520
vvi 100242778
OrysaPIN5a
zma 100281763
OrysaPIN5b
SorbiPIN1
zma 100273056
OrysaPIN5c
SorbiPIN8
OrysaPIN9
SorbiPIN4
zma 100191787
PhypaPIND
MarpoPIN2
MarpoPIN3
MarpoPIN4
CpslcPIN1
1021 1030 1040 1050 1060 1070 1080
PoptrPIN7
PoptrPIN1
vvi 100249181
ArathPIN1
OrysaPIN1a
OrysaPIN1c
SorbiPIN10
PoptrPIN2
PoptrPIN8
vvi 100258578
vvi 100263725
OrysaPIN1b
OrysaPIN1d
SorbiPIN7
zma 100285745
GinbiPIN1
PinthPIN1
ArathEIR1/ PIN2
RCOM 0583560
PoptrPIN9
vvi 100256460
OrysaPIN2
SorbiPIN11
ArathPIN4
ArathPIN7
ArathPIN3
PoptrPIN6
PoptrPIN3
RCOM 0843030
vvi 100268124
OrysaPIN3a
SorbiPIN2
zma 100383548
OrysaPIN3b
SorbiPIN9
SelmoPIN1-1
SelmoPIN1-2
SelmoPIN3-1
SelmoPIN3-2
SelmoPIN2-1
SelmoPIN4-1
SelmoPIN5-1
PhypaPINA
PhypaPINB
PhypaPINC
MarpoPIN1
ArathPIN6
PoptrPIN5
PoptrPIN4
vvi 100250503
PoptrPIN15
vvi 100253234
PoptrPIN10
PoptrPIN13
vvi 100259491
ArathPIN8
OrysaPIN8
SorbiPIN3
zma 100381964
ArathPIN5
PoptrPIN12
PoptrPIN11
RCOM 1437510
vvi 100244520
vvi 100242778
OrysaPIN5a
zma 100281763
OrysaPIN5b
SorbiPIN1
zma 100273056
OrysaPIN5c
SorbiPIN8
OrysaPIN9
SorbiPIN4
zma 100191787
PhypaPIND
MarpoPIN2
MarpoPIN3
MarpoPIN4
CpslcPIN1

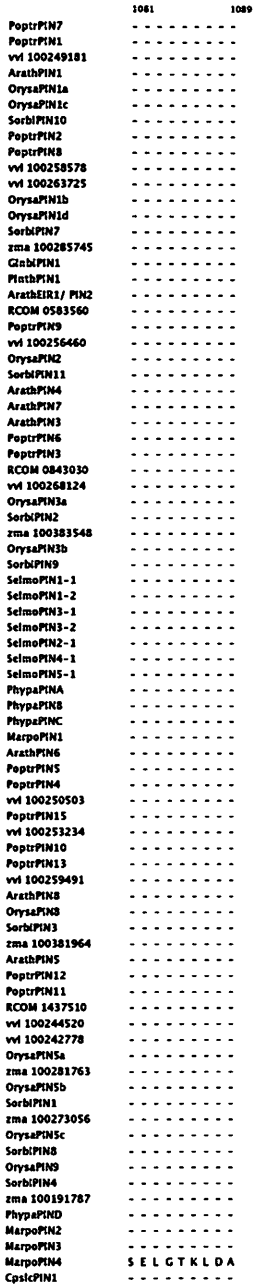


Fig. 11. Alignment of deduced amino acid sequences of *PIN* genes.

All sites are used for the phylogenetic analyses shown in Fig. 14 and 15. Dashes indicate gaps. The amino acid motifs in central hydrophilic loop domain are indicated: (S/T)XX(D/E), casein kinase II phosphorylation site (CK2), is broken-line boxed; (S/T)X(R/K), protein kinase C phosphorylation site (PKC), is solid-line boxed; NX(S/T)X, N-glycosylation site (Gly), is red-line boxed; bold-lined box is PKC and CK2 site (P/C), (S/T)X(R/K)(D/E); TPRX(N/S) is light-gray boxed; and NPXXY is bold-red-line boxed. Framed Ser and Thr are phosphorylation sites that have been revealed in AtPIN1 by Zhang et al. (2010) and Michniewicz et al. (2007). Asterisks above the alignment indicate the putative phosphorylation sites that have been predicted by Huang et al. (2010).

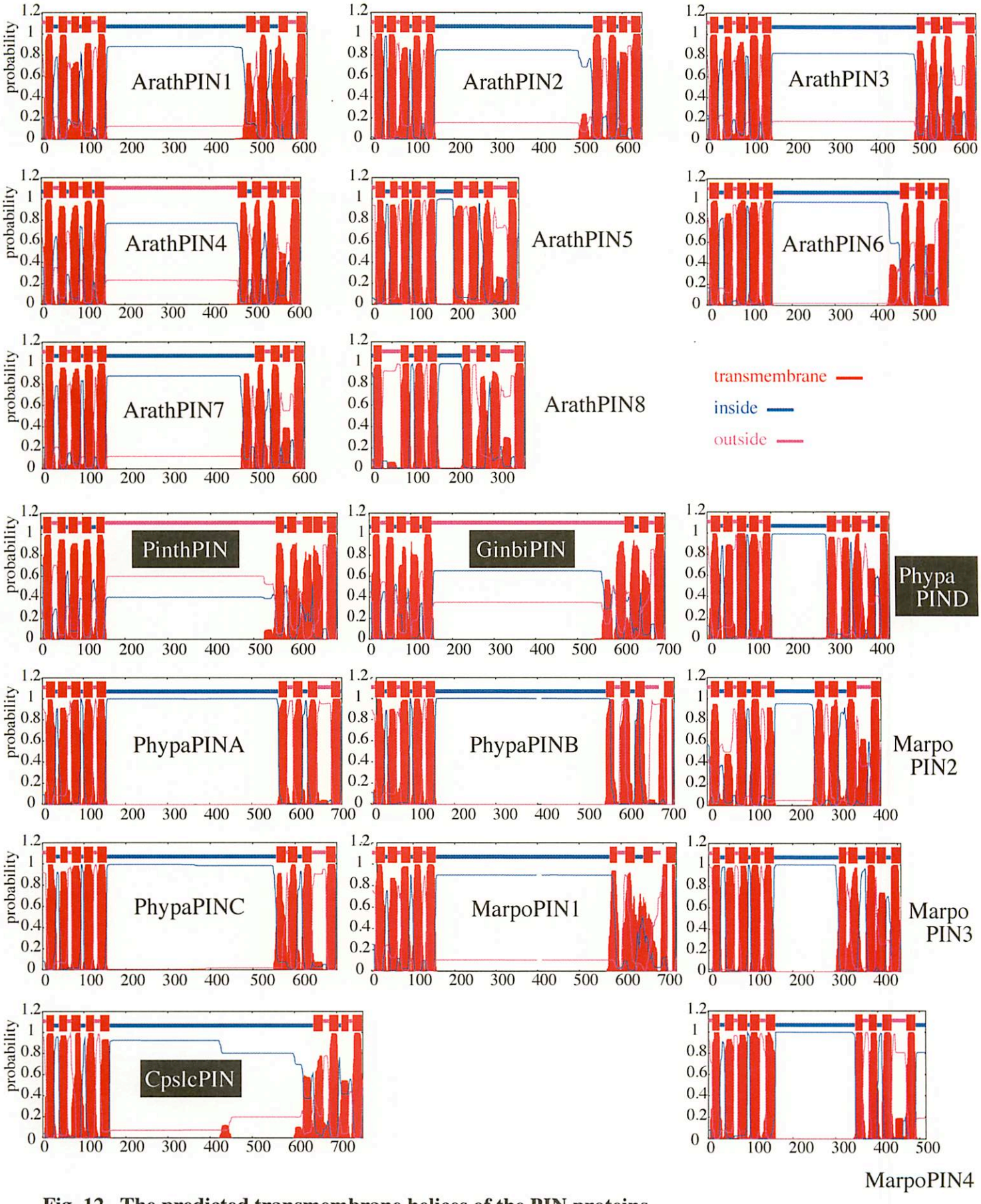


Fig. 12. The predicted transmembrane helices of the PIN proteins.

The transmembrane domains were estimated using TMHMM2: www.cbs.dtu.dk/services/TMHMM/, and the red peaks show the predicted transmembrane regions of proteins. The membrane hydrophobic regions at the N- and C-termini and the central hydrophilic region are indicated. The four *PIN* genes from this study are indicated with black boxes.

