Phylogeny and distinct properties of major intrinsic proteins in the genomes of six *Phytophthora* species suggest their novel functions in *Phytophthora*

Abul Kalam Azad[1,*], Jahed Ahmed[1,†], Al-Hakim[1,†], Md. Mahbub Hasan[2], Md. Asraful Alum[3], Mahmudul Hasan[1,†], Takahiro Ishikawa[4], Yoshihiro Sawa[4]

[1]*Department of Genetic Engineering and Biotechnology, Shahjalal University of Science and Technology, Sylhet 3114, Bangladesh*

[2]*Department of Genetic Engineering and Biotechnology, University of Chittagong, Chittagong-4331, Bangladesh*

[3]*Forensic DNA Laboratory of Bangladesh Police, Malibagh, Dhaka-1000, Bangladesh*

[4]*Department of Life Science and Biotechnology, Shimane University, Shimane 690-8504, Japan*

†Present Addresses: JA, *Louvain Institute of Biomolecular Science and Technology, Université catholique de Louvain, Louvain-la-Neuve-1348, Belgium; AH, Centre for Vaccine Sciences, International Centre for Diarrhoeal Disease Research, Bangladesh; MH, Department of Pharmaceutical and Industrial Biotechnology, Sylhet Agricultural University, Sylhet, Bangladesh*

***Corresponding author:** Department of Genetic Engineering & Biotechnology, Shahjalal University of Science and Technology, Sylhet 3114, Bangladesh. Tel.: + 88 0821 717850x411; fax: + 88 0821 725050; E-mail: dakazad-btc@sust.edu

Highlights:

- Major intrinsic proteins (MIPs) family in *Phytophthora* species are divergent.
- *Phytophthora* MIPs (PMIPs) were phylogenetically and structurally distinct from their counterparts in other taxonomic domains.
- PMIPs might have novel functions.
- The MIPs are suggested to be involved in host-pathogen interactions and could be considered attractive anti-*Phytophthora* targets.

ABSTRACT

Major intrinsic proteins (MIPs), commonly known as aquaporins, transport water and other non-polar solutes across membranes. MIPs are believed to be involved in host-pathogen interactions. Herein, we identified 17, 24, 27, 19, 19, and 22 full-length MIPs, respectively, in the genomes of six *Phytophthora* species, *P. infestans*, *P. parasitica*, *P. sojae*, *P. ramorum*, *P. capsici*, and *P. cinnamomi*, which are devastating plant pathogens and members of oomycetes, a distinct lineage of fungus-like eukaryotic microbes. Phylogenetic analysis showed that the *Phytophthora* MIPs (PMIPs) formed a completely distinct clade from their counterparts in other taxa and were clustered into nine subgroups. Sequence and structural analysis of homology models indicated that the primary selectivity-related constrictions, including aromatic arginine (ar/R) selectivity filter and Froger's positions in PMIPs were distinct from those in other taxa. The substitutions in the conserved Asn-Pro-Ala motifs in loops B and E of many PMIPs were also divergent from those in plants. We further deciphered group-specific consensus sequences/motifs in different loops and transmembrane helices of PMIPs, which were distinct from those in plants, animals, and microbes. The data collectively supported the notion that PMIPs might have novel functions and could be considered attractive anti-*Phytophthora* targets.

Keywords: Major intrinsic proteins, *Phytophthora*, Phylogeny, Oomycetes and Substrate specificity

1. Introduction

The super family of major intrinsic proteins (MIPs) possesses channel-forming integral membrane proteins that transport water and other non-polar small solutes, such as ammonia, urea, boron, silicon, carbon dioxide, glycerol, hydrogen peroxide, antimony, and arsenite (Gomes et al., 2009; Ishibashi et al., 2009; Azad et al., 2012; Verkman, 2012; Maurel et al., 2015; Pommerrenig et al., 2015; Azad et al., 2016; Azad et al. 2018). They are found in all living organisms from bacteria to mammals and are abundant in plants (Gomes et al., 2009; Azad et al., 2011; Maurel et al., 2015). Orthodox aquaporins (AQPs), which transport only water and aquaglyceroporins (AQGPs), that can transport other uncharged small solutes in addition to water or without water, are prototype members of MIPs (Mukhopadhyay et al., 2014; Maurel et al., 2015; Pommerrenig et al., 2015).

All AQPs share some structural features, although their amino acid sequences differ substantially. Each monomer of MIP is composed of six transmembrane (TM) α-helices (H1-H6) with five connecting loops (loops LA–LE) and cytoplasmic N- and C-termini. The pore of the channel is characterized by two constrictions that theoretically specify the profile of transport selectivity. The first constriction is formed at the center of the pore by two highly conserved Asn-Pro-Ala (NPA) motifs on loops B and E because of the close opposition of their asparagine residues (Wallace and Roberts, 2004). This constriction is involved in proton exclusion (Tajkhorshid et al., 2002). The second constriction, called the aromatic/arginine (ar/R) constriction or the selectivity filter, is formed toward the luminal side of the membrane by four residues from helix 2 (H2), helix 5 (H5), and loop E (LE1 and LE2) (Fu et al., 2000; Sui et al., 2001). Mutation at this ar/R selectivity filter is thought to determine the broad spectrum of substrate conductance in plant AQPs (Wallace and Roberts, 2004; Gupta and

Sankararamakrishnan, 2009; Azad et al., 2012; Azad et al., 2016). Five relatively conserved amino acid residues known as Froger's positions (FPs) and designed P1–P5 play roles in MIP sub-grouping and substrate selectivity (Froger et al., 1998; Heymann and Engel, 2000; Azad et al., 2016).

Although 13 different MIPs identified in mammals are divided into three major subfamilies, the genomes of plants encode 2–5-fold or more MIPs, which are grouped into 4–7 subfamilies (Maurel et al., 2015; Azad et al., 2016; Potokar et al., 2016). However, fungi genomes have up to five MIPs with diversified subgroups (Pettersson et al., 2005; Verma et al., 2014). Algae have 1–6 MIPs, but are highly divergent and share only limited similarities with land plant MIPs (Anderberg et al., 2011). In humans, MIPs play significant roles in brain-water balance, kidney nephrons, cell migration, cell proliferation, neural activity, pain, epithelial fluid secretion, skin hydration, adipocyte metabolism, and ocular function (Ishibashi et al., 2009; Verkman, 2012). They are associated with many human diseases, such as glaucoma, cancer, epilepsy, nephrogenic diabetes insipidus, and obesity (Ishibashi et al., 2009; Verkman, 2012), and therefore, they have been potential drug targets (Soveral and Casini, 2017). In plants, MIPs are involved in many physiological processes, such as motor cell movement, root and leaf hydraulic conductance, diurnal regulation of leaf movements, rapid internode elongation, responses to numerous biotic and abiotic stresses, temperature-dependent petal movement, and petal development (Azad et al., 2004; Peng et al., 2007; Azad et al., 2008; Uehlein and Kaldenhoff, 2008; Gomes et al., 2009; Gao et al., 2010; Muto et al., 2011; Azad et al., 2013; Maurel et al., 2015; Afzal et al., 2016; Sonah et al., 2017). MIPs also play important roles in host-parasite interactions and numerous MIPs have been reported in several protozoan parasites, such as *Plasmodium*, *Trypanosoma*, and *Leishmania* species (Hansen et al., 2002; Montalvetti et al., 2004; Beitz, 2005; Fadiel et al., 2009; Kun and de Carvalho, 2009; Baker et al., 2012). These

MIPs are considered potential drug targets (Beitz, 2005; Fadiel et al., 2009; Kun and de Carvalho, 2009). In mycorrhized plants, both plant and fungal MIPs play significant roles in water and nutrient transport and in the drought resistance of plants (Giovannetti et al., 2012; Li et al., 2013; Maurel et al., 2015). MIPs in pathogenic fungi may act as attractive targets for antifungal drugs (Verma et al., 2014). However, no study has been conducted on the MIPs of oomycetes, a distinct lineage of fungus-like eukaryotes with diverse microorganisms that are related to organisms such as brown algae and diatoms (Beakes et al., 2012; Thines, 2014; Fawke et al., 2015; Wang et al., 2016). Oomycetes are globally distributed and ubiquitous in marine, freshwater, and terrestrial environments (Thines, 2014), and cause devastating diseases to both plants and animals (Derevnina et al., 2016).

Among oomycetes, the *Phytophthora* genus comprises more than 117 species, which are highly devastating to a wide range of agriculturally and ornamentally important plants, causing severe economic losses (Martin et al., 2012; Kroon et al., 2004; Tyler et al., 2006; Beakes et al., 2012; Thines, 2014; Fawke et al., 2015; Wang et al., 2016). *P. infestans*, the cause of late blight of potato and tomato, resulted the Irish Potato Famine in the mid-19th century (Haas et al., 2009; Yoshida et al., 2013; Thines, 2014; Wang et al., 2016), and *P. sojae* costs the soybean industry millions of dollars each year (Tyler et al., 2006; Sahoo et al., 2017). *P. ramorum*, the cause of death of oak trees in North America, has severe impact on natural ecosystem, and *P. cinnamomi* is a substantial threat to natural eucalyptus forests in Australia (Tyler et al., 2006; Lamour et al., 2012; Thines, 2014). *P. capsici* causes foot rot disease in black pepper and *Phytophthora* blight in vegetable crops, which causes serious effects on the production of cucurbits, peppers, eggplants, and numerous other important vegetables worldwide (Jackson et al., 2012; Hao et al., 2016; Johnson et al., 2016; Liu et al., 2016). *P. parasitica* infects a broad range of plants, being capable of infecting over 72 plant genera worldwide (Blackman et al., 2015; Meng et al., 2015).

The genome sequences of these six plant pathogenic *Phytophthora* species have been completed (Tyler et al., 2006; Haas et al., 2009; Lamour et al., 2012; Fawke et al., 2015) and are available in 'FungiDB' (http://fungidb.org/fungidb/). In the study reported herein, we identified and characterized *MIP* genes in the genomes of these six *Phytophthora* species (*PinMIP, PpaMIP, PsoMIP, PraMIP, PcaMIP,* and *PciMIP* genes of *P. infestans, P. parasitica, P. sojae, P. ramorum, P. capsici,* and *P. cinnamomi,* respectively) by using bioinformatics tools. To the best of our knowledge, this is the first report on MIPs of any organism of oomycetes. This report showed that the genomes of *Phytophthora* species have several-fold greater number of MIP homologues than those of algae, fungi, other parasites, and even more than in mammals. The numbers of their MIP homologues are comparable to that of plants. Comprehensive analysis with different bioinformatics tools revealed that the MIPs in *Phytophthora* species are phylogenetically and structurally distinct from their counterparts in other taxa.

2. Materials and methods

2.1. Identification of PinMIP, PpaMIP, PsoMIP, PrMIP, PcaMIP, and PciMIP genes

The genomes of *P. infestans* T30-4, *P. parasitica* INRA-310, *P. sojae* strain P6497, *P. ramorum* strain Pr102, *P. capsici* LT1534, and *P. cinnamomi* CBS 144.22 are available at FungiDB (http://fungidb.org/fungidb/). In FungiDB, a search with the keyword 'aquaporin' revealed that 21, 26, 35, 32, 25, and 32 aquaporin hits were available for each strain of *P. infestans, P. parasitica, P. sojae, P. ramorum, P. capsici,* and *P. cinnamomi,* respectively. We retrieved the protein sequences of these available AQPs for every organism and checked the typical features of AQPs in their primary protein structures. All of the AQP protein sequences available in every *Phytophthora* species were used as queries for MIPs in the genome of a particular *Phytophthora*

153 species using TBLASTN and BLASTp tools. Furthermore, the genomes of the *Phytophthora*
154 species were searched for MIPs using TBLASTN and BLASTp tools with the protein sequences
155 of the complete set of 55 MIPs from *Populus trichocarpa* (PtMIP), 22 MIPs from
156 *Physcomitrella patens* (PpMIP), and 13 MIPs from humans (MAQPs) as queries so that XIPs
157 (uncharacterized X intrinsic proteins), GIPs (GlpF-like intrinsic proteins), HIPs (Hybrid intrinsic
158 proteins) or superaquaporins could be detected if they were encoded in the genomes of the six
159 *Phytophthora* species. MIPs in *P. infestans* (PinMIPs), *P. parasitica* (PpaMIPs), *P. sojae*
160 (PsoMIPs), *P. ramorum* (PraMIPs), *P. capsici* (PcaMIPs), and *P. cinnamomi* (PciMIPs) were
161 included until no more MIPs could be found from the corresponding species. Every sequence
162 from each *Phytophthora* species was individually compared to identify the maximum number of
163 MIPs for further analyses. Some of the MIP sequences might have been partial or might not
164 have had all the features associated with its MIP channel. To investigate this, we used the
165 multiple alignment program Clustal Omega (http://www.ebi.ac.uk/Tools/msa/clustalo/) to align
166 all the sequences in an individual species. The multiple sequence alignment was used to
167 determine the following features specific to the MIP family: (i) presence of two NPA or NPA-
168 like motifs, (ii) presence of six TM α-helices, and (iii) two functionally important loops
169 possessing the features characteristically present in MIP channels. The TM α-helices were
170 predicted by SOSUI (http://bp.nuap.nagoya-u.ac.jp/sosui/), TMpred
171 (http://www.ch.embnet.org/software/TMPRED_form.html), and the tools of ExPASy
172 (http://kr.expasy.org/tools/). The genomic regions containing *MIP* genes were further used to
173 determine the gene structure using the program GeneMark.hmm ES-3.0 (Lomsadze et al., 2005)
174 (http://exon.gatech.edu/GeneMark), a self-training based algorithm for prediction of genes from
175 novel eukaryotic genomes. When short genes were found, their sequences with 1000 base
176 flanking regions were subjected to Genetyx_SV_RC_version 7 to investigate their protein

sequences.

2.2. Phylogenetic and domain analysis of PinMIPs, PpaMIPs, PsoMIPs, PrMIPs, PcaMIPs, and PciMIPs

To understand the diversity and evolution of PMIPs and to compare them with those in plants, animals, fungi, and algae, we performed phylogenetic analysis of all MIPs in six *Phytophthora* species using Molecular Evolution Genetic Analysis (MEGA) version 5.0 (Tamura et al. 2011). PinMIPs, PpaMIPs, PsoMIPs, PrMIPs, PcaMIPs, and PciMIPs were aligned with all PtMIPs, PpMIPs, and MAQPs, and five representative members from each of 10 subfamilies of MIPs in fungi (Verma et al., 2014) and all MIPs in the genomes of nine algae (Anderberg et al., 2011) using the Clustal Omega program (http://www.ebi.ac.uk/Tools/msa/clustalo/) and a phylogenetic tree was constructed with MEGA. The evolutionary history was inferred using two different clustering algorithms, namely neighbor-joining and maximum parsimony methods which are generally used for phylogentic analysis of MIPs (Abascal et al. 2014; Verma et al. 2014; Danielson and Johanson 2010), and the genetic distance was estimated by the p-distance method. Reliability of individual branches of the tree was estimated by performing bootstrapping with 1000 replicates. To construct the phylogenetic tree with the MIPs in the six *Phytophthora* species, all of their MIPs were aligned as above. The identified PinMIPs, PpaMIPs, PsoMIPs, PrMIPs, PcaMIPs, and PciMIPs were classified into different subfamilies and groups based on the phylogenetic tree constructed from them.

