HAL
open science

Genome-wide annotation of remorins, a plant-specific protein family: evolutionary and functional perspectives

Sylvain Raffaele, Sébastien Mongrand, Pascal Gamas, Andréas Niebel, Thomas Ott

To cite this version:

Sylvain Raffaele, Sébastien Mongrand, Pascal Gamas, Andréas Niebel, Thomas Ott. Genome-wide annotation of remorins, a plant-specific protein family: evolutionary and functional perspectives. Plant Physiology, 2007, 145 (3), pp.593-600. 10.1104/pp.107.108639 . hal-01601863

HAL Id: hal-01601863

https://hal.science/hal-01601863v1

Submitted on 31 May 2020

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Genome-Wide Annotation of Remorins, a Plant-Specific Protein Family: Evolutionary and Functional Perspectives[1[W]]

Sylvain Raffaele, Sébastien Mongrand, Pascal Gamas, Andreas Niebel, and Thomas Ott*

Laboratoire de Biogenèse Membranaire, Centre National de la Recherche Scientifique/Université Victor Segalen Bordeaux 2, 33076 Bordeaux cedex, France (S.R., S.M.); and Laboratoire des Interactions Plantes Micro-organismes, Institut National de la Recherche Agronomique-Centre National de la Recherche Scientifique, 31326 Castanet Tolosan cedex, France (P.G., A.N., T.O.)

Remorins were discovered in a screen for plasma membrane (PM) proteins differentially phosphorylated in the presence of oligogalacturonides (Farmer et al., 1989). The first remorin was initially designated as pp34, as it corresponded to a phosphorylated protein with a molecular mass of 34 kD in purified PM preparations from potato (*Solanum tuberosum*; Jacinto et al., 1993). After a section in J.L. Borge's "Book of Imaginary Beings," in which the remora fish is described to attach itself to larger fish and ships, this protein was renamed remorin (Reymond et al., 1996) due to its hydrophilic profile and its ability to attach to PMs (E. Farmer, personal communication).

Several proteomic approaches on PM preparations from tobacco (*Nicotiana tabacum*) leaves and Arabidopsis (*Arabidopsis thaliana*) seedlings now suggest localization of remorins in association with PMs (Watson et al., 2003; Marmagne et al., 2004; Mongrand et al., 2004; Sazuka et al., 2004; Nelson et al., 2006; Valot et al., 2006). Interestingly, remorins have also been found in detergent-resistant membrane fractions, called lipid rafts (Mongrand et al., 2004; Shahollari et al., 2004; Bhat and Panstruga, 2005; Laloi et al., 2007; Lefebvre et al., 2007). Lipid rafts are characterized by an enrichment of sterols, sphingolipids, and steryl-glycosides, and are believed to be functional microdomains in PMs (Peskan et al., 2000; Mongrand et al., 2004; Borner et al., 2005; Laloi et al., 2007). Most recently, Lefebvre and coworkers demonstrated the presence of remorins in lipid raft fractions from root PMs of the model legume *Medicago truncatula* (Lefebvre et al., 2007). While little is known about the function of these microdomains in plants, key signaling processes during HIV-1 infection (Manes et al., 2000; Ono and Freed, 2001), spreading of malarial parasites and the Newcastle disease virus, as well as prion conversion occur at lipid rafts in mammalian systems (Samuel et al., 2001; Nishina et al., 2004; Laliberte et al., 2006).

Some remorins share physical properties with viral movement proteins due to their hydrophobic N-terminal region. The C-terminal domain is generally predicted to consist of coiled-coil domains, a feature that appears to be common to all remorin proteins. This region was found to be responsible for the ability of remorins to form oligomeric filamentous structures in vitro, suggesting that these proteins might be structural components of plant cytoskeletons and/or the membrane skeletons (Bariola et al., 2004). In addition, expression of a remorin from tomato (*Solanum lycopersicum*) was observed in embryonic, apical, and vascular tissue (Bariola et al., 2004). However, astonishingly, there is no published genetic evidence in any plant species that can help us to decipher putative functions of remorins. The few reported attempts to generate plant mutants in which one or more of the ubiquitously expressed remorin genes were knocked out either failed or did not lead to an obvious phenotype (Reymond et al., 1996; Bariola et al., 2004).

Searches in databases of fully sequenced plant species revealed the existence of eight, 16, and 19 remorin genes in *Populus trichocarpa*, Arabidopsis, and rice (*Oryza sativa*), respectively (Supplemental Table S1). Patterns of differentially regulated remorin genes and proteins have been reported in an increasing number of transcriptomic and proteomic studies. Since many of these are annotated as "unknown function," they need to be properly classified and given suitable nomenclature. Here, we report on a genome-wide survey of the remorin family throughout fully sequenced plant genomes. We combined a phylogenetic approach with various in silico sequence-scanning tools (alignments, and motif and domain predictions) and analysis of available genetic and genomic data to identify a specific remorin signature and to define several

[1] This work was supported by the French Agence Nationale de la Recherche (contract ANR–JC05–45555 "Plant rafts"; S.R. and S.M.) and the European Union (Marie Curie Intra-European Fellowship grant no. MEIF–CT–2005–024587; T.O.).

* Corresponding author; e-mail thomas.ott@toulouse.inra.fr.

The author responsible for distribution of materials integral to the findings presented in this article in accordance with the policy described in the Instructions for Authors (www.plantphysiol.org) is: Thomas Ott (thomas.ott@toulouse.inra.fr).

[W] The online version of this article contains Web-only data.

www.plantphysiol.org/cgi/doi/10.1104/pp.107.108639

Downloaded from www.plantphysiol.org on June 12, 2017 - Published by www.plantphysiol.org
Copyright © 2007 American Society of Plant Biologists. All rights reserved.

groups within the remorin family. We used this classification to propose a general nomenclature for remorins that should help to differentiate between the different members in future studies.

IDENTIFICATION OF A GENERIC REMORIN SIGNATURE

Keyword and BLAST searches among available data allowed us to retrieve approximately 1,000 remorin sequences corresponding to more than 130 remorin proteins from 55 different plant species (Supplemental Table S1). Importantly, no significant hits were identified among sequence data from other kingdoms (fungi, animals, archaeabacteria, and eubacteria). Unexpectedly, no remorin genes were found in algae (neither by BLAST against publicly available data nor in the *Ectocarpus siliculosus* full-genome sequence; M. Cock, personal communication), while remorins were identified in mosses and ferns.