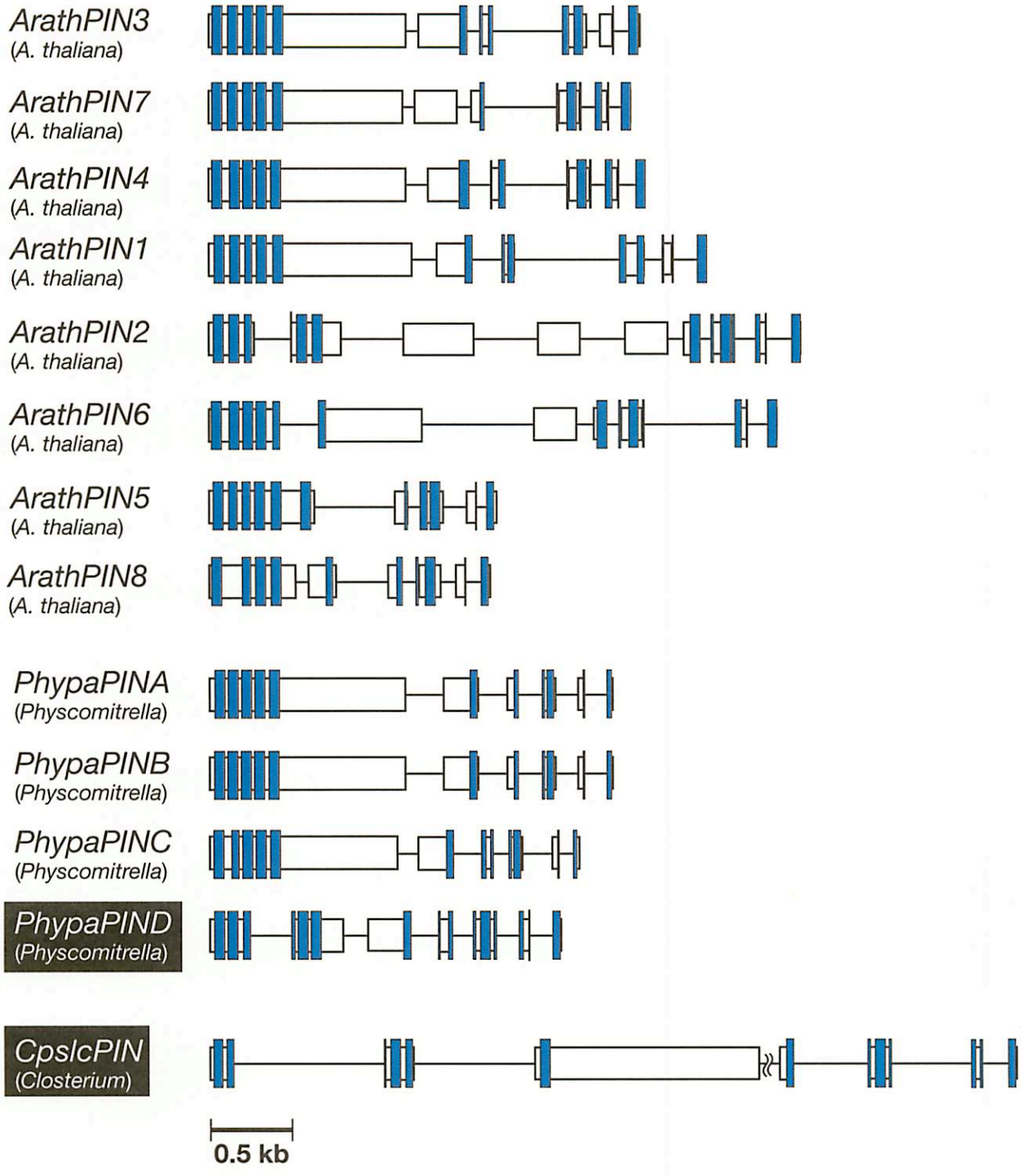


Fig. 13. The exon-intron structures of the *PIN* genes in charophycean *C. pslc* and moss *P. patens*, and angiosperm *A. thaliana*.

Gene structures for *A. thaliana* and *P. patens* were taken from http://www.Arabidopsis.org and NCBI respectively: *ArathPIN1*, At1g73590; *ArathPIN2*, At5g57090; *ArathPIN3*, At1g70940; *ArathPIN4*, At2g01420; *ArathPIN5*, At5g16530; *ArathPIN6*, At1g77110; *ArathPIN7*, At1g23080; *ArathPIN8*, At5g15100; *PhypaPINA*, NW_001865266.1; *PhypaPINB*, NW_001865274.1; *PhypaPINC*, NW_001865288.1. The *PhypaPIND* structure was constructed by sequences of cDNA and genome from this study. *CpslcPIN* was obtained from the genome sequences data in *C. pslc* (Sekimoto et al. unpublished data). The blue boxes represent transmembrane domains. The transmembrane domains were estimated using TMHMM: http://www.cbs.dtu.dk/services/TMHMM/. White boxed domains indicate translated exon sequence. Introns are depicted as black lines. The four *PIN* genes from this study are indicated with black boxes.

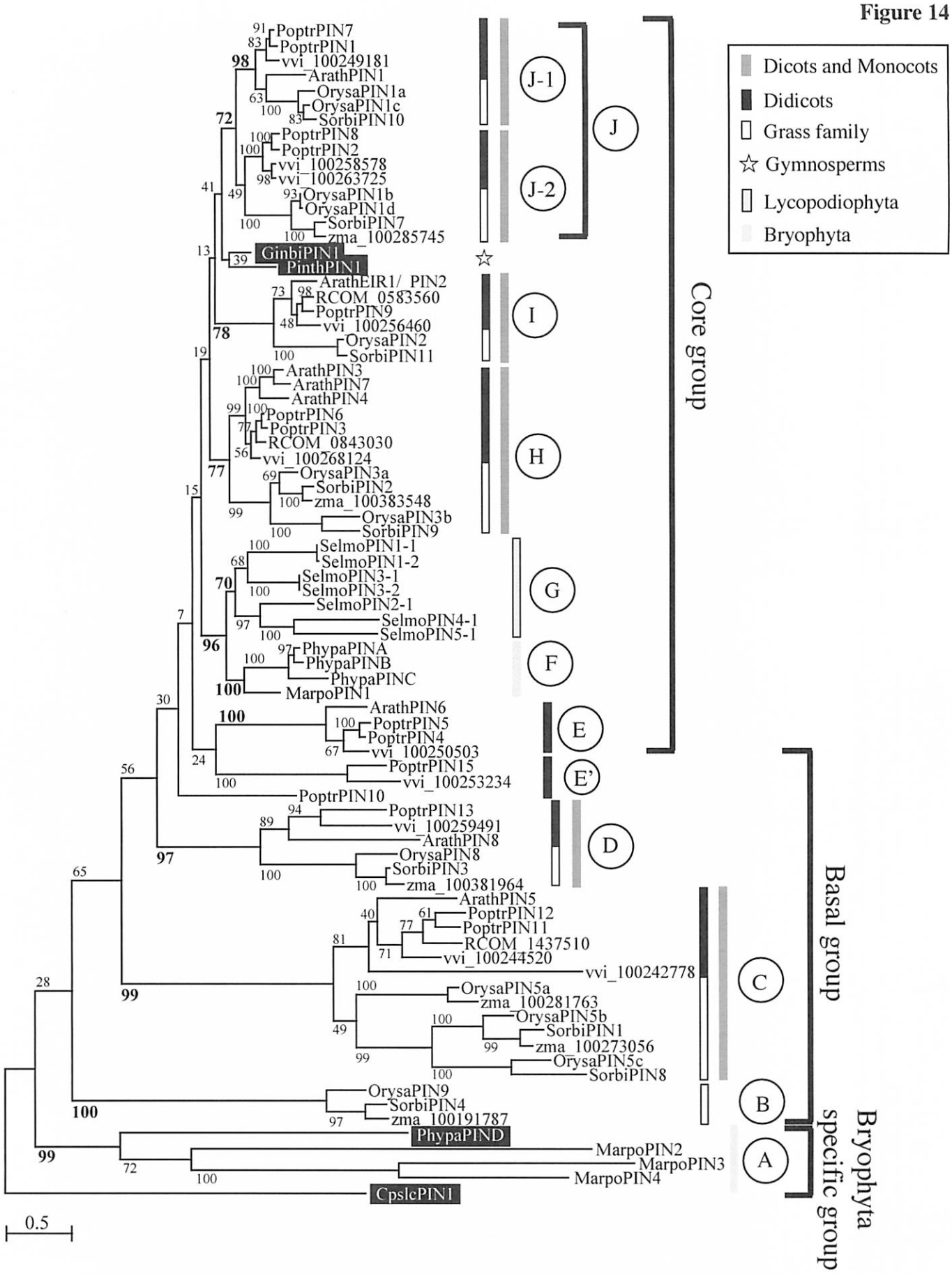


Fig. 14. Maximum-likelihood (ML) tree for *PIN* genes.