2.3. Prediction of subcellular localization and computation of Ka/Ks value

The subcellular localizations of PMIPs were predicted *in silico* by using tools of WoLF PSORT (http://www.genscript.com/wolf-psort.html) and Cello prediction system

(http://cello.life.nctu.edu.tw/) as described previously (Azad et al., 2016). The Ka/Ks values of the *PMIPs* were calculated using an online Ka/Ks calculation tool at http://services.cbu.uib.no/tools/kaks. A Ka/Ks value greater than one implied gene evolution under positive or Darwinian selection; less than one indicated purifying (stabilizing) selection, and a Ka/Ks value of one suggested a lack of selection or possibly a combination of positive and purifying selections at different points that canceled each other out (Zhang et al., 2013).

2.4. Homology modeling

Homology models were constructed using the Molecular Operating Environment software (MOE 2009.10; Chemical Computing Group, Quebec, Canada), based on a segment-matching procedure (Levitt, 1992) and a best-intermediate algorithm with the option to refine each individual structure enabled. The sequence of each MIP homolog was aligned with the open conformation of spinach PIP, SoPIP2;1 (PDB, Protein Data Bank ID: 2B5F) (Törnroth-Horsefield et al., 2006), or other AQP templates as indicated, using the MOE software as previously described (Azad et al., 2012). The alignment of each MIP homolog was based on both sequence and structural homology with the structure of SoPIP2;1. The 3D structure models were formed using the MOE homology program and the stereochemical quality of the templates and models were assessed as previously described (Azad et al., 2011).

2.5. Determination of pore diameter and pore lining residues

To analyze the MIP channels, the poreWalker server (http://www.ebi.ac.uk/thornton-srv/software/PoreWalker/ (Pellegrini-Calace et al., 2009) was used, which is a fully automated method designed to detect and characterize transmembrane protein channels from their 3D structures. The 3D structure of each MIP was uploaded to the server, which generated its

specific pore characteristics; in particular, the conformation and the regularity of the channel cavity, the corresponding pore lining residues and atoms, and the location of pore centers along the channel. The pore diameter at the ar/R selectivity filter was determined from the PoreWalker outputs as previously described (Azad et al., 2016). The pore lining residues, which are essential for channel formation, were identified using the PoreWalker server.

3. Results

3.1. Genome-wide identification of PinMIP, PpaMIP, PsoMIP, PraMIP, PcaMIP, and PciMIP genes

The whole genome shotgun sequence (WGS) of *P. infestans*, *P. parasitica*, *P. sojae*, *P. ramorum*, *P. capsici*, and *P. cinnamomi* available at FungiDB was searched for *PinMIP*, *PpaMIP*, *PsoMIP*, *PraMIP*, *PcaMIP*, and *PciMIP* genes using TBLASTN. The query PinMIP, PpaMIP, PsoMIP, PtMIPs, PpMIPs and MAQPs sequences resulted in 21, 27, and 33 hits for *PinMIP*, *PpaMIP*, and PsoMIP, respectively. We further analyzed the PinMIP, PpaMIP, and PsoMIP sequences for manual inspection of their amino acid sequences, TM domains, and homology models. After all these analysis, we found that out of 21, 27, and 33 hits for *PinMIP*, *PpaMIP*, and *PsoMIP*, respectively, 4, 3, and 6 were deemed to be pseudo *MIP* genes in *P. infestans*, *P. parasitica*, *P. sojae*, respectively, and were discarded. Characteristics, such as short sequences, N- or C-termini less, addition or deletion of sequences, and combinations thereof were observed in the discarded *MIPs*. However, the query PraMIP, PcaMIP, and PciMIP returned no hits in the TBLASTN and BLASTp searches. Therefore, we retrieved the MIP sequences of *P. ramorum*, *P. capsici*, and *P. cinnamomi* from the FungiDB and analyzed them as mentioned above. Out of 32 MIP sequences in each of *P. ramorum*, *P. capsici*, and *P. cinnamomi*, 13 MIPs in the former two and 10 MIPs in the latter were deemed pseudo *MIPs* for

the reasons noted above and were discarded. We ultimately obtained 17, 24, 27, 19, 19, and 22 full-length PinMIP, PpaMIP, PsoMIP, PraMIP, PcaMIP, and PciMIP protein sequences from the WGS of *P. infestans*, *P. parasitica*, *P. sojae*, *P. ramorum*, *P. capsici*, and *P. cinnamomi*, respectively (Tables 1-6). The amino acid lengths of PinMIP, PpaMIP, PsoMIP, PrMIP, PcaMIP, and PciMIP homologues with their maximum sequence identity with MIPs in other *Phytophthora*, humans (taxid 9606), *Arabidopsis thaliana* (taxid 3702), fungi (taxid 4751), and algae (taxid 3027) are tabulated in Tables 1–6. Although the MIPs of one *Phytophthora* species exhibited a maximum 80–95% identity with those of other *Phytophthora* species (Tables 1-6), the TBLASTN search revealed that their highest identity with MIPs of *Arabidopsis thaliana* (taxid: 3702), *Homo sapiens* (taxid: 9606), fungi (taxid: 4751), and algae (taxid: 3027) was within 33, 45, 52, and 32%, respectively (Tables 1-6). This result indicated that PMIPs have higher identity with MIPs of fungi compared to those of animals, plants, and algae. This result further supported the notion that PMIPs might have extensive divergence in their sequence and structural properties compared to those in other taxonomic groups.

The Ka/Ks value was >1 only for *PinMIPC1;2* and *PsoMIPG1;*9, indicating that the evolution of these genes in *P. infestans* and *P. sojae*, respectively, was likely under positive or Darwinian selection (Tables 1-6). The Ka/Ks value 0 for *PpaMIPC1;3*, *PsoMIPG1;8*, *PraMIPG1;2*, and *PraMIPG1;3* indicated neutral selection. The remaining *PMIP*s showed Ka/Ks values <1, demonstrating purifying selection. The score of the Cello prediction system showed that all PMIPs might have been localized in the plasma membrane (Supplementary Table S1). However, WoLF PSORT scores suggested that some PMIPs have possibility to localize in other intragranular membranes, such as the cytoskeleton, endoplasmic reticulum, vacuole, golgi, and mitochondria.

3.2. Phylogenetic distribution and nomenclature of PMIPs

The sequence of PMIPs was routinely compared with the subfamilies of those in other kingdoms by constructing multiple and/or pair-wise alignments using Clustal Omega and EMBOSS Needle, respectively. To investigate the structural conservation of PMIPs with other MIPs in plants, animals, and microbes, a 3D structural alignment was constructed with the templates of human AQP1 (PDB ID, 1J4N), spinach plasma membrane intrinsic protein, SoPIP2;1 (PDB ID, 2B5F), and *E. coli* glycerol facilitator, GlpF (PDB ID, 1FX8). All 3D models constructed with these three templates showed the typical hourglass shaped AQPs with six TM helices and five connecting loops (Figure 1). Superposition of the three models of each PMIP showed that the helices superposed very closely. Deviation was observed only in the loops. This structural alignment was used as a guide for sequence alignment. To classify the PMIPs, their protein sequences were analyzed phylogenetically with representatives of subfamilies of MIPs in two plant genomes, and a human, fungi, and algae genome, as described in Materials and Methods. Although PtMIPs and PpMIPs were divided into five and seven subfamilies, respectively (Danielson and Johanson, 2008; Gupta and Sankararamakrishnan, 2009), and those in fungi and algae were separated into 10 and seven subfamilies, respectively (Anderberg et al., 2011; Verma et al., 2014), PMIPs formed a completely distinct clade (Figure 2). Despite one MIP in *P. parasitica*, which was associated with Delta AQGPs of fungi (Verma et al., 2014), no other PMIP was associated with any subfamilies of MIPs in other organisms in the three domains of life. We therefore constructed a phylogenetic tree with all 126 MIP homologues in the six *Phytophthora* species (Figure 3 and supplementary figure S1). The 126 MIPs were clustered into nine subgroups. Because the PMIPs were not phylogenetically

297 clustered with any subfamily or group of MIPs reported in plants, animals, fungi, algae, or
298 bacteria, these groups were arbitrarily named MIPA-MIPI. Each MIP homologue in every
299 *Phytophthora* species was named by taking the first letter (Upper case) from the genus and the
300 first two letters (lower case) of the species names with the MIP group (A to I), and the number
301 of the homologues in each group was stated consecutively, i.e., PsoMIPA1;1 for MIPA in *P.*
302 *sojae* and so on. MIPAs and MIPCs-MIPGs included one to several MIPs from each of the six
303 *Phytophthora* species. MIPBs contained five homologues of five species, except *P. capsici*, and
304 MIPHs consisted of four MIPs from three *Phytophthora* species. The MIPI group had only one
305 homologue, which was Delta AQGP, as previously mentioned. However, although these groups
306 included a small number of MIP homologues, they had a distinct ar/R constriction. The
307 characteristics of each of the nine PMIPs groups are detailed in the next sections.

308
309
310 *3.3. Sequence analysis of PMIPs*

311 We calculated the pair-wise sequence identity and similarity among the intragroup and
312 intergroup PMIPs by using EMBOSS Needle (http://www.ebi.ac.uk/Tools/psa/emboss_needle/).
313 The average identity and similarity among the intragroup PMIPs was 76% and 83%,
314 respectively, whereas those in the intergroup PMIPs were 40 and 52%, respectively
315 (Supplementary Table S2). However, the sequence identity and similarity of MIPHs to MIPA-
316 MIPGs was only 34 and 48%, respectively. Nevertheless, the intergroup average sequence
317 identity among MIPAs-MIPDs and MIPEs-MIPGs was 46 and 50% and similarity was 64 and
318 62%, respectively. In contrast, the intergroup average sequence identity and similarity of PMIPs
319 from MIPAs-MIPDs to MIPEs-MIPGs was 48 and 63%, respectively. These results indicated
320 that PMIPs of MIPAs-MIPDs were closer compared to those of MIPEs-MIPGs and vice versa.
321 The intergroup sequence identity varied from 2 to 30% and similarity varied from 1 to 56%,

indicating that each group was divergent from the others.

3.4. Gene structure of MIPs in Phytophthora species

All the full-length *MIP* sequences found in *P. infestans*, *P. parasitica*, *P. sojae*, *P. ramorum*, *P. capsici*, and *P. cinnamomi* were analyzed for introns and exons (Figure 4). Interestingly, among the 126 PMIPs, 107 homologues had no introns. This characteristic might be common to prokaryotic genes. Thirteen PMIPs had one introns. Two introns were observed in four *PMIP* genes, namely *PciMIPH1;1*, *PinMIPE1;1*, *PinMIPG1;2*, and *PsoMIPH1;1*. Three and five introns were found only in PciMIPA1;4 and PinMIPC1;3, respectively. Despite a few disparities, conserved intron positions were not found in *PMIPs* (Figure 4).

3.5. Analysis of ar/R selectivity filter and FPs of PMIPs

The amino acid residues in the ar/R selectivity filter and FPs are crucial for functional grouping and substrate selectivity of MIPs (Froger et al., 1998; Heymann and Engel, 2000; Hove and Bhave, 2011; Pommerrenig et al., 2015; Azad et al., 2016; Azad et al. 2018). To determine the residues in the ar/R selectivity filter and FPs, we constructed 3D models of all PMIPs. The structure-based alignments and multiple sequence alignments of PMIPs helped us to identify the four amino acid residues at the ar/R selectivity filter, and the five residues in the FPs. The residues at the ar/R selectivity filter and FPs in nine groups are shown in Figure 3. Although MIPs in plants and fungi conserve group-specific ar/R selectivity filters and FPs (Wallace and Roberts, 2004; Gupta and Sankararamakrishnan, 2009; Verma et al., 2014; Pommerrenig et al., 2015; Azad et al., 2016), the ar/R filter and FPs were identical in several groups of PMIPs (Figure 3). Despite a few disparities (PciMIPA1;4, PsoMIPA1;4, PpaMIPA1;3, and

PinMIPE1;2), the tetrad in the ar/R selectivity filter in MIPAs and MIPDs-MIPGs was WGYR, which was found in β-AQPGs in fungi (Verma et al., 2014). However, MIPBs, MIPCs, MIPHs, and MIPI contained group-specific ar/R selectivity filters, having the tetrad WGCR, WALR, WSLR, and WTAR, respectively, which are not usually observed in MIPs of other taxonomic groups. Nevertheless, the residue in the H5 position of the ar/R filter in PciMIPH1;1 and PsoMIPH1;1 was deleted. The ar/R selectivity filter, as the name suggests, usually consists of an aromatic residue and R, was found in the H2 and LE2 position, respectively. This was true for all PMIPs, although exceptions have been reported in plants, fungi, and algae (Gupta and Sankararamakrishnan, 2009; Anderberg et al., 2011; Verma et al., 2014; Azad et al., 2016). The H5 position in PMIPs is occupied by small residues (G/A/S), which is generally the case in many MIPs in plants (TIPs, NIPs, and SIPs), fungi, and algae (Anderberg et al., 2011; Verma et al., 2014; Azad et al., 2016). However, all PIPs in plants, most of the AQPs in mammals, and some groups of MIPs in algae conserved H, and some fungal, algal, and plant MIPs conserved L/I/V in the corresponding position ((Anderberg et al., 2011; Azad et al., 2012; Verma et al., 2014; Azad et al., 2016), Supplementary Figure S2). A/C/G/S/T residues are usually found in the LE1 position in MIPs of plants, animals, fungi, and algae (Anderberg et al., 2011; Verma et al., 2014; Azad et al., 2016). Except for some fungal MIPs, an aromatic residue is not usually available at the LE1 position in eukaryotes. Interestingly, most of the PMIPs (MIPAs, MIPDs-MIPGs) conserved the bulky hydrophobic aromatic Y in the LE1 position (Figure 3). The bulky aromatic W and Y in the H2 and LE1 positions in many PMIPs in the five groups collectively might have influenced the channel properties and their transport profile. Hydrophobic L, which is not generally found in MIPs of plants, animals, and algae, but available in some fungal MIPs, is conserved in homologues of MIPC and MIPH. This hydrophobic larger amino acid would have changed the channel property and transport profile compared to MIPs that have A/C/G/S/T

in the same position. Superposition of the 3D models of PMIPs with crystal structures of bacterial glycerol facilitator (Glps), bovine aquaporin 1 (AQP1), and spinach PIP, SoPIP2;1, or that of intergroup homologues revealed that the architecture of the ar/R selectivity filter in PMIPs would be influenced by the residue at the H5 and LE1 positions (Figure 5). However, this might be one of the reasons for transport selectivity. The conserved residue WGYR in the ar/R selectivity filter of PMIPAs and MIPDs-MIPGs indicated that their transport selectivity might be caused by another mechanism or they have the same transport profile.

The FPs, P1 from H3, P2 and P3 from LE, and P4 and P5 from H6 as described by Froger et al. (Froger et al., 1998) were YDRFW in MIPAs-MIPDs, excluding two homologues, PsoMIPA1;4 and PsoMIPA1;5, in which the F was substituted by G and W, respectively (Figure 3). Although these FPs are diverse from MIPs in plants, animals, and algae, most of the groups of fungal MIPs conserve them (Supplementary Figure S2). In MIPEs-MIPHs, the P2-P5 positions were conserved with DRFW, except PsoMIPF1;2, PpaMIPF1;3, PsoMIPF1;3, PcaMIPF1;2, and PciMIPF1;3, in which the W in the P5 position was substituted by L. The P1 position in these groups was diverse, Q/D in MIPEs; D/N in MIPFs; Q in MIPGs; and H/N in MIPHs. The FPs in the MIPI (PpaMIPI1;1) was FDRCW.

