African Journal of Biotechnology

Full Length Research Paper

In-silico identification and phylogenetic analysis of auxin efflux carrier gene family in Setaria italica L.

Tapan Kumar Mohanta¹*, Mickael Malnoy², Nibedita Mohanta³ and Chidananda Nagamangala Kanchiswamy⁴*

¹National Institute of Plant Genome Research, Aruna Asaf Ali Marg, 110067 New Delhi, India.
²Istituto Agrario San Michele all'Adige, Research and Innovation Centre, Foundation Edmund Mach, Trento, Italy.
³Departent di Biologia Vegetale, Viale Mattioli, 10125, University of Turin, Italy.
⁴Istituto Agrario San Michele all'Adige, Research and Innovation Centre, Foundation Edmund Mach, Trento, Italy.

Accepted 07 August, 2013

The phytohormone auxin is crucial for plant growth and development. Transportation and distribution of auxin throughout the plant is very important to maintain the auxin homeostasis. Auxin efflux carrier genes play significant roles in auxin transport. In this study, we identified 12 auxin efflux carrier genes from the Setaria italica genome, which are similar in number with that of monocotyledonous plant Oryza sativa. Twelve (12) SiPIN genes are distributed in eight scaffolds. All the identified genes possess the transmembrane auxin efflux carrier domain. Phylogenetic analysis shows that SiPIN genes are much closer to Sorghum bicolor and O. sativa PIN genes of the grass family.

Key words: Auxin efflux carrier, auxin efflux carrier domain, phosphorylation, transmembrane domain.

INTRODUCTION

In the model plant Arabidopsis thaliana, auxin plays a crucial role in regulating and coordinating plant growth and is involved in many developmental processes, meristem includina embryogenesis, maintenance. organogenesis, lateral root initiation, vascular tissue differentiation and tropisms. Specific auxin influx carriers (AUX/LAX proteins) and efflux carriers (PIN and PGP/MDR proteins) mediate a directional, active, cell-tocell auxin transport, creating auxin concentration maxima in specific tissues or cells. PIN auxin efflux carriers play a major role in mediating and regulating polar auxin transport (PAT), creating the auxin gradients that provide positional information for cells and tissues development (Benkova et al., 2003; Michniewicz et al., 2007; Reinhardt et al., 2000).

In *A. thaliana*, there are eight *PIN* genes (At*PIN1*–*AtPIN8*) coding for proteins that differ in the length of the hydrophilic loop in the middle of their polypeptide chain (Krecek et al., 2009a; Zazimalova et al., 2007). The long PIN proteins of *Arabidopsis* viz., *PIN1,PIN4* and *PIN7* show plasma membrane localization and their polar localization determines direction of auxin flux(Friml 2010). The three *PIN* proteins *PIN5*, *PIN6*, and *PIN8*, have a shorter central hydrophilic domain and both *PIN5* and *PIN8*, have been shown to localize in the endoplasmic reticulum, suggesting a possible role in regulating intracellular auxin homeostasis(Wabnik et al., 2010; Wabnik et al., 2011). The classification of *AtPIN6* is more controversial since it has a partially reduced hydrophilic loop with high sequence similarities at trans-membrane

*Corresponding author. E-mail:nostoc.tapan@gmail.com. and chidananda.nagamangala@fmach.it.

Abbreviations: AtPIN, Arabidopsis thaliana auxin efflux carrier; OsPIN, Oryza sativa auxin efflux carrier; PpPIN, Physcomitrella patens auxin efflux carrier; PtPIN, Populus trichocarpa auxin efflux carrier; SbPIN, Sorghum bicolor auxin efflux carrier; SiPIN, Setaria italica auxin efflux carrier.

regions (Krecek et al., 2009a; Mravec et al., 2009). In addition to the eight *AtPIN* proteins, *Arabidopsis* encodes seven *PIN* like genes and they form a different clusters and the role of these is yet to find out (Paponov et al., 2005).

Many homologous PIN genes were well characterized in monocot species like rice (Oryza sativa) and maize (Zea mays). Both specific features and homologies between monocots and Arabidopsis (eudicot) PIN families have been shown. Monocot-specific features comprise both sequence clustering in phylogenetic analyses and expression pattern at transcript and protein level. In rice, the sequence analysis of the 12 PIN genes present in the genome showed that rice has four PIN1 genes and one OsPIN2, while no OsPIN protein was grouped into the AtPIN3, AtPIN4 and AtPIN7 cluster. Four OsPIN genes encode for rice PIN proteins with a short central hydrophilic domain: three OsPIN5 and one OsPIN8. Furthermore, three OsPIN proteins appear monocot-specific: OsPIN9, OsPIN10a, and OsPIN10b. OsPIN9 has a central hydrophilic domain intermediate in length between long and short PINs of Arabidopsis and its expression analysis at transcription level suggests a possible function in adventitious root differentiation. OsPIN10a and OsPIN10b have a long central hydrophilic domain (Carraro et al., 2006; Forestan et al., 2012; Forestan and Varotto 2010; 2012; Xu and Scheres 2005). So far, three PIN1 genes were described in maize using an antibody raised against AtPIN1 protein (Forestan and Varotto, 2010). Recent studies of PIN genes on Sorghum bicolor revealed the presence of 11 PIN genes; at least there members were grouped in the AtPIN1 cluster and another three in the AtPIN5 cluster (Shen et al., 2010; Wang et al., 2010).

S. italica [(L.) P. Beauv.] commonly known as foxtail millet is one of the most cultivated millet species grown worldwide including India, China, Japan, Australia North and South America (Devos et al., 1998). Foxtail millet is a diploid grass with small genome (≈515 Mb) and its draft sequences has been published recently (Bennetzen et al., 2012). The major phytohormone auxin is central to plant growth and development. Availability of publicly accessible genome sequences of S. italica lead us to find the auxin efflux carrier genes (PIN) using an in-silico approach. Here, we used bioinformatics and comparative genomics approaches to find auxin efflux carrier genes in S. italica.

MATERIALS AND METHODS

Auxin efflux carrier (*PIN*) genes of *S. italica* were identified from plant genome database (http://www.plantgdb.org) and phytozome (www.phytozome.net) database (Dong et al., 2004; Duvick et al., 2008; Goodstein et al., 2012). To identify *PIN* genes, orthologous auxin efflux carrier genes from *A. thaliana* were used as search query. *Arabidopsis PIN* genes were downloaded from "The Arabidopsis Information Resources" (http://www.arabidopsis.org/). Hidden markov model approach was carried out to find the auxin

efflux carrier genes of S. italica (Altschul et al., 1997). Identified StPIN genes were again confirmed by running BLASTP searches in "The Arabidopsis Information Resources" and presence of auxin efflux carrier domains were confirmed by SWISS MODEL Workspace (www.swissmodel.expasy.org/workspace/). Nomenclature of identified StPIN genes were carried out according to BLASTP similarity found with A. thaliana AtPIN genes. TMMOD (The Hidden Markov Model for Transmembrane Protein Topology (http://www.cbs.dtu.dk/services/TMHMM/) analyses were carried out to confirm the presence of transmembrane domains in SiPIN proteins(Kahsay et al., 2005; Kahsay et al., 2004). Orthologous PIN genes from A. thaliana (AtPIN), O. sativa (OsPIN), Physcomitrella patens (PpPIN), Populus trichocarpa (PtPIN), and S. bicolor (SbPIN) were used to analyze protein sequence similarity and construction of phylogenetic trees. OsPIN genes were downloaded from The TIGR Rice Genome Annotation Resources (Ouyang et al., 2007) whereas, PpPIN, PtPIN, SbPIN genes were downloaded from plant genome database and phytozome database. Multiple alignments of PIN genes from the above mentioned species were carried out by using the online available software Multalin (http://multalin.toulouse.inra.fr/multalin/). Phylogenetic tree was constructed by using MEGA5.2 software.

RESULTS AND DISCUSSION

Genome wide analysis of the S. italica genome led to the identification of 12 auxin efflux carrier (SiPIN) genes (Table 1). This result shows, Setaria has the same number of SiPIN genes as of rice and has four more SiPIN genes than A. thaliana. The major genome assembly of S. italica is arranged in 336 scaffolds. The first nine scaffolds are pseudomolecules and 98.9% of sequence data is presented in the nine pseudomolecule. Besides, the Setaria genome has 35,471 loci containing 40,599 protein coding transcripts (Bennetzen et al., 2012). The 12 identified S. italica auxin efflux carrier genes are distributed in eight scaffolds. Scaffold five contains four auxin efflux carrier genes (SiPIN4a, SiPIN5a, SiPIN5b and SiPIN8). The biggest SiPIN gene was SiPIN2 with an ORF (open reading frame) length of 1890 nucleotides present in scaffold 4, whereas the smallest one was SiPIN5c which was present in scaffold 6. Among the 12 SiPIN genes, seven SiPIN genes (SiPIN1a, SiPIN1b, SiPIN4a, SiPIN4b, SiPIN4c, SiPIN4d and SiPIN8) contained five introns each and SiPIN2 and SiPIN5d contained six introns each (Figure 1). SiPIN1 transcript organization matched with that of OsPIN1 and AtPIN1 indicating their close homology (Wang et al., 2009).

Auxin efflux carrier genes are characterized by the presence of a transmembrane auxin efflux carrier domain (Carraro et al., 2006; Forestan et al., 2012). All SiPIN genes contain transmembrane auxin efflux carrier domains (Figure 2). Among the 12 SiPINs, seven SiPIN genes (SiPIN1a, SiPIN1b, SiPIN2, SiPIN4a, SiPIN4b, SiPIN4c and SiPIN4d) contain a long transmembrane domain. The other SiPIN genes (SiPIN5a, SiPIN5b, SiPIN5c, SiPIN5d and SiPIN8) contained a short transmembrane domain. In Arabidopsis, AtPIN4 has a short transmembrane domain which is different from that

Table 1. Phytozome locus ID and transcript information of	SiPIN.	Naming of SiPIN were done as found by BLASTP against
the Arabidopsis Information Resources database.		