ML tree showing the phylogenetic relationships among *PIN* genes of land plants and *C. pslc* (outgroup) using 80 amino acid sequences and 1089 sites (amino acid residues and gaps in Fig. 11) as final dataset. The evolutionary history was inferred by using the maximum-likelihood based on the WAG+I+G4 substitution model. Local bootstrap probabilities are shown on or below the branches. Scale bar corresponds to 0.5 amino acid substitutions per residue. The four *PIN* genes from this study are indicated with black boxes.

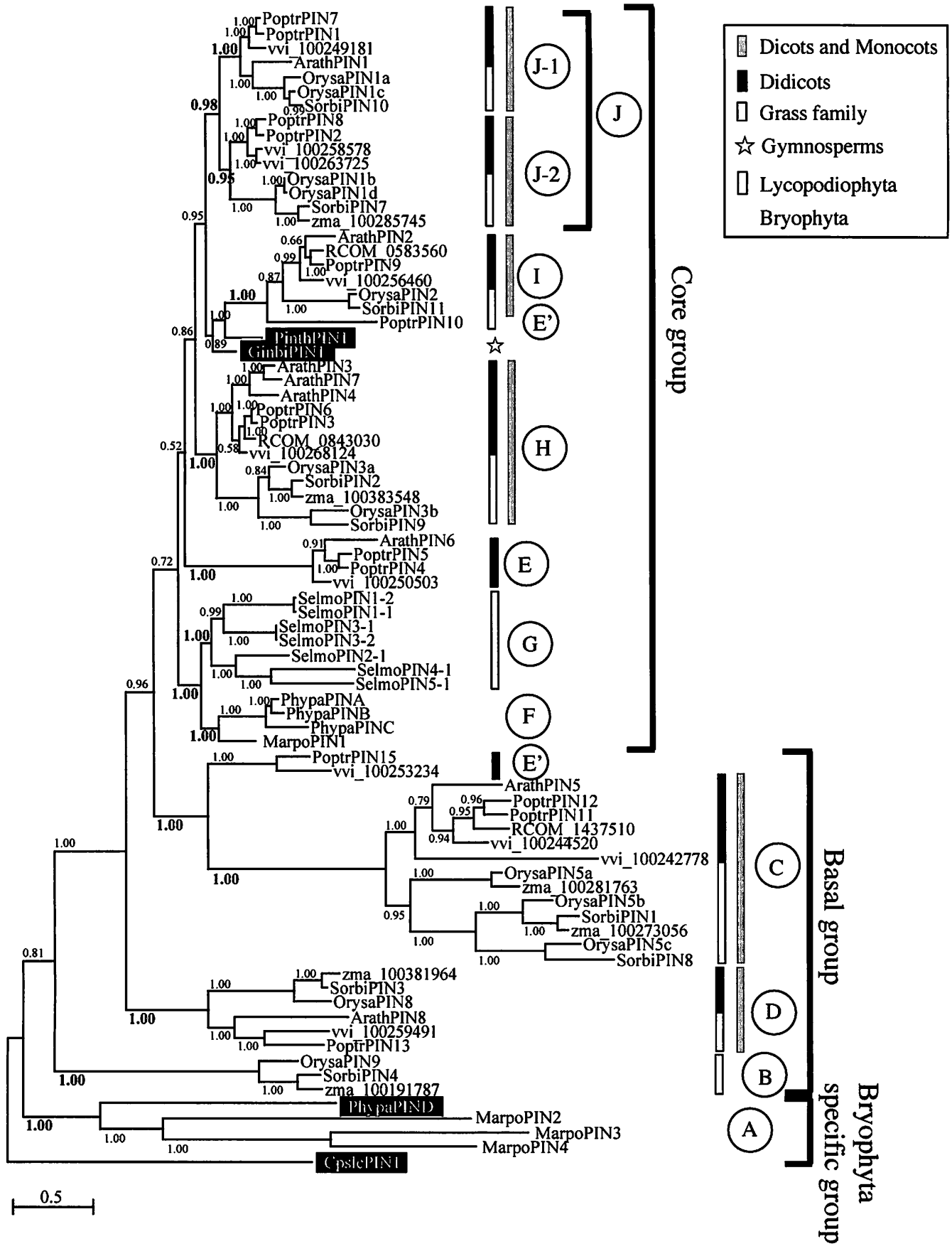
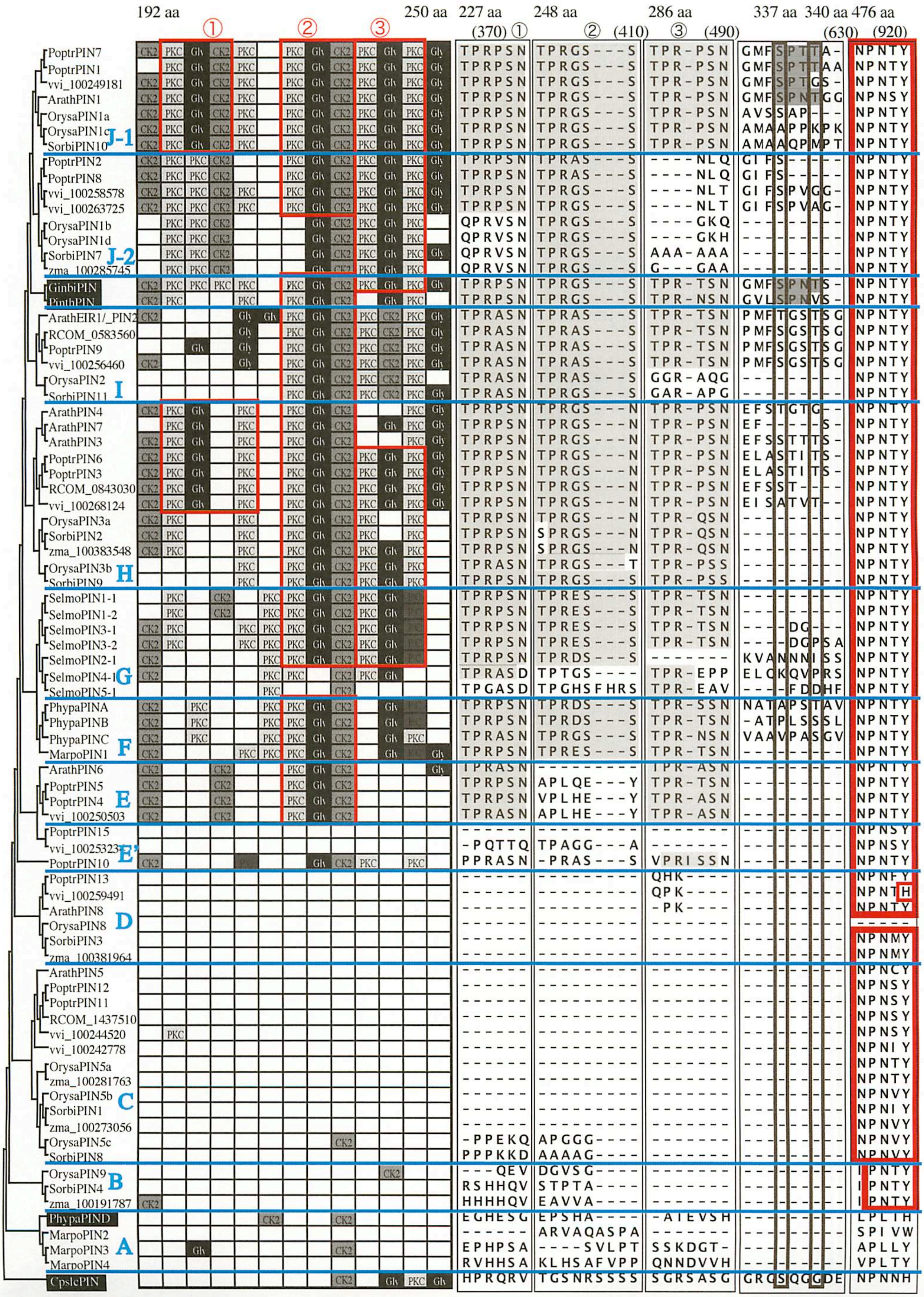


Fig. 15. Bayesian inference (BI) tree for *PIN* genes.