3.6. PMIPs with substituted NPA motifs

The conserved NPA motifs in LB and LE were found in MIPAs, MIPBs, MIPDs, MIPI, and a subgroup of MIPCs, except PciMIPA1;4, in which the P of the NPA in LE was substituted by S (Figure 3). As reported for the small basic intrinsic proteins (SIPs) in all plants (Gomes et al., 2009; Azad et al., 2016), the homologues of MIPEs-MIPGs conserved a usual NPA motif in the LE, but an unusual NPA motif in the LB, apart from PpaMIPG1;1 that contained the usual NPA in both loops and PinMIPG1;1, in which the N of the NPA in LE was substituted by K. However,

393 unlike SIPs, MIPEs-MIPGs have high molecular weight with longer amino acid sequences.
394 Moreover, although substitution of A by T or L in the NPA motif of LB is observed in SIPs
395 (Verma et al., 2014; Azad et al., 2016), P was substituted by T in all MIPEs and MIPGs. In seven
396 homologues of MIPFs, P was substituted by C or S, and in six homologues of the same group, A
397 was substituted by S. Interestingly, two MIP homologues from each of the six *Phytophthora*
398 species had unusual NPA motifs both in LB and LE, and they were clustered as a subgroup in
399 MIPCs. More intriguingly, the NPA motifs of LB and LE in MIPHs were substituted with ISV
400 and S(P/V/I)N, respectively. The NIPs with unusual NPA motifs, where A was in LB and that in
401 LE were substituted by S and V, respectively, had characteristic R-rich C-termini (Azad et al.,
402 2016; Azad et al. 2018). However, MIPCs with both unusual NPA motifs had a RSxGPYE(Y/F)
403 C-terminus followed by the group-conserved GYHH motif. Although MIPHs' C- termini have no
404 such conserved sequences before the ExQH motif, they contain comparatively more charged
405 residues (Supplementary Figure S3).
406
407 *3.7. Group-specific consensus sequences/motif of PMIPs*
408 We further compared the PMIPs sequences for group-specific consensus sequences. In the TM
409 regions, the intergroup similarities among the PMIPs were very high. The group-specific
410 deviation was particularly observed in the loops and N- and C-termini (Table 7). Most of the
411 homologues in MIPEs-MIPGs had longer N- termini compared to MIPAs-MIPCs
412 (Supplementary Figure S3). Interestingly, MIPDs had short N- and C-termini. It is more
413 intriguing that excluding only three PMIPs in group E (PinMIPE1;1, PcaMIPE1;1, and
414 PpaMIPE1;1), all PMIPs conserved a positively charged histidine residue in the C-terminus.
415 With this conserved H residue, all PMIPs conserved a group-specific motif (Table 7). Except
416 MIPEs and MIPHs, the C-termini in all PMIPs were H-enriched, to a greater extent in MIPAs

and MIPBs. The C-termini of MIPEs were enriched with D and E and had di- or tri-acidic residues. Despite the NPA motifs in LB and LE, every loop (LA-LE) had group-conserved sequences (Table 7). However, the sequence shown in LB shared both TM2 and LB. In all PMIPs, the LE was longer, where two group-conserved sequences were observed. The first one composed of eight residues included the P2 position. The second consisted of six residues and was located downstream of the P3 position. We further identified group-specific consensus sequences, with one in each of TM3 and TM5, and two in TM4. In the upstream region of TM3, there was a YxxxQ motif in all PMIPs, in which xxx contained group-specific residues. In the consensus sequence composed of eleven residues (Pxxxxxxxxx(M/S) in TM5, the interior positions were blocked with group-wise preserved residues. Nevertheless, the first group-wise conserved sequence located at the start of TM4 was composed of eight residues; and the second one located before the end of the TM4 was composed of seven residues.

4. Discussion

MIPs have a main role in water and solute transport and aid in homeostasis during plant stress responses (Afzal et al., 2016). Recently, transcriptome data provided important clues about the involvement of MIPs in host-pathogen interactions (Galindo-González and Deyholos, 2016; Reeksting et al., 2016; Guo et al., 2017). Although there has been no study on PMIPs, in the present study, we identified and characterized a total of 126 MIP homologues from the genomes of six *Phytophthora* species, which cause severe economic losses because of devastating effects on numerous agriculturally and ornamentally important plants (Tyler et al., 2006; Haas et al., 2009; Lamour et al., 2012; Thines, 2014; Fawke et al., 2015; Derevnina et al., 2016; Wang et al., 2016). This study provided comparative information in context of genome-wide number, subclasses or groups, structural insight, and evolution of PMIPs relative to those in taxonomic

441 groups.

442

443 *4.1. PMIPs are a new paradigm in microbial aquaporins*

444 Although the number of MIP homologues varies from organism to organism, plants
445 comparatively have more homologues than animals and microbes. Although bacterial genomes
446 have 1–2 and fugal and/or algal genomes have only 1–6 *MIP* genes (Wang et al., 2005;
447 Anderberg et al., 2011; Azad et al., 2011; Verma et al., 2014), this study showed that the
448 genomes of fungi-like *Phytophthora* species of oomycetes had 17–27 *MIP* genes (Tables 1-6).
449 The number of *MIP* genes in the genomes of *Phytophthora* species was higher than that even in
450 the human genome and almost similar to that in many plants (Maurel et al., 2015; Azad et al.,
451 2016; Potokar et al., 2016). The increase in the number of PMIPs might be have been caused by
452 gene duplication or horizontal gene transfer in addition to the polyploidy nature of *Phytophthora*
453 spp. (Bancroft, 2001; Moore and Purugganan, 2003; Bertier et al., 2013). More interestingly,
454 systematic searching and phylogenetic and structural analysis revealed that PMIPs were distinct
455 and did not cluster to those in taxonomic groups (Figures 2, 3). Data collectively indicated that
456 despite some fungal MIPs (Verma et al., 2014), PMIPs had distinct ar/R filter and FPs, and
457 substitution in the conserved NPA motifs in comparison with those in plants, animals, and algae
458 were divergent. The large numbers, phylogenetical and structural novelty of PMIPs reflected
459 their wide diversity in function and physiological relevance.

460

461 *4.2. Group-specific characteristic C-termini and consensus motifs likely to be associated with*
462 *novel functions*

463 The uneven length of N- and C-termini of PMIPs groups (Table 7, Supplementary Figure 3)
464 might have affected their interaction with other molecules or physical interaction for

heteromerization of PMIPs (Fetter et al., 2004; Yaneff et al., 2014) because the protein termini were generally exposed on the surface of protein structures making them available for interaction (Jacob and Unger, 2007; Tanco et al., 2015). The physical interaction through heteromerization is one of the mechanisms for regulation of intrinsic permeability of MIPs (Yaneff et al., 2014). The C-termini of proteins have been associated with diverse biological functions and processes, such as membrane integration of proteins, protein activity, protein sorting, post-translational modification, protein-protein interaction, or formation of protein complexes (Chung et al., 2002; Azad et al., 2008; Tanco et al., 2015). The conserved positively charged H residue in all PMIPs, which was included in the group-specific C-terminal motif (Table 7), is a novel character not observed in other MIPs. Furthermore, despite MIPEs that are enriched with negatively charged D and E residues, the C-termini in most PMIPs were enriched with positively charged H, although its extent is group-specific (Supplementary Figure 3). Therefore, the distinctive C-termini in PMIPs might be associated with protein sorting, protein-protein interaction, post-translation modification, or other novel functions. KDEL, HDEL, or KKXX motifs in the C-termini of proteins are involved in the retention of protein in the endoplasmic reticulum and prevent them entering into the secretory pathway (Chung et al., 2002; Capitani and Sallese, 2009; Tanco et al., 2015). The NIPs with unusual NPA motifs in LB and LE, have a characteristic R-rich C-termini, which are not seen in NIPs with only one unusual NPA motif, and the R-rich C-termini are thought to be involved in structural stabilization of MIPs (Ishibashi, 2006; Worth and Blundell, 2010; Azad et al., 2016). Although a subgroup of MIPCs with both unusual NPA motifs have no such R-rich C-termini, their conserved RSxGPYE(Y/F) sequence would have been involved in the same function in addition to the other aforementioned functions of C-termini because of having charged and phosphorylatable residues. Similarly, MIPHs having both unusual NPA motifs would have

different functions because of the presence of comparatively more charged (20–40%) residues in the last 10 amino acids of the C-termini (Supplementary Figure 3). In contrast, SIPs with only one unusual NPA motif have K-rich C termini (Azad et al., 2016), which is a potential endoplasmic retention signal (Ishibashi, 2006; Gomes et al., 2009). However, MIPEs-MIPGs have no such K-rich C-termini. It is usually supposed that MIPs with unusual NPA motifs are involved in non-aqua transport rather than water transport (Pommerrenig et al., 2015; Azad et al., 2016). However, water transport activity has been demonstrated in two AtSIP1s, but not in AtSIP2;1, and the latter is supposed to have non-aqua transport activity (Fetter et al., 2004).

Comparison of PMIPs with MIPs in other taxonomic groups revealed that the group-specific motifs or consensus sequences in PMIPs shown in Table 7 are distinct from the corresponding positions in MIPs of other domains of life (Supplementary Figure S2). However, some of these motifs or consensus sequences are partly similar to some of the fungal MIPs. Interestingly, the YxxxQ motif in TM3 is largely conserved in most MIPs of other taxonomic groups. This might have important structural roles in MIPs as is reported for NPA motifs. Some of the residues in the group-specific motif or consensus sequences are pore-lining (Supplementary Figure 3), which in turn, may influence the transport profiles of PMIPs. Experimentally proven non-aqua transporter MIPs in plants have substrate-specific signature sequences (SSSS) or specificity-determining positions (SDPs) in the NPA regions, ar/R filters, and FPs (Hove and Bhave, 2011; Azad et al., 2016). These SSSS and SDPs have been shown to be used as tools to predict non-aqua transport profiles of plant MIPs (Hove and Bhave, 2011; Azad et al., 2016; Azad et al. 2018). Based on the SSSS or SDPs in individual constrictions, several PMIPs were predicted to be non-aqua transporters (Supplementary Table S3). However, considering only the PMIPs commonly predicted by the SSSS or SDPs for all three constrictions, no PMIP was predicted to be non-aqua transporter. This result indicated that the

SSSS and SDPs of PMIPs in the three constrictions might be structurally distinct, which in turn, would have imparted novel functions to PMIPs. It will be interesting to conduct wet-lab experiments to determine the transport activities and elucidate the structure-function relationship of PMIPs.

4.3. PMIPs could be attractive targets for anti-Phytophthora disease

As previously mentioned, *Phytophthora* spp. have been identified to cause devastating diseases in wide range of plants, including agriculturally and economically important crops, such as rice, potato, tomato, wheat, rye, barley, and fruits (Fawke et al., 2015; Reeksting et al., 2016; Wang et al., 2016). These plant pathogens display resistance to existing fungicides, indicating the necessity of new drugs to eradicate *Phytophthora* from crops. To develop new drugs and treatments, the physiology of these pathogenic *Phytophthora* need to be studied in detail, specifically the significance of membrane transporters such as MIPs in plant-*Phytophthora* interactions. Understanding the plant-*Phytophthora* interactions and the infection mechanism is tremendously important and will aid in developing anti-*Phytophthora* drugs. MIPs are known to transport water, glycerol, and other physiologically important small-uncharged solutes (Chaumont and Tyerman, 2014; Afzal et al., 2016; Azad et al., 2016). Our present study revealed that PMIPs are phylogenetically and structurally distinct from their counterpart in other taxonomic domains (Figures 2, 3). The ar/R selectivity filter of PMIPs is unique and hydrophobic in nature and is likely to transport novel hydrophobic solutes. The distinctive FPs and group-specific consensus motifs (Figure 3, Table 7) further support the same notion that PMIPs might have novel transport profiles, which might include essential interactors for plant-*Phytophthora* interactions. The nature of molecules that are transported through PMIPs and the role of group-specific consensus motifs that might have regulated the channel function must first

be experimentally investigated. Hence, PMIP channels in disease causing *Phytophthora* can be considered as attractive drug target like some AQP modulators that are promising agents for the treatment of human disorders (Huber et al., 2012). However, the revelation of 3D structures and crystallographic information of PMIPs will broaden the understanding of drug design against *Phytophthora*-mediated plant diseases.

5. Conclusions

After its discovery in the last decade of the twentieth century, the MIP has fascinated scientists with its potential to aid in understanding the molecular physiology of organisms and development of new innovative pharmacological agents for the treatment of human and plant diseases, owing to their potential structure and functions. The host-pathogen interactions at *Phytophthora*-induced infection require understanding at the molecular level. In this context, understanding the evolution, structure, function, and diversity of PMIP channels might be very significant. In the present study, a genome-wide identification of MIPs has been performed in six *Phytophthora* species, which are devastating plant pathogens and members of oomycetes, a distinct lineage of fungus-like eukaryotic microbes. Every *Phytophthora* genome encodes several-fold MIP homologues compared to other eukaryotic microbes, such as fungi and algae, even more than in the human genome, and a similar number of homologues as found in many plants. The PMIPs were phylogenetically and structurally distinct from their counterparts in other taxonomic domains. Sequence analysis and homology modeling studies indicated that the ar/R selectivity filter, Froger's positions, and group-specific consensus sequences/motifs in different loops and TM helices of PMIPs are distinct from those in other taxonomic domains. The substitutions in the conserved NPA motifs of many PMIPs were also unique. The data collectively support the notion that PMIPs might have novel functions and could be considered

attractive anti-*Phytophthora* targets. Currently, no functional studies are available on PMIPs. The present study has provided a picture of PMIPs with distinct evolutionary relationships and structural properties. However, wet-lab experiments are needed to find out the possible solutes that are transported through the PMIP channels and the importance of these channels in the *Phytophthora* life cycles.

Acknowledgements

JA and AH were supported by stipends from The University Grant Commission of Bangladesh. There was no additional external funding received for this study.

Author's contributions

AKA and JA conceived the project. AKA designed the work. JA, AH, MAA, MMH and MH carried out the work and participated in analysis with the close supervision of AKA. JA participated in draft preparation of some parts of the manuscript. AKA supervised all procedures and wrote the manuscript. TI and YS critically read the manuscript. All authors approved the final version of the manuscript.

Competing interests

The authors declare that they have no competing interests.

References

Abascal, F., Irisarri, I. and Zardoya, R., 2014. Diversity and evolution of membrane intrinsic proteins. Biochimica et Biophysica Acta (BBA)-General Subjects, 1840(5), pp.1468-1481.

Afzal, Z., Howton, T., Sun, Y., Mukhtar, M.S., 2016. The roles of aquaporins in plant stress responses. Journal of Developmental Biology 4, 9.

Anderberg, H.I., Danielson, J.Å., Johanson, U., 2011. Algal MIPs, high diversity and conserved motifs. BMC Evolutionary Biology 11, 110.

Azad, A.K., Ahmed, J., Alum, M.A., Hasan, M.M., Ishikawa, T., Sawa, Y., Katsuhara, M., 2016. Genome-wide characterization of major intrinsic proteins in four grass plants and their non-aqua transport selectivity profiles with comparative perspective. PloS One 11, e0157735.

Azad AK, Ahmed J, Alum MA, Md. Hasan M, Ishikawa T, Sawa Y (2018) Prediction of Arsenic and Antimony Transporter Major Intrinsic Proteins from the Genomes of Crop Plants. International Journal of Biological Macromolecules, 107: 2630-2642.

Azad, A.K., Hanawa, R., Ishikawa, T., Sawa, Y., Shibata, H., 2013. Expression profiles of aquaporin homologues and petal movement during petal development in *Tulipa gesneriana*. Physiologia plantarum 148, 397-407.

Azad, A.K., Katsuhara, M., Sawa, Y., Ishikawa, T., Shibata, H., 2008. Characterization of four plasma membrane aquaporins in tulip petals: a putative homolog is regulated by phosphorylation. Plant and Cell Physiology 49, 1196-1208.