Locus ID	Gene name	ORF Length	Number Of a.a	Number of Introns	5'-3' Coordinates
Si016714m	SiPIN1a	1785	594	5	scaffold_1: 37807344 - 37810727
Si006110m	SiPIN1b	1794	597	5	scaffold_4: 8906772 - 8909697
Si006060m	SiPIN2	1890	629	6	scaffold_4: 30687696 -30691064
Si000693m	SiPIN4a	1863	620	5	scaffold_5: 30444777 - 30448448
Si009737m	SiPIN4b	1677	558	5	scaffold_7: 34823548 - 34826535
Si025109m	SiPIN4c	1773	590	5	scaffold_3: 9460738 - 9463047
Si026177m	SiPIN4d	1683	560	5	scaffold_8: 1266812 - 1269757
Si001955m	SiPIN5a	1095	364	4	scaffold_5: 44864303 - 44869298
Si003879m	SiPIN5b	1272	423	4	scaffold_5: 39347962 - 39350295
Si015697m	SiPIN5c	1029	342	2	scaffold_6: 33597186 - 33599056
Si033365m	SiPIN5d	1134	377	6	scaffold_2: 36679971 - 36681750
Si003769m	SiPIN8	1081	361	5	scaffold_5: 35148382 - 35150803

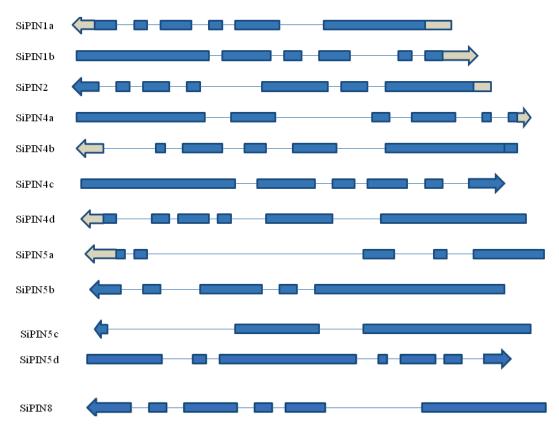


Figure 1. Transcript organization of *SiPIN* genes. Blue color boxes indicate the exons and lines indicate the introns of respective *SiPIN* genes. The arrow mark indicate the direction of expression of transcript.

of *SiPIN* genes showing diversification of PIN genes among species. *Arabidopsis AtPIN1*, rice *OsPIN1* and *OsPIN2* were grouped under long transmembrane auxin efflux carrier domain (Wang et al., 2009). *SiPIN1* and *SiPIN2* genes contain a long transmembrane domain, showing close relationship with that of *AtPIN1*, *OsPIN1*

and OsPIN2.

Phylogenetic analysis with orthologous PIN genes from A. thaliana (AtPIN), O. sativa (OsPIN), Physcomitrella patens (PpPIN), Populus trichocarpa (PtPIN), S. bicolor (SbPIN) resulted in clustering into three major groups (group I, II and III) (Figure 3). In group I, SiPIN1a and

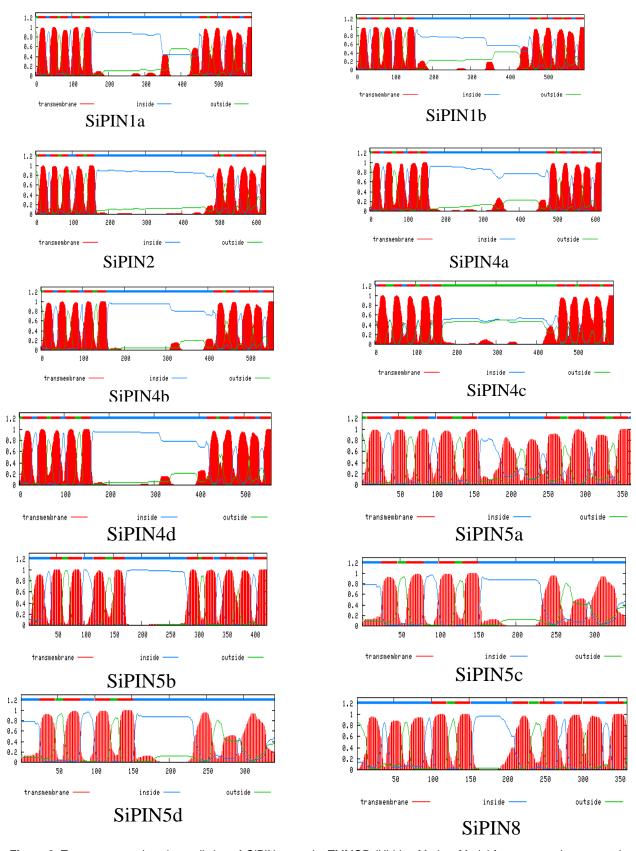


Figure 2. Transmenrane domain prediction of *SiPIN* genes by TMMOD (Hidden Markov Model for transmembrane protein topology prediction). *SiPIN1a*, *SiPIN1b*, *SiPIN2*, *SiPIN4a*, *SiPIN4b*, *SiPIN4c* and *SiPIN4d* contains long transmembrane domain where as *SiPIN5a*, *SiPIN5b*, *SiPIN5c*, *SiPIN5d* and *SiPIN8* contains short transmembrane domain.

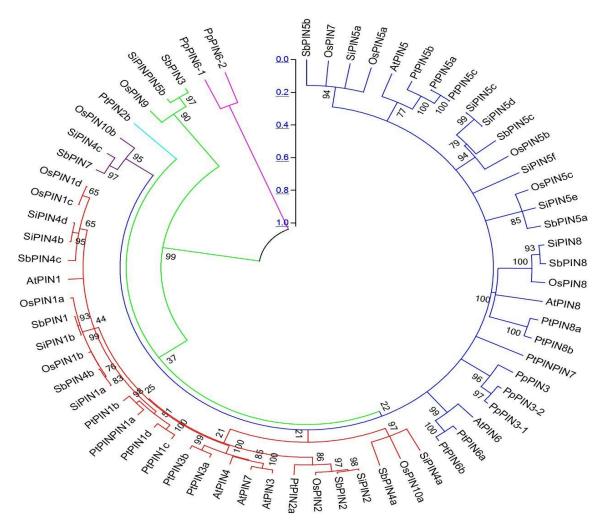


Figure 3. Phylogenetic tree of *SiPIN* genes with orthologous PIN genes of *Arabidopsis thaliana* (*AtPIN*), *Oryza sativa* (*OsPIN*), *Sorghum bicolor* (*SbPIN*), *Populus trichocarpa* (*PtPIN*) and *Physcomitrella patens* (*PpPIN*). Statistical method used to construct the phylogenetic tree was neighbor joining method; test phylogeny-boot strap method; no. of boot strap replication -500; substitution type-amino acids and model used was Jones-Taylor-Thornton (JTT).

SiPIN1b clustered with OsPIN1; SiPIN4b and SiPIN4d clustered with SbPIN4c; SiPIN4a clustered with SbPIN4a; SiPIN4c clustered with SbPIN7 and OsPIN10b; SiPIN2 clustered with ObPIN2 and OsPIN2; SiPIN8 clustered with SbPIN8 and OsPIN8. In group II, SiPIN5a clustered with SbPIN5b and OsPIN5b; SiPIN5c clustered with SbPIN5c and OsPIN5c; SiPIN5d clustered with SbPIN5a and OsPIN5c: SiPIN5b clustered with SbPIN3 and OsPIN9. In group III, AtPIN6 clustered with PtPIN6. There is no gene of Setaria or any other grass in this cluster, showing diversification of PIN genes. Cluster analysis reflects, S. italica PIN genes are much closer to PIN genes of grasses Sorghum bicolor and Oryza sativa. Multiple alignment of amino acid sequences shows conserved N and C-terminal domains (Supplementary Figure 1). The N-terminal region shows a conserved S-P/T-P motif, a potential target phosphorylation site for mitogen activated protein kinases (MAPK) (Sinha et al., 2011). The central hydrophilic loop is dynamic in nature and differs from each other in terms of sequence homology, but some *PIN* genes are conserved in this dynamic region with a T-P-R motif (Supplementary Figure 1). The T-P-R motif is a target phosphorylation site of mitogen activated protein kinase 3 and mitogen activated protein kinase 6 (Sorensson et al., 2012). The T-P-R motif is conserved only in the case of long transmembrane auxin efflux carrier domains. This shows that, although the central hydrophilic loop is diverse in nature, its phosphorylation events are conserved to carry out specific function suggesting that evolution of protein phosphorylation is conserved.

The *PIN*-Formed (*PIN*) proteins are a plant-specific family of transmembrane proteins that transport the phytohormone auxin as substrate molecule. There is very

limited data available which suggests auxin is a signaling molecule of ancient origin. The PIN gene family is found only in genomes of land plants. They act as regulator and play key roles in developmental process including embryogenesis, morphogenesis and organogenesis (Krecek et al., 2009b). The number of PIN genes present in S. italica (12) is equal to that of rice (12) and more than that of Arabidopsis (8) suggesting that the presence of more PIN genes may have some extra role in development and morphogenesis. The predicted structure of a PIN protein is similar to the structure of membrane transport proteins that use the electrochemical gradient across the membrane to transport molecules. All the identified PIN proteins have two hydrophobic domains with cytoplasmic orientation. The transmembrane helices of hydrophobic domains are highly conserved in their amino acid sequence. But substantial differences are present between the long and short PINs. The hydrophobic domains of all long PIN proteins contain the amino acids at invariant position, but these positions are not invariants in short PINs. The presence of invariant amino acid sequences in long PINs may play major roles, which has not been retained in short PIN. The loop between the transmembrane helices being present within the hydrophobic domain exhibits dynamic variability in size and sequence.

Conclusion

S. italica popularly known as foxtail millet is one of the best studied millet species in the world. The genome sequencing project of this plant is going to be completed in the near future. This will open the door for progressing research of this plant at the molecular level. Auxin efflux carrier genes identified in this report will help to understand the role of auxin signaling and its role in growth, development as well as response to different biotic and abiotic stresses. Phylogenetic analysis shows that auxin efflux carrier genes in species of grass family are conserved.

