Plant Omics Journal POJ

POJ 5(6):590-596 (2012) ISSN:1836-3644

Rhizobiales-like protein phosphatases (Rhilphs): A role in plant defence responses?

Mikhail A. Kutuzov[1*], Alexandra V. Andreeva[1*]

[1]Department of Pharmacology, University of Illinois at Chicago, Chicago, USA
[1]Department of Medicine, Northwestern University, Chicago, USA
[1]Department of Medicine, Northwestern University, 240 E. Huron St., M-300, Chicago, IL 60611, USA

***Corresponding authors: m.kutuzov@usa.net; alexandravandreeva@gmail.com**

Abstract

We previously identified a novel group of protein phosphatases ("*Rhilphs*"), shared by plants and some α-Proteobacteria, including purple photosynthetic bacteria. In this work, we (1) identified genes that show expression correlation with *Rhilphs*; (2) examined the physiological stimuli affecting *Rhilph* expression, and (3) examined characterised *Arabidopsis thaliana* (L.) Heynh. mutants with altered *Rhilph* expression. We found that *Rhilph* expression correlated best with the genes associated with defence responses, carbohydrate metabolism, membrane trafficking and cell wall modification. Using available expression profiling data, we found that *Rhilph-1* (but not *Rhilph-2*) is induced by flg22 treatment (but not by its inactive analogue from Agrobacterium tumefaciens). Both isoforms are induced in response to *Pseudomonas syringae* infection. This induction is impaired in *A. thaliana* mutants deficient in salicylic acid production. Examination of available data for characterised *A. thaliana* mutants showed that Rhilph-1 expression is elevated in plants lacking MPK4 or both MKK1 and MKK2, components of MAP kinase signalling that regulates innate immune responses. Taken together, these data suggest that Rhilph functions are likely associated with defence responses / innate immunity and cell wall-related processes, possibly cell wall remodelling during pathogen attack.

Keywords: cell wall; defence; expression profiling; protein phosphatase, protein phosphorylation.
Abbreviations: GO, gene ontology; HMA, heavy metal-associated; LRR, leucine-rich repeats; MAS5, Microarray Suite 5.0; NBS, nucleotide binding site; RBOH, respiratory burst oxidase homologues; RMA, Robust Multi-Array Average; SA, salicylic acid.

Introduction

Reversible protein phosphorylation plays an essential role in regulation of all cellular functions and requires a finely tuned balance between the activity of protein kinases and phosphatases. Several years ago, we reported the existence in plants of unusual protein phosphatases of the PPP superfamily, which are structurally more similar to bacterial phosphatases than to canonical phosphatases from eukaryotes (Andreeva and Kutuzov, 1999; Andreeva and Kutuzov, 2004). One of the groups of these "bacterial-like" phosphatases is most closely related to PPP phosphatases from some α -Proteobacteria, including Rhizobiales and purple bacteria, and was therefore termed Rhizobiales /Rhodobacterales / Rhodospirillaceae-like phosphatases, or *Rhilphs*. *Rhilphs* are conserved in all land plants (Andreeva and Kutuzov, 2004). Since *Rhilphs* could not be found in any other organisms, including algae, we hypothesised that they may have been acquired by horizontal gene transfer after plants started colonising land and developed symbiosis with N_2-fixing bacteria (Andreeva and Kutuzov, 2004). To date, *Rhilphs* remain completely uncharacterised. In this report, we re-assessed Rhilph phylogeny and aimed to predict possible biological functions of these protein phosphatases using publicly available expression profiling data, the same strategy as we recently used to predict roles and functional partners of Shelphs, another uncharacterised group of protein phosphatases that crosses the prokaryote-eukaryote boundary (Kutuzov and Andreeva, 2012).

Results and discussion

Phylogenetic analysis of Rhilphs

As our previous analysis that led to identification of *Rhilphs* (Andreeva and Kutuzov, 2004) was performed almost a decade ago, we took advantage of much more representative sequence data now available, and re-assessed the inventory of *Rhilphs* throughout different domains of life. In line with our previous observations, *Rhilphs* could be detected in all land plants and in a number of α-proteobacteria (Fig 1). In addition, we identified *Rhilphs* in a α-proteobacterium, as well as in plactomycetes. *Rhilphs* were also detected in a unicellular green alga *Micromonas* and in two non-plant unicellular eukaryotes, an amoeba *Naegleria gruberi* and a choanoflagellate *Salpingoeca sp*. (Fig 1). The originally detected presence of *Rhilphs* in land plants (but not in green algae) and in Rhizobiales, bacterial symbionts or parasites of land plants, prompted us to hypothesise that *Rhilphs* may have been acquired by early land plants from bacteria. The presence of *Rhilphs* in *Micromonas* species, which belong to unicellular green algae (Chlorophyta) seems to contradict this hypothesis. However, to the best of our knowledge, the evolutionary origin of symbiosis between green plants and Rhizobiales is not well defined. The presence of several representatives of Rhizobiales in lichens together with

Chlorophyta has been reported (Bates et al., 2011), which suggests that Chlorophyta may be capable of a symbiotic relationship with Rhizobiales. This may indicate a more ancient origin of plant *Rhilphs*. As to the presence of *Rhilphs* in *Naegleria* and *Salpingoeca*, both of which are phagotrophic, it might be due to horizontal gene transfer from phagocytosed algae. Such transfer from algae to phagotrophic unicellular eukaryotes has recently been documented (Sun et al., 2010). Many plant species have 2 closely related Rhilph isoforms (Fig 1, red brackets), which appear to have originated by recent independent duplications in different lineages. This would be consistent with an on-going functional diversification.