BI tree showing the phylogenetic relationships among *PIN* genes of land plants and *C. pslc* (outgroup) using 80 amino acid sequences and 1089 sites (amino acid residues and gaps in Fig. 11). The smallest ESS value was about 265. Numbers on or below branches are posterior probabilities. Scale bar indicates 0.5 amino acid substitutions per residue. The four *PIN* genes from this study are indicated with black boxes.

Figure 16



192 aa
250 aa
227 aa
(370)
248 aa
(410)
286 aa
(490)
337 aa
340 aa
(630)
476 aa
(920)
PoptrPIN7
PoptrPIN1
vvi_100249181
ArathPIN1
OrysaPIN1a
OrysaPIN1c
SorbiPIN10
J-1
PoptrPIN2
PoptrPIN8
vvi_100258578
vvi_100263725
OrysaPIN1b
OrysaPIN1d
SorbiPIN7
zma_100285745
J-2
GinbiPIN
PinthPIN
ArathEIR1/_PIN2
RCOM_0583560
PoptrPIN9
vvi_100256460
OrysaPIN2
SorbiPIN11
I
ArathPIN4
ArathPIN7
ArathPIN3
PoptrPIN6
PoptrPIN3
RCOM_0843030
vvi_100268124
OrysaPIN3a
SorbiPIN2
zma_100383548
OrysaPIN3b
SorbiPIN9
H
SelmoPIN1-1
SelmoPIN1-2
SelmoPIN3-1
SelmoPIN3-2
SelmoPIN2-1
SelmoPIN4-1
SelmoPIN5-1
G
PhypaPINA
PhypaPINB
PhypaPINC
MarpoPIN1
F
ArathPIN6
PoptrPIN5
PoptrPIN4
vvi_100250503
E
PoptrPIN15
vvi_100253237
PoptrPIN10
E'
PoptrPIN13
vvi_100259491
ArathPIN8
OrysaPIN8
SorbiPIN3
zma_100381964
D
ArathPIN5
PoptrPIN12
PoptrPIN11
RCOM_1437510
vvi_100244520
vvi_100242778
OrysaPIN5a
zma_100281763
OrysaPIN5b
SorbiPIN1
zma_100273056
OrysaPIN5c
SorbiPIN8
C
OrysaPIN9
SorbiPIN4
zma_100191787
B
PhypaPIND
MarpoPIN2
MarpoPIN3
MarpoPIN4
A
CpslePIN

Fig. 16. Summary of motifs in multiple alignment of deduced amino acid sequences of *PIN* genes.

It shows relationship between the motifs in Fig. 11 and gene tree in Fig. 14. The red line boxes indicate cluster of N-glycosylation site and two phosphorylation sites (Gly/2P). CK2 is casein kinase II phosphorylation site {(S/T)XX(D/E)}. PKC is protein kinase C phosphorylation site {(S/T)X(R/K)}. Gly is N-glycosylation site {NX(S/T)X}. P/C is PKC and CK2 {(S/T)X(R/K)(D/E)}. Light gray indicates TPRXS(N/S) motif. Framed Ser and Thr are phosphorylation sites that have been revealed in AtPIN1 by Zhang et al. (2010) and Michniewicz et al. (2007). Bold-red-lined box is NPXXY motif. Dashes indicate gaps. Grouping is provided by gene tree in Fig. 14, which are indicated by blue lines. The amino acid number corresponding to ArathPIN1 is above. The site number in Fig. 11 is in parenthesis. The four *PIN* genes from this study are indicated with black boxes.

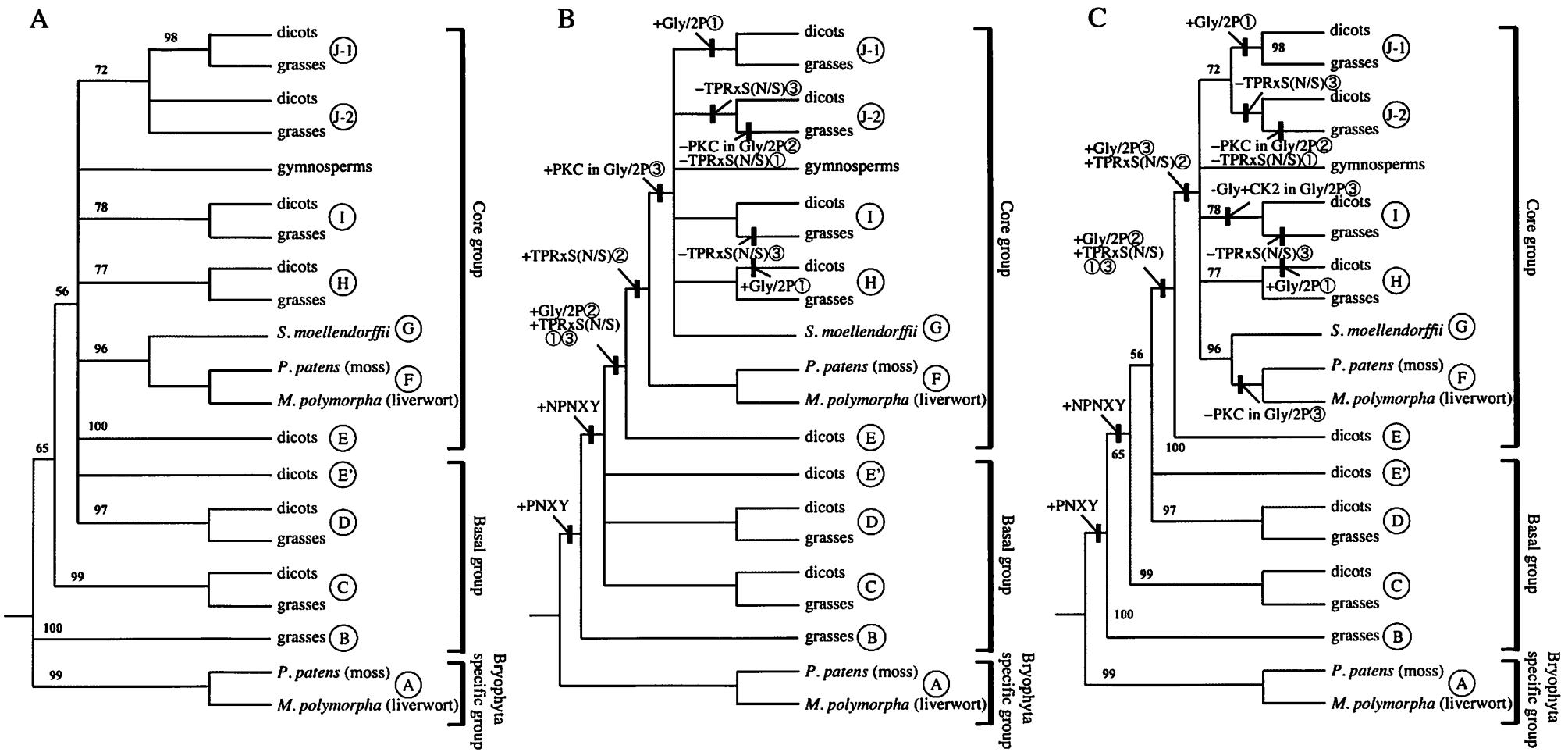


Fig. 17. The summary of phylogenetic relationships among *PIN* genes in land plants.

Phylogenetic relationships among *PIN* genes in land plants was constructed by based on the gene tree and signature motifs shown in Fig. 14 and Fig. 16. (A) A summary of the ML tree. Unresolved issues are indicated by multifurcating nodes, because of the low bootstrap support (less than 50%) in Fig. 14. (B) A phylogenetic tree was constructed by using the signature motifs in Fig. 16. The multifurcating node indicates the unresolved issue. (C) A inference of phylogenetic relationships among *PIN* genes in land plants was constructed based on (A) and (B). Numbers on or below branches are the local bootstrap probabilities. Black bar on branch indicates the molecular synapomorphy. Plus (+) indicates gain. Minus (-) indicates absence.