Azad, A.K., Sato, R., Ohtani, K., Sawa, Y., Ishikawa, T., Shibata, H., 2011. Functional characterization and hyperosmotic regulation of aquaporin in *Synechocystis* sp. PCC 6803. Plant Science 180, 375-382.

Azad, A.K., Sawa, Y., Ishikawa, T., Shibata, H., 2004. Phosphorylation of plasma membrane aquaporin regulates temperature-dependent opening of tulip petals. Plant and Cell Physiology 45, 608-617.

Azad, A.K., Yoshikawa, N., Ishikawa, T., Sawa, Y., Shibata, H., 2012. Substitution of a single amino acid residue in the aromatic/arginine selectivity filter alters the transport profiles of tonoplast aquaporin homologs. Biochimica et Biophysica Acta (BBA)-Biomembranes 1818, 1-11.

Baker, N., Glover, L., Munday, J.C., Andrés, D.A., Barrett, M.P., De Koning, H.P., Horn, D., 2012. Aquaglyceroporin 2 controls susceptibility to melarsoprol and pentamidine in African trypanosomes. Proceedings of the National Academy of Sciences 109, 10996-11001.

Bancroft, I., 2001. Duplicate and diverge: the evolution of plant genome microstructure. Trends in Genetics 17, 89-93.

Beakes, G.W., Glockling, S.L., Sekimoto, S., 2012. The evolutionary phylogeny of the oomycete "fungi". Protoplasma 249, 3-19.

Beitz, E., 2005. Aquaporins from pathogenic protozoan parasites: structure, function and potential for chemotherapy. Biology of the Cell 97, 373-383.

Bertier, L., Leus, L., D'hondt, L., de Cock, A.W., Höfte, M., 2013. Host adaptation and speciation through hybridization and polyploidy in Phytophthora. PloS one 8, e85385.

Blackman, L.M., Cullerne, D.P., Torreña, P., Taylor, J., Hardham, A.R., 2015. RNA-Seq analysis of the expression of genes encoding cell wall degrading enzymes during infection of lupin (Lupinus angustifolius) by Phytophthora parasitica. PloS one 10, e0136899.

Capitani, M., Sallese, M., 2009. The KDEL receptor: new functions for an old protein. FEBS Letters 583, 3863-3871.

Chaumont, F., Tyerman, S.D., 2014. Aquaporins: highly regulated channels controlling plant water

relations. Plant Physiology 164, 1600-1618.
Chung, J.-J., Shikano, S., Hanyu, Y., Li, M., 2002. Functional diversity of protein C-termini: more than zipcoding? Trends in cell biology 12, 146-150.
Danielson, J.Å., Johanson, U., 2008. Unexpected complexity of the aquaporin gene family in the moss *Physcomitrella patens*. BMC Plant Biology 8, 45.
Danielson, J.Å., and Johanson, U. (2010) Phylogeny of major intrinsic proteins. MIPs and their role in the exchange of metalloids, edited by Tomas P. Jahn and Gerd P. Bienert. Landes Bioscience and Springer Science+Business Media, PP 19-31.
Derevnina, L., Petre, B., Kellner, R., Dagdas, Y.F., Sarowar, M.N., Giannakopoulou, A., De la Concepcion, J.C., Chaparro-Garcia, A., Pennington, H.G., van West, P., 2016. Emerging oomycete threats to plants and animals. Phil. Trans. R. Soc. B 371, 20150459.
Fadiel, A., Isokpehi, R.D., Stambouli, N., Hamza, A., Benammar-Elgaaied, A., Scalise, T.J., 2009. Protozoan parasite aquaporins. Expert review of proteomics 6, 199-211.
Fawke, S., Doumane, M., Schornack, S., 2015. Oomycete interactions with plants: infection strategies and resistance principles. Microbiology and Molecular Biology Reviews 79, 263-280.
Fetter, K., Van Wilder, V., Moshelion, M., Chaumont, F., 2004. Interactions between plasma membrane aquaporins modulate their water channel activity. The Plant Cell 16, 215-228.
Froger, A., Thomas, D., Delamarche, C., Tallur, B., 1998. Prediction of functional residues in water channels and related proteins. Protein Science 7, 1458-1468.
Fu, D., Libson, A., Miercke, L.J., Weitzman, C., Nollert, P., Krucinski, J., Stroud, R.M., 2000. Structure of a glycerol-conducting channel and the basis for its selectivity. Science 290, 481-486.
Galindo-González, L., Deyholos, M.K., 2016. RNA-seq Transcriptome Response of Flax (*Linum usitatissimum* L.) to the Pathogenic Fungus *Fusarium oxysporum* f. sp. lini. Frontiers in plant science 7.
Gao, Z., He, X., Zhao, B., Zhou, C., Liang, Y., Ge, R., Shen, Y., Huang, Z., 2010. Overexpressing a putative aquaporin gene from wheat, TaNIP, enhances salt tolerance in transgenic Arabidopsis. Plant and Cell Physiology 51, 767-775.
Giovannetti, M., Balestrini, R., Volpe, V., Guether, M., Straub, D., Costa, A., Ludewig, U., Bonfante, P., 2012. Two putative-aquaporin genes are differentially expressed during arbuscular mycorrhizal symbiosis in Lotus japonicus. BMC plant biology 12, 186.
Gomes, D., Agasse, A., Thiébaud, P., Delrot, S., Gerós, H., Chaumont, F., 2009. Aquaporins are multifunctional water and solute transporters highly divergent in living organisms. Biochimica et Biophysica Acta (BBA)-Biomembranes 1788, 1213-1228.
Guo, S., Zuo, Y., Zhang, Y., Wu, C., Su, W., Jin, W., Yu, H., An, Y., Li, Q., 2017. Large-scale transcriptome comparison of sunflower genes responsive to *Verticillium dahliae*. BMC Genomics 18, 42.
Gupta, A.B., Sankararamakrishnan, R., 2009. Genome-wide analysis of major intrinsic proteins in the tree plant *Populus trichocarpa*: characterization of XIP subfamily of aquaporins from evolutionary perspective. BMC Plant Biology 9, 134.
Haas, B.J., Kamoun, S., Zody, M.C., Jiang, R.H., Handsaker, R.E., Cano, L.M., Grabherr, M., Kodira, C.D., Raffaele, S., Torto-Alalibo, T., 2009. Genome sequence and analysis of the Irish potato famine pathogen *Phytophthora infestans*. Nature 461, 393-398.
Hansen, M., Kun, J.F., Schultz, J.E., Beitz, E., 2002. A Single, Bi-functional Aquaglyceroporin in Blood-stagePlasmodium falciparum Malaria Parasites. Journal of Biological Chemistry 277, 4874-4882.

679 Hao, C., Xia, Z., Fan, R., Tan, L., Hu, L., Wu, B., Wu, H., 2016. De novo transcriptome sequencing
680 of black pepper (Piper nigrum L.) and an analysis of genes involved in phenylpropanoid
681 metabolism in response to Phytophthora capsici. BMC Genomics 17, 822.
682 Heymann, J.B., Engel, A., 2000. Structural clues in the sequences of the aquaporins. Journal of
683 Molecular Biology 295, 1039-1053.
684 Hove, R.M., Bhave, M., 2011. Plant aquaporins with non-aqua functions: deciphering the signature
685 sequences. Plant molecular biology 75, 413-430.
686 Huber, V.J., Tsujita, M., Nakada, T., 2012. Aquaporins in drug discovery and pharmacotherapy.
687 Molecular aspects of medicine 33, 691-703.
688 Ishibashi, K., 2006. Aquaporin subfamily with unusual NPA boxes. Biochimica et Biophysica Acta
689 (BBA)-Biomembranes 1758, 989-993.
690 Ishibashi, K., Hara, S., Kondo, S., 2009. Aquaporin water channels in mammals. Clinical and
691 Experimental Nephrology 13, 107-117.
692 Jackson, K., Yin, J., Ji, P., 2012. Sensitivity of Phytophthora capsici on vegetable crops in Georgia
693 to mandipropamid, dimethomorph, and cyazofamid. Plant disease 96, 1337-1342.
694 Jacob, E., Unger, R., 2007. A tale of two tails: why are terminal residues of proteins exposed?
695 Bioinformatics 23, e225-e230.
696 Johnson, G.K., Rosana, B.O., Vijesh, K.I., Santhosh, E.J., Anandaraj, M., 2016. Interplay of genes
697 in plant-pathogen interactions: In planta expression and docking studies of a beta 1, 3
698 glucanase gene from *Piper colubrinum* and a glucanase inhibitor gene from *Phytophthora*
699 *capsici*. Physiology and molecular biology of plants: an international journal of functional
700 plant biology 22, 567-573.
701 Kroon, L., Bakker, F., Van Den Bosch, G., Bonants, P., Flier, W., 2004. Phylogenetic analysis of
702 Phytophthora species based on mitochondrial and nuclear DNA sequences. Fungal Genetics
703 and Biology 41, 766-782.
704 Kun, J.F., de Carvalho, E.G., 2009. Novel therapeutic targets in *Plasmodium falciparum*:
705 aquaglyceroporins. Expert opinion on therapeutic targets 13, 385-394.
706 Lamour, K.H., Mudge, J., Gobena, D., Hurtado-Gonzales, O.P., Schmutz, J., Kuo, A., Miller, N.A.,
707 Rice, B.J., Raffaele, S., Cano, L.M., 2012. Genome sequencing and mapping reveal loss of
708 heterozygosity as a mechanism for rapid adaptation in the vegetable pathogen *Phytophthora*
709 *capsici*. Molecular Plant-Microbe Interactions 25, 1350-1360.
710 Levitt, M., 1992. Accurate modeling of protein conformation by automatic segment matching.
711 Journal of Molecular Biology 226, 507-533.
712 Li, T., Hu, Y.J., Hao, Z.P., Li, H., Wang, Y.S., Chen, B.D., 2013. First cloning and characterization
713 of two functional aquaporin genes from an arbuscular mycorrhizal fungus *Glomus*
714 *intraradices*. New Phytologist 197, 617-630.
715 Liu, P., Gong, J., Ding, X., Jiang, Y., Chen, G., Li, B., Weng, Q., Chen, Q., 2016. The L-type Ca2+
716 Channel Blocker Nifedipine Inhibits Mycelial Growth, Sporulation, and Virulence of
717 Phytophthora capsici. Frontiers in Microbiology 7.
718 Lomsadze, A., Ter-Hovhannisyan, V., Chernoff, Y.O., Borodovsky, M., 2005. Gene identification in
719 novel eukaryotic genomes by self-training algorithm. Nucleic acids research 33, 6494-6506.
720 Maurel, C., Boursiac, Y., Luu, D.-T., Santoni, V., Shahzad, Z., Verdoucq, L., 2015. Aquaporins in
721 plants. Physiological reviews 95, 1321-1358.
722 Meng, Y., Huang, Y., Wang, Q., Wen, Q., Jia, J., Zhang, Q., Huang, G., Quan, J., Shan, W., 2015.
723 Phenotypic and genetic characterization of resistance in *Arabidopsis thaliana* to the
724 oomycete pathogen *Phytophthora parasitica*. Frontiers in Plant Science 6.
725 Montalvetti, A., Rohloff, P., Docampo, R., 2004. A functional aquaporin co-localizes with the

vacuolar proton pyrophosphatase to acidocalcisomes and the contractile vacuole complex of *Trypanosoma cruzi*. Journal of Biological Chemistry 279, 38673-38682.

Moore, R.C., Purugganan, M.D., 2003. The early stages of duplicate gene evolution. Proceedings of the National Academy of Sciences 100, 15682-15687.

Mukhopadhyay, R., Bhattacharjee, H., Rosen, B.P., 2014. Aquaglyceroporins: generalized metalloid channels. Biochimica et Biophysica Acta (BBA)-General Subjects 1840, 1583-1591.

Muto, Y., Segami, S., Hayashi, H., Sakurai, J., Murai-Hatano, M., Hattori, Y., Ashikari, M., Maeshima, M., 2011. Vacuolar proton pumps and aquaporins involved in rapid internode elongation of deepwater rice. Bioscience, Biotechnology, and Biochemistry 75, 114-122.

Pellegrini-Calace, M., Maiwald, T., Thornton, J.M., 2009. PoreWalker: a novel tool for the identification and characterization of channels in transmembrane proteins from their three-dimensional structure. PLoS Computional Biology 5, e1000440.

Peng, Y., Lin, W., Cai, W., Arora, R., 2007. Overexpression of a *Panax ginseng* tonoplast aquaporin alters salt tolerance, drought tolerance and cold acclimation ability in transgenic Arabidopsis plants. Planta 226, 729-740.

Pettersson, N., Filipsson, C., Becit, E., Brive, L., Hohmann, S., 2005. Aquaporins in yeasts and filamentous fungi. Biology of the Cell 97, 487-500.

Pommerrenig, B., Diehn, T.A., Bienert, G.P., 2015. Metalloido-porins: Essentiality of Nodulin 26-like intrinsic proteins in metalloid transport. Plant Science 238, 212-227.

Potokar, M., Jorgačevski, J., Zorec, R., 2016. Astrocyte Aquaporin Dynamics in Health and Disease. International Journal of Molecular Sciences 17, 1121.

Reeksting, B.J., Olivier, N.A., Van den Berg, N., 2016. Transcriptome responses of an ungrafted Phytophthora root rot tolerant avocado (*Persea americana*) rootstock to flooding and Phytophthora cinnamomi. BMC Plant Biology 16, 205.

Sahoo, D.K., Abeysekara, N.S., Cianzio, S.R., Robertson, A.E., Bhattacharyya, M.K., 2017. A Novel *Phytophthora sojae* Resistance Rps12 Gene Mapped to a Genomic Region That Contains Several Rps Genes. PloS one 12, e0169950.

Sonah, H., Deshmukh, R.K., Labbé, C., Bélanger, R.R., 2017. Analysis of aquaporins in Brassicaceae species reveals high-level of conservation and dynamic role against biotic and abiotic stress in canola. Scientific Reports 7.

Soveral, G., Casini, A., 2017. Aquaporin modulators: a patent review (2010–2015). Expert opinion on therapeutic patents 27, 49-62.

Sui, H., Han, B.-G., Lee, J.K., Walian, P., Jap, B.K., 2001. Structural basis of water-specific transport through the AQP1 water channel. Nature 414, 872-878.

Tajkhorshid, E., Nollert, P., Jensen, M.Ø., Miercke, L.J., O'connell, J., Stroud, R.M., Schulten, K., 2002. Control of the selectivity of the aquaporin water channel family by global orientational tuning. Science 296, 525-530.

Tanco, S., Gevaert, K., Damme, P., 2015. C-terminomics: Targeted analysis of natural and posttranslationally modified protein and peptide C-termini. Proteomics 15, 903-914.

Thines, M., 2014. Phylogeny and evolution of plant pathogenic oomycetes--a global overview. European Journal of Plant Pathology 138, 431.

Törnroth-Horsefield, S., Wang, Y., Hedfalk, K., Johanson, U., Karlsson, M., Tajkhorshid, E., Neutze, R., Kjellbom, P., 2006. Structural mechanism of plant aquaporin gating. Nature 439, 688-694.

Tyler, B.M., Tripathy, S., Zhang, X., Dehal, P., Jiang, R.H., Aerts, A., Arredondo, F.D., Baxter, L., Bensasson, D., Beynon, J.L., 2006. Phytophthora genome sequences uncover evolutionary origins and mechanisms of pathogenesis. Science 313, 1261-1266.

Uehlein, N., Kaldenhoff, R., 2008. Aquaporins and plant leaf movements. Annals of botany 101, 1-4.

Verkman, A., 2012. Aquaporins in clinical medicine. Annual Review of Medicine 63, 303-316.

