Distinct Patterns of Symbiosis-Related Gene Expression in Actinorrhizal Nodules from Different Plant Families

Katharina Pawlowski,^{1,2} Susan Swensen,³ Changhui Guan,¹ Az-Eddine Hadri,¹ Alison M. Berry,⁴ and Ton Bisseling¹

¹Department of Molecular Biology, Agricultural University, Dreijenlaan 3, 6703 HA Wageningen, The Netherlands; ²Albrecht von Haller Institute for Plant Sciences, Plant Biochemistry, Göttingen University, Justus-von-Liebig-Weg 11, 37077 Göttingen, Germany; ³Department of Biology, Ithaca College, 953 Danby Road, Ithaca, NY 14850, U.S.A.; ⁴Department of Environmental Horticulture, University of California, Davis 95616, U.S.A.

Submitted 19 August 2002. Accepted 2 June 2003.

Phylogenetic analyses suggest that, among the members of the Eurosoid I clade, nitrogen-fixing root nodule symbioses developed multiple times independently, four times with rhizobia and four times with the genus *Frankia*. In order to understand the degree of similarity between symbiotic systems of different phylogenetic subgroups, gene expression patterns were analyzed in root nodules of *Datisca glomerata* and compared with those in nodules of another actinorrhizal plant, *Alnus glutinosa*, and with the expression patterns of homologous genes in legumes. In parallel, the phylogeny of actinorrhizal plants was examined more closely. The results suggest that, although relationships between major groups are difficult to resolve using molecular phylogenetic analysis, the comparison of gene expression patterns can be used to inform evolutionary relationships. In this case, stronger similarities were found between legumes and intracellularly infected actinorrhizal plants (*Alnus*) than between actinorrhizal plants of two different phylogenetic subgroups (*Alnus/Datisca*).

Additional keywords: *ag164*, *agNod-GHRP*, major latex protein, PR10.

Phylogenetic analyses based on the comparison of *rbcl* gene sequences have shown that all plants able to establish a root nodule symbiosis are members of a single clade, the Eurosoid I clade. Therefore, it has been proposed that the underlying capacity to enter a root nodule symbiosis with nitrogen-fixing bacteria has been acquired by the ancestor of this clade, possibly in a single evolutionary event (Soltis et al. 1995). More detailed phylogenetic studies suggested that, among the descendants of the Eurosoid I ancestor, the ability to enter a symbiosis with rhizobia was acquired three times independently within the legume family (Doyle et al. 1997), and once for the genus *Parasponia* (*Ulmaceae*), the only nonlegume able to form a symbiosis with rhizobia. Similarly, the ability to enter a symbiosis with filamentous nitrogen-fixing actinomycetes of the genus *Frankia* (actinorrhizal symbioses) may have arisen three or four times independently in the Eurosoid I clade (Swensen 1996). Altogether, plant species able to enter an endosymbiotic nitrogen-fixing symbiosis are found in four subclades of Eurosoid I, interspersed with plant species lacking this ability (Fig. 1).

These hypothetical independent origins of the symbiotic interaction with either rhizobia or *Frankia* spp. are reflected in differences in plant-determined characteristics of the symbiosis, such as nodule morphology and infection mechanism (Swensen 1996). For instance, nodules induced by rhizobia on members of *Leguminosae* represent stem-like organs with a peripheral vascular system and infected cells in the central tissue, whereas nodules induced on other plants are actually modified lateral roots with a central vascular system and infected cells in the expanded cortex. Plants of subclade B (Fig. 1; *Alnus*, *Casuarinaceae*, *Myricaceae*) are the only hosts of *Frankia* to be infected intracellularly, via root hairs (Benson and Silvester 1993). Actinorrhizal nodules of *Datisca* (*Datisceae*) and *Coriaria* (*Coriariaceae*), the only symbiotic members of subclade E, show a unique distribution of the infected cortical cells which form a continuous patch, not interspersed with uninfected cells, on one side of the acentric stele, kidney-shaped in cross section (Hafeez et al. 1984; Newcomb and Pankhurst 1982).