Figure 17

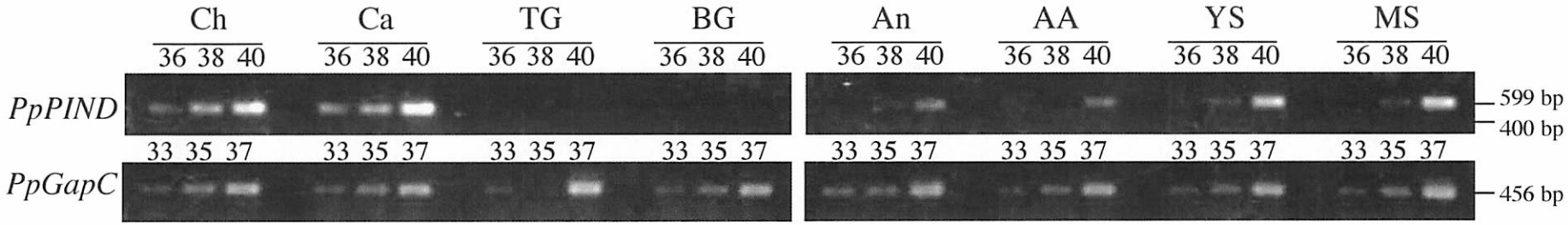


Fig. 18. Semi quantitative RT-PCR analysis of *P. patens PIND* gene.

(Ch) chloronemata; (Ca) caulonemata; (TG) shoot tips of adult gametophores without rhizoids at 25°C; (BG) base of adult gametophores with rhizoids at 25°C; (An) shoot tips of gametophores bearing antheridia, 2 weeks after induction; (AA) shoot tips of gametophores bearing antheridia and archegonia, 3weeks after induction; (YS) shoot tips of gametophores bearing young sporophyte, 3.5 weeks after induction; (MS) shoot tips of gametophores bearing mature sporophyte, 5 weeks after induction. The *PpGapC* was used as a quantifying control.

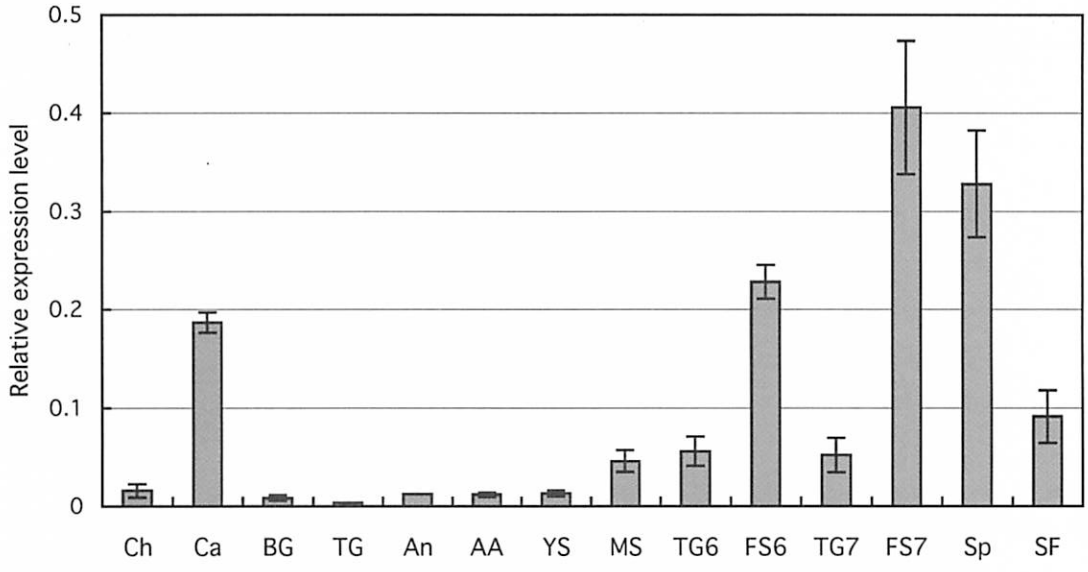


Fig. 19. Quantitative real-time RT-PCR analysis of *P. patens PIND* gene.

(Ch) chloronemata; (Ca) caulonemata; (BG) base of adult gametophores with rhizoids at 25°C; (TG) shoot tips of adult gametophores without rhizoids at 25°C; (An) shoot tips of gametophores bearing antheridia, 2 weeks after induction; (AA) shoot tips of gametophores bearing antheridia and archegonia, 3weeks after induction; (YS) shoot tips of gametophores bearing young sporophyte, 3.5 weeks after induction; (MS) shoot tips of gametophores bearing mature sporophyte, 5 weeks after induction; (TG6) shoot tips of gametophores excluding fully matured sporophytes, 6 weeks after induction; (FS6) fully matured sporophytes in early period, 6 weeks after induction; (TG7) shoot tips of gametophores excluding fully matured sporophytes, over 7 weeks after induction; (FS7) fully matured sporophytes in latter period, over 7 weeks after induction; (Sp) sporangium of fully matured sporophytes in latter period, over 7 weeks after induction; (SF) shoot tips of gametophores including foot from fully matured sporophytes, over 7 weeks after induction. The *PpTUA1* was used as a quantifying control to normalize the amount of templates in RT-PCR.

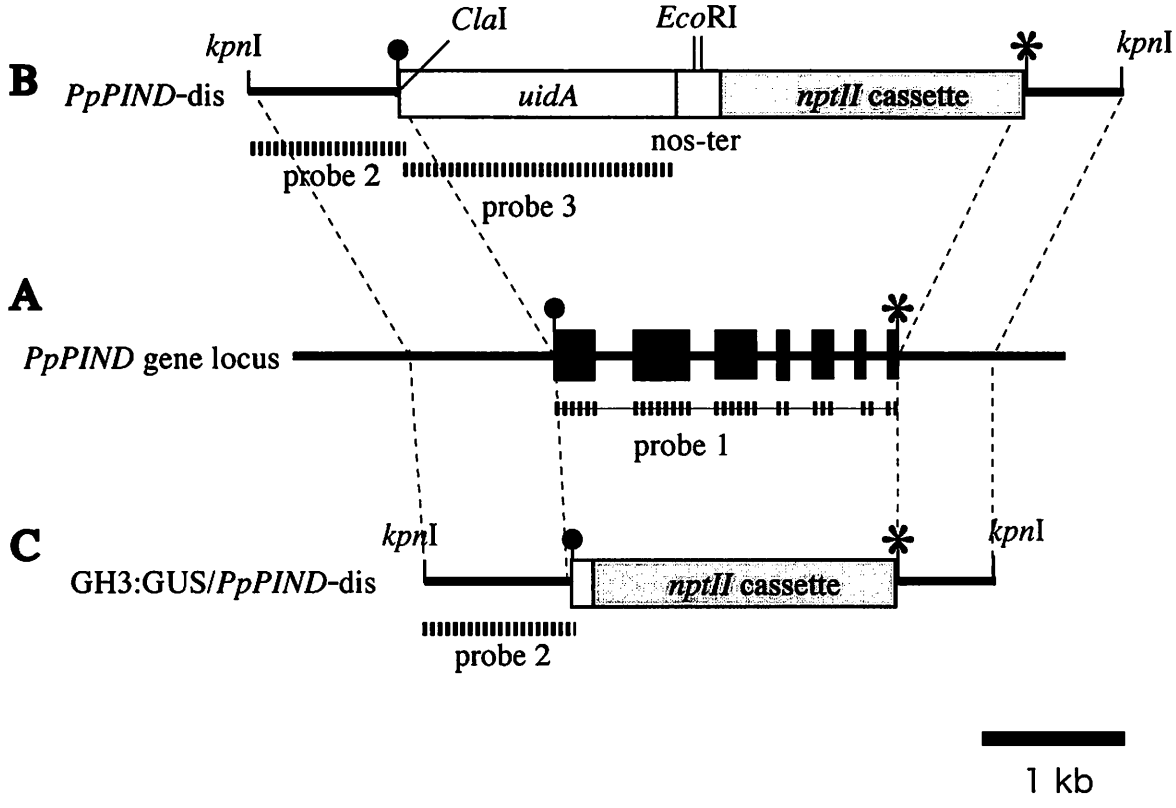


Fig. 20. Genomic structures of *PpPIND* in wild type and transformant lines.
Genomic structures of *PpPIND* gene in wild type (A), *PpPIND* disruptant, PpPIND-dis (B), and GH3:GUS/*PpPIND* disruptants, GH3:GUS 1-1/PpPIND-dis and GH3:GUS 36-1/PpPIND-dis. The boxes indicate the translated regions. The lines indicate the untranslated regions. The circle and asterisc indicate putative start and stop codons. The thin-broken lines indicate same region. The probes for genomic southern analyses in Fig. 21, 22, and 23 are shown as bold-broken lines. The *uidA*, *nos-ter*, and *nptII* designate the uidA-coding region, nopaline synthase polyadenylation signal, and NPTII expression cassette, respectively. *Kpn*I, *Cla*I, and *Eco*RI indicate the restrict enzyme sites. Scale bar corresponds to 1 kb.