Expression of A. thaliana Rhilph isoforms

A.thaliana has 2 Rhilph isoforms: Rhilph-1 (The *Arabidopsis* Information Resource (TAIR) ID: At3g09960) and Rhilph-2 (TAIR ID: At3g09970); they are encoded by the genes arranged in tandem (Andreeva and Kutuzov, 2004). Phylogenetic analysis indicates that this duplication occurred before the divergence of *A.thaliana* and *A.lyrata* (Fig 1). Overall, microarray data compiled in Genevestigator (Grennan, 2006) indicate that *Rhilph-2* is expressed at higher levels as compared to *Rhilph-1*. The highest levels of *Rhilph-1* transcript are detected in roots (especially in the root tips and elongation zone), xylem, pollen and senescent leaves. The highest expression *Rhilph-2* is observed in pollen, seeds, as well as root tips and elongation zone. According to the PRIDE proteomics database (Vizcaino et al., 2009), expression of Rhilph-1 has been documented in roots as well as in extending leaves under drought stress, whereas Rhilph-2 has been detected in roots, cotyledons, leafs, flowers, pollen and seeds, consistent with a higher level of its expression compared to Rhilph-1. We also examined protein-protein interaction databases (Arabidopsis Interactions Viewer, University of Toronto; Plant Interactome Database, Harvard University; AtPIN, University of São Paulo). These databases contain no information on *Rhilph* interaction partners.

A. thaliana genes co-expressed with Rhilphs

Expression correlation has been reported to be a good predictor of the functional associations in Arabidopsis (Wei et al., 2006). To identify genes co-expressed with *A.thaliana Rhilphs*, we used CressExpress (Srinivasasainagendra et al., 2008) and Arabidopsis Co-expression Data Mining Tool (ACT) (Jen et al., 2006), tools for co-expression analysis of publicly available microarray expression data. CressExpress analysis revealed 285 genes that showed positive correlation with Rhilph-1 (Pearson correlation coefficient (R) >0.65). Yet, only 9 genes correlated with Rhilph-2 (Supplementary Table 1). ACT returned a smaller gene set for Rhilph-1 and a larger set for Rhilph-2 (Supplementary Table 1). Possible reasons for these differences may include different sets of expression profiling experiments used by CressExpress and ACT, outlier chip filtering in CressExpress and different microarray processing algorithms (Robust Multi-Array Average (RMA) and Microarray Suite 5.0 (MAS5), respectively). We used the ACT Gene Ontology (GO) term count feature for functional classification of 50 most correlated genes for each Rhilph. Most represented GO terms for both isoforms were related to carbohydrates / cell wall / extracellular region (Supplementary Table 1). This is in line with a study that found that *Rhilph-1* is among 20 genes most highly up-regulated in the mature stem in *A. thaliana* (Ko and Han, 2004), which suggests its involvement in secondary growth and/or xylem formation. Other prominent groups of GO terms were related to responses to pathogens for *Rhilph-1*, and cytoskeleton and membrane trafficking, as well as transmembrane transport, for *Rhilph-2* (Supplementary Table 1). Since CressExpress does not include a feature for GO analysis, preliminary functional enrichment analysis was performed with the Gene Functional Classification tool within the DAVID package (Huang et al., 2009). The analysis revealed several clusters, the most prominent of which included several receptor protein kinases and disease resistance proteins, consistent with the ACT results. Other clusters included lipases/esterases/carboxypeptidase, peroxidases and various transcription factors. To refine these data, we analysed the CressExpress dataset for Rhilph-1 manually for the presence of enriched keywords in gene annotations (Fig 2). This analysis showed that expression of a number of genes implicated in defence responses correlates with *Rhilph-1*. Of the 10 members of the *A. thaliana* respiratory burst oxidase homologues (AtRBOH), expression of *AtRBOHF* (*At1g64060)*, *AtRBOHG* (*At4g25090*) and *AtRBOHI* (*At4g11230*) correlates with *Rhilph-1*, and 4 other members show R values slightly below the cut-off (Fig 2A, Supplementary Table 1). Respiratory burst oxidases generate reactive oxygen species (ROS) in response to pathogens, and are also involved in development and hormone biosynthesis (Sagi and Fluhr, 2006). AtRBOHF is required for ROS production during defence response (Torres et al., 2002). AtRBOHD is phosphorylated in response to elicitors (Benschop et al., 2007) and thus would be a direct substrate of an unidentified phosphatase. Expression of two classes of protein kinases correlates with *Rhilph-1* (Fig 2A). LRR transmembrane protein kinases represent a large receptor family in plants, and many of them are involved in responses to pathogens (Torii, 2004). None of the detected LRR kinases is functionally characterised. Lectin receptor kinases (including S-locus kinases) are involved in interactions between the plasma membrane and the cell wall, and some of them function to strengthen these interactions as a defence mechanism against invading pathogens (Bouwmeester et al., 2011). They may also play a role in symbiosis with Rhizobiales (Navarro et al., 2006). Representatives of several families of transcription factors are co-expressed with *Rhilph-1*, the most overrepresented one being the WRKY family (Fig 2B), members of which play a central role in plant immune responses (Agarwal et al., 2011). We also noticed overrepresentation of some proteins that have particular structural elements, namely DC1 (divergent C1) domain, VQ motif, NBS/LRR domains (nucleotide binding site / leucine-rich repeats) and HMA (heavy metal-associated) domain (Fig 2C). Precise functions of the plant-specific DC1 domain proteins are unknown, but at least in one instance their induction by a fungal elicitor has been reported, suggesting a role in defence responses (Shinya et al., 2007). Likewise, VQ motif-containing proteins are only found in plants and their function is generally unknown. A VQ motif containing protein MKS1 (MAP kinase 4 substrate 1) is a substrate of MPK4 (mitogen activated protein kinase 4), which regulates defence responses, and may couple defence-related MAP kinase signalling to WRKY transcription factors (Andreasson et al., 2005). One of the VQ motif proteins that correlate with Rhilph-1, At4g20000, has been associated with early defence response under high light conditions (González-Pérez et al., 2011). The NBS-LRR proteins are well known to play a central role in plant innate immunity by detecting pathogen-associated signals (DeYoung and Innes, 2006). Correlation with several other protein groups points at possible functions related to