The fact that species of *Parasponia*, which enter a symbiosis with rhizobia, group among actinorrhizal plants (Swensen 1996), suggests that the taxonomic affiliation of the microsymbiont may not be the most fundamental difference between these symbiotic interactions. The close relationship between actinorrhizal and rhizobial symbioses also is underscored by the discovery of similar gene expression systems. Jacobsen-Lyon and associates (1995) have shown that the promoter of a nodule-specific hemoglobin gene from the actinorrhizal plant *Casuarina glauca* also directs expression in transgenic nodules of the legume *Lotus corniculatus* in the same cells where the endogenous leghemoglobin genes are expressed (the infected cells), in spite of the differences in nodule morphology. In contrast, the *Parasponia andersonii* hemoglobin promoter did not direct expression specifically in the infected cells of transgenic *Lotus* nodules (Andersson et al. 1997). Thus, some infected cell-specific transcription factors seem to be conserved among *C. glauca* and legumes, but not *Parasponia* spp., suggesting that their evolution has predated the development of microsymbiont specificity. Furthermore, the existence of a nodule-specific and a vegetative hemoglobin gene family is a feature shared by legumes and *C. glauca*, but not by *P. andersonii*, where a single gene fulfills both functions (Landsmann et al. 1986), again suggesting an ancestor common to legumes and *C. glauca*, but not to *P. andersonii* (Jacobsen-Lyon et al. 1995). Hence, the data on transcription factor conservation and hemoglobin gene duplication suggest a common evolutionary origin of the symbiotic syndrome in legumes and intracellularly infected actinorrhizal plants.

Corresponding author: K. Pawlowski; Telephone: +49-551-3914430; Fax: +49-551-395749; E-mail: kpawlow@gwdg.de.

If legumes are more closely related to the actinorhizal *C. glauca* than to (*Brady*-)*Rhizobium*-nodulated *P. andersonii*, then which symbiotic features would be shared among plants from a common subclade? The genus *Parasponia* is found in the same subclade as the actinorhizal families Rosaceae, Elaeagnaceae, and Rhamnaceae according to *rbcL* phylogeny (Swensen 1996); however, the different microsymbionts (rhizobia or *Frankia* spp.) involved in these symbioses have different influences on their plant hosts. For example, *Frankia* can fix nitrogen under atmospheric oxygen tension by forming their own specialized vesicles where nitrogenase is protected from oxygen (Parsons et al. 1987), but rhizobia depend on plant-derived structures to form oxygen barriers within the nodule. These different strategies to control nodule oxygen also affect carbon and nitrogen metabolism. We are interested in knowing whether differences between actinorhizal and rhizobial nodules of the same subclade exceed those between actinorhizal nodules formed by plants belonging to different subclades.

To address this question, we have compared tissue specialization for nitrogen assimilation and expression patterns of nodule-enhanced genes in actinorhizal nodules formed by plants from different subclades, and in legume nodules. The acti-

norhizal plants *Alnus glutinosa* and *Datisca glomerata* were chosen for this comparison. *A. glutinosa*, a representative of the intracellularly infected actinorhizal plants (Fig. 1, group B), is well studied on the molecular level (Pawlowski and Bisseling 1996); whereas the nodules of *D. glomerata* show structural features found only in nodules of group E (Fig. 1) (Hafeez et al. 1984). Published legume data were used for comparison, except for the *Pisum sativum* homologue of actinorhizal *PR10* gene which we obtained for comparison.

RESULTS

Phylogenetic analysis of symbiotic plants.