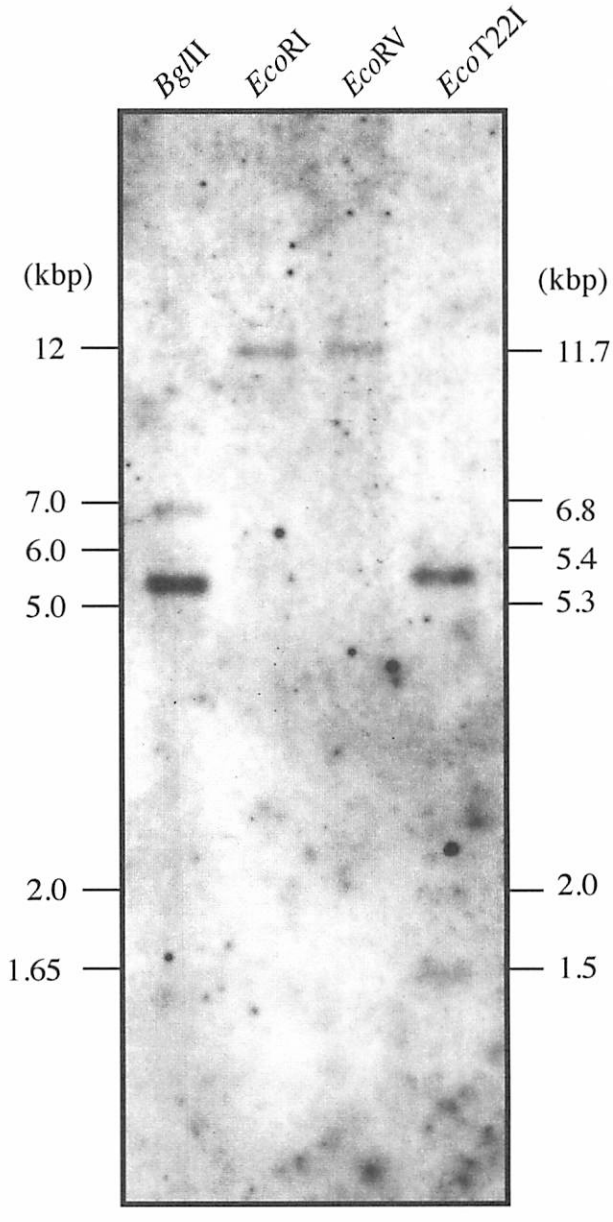


Fig. 21. Southern analysis of *PpPIND* gene in the wild type.

Genomic DNA of wild type was digested with *Bgl*II, *Eco*RV, *Eco*T22I, and *Eco*RI, and the probe1 were used for hybridization.

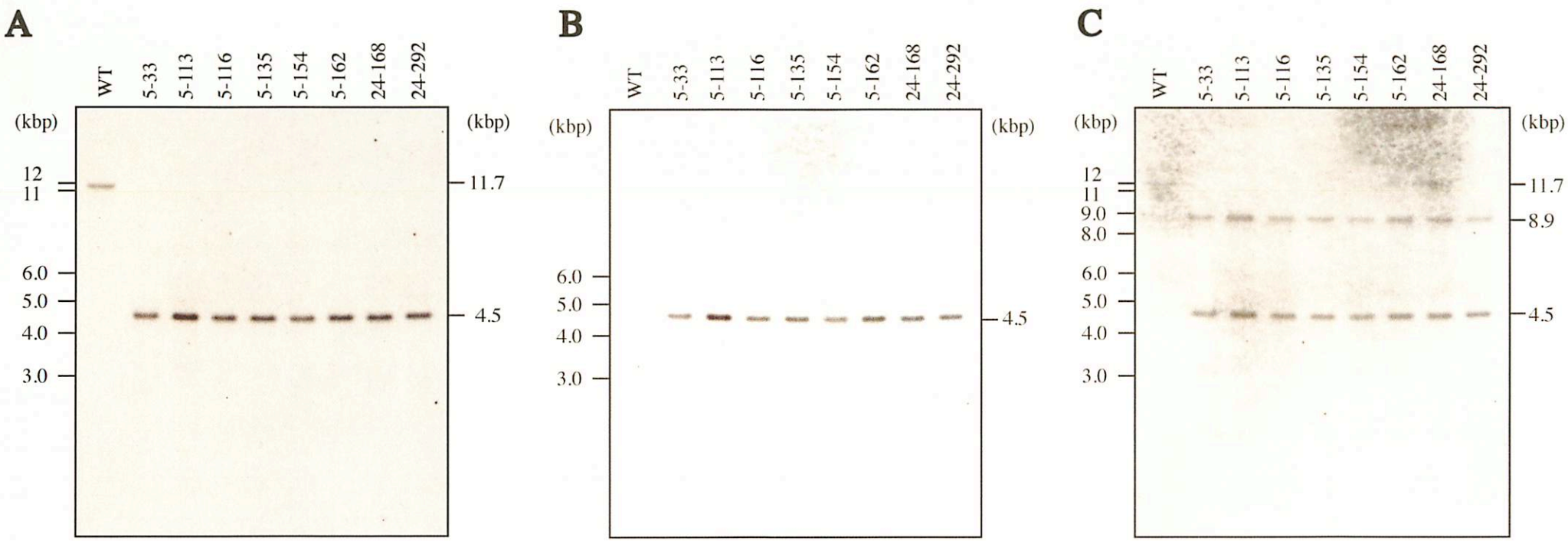


Fig. 22. Southern analyses to confirm gene targeting in PpPIND-dis lines.

Genomic DNA of wild type (wt) and transformants PpPIND-dis lines were digested with *Eco*RI, and the probe2 (A), probe3 (B), or plasmid including construct as probe (C) were used, respectively. (A) The membrane for hybridized with probe2. (B) The membrane for (A) re-hybridized with probe3. (C) The membrane for (B) re-hybridized with plasmid pPpPIND-dis probe.

Figure 22

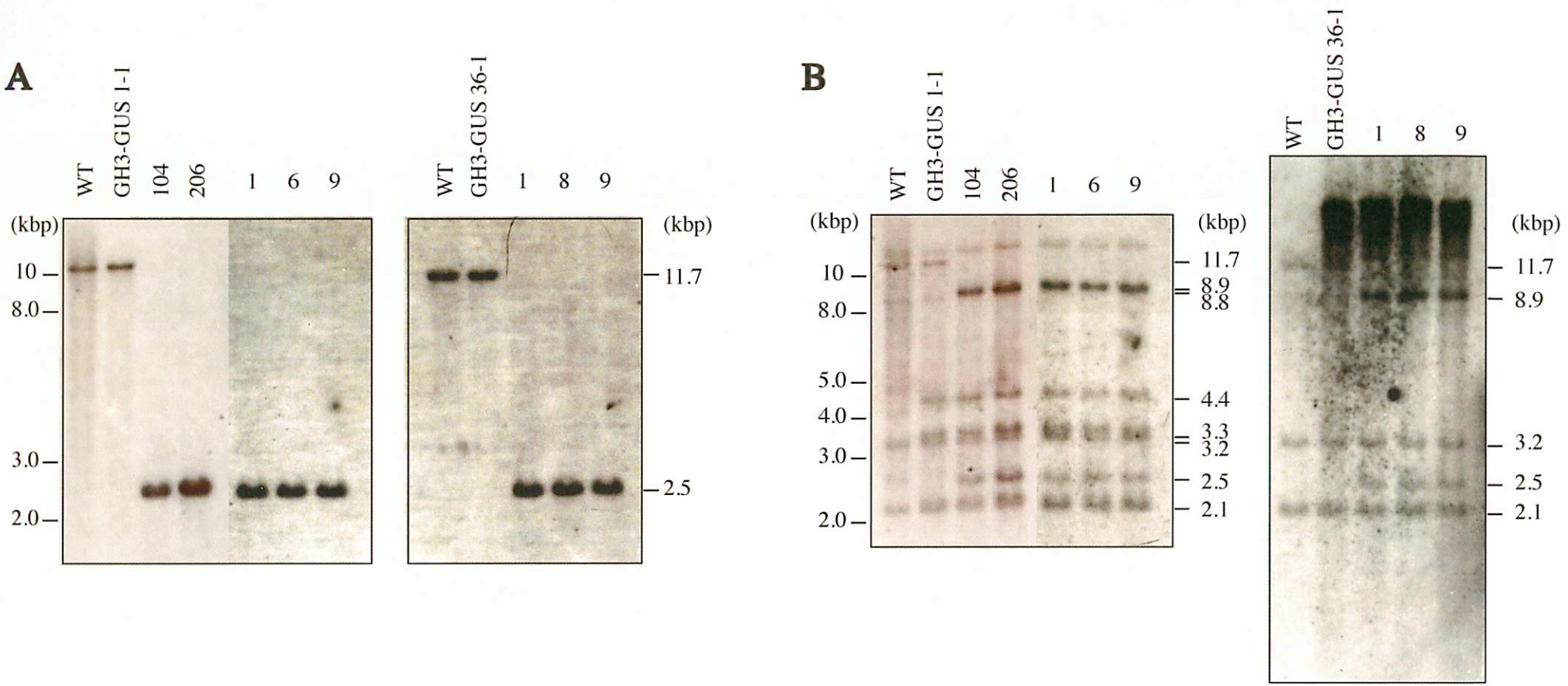


Fig. 23. Southern analyses to confirm gene targeting in GH3:GUS/PpPIND-dis lines.