Canonical remorins, such as StREM (Reymond et al., 1996; now StREM1.3), are characterized by the presence of an N-terminal and a C-terminal PFAM domain (Remorin_N: PF03766; Remorin_C: PF03763). Higher sequence variability in N-terminal regions of remorins suggests structural and functional divergence and indicates that a suitable classification and nomenclature are needed. In contrast, the Remorin_C domain, identified by InterProScan using default parameters, appeared to be well conserved among all remorin proteins. It was thus considered as the remorin signature and was used to validate potential family members.

A striking property of the Remorin_C domain is the high content of charged residues (e.g. 45% for StREM1.3), illustrated by well-conserved charged regions as shown in Figure 1. The presence of remarkably hydrophilic domains with GRAVY (grand average of hydropathy) indices between −1.4 and −0.5 (Supplemental Table S1) results in a high hydrophilicity, posing the question how these proteins are tightly associated with membranes. Neither transmembrane domains nor membrane anchor signatures were found in any of the studied sequences (Supplemental Table S2).

Some particularly conserved residues can be outlined: position 11 of Remorin_C is occupied by an aromatic residue (mainly Trp); positions 18, 22, and 25 by positively charged residues; and position 30 by an aliphatic residue for all but one sequence (Fig. 1). Position 33 harbors a Trp in all sequences, except in *Physcomitrella patens* and *Ceratopteris richardii*, and positions 75 and 80 are Alas in most of the proteins. Because of their very high conservation, these amino acids can be considered as identifying features of the Remorin_C domain. In addition, a 23-amino-acid-long coiled-coil structure between Remorin_C residues 40 to 63 is predicted with a very high probability (Marcoils probability >90%). Prediction of the different positions in heptad repeats (Fig. 1, top diagram) reveals aliphatic residues for sites a and d that form the core of the coiled-coil region. In comparison with coiled-coil domains from other proteins, in remorins these motifs consist of a higher number of charged residues (56% for StREM1.3) in positions b, c, d, and e. This characteristic could be responsible for the high stability of protein interactions mediated by this domain (Burkhard et al., 2001).

THE REMORIN FAMILY CONSISTS OF SIX DIFFERENT GROUPS

With the aim of considering only full-length proteins and to avoid a bias of the study toward only highly expressed family members, we focused mainly on genes/proteins from plants with annotated genomes. As recent proteomic studies report the identification

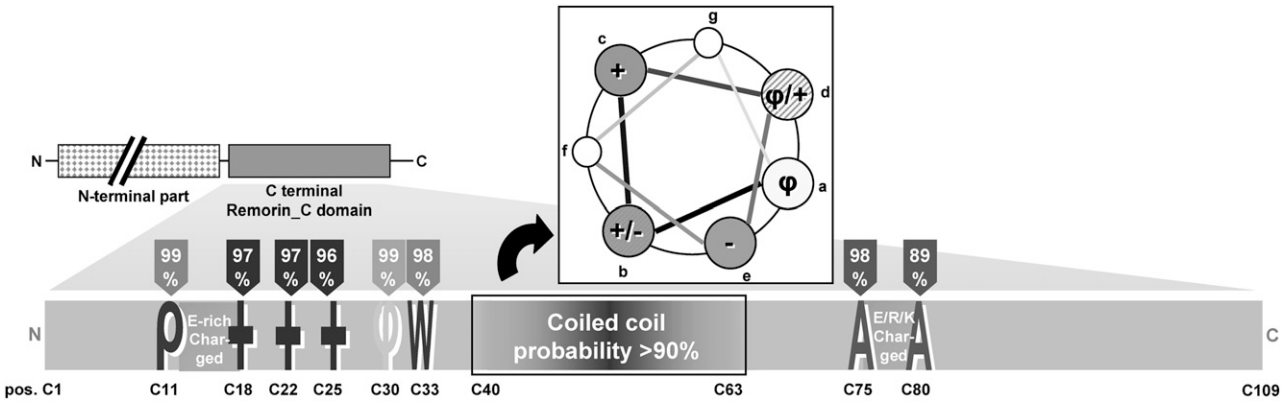

Figure 1. The remorin signature. Highly conserved residues within the Remorin_C domain of the 85 remorin proteins that were considered for phylogenetic analysis are shown. Prediction of the position along the heptad repeat of the coiled-coil structure is given in the top panel. pos., Position in Remorin_C domain alignment; ρ, aromatic residues; +, positively charged residues; ϕ, aliphatic residues; % refers to the number of sequences showing the indicated residue at the indicated position, among the list of 85 remorins given in Supplemental Table S1.

Downloaded from www.plantphysiol.org on June 12, 2017 - Published by www.plantphysiol.org
Copyright © 2007 American Society of Plant Biologists. All rights reserved.

of remorins in legumes (Watson et al., 2003; Wienkoop and Saalbach, 2003; Valot et al., 2006; Lefebvre et al., 2007), all 10 remorin sequences of the partially sequenced model legume *M. truncatula* have been included in the analysis as representative remorins for the Fabaceae family. Considering that plant species that diverged early in evolution are likely to contain ancestral remorin sequences, all available remorins from ferns, mosses, and gymnosperms were also incorporated. This led to a list of 85 remorins (Supplemental Table S1) on which phylogenetic analysis was performed, as described in detail in Supplemental Materials and Methods S1.

Two distinct phylogenetic trees were built, one using the full-length amino acid sequences (Fig. 2) and the other using only Remorin_C domain sequences (data not shown). Both trees were very similar, with the main exception that using full-length sequences allows clustering of remorins that lack an N-terminal domain in a single group. Considering that N-terminal regions may confer functional specificity to the proteins, we favored the use of full-length sequences in the following analyses.