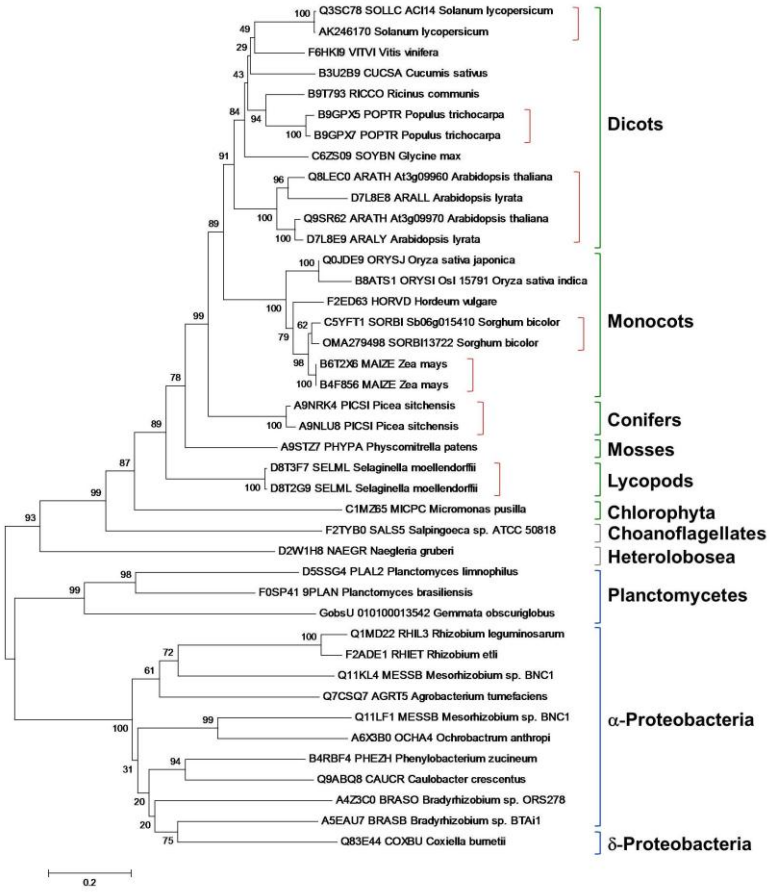

Fig 1. Phylogenetic analysis of bacterial and eukaryotic *Rhilphs*. The neighbor-joining (Saitou and Nei, 1987) tree was constructed in MEGA 5.0 (Tamura et al., 2011). Scale bar, the number of amino acid substitutions per site. Bootstrap values (out of 1,000 replicates) are shown. Red brackets, duplicated Rhilph isoforms in plants.