Genomic DNA of wild type (wt), GH3:GUS1-1, 36-1, and GH3:GUS/PpPIND-dis lines were digested with *Eco*RI, and the probe2 or plasmid including construct as probe were used, respectively. (A) The membrane for hybridized with probe2. (B) The membrane for (A) re-hybridized with plasmid pGH3GUS/PpPIND-dis probe.

Figure 23

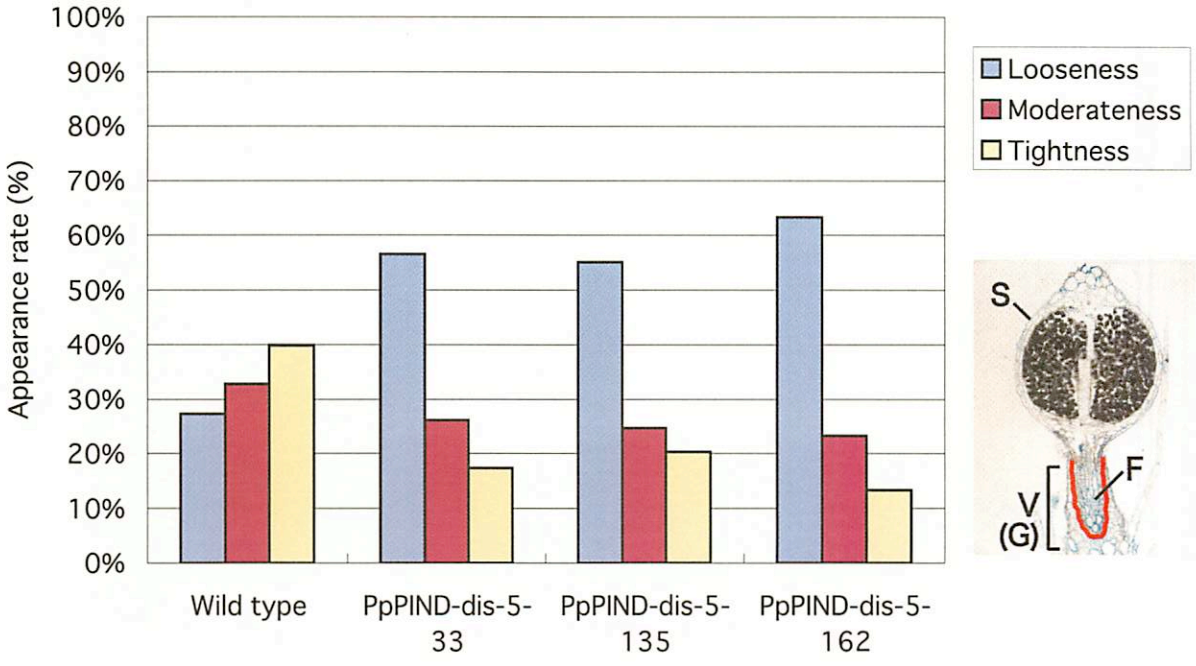


Fig. 24. Comparison of frequency of junction strength between sporophyte and gametophyte of wild type and *PpPIND* disruptant lines.

The untransformed line [Wild type (n = 128)] and transformed lines [PpPIND-dis-5-33 (n = 23), PpPIND-dis-5-135 (n = 69) and PpPIND-dis-5-162 (n = 30)] were measured. Blue, red, and yellow bars indicate the loose, moderate, and tight junction, respectively. The inset indicates longitudinal microtome section of sporophyte-gametophyte junction in *P. patens*. Red line indicates the place of sporophyte-gametophyte junction. F foot, G gametophyte, S sporophyte, V vaginula.

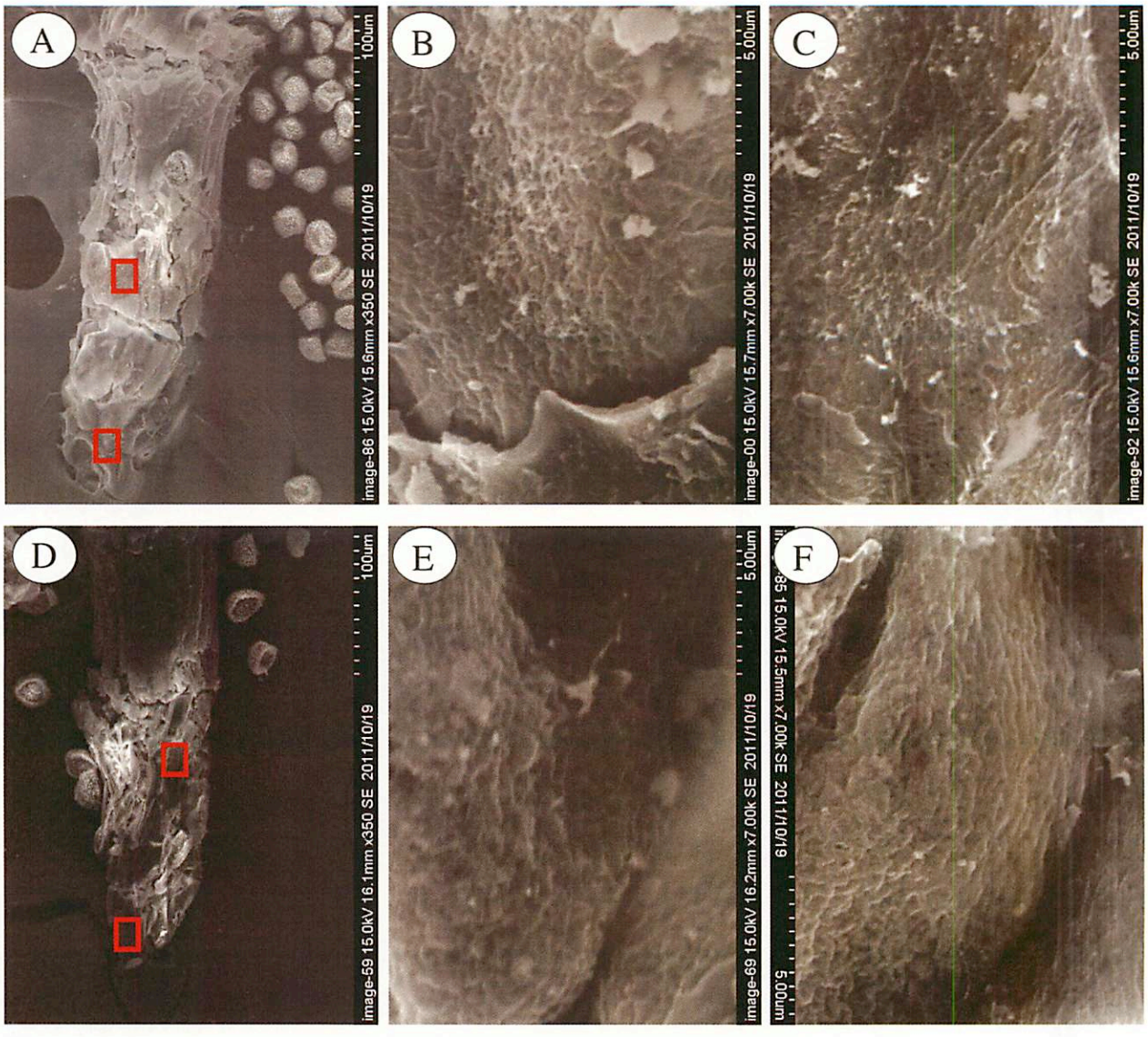


Fig. 25. The scanning electron microscopy (SEM) micrographs of sporophyte foot of the wild type and *PpPIND* disruptant line.