For the final grouping leading to the proposed nomenclature, results from the phylogenetic analysis were combined with domain features and similarities between different remorin proteins. Due to significant differences, mainly in the N-terminal regions, the remorin family was subdivided into six separate groups. Such subdivision based on domain features

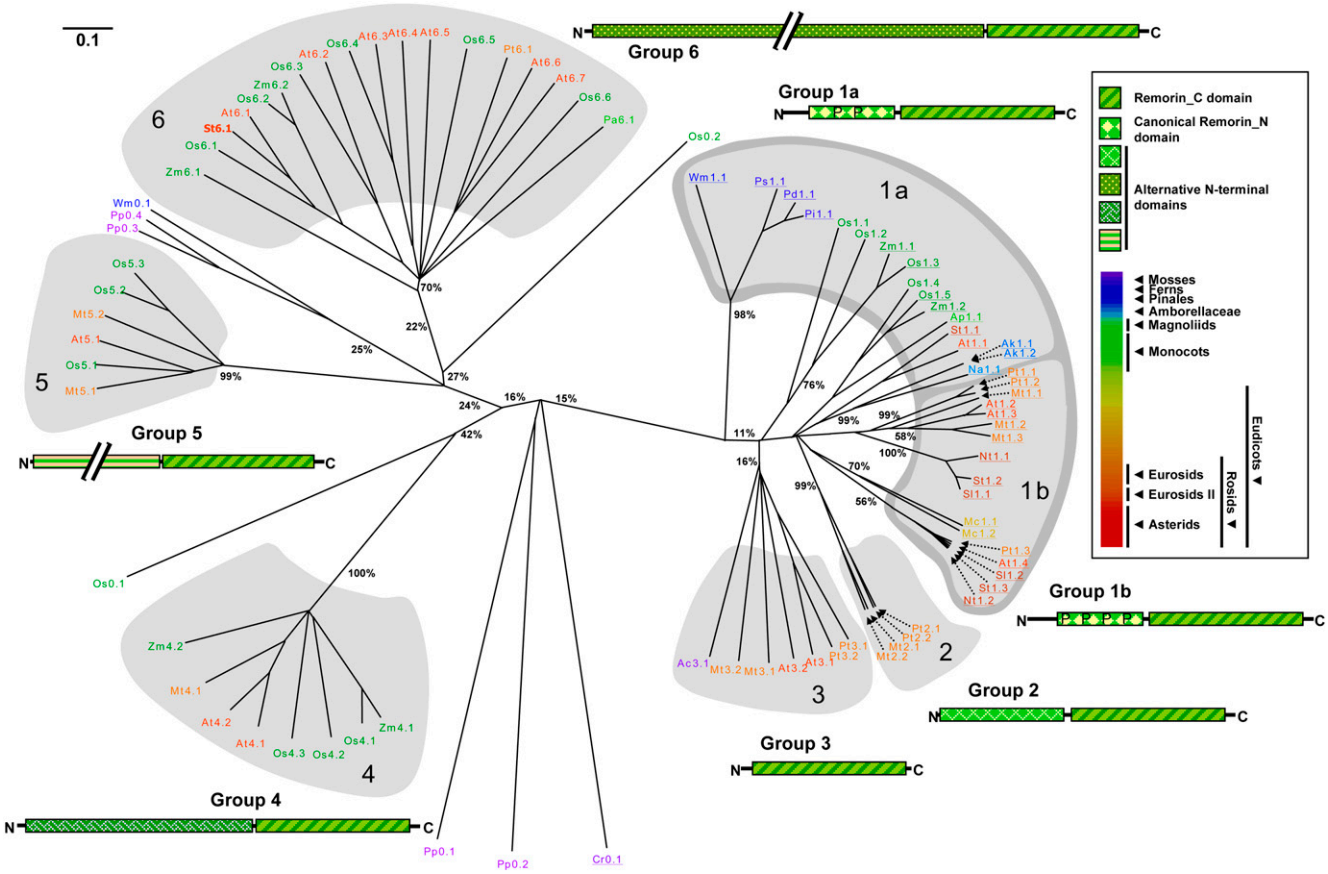

Figure 2. Phylogenetic and domain organization of the remorin family based on their protein sequences. The right box shows the color code used to indicate evolutionary positions within the plant kingdom (according to the Angiosperm Phylogeny Group, 2003) and domain elements used for classification. In the diagrams domain lengths are proportional to the average protein sequence length (except for groups 5 and 6, for which the representation of the N-terminal region is intercepted by "//," indicating a variable length of this module within these two groups). Underlined sequences contain Remorin_N domains identified by InterProScan. Bootstrap values (%) are given along the most important branch points. True branch lengths were inferred by maximum-likelihood analysis; the upper-left scale represents a probability of 0.1 amino acid changes per site. The tree was reconstructed using parsimony analysis. For clarity of the figure, a shortened nomenclature was used (e.g. At1.1 for AtRem1.1). The first number in the proposed name indicates the corresponding group (e.g. At1.3 is the third member of group 1). Abbreviations: Ac, *Adiantum capillus-veneris*; Ak, *Amborella trichopoda*; Ap, *Allium cepa*; At, Arabidopsis; Cr, *C. richardii*; Mc, *Mesembryanthemum crystallinum*; Mt, *M. truncatula*; Na, *Nuphar advena*; Nt, tobacco; Os, rice; Pa, *Persea americana*; Pd, *Pinus taeda*; Pi, *Pinus pinaster*; Pp, *P. patens*; Ps, *Picea sitchensis*; Pt, *P. trichocarpa*; Sl, tomato; St, *Solanum tuberosum*; Wm, *Welwitschia mirabilis*; Zm, *Zea mays*.

Downloaded from www.plantphysiol.org on June 12, 2017 - Published by www.plantphysiol.org
Copyright © 2007 American Society of Plant Biologists. All rights reserved.

enabled us to discriminate between groups 1, 2, and 3 on those parts of the tree where phylogenetic separation was less clear (Fig. 2; Table I).

Group 1 (Subgroups 1a and 1b): Canonical Remorins with a Pro-Rich N-Terminal Region

This group gathers all remorins that possess both Remorin_N and Remorin_C domains identified by InterProScan. Genomic data support coherence of this group as most genes consist of five exons, except for AtREM1.3 and OsREM1.2, in which the last two exons were fused by deletion of an intron. Generally, group 1 remorins are characterized by a high Pro content in the N-terminal region (only group 4 members have similar numbers of Pros in this region). Group 1 remorins were subdivided into group 1a (8.9% Pros) and group 1b (14.4% Pros) based on their N-terminal amino acid composition (Table I; Supplemental Fig. S2).