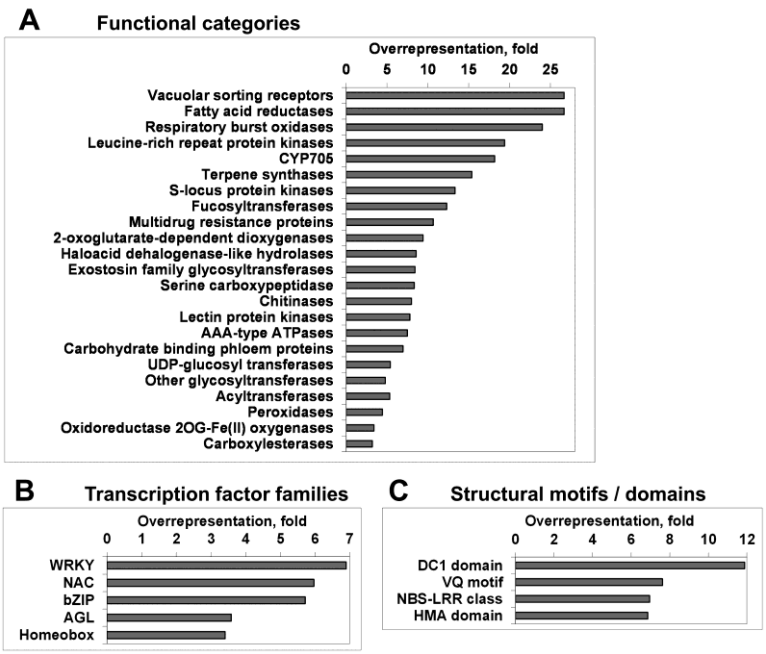

Fig 2. Overrepresented groups among genes co-expressed with *A.thaliana Rhilph-1*. (**A**), General functional groups; (**B**), Transcription factors; (**C**), Proteins containing specific structural motifs or domains.

membrane trafficking and cell wall organisation (Fig 2A, Supplementary Table 1). In addition to lectin receptor kinases, these include vacuolar sorting receptors, fatty acid reductases involved in suberin deposition (Domergue et al., 2010), and various groups of glycosyltransferases (Fig 2A). Unlike *Rhilph-1*, there was no clear indication for functional associations of Rhilph-2 in the CressExpress analysis. No overlap was found among the genes co-expressed with the two *Rhilphs*, however both sets included genes for functionally related proteins, namely glutathione S-transferases, enzymes involved in steroid metabolism, zinc finger components of ubiquitin ligases, and ATPases that function in transmembrane transport (Supplementary Table 1). Thus, despite different lists of co-expressed genes produced by CressExpress and ACT, both algorithms indicate correlation of Rhilph expression with genes involved in defence responses or carbohydrate metabolism / membrane trafficking / cell wall modification. In addition, we examined genes co-expressed with barley *Rhilph* homologue using GeneCAT tool (Mutwil et al., 2008). The only gene that showed reciprocal expression correlation with Rhilph encoded an aspartate transaminase. In addition to its role in amino acid metabolism, aspartate transaminase has recently been implicated in defence responses *in A. thaliana* (Brauc et al., 2011). Since most barley genes are not yet properly annotated, a detailed GO analysis of retrieved genes was not attempted. However, the presence of genes encoding homologues of Sec14 and ADP ribosylation factors (ARF) as well as transmembrane transporters is consistent with a possible involvement of barley Rhilph in membrane trafficking and events at the cell surface.

Physiological stimuli affecting Rhilph expression

To obtain further insight into possible biological roles of *Rhilphs*, we examined the physiological stimuli that affect their expression. Most studies where *Rhilph-1* up-regulation was detected used either *A. thaliana* infection with *Pseudomonas syringae*, or treatment with salicylic acid or elicitors such as a flagellin fragment flg22 or syringolin. This is in line with a report that tomato *Rhilph* is induced downstream of Cf-4/Cf-9 signalling among other genes presumably involved in hypersensitive response (Wang et al., 2008), and with the presence of a "W-box" motif (TTGAC) in *Rhilph-1* promoter, which is recognised by salicylic acid (SA)-induced WRKY transcription factors (Agarwal et al., 2011). We analysed manually the data from two transcript profiling experiments (Fig 3). In the experiment E-GEOD-17464, *Rhilph-1* was induced by flg22, but not by its inactive analogue from *Agrobacterium tumefaciens* (Fig 3A). Expression of *Rhilph-2* was not affected under these conditions. Induction of *Rhilph-1* by flg22 was suppressed in the leaves of plants overexpressing AFB1, an F-box protein and one of the auxin receptors (E-GEOD-17479). This would be consistent with the proposed inverse interdependence between auxin signalling and immune responses (Navarro et al., 2006). In the experiment E-GEOD-18978 (Wang et al., 2008), Rhilph-1 was strongly induced upon *P. syringae* infection (Fig 2B). This up-regulation was unaffected in the ein2 (*ethylene-insensitive 2*) mutant defective in ethylene signalling (Alonso et al., 1999), and was moderately impaired in the *coi1* (*coronatine-insensitive 1*), a mutant insensitive to jasmonate, and in *pad2* (*phytoalexin deficient 2*), which is deficient in glutathione biosynthesis and has enhanced susceptibility to pathogens (Parisy et al., 2007). The strongest suppression of Rhilph-1 induction was observed in the *npr1* (non-expresser of pathogenesis-related genes 1), pad4, and