Verma, R.K., Prabh, N.D., Sankararamakrishnan, R., 2014. New subfamilies of major intrinsic proteins in fungi suggest novel transport properties in fungal channels: implications for the host-fungal interactions. BMC Evolutionary Biology 14, 173.

Wallace, I.S., Roberts, D.M., 2004. Homology modeling of representative subfamilies of Arabidopsis major intrinsic proteins. Classification based on the aromatic/arginine selectivity filter. Plant Physiology 135, 1059-1068.

Wang, X.-W., Guo, L.-Y., Han, M., Shan, K., 2016. Diversity, evolution and expression profiles of histone acetyltransferases and deacetylases in oomycetes. BMC Genomics 17, 927.

Wang, Y., Schulten, K., Tajkhorshid, E., 2005. What makes an aquaporin a glycerol channel? A comparative study of AqpZ and GlpF. Structure 13, 1107-1118.

Worth, C.L., Blundell, T.L., 2010. On the evolutionary conservation of hydrogen bonds made by buried polar amino acids: the hidden joists, braces and trusses of protein architecture. BMC Evolutionary Biology 10, 161.

Yaneff, A., Sigaut, L., Marquez, M., Alleva, K., Pietrasanta, L.I., Amodeo, G., 2014. Heteromerization of PIP aquaporins affects their intrinsic permeability. Proceedings of the National Academy of Sciences 111, 231-236.

Yoshida, K., Schuenemann, V.J., Cano, L.M., Pais, M., Mishra, B., Sharma, R., Lanz, C., Martin, F.N., Kamoun, S., Krause, J., 2013. The rise and fall of the Phytophthora infestans lineage that triggered the Irish potato famine. elife 2, e00731.

Zhang, D.Y., Ali, Z., Wang, C.B., Xu, L., Yi, J.X., Xu, Z.L., Liu, X.Q., He, X.L., Huang, Y.H., Khan, I.A., 2013. Genome-wide sequence characterization and expression analysis of major intrinsic proteins in soybean (*Glycine max* L.). PLoS One 8, e56312.

Figure legends

Figure 1. Superposition of representative 3D models of PcaMIPC1;1 constructed with the template AQP1 (red), GlpF (violet), and SoPIP2;1 (green). A and B are side and top views of the superposed models, respectively. The 3D models of all PMIPs were first constructed separately with AQP1, GlpF, and SoPIP2;1, and then the model of PcaMIPC1;1 as representative of PMIPs was superposed. The TM α-helices (H1-H6) and loops (LA-LE) are indicated. The NPA motifs in LB and LE are shown as sticks. The residues that form the aromatic/arginine (ar/R) tetrad in the superposed structures are shown as sticks (B). The TM α-helices and the loops to which these residues belong are indicated.

Figure 2. Evolutionary relationship of MIPs in the six *Phytophthora* species. Phylogenetic analysis of all PMIPs from the six *Phytophthora* species is shown along with MIPs from *Populus trichocarpa* (PtMPs; PtPIPs, PtTIPs, PtNIPs, PtSIPs, and PtXIPs), *Physcomitrella patens* (PpMIPs), and human (MAQP), and five representative members from each of 10 subfamilies of MIPs in fungi (Verma et al. 2014; AQPs, Alpha AQGPs, Beta AQGPs, Gamma 1 AQGPs, Gamma 2 AQGPs, Delta AQGPs, Fps like AQGPs, XIPs, SIP-like Fungus MIP, and Yfl054-like AQGPs) and all MIPs in the genomes of nine algae (Anderberg et al. 2011; all homologues start with 'A'). The deduced amino acid sequences of MIPs were aligned using the Clustal Omega computer program and a phylogenetic tree was constructed using MEGA. The evolutionary history was inferred using the Bootstrap Neighbor-Joining (1000 replicates) method and the genetic distance was estimated by the p-distance method. PMIPs are shown with cyan background. All PMIPs, except PpaMIPI1;1, formed a distinct clade and did not cluster with any subfamily or group of MIPs in plants, humans, fungi, or algae. Only PPaMIPI1;1 clustered with fungal MIP Delta AQGPs.

844

845 **Figure 3.** Grouping of MIPs from *Phytophthora* species based on phylogeny showing ar/R
846 selectivity filter, NPA motifs, FPs, and diameter in the ar/R selectivity filter region. The
847 phylogenetic tree was generated as described in Figure 2. The residues in the ar/R selectivity
848 filter, NPA motifs, and the FPs were selected from the 3D models, as well as from the alignment
849 shown in Supplementary Figures S3. The pore diameter at the ar/R region, which is considered
850 one of the reasons for transport selectivity, was determined as described in the methods and
851 tabulated on the right side.

852

853 **Figure 4.** Gene structure of MIPs from *Phytophthora* species. Exon-intron organizations of *MIP*
854 genes from *Phytophthora* species were depicted as described in the methods. The *PMIPs* having
855 introns are mentioned in the second column. The six TM regions are shown as black bars and
856 the loops B and E are shown in diamond shapes. The intron positions are indicated by inverted
857 triangles. The number within the parenthesis in the first column indicates the total number of
858 PMIP homologues.

859

860 **Figure 5.** Superposition of 3D models of representative PMIP homologues of group A, B, C,
861 and H. The 3D models of all PMIPs were first constructed separately with SoPIP2;1, and then
862 the models of PciMIPA1;1 (red), PciMIPB1;1 (green), PciMIPC1;1 (yellow), and PciMIPH1;2
863 (cyan) were superposed (A). The 3D models of PciMIPA1;1 (B), PciMIPB1;1 (C),
864 PciMIPC1;1(D), and PciMIPH1;2 (E) were separately superposed with the template SoPIP2;1.
865 The SoPIP2;1 is shown in green and the PMIP models are in red (B-E). The residues that form
866 the aromatic/arginine (ar/R) tetrad in the superposed structures are shown as sticks. The TM α-
867 helices and the loops to which these residues belong are indicated.

Supplementary Figure S1: Grouping of MIPs from *Phytophthora* species based on phylogeny. The phylogenetic tree was generated as described in Figure 2. Each MIP group is shown with a specific background color to distinguish them from others.

Supplementary Figure S2: Multiple alignment of MIPs shown in Figure 2. The amino acid sequences were aligned using the Clustal Omega program. The TM helices and the dual NPA motifs are shown as gray and yellow, respectively. The residues in the ar/R selectivity filter and FP are green and cyan boxed, respectively.

Supplementary Figure S3: Multiple sequence alignment of MIPs in the six *Phytophthora* species. The amino acid sequences were aligned using the Clustal Omega program. Each MIP group is shown with a specific background color to distinguish them from others. The TM helices are shown within boxes with black lines and the dual NPA motifs are shown as gray. The residues in the ar/R selectivity filter and FP are red and blue boxed, respectively. The region of the TM helices and loops from which consensus sequences or motifs were depicted in Table 7 are green boxed. The pore-lining residues are indicated by arrows above the alignment and the conserved residues are indicated by stars (*) at the bottom of the alignment.

Supplementary Table S1: Predicted sub-cellular localization of PMIPs

Supplementary Table S2: Average pairwise sequence identity and similarity between different PMIP groups.

Table 1: *MIP* genes in *P infestans* identfed from the whole genome shortgun (WGS).

Gene Name	Accession No.		Genomic location	PPL (aa)	Maximum identity with MIP (%) in					Ka/Ks Value
	Fungi DB	NCBI			Phytophthora	Human taxid 9606	*A. thaliana* taxid 3702	Fungi taxid 4751	Algae taxid 3027	
PinMIPA1;1	PITG_09354	XP_002903624	DS028131:2,512,989..2,513,854(+)	263	KUF83438 (83)[c]	BAA24864 (38)	OAO95087 (33)	XP_016613015 (47)	XP_005824571 (26)	0.2387
PinMIPA1;2	PITG_09355	XP_002903625	DS028131:2,546,498..2,547,370(+)	290	ETI40982 (96)[a]	BAA24864 (43)	NP_198598 (31)	XP_016613015 (51)	XP_005824571 (30)	0.2434
PinMIPA1;3	PITG_09356	XP_002903626	DS028131:2,548,127..2,549,032(+)	301	ETP38902 (93)[a]	EAW53211(43)	AAF30303 (31)	XP_016613015 (51)	XP_005824571 (29)	0.4144
PinMIPB1;1	PITG_09350	XP_002903621	DS028131:2,486,995..2,487,897(+)	300	XP_008894610 (95)[a]	XP_016870189(40)	NP_198597 (28)	XP_016613015 (47)	XP_005826812 (29)	0.5579
PinMIPC1;1	PITG_08296	XP_002903694	DS028130:793,924..794,826(+)	300	XP_002903695 (99)[b]	EAW53211 (39)	NP_174472 (29)	XP_016613015 (40)	XP_005824571 (28)	0.1783
PinMIPC1;2	PITG_08297	XP_002903695	DS028130:798,418..799,320(-)	300	XP_002903694 (99)[b]	EAW53211 (39)	NP_174472 (30)	XP_016613015 (41)	XP_005824571 (27)	3.4424
PinMIPC1;3	PITG_01100	XP_002909619	DS028118:6,345,258..6,347,464(-)	364	XP_002909620(81)[a]	XP_011508406 (37)	NP_198597 (29)	KIK26743 (40)	XP_005824571 (26)	0.8084
PinMIPC1;4	PITG_01101	XP_002909620	DS028118:6,347,827..6,348,720(+)	297	ETP27623 (98)[a]	EAW53211 (38)	NP_178191 (32)	KNE66934 (42)	XP_005824571 (26)	0.4934
PinMIPD1;1	PITG_09383	XP_002903644	DS028131:2,824,064..2,824,888(-)	274	XP_008894631(96)[a]	NP_066190 (42)	AAF30303 (33)	XP_016613015 (50)	XP_005824571 (30)	0.1845
PinMIPD1;2	PITG_09384	XP_002903645	DS028131:2,825,233..2,826,057(-)	274	ETO69638 (98)[a]	NP_066190 (42)	OAO95087 (32)	KNE66934 (47)	XP_005824571 (27)	0.1479
PinMIPD1;3	PITG_09386	XP_002903646	DS028131:2,830,938..2,831,762(+)	274	ETK81038 (98)[a]	NP_066190 (43)	NP_198598 (28)	KNE66934 (46)	XP_005824571 (27)	0.118
PinMIPE1;1	PITG_09203	XP_002903501	DS028131:1,363,292..1,364,282(-)	295	KUF89515 (93)[c]	NP_004916 (36)	OAO95087 (30)	XP_016613015 (40)	XP_005824571 (26)	0.2632
PinMIPE1;2	PITG_09204	XP_002903502	DS028131:1,364,676..1,365,540(-)	270	XP_008894432(89)[a]	BAA24864 (40)	NP_174472 (31)	XP_016613015 (44)	XP_005824571 (30)	0.2192
PinMIPF1;1	PITG_09091	XP_002903420	DS028131:283,842..284,768(+)	308	ETP11188 (96)[a]	BAA24864 (40)	OAP08944 (33)	XP_016613015 (43)	XP_005824571 (28)	0.4186
PinMIPG1;1	PITG_09197	XP_002903496	DS028131:1,281,315..1,282,223(-)	302	ETO69869 (87)[a]	NP_066190 (39)	AAM61294 (30)	XP_016613015 (49)	XP_005824571 (27)	0.2059
PinMIPG1;2	PITG_09196	XP_002903495	DS028131:1,266,835..1,267,745(-)	272	KUF89527 (85)[c]	CAG46822 (39)	CAA68906 (30)	CDO57616 (46)	XP_005824571 (25)	0.2134
PinMIPG1;3	PITG_09193	XP_002903492	DS028131:1,254,794..1,255,735(-)	313	ETP10991(95)[a]	EAW53211 (39)	AAM61294 (31)	XP_016613015 (50)	XP_005826794 (28)	0.1653

PPL: polypeptide length, aa: amino acid.

[a]*P. parasitica*, [b]*P. infestans*, [c]*P. nicotianae*

Parenthesis indicates the percentage of identity at the amino acid level

898 **Table 2: *MIP* genes in *P. parasitica* identfed from the whole genome shortgun (WGS).**

899

Gene Name	Accession No.		Genomic location	PPL (aa)	Maximum identity with MIP (%) in					Ka/Ks Value
	Fungi DB	NCBI			Phytophthora	Human taxid 9606	*A. thaliana* taxid 3702	Fungi taxid 4751	Algae taxid 3027	
PpaMIPA1;1	PPTG_21368	XP_008894870	KI669564:2,505,888..2,508,408(-)	294	ETM40999 (99)[b]	BAA24864 (43)	OAO95087 (33)	XP_016613015 (51)	XP_005824571 (31)	0.242
PpaMIPA1;2	PPTG_03679	XP_008894612	KI669564:2,503,959..2,504,957(-)	290	ETI40982 (99)[b]	BAA24864 (43)	NP_198598 (31)	XP_016613015 (51)	XP_005824571 (29)	0.3506
PpaMIPA1;3	PPTG_03683	XP_008894614	KI669564:2,516,194..2,517,439(+)	311	KUF81753 (99)[a]	BAA24864 (43)	NP_174472 (33)	XP_016613015 (51)	XP_005824571 (30)	0.28
PpaMIPA1;4	PPTG_03678	XP_008894611	KI669564:2,502,439..2,503,474(-)	301	ETP38902 (99)[b]	BAA24864 (42)	OAO95087 (31)	XP_016613015 (50)	XP_005824571 (29)	0.2414
PpaMIPB1;1	PPTG_03677	XP_008894610	KI669564:2,500,793..2,501,695(+)	300	XP_002903621(99)[c]	NP_066190 (38)	NP_198597 (29)	XP_016613015 (47)	XP_005826812 (29)	0.4847
PpaMIPC1;1	PPTG_12775	XP_008908228	KI669595:679,441..680,883(-)	300	ETI42987 (99)[b]	EAW53211 (42)	NP_174472 (31)	XP_016613015 (42)	XP_005824571 (28)	0.3463
PpaMIPC1;2	PPTG_10261	XP_008903172	KI669579:270,346..271,914(-)	296	XP_008903173 (99)[b]	EAW53211 (39)	NP_178191 (31)	KNE66934 (42)	XP_005824571 (26)	0.6476
PpaMIPC1;3	PPTG_10262	XP_008903173	KI669579:273,175..274,136(+)	296	ETP27623 (99)[b]	EAW53211 (39)	NP_178191 (31)	KNE66934 (42)	XP_005824571 (26)	0
PpaMIPD1;1	PPTG_03700	XP_008894631	KI669564:2,594,866..2,595,834(-)	274	ETK81036 (99)[b]	BAA24864 (43)	NP_174472 (33)	XP_015224218 (46)	XP_005824571 (30)	0.226
PpaMIPD1;2	PPTG_03701	XP_008894632	KI669564:2,595,983..2,596,984(-)	274	ETP38884 (99)[b]	BAA24864 (41)	AAM65333 (31)	KNE66934 (48)	XP_005824571 (26)	0.1461
PpaMIPD1;3	PPTG_03702	XP_008894633	KI669564:2,598,571..2,599,395(+)	274	KUF96448 (99)[a]	BAA24864 (43)	NP_198598 (29)	KNE66934 (46)	005824571 (27)	0.1638
PpaMIPE1;1	PPTG_03509	XP_008894431	KI669564:1,787,887..1,789,461(-)	328	ETL87939 (99)[b]	CAG46822 (40)	NP_198597 (30)	XP_016613015 (44)	XP_005824571 (26)	0.2602
PpaMIPE1;2	PPTG_03510	XP_008894432	KI669564:1,789,897..1,790,946(-)	349	XP_002903502 (89)[c]	NP_066190 (42)	NP_174472 (29)	XP_016613015 (45)	XP_005824571 (29)	0.2728
PpaMIPF1;1	PPTG_03354	XP_008894280	KI669564:1,303,524..1,305,303(+)	309	KUF92657 (99)[a]	NP_536354 (41)	AAM61294 (30)	XP_016613015 (42)	XP_005824571 (29)	0.3905
PpaMIPF1;2	PPTG_03358	XP_008894285	KI669564:1,314,285..1,315,223(+)	312	ETP39322 (99)[b]	NP_536354 (40)	OAP08944.1 (32)	XP_018224918 (43)	XP_005824571 (28)	0.2058
PpaMIPF1;3	PPTG_03512	XP_008894434	KI669564:1,792,870..1,793,826(-)	318	ETL87944 (99)[b]	NP_066190 (41)	AAM61294 (31)	XP_016613015 (41)	XP_005824571 (28)	0.3234
PpaMIPG1;1	PPTG_03508	XP_008894430	KI669564:1,785,403..1,786,911(-)	311	ETO69869 (99)[b]	EAW53211 (41)	AAM61294 (31)	XP_016613015 (50)	XP_005824571 (27)	0.2455
PpaMIPG1;2	PPTG_03502	XP_008894424	KI669564:1,774,974..1,776,155(-)	298	ETO69896 (99)[b]	EAW53211 (43)	NP_174472 (32)	XP_016613015 (49)	XP_005836263 (32)	0.286
PpaMIPG1;3	PPTG_03507	XP_008894429	KI669564:1,784,034..1,785,409(-)	346	ETI41209 (99)[b]	NP_004916 (43)	NP_198597 (27)	XP_016613015 (45)	XP_005824571 (25)	0.425
PpaMIPG1;4	PPTG_03504	XP_008894426	KI669564:1,780,246..1,781,526(+)	317	ETP10991 (99)[b]	EAW53211 (38)	AAM61294 (30)	XP_016613015 (46)	XP_005826794 (28)	0.2374
PpaMIPH1;1	PPTG_03511	XP_008894433	KI669564:1,791,348..1,792,473(-)	322	KUF89513 (95)[a]	BAA24864 (35)	NP_198597 (26)	XP_016613015 (36)	XP_005840893 (25)	0.6009
PpaMIPI1;1	PPTG_14248	XP_008909057	KI669599:124,031..125,191(-)	386	ETL88792 (99)[b]	NP_536354 (26)	CAA16760 (27)	OAJ43010 (34)	XP_005826812 (26)	0.6177