Group 1a remorins are about 200 amino acids long with a very homologous N-terminal region (about 82 amino acids). Four well-conserved motifs were identified by MEME/MAST (Multiple Em for Motif Elicitation/Motif Alignment and Search Tool; Bailey and Elkan, 1994) in these N-terminal regions (Supplemental Table S2). They bear two ELMs (Eukaryotic Linear motifs; Puntervoll et al., 2003): a casein kinase Ser/Thr phosphorylation site and a motif found in cargo adaptor proteins that bind clathrin heavy chains. These elements, although not functionally characterized in remorins, underline the potential importance of phosphorylation events and membrane association for these proteins. Phylogenetic clustering of this group is remarkable as it assembles branches of ancestral plants: One is comprised of sequences from ferns and gymnosperms, the second consists of sequences from early angiosperms, and two other branches contain exclusively monocot sequences. Only two dicot remorin sequences from Arabidopsis and potato belong to this group.

One of them, AtREM1.1, shows increased expression in senescing leaves, roots, and xylem, and after salt and osmotic stress and abscisic acid (ABA) treatment (Supplemental Table S3). The set of Arabidopsis

Table I. *Structural characteristics of remorins*

Structural features were identified for all remorins and were combined with phylogenetic clustering of the different remorin isoforms to propose a group classification. Numbers of exons were obtained from public databases for genes of all sequenced plant species or newly assembled for *M. truncatula* sequences (T. Ott, unpublished data). Details of MEME motifs can be found in Supplemental Table S2. *, Number of amino acids. **, Average *e* value for domain recognition: +, $>10^{-10}$; ++, $>10^{-15}$; +++, $>10^{-30}$; ++++, $<10^{-30}$ (3). ***, Two exceptions with four exons: OsREM1.2 for group 1a and AtREM1.3 for group 1b. NA, Not applicable; Rem_C, Remorin_C domain; Rem_N, Remorin_N domain.

Group	No. of Exons	N-Terminal Part Properties				Remorin_C		Plant Groups Covered
		Length*	Pro Content	Rem_N Domain?	Remarkable Features	Rem_C Length*	Homology to Rem_C**	
			% of N terminus					
1a	5***	79.6 ± 10.7	8.9 ± 3.2	Yes	MEME motifs 1, 2, 3, 4	107 ± 4.7	+++	► Gymnosperms (Pinales, Gnetales) ► Archaic angiosperms (Amborellaceae, Nympheaceae) ► Monocots (Liliales, Poales)
1b	5***	82.6 ± 11.5	14.4 ± 3.1	Yes	MEME motifs 5, 6, 7, 8	109.2 ± 0.7	++++	► Core eudicots (Caryophyllales, Fabales, Malpighiales, Brassicales, Solanales)
2	5	86.2 ± 7.3	5.8 ± 1.3	No	MEME motif 9	109 ± 0.0	++++	► Eurosids I (Malpighiales, Fabales)
3	3	12.6 ± 4.9	4.6 ± 4.3	No	NA	111 ± 0.8	++	► Rosids (Malpighiales, Fabales, Brassicales)
4	2	159.8 ± 30.5	10.6 ± 3.2	No	Ser rich (11.7% ± 4.2%) MEME motifs 10, 11, 12, 13	111 ± 0.8	++++	► Monocots (Poales) ► Dicots (Fabales, Brassicales)
5	4–7	90.6 ± 45.3	4.0 ± 2.9	No	MEME motifs 14, 15	110 ± 0.7	+	► Monocots (Poales) ► Dicots (Fabales, Brassicales)
6	2–7	278.5 ± 88.7	7.8 ± 2.2	No	Ser rich (13.2% ± 2.2%) and Gly rich (6.5% ± 2.7%) MEME motifs 16, 17, 18, 19	110 ± 6.0	+++	► Monocots (Poales) ► Dicots (Fabales, Brassicales, Solanales)

Downloaded from www.plantphysiol.org on June 12, 2017 - Published by www.plantphysiol.org
Copyright © 2007 American Society of Plant Biologists. All rights reserved.

genes showing the best expression correlation to AtREM1.1 expression is markedly enriched in genes associated with cell wall biology and extracellular functions (Fig. 3). Several of the well-correlated genes encode proteins with structural, signaling, or kinase activity. These data point to a possible link of group 1 remorins with cell wall biology and/or assembly.

Group 1b is the second group of proteins comprising both Remorin_N and Remorin_C domains, but their N-terminal regions contain twice the number of Pro residues compared to group 1a remorins. Only dicotyledonous plant species fall into this subgroup, suggesting that the Pro enrichment of the Remorin_N domain occurred after divergence of dicots. Group 1b contains the first discovered remorins, StREM1.3 (Reymond et al., 1996) and LeREM1 (Bariola et al., 2004; now named SlREM1.2), to which most of the current literature refers. The described Arabidopsis remorin AtDBP (AtREM1.3) and most remorins currently identified in proteomic screens of membrane preparations also cluster in this group (Watson et al., 2003; Marmagne et al., 2004; Mongrand et al., 2004; Sazuka et al., 2004; Nelson et al., 2006; Valot et al., 2006). Association with membranes thus seems to be a common feature of group 1b remorins. MEME motifs found in the N-terminal part of group 1b remorins are close to those found for group 1a (Supplemental Table S2). In addition to the Eukaryotic Linear Motif identified in group 1a, group 1b well-conserved motifs also notably bear an endoplasmic reticulum retention and retrieving signal.

Abundance of group 1b remorins in proteomic studies may be explained by the high expression level of the corresponding genes. Indeed, AtREM1.2 and AtREM1.3 are among the 10% most highly and ubiquitously expressed genes in Arabidopsis (Supplemental Table S3). AtREM1.2 and AtREM1.4 are likely to represent genes derived from a recent duplication event (Supplemental Fig. S1), and their function may therefore be redundant.