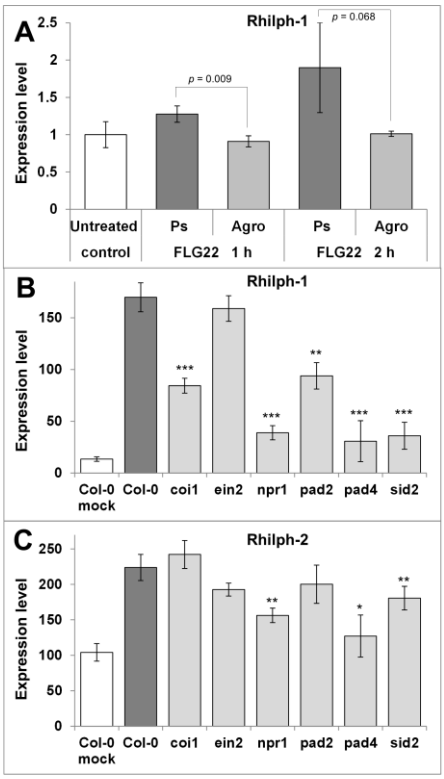

Fig 3. Effects of biotic stresses on expression of *Rhilph-1* (TAIR ID: At3g09960) and *Rhilph-2* (TAIR ID: At3g09970) in *A. thaliana*. (A), *Rhilph-1* expression in leaf disks from 5 weeks old short day grown plants (*Landsberg erecta*) treated or not with flg22 from *Pseudomonas syringae* or inactive flg22 from *Agrobacterium tumefaciens* for 1 or 2 h. Data are from the experiment E-GEOD-17464. *p* values from a two-tailed *t*-test are shown. (B, C), Expression of *Rhilph-1* (B) and *Rhilph-2* (C) in leaves of wild type (Col-0) or mutant *A. thaliana*, inoculated with *P. syringae* ES4326. Data are from the experiment E-GEOD-18978 (Wang *et al*., 2008). Error bars show standard deviation (n=3). *, $p<0.05$; **, $p<0.01$; ***, $p<0.001$.

sid2 (*salicylic acid induction deficient 2*) mutants. PAD4 is thought to act upstream of SA production (Vlot et al., 2009). SID2 (ICS1) is responsible for production of the bulk of SA in response to pathogens, and NPR1 is a master regulator of the signalling downstream of SA (Vlot et al., 2009). Thus, *P. syringae*-induced *Rhilph-1* expression appears to be predominantly regulated through the SA pathway, regulated in part by the jasmonate pathway and the redox status, and unaffected by ethylene signalling. Notably, a redox-dependent NPR1 effector transcription factor *TGA1* (Després et al., 2003) shows a good expression correlation with *Rhilph-1* (Supplementary Table 1).Rhilph-2 showed a weaker relative response to *P. syringae* infection, which was unaffected in *coi1*, insignificantly reduced in the *ein2*, *pad2* and *sid2* mutants, and considerably suppressed in the *npr1* and *pad4* mutants (Fig 2C). This suggests that *Rhilph-2* induction depends in part on SA signalling and does not require the jasmonate pathway. In line with the above observations, *Rhilph-1* promoter contains 6 copies of a binding motif (W-box) for WRKY transcription factors, whereas no such motifs are found in the Rhilph-2 promoter (data from AGRIS; Yilmaz et al., 2011).

A. thaliana mutants with altered expression of Rhilphs

We also examined whether characterised *A. thaliana* mutations affect *Rhilph* expression (Supplementary Table 2). The highest increase in *Rhilph-1* expression was found in a double knockout mutant lacking both MKK1 and MKK2 (mitogen activated protein kinase kinases 1 and 2), but not in the mutants lacking only one isoform (E-GEOD-18978; Qiu et al., 2008). This indicates that MKK1 and MKK2 redundantly suppress *Rhilph-1* expression in non-stimulated plants. MKK1 and MKK2 are MAP kinase kinases that play partially redundant roles in jasmonate-and salicylate-dependent defence responses (Qiu et al., 2008). Rhilph-1 expression was also elevated in MPK4 mutants (E-MEXP-174; Brodersen et al., 2006). *MPK4* is a MAP kinase that functions downstream of MKK1/2 in a cascade that regulates innate immune responses (Gao et al., 2008; Pitzschke et al., 2009). Since mkk1/2and mpk4 mutants are in different backgrounds (Columbia and Landsberg, respectively), regulation of *Rhilph-1* expression by the MKK1/2-MPK4 pathway appears not to be ecotype-dependent. *Rhilph-1* was up-regulated in the *rpp7* mutant, which is resistant to Peronospora parasitica infection (E-NASC-16; Can et al., 2003). In line with a possible functional link between *Rhilph-1* and respiratory burst oxidases and ROS production, *Rhilph-1* was up-regulated in an *A. thaliana* mutant lacking AtRBOHC (E-GEOD-6165) and in the *upb1* (*upbeat1*) mutant, which has altered ROS levels (E-GEOD-21876). Down-regulation of *Rhilph-1* was observed in plants transfected with a microRNA targeting an allele of At5g41750 (E-ATMX-26; Bomblies et al., 2007), which encodes a disease resistance protein of the TIR (Toll/interleukin-1 receptor-like)-NBS-LRR class. The above data support the involvement of *Rhilph-1* in defence responses. In addition, *Rhilph-1* expression was affected in the mutants with no obvious links to defence responses, including two cyclic nucleotide gated cation channels (E-GEOD-20222) and the *brx* (*brevis radix*) mutant, which has altered cell proliferation and elongation in the growth zone of the root tip (Mouchel et al., 2004). The latter effect is consistent with elevated *Rhilph-1* expression in the root elongation zone.