The phylogenetic analysis of 107 taxa using parsimony indicates that legumes (Fig. 1, group A) are basal to other root-nodulated host plants. The basal position of legumes is invariant among the 72 most parsimonious trees obtained in the parsimony searches. However, this position was not found in the parsimony jackknife tree; instead, the legumes were part of a basal polytomy along with other actinorhizal groups (*tree not shown*). The basal position of legumes also has been observed in phylogenetic analyses of angiosperms using combined data

Fig. 1. Phylogenetic analysis. Summary diagram of a strict consensus tree based on 72 equally parsimonious trees of length 4077, consistency index = 0.284 (excluding autapomorphies); rescaled consistency index = 0.241. Bold lines or triangles represent taxa containing species with root nodules. Numbers below the branches are support values obtained from parsimony jackknife analysis. Branches with values <50% were not found in the parsimony jackknife analysis. For simplicity, closely grouped species are represented by group, family, or subfamily names followed by the number of analyzed species in parentheses.

from three different genes (Soltis et al. 2000); however, the legume position in that case also was not supported by parsimony jackknife analysis.

Among the actinorhizal species (Fig. 1, groups B through E), group E (*Datisca*, *Coriaria*) appears basal to other actinorhizal groups, but this position lacks jackknife support. Group B (Betulaceae, Casuarinaceae, Myricaceae), group C (Rhamnaceae, Elaeagnaceae), and group D (Rosaceae) appear more closely related to each other than to group E (*Coriaria*, *Datisca*), but again without support from jackknife values. Groups C and D appear as sister groups in parsimony analyses

and also have jackknife support of 70%. The phylogenetic analysis provides a view of the composition of various actinorhizal and legume groups, but does not provide a clear picture of how these different groups are related evolutionarily. Thus, although *Datisca* and *Coriaria* (group E) appear as the basal-most actinorhizal group after the legumes, there is no support for this position; group B or group C/D are equally plausible candidates for this position.

Isolation of *D. glomerata* cDNAs representing genes expressed at elevated levels in nodules compared to roots.

A cDNA library was constructed from *Datisca glomerata* nodules and screened differentially with nodule and root cDNA to isolate cDNA clones representing genes expressed at elevated levels in nodules (Okubara et al. 1999). These genes will be termed "actinorhizal nodulin genes", in analogy to legume nodulin genes, in this article. The expression patterns of the genes represented by the isolated cDNAs, as well as the expression pattern of glutamine synthetase (GS), were confirmed by hybridization with RNA gel blots containing RNA from roots, nodules, leaves, flowers, and developing fruit of *D. glomerata* plants. Two cDNA clones, pDg200 and pDg241, representing full-length cDNAs of genes expressed at low levels in roots and at high levels in nodules (Fig. 2A and B), were selected for further analysis.

Sequence analysis revealed that *dg200* represents a gene similar to *ag164/agNt84*, a family of nodule-specific genes encoding small glycine- and histidine-rich proteins identified from *A. glutinosa* and collectively called *AgNod-GHRPs* (Fig. 3A; EMBL accession number AJ489322) (Pawlowski et al. 1997). Accordingly, *dg200* was renamed *dgGHRP1*. Sequence analysis revealed that *dg241* encodes a member of the so-called PR10 protein family, which includes pollen allergens, stress-induced proteins, and pathogenesis related proteins (Fig. 3B; EMBL accession number AJ489320) (Breiteneder et al. 1989; Walter et al. 1990; Warner et al. 1992). Some members of this family are expressed at high levels in legume roots but at low levels in nodules (Gamas et al. 1998; Mylona et al. 1994; Sikorski et al. 1999). Accordingly, *dg241* was renamed *dgPR10A*.

Isolation of *Alnus* and *Datisca* cDNAs encoding members of the PR10 family.