Expression of group 1b genes in Arabidopsis is markedly increased in mature and/or senescing tissues (mature leaves, roots, mature pollen, flowers) and in vascular tissue (Bariola et al., 2004; AtGenExpress). Transient gene induction was also found following abiotic stress treatment, such as cold, osmotic, or drought (Bray, 2002; Reddy et al., 2002), and by addition of the phytohormone ABA (AtGenExpress; Lin et al., 2003). Another link with plant hormones is also suggested by remorin AtREM1.3 (Alliotte et al., 1989), which was copurified with the cytokinin-induced response regulator AtARR4 (Yamada et al., 1998). Differential expression of group 1b remorins was also observed during plant-microbe interactions—in Arabidopsis

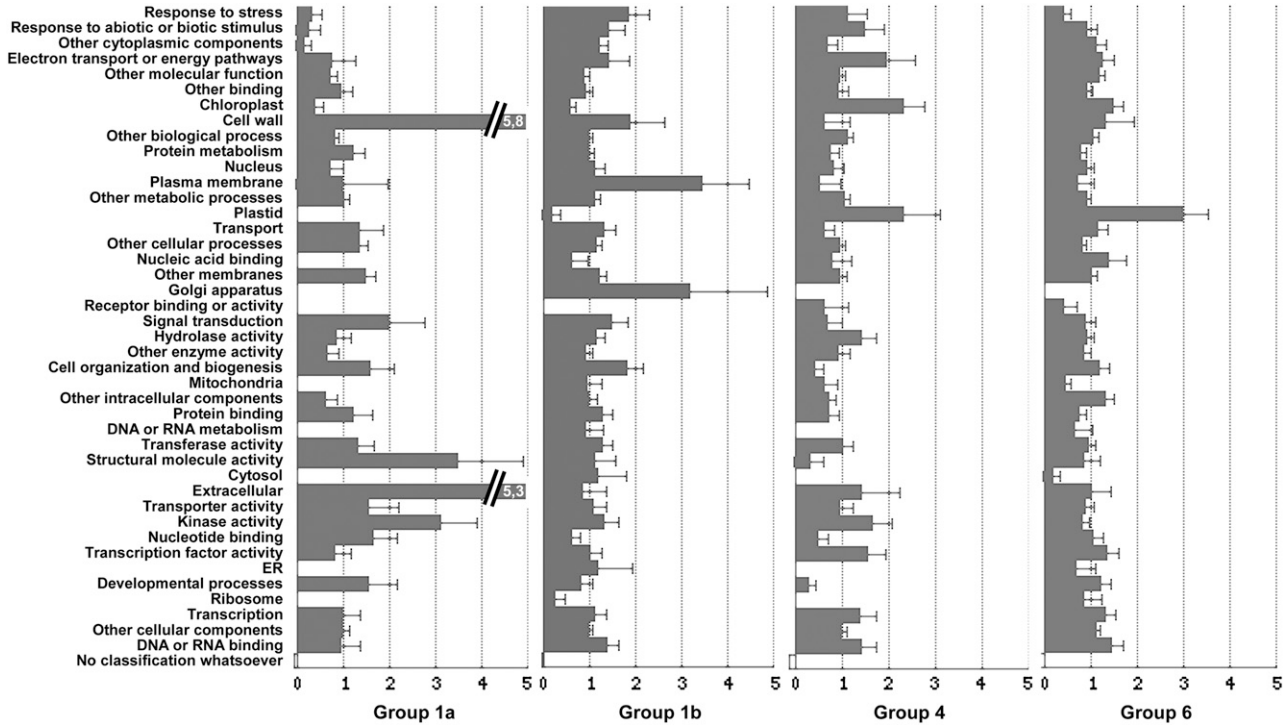

Figure 3. Coexpression analysis of remorins throughout the Arabidopsis transcriptome. Expression data from hybridizations using At22k Affymetrix GeneChips were used to retrieve a list containing 100 genes showing highest Pearson correlation coefficients for different remorin genes at http://www.bar.utoronto.ca/. Graphs represent the normalized frequency of each function in the list of well-correlated genes compared to randomly selected genes among the Arabidopsis genome (Toufighi et al., 2005). Error bars indicate variation over 100 different sets of randomly selected genes.

Downloaded from www.plantphysiol.org on June 12, 2017 - Published by www.plantphysiol.org
Copyright © 2007 American Society of Plant Biologists. All rights reserved.

upon infection with *Pseudomonas syringae* (Journot-Catalino et al., 2006) and in *Lycopersicon hirsutum* after inoculation with *Clavibacter michiganensis* (Coaker et al., 2004). Whether a direct correlation exists between abiotic/biotic factors and hormone-mediated responses remains to be elucidated. Several genes that were found to be coexpressed with group 1b remorins in Arabidopsis encode proteins associated with PMs, the Golgi network, and cell wall components. Many of them have functions related to response to stress, cell organization, and signaling (Fig. 3).

Group 2: Remorins Hitherto Found in Legumes and Poplar

Group 2 remorins contain the canonical Remorin_C domain (Fig. 1), while their N-terminal domain shows no significant protein homology to the Remorin_N domain of group 1 remorins and contains fewer Pro residues. Moreover, the MEME motifs found in their N-terminal regions differ from those of group 1. So far, group 2 contains only proteins from legumes (exemplified by sequences from *M. truncatula*) and *P. trichocarpa* (no group 2 remorins being found in Arabidopsis or rice).

Interestingly, a group 2 remorin was found during transcriptome and proteome approaches to be induced during symbiotic interactions between legumes and rhizobial bacteria (Fedorova et al., 2002; Wienkoop and Saalbach, 2003; Colebatch et al., 2004; El Yahyaoui et al., 2004). Thus, this gene was named symREM (MtREM2.2; T. Ott, unpublished data). Orthologs of this gene can only be found in other legumes, such as *Lotus japonicus* and *Glycine max*.

In addition, a gene encoding a group 2 remorin was found to be strongly induced during the symbiotic interaction between the model legume *L. japonicus* and the mycorrhizal fungus *Glomus intraradices* (Kistner et al., 2005). However, the absence of orthologous sequences in rice or other plants that are able to undergo a mycorrhizal symbiosis suggests that this remorin isoform is not essential for such interaction.