ACKNOWLEDGEMENT

These data were produced from Plant genome database (http://www.plantgdb.org/) and the US Department of Energy Joint Genome Institute (http://www.jgi.doe.gov) in collaboration with user community. Authors are highly thankful to the plant genome database and Joint Genome Institute.

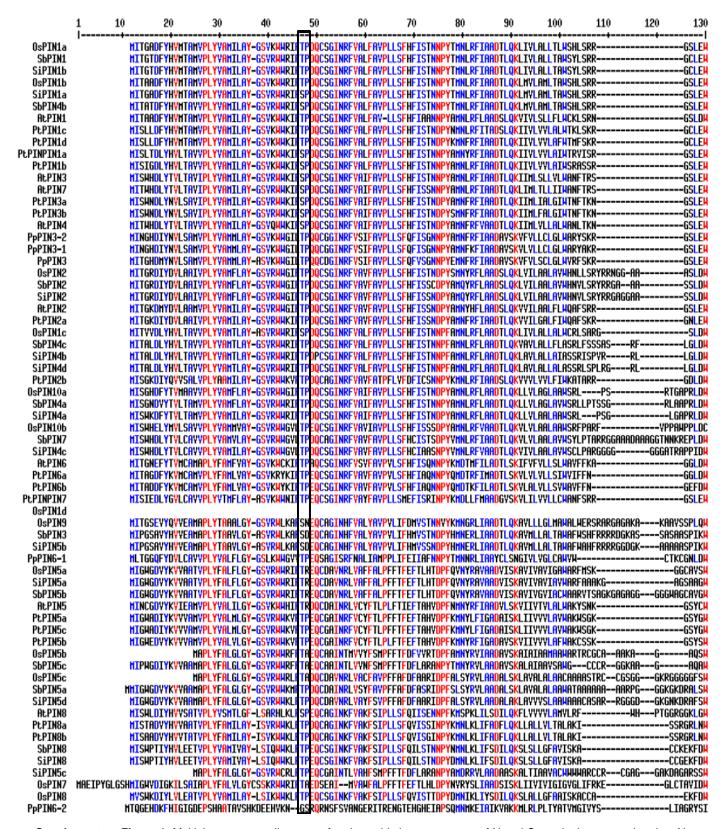
REFERENCES

Altschul SF, Madden TL, Schaffer AA, Zhang JH, Zhang Z, Miller W, Lipman DJ (1997) Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. Nucleic Acids Res. 25:3389-3402.

- Benkova E, Michniewicz M, Sauer M, Teichmann T, Seifertova D, Jurgens G, Friml J (2003) Local, efflux-dependent auxin gradients as a common module for plant organ formation. Cell 115:591-602
- Bennetzen JL, Schmutz J, Wang H, Percifield R, Hawkins J, Pontaroli AC, Estep M, Feng L, Vaughn JN, Grimwood J, Jenkins J, Barry K, Lindquist E, Hellsten U, Deshpande S, Wang X, Wu X, Mitros T, Triplett J, Yang X, Ye C-Y, Mauro-Herrera M, Wang L, Li P, Sharma M, Sharma R, Ronald PC, Panaud O, Kellogg EA, Brutnell TP, Doust AN, Tuskan GA, Rokhsar D, Devos KM (2012) Reference genome sequence of the model plant Setaria. Nat. Biotechnol. 30:555-+
- Carraro N, Forestan C, Canova S, Traas J, Varotto S (2006) ZmPIN1a and ZmPIN1b encode two novel putative candidates for polar auxin transport and plant architecture determination of maize. Plant Physiol. 142:254-264
- Devos KM, Wang ZM, Beales J, Sasaki Y, Gale MD (1998) Comparative genetic maps of foxtail millet (Setaria italica) and rice (Oryza sativa). Theor. Appl. Genet. 96:63-68
- Dong QF, Schlueter SD, Brendel V (2004) PlantGDB, plant genome database and analysis tools. Nucleic Acids Res. 32:D354-D359
- Duvick J, Fu A, Muppirala U, Sabharwal M, Wilkerson MD, Lawrence CJ, Lushbough C, Brendel V (2008) PlantGDB: a resource for comparative plant genomics. Nucleic Acids Res. 36:D959-D965
- Forestan C, Farinati Ś, Varotto S (2012) The Maize PIN Gene Family of Auxin Transporters. Front. Plant Sci. 3: 16
- Forestan C, Varotto S (2010) PIN1 auxin efflux carriers localization studies in Zea mays. Plant Signal. Behav. 5: 436-439
- Forestan C, Varotto S (2012) The role of PIN auxin efflux carriers in polar auxin transport and accumulation and their effect on shaping maize development. Mol. Plant 5:787-798
- Friml J (2010) Subcellular trafficking of PIN auxin efflux carriers in auxin transport. Europ. J. Cell Biol. 89:231-235
- Goodstein DM, Shu S, Howson R, Neupane R, Hayes RD, Fazo J, Mitros T, Dirks W, Hellsten U, Putnam N, Rokhsar DS (2012) Phytozome: a comparative platform for green plant genomics. Nucleic Acids Res. 40:D1178-D1186
- Kahsay RY, Gao G, Liao L (2005) An improved hidden Markov model for transmembrane protein detection and topology prediction and its applications to complete genomes. Bioinformatics 21:1853-1858
- Kahsay RY, Liao L, Gao G (2004) An improved hidden Markov model for transmembrane topology prediction. In: Khoshgoftaar TM (ed) lctai 2004: 16th leee Internationalconference on Tools with Artificial Intelligence, Proceedings. pp. 634-639
- Krecek P, Skupa P, Libus J, Naramoto S, Tejos R, Friml J, Zazimalova E (2009a) The PIN-FORMED (PIN) protein family of auxin transporters. Genome Biol. 10:249
- Krecek P, Skupa P, Libus J, Naramoto S, Tejos R, Friml J, Zazimalova E (2009b) The PIN-FORMED (PIN) protein family of auxin transporters. Genome Biol. 10:249
- Michniewicz M, Zago MK, Abas L, Weijers D, Schweighofer A, Meskiene I, Heisler MG, Ohno C, Zhang J, Huang F, Schwab R, Weigel D, Meyerowitz EM, Luschnig C, Offringa R, Friml J (2007) Antagonistic regulation of PIN phosphorylation by PP2A and PINOID directs auxin flux. Cell 130:1044-1056
- Mravec J, Skupa P, Bailly A, Hoyerova K, Krecek P, Bielach A, Petrasek J, Zhang J, Gaykova V, Stierhof YD, Dobrev PI, Schwarzerova K, Rolcik J, Seifertova D, Luschnig C, Benkova E, Zazimalova E, Geisler M, Friml J (2009) Subcellular homeostasis of phytohormone auxin is mediated by the ER-localized PIN5 transporter. Nature 459:1136-1140.
- Ouyang S, Zhu W, Hamilton J, Lin H, Campbell M, Childs K, Thibaud-Nissen F, Malek RL, Lee Y, Zheng L, Orvis J, Haas B, Wortman J, Buell CR (2007) The TIGR Rice Genome Annotation Resource: Improvements and new features. Nucleic Acids Res. 35:D883-D887
- Paponov IA, Teale WD, Trebar M, Blilou I, Palme K (2005) The PIN auxin efflux facilitators: evolutionary and functional perspectives. Trends Plant Sci. 10:170-177
- Reinhardt D, Mandel T, Kuhlemeier C (2000) Auxin regulates the initiation and radial position of plant lateral organs. Plant Cell 12:507-518
- Shen C, Bai Y, Wang S, Zhang S, Wu Y, Chen M, Jiang D, Qi Y (2010) Expression profile of PIN, AUX/LAX and PGP auxin transporter gene families in *Sorghum bicolor* under phytohormone and abiotic stress.

- FEBS J. 277:2954-2969
- Sinha AK, Jaggi M, Raghuram B, Tuteja N (2011) Mitogen-activated protein kinase signaling in plants under abiotic stress. Plant Signal. Behav. 6:196-203
- Sorensson C, Lenman M, Veide-Vilg J, Schopper S, Ljungdahl T, Grotli M, Tamas MJ, Peck SC, Andreasson E (2012) Determination of primary sequence specificity of Arabidopsis MAPKs MPK3 and MPK6 leads to identification of new substrates. Biochem. J. 446:271-278
- Wabnik K, Kleine-Vehn J, Balla J, Sauer M, Naramoto S, Reinohl V, Merks RM, Govaerts W, Friml J (2010) Emergence of tissue polarization from synergy of intracellular and extracellular auxin signaling. Mol. Syst. Biol. 6:447
- Wabnik K, Kleine-Vehn J, Govaerts W, Friml J (2011) Prototype cell-tocell auxin transport mechanism by intracellular auxin compartmentalization. Trends Plant Sci. 16:468-475

- Wang J-R, Hu H, Wang G-H, Li J, Chen J-Y, Wu P (2009). Expression of PIN Genes in Rice (*Oryza sativa* L.): Tissue specificity and regulation by hormones. Mol. Plant 2:823-831
- Wang S, Bai Y, Shen C, Wu Y, Zhang S, Jiang D, Guilfoyle TJ, Chen M, Qi Y (2010) Auxin-related gene families in abiotic stress response in Sorghum bicolor. Funct. Integr. Genomics 10:533-546
- Xu J, Scheres B (2005) Cell polarity: ROPing the ends together. Curr. Opin. Plant Biol. 8:613-618
- Zazimalova E, Krecek P, Skupa P, Hoyerova K, Petrasek J (2007) Polar transport of the plant hormone auxin the role of PIN-FORMED (PIN) proteins. Cell. Mol. Life Sci. 64:1621-1637



Supplementary Figure 1. Multiple sequence alignment of amino acid shows presence of N and C-terminal conserved region. N-terminal region shows presence of conserved S-P/T-P motif, target phosphorylation site of mitogen activated protein kinase. The central hydrophilic region is very dynamic in nature and differ from each other, but still contains conserved S-P-R/T-P-R motif, target phosphorylation site of mitogen activated protein kinase 3 and mitogen activated protein kinase 6. S-P-R/T-P-R motifs are conserved only in long transmembrane domain containing PIN gene.