To compare the expression levels of a PR10 gene in roots versus nodules of *Alnus glutinosa*, a degenerate oligonucleotide primer was constructed based on a conserved region of PR10 family protein sequences (Fig. 3B). This primer was used in combination with an oligo-dT primer to isolate a homolog from *Alnus* spp. by reverse-transcription polymerase chain reaction (RT-PCR) with nodule RNA. The identity of the PCR product, which was termed *agPR10A*, was confirmed by sequencing. The deduced amino acid sequence is shown in Figure 3B (EMBL accession number AK489323). *agPR10A* is expressed at high levels in roots and at low levels in nodules (Fig. 2C), similar to the expression of pea *RH2* (a PR10 gene) and lupine *PR10.1A/PR10.1B* (Sikorski et al. 1998). Low-level hybridization with RNA of male flowers from *A. glutinosa* is likely to represent cross hybridization with mRNA of the pollen allergen (Breiteneder et al. 1992), which represents another member of the PR10 protein family in *A. glutinosa* (Fig. 3B). Studies by Swoboda and associates (1995) also have shown that in birch, a close relative of the genus *Alnus*, different members of the PR10 protein gene family are expressed in pollen as pollen allergens and in somatic cells in response to microorganisms.

In an independent study, a partial cDNA was amplified from *D. glomerata* nodule cDNA that encoded a major latex protein 1-like protein (MLP1). This group of latex proteins shares

Fig. 2. Gene expression levels in different organs of actinorhizal plants. **A**, *dgGHRP1* (*dg200*), **B**, *dg241* (*dgPR10A*), **C**, *ag241* (*agPR10A*), and **D**, *dgMPL1*. RNA gel blots containing approximately 10 µg of total RNA isolated from different plant organs per slot were hybridized with ³²P-labeled cDNA probes. The amount of mRNA on the blot was determined by hybridization with a heterologous ubiquitin probe (Kouchi and Hata 1993). R, roots; N, nodules; L, leaves (in C, shoot tips); F, flowers; ff, female flowers; mF, male flowers; dF, developing fruit; dA, developing fruits harvested in April; dJ, developing fruits harvested in June; dS, developing fruits harvested in September.

structural and sequence similarity with PR10 proteins (Osmark et al. 1998) (Fig. 3B). Hence, another member of the PR10 family was available from *D. glomerata* spp. The full-length cDNA sequence of *dgMLP1* was obtained by 5'-RACE (EMBL accession number AJ489321). *dgMLP1* was expressed at comparable levels in roots, nodules, leaves, and flowers of *D. glomerata* plants (Fig. 2D).

Comparison of gene expression patterns in *Alnus*, *Datisca*, and legume nodules.

Two types of actinorhizal nodulin genes were identified in both *A. glutinosa* and *D. glomerata*; the PR10 genes (*dgPR10* and *agPR10A*) and the glycine- and histidine-rich protein genes (*dgGHRP*, *ag164*, and *agNt84*). In addition, PR10 genes from pea and lupine nodules were available. To find out whether the expression of these genes was similarly regulated in the different symbiotic systems, we used in situ hybridization of antisense mRNA to longitudinal nodule sections. Hybridization with *Frankia nifH* was performed on an adjacent nodule section to visualize the developmental state of the infected cells, with results as shown (Fig. 4A).

The expression of *dgPR10A* in *D. glomerata* nodules was confined to the infected cells, where it followed a complicated time course (Fig. 4B). *dgPR10A* expression was induced upon infection of cortical cells and was maintained throughout the development of infected cells and throughout the nitrogen-fixation zone. However, *dgPR10A* expression was reduced significantly once the infected cells were visibly filled with *Frankia* hyphae, and reinduced when vesicles had differentiated and nitrogen fixation began, as denoted by *Frankia nifH* expression (Fig. 4B). Concomitant with evidence of microsymbiont degradation in the senescence zone, *dgPR10A* expression was turned off (*data not shown*). The PR10 gene from pea, *RH2*, was expressed in nodules (Fig. 4H and I), though at much lower levels than in roots (*data not shown*). Expression of *RH2* in nodules was confined to the nodule cortex. Unfortunately, expression of *agPR10A* in *A. glutinosa* nodules was too low to be detected using in situ hybridization (*data not shown*). *dgMPL1* was expressed in the cortical cell layers of *Datisca* nodules just below the nodule periderm, in external nodule tissue roughly equivalent to the legume nodule cortex (Fig. 4J). Sometimes, *dgMLP1* expression also was detected in individual uninfected cells of the nodule cortex