Group 3: Short Remorins

Remorins from group 3 only consist of a Remorin_C domain that is more divergent with some conserved, positively charged residues being replaced by aliphatic amino acids. Nevertheless, it includes the essential residues of the remorin signature. The absence of an N-terminal domain suggests that group 3 may serve basic functions. AtREM3.1 and AtREM3.2 show high sequence similarities with the C-terminal regions of AtREM6.6 (group 6) and AtREM1.2 (group 1), respectively. These genes are believed to result from ancient gene duplications (Supplemental Fig. S1). Differential expression of AtREM3.2 has been found under abiotic stress conditions, while for AtREM3.1 no expression data have been deposited in the AtGenExpress database.

Group 4: Remorins with Alternative Pro-Rich N-Terminal Regions

The Remorin_C domain of group 4 remorins is very similar to that of canonical remorins, and this group also contains genes comprising only two exons. Group 4 remorins are characterized by a long Pro- and Ser-rich N-terminal region (about 160 amino acids), that does not show any significant homology to the group 1 Remorin_N domain. Conserved motifs found in group 4 N-terminal regions (Supplemental Table S2) indicate that most of these remorins possess (1) two positions resembling phosphopeptide-binding sites from proteins associated with cell-cycle checkpoint and DNA repair functions, and (2) a site predicted to mediate protein-protein interaction through binding of short Pro-rich regions within proteins.

Sequence similarities suggest that AtREM4.1 and AtREM4.2 were derived from each other by a recent duplication event within the Arabidopsis genome (Supplemental Fig. S1). These two genes are predominantly expressed in green, aerial plant organs, namely, cauline and senescing leaves, as well as stems. They are strongly stress responsive, showing a marked increase after osmotic, salt, and drought stresses, and treatment of plants with phytohormones such as ABA and brassinosteroids. Furthermore, these genes are transcriptionally induced upon pathogen infections (Supplemental Table S3). Such links are supported by the expression of group 4 remorins being correlated with those of genes associated with chloroplasts and plastids, energy pathways, as well as abiotic or biotic stimuli (Fig. 3).

Group 5: Remorins with Low N-Terminal Pro Content

Group 5 remorins show a more divergent Remorin_C domain, as indicated by lower score values given by protein motif-scanning programs (Table I). This group consists of a highly variable N-terminal region, although all members contain two conserved MEME motifs that bear putative phosphorylation sites (Supplemental Table S2). Due to the absence of the group 5 remorin (AtREM5.1) from the ATH1 GeneChip and the lack of any significant sequence similarity of their N-terminal regions, it remains difficult to predict the functionality of remorins from this group. However, MtREM5.1 of *M. truncatula*, the putative ortholog of AtREM5.1, is one of the major remorins expressed in seeds (V. Benedito and M. Udvardi, personal communication).

Group 6: Long Remorins

The last group contains remorins ranging from 240 to 522 amino acids in length. The size of their N-terminal regions is variable and may reflect the existence of several subgroups, which are difficult to define based on currently available data. Interestingly, sequence similarities in the N-terminal regions of

Downloaded from www.plantphysiol.org on June 12, 2017 - Published by www.plantphysiol.org
Copyright © 2007 American Society of Plant Biologists. All rights reserved.

some members of group 6 were found with a xylulokinase homolog from *Strongylocentrotus purpuratus* (involved in pentose and glucuronate interconversion; GI:72085098) and MurB reductase from *P. patens* (involved in amino-sugar conversion; GI:76880168). In addition, four conserved MEME motifs can be detected in their N-terminal regions, indicating that many remorins of this group contain (1) conserved putative phosphorylation sites, (2) a predicted protein-protein interaction motif, and, for a subgroup of them, (3) a predicted Tyr-based sorting signal responsible for the interaction with an adaptor protein complex.

Transcript levels of Arabidopsis remorins AtREM6.5 and AtREM6.1 are highest in meristematic tissues. A remarkable consistency in repression of these genes after pathogen infection, cold and osmotic stress, as well as induction after salt stress was also noted (see Supplemental Table S3). Interestingly, another study predicted a putative recognition site for the bacterial effector protein AvrRpt2, encoding a Cys protease, within a group 6 remorin (AtREM6.1; Chisholm et al., 2005). However, such direct targeting of remorins by bacterial proteases and other roles during plant-microbe interactions remain to be proven.

POSSIBLE BIOLOGICAL ROLES OF REMORINS

Remorin proteins appear to be specific to the plant kingdom. The absence of remorins in algae, but their presence in mosses, ferns, and higher plants, suggests that the emergence of remorins coincided with the colonization of land. It is very likely that this colonization required chemical adaptation of cell walls to cope with particularly rigorous selection pressure. Evolution in cell wall composition is probably related to the emergence of an upright position and the acquisition of vascular tissue. Major differences in primary cell wall components were found by comparing taxa from charophytes (the closest extant algal ancestor of land plants), bryophytes, and pteridophytes to vascular plants. For example, major components of cell walls, such as xyloglucan (the main hemicellulose) and Hyp-rich glycoproteins (major components of cell walls), are present in bryophytes, pteridophytes, and vascular plants, but are not detectable in charophytes. Similarly, acid hydrolysis of cell wall-rich material revealed that uronic acids were more abundant in ancestral plants than in more recent vascular ones, and that GalUA is more abundant in bryophytes and charophytes than in vascular plants. Finally, mixed-linkage glucan was absent from all algae and bryophytes tested (for review, see Popper and Fry, 2003).

Interestingly, the canonical potato StREM1.3 was found to be able to bind both simple and complex galacturonides, and interaction with polyanionic cell or cell wall components can be hypothesized (Reymond et al., 1996). It remains an intriguing question as to whether remorins were positively selected together with these molecules, thus giving an adaptive advantage that allowed aquatic-to-land transition.