	131	140	150	160	170	180	190	200	210	220	230	240	250	260
OsPIN1a	TITLE	SLSTLPNTI	VHGIPL	LKGMYGEFS	GSLHVQI	VVLQCIINYT	LHLFHFE	YRGARILITEQFF	PDTAG-A	-IRSIVVDADVV	SL-DGRRDMI	ETEREVK	EDGKIHVT	
SbPIN1	TITLE	SLSTLPNTI	YMGIPL	LKGMYGDFS	GSLMVQI	YYLQCIINYT	LHLFHFE	YRGARILITEQFF	PDTAG-A	-IASIVVDPDVV	SL-DGRNDAI	ETEAEYK	EDGKIHYT	VRRSNAS
								YRGARILITEOFF						
OsPIN1b SiPIN1a								YRGARMLITEQFF YRGARMLITEQFF						
SbPIN4b								YRGARLLITEQFI						
ALPIN1								YRGAKLLISEQFF						
PtPIN1c	TITLE	SLSTLPNTI	YHGIPL	LKG <mark>myg</mark> dys	GSLHVQV	YYLQCIIHYT	LHLFHFE	YRGAKLLISEQFF	PDTAG-S	-IVSIHVDSDIM	SL-DGRQP-L	ETEAAIK	EDGKLHYT'	YRKSNAS
								YRGAKLLISEQFF						
tPINPIN1a								YRGARILIGEQFF						
PtPIN1b AtPIN3								YRGARILIGEQFF Frgakmlimeqff						
ALPIN7								Y <mark>rga</mark> kilineqfi						
PtPIN3a								YRGAKHLIHEQFF						
PtPIN3b								Y <mark>rga</mark> khliheoff						
ALPIN4								Y <mark>rga</mark> kllimeqfi						
								YRAAKILIMQQFF						
PpPIN3-1 PpPIN3								YRAAKILIMQQFF Yraarilimhrff						
								YRGAKALISEQFF						
SbPIN2								YRGAKALISEQFI						
SiPIN2								YRGAKALISEQFF						
								F <mark>rga</mark> kllise q ff						
PtPIN2a								YRGARLLISEQFF						
OsPIN1c SbPIN4c								YRGARLLYMEQFF Yraaralyldoff						
SiPIN4b								YRAARALYHDQFI						
SiPIN4d								YRAARAL YMDQFI						
PtPIN2b	TITLE	SLSTLPNTI	VHGVPL	LKSMYGEFT	SPLHIQV	CFHQSYLHYT	LLLSHFE	Y <mark>rga</mark> krlyagoff	PETAA-S	-ISSFKYDSAYY	SLGGHEPL	.ETDAEID	DDGKLRYY	VRRSSAT
OsPIN10a								F <mark>raa</mark> rmliadqff						
SbPIN4a								FRAARHLIADQFF						
SiPIN4a Naptwink	STILL	SYSTLPNII	VMCTD	THULPLH	GSLMYQY	TALMCTTHAL		FRAARMLIADOFF	7USHHH	-TH2FKAFLDAA -TOOMNADDAA	LEGGO	1E 1 EHEYH==	EUGKLHYT	YKK55Y5
SbPIN7								FRAARVLIAGQFP						
SiPIN4c	SITLE	SLSTLPNT	LIHGIPL	LVAHYGRYS	GDLLYQY	VVLQCIVAYT	LLLYLFE	F raa raliagocp	DTAAA	-IADVRVDPDVVS	LAGSQAI	EAQAEYA	PDGRYRLYY	RRSTSY
ALPING								LRAARLLIRAEFP						
PtPINGa PtPINGb								Y <mark>RAAT</mark> LLIKTQFP YRAATLLIKAQFP						
PtPINPIN7	AITLE	SLSTLPNT	LYMGIPL	LKSMYGDDK	EGLMIQY	YYLQCIINYT	LLLFLFE	YREARLAMLKNEK	655555	SFSNSERSKG	SFKGYGEYL	GSRGSNA	DEYYNYIY	STPTSQ
OsPIN1d								MEQFP	DTAAS	-IYSFRYDSDYY <mark>s</mark>	LAGGGGGAAI	ELQAEAEY	-GDDGKMRYTY	RKSTSS
OsPIN9	VITCE	SVASLPNT	IIMGVPL	LNGMYGPYS	KDLMKQI	VVMQFCIHYN	VIIFLYE	YMAARRSASAPPP FMAARDGTAKISP	ASS			EGSAKIS	PSSPYKAA	AAAADT
SbPIN3 SiPIN5b	VYINE	SYNSLPNI.	TTUGVPI TTUGVPL	LUUNTUSYS I NGHYGSVS	GGI MKOT	VVHOECTUYN VVHOECTUYN	MATEL AEI	FMAARDGSAKISP	'HIY 'V			22444444 5000000	EKBUBTI	DENGGS GENGGS
PoPTNG-1	VTTLE	OL CUMPNT	TTUGTPU	SPI VSVTE		TETGOVI UI E	PTI EI YEI	KEABKACOBOAC	SVA			OVODDGP	PSI HSGET	TESAAE
OsPTN5a	STTS	CI STI TNS	VVGVPH	ARAHYGEUA	001 VVOI	SVEOATVUI T	I I I FVI F	V <mark>RKA</mark> ATGH						YV
SiPTN5a	STTSF	SLSTLTNS	VVGVP	ARAMYGFUA	DDI YVDI	SVEDATVII T	I I I FVI F'	V <mark>rka</mark> atgm						YV
SPLINSP SPLINSP	STIG	SLSTLINS	LYYGYPH	HKHMYGENH	VOLVYOC	SYFUHIYALI SVEQOTVULT	LLLFYLE.	IRKAAIGH Frkagfs						YY
Pt.PTN5a	STTSF	SI CTI TNS	VI GVP	TKAHYGPTA	VDI VVOS	SVTOSTTULT	HILL VI FI	FRRSGI GV						55
PtPIN5c	STTS	SI CTI THS	VI GVP	TKANYGPTA	VDI VVOS	SVTOSTTULT	HIII VI FI	FRRSGI GV						55
PtPIN5b	CTTC	CI CTI TNC	VUCUPT	TKOHYCDOO	VDI VVOS	T INTTANTV2:	I I I I VI E	EDDTGI GE						22
OsPIN5b	SITGE	SLAALNNT	LYYGYPL	LDAMYGRHA	QDLYYQI	AVVQSMVNFP	LLLMAFE	LRKAHYYG Lrkahhygipsaq			666			YG
SbPIN5c OsPIN5c	CTTG	SI ATI NNT	VVGVP	I DAMACKAB	RNI TVOT	SVVOTTVYFP	IIII AFF'	VRRATTAA						88
SbPIN5a	CTTVE	SLATI NAT	VVGVP	LDAMYGKHA	RDI TYOT	SVVOFTVYFP	ALLLALE	ARRAGKHA						AA
SiPIN5d	CTTGE	SI AAI NNT	VVGVP	I DAMYGKUA	RDI TVOT	SVVOTTVYFP	IIII TEE	ARRAUGAG						KR
ALPINS	VITGL	SISYLPHT	LILGHPI	LSAIYGDEA	ASILEQI	YYLQSLIHYT	ILLFLFE	LNAARALPSSGA- Lnatkeatvaps-		<u>S</u>	LEH			TG
PtPIN8a PtPIN8b	TTIEL	SLSTLPNI SLSTLPNI	LTLULPL	LKHNTUHEH=== RAMHGAFA===	EPLLSUI	VALTICATURA NATA 150 IUA	LLLTLT E 	LNATKEATYAPS- LNATKEATYAPS-						EI
SEPTINS	LITTG	SI STI PNT	TVGTP	I KGMYGDFA	VKI I SOT	vai osi tuyt	THEFF	FRAARGHATTTS-		<u>9</u>				FT
SiPIN8	LITGE	SLSTLPNT	IVGIP	LKGMYGNEA	YKLLSOT	VALOSLIMYT	LLLFLFE	F <mark>raa</mark> rglattts-						GT
SiPIN5c	CYTGE	SLAAFNNT	LVVGVPL	LDAMYGAHA	RDLYVOL	AVYOAL VHYP	LLLLGFE	L <mark>rka</mark> hyvrngias	AAAGDGQAA	AAGGGRVEPAVS <mark>S</mark>	QAK			KA
OsPIN7 OsPIN8								Y <mark>rka</mark> fysdahdes L <mark>raan</mark> ghatttss						
PpPIN6-2			LITUIFL	ENGIT OCUIT	unilogi	TTEQUELMIT	LLLTLTC	LININGIINT I I 33	FII GETWU.	CTUI KMITKUL	TTONOTKIIL	JOUGLUI III ()	SCULF INLUIT	VITHUU
. F. 200 E														