Methods

Phylogenetic analysis

Rhilph sequences were identified by BLAST (Altschul et al., 1997) at NCBI (http://blast.ncbi.nlm.nih.gov/) and retrieved from UniProt (http://www.uniprot.org/). Sequences were aligned with ClustalW (with default parameters, except that Blosum matrix was used) as implemented in MEGA 5.0 (Tamura et al., 2011). Phylogenetic trees were constructed using the neighbour-joining method (Saitou and Nei, 1987) and bootstrap analysis (Felsenstein, 1985) was performed in MEGA 5.0. Positions corresponding to gaps in >20% sequences were disregarded.

Identification of co-expressed genes

Co-expression of *A. thaliana* genes was examined by using data mining tools CressExpress v. 3.0 (http://cressexpress.org/; Srinivasasainagendra et al., 2008) and Arabidopsis Co-expression Data Mining Tool ACT (http://www.arabidopsis.leeds.ac.uk/act ; Jen et al., 2006). All available experiments were used, which include the Nottingham Arabidopsis Stock Center (NASC) expression data from Affymetrix ATH1 arrays (22,810 probes). RMA processing was used in CressExpress; the cut-off value for Kolmogorov-Smirnov quality-control statistic was set at 0.15 to filter for potential outlier chips (Srinivasasainagendra et al., 2008). ACT was used with default parameters ("Co-expression analysis over available array experiments" option). Pearson correlation coefficient (R) cut-off of $R \geq 0.65$ was used, except for Rhilph-2 in the ACT analysis, where a more stringent cut-off of $R \geq 0.70$ was used due to a larger number of strongly co-expressed genes.

Functional enrichment analysis

Preliminary functional enrichment analysis of the CressExpress results was performed with the Gene Functional Classification tool within the DAVID package (Huang et al., 2009) at NIH (http://david.abcc.ncifcrf.gov/ home.jsp) using high stringency setting in the cluster analysis. Keyword search was then performed manually in Excel spreadsheets. Annotations of the genes represented on total *A.thaliana* ATH1 microarrays and those of the transcripts co-expressed with *Rhilph-1* ($R \geq 0.65$) were searched with appropriate keywords. Overrepresentation was calculated as a ratio of the percentage of matching annotations in the co-expressed subset to that in total ATH1 microarrays. For the ACT results, automated GO annotation search in the lists of co-expressed genes was used ("Gene Ontology term count of Co-expression analysis" option).

Assessment of the effects of physiological stimuli on Rhilph expression

The effects of stimuli on *Rhilph* expression were assessed by using the 'Response Viewer' tool in Genevestigator. Data from individual transcription profiling experiments, which showed changes in *Rhilph* expression, were downloaded from NCBI Gene Expression Omnibus (http://www.ncbi.nlm. nih.gov/geo/) or ArrayExpress (http://www.ebi.ac.uk/ microarray-as/ae), and imported into Excel. In the experiments where expression data were represented as logarithmic values, these were converted into arithmetic values. Only experiments that have at least 3 biological replicates were analysed. Expression levels of *Rhilph-1* (probe set 258942_at) and *Rhilph-2* (258883_at) was determined as means±S.D., and was either plotted in the original signal intensity units, or was normalised to untreated control (see figure for details). Statistical significance was assessed by using a two-tailed Student's test. Differences were considered significant at $p<0.05$.

Assessment of the effects of mutations on Rhilph expression

The effects of mutations on *Rhilph* expression were assessed by using the 'Mutant Surveyor' tool in the Plant Biology section of Genevestigator (https://www.genevestigator.com/ gv/plant.jsp; Grennan, 2006). Information on the anatomical parts used and the number of biological replicates was obtained from Genevestigator. Annotation of the genes affected in particular mutants was compiled from TAIR (https://www.arabidopsis.org/) and from the summaries of respective profiling experiments, available at the European Bioinformatics Institute ArrayExpress database (http://www. ebi.ac.uk/arrayexpress/).

Conclusions

In this report, we used several data mining approaches to predict the biological processes that may involve *Rhilphs*, enigmatic protein phosphatases probably acquired by plants from α-proteobacteria. The data suggest that *A. thaliana Rhilphs*, especially *Rhilph-1*, are induced by pathogen-related signals and their expression is altered in several defence signalling mutants. Moreover, functional enrichment analysis indicates that genes associated with responses to pathogens and innate immunity are overrepresented among those that show expression correlation with *Rhilph-1*. *Rhilph-2* shows similar but less pronounced transcriptional responses to pathogen-related signals. Since it is more constitutively expressed, transcription profiling alone may not be sufficient to assess its possible involvement in defence responses. Another category of genes co-expressed with both *Rhilphs* are associated with carbohydrate metabolism and cell wall, and some (like lectin receptor kinases) may be involved in both defence responses and cell wall-related processes. This suggests that some *Rhilph* functions may be related to cell wall remodelling during a defence response, which would define *Rhilphs* (or at least *Rhilph-1*) as novel potential defence-related proteins, and provide a framework for experimental assessment of their functions.