Precise biological roles of different remorins from the various groups remain to be investigated, but gene expression data suggest that some of these proteins might have key functions during responses to biotic and abiotic stimuli and might possibly be involved in hormone-mediated responses and signal transduction. Transient and rapid induction of gene expression (e.g. AtREM1.2, AtREM1.4, and AtREM6.1) upon biotic stimuli, such as powdery mildew infection, suggests possible roles in early stages of plant-microbe interactions. Whether such roles can be directly linked to cell wall- and/or PM-associated functions remains to be elucidated.

With respect to the variability of the N-terminal parts among the different groups and the hypothesis that this region confers functional specificity of the protein, different functions might be expected.

Interestingly, no homozygous insertion mutants are available for any remorin gene in Arabidopsis, possibly indicating lethal phenotypes when remorin gene expression is completely abolished. Other model systems, such as symbiotic root nodulation, as described above, may therefore be helpful to elucidate the biological functions of certain remorins. The nomenclature suggested here should simplify analysis and classification of detected remorin transcripts and proteins during future experiments and lead to a more comprehensive view of this novel protein family.

Supplemental Data

The following materials are available in the online version of this article.

Supplemental Figure S1. Chromosomal positions and duplication events for remorin genes within the Arabidopsis genome.

Supplemental Figure S2. Global alignment of 85 remorin proteins used for phylogenetic analysis.

Supplemental Table S1. List of remorin sequences retrieved from publicly available databases as described in Supplemental Materials and Methods S1.

Supplemental Table S2. Identification of characteristic MEME motifs leading to specific remorin signatures.

Supplemental Table S3. A set of expression data of remorin genes from Arabidopsis.

Supplemental Materials and Methods S1. Detailed description of methods used for this study.

ACKNOWLEDGMENTS

We thank Steven Ball for fruitful discussion and Mark Cock for the BLAST analysis on *E. siliculosus* sequences. Furthermore, we would like to thank Julie Cullimore and Clare Gough for critical reading of the manuscript.

Received September 5, 2007; accepted September 13, 2007; published November 7, 2007.

LITERATURE CITED

Alliotte S, Tire C, Engler G, Peleman J, Caplan A, Van Montagu M, Inzé D (1989) An auxin-regulated gene of *Arabidopsis thaliana* encodes a DNA-binding protein. Plant Physiol **89:** 743–752

Downloaded from www.plantphysiol.org on June 12, 2017 - Published by www.plantphysiol.org
Copyright © 2007 American Society of Plant Biologists. All rights reserved.

Angiosperm Phylogeny Group (2003) An update of the Angiosperm Phylogeny Group classification for the orders and families of flowering plants: APG II. Bot J Linn Soc **141:** 399–436

Bailey TL, Elkan C (1994) Fitting a mixture model by expectation maximization to discover motifs in biopolymers. *In* Proceedings of the Second International Conference on Intelligent Systems for Molecular Biology. AAAI Press, Menlo Park, CA, pp 28–36

Bariola PA, Retelska D, Stasiak A, Kammerer RA, Fleming A, Hijri M, Frank S, Farmer EE (2004) Remorins form a novel family of coiled coil-forming oligomeric and filamentous proteins associated with apical, vascular and embryonic tissues in plants. Plant Mol Biol **55:** 579–594

Bhat RA, Panstruga R (2005) Lipid rafts in plants. Planta **223:** 5–19

Borner GH, Sherrier DJ, Weimar T, Michaelson LV, Hawkins ND, Macaskill A, Napier JA, Beale MH, Lilley KS, Dupree P (2005) Analysis of detergent-resistant membranes in Arabidopsis. Evidence for plasma membrane lipid rafts. Plant Physiol **137:** 104–116

Bray EA (2002) Abscisic acid regulation of gene expression during water-deficit stress in the era of the Arabidopsis genome. Plant Cell Environ **25:** 153–161

Burkhard P, Stetefeld J, Strelkov SV (2001) Coiled coils: a highly versatile protein folding motif. Trends Cell Biol **11:** 82–88

Chisholm ST, Dahlbeck D, Krishnamurthy N, Day B, Sjolander K, Staskawicz BJ (2005) Molecular characterization of proteolytic cleavage sites of the Pseudomonas syringae effector AvrRpt2. Proc Natl Acad Sci USA **102:** 2087–2092

Coaker GL, Willard B, Kinter M, Stockinger EJ, Francis DM (2004) Proteomic analysis of resistance mediated by Rcm 2.0 and Rcm 5.1, two loci controlling resistance to bacterial canker of tomato. Mol Plant Microbe Interact **17:** 1019–1028

Colebatch G, Desbrosses G, Ott T, Krusell L, Montanari O, Kloska S, Kopka J, Udvardi MK (2004) Global changes in transcription orchestrate metabolic differentiation during symbiotic nitrogen fixation in Lotus japonicus. Plant J **39:** 487–512

El Yahyaoui F, Kuster H, Ben Amor B, Hohnjec N, Puhler A, Becker A, Gouzy J, Vernie T, Gough C, Niebel A, et al (2004) Expression profiling in *Medicago truncatula* identifies more than 750 genes differentially expressed during nodulation, including many potential regulators of the symbiotic program. Plant Physiol **136:** 3159–3176

Farmer EE, Pearce G, Ryan CA (1989) In vitro phosphorylation of plant plasma membrane proteins in response to the proteinase inhibitor inducing factor. Proc Natl Acad Sci USA **86:** 1539–1542

Fedorova M, van de Mortel J, Matsumoto PA, Cho J, Town CD, VandenBosch KA, Gantt JS, Vance CP (2002) Genome-wide identification of nodule-specific transcripts in the model legume *Medicago truncatula.* Plant Physiol **130:** 519–537

Jacinto T, Farmer EE, Ryan CA (1993) Purification of potato leaf plasma membrane protein pp34, a protein phosphorylated in response to oligogalacturonide signals for defense and development. Plant Physiol **103:** 1393–1397

Journot-Catalino N, Somssich IE, Roby D, Kroj T (2006) The transcription factors WRKY11 and WRKY17 act as negative regulators of basal resistance in *Arabidopsis thaliana.* Plant Cell **18:** 3289–3302

Kistner C, Winzer T, Pitzschke A, Mulder L, Sato S, Kaneko T, Tabata S, Sandal N, Stougaard J, Webb KJ, et al (2005) Seven *Lotus japonicus* genes required for transcriptional reprogramming of the root during fungal and bacterial symbiosis. Plant Cell **17:** 2217–2229