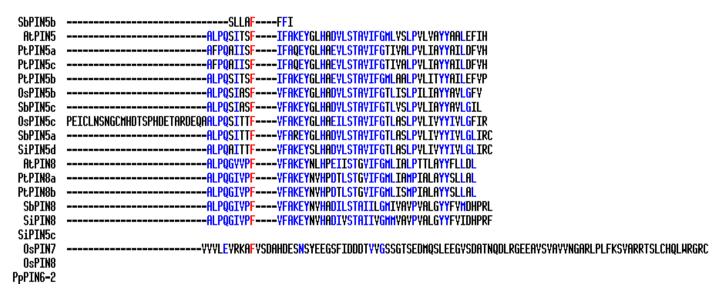
	904	970	280	200		300		210	320		330	-	40	350	360		370	380	390
	261 	270 +		290		+-		310			+-		-+	+	+-		+	+	
OsPIN1a	RSDYY	SRRSHGFS-	S]	TPRE	SNLTNA	EIYSL	.QSSRNI	TPR	GSS-FNHTDF	YSHVG-		RSSNFAA	G	DAFGY	RTGAT	PRPSNY	EEDAA-	APNK	AGS
SbPIN1 SiPIN1b	KPDTI	5KK5NuF5• coocucec.	S1 S1	TDD	SNL INH	FT12F	.Ų55КМ песры	TDD	255-FNHTUF	150YU- Venue.		KSSNEHH DeenEag	ն	DHF6L	KIGHI DTGG1	PKP5NT	'EEEH Ų= '	bbKH 	NNK
OsPIN1b			S1																
SiPIN1a	RSDIY	SRRSMGFS-	S1	TPR	SNLTNA	EIYSL	.OSSRNI	TPR	GSS-FNHNDF	YSMVG-		RSSNFGA	A	DAFGI	RTGA1	PRPSNY	'EDDAS-	KP-K	YPLPVVN
SbPIN4b	RSDIY	SRRSHGFS-	S]	TPR	SNLTNA	EIYSL	.QSSRNI	TPR	GSS-FNHNDF'	YSHVG-		RSSNFGA	A	DAFGI	RTGAT	PRPSNY	'EDDAA-	KPAK	YPLPYYN
Atpini	RSDIY	SRRSQGLS-	A1 S1	-PRP	SNLTNA	EIYSL	.QSSRNI	TPR	GSS-FNHTDF1	YSMMA-	SGGE	RNSNFGP	G	EAY	FGSKGP1	PRPSNY	EEDGGP1	AKPTAAGTAA	GAGRFHY
PtPIN1c PtPIN1d	KOUTE	SRRSQGI S.	S	TPR	SNL INN SNI TNA	ETACI	NASSU. Nassu	TPR	GSS-ENHTDE	TSNNH- YSMMA.	AG	KNSNFGA RNSNFGA	S	DYTGL DVYGI	SASRGP1	PRPSNE	FFFNG=	65 69	NKPRFHH
PtPINPIN1a	RSDYE	SRHSHGL N-	SGL St	TPR	SNLTNA	ETYSI	OSSRNI	TPR	ASS-ENHTDE	YSK-N-					ASNAS	SPRHSNE	SNLOF-	D	FESGGL G
PtPIN1b	RSEVE	SHMSHGLN-	SGLSL	TPRP	SNLTNA	EIYSL	.QSSRNI	TPR	RSS-FNHTDF'	YSHVN-					GKNAS	SPRHSNF	THLQF-	D	EESGGLG
AtPIN3																			
AtPIN7 PtPIN3a																			
PtPIN3b																			
ALPIN4	RRSLH	H		TPRP	SNLTGA	EIYSL	.55	TPR	GSN-FNHSDF'	YSYMGF	PG6	RLSNFGP	A	DLYSY	QSSRGP1	PRPSNF	EENNA-	YK	YGFYNNT
PpPIN3-2			SSKAL																
PpPIN3-1 PpPIN3	RSQUIT	HSHHHS M P- UCONUCTO.	SSKAL SSKAL	TDD	SNL I GH GNI GNO	FT12L	HSSYNI Meeval	.IPK	USS-FNUGEFI ccc_cnoccn	HSMMSI POTMOL	(KS=P) JDD_D1	IKŲ5NFU I IDVONENT	S	DYYSL	USSKUP I Neededt	PKSSNE	NEENSKI	CANNIDGOI N NTH I HHKATN	MNSPKEH
OsPIN2	STTGG	GGGAARSG\	/Sraygasnah	TPR	SNLTGV	EIYSL	.OTSREI	TPR	ASS-FNOADF'	YAMFS0	IKD-FT ISKMAS	1863AFD1 1914		ASPM	AOHGGAG	GRAOGL	DEOYTH:	K	FASGK
SbPIN2	STTGG	GHGGARSG\	/GGYRPYGPSSAN	TPRO	SNLTGV	EIYSL	OTSREI	TPR	GSS-FNOSDFY	YAMENO	SKMA9	PL		AOP-	G	-ARAPGL	DEOVAN-	K	FASGK
SiPIN2	STTG-	-HGAARSG\	/GGYRPYGASSAI	TPR	SNLTGY	EIYSL	.QTSREI	TPR	GSS-FNQSDF	YAHENG	SKMAS	PL		AQP-	GSG	ARAPGL	DEQYAN:	K	FASGK
AtPIN2	SSMIS	SFNKSHGG0	GLNSSMI NSITSM-	TPRO	SNLTGY	EIYSY	QSSREI	TPR	RSS-FNQTDF'	YAMFNE	1SKAPS	PRHGYTN	SYGGAG	AGPGGDYYSL	QSSKGYT	PRTSNE	DEEAWK	TAKKAGRGGR	SMSGELY
PtPIN2a OsPIN1c	BSEAU	SFNKSHULM CSHGTOSH.	SQSM0	-PRY	ISNL TUY ISNI SGV	ETYSI	USSKEI	TPR	HSS-FNUTUF RSS-FNHAFFI	THNF-F FNTVG-	15KHH:	PKHUTIN	21.AG0T	GDYTSL	NGKOG	PKISNT	DEENLKI DEEKG=:		NU20FF
SbPIN4c	RSEAA	CSH-SHSH-	SOSMO	-PR*	SNLSGY	EIYSL	OSSRNI	TPR	GSS-FNHTDFI	FNIVG-					AGAKGGF	AAAAAAG	DEEKG-		
SiPIN4b	RSEAA	CSH-SHSH-	SQSHC	Q <mark>P</mark> R*	SNLSGV	EIYSL	.QSSRNI	TPR	GSS-FNHAEFI	FNIVG-					AKG	AGG	DEEKG-		
SiPIN4d	RSEAA	CSH-SHSH-	SQSHQ	QPR	SNLSGY	EIYSL	.QSSRNI	TPR	GSS-FNHAEFI	FNIYG-				DTVCI	-GAKG	AGG	DEEKG-		
PtPIN2b OsPIN10a	SSNES _DDCI	SKUKFUGAN I VT	IPYLSYHL	_DD	SNESSY CULTCO	ETVEL	QSS	TDD	HSS-YKUTULI CCN_CNUONCI	PNL INS FO	չի կ (Vecer	DDDTDQQ	U_DCCC	DITSL	USSKNSY CC_DCD1	LDDUGNE LDDUGNE	ILEEENK:	D_DVDDQTTT	COI NUNO
SbPIN4a	SRRSL	LHVTT		TPR	SNLTGA	ETYSE	.ssskni (SSSROI	ISPR	GSN-FNHADFI	FAAMAN	(VDGAF	PPPPPAG	GARASS	FGAAELYSMH	SS-RGP1	PROSNE	DERSASI	ARSSSRPAGA	VPSCHDA
SiPIN4a	-RRSH	LGYT		-PRP	SNLTGA	EIYS	ISSSRQI	ISPR	GSN-FNHADFI	FAh	IYDGAF	PPPTPAG	G-RGSS	FGAAEYFSHH	SS-RGP1	TPRQSNF	DEHSASI	arsskpaaaa	AYPSHDA
OsPIN10b	-SRRS	AAAA	f	TPRO	SNLTGV	EIYSI	SSSRNI	ITPR	GST-FTLADI		PGł	IQPPNS	ALRASS	FGAADLFSLH	SSSRQHT	PRPSSF	DEHAAA	RARASAT	VAPTNDL
SbPIN7 SiPIN4c	PSRRS _cppc	LAAAT	Af	TPRE	SNLTGY	EIYSh	ISSSRNI Jecepu	ITPR	6SSGIAHADI(GH	PAF	IAPLHGAS	SLRMSS CMDMCC	FGAADLFSLH ECOON ECLU	P1 cc_pout	PRPSSF	DEQAY-	RARSAAT	YAPSNDP
AFLING	VPNSV	HSSSLCI		TPR	SNL TUY ISNL SNA	FTFSV	'эээкні 'N	TPN	N-REFHAGAG	SGTI OF	YNGSK	INTTLNUN IETHECNG	011K1100	66F6FT	RPGI GAS	IPRRI SG	YASSD-	кикэппп	YNF310F
PtPIN6a	apdsa	LSSSICL		TPRP	SNLSNA	EYFSY	'S	·TPA	PLQEYHGYNGI	RFS	-HGPNN	IEIHLCNG	DL	G-FGYH	RSGTS	SPR-LSG	YASSD-		
PtPIN6b								·TPY											
PtPINPIN7 Ospin1d	ETTEN	YNTKY ceuetneu.	SQS	(APD) (ADD)	QQFRSM CNL CCV	YAAAY	06	TDD											
OsPIN9	NGNAV	cona i yon- AANR		PO	SMLSUY FVAVNT	FTTF	. WSSKNI IAASTAL	nrk	199-LMUUELI	LMTAG-							IDEEKG-		
SPIN3	STHEH	DHRS		HH	OVVVNI	EITEV	'AAAAYF	IST-											
SiPIN5b	TH	AFRS		QN	DVVVNT	FTTEV	'AAAAADE	ISP-											
PpPIN6-1	HGFEG	HESG	AGKDVE	IP	ALLTQG	EHNDF	RIGIGE)EP-											
USMINDA SIPTN5=	-AAFP	nnn PVK	DYEAAS	:nnuH :ANAA	MMMHU I PAAAAV	yyyHh VVVPL	MH==== MS====												
SbPIN5b	DYPDS	PYK	DYEAAA	ADAAS	AATYYY	γργγι	'AS												
ALPIN5	NISDY	QYD	NINIES	GKRE	T	\	'γγ												
PEPTNS6	SMPUK	ULE NI F		.GCVN	N	 1	75												
OsPIN5b	PAVHS	SSS	PPEK09	DVEN	NGAYYA	APGGG	iGG												
ShPTN5c	PVSSS	SASSSSPPF	PSPPPKKDAAAF	OVEN	NAFETA	AAAAA	iTT												
OsPIN5c	PPPPP	TGT	[)-DDD	YEDGAA	AAATA	1AA												
SPLINPA	HHYEP	HGD	[JYUES	ննննեն coccny	PET II	HH====												
ALPIN8	NDQEE	ANI	[DEPK	EEEDEE	EVAIV	'RT												
PtPIN8a	TGDQE	ALQ	[S	QYKEGG	EYHSR	:TT												
PtPIN8b	TGDI F	AI 0	F	A	OHKNDF	GVORE	?T												
SbPIN8	TNEAE	SGT		GPMQ	QRHEEG	QEKRY DOVC	'SA												
SIPINS	-MEUE	301 AAF	K('UI'NE INVEK	EKHEEU NAAAFD	KNKUY NAPPO	3H												
OsPIN7	GARLP	LFKSYAR	KLAC	:NPNL	HASYIG	ISHAC	ISHRSH	łLTL	PPALEGSYQII	HSR									
OsPIN8	AKRHL	LALAIRFFL	_GPALHGHSSYA]	GHRG	YLLKIA	IVQAF	LPQGIV	/PFV	FAKEYNYQAD:	[LSTA]	CIVGHI	IYAYPYAL	AYYFAH	IIPAIK					
PpPIN6-2																			