900 PPL: polypeptide length, aa: amino acid; [a]*P. nicotianae*, [b]*P. parasitica*, [c]*P. infestans;* parenthesis indicates the percentage of identity at the amino acid level

901

902

903

904

905 **Table 3: *MIP* genes in *P. sojae* identfed from the whole genome shortgun (WGS).**

Gene Name	Accession No.		Genomic location	PPL (aa)	Maximum identity with MIP (%) in					Ka/Ks Value
	Fungi DB	NCBI			Phytophthora	Human taxid 9606	*A. thaliana* taxid 3702	Fungi taxid 4751	Algae taxid 3027	
PsoMIPA1;1	PHYSODRAFT_339641	XP_009535925	JH159160:2,208,709..2,209,590(-)	293	XP_009535926 (92)[b]	CAG46822 (48)	OAO95087 (30)	XP_016613015 (52)	XP_005824571 (31)	0.1184
PsoMIPA1;2	PHYSODRAFT_288735	XP_009535926	JH159160:2,210,235..2,211,326(-)	294	KUF83438 (92)[d]	NP_004916(45)	OAO95087 (32)	XP_016613015 (52)	XP_005824571 (30)	0.247
PsoMIPA1;3	PHYSODRAFT_525185	XP_009535924	JH159160:2,206,900..2,207,772(-)	290	ETI40982 (96)[a]	CAG46822 (48)	XP_016613015 (52)	XP_016613015 (52)	XP_005824571 (31)	0.373
PsoMIPA1;4	PHYSODRAFT_525056	XP_009535927	JH159160:2,214,127..2,215,062(+)	311	ETP38886 (86)[a]	EAW53211(42)	NP_198597 (31)	XP_016613015 (46)	XP_005824571 (30)	0.2643
PsoMIPA1;5	PHYSODRAFT_525520	XP_009535923	JH159160:2,205,229..2,206,137(-)	302	XP_002903626(80)[c]	NP_004916(45)	OAO95087(31)	XP_016613015(49)	XP_005824571(30)	0.2899
PsoMIPB1;1	PHYSODRAFT_355956	XP_009535922	JH159160:2,203,313..2,204,352(+)	300	XP_008894610(88)[a]	NP_001161 (38)	NP_198597 (28)	XP_016613015 (44)	XP_005826812 (29)	0.6376
PsoMIPC1;1	PHYSODRAFT_330305	XP_009525250	JH159153:10,045,045..10,045,959(+)	304	XP_009529571(75)[b]	EAW53211(40)	NP_174472 (33)	XP_016613015 (39)	XP_005824571(30)	0.2843
PsoMIPC1;2	PHYSODRAFT_334026	XP_009529571	JH159155:4,679,293..4,680,189(+)	298	XP_008908228 (95)[a]	NP_536354(40)	CAA68906 (31)	XP_016613015 (40)	XP_005824571 (28)	0.2371
PsoMIPC1;3	PHYSODRAFT_468333	XP_009514460	JH159151:2,362,247..2,363,155(-)	302	XP_009514453 (75)[b]	CAG46822 (42)	NP_198597 (27)	XP_016613015 (40)	XP_005824571 (26)	0.3452
PsoMIPC1;4	PHYSODRAFT_477070	XP_009514453	JH159151:2,342,386..2,343,285(-)	299	XP_002909620 (83)[a]	EAW53211(38)	NP_198597 (27)	KNE66934 (41)	XP_005824571 (25)	0.461
PsoMIPD1;1	PHYSODRAFT_525349	XP_009535952	JH159160:2,282,738..2,283,562(-)	274	XP_008894631 (97)[a]	BAA24864 (43)	AAF30303 (32)	XP_016613015 (50)	XP_005824571 (30)	0.4964
PsoMIPD1;2	PHYSODRAFT_524483	XP_009535953	JH159160:2,283,997..2,284,821(-)	274	XP_002903645 (96)[c]	BAA24864 (42)	OAO95087 (32)	KNE66934 (49)	XP_005824571 (25)	0.2716
PsoMIPD1;3	PHYSODRAFT_339665	XP_009535954	JH159160:2,286,447..2,287,271(+)	274	ETK81038 (95)[a]	BAA24864 (41)	NP_198598 (29)	KNE66934 (47)	XP_005824571 (26)	0.158
PsoMIPE1;1	PHYSODRAFT_339490	XP_009535744	JH159160:1,691,393..1,692,387(-)	319	ETL87939 (82)[a]	NP_536354(36)	NP_198597 (29)	XP_016613015 (47)	XP_005826794 (25)	0.361
PsoMIPF1;1	PHYSODRAFT_252268	XP_009535532	JH159160:1,090,627..1,091,388(+)	253	ETP11188 (88)[a]	CAG46822 (41)	AAM61294 (32)	XP_016613015 (47)	XP_005826812 (31)	0.1857
PsoMIPF1;2	PHYSODRAFT_259896	XP_009535755	JH159160:1,719,459..1,720,510(+)	340	XP_009535748 (71)[b]	BAA24864 (39)	OAP08944 (32)	XP_016613015 (39)	XP_005826812 (31)	0.2486
PsoMIPF1;3	PHYSODRAFT_549775	XP_009535748	JH159160:1,700,410..1,703,096(-)	323	XP_008894434 (88)[a]	EAW53211(41)	NP_198597 (30)	XP_016613015 (49)	XP_005826794 (25)	0.3053
PsoMIPG1;1	PHYSODRAFT_355937	XP_009535743	JH159160:1,687,843..1,688,930(-)	304	ETO69869 (89)[a]	NP_066190(40)	NP_174472 (30)	XP_016613015 (52)	XP_005824571 (27)	0.6348
PsoMIPG1;2	PHYSODRAFT_524694	XP_009535734	JH159160:1,659,268..1,660,167(-)	299	KUF84847 (91)[d]	EAW53211(41)	NP_180986 (29)	XP_016613015 (48)	XP_005824571 (25)	0.3263
PsoMIPG1;3	PHYSODRAFT_549767	XP_009535742	JH159160:1,685,479..1,686,495(-)	265	ETI41209 (93)[a]	CAG46822 (44)	NP_198597 (28)	XP_016613015 (46)	XP_005824571 (24)	0.1758
PsoMIPG1;4	PHYSODRAFT_355919	XP_009535727	JH159160:1,644,341..1,647,821(-)	312	KUF89525 (89)[d]	EAW53211(40)	AAM61294 (30)	XP_016613015 (46)	XP_005826794 (28)	0.4995
PsoMIPG1;5	PHYSODRAFT_339474	XP_009535728	JH159160:1,647,826..1,648,716(-)	296	KUG01632 (89)[d]	EAW53211 (39)	AAM61294 (30)	XP_016613015 (48)	XP_005826794 (28)	0.5854
PsoMIPG1;6	PHYSODRAFT_549764	XP_009535739	JH159160:1,671,210..1,672,275(-)	322	XP_009535740 (97)[b]	EAW53211 (39)	NP_192776 (30)	XP_016613015 (51)	XP_005826794 (28)	0.3187
PsoMIPG1;7	PHYSODRAFT_524365	XP_009535740	JH159160:1,682,070..1,683,032(+)	320	XP_009535739 (97)[b]	EAW53211(40)	NP_198597 (29)	XP_016613015 (50)	XP_005826794 (27)	0.1838
PsoMIPG1;8	PHYSODRAFT_451744	XP_009535736	JH159160:1,664,850..1,665,770(-)	307	XP_009535737 (99)[b]	CAG46822 (40)	NP_198597 (29)	XP_016613015 (49)	XP_005824571 (25)	0
PsoMIPG1;9	PHYSODRAFT_452243	XP_009535737	JH159160:1,666,625..1,667,545(+)	307	XP_009535736 (99)[b]	CAG46822 (40)	NP_198597 (29)	XP_016613015 (49)	XP_005824571 (25)	1.6327
PsoMIPH1;1	PHYSODRAFT_259868	XP_009535754	JH159160:1,717,811..1,718,779(+)	247	XP_008894433 (54)[c]	BAA24864 (33)	AAM61294 (28)	XP_016613015 (33)	XP_005824571 (26)	0.4791

906 PPL: polypeptide length, aa: amino acid; [a]*P. parasitica*, [b]*P. sojae*, [c]*P. infestans* ,[d]*P. nicotianae*, parenthesis indicates the percentage of identity at the amino acid level

907 **Table 4: *MIP* genes in *P. ramorum* identfed from the whole genome shortgun (WGS).**
908

Gene Name	Accession No.		Genomic location	PPL (aa)	Maximum identity with MIP (%) in					Ka/Ks Value
	Fungi DB	NCBI			Phytophthora	Human taxid 9606	*A. thaliana* taxid 3702	Fungi taxid 4751	Algae taxid 3027	
PraMIPA1;1	PSURA_84475	-	PramPr-102_SC0114:47,366..48,247(-)	293	KUF83438 (92)[d]	BAA24864 (41)	NP_198598 (31)	XP_016613015 (31)	XP_005824571 (29)	0.1842
PraMIPA1;2	PSURA_84474	-	PramPr-102_SC0114:45,578..46,447(-)	289	ETI40982 (96)[a]	BAA24864 (35)	NP_198598 (31)	XP_016613015 (49)	XP_005824571 (27)	0.171
PraMIPA1;3	PSURA_84476	-	PramPr-102_SC0114:50,260..51,177(-)	305	ETP38886.1 (88)[a]	EAW53211 (41)	NP_198597 (30)	XP_016613015 (45)	XP_005826794 (27)	0.2641
PraMIPB1;1	PSURA_72233	-	PramPr-102_SC0114:42,371..43,273(+)	300	XP_008894610 (94)[a]	NP_066190 (38)	NP_198597 (29)	XP_016613015 (46)	XP_005826812 (30)	0.5878
PraMIPC1;1	PSURA_71169	-	PramPr-102_SC0006:392,090..392,875(+)	261	XP_008908228 (94)[a]	EAW53211 (40)	AAM61294 (31)	XP_016613015 (41)	XP_005824571 (27)	0.439
PraMIPC1;2	PSURA_72183	-	PramPr-102_SC0102:39,914..40,669(-)	251	XP_009514460 (94)[b]	CAG46822 (41)	AAM61294 (31)	XP_016613015 (39)	XP_005824571 (25)	0.1716
PraMIPC1;3	PSURA_72182	-	PramPr-102_SC0102:33,841..34,596(+)	251	XP_009514453 (84)[b]	CAG46822 (41)	AAM61294 (30)	KNE66934 (42)	XP_005824571 (28)	0.212
PraMIPD1;1	PSURA_72235	-	PramPr-102_SC0114:104,901..105,626(-)	241	XP_009535952 (96)[b]	BAA24864 (44)	NP_174472 (33)	XP_016613015 (50)	XP_005824571 (30)	0.1736
PraMIPD1;2	PSURA_72236	-	PramPr-102_SC0114:106,106..106,930(-)	274	XP_002903645 (96)[c]	NP_066190 (40)	OAO95087 (31)	XP_016613015 (49)	XP_005824571 (26)	0.1811
PraMIPD1;3	PSURA_72237	-	PramPr-102_SC0114:108,512..109,339(+)	275	XP_009535954 (95)[b]	NP_066190 (41)	NP_198598 (29)	KNE66934 (47)	XP_005824571 (26)	0.2497
PraMIPE1;1	PSURA_80174	-	PramPr-102_SC0046:362,803..363,804(+)	333	XP_008894431 (87)[a]	BAA24864 (40)	NP_198597 (32)	XP_016613015 (50)	XP_005824571 (26)	0.4705
PraMIPF1;1	PSURA_95006	-	PramPr-102_SC0029:82,759..83,706(+)	315	ETP11188 (80)[a]	BAA24864 (39)	AAM61294 (31)	XP_016613015 (47)	XP_005824571 (27)	0.1983
PraMIPF1;2	PSURA_80164	-	PramPr-102_SC0046:326,376..327,332(+)	318	ETP11188 (83)[a]	BAA24864 (38)	AAM61294 (31)	XP_016613015 (44)	XP_005824571 (28)	0.2214
PraMIPG1;1	PSURA_71778	-	PramPr-102_SC0046:367,948..368,772(-)	274	ETO69869 (93)[a]	NP_536354 (42)	AAM61294 (30)	XP_016613015 (51)	XP_005824571 (27)	0.2148
PraMIPG1;2	PSURA_87677	-	PramPr-102_SC2368:1,241..2,140(-)	299	XP_009535734 (93)[b]	EAW53211 (41)	NP_180986 (31)	XP_016613015 (49)	XP_005826794 (25)	0
PraMIPG1;3	PSURA_72426	-	PramPr-102_SC0716:5,808..6,707(-)	299	XP_009535734 (93)[b]	EAW53211(41)	NP_180986 (31)	XP_016613015 (49)	XP_005824571 (25)	0
PraMIPG1;4	PSURA_71777	-	PramPr-102_SC0046:365,813..366,649(-)	278	ETI41209 (91)[a]	CAG46822 (42)	NP_198597 (31)	XP_016613015 (46)	XP_005841067 (30)	0.29
PraMIPG1;5	PSURA_72419	-	PramPr-102_SC0660:6,538..7,374(+)	278	KUF89525.1 (91)[d]	EAW53211(41)	AAM61294(29)	XP_016613015 (48)	XP_005826794 (28)	0.252
PraMIPG1;6	PSURA_72425	-	PramPr-102_SC0716:955..1,704(-)	249	XP_009535728 (92)[b]	NP_004916 (43)	NP_192776 (31)	XP_016613015 (49)	XP_005826794 (28)	0.4837

909 PPL: polypeptide length, aa: amino acid; [a]*P. parasitica*, [b]*P. sojae*, [c]*P. infestans* , [d]*P. nicotianae;* Parenthesis indicates the percentage of identity at the amino acid level
910
911
912

Table 5: *MIP* genes in *P. capsici* identfed from the whole genome shortgun (WGS).