Laliberte JP, McGinnes LW, Peeples ME, Morrison TG (2006) Integrity of membrane lipid rafts is necessary for the ordered assembly and release of infectious Newcastle disease virus particles. J Virol **80:** 10652–10662

Laloi M, Perret AM, Chatre L, Melser S, Cantrel C, Vaultier MN, Zachowski A, Bathany K, Schmitter JM, Vallet M, et al (2007) Insights into the role of specific lipids in the formation and delivery of lipid microdomains to the plasma membrane of plant cells. Plant Physiol **143:** 461–472

Lefebvre B, Furt F, Hartmann MA, Michaelson LV, Carde JP, Sargueil-Boiron F, Rossignol M, Napier JA, Cullimore J, Bessoule JJ, et al (2007) Characterization of lipid rafts from *Medicago truncatula* root plasma membranes: a proteomic study reveals the presence of a raft-associated redox system. Plant Physiol **144:** 402–418

Lin F, Xu SL, Ni WM, Chu ZQ, Xu ZH, Xue HW (2003) Identification of ABA-responsive genes in rice shoots via cDNA macroarray. Cell Res **13:** 59–68

Manes S, del Real G, Lacalle RA, Lucas P, Gomez-Mouton C, Sanchez-Palomino S, Delgado R, Alcami J, Mira E, Martinez AC (2000) Membrane raft microdomains mediate lateral assemblies required for HIV-1 infection. EMBO Rep **1:** 190–196

Marmagne A, Rouet MA, Ferro M, Rolland N, Alcon C, Joyard J, Garin J, Barbier-Brygoo H, Ephritikhine G (2004) Identification of new intrinsic proteins in Arabidopsis plasma membrane proteome. Mol Cell Proteomics **3:** 675–691

Mongrand S, Morel J, Laroche J, Claverol S, Carde JP, Hartmann MA, Bonneu M, Simon-Plas F, Lessire R, Bessoule JJ (2004) Lipid rafts in higher plant cells: purification and characterization of Triton X-100-insoluble microdomains from tobacco plasma membrane. J Biol Chem **279:** 36277–36286

Nelson CJ, Hegeman AD, Harms AC, Sussman MR (2006) A quantitative analysis of Arabidopsis plasma membrane using trypsin-catalyzed (18)O labeling. Mol Cell Proteomics **5:** 1382–1395

Nishina K, Deleault NR, Lucassen RW, Supattapone S (2004) In vitro prion protein conversion in detergent-solubilized membranes. Biochemistry **43:** 2613–2621

Ono A, Freed EO (2001) Plasma membrane rafts play a critical role in HIV-1 assembly and release. Proc Natl Acad Sci USA **98:** 13925–13930

Peskan T, Westermann M, Oelmuller R (2000) Identification of low-density Triton X-100-insoluble plasma membrane microdomains in higher plants. Eur J Biochem **267:** 6989–6995

Popper ZA, Fry SC (2003) Primary cell wall composition of bryophytes and charophytes. Ann Bot (Lond) **91:** 1–12

Puntervoll P, Linding R, Gemünd C, Chabanis-Davidson S, Mattingsdal M, Cameron S, Martin DM, Ausiello G, Brannetti B, Costantini A, et al (2003) ELM server: a new resource for investigating short functional sites in modular eukaryotic proteins. Nucleic Acids Res **31:** 3625–3630

Reddy AR, Ramakrishna W, Sekhar AC, Ithal N, Babu PR, Bonaldo MF, Soares MB, Bennetzen JL (2002) Novel genes are enriched in normalized cDNA libraries from drought-stressed seedlings of rice (Oryza sativa L. subsp. indica cv. Nagina 22). Genome **45:** 204–211

Reymond P, Kunz B, Paul-Pletzer K, Grimm R, Eckerskorn C, Farmer EE (1996) Cloning of a cDNA encoding a plasma membrane-associated, uronide binding phosphoprotein with physical properties similar to viral movement proteins. Plant Cell **8:** 2265–2276

Samuel BU, Mohandas N, Harrison T, McManus H, Rosse W, Reid M, Haldar K (2001) The role of cholesterol and glycosylphosphatidylinositol-anchored proteins of erythrocyte rafts in regulating raft protein content and malarial infection. J Biol Chem **276:** 29319–29329

Sazuka T, Keta S, Shiratake K, Yamaki S, Shibata D (2004) A proteomic approach to identification of transmembrane proteins and membrane-anchored proteins of Arabidopsis thaliana by peptide sequencing. DNA Res **11:** 101–113

Shahollari B, Peskan-Berghoefer T, Oelmueller R (2004) Receptor kinases with leucine-rich repeats are enriched in Triton X-100 insoluble plasma membrane microdomains from plants. Physiol Plant **122:** 397–403

Toufighi K, Brady SM, Austin R, Ly E, Provart NJ (2005) The Botany Array Resource: e-Northerns, Expression Angling, and promoter analyses. Plant J **43:** 153–163

Valot B, Negroni L, Zivy M, Gianinazzi S, Dumas-Gaudot E (2006) A mass spectrometric approach to identify arbuscular mycorrhiza-related proteins in root plasma membrane fractions. Proteomics (Suppl 1) **6:** S145–S155

Watson BS, Asirvatham VS, Wang L, Sumner LW (2003) Mapping the proteome of barrel medic (*Medicago truncatula*). Plant Physiol **131:** 1104–1123

Wienkoop S, Saalbach G (2003) Proteome analysis. Novel proteins identified at the peribacteroid membrane from *Lotus japonicus* root nodules. Plant Physiol **131:** 1080–1090

Yamada H, Hanaki N, Imamura A, Ueguchi C, Mizuno T (1998) An Arabidopsis protein that interacts with the cytokinin-inducible response regulator, ARR4, implicated in the His-Asp phosphorylay signal transduction. FEBS Lett **436:** 76–80

Downloaded from www.plantphysiol.org on June 12, 2017 - Published by www.plantphysiol.org
Copyright © 2007 American Society of Plant Biologists. All rights reserved.