	391	400	410	420	430	440	450	460	470	480	490	500	510	520
OsPIN1a		KYGOYP	AP-NPA						+ MAAPPK-	PKKAAN		-GOAKGEDGKI	LHMEVHSSSE	SPVSD=
SbPIN1		YGGÖYP	AP-NPA						-HAAOPHPTKO	GLKKAAAN		-GOAKGEDGKI	LHMFYNSSSA	SPYSD-
SiPIN1b		YG-OYP	AP-NPA						-HAAPAK-	GOPKKAAN		-GOAKGENGKI	I HMEVUSSSA	SPVSD-
OsPIN1b	AAP	MAGHYP	AP-NPA						-VSSAPK	GAKKAATN		-GOAKGED	LHMEVHSSSE	SPVSD-
SiPIN1a	AAP	GAGHYP	AP-NPA						YAAAPK	GAKKAATN		-GQAKGED	LHMFYNSSSA	SPYSD-
SbPIN4b	AAAPYI	PGAAAGHYP	AP-NPA						YAAAPK	-KKAAGTN		-GQAKGED	LHMFYKSSSR	SPYSD-
ALPIN1	QSGGS	GGGGGAHYP	AP-NPG						-MFSPNTGGGG	GTAAKGN	APYY	GGKRQDGNGRE	LHMFYHSSSA	SPYSD-
PtPIN1c	YHAP-	GGATHYP	AP-NPG						-HFSPTTAASK	GYSANANNTA	IAAAAKKPNG	QAQQKAEDGRE	LHMFYHSSSA	SPYSD-
PtPIN1d	YHAP-	GGATHYP	AP-NPG						-HFSPTTA-SK	GYAANANN	-AAAKKPNG	QAQQKAEDGRI	LHMFYNSSSA	SPYSD-
PtPINPIN1a	VFGNV	PRANGSAYP	TPPNAG						-IFSP	GGKKKANG		TENGKE	LHMFYKSSSR	SPYSE-
PtPIN1b	YFGNY	PRANGSAYP	APPNAG						-IFSP	GGKKKANG		AENGKE	LHMFYNSSSA	SPYSE=
ALPIN3	FGYYP	GGG-AGSYP	APNPEF						-SSTTTSTANK	SYNKNPKDYN	ITNQQTTLPT	GGKSNSHDAKE	LHMFYKSSNG	SPYSDR
AŁPIN7	FGYYP	GGA-PGSYP	APNPEF						STGNK	TGSKAPKENH	HHY	-GKSNSNDAKE	LHMFYNGSNG	SPYSDR
PtPIN3a	FGFYP	AQTYPTSYP	APNPEF						-ASTYTTKT	AKNQQQQN		SKANHDAKE	LHMFYNSSSA	SPYSEG
PtPIN3b	FGFYP	AQTYPTSYP	APNPEL						-ASTITSKT stgtgys1	TKNQQQQNHQ	QQLLQPQPQ	QNSKYNHDAKE	LHMFYNSSSA	SPYSEG
Atpina	NSSVE	PAAGSYF	APNPEF	UCNDDICUCC	L CTL CTDCM	PDCDTTVDCC	OTOTOL UTD	CCTCUT_OT	STGTGYS1 PLSSSLNTQIVN	KPNKIPKEN	QQLQE	KDSKASHDAKI	LHMFVHSSSF	SPVSD-
PpPIN3-2 PpPIN3-1	PPLYR	RNGMGARMFT	PRPGLGGLP	/PGTDCTGHGT	LSTLGAPGMO	3PDGRTIYPGS 3PDGRTIYPGS	QTAINILTL	.GGAANYNATI	PESTAYNTQIYI	(PYYSPQASQ)	CAKKYKDPKA	S-PRADEDAKI	ELHNFYASHNI ELHMFYASANI	ISPYSE-
PpPIN3 OsPIN2	PPLYR	RNGSGGRLFI	ARSDLGGYGI	ALSFEPAAHS-	H	PDGRTIYPG-	ITVVT-	NSVAAY	APSTAVNTQIVI VPASGVSTHIII MMPA	PVFSPLVSQ\	/AKKYNDPRA	SIPKTDEEAKI	LHMFVSSANF	TSVSE-
SbPIN2	-66		PAPNPG						HHPP	-DDANEI CCCI	1CM	SNVI	TI UMEVUÇÇÇE	ICDUCE_
SiPIN2 AtPIN2	-QG	SDATAYF S-VPSYF	PAPNPG						HMPP FTGSTSGAS(-PRKKELGGSI	(SN :SG	eeneneeunki SNKI	ELHMFYASSSF Hnnfyusssf	ISPVSE-
PtPIN2a	NGG	SLYSSYF	PPNPH						FSGSTSGC	PKKKESGSG	H	PNKI	LHMFYHSSSF	SPYSE-
OsPIN1c SbPIN4c		AAGGGGI GAGGHSF	ISPQPYV POPOAVA									GKRKI AVAAKRKI)LHMFVASSSF 11 HMI VASSSF	ISPVSE- ISPVSE-
SiPIN4b		AGGGYSF	QPHAQA									AKRKI	DLHMFYASSSF	SPYSE-
SiPIN4d PtPIN2b		RKNGVAFF	GSPSCA							-VPOKEGGG		APAPNKI	11 HMEVUSSS1	SSNTS-
OsPIN10a	KELHM	(EVUSSSASE	VSEVSG						l l	PVFSGGGG		-GGAL DYGAKI	THMVTPADLE	ONN-
SbPIN4a SiPIN4a	KEI HH	(EVUSSSASE	PVSEVSG							PVFTGGAA		VNVGAKI	TPHVVHADI F	ONG-
OsPIN10b SbPIN7	KDTHM	(TEUSSGASA	ASEVTG						[PYFRSG		RI	TRRLYPSDAF	SIASS-
SiPIN4c	KOHPT	rfeussgasa	ARSEVSG							_PYFRGGE		-HRAKI	DVRRLVPSEAF	PYGMS-
ALPING PLPINGa		AYSLQF AYSL OF	TPRASN									FNELDY-NGN(FNEUDI TNATI	STPYNMKSPAF ITPENARSPVF	IGRIYR- IGKTSR-
PtPIN6b		AVSL OF	TPPACN									ENEL DI TNATI	ITPEUVPSPVA	ICKTAD"
PtPINPIN7 OsPIN1d		-AAGGGGHSF	OPVV6									DDKI	EYHLFIARCYO OLHMFYASSSF	CTSQG- ISPVSE-
OsPIN9														
SBPIN3 SiPIN5b														YDGAA- ESGAE-
PpPIN6-1 OsPIN5a														
SiPIN5a														
SbPIN5b AtPIN5														
PtPIN5a														
PtPIN5c PtPIN5b														
0sPIN5b														
SbPIN5c OsPIN5c														
SbPIN5a														
SiPIN5d AtPIN8														
PtPIN8a PtPIN8b														
SbPIN8														
SiPIN8 SiPIN5c														
OsPIN7														
0sPIN8 PpPIN6-2														
. F. ZIIO E														