Gene Name	Accession No.		Genomic location	PPL (aa)	Maximum identity with MIP (%) in					Ka/Ks Value
	Fungi DB	NCBI			Phytophthora	Human taxid 9606	*A. thaliana* taxid 3702	Fungi taxid 4751	Algae taxid 3027	
PcaMIPA1;1	PHYCA_530062	-	PcapLT1534_SC054:141,263..142,555(-)	293	KUF83438 (95)[d]	CAG46822 (45)	OAO95087 (32)	XP_016613015 (51)	XP_005824571 (32)	0.288
PcaMIPA1;2	PHYCA_510138	-	PcapLT1534_SC054:139,634..140,630(-)	288	XP_009535925 (94)[b]	CAG46822 (44)	OAO95087 (33)	XP_016613015 (51)	XP_005824571 (31)	0.2203
PcaMIPA1;3	PHYCA_553594	-	PcapLT1534_SC054:138,017..138,944(-)	291	XP_009535924 (96)[b]	NP_066190 (43)	OAO95087 (32)	XP_016613015 (51)	XP_005824571 (31)	0.1602
PcaMIPA1;4	PHYCA_553588	-	PcapLT1534_SC054:136,640..137,545(-)	301	ETP38902 (90)[a]	EAW53211 (43)	NP_178191 (32)	XP_016613015 (48)	XP_005824571 (30)	0.4023
PcaMIPC1;1	PHYCA_6911	-	PcapLT1534_SC015:427,750..438,233(-)	280	XP_008908228 (87)[a]	EAW53211 (35)	NP_567572 (30)	XP_016613015 (39)	XP_005824571 (29)	0.425
PcaMIPC1;2	PHYCA_123018	-	PcapLT1534_SC049:168,741..169,683(+)	297	XP_009514460 (90)[b]	EAW53211 (38)	NP_198597 (26)	XP_016613015 (40)	XP_005824571 (26)	0.2412
PcaMIPC1;3	PHYCA_509760	-	PcapLT1534_SC049:192,973..193,950(+)	299	XP_002909620 (91)[c]	EAW53211 (39)	NP_198597 (27)	KNE66934 (40)	XP_005824571 (26)	0.3848
PcaMIPD1;1	PHYCA_553631	-	PcapLT1534_SC054:213,802..214,626(-)	274	XP_009535952 (97)[b]	BAA24864 (42)	AAF30303 (32)	XP_016613015 (50)	XP_005824571 (30)	0.2115
PcaMIPD1;2	PHYCA_510151	-	PcapLT1534_SC054:214,814..215,848(-)	274	XP_002903645 (97)[c]	NP_066190 (42)	NP_198597 (33)	KNE66934 (48)	XP_005824571 (27)	0.1879
PcaMIPD1;3	PHYCA_124758	-	PcapLT1534_SC054:222,559..223,383(+)	274	ETK81038 (96)[a]	BAA24864 (42)	NP_001330234 (29)	KNE66934 (47)	XP_005824571 (27)	0.2442
PcaMIPE1;1	PHYCA_130684	-	PcapLT1534_SC096:30,774..31,649(-)	291	XP_008894431 (95)[a]	CAG46822 (41)	OAO95087(31)	XP_016613015 (46)	XP_005824571 (27)	0.2135
PcaMIPE1;2	PHYCA_130620	-	PcapLT1534_SC096:32,191..33,045(-)	284	XP_008894432 (92)[a]	BAA24864 (39)	NP_174472 (29)	XP_016613015 (46)	XP_005824571 (28)	0.2094
PcaMIPF1;1	PHYCA_125677	-	PcapLT1534_SC059:215,629..216,567(+)	312	ETP11193 (88)[a]	NP_001161 (38)	OAP08944 (33)	XP_016613015 (41)	XP_005824571 (30)	0.2323
PcaMIPF1;2	PHYCA_130673	-	PcapLT1534_SC096:35,031..35,984(-)	317	XP_008894434 (86)[a]	BAA24864 (42)	AAM61294 (29)	XP_016613015 (42)	XP_005826812 (31)	0.2055
PcaMIPG1;1	PHYCA_511729	-	PcapLT1534_SC096:27,663..28,733(-)	274	ETO69869 (93)[a]	EAW53211 (41)	AAM61294 (32)	XP_016613015 (52)	XP_005824571 (28)	0.188
PcaMIPG1;2	PHYCA_511737	-	PcapLT1534_SC096:84,419..89,905(-)	300	XP_009535734 (92)[b]	EAW53211 (41)	AAM61294 (31)	XP_016613015 (51)	XP_005824571 (25)	0.2582
PcaMIPG1;3	PHYCA_511725	-	PcapLT1534_SC096:15,964..17,034(-)	304	ETI41209 (87)[a]	NP_004916 (39)	CAA68906 (29)	XP_016613015 (44)	XP_005841067 (32)	0.1959
PcaMIPG1;4	PHYCA_512201	-	PcapLT1534_SC669:271..1,219(+)	297	XP_009535 27 (80)[b]	EAW53211 (38)	AAM61294 (29)	CDO57616 (46)	XP_005826794 (28)	0.2401
PcaMIPG1;5	PHYCA_130634	-	PcapLT1534_SC096:96,478..97,393(+)	268	XP_009535727 (85)[b]	EAW53211 (38)	NP_192776 (27)	XP_016613015 (42)	XP_005826794 (27)	0.6911

PPL: polypeptide length, aa: amino acid; [a]*P. parasitica*, [b]*P. sojae*, [c]*P. infestans*, [d]*P. nicotianae;* parenthesis indicates the percentage of identity at the amino acid level

928 **Table 6: *MIP* genes in *P. cinnamomi* identfed from the whole genome shortgun (WGS).**
929

Gene Name	Accession No.		Genomic location	PPL (aa)	Maximum identity with MIP (%) in					Ka/Ks Value
	Fungi DB	NCBI			Phytophthora	Human taxid 9606	*A. thaliana* taxid 3702	Fungi taxid 4751	Algae taxid 3027	
PciMIPA1;1	PHYCI_94066	-	PcinCBS144-22_SC0062:223,255..224,286(+)	294	XP_009535926 (94)[a]	CAG46822 (45)	NP_198598(30)	XP_016613015 (52)	XP_005824571 (29)	0.3049
PciMIPA1;2	PHYCI_94067	-	PcinCBS144-22_SC0062:224,851..225,867(+)	293	XP_009535925 (91)[a]	BAA24864 (44)	OAO95087 (32)	XP_016613015 (51)	XP_005824571 (30)	0.1199
PciMIPA1;3	PHYCI_229810	-	PcinCBS144-22_SC0062:226,809..227,681(+)	290	ETI40982 (97)[c]	CAG46822 (47)	NP_198598 (31)	XP_016613015 (52)	XP_005824571 (30)	0.1787
PciMIPA1;4	PHYCI_9455	-	PcinCBS144-22_SC0062:220,189..222,911(+)	381	XP_009535927 (76)[a]	NP_004916 (39)	OAO95087 (28)	XP_016613015 (44)	XP_005824571 (28)	0.3056
PciMIPA1;5	PHYCI_313549	-	PcinCBS144-22_SC0062:228,193..229,286(+)	301	XP_009535923 (89)[a]	BAA2486 (45)	OAO95087 (31)	XP_016613015 (49)	XP_005824571 (30)	0.3364
PciMIPB1;1	PHYCI_229797	-	PcinCBS144-22_SC0062:231,204..232,106(-)	300	XP_009535922 (91)[a]	NP_066190 (37)	NP_198597 (29)	XP_016613015 (46)	XP_005826812 (29)	0.8309
PciMIPC1;1	PHYCI_93687	-	PcinCBS144-22_SC0055:271,156..272,257(-)	298	XP_009529571 (96)[a]	EAW53211 (39)	NP_567572 (29)	XP_016613015 (39)	XP_005824571 (26)	0.331
PciMIPC1;2	PHYCI_91871	-	PcinCBS144-22_SC0024:340,304..341,393(-)	302	XP_009514460 (95)[a]	NP_004916 (41)	NP_198597 (25)	ESA22038(40)	XP_005824571 (26)	0.1932
PciMIPC1;3	PHYCI_91869	-	PcinCBS144-22_SC0024:333,891..335,020(+)	299	XP_009514453 (97)[a]	EAW53211 (38)	NP_198597 (26)	KNE66934 (39)	XP_005824571 (24)	0.486
PciMIPD1;1	PHYCI_313498	-	PcinCBS144-22_SC0062:137,089..138,034(+)	274	XP_009535952 (99)[a]	BAA24864 (43)	AAF30303 (32)	XP_016613015 (50)	XP_005824571 (30)	0.2961
PciMIPD1;2	PHYCI_94042	-	PcinCBS144-22_SC0062:135,806..136,774(+)	275	XP_002903645 (97)[b]	NP_066190 (41)	OAO95087 (33)	KNE66934 (48)	XP_005824571 (26)	0.2372
PciMIPD1;3	PHYCI_9427	-	PcinCBS144-22_SC0062:133,313..134,252(-)	289	XP_009535954 (97)[a]	NP_066190 (41)	NP_198598 (30)	KNE66934 (47)	XP_005824571 (26)	0.1816
PciMIPE1;1	PHYCI_254750	-	PcinCBS144-22_SC0271:71,261..72,250(+)	329	XP_009535744 (88)[a]	EAW53211 (40)	NP_198597 (32)	XP_016613015 (44)	XP_005824571(27)	0.2887
PciMIPE1;2	PHYCI_254766	-	PcinCBS144-22_SC0271:69,814..70,674(+)	265	XP_002903502 (88)[b]	BAA24864 (38)	NP_174472 (31)	XP_016613015 (43)	XP_005824571 (29)	0.194
PciMIPF1;1	PHYCI_92025	-	PcinCBS144-22_SC0027:233,171..234,286(-)	314	ETP11188.1 (83)[c]	BAA24864 (38)	OAP08944(32)	XP_016613015 (46)	XP_005824571 (27)	0.243
PciMIPF1;2	PHYCI_437341	-	PcinCBS144-22_SC0271:58,255..59,614(-)	311	ETP11188 (79)[c]	NP_066190 (39)	AAM61294 (32)	XP_016613015 (46)	XP_005824571 (25)	0.2497
PciMIPF1;3	PHYCI_328142	-	PcinCBS144-22_SC0271:66,502..67,452(+)	316	XP_009535748 (88)[a]	BAA24864 (40)	AAM61294 (30)	XP_016613015 (41)	XP_005824571 (26)	0.3537
PciMIPG1;1	PHYCI_98973	-	PcinCBS144-22_SC0271:77,191..78,471(+)	360	ETI41209 (84)[c]	NP_004916 (39)	NP_198597 (29)	XP_016613015 (43)	XP_005841067 (30)	0.2493
PciMIPG1;2	PHYCI_319932	-	PcinCBS144-22_SC0114:181,147..182,144(+)	302	XP_009535727 (88)[a]	NP_004916 (43)	AAM61294 (30)	XP_016613015 (45)	XP_005824571 (25)	0.211
PciMIPG1;3	PHYCI_319929	-	PcinCBS144-22_SC0114:178,016..179,052(-)	314	XP_008894426 (86)[c]	EAW53211 (39)	NP_192776 (29)	XP_016613015 (44)	XP_005826794 (28)	0.1665
PciMIPH1;1	PHYCI_254774	-	PcinCBS144-22_SC0271:47,972..49,023(-)	281	XP_009535754 (81)[a]	BAA24864 (32)	AAM61294 (29)	KJX95442 (33)	XP_005824571 (28)	0.2739
PciMIPH1;2	PHYCI_34767	-	PcinCBS144-22_SC0271:68,054..69,058(+)	321	XP_008894433 (82)[b]	BAA24864 (38)	NP_566271 (28)	XP_013271045(38)	XP_005825107 (33)	0.4769

930 PPL: polypeptide length, aa: amino acid; [a]*P. sojae*, [b]*P. infestans*, [c]*P. parasitica*; parenthesis indicates the percentage of identity at the amino acid level
931

Table 7: PMIP group-specific consensus sequences/motif in the corresponding regions in different TM helices and loops.

PMIPs Group	N-termini	LA	LB*	LC	LD	LE	TM3	TM4	TM5	C-termini
MIPAs	-	QV(V/T)LS	GIH(V/A)(S/C)GGVS	P(M/L)(F/I)x(H/V/T)(I/T/V)D	GDE(L/M)NKP	GMN***(T/S)***G(***Y/F***)AIAGWGS(R/E)	Y(I/V)(I/V)(A/S)Q	(V/I/T)GTCF(L/M)TELL(L/M/V)(G/C)(G/C)(I/L)FA	P(G/S)AVALLVVAIGM	E(H/Q)HH
MIPBs	-	QSILS	GIHI(A/G)GGVS	PLLNV(A/V)D	LDQHNRP	***S****V*NTG***C***AIAGWGSH	YVAAQ	NF(I/V)AFLTELVGGIFA	PSAVALLVVGIGM	GHHH
MIPCs	-	QV(T/V)LS	GI(T/S)(V/I)(G/V)GGVS	P(M/L)(E/L)N(E/I)VD	(L/Q)DQHNRP	(G/A)MNTG***L***AI(***A/V***)G***W***GS(***R/H***)	Y(I/L)L(V/N)Q	NFTCFLTEL(V/I)(V/L)G((I/F)LA	P(P/A)A(I/V)G(V/A)LVS(A/T)IAM	GYHH
MIPDs[a]	Short	QFVLS	G(V/I)HF(S/A)GGVS	PW(F/L)D(I/L/V)(V/H/Q)D	(G/C)DQ(L/I)NKP	G***L***(D/N)TGYALA(A/G)GWGWK	Y(F/I)(V/I)AQ	NWx(A/G)(L/F)ANELV(S/G)GIFA	P(G/A)AV(A/G)L(M/L)LT(C/G)(I/V)GM	Ex(H/F)H
MIPEs[b]	Long	QVNLS	GVYVAEGIS	Q(R/N)(L/I)M(D/E)ED	TDERNRG	GMNTGYA***L***AG(***F***/***Y)***GPK	YAxAQ	NLTAFYSELL(M/I/L)AIYA	PFAF(A/S)(L/M)LFMGLGM	Q(L/I)QH
MIPFs	-	QVxLS	GVYCSEGVS	Q(K/R)(I/L)xKxD	TDQ(N/R)NRS	GMNTGYA***M***AG***Y***GSK(***W***/F)	Y(W/C/V)(L/V/A)(S/A)Q	NxTAFYTELLLC(I/V)YA	PFAFALMIMALGM	QxQH
MIPGs	Long	QVTNS	GVYCSEGIS	QNLNV(I/V)D	TD(K/Q/T)RNR(P/S/A)	GMNTGYA***(V/I)***AG***W***GSK	YMxSQ	NYTAFYTEL(V/L)L(G/S/A)(I/V)YA	(P/A)FAF(C/A)L(M/L)I(M/W)ALGM	E(I/M)HH
MIPHs		QVA(L/I)(S/W)	G(T/V)x(V/T)x(V/T)A(D/E)VS	(Q/E)((K/N)I(R/S)(R/A) (E/G)D	KD(K/R)RNRW	SLNTGLGx(A/I)GY(H/K)-(M/E)	YMxAQ	NYTAFYTE(I/L)MLASYA	PFALALLVTAIS	ExQH
MIPI	-		ATFIASPGS		R-QPVQP	TPFTQACL	YMLFQ	AGGAFVLEFVLVHRV	PLYIGSALAALTM	