Acknowledgements

The authors are indebted to all researchers who have deposited their primary data into open access repositories, and in particular to developers of CressExpress and ACT used in this work.

Supplementary data

Supplementary Table 1. Co-expression of *A. thaliana* genes with *Rhilph-1* and *Rhilph-2*.
Supplementary Table 2. *A. thaliana* mutants with altered *Rhilph-1* expression.

References

Agarwal P, Reddy MP, Chikara J (2011) WRKY: its structure, evolutionary relationship, DNA-binding selectivity, role in stress tolerance and development of plants. Mol Biol Rep 38:3883–3896

Alonso JM, Hirayama T, Roman G, Nourizadeh S, Ecker JR (1999) EIN2, a bifunctional transducer of ethylene and stress responses in *Arabidopsis*. Science 284:2148–2152

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

Andreasson E, Jenkins T, Brodersen P, Thorgrimsen S, Petersen NH, Zhu S, Qiu JL, Micheelsen P, Rocher A, Petersen M, Newman MA, Bjorn Nielsen H, Hirt H, Somssich I, Mattsson O, Mundy J (2005) The MAP kinase substrate MKS1 is a regulator of plant defense responses. EMBO J 24:2579–2589

Andreeva AV, Kutuzov MA (1999) *Physcomitrella* patens gene/cDNA fragments related to genes encoding protein Ser/Thr phosphatases. J Plant Physiol 155:153–158

Andreeva AV, Kutuzov MA (2004) Widespread presence of "bacterial-like" PPP phosphatases in eukaryotes. BMC Evol Biol 4:47

Bates ST, Cropsey GW, Caporaso JG, Knight R, Fierer N (2011) Bacterial communities associated with the lichen symbiosis. Appl Environ Microb 77:1309–1314

Benschop JJ, Mohammed S, O'Flaherty M, Heck AJ, Slijper M, Menke FL (2007) Quantitative phosphoproteomics of early elicitor signaling in *Arabidopsis*. Mol Cell Proteomics 6:1198–1214

Bomblies K, Lempe J, Epple P, Warthmann N, Lanz C, Dangl JL, Weigel D (2007) Autoimmune response as a mechanism for a Dobzhansky-Muller-type incompatibility syndrome in plants. PLoS Biol 5:e236

Bouwmeester K, de Sain M, Weide R, Gouget A, Klamer S, Canut H, Govers F (2011) The lectin receptor kinase LecRK-I.9 is a novel phytophthora resistance component and a potential host target for a RXLR effector. PLoS Pathog 7:e1001327

Brauc S, De Vooght E, Claeys M, Höfte M, Angenon G (2011) Influence of over-expression of cytosolic aspartate aminotransferase on amino acid metabolism and defence responses against *Botrytis cinerea* infection in *Arabidopsis thaliana*. J Plant Physiol 168:1813–1819

Brodersen P, Petersen M, Bjorn Nielsen H, Zhu S, Newman MA, Shokat KM, Rietz S, Parker J, Mundy J (2006) *Arabidopsis* MAP kinase 4 regulates salicylic acid- and jasmonic acid/ethylene-dependent responses via EDS1 and PAD4. Plant J 47:532–546

Can C, Özaslan M, Holub E (2003) Characterisation of RPP7 mutant lines of the col-5 ecotype of *Arabidopsis thaliana*. J Cell Mol Biol 2:25–30

Després C, Chubak C, Rochon A, Clark R, Bethune T, Desveaux D, Fobert PR (2003) The *Arabidopsis* NPR1 disease resistance protein is a novel cofactor that confers redox regulation of DNA binding activity to the basic domain/leucine zipper transcription factor TGA1. Plant Cell 15:2181–2191

DeYoung BJ, Innes RW (2006) Plant NBS-LRR proteins in pathogen sensing and host defense. Nat Immunol 7:1243–1249

Domergue F, Vishwanath SJ, Joubés J, Ono J, Lee JA, Bourdon M, Alhattab R, Lowe C, Pascal S, Lessire R, Rowland O (2010) Three *Arabidopsis* fatty acyl-coenzyme A reductases, FAR1, FAR4, and FAR5, generate primary fatty alcohols associated with suberin deposition. Plant Physiol 153:1539–1554 1

Felsenstein J (1985) Confidence limits on phylogenies: An approach using the bootstrap. Evolution 39:783–791

Gao M, Liu J, Bi D, Zhang Z, Cheng F, Chen S, Zhang Y (2008) MEKK1, MKK1/MKK2 and MPK4 function together in a mitogen-activated protein kinase cascade to regulate innate immunity in plants. Cell Res 18:1190–1198

González-Pérez S, Gutiérrez J, García-García F, Osuna D, Dopazo J, Lorenzo O, Revuelta JL, Arellano JB (2011) Early transcriptional defense responses in *Arabidopsis* cell suspension culture under high-light conditions. Plant Physiol 156:1439–1456