	521	530	540	550	560	570	580	590	600	610	620	630	640	650
OsPIN1a	U	EGNG-REYNI	+	+ AVASPRKAI	-	VFRNC	ESEGNR	-GYAERDAEAGI	TEKSVAAAV.	SGEHGKPGI	 TPAPTA	-HPPTSUHTRI	TI THVURKI	TRNPNTY
SbPIN1								-GAAERDAEAGI						
SiPIN1b	Y	FGNGNAEYNI	DAAAYKEYRH	ayasprkyaai	DGRKERGE	DFAERDO	FSFGNR	-GGAERDAEAGI)EKAAY-	AGQGNA-GY	GA-PAA	MPPTSYMTRL	ILIHYARKI	IRNPHTY
OsPIN1b								-GYMDRDAEAGI						
SiPIN1a								-GAMDRDAEAGO						
SbPIN4b								-GAMDRDAEAGI						
ALPIN1 Plpin1c	-							-D-DDSKYLATI -G-LDRERNSHE	-					
PtPIN1d								-G-HDREHNNPE						
PtPINPIN1a								PNGYDRDGPYLS						
PtPIN1b								PNGVDRDGPVLS						
ALPIN3								EEEAERPKDAEN						
ALPIN7								-EESERYKEYPI						
PtPIN3a								GDDDQREKEGP1						
PŁPIN3b AŁPIN4								GEDDQREKEGP1 Gegerejekata						
PpPIN3-2								TKTEDTDKDGL						
PpPIN3-1								LKLEDLDKDGPF						
PpPIN3								LKYEDYDNNGSI						
OsPIN2	ANL	rnaynhaas	TDFASAPPPA	AVPYGGATPKI	GYSGSYTPAAI	<nggge< th=""><th>LEIEDGLK</th><th>SPAAGLAAKFP\</th><th>/SGSPYVAPR</th><th>KKGGGADYP</th><th>GLAEAAH</th><th>PHPPTSYMTRL</th><th>ILIHYARKI</th><th>IRNPNTY</th></nggge<>	LEIEDGLK	SPAAGLAAKFP\	/SGSPYVAPR	KKGGGADYP	GLAEAAH	PHPPTSYMTRL	ILIHYARKI	IRNPNTY
SbPIN2								SPATGLAAKFP\						
SiPIN2								SPATGLGAKFPY						
ALPIN2 PLPIN2a								GKSP1 CKFP0						
OsPIN1c								GNKNEKDGF						
SbPIN4c								TKNGNGADKGGF						
SiPIN4b								TKNGNGADKGGF						
SiPIN4d			RA	AYHYI	FGAGA-EHGD'	/LAKGTQAYDE	YGRDDFSR	TKNGNGADKGGF	PTLSKFGSNS	TAQLYPKDD	GEGRPAA	MPPASYMTRL	ILIHYARKI	IRNPNTY
PtPIN2b								DPFNGADPQEDI						
OsPIN10a								TAELHPK\						
SbPIN4a								CSTTAAEYRYKI						
SiPIN4a OsPIN10b								TAELRYKI kteootta						
SbPIN7								TEEERK						
SiPIN4c								ADQRQNYI						
ALPIN6			Q	SSPKMI	HAESGQRHAAI	KDINGSYPEKE	ISFRDALK	AAPQATAAGGGA	RSHEEGAAGK	DTTPYAAIG	KQE	MPSAIVHMRL	ILTYYGRKI	SRNPNTY
PtPIN6a								MP						
PtPIN6b			Q					MP						
PtPINPIN7 Ospin1d			D00					SMKRGVESEKTE NEKDGF						
OSPIN10			KNN		. aaaannnan.	nkanyntuc KFVS9	:T3FUNK :GEVAP	\rugh	TLSKLUSNS JEFFFASAPA			PSMKHVIN		
SPIN3						AAKEL	MSADDAGD	GDADANNNKAA	/DVVPPPTPV	0		POVPSVMHVVA	HATKK	LOIPNTY
SiPIN5b						EAAAA	AAAAEE	\	/AAYPPPAPA	0		QYPPYMHYYAI	MAYKKI	LQIPNTY
PpPIN6-1						KDEEH	IYKNGSRQR	NSFSYANGESGT	FRENGTEHGH	É				
OsPIN5a												GKPSLAAI	TAKAAUHK	LARNPNTY
SiPIN5a												GKPSLAH	LAKAAHHK	LSRNPNTY
SbPIN5b Alpin5												CERCELE	LAKAAHHKI	LOKNPHIT
PtPIN5a												SRPAFUC	HKTVGVK	AMNPNSY
PtPIN5c												SRPAFACI	LMKTYGYK	AMNPNSY
PtPIN5b												SRPAFACI	LMRTVAVK	.AMNPNSY
OsPIN5b												YRLPFHA	TARTYGLKI	LARNPNYY
SbPIN5c												IRLHPI	LYRTYGLK	AGNPNYY
OsPIN5c												ARRSLAPI	LYRHYHLK	VARNPNVY
SbPIN5a SiPIN5d												Q555FMPI	LYKHYGIKI	NUNDUNA
AFLINS 2151430												RSVGTHK	LI I KUND <mark>KI</mark>	TTNPNTY
PtPIN8a												RKVKAHL	VLLTVGKK	HSNPNFY
PtPIN8b												RKYKAHY	TLLTVGRK	HSNPNFY
SbPIN8												RCYGAFQI	FLLVVGKK	VTNPNHY
SiPIN8												RCYSAFCI	FLLYYGRK	VINPNHY
SiPIN5c								UDUDI TEDUET						
OsPIN7 OsPIN8					ITHIC	FOTHTHCKKLF	:TSKTUUHK	VPVRLICRHGTF	IMKUH I KUHY	KEKNTUTUH	ANTRKTEZHI	LILLTEHLHLGY	L92KKHHŲ.	TL LHE <mark>n</mark> 2F
PpPIN6-2														
i bi Tiin-c														

	651	660	670	680	690	700	710	720	730	740	750	760	770	780
OsPIN1a	SSLIGI	IHSL	YCFRANFEMPA	IILK <mark>SI</mark> SI	LSDAGLGH	AMFSLGLFHALI	+- Q		TFAMAYRE	LTGPAYMAI	ASIAYGLRGT	LLHYAIYQA-		
SbPIN1	SSLIG	YHSL	YCFRANFEMPA	IILKSISI	LSDAGLGH	AMFSLGLFMAL	Q	-PRIIACGNKYA	TFAMAYRE	LTGPAYMAI	ASLAYGLRGT	LLHYAIYQA-		
SiPIN1b OsPIN1b	SSLIG	IHSL	VCFRANFEMPR	IILKSISI	LSDAGLGM	AMFSLGLFMALI	Q	-PRIIACGNKVA -PHIIACGNKVA	TFAMAYRF	LTGPAVMAI	AASFAYGLRGT	LLHVAIVQA-		
USPINID SiPIN1a	SSI TGI	THSI	YCERUNFENPA VCERUNFEMPA	TVLKSTST	LSDAGLGA LSDAGLGA	AMESLOLFIALI AMESLOLFAALI	ų Ո		I THNHYKE TEAMAVRE	LHUPHYNNI I TGPAVNAI	INSFNYULKU I IASFAVGI RGT			
SbPIN4b	SSLIG'	VIHSL	YCFRANFQMPA	IVLQSISI	LSDAGLGH	AMFSLGLFMALI	Q	-PRIIACGNKVA	TFAMAVRE	LTGPAVMAI	aasfayglrgt	LLHVAIVQA-		
ALPIN1	SSLF6:	ITHSL	ISFKUNIEMPR	LIAK <mark>SI</mark> SI	LSDAGLGM	AMFSL G LFMALI	N	-PRIIACGNRRA	AFAAAMRE	VYGPAYMLY	/ASYAYGLRGY	LLHYAIIQA-		
PtPIN1c PtPIN1d	SSLIG	IHSL	YSFRANYKMPL Veedudvoado	IIAKSISI	LSDAGLGM	AMFSL G LFMALI OMECL C LEMOLI	Q 0	-PRIIACGNSIA -PRIIACGNSVA	AFSMAYRF	LTGPAIMAI TCDOVMOI	AASIAYGIRGT Doctovci Dct	LLHIAIYQA		
PtPINPIN1a	SSLIG	THSI	YSEKHOLEMPO	TTAHSTST	LSDAGLGN	AMFSL G LFHALI	Q D	-PRIIACGNSYA	AFAMSVRF	LTGPAVNA	AASEAVGLRGV	LLHTATYOS-		
PtPIN1b	SSLIG	LTHSL	YSFKUNIMMPR	IIADSIAI	LSNAGLGM	amfsl g lfhali	Q	-PRIIACGNSIA	SFAMGYRF	LTGPAYMAI	ASFAYGLRGD	LLRIAIYQA-		
ALPIN3								-PKLIACGNSVA						
AtPIN7 PtPIN3a	SSLIG CCL TC	H_ H_	YHFRAUYHMPK VOCDUUVCHDA	11005151	LSDAGLGA	AMESLGLEMALI OMEGLOLEMOLI	Q O	-PKLIACGNSTA -PKLIACGNSYA	TEOMOVDE	FTGPHYMHY TTCDOVMOI	/HHMHIGLRGD	LLKYHIYQH		
PtPIN3b	SSLIG	IHSL	IAFRUHVENPK	TIKQSISI	LSDAGLGN	AMFSL G LFHALI	Q	-PKIIACGNSYA	TFAMAVRE	LTGPAVNA	AASIAYGLRGT	LLHVAIVOA-		
AtPIN4	SSLIG	IHAL	ya <mark>yru</mark> hyampk	ILQQSISI	LSDAGLGH	AMFSLGLFMALI	Q	-PKIIACGNSYA	TFAMAYRE	ITGPAINA'	/AGI <mark>AIGL</mark> HGD	LLRIAIYQA-		
PpPIN3-2								-DRIVYCGRKHA						
PpPIN3-1								-DRIIYCGTKHA -Driiacgtkha						
PpPIN3 OsPIN2	SSI TA	VHSI	VSERUNTOMPS	TTK6STST	I SDAGI GH	AMESLALFNAN AMESLALFMALI	u ()	-PKIISCGKTYA	LFAMAVRF	I TGPAVTAI	INSTLYULKUY ATSTATGI RGV	3LMY31YQN- 11HVATVQA-		
SbPIN2	SSLIG	_Y <mark>H</mark> SL	VSFRFNIQMP9	IIKGSISI	LSDAGLGH	amfslglfmali	Q	-PKIISCGKRYA	TFAMAYRF	LTGPAVIA	ATSIAIGLRGV	LLHVAIVQA-		
SiPIN2	SSLYG	LYHAL	YSFRANIQMPS	IIKGSISI	LSDAGLGM	AMFSL G lfmali	Q	-PKIISCGKRYA	TFAMAVRE	LTGPAVIA	ATSIAIGLRGY	LLHYAIYQA-		
AtPIN2								-PKIIACGKSYA						
PtPIN2a OsPIN1c								-PKIIACGKSVA -PRIIACGNSLA						
SbPIN4c	SSLI6	YI <mark>H</mark> SL	YSYRHGIQHPA	IIARSISI	LSDAGLGM	amfsl g lfmali	Q	-PRIIACGNKFA	YIAMAYRF	LAGPAYMAI	AASIAYGLRGY	LLHIAIYQA-		
SiPIN4b	SSLIG'	VIHSL	VSFRAGIEMPR	IIARSISI	LSDAGLGH	AMFSLGLFMAL I	Q	-PRIIACGNKLA	AYAMAYRE	LYGPAYMAI	AASIAYGLRGY	LLHIAIVQA-		
SiPIN4d	SSLIG	/IHSL	VSFRHGIEMPR	IIARSISI	LSDAGLGM	AMFSLGLFMALI	Q	-PRIIACGNKLA	AVAMAVRF	LYGPAYMA	AASIAYGLRGY	LLHIAIYQA-		
PtPIN2b OsPIN10a								-PKYTASGKYLA APARRHCSGGSTTRI						
SbPIN4a								-PNIIACGURAT						
SiPIN4a	SSLIG	AHSL	IAFRUHISHPA	YYAKSISI	LSDAGLGH	AMFSLGLFHALI	Q	-PNLIACGHRAT	GISMGYRF	LAGPAYMAI	AASLAIGLRGS	LLQYAIYQA-		
OsPIN10b	ASLIG	_THSL	IAFRFHITMPI	IVAKSISI	LSDAGLGH	AMFSLGLFMAT	Q	-PKIIACGYSVA	AASMGVRF	FFGPAINA	ASAAYGIRGT	LLRIAIYQA-		
SbPIN7 SiPIN4c	HSYYG	_	TSEREHAHULA TEEDENTONDI	LYKNSISI	LSVHGLGM	HMFSLGLFMHTI Omeglelenom	Ų	-PKIIACGNKYA -PKIYACGNSYA	HT1UHTKI	LEGPHYMHI LLCDOVMOI	HISHHIGERGI HISHHIGERGI	TTKAHTANH-		
ALPING								-PKMIPCGAKKA						
PtPIN6a	SSYLG	LHS	YSFKUNYGMPS	LVKY <mark>SI</mark> KI	ISDAGLGM	AMFSL G LFMALI	Q	-PRIIYCGKKRA	THAMAIRE	ICGPYVMS	FASLAYGMRGE	RLRAAIYQA-		
PtPIN6b	SSYLG	LHSL	ASFKUNYGMPS	LYKYSIKI	ISDAGLGM	AMFSLGLFMALI	Q	-PRIIYCGKKRA	THAMAIRE	ICGPIYMS	rtsyaygmrgy	RLRAAIYQA-		
PtPINPIN7 Ospin1d	H2FTPI	LSHHL VTUSI	ASCKIPTUKA ASCKIPTUKA	TTAPSTST	LSKHGLGN I SNAGI GN	HMFSLGLFMHLI AMESLGLENALI	Ų	-PRIIACGNRMA -PRIIACGNSLA	110ULHKI CAUNUADE	LHGPHYNN I VGPAVMAI	AHSTRARI DEA BUSTUARI DEA	1 H181708		
OsPIN9	ASFLGI	IHS	IAFKCGFSMPK	IVEDSLFT	IRTTAVGL	SMFSSGTFIAR	Q	-SRFYPCGYKIA	SFSMYIKE	LIGPYYHLI	ASLVIGHHGT	LLHIAYYOA-		
SbPIN3	ASFLGI	IHSL	MAFKIGFSMPK	IYGDSLFI	IYTTAY <mark>GL</mark>	SMFASGTFIAR	Q	-SRFYPCGYTIA	SLSMYLKF	LIGPIIMLI	_aslaygmhgt	LLHIAYYQA-		
SiPIN5b	ASFLGI	IHSL	MAFKIGFSMPK	IVDDSLFT	IHTTAVGL	SMFASGTFIAR	Q	-SRFVPCGYTIA	SLSMYLKF	LIGPVIMLI	_VSLAIGMHGT	LLHIAVVQA-		
PpPIN6-1	HIYMG.	LYYSL FTUOC	THRKMALDER TONDI MEDI DO	DEECCH T	NCKCCTCH	OMECHOI EMOU	Ų N	-KKLY-ASHHYA -EKIIACGTSFA		TAPATUR.	/HSLLLGLKGD reetavel den	ILKHHELUH- VI DVOTTOO-		
SiPIN5a	ASFV6	ETHAC	VANRLHLELPS	AFEGSVLI	MSKSGTGM	AMFSHGLFHAQ	ŭ	-EKYLACGPCFA	ALGLYLK	ALGPAANA	CGSIAVGLRGD	VLRVAIIOA-		
SbPIN5b	asfyg:	ET <mark>H</mark> AC	YANRLHIELPS	AFEGSYLI	HSKSGTGM	AMFSM G LFMA Q i	Q	-EKYLACGPSFA	ALGLYLK	'ALGPAAMA'	[GSIAYGLRGD	YLRYATIQY-		
ALPIN5								-EKLIYCGTSLT						
PtPIN5a PtPIN5c								-EKYIACGASLT -EKYIACGASLT						
PtPIN5b	ACIIG	Y H AF	ISNRUHFEMPA	MMEGSILI	MSKAGTGT	AMFSHGIFMAL	Q	-EKYISCGASLA	VIGMILRE	IAGPAAHA	[GSIAYGLOGD	VLRVAIIOA-		
OsPIN5b	asylg'	/V <mark>H</mark> AC	IAYRUHLSLP 6	IYTGSLQY	HSRTGTGH	SHFSHGLFHGQ	Q	-ERVIACGAGLT	ALGMALRF	VAGPLATL'	/GAAAL <mark>GLRG</mark> D	VLHLAIIQA-		
SbPIN5c								-EKMIACGSGLA						
OsPIN5c SbPIN5a								-DKITYCGAGLT -EKITYCGAGPT						
SiPIN5d								-EKITYCGAGPT						
AtPIN8	ATLIG:	II <mark>H</mark> AT	LHFRLGANLPE	MIDK <mark>SI</mark> HL	LSDGGLGH	AMFSLGLFHAS	Q	-SSIIACGTKMA	IITMLLKF	YLGPALMI	ASAYCIRLKST	LFKY <mark>AILQA-</mark>		
PtPIN8a	ATHVA	IHAS	IHFRUGYKLPD	IVDKSIVI	LSTGGLGH	AMFSLGLFHALI	R	-PSIISCGIRMA	YVAMANKE	IVGPALMA'	/ASFAVRLEGT	YLKYAIYQA-		
PtPIN8b SbPIN8	HILYA	NNS	THERMOTOL DO	TVNNSTRI	LSTGGLGM LSTGGLGM	HMFSLGLFMASI AMEGLOLETALI	N	-PSIIACGIRMA -TKIVACGTKKM	NYHMAMKE	TAPATTU	/HSTHARFKRL ARZANTEMBET	ATKAHTAÓH-		
SiPIN8	ASLIG	IH	ISFRURVOLPS	IISNSTRI	LSDGGLGH	AMFSLGLFTALI	0	-TKITACGTKKH	LLSLGIRF	FLGPALMV	ESSYATGMHGT	TTKAUTIOU-		
SiPIN5c	ASYLG Y	YY <mark>H</mark> AC	IAFRADLRMP6	YYSESLQY	MSRTGTGH	SHFSHGLFHAQ	Q	-EKTYACGAGLA	ALGMALRF	VAGPLATLI	AGAAAF <mark>GLR</mark> SD	YMRFAIIQTY	LGGY	
OsPIN7	AINRH!	YY <mark>H</mark> FAFPF	FTFEFTLHLDP	YNYRCSL <mark>i</mark> aai	DSIAKLII-Y	AAISIGYMLKFI	RK	-EGLCAAYTDACIS	GESLASLTNS	LYYGHP <mark>ha</mark> i	RAHYGNHAGQI	YYQLS <mark>IFQA</mark> I'	YHLTSL	
OsPIN8														
PpPIN6-2														