[a]PMIP group containing MIPs of short C-termini.
[b]PMIP group containing MIPs of D and E-enrich C-termini
*Sequence shares TM2 and LB
LE, the first sequence is from the upstream of NPA motif and the second sequence is from five residues downstream of the P3.
LC motif starts one residue downstream of P1
TM4, the first sequence starts the TM4 and the second sequence is just before the last residue the TM4.
C-termini, adjacent to the TM6

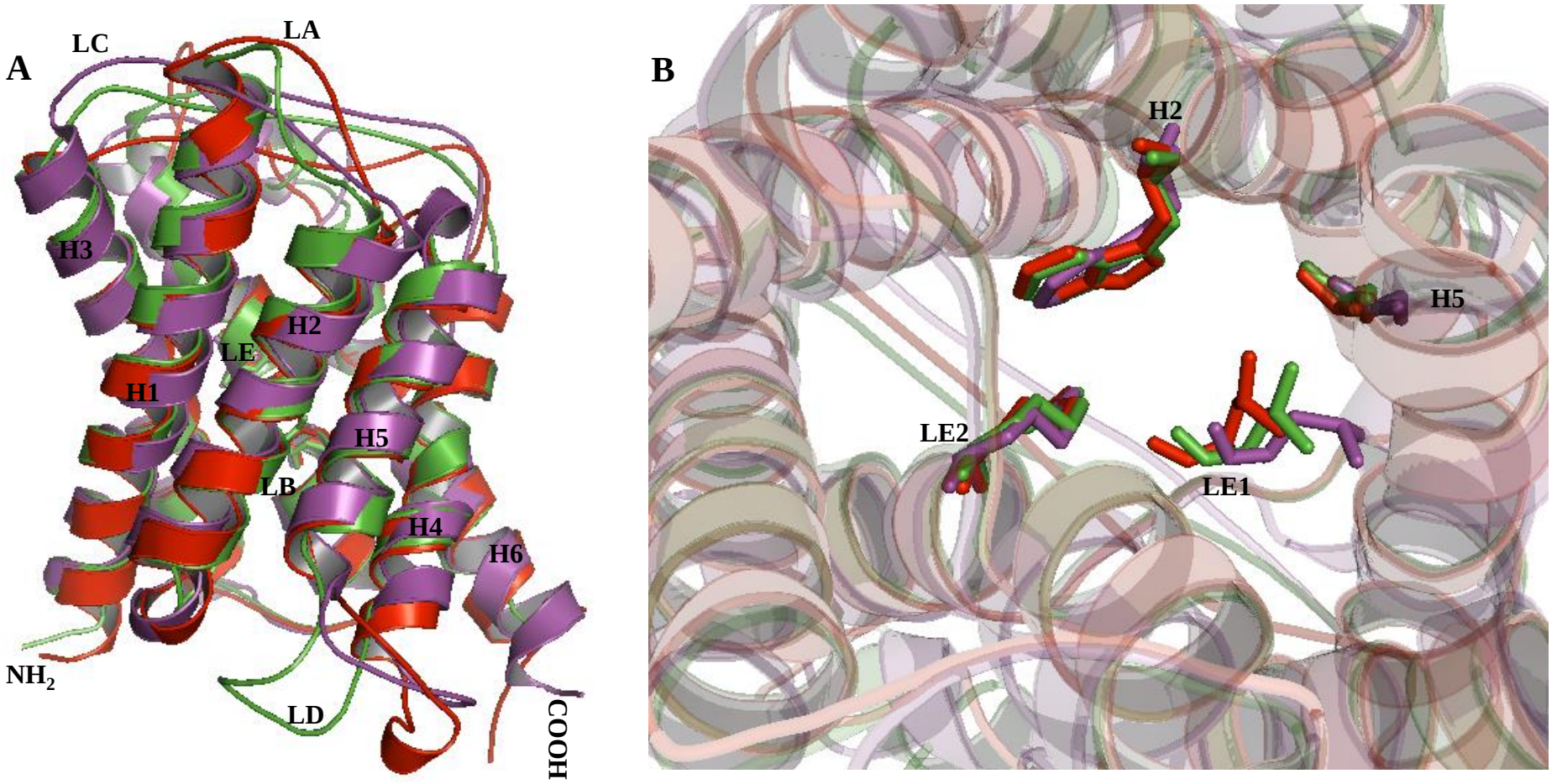


Figure 1

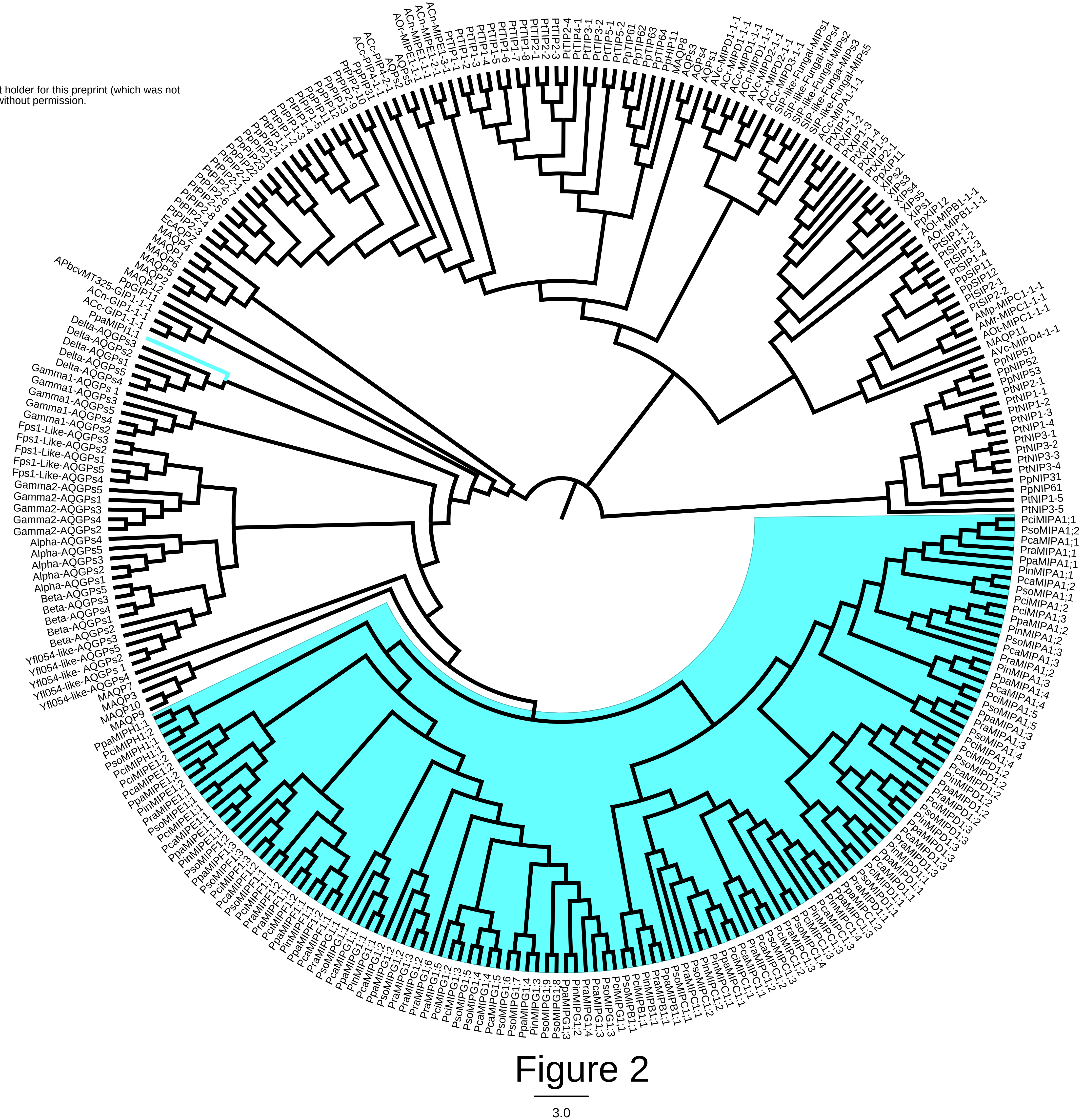


Figure 2

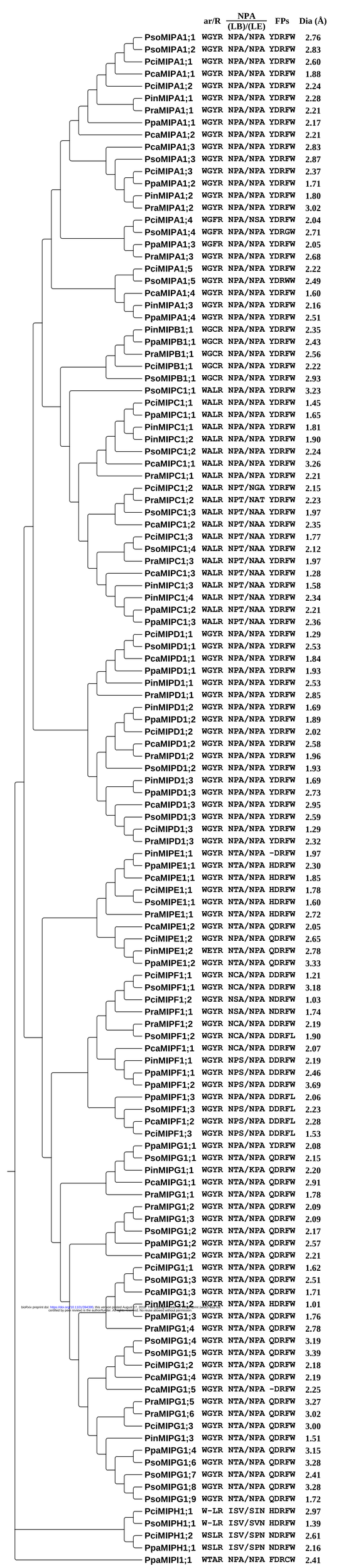


Figure 3

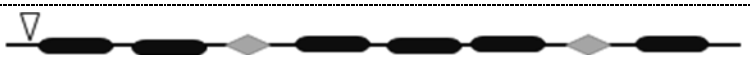
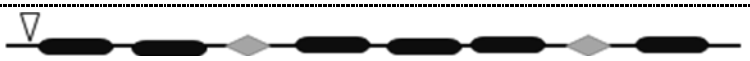
Intron Number	Gene name	General schematic structure of *Phyotpthora MIP* genes with introns' position
No intron (107)	*PMIPs*	
Single Intron (13)	*PcaMIPC1;1*	
	PcaMIPC1;2 , PcaMIPG1;5	
	PcaMIPG1;2, PciMIPD1;3	
	PcaMIPG1;4	
	PciMIPE1;2	
	PinMIPE1;2	
	PinMIPA1;1	
	PpaMIPA1;1	
	PsoMIPE1;1	
	PsoMIPF1;3	
	PsoMIPF1;2	
Two Introns (4)	*PciMIPH1;1*	
	PinMIPG1;2	
	PinMIPE1;1	
	PsoMIPH1;1	
Three Introns (1)	*PciMIPA1;4*	
Five Introns (1)	*PinMIPC1;3*	

Figure 4

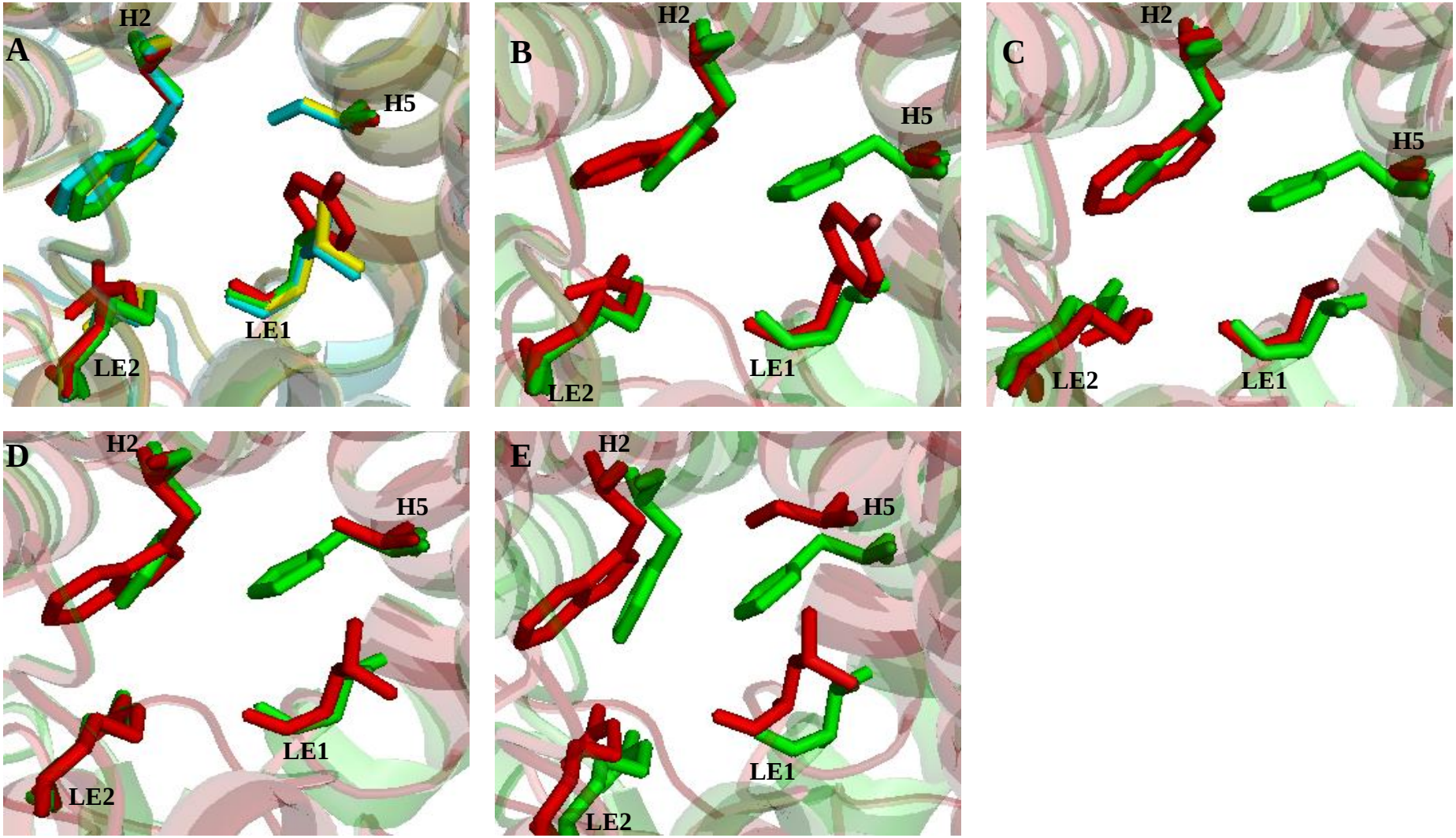

A
H2
H5
LE1
LE2
B
H2
H5
LE1
LE2
C
H2
H5
LE1
LE2
D
H2
H5
LE1
LE2
E
H2
H5
LE1
LE2

Figure 5