Grennan AK (2006) Genevestigator. Facilitating web-based gene-expression analysis. Plant Physiol 141:1164–1166

Huang DW, Sherman BT, Lempicki RA (2009) Systematic and integrative analysis of large gene lists using DAVID Bioinformatics Resources. Nat Protoc 4:44-57

Jen CH, Manfield IW, Michalopoulos I, Pinney JW, Willats WG, Gilmartin PM, Westhead DR (2006) The *Arabidopsis* co-expression tool (ACT): a WWW-based tool and database for microarray-based gene expression analysis. Plant J 46:336–348

Ko JH, Han KH (2004) *Arabidopsis* whole-transcriptome profiling defines the features of coordinated regulations that occur during secondary growth. Plant Mol Biol 55:433–453

Kutuzov MA, Andreeva AV (2012) Prediction of biological functions of *Shewanella*-like protein phosphatases (Shelphs) across different domains of life. Funct Integr Genomic 12:11–23

Mouchel CF, Briggs GC, Hardtke CS (2004) Natural genetic variation in *Arabidopsis* identifies BREVIS RADIX, a novel regulator of cell proliferation and elongation in the root. Gene Dev 18:700–714

Mutwil M, Obro J, Willats WG, Persson S (2008) GeneCAT – novel webtools that combine BLAST and co-expression analyses. Nucleic Acids Res 36:W320–W326

Navarro L, Dunoyer P, Jay F, Arnold B, Dharmasiri N, Estelle M, Voinnet O, Jones JD (2006) A plant miRNA contributes to antibacterial resistance by repressing auxin signaling. Science 312:436–439

Parisy V, Poinssot B, Owsianowski L, Buchala A, Glazebrook J, Mauch F (2007) Identification of PAD2 as a α-glutamylcysteine synthetase highlights the importance of glutathione in disease resistance of *Arabidopsis*. Plant J 49:159–172

Pitzschke A, Djamei A, Bitton F, Hirt H (2009) A major role of the MEKK1-MKK1/2-MPK4 pathway in ROS signalling. Mol Plant 2:120–137

Qiu JL, Zhou L, Yun BW, Nielsen HB, Fiil BK, Petersen K, Mackinlay J, Loake GJ, Mundy J, Morris PC (2008) *Arabidopsis* mitogen-activated protein kinase kinases MKK1 and MKK2 have overlapping functions in defense signaling mediated by MEKK1, MPK4, and MKS1. Plant Physiol 148:212–222

Sagi M, Fluhr R (2006) Production of reactive oxygen species by plant NADPH oxidases. Plant Physiol 141:336–340

Saitou N, Nei M (1987) The neighbor-joining method: a new method for reconstructing phylogenetic trees. Mol Biol Evol 4:406–425.

Shinya T, Galis I, Narisawa T, Sasaki M, Fukuda H, Matsuoka H, Saito M, Matsuoka K (2007) Comprehensive analysis of glucan elicitor-regulated gene expression in tobacco BY-2 cells reveals a novel MYB transcription factor involved in the regulation of phenylpropanoid metabolism. Plant Cell Physiol 48:1404–1413

Srinivasasainagendra V, Page GP, Mehta T, Coulibaly I, Loraine AE (2008) CressExpress: a tool for large-scale mining of expression data from *Arabidopsis*. Plant Physiol 147:1004–1016

Sun G, Yang Z, Ishwar A, Huang J (2010) Algal genes in the closest relatives of animals. Mol Biol Evol 27:2879–2889

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

Torii KU (2004) Leucine-rich repeat receptor kinases in plants: structure, function, and signal transduction pathways. Int Rev Cytol 234:1–46

Torres MA, Dangl JL, Jones JD (2002) *Arabidopsis* gp91phox homologues AtrbohD and AtrbohF are required for accumulation of reactive oxygen intermediates in the plant defense response. P Natl Acad Sci USA 99:517–522

Vizcaino JA, Cote R, Reisinger F, Foster JM, Mueller M, Rameseder J, Hermjakob H, Martens L (2009) A guide to the Proteomics Identifications Database proteomics data repository. Proteomics 9:4276–4283

Vlot AC, Dempsey DA, Klessig DF (2009) Salicylic Acid, a multifaceted hormone to combat disease. Annu Rev Phytopathol 47:177–206

Wang L, Mitra RM, Hasselmann KD, Sato M, Lenarz-Wyatt L, Cohen JD, Katagiri F, Glazebrook J (2008) The genetic network controlling the *Arabidopsis* transcriptional response to *Pseudomonas syringae* pv. *maculicola*: roles of major regulators and the phytotoxin coronatine. Mol Plant Microbe In 21:1408–1420

Wei H, Persson S, Mehta T, Srinivasasainagendra V, Chen L, Page GP, Somerville C, Loraine A (2006) Transcriptional coordination of the metabolic network in *Arabidopsis*. Plant Physiol 142:762–774

Yilmaz A, Mejia-Guerra MK, Kurz K, Liang X, Welch L, Grotewold E (2011) AGRIS: the *Arabidopsis* Gene Regulatory Information Server, an update. Nucleic Acids Res 39:D1118–D1122