	781	790	800	810	820	830	840	850	860	870	880	890	899
OsPIN1a			-	-		-	STAVTEGHI TI	RLPITLVYYILLG	,				
SbPIN1								RLPITLYYYILMG	_				
SiPIN1b								RLPITLYYYILLG					
OsPIN1b								RLPITLVYYILLG	_				
SiPIN1a								RLPITLVYYILLG					
SbPIN4b								RLPITLYYYILLG					
ALPIN1				ALPQGIYPF	VFAK	EYNYHPDIL!	STAVIFGHLI	ALPITLLYYILLG	L				
PtPIN1c				ALPQGIYPF	VFAK	EYNYHPEIL!	STGVIFGHLI	RLPITLVYYILLG	L				
PtPIN1d				ALPQGIYPF	YFAK	EYNYHPDIL!	STAVIFGHLI	RLPITLYYYILMG	L				
tPINPIN1a				ALPQGIYPF	VFAK	EYNYHPDIL!	STGYIFGMLII	ALPITLYYYILLG	L				
PtPIN1b								RLPITLYYYILLG	_				
ALPIN3								ALPITLYYYILLG					
ALPIN7								ALPITLYYYILLG					
PtPIN3a								RLPITLVYYIFLG					
PtPIN3b				ALPQGIYPF	YFHK	EYNYHPHIL	STGYIFGMLII	ALPITLYYYILLG	L				
ALPIN4								ALPITLYYYILLG					
PpPIN3-2								RLPITHVYYILLG					
PpPIN3-1								RLPITMYYYILLG					
PpPIN3								TLPTALLYYILLG Alpitilyyvllg					
OsPIN2 SbPIN2								ALPITILYYYLLG BLPITILYYYLLG					
SiPIN2								ALPITILYYYLLG	_				
ALPIN2								ALPYTYLYYLLG					
PtPIN2a								ALPITYLYYYLLG	_				
OsPIN1c								RLPITLYYYILLG					
SbPIN4c								RLPITLVYYILLG	_				
SiPIN4b								RLPITLYYYILLG					
SiPIN4d								RLPITLYYYILLG					
PtPIN2b								RLPYTIIYYYLLD					
OsPIN10a								LICIACCRTLPM		RAIYLCDICTRO	INDRHYSNITI)IRYYFC	
SbPIN4a				ALPQGIYPF	YFAK	EYNYHPAIL!	STHYIFGHLI	ALPITLLYYILLG	LKPY				
SiPIN4a				ALPQGIVPF	VFAK	EYNYHPAIL!	STHVIFGHLI	RLPITLLYYIVLG	LGPV				
OsPIN10b				ALPQGIYPF	YFAK	EYNLHATIL	CTLVIFGMLII	RLPITLYYYIILG	LL				
SbPIN7								ALPITLYYYIILG					
SiPIN4c								<u>GLPIALYYYIILG</u>					
AtPIN6								SLPYTILYYYLLG					
PtPIN6a				ALSQGIYPF	YFAR	EYGLHPDIM	STGYIFGMLY	SLPYTLLYYIFIG	L				
PtPIN6b								SLPYTLIYYIFLG					
PtPINPIN7								SLPITILYYIFLG					
OsPIN1d								ALPITLYYYILLG					
OsPIN9								SLPYTIYYYILLG					
SPLIN3								SLPYTIYYYILLG Slpytiyyyillg					
SiPIN5b				ALPONVISE				PLAITALLTER	L				
PpPIN6-1 OsPIN5a								SLPLLVGFYIVLE	TD				
SiPIN5a								SLPLLYGFYIYLE					
SbPIN5b				SLLAF		L (ULINDYL)	STITTE UNLY	JEI EETUF I I TEE	LTV				
ALPIN5						EYGI HANVI (STAVTEGHI V	S <mark>lp</mark> ylyayyaale	FTH				
PtPIN5a				AFPOATTSF	TFAO	EYGL HAEVI	STAVTEGTTV	ALPYLIAYYAILD	FVH				
				4112201	21.119								



Supplementary Figure 1. Contd.