(A-C) The wild type. (D-F) *PpPIND* disruptant line, PpPIND-dis-5-135. (B, E) Close-up (x7000) of a cell of the foot tip in (A) and (D). (C, F) Close-up (x7000) of a cell of middle region of the foot in (A) and (D). Red frames indicate position of (B) or (C) or (E) or (F).

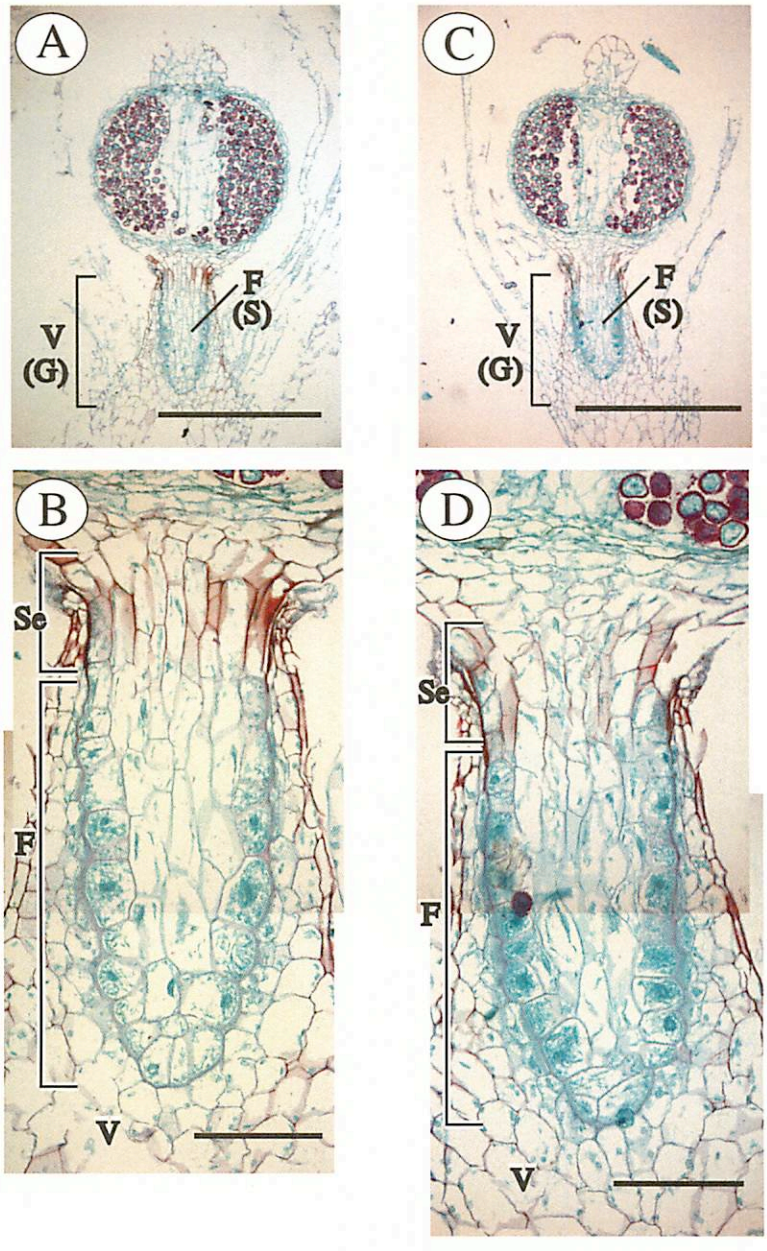


Fig. 26. The microtome section (MS) micrographs of longitudinal section of sporophyte-gametophyte junction of the wild type and *PpPIND* disruptant lines.
(A, B) The wild type. (C, D) *PpPIND* disruptant line, PpPIND-dis-5-135. (B, D) Magnification of (A) and (C) showing sporophyte-gametophyte junction. F foot, G gametophyte, S sporophyte, Se seta, V vaginula. Scale bars are 500 μm in (A, C), and 100 μm in (B, D).

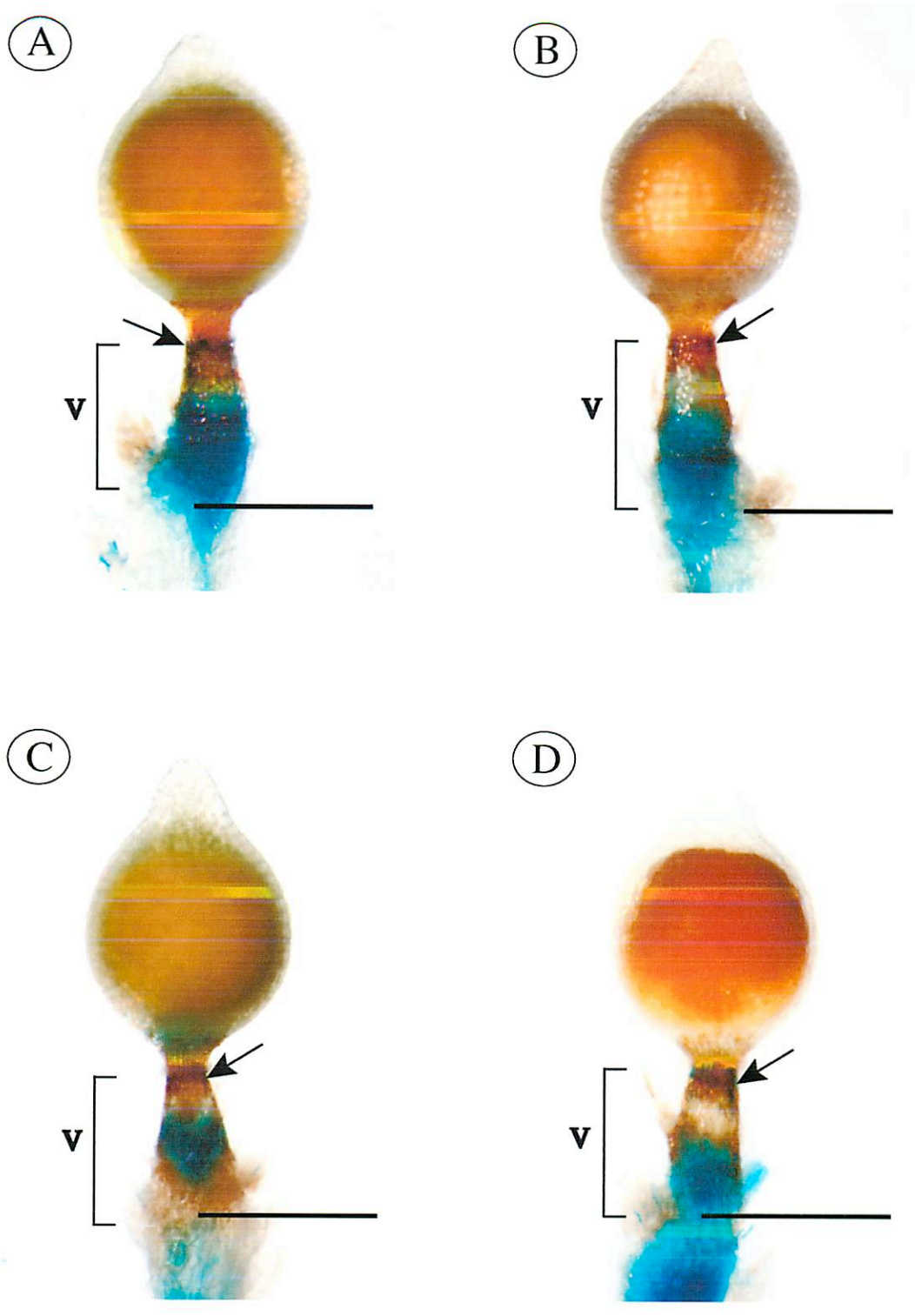


Fig. 27. Histochemical detection of GUS activity in the control strains and GH3:GUS/*PpPIND* disruptant lines.

(A, C) The control strains, GH3:GUS1-1 (A) and GH3:GUS36-1 (C). (B, D) GH3:GUS/*PpPIND* disruptant lines, GH3:GUS1-1/PpPIND-dis-9 (B) and GH3:GUS36-1/PpPIND-dis-8 (D). A change of blue staining was observed in upper region of vaginula (arrows): (A) was stained but not, or very weak, in (B); and (C) was very weak but (D) was sufficient. V vaginula. Scale bars are 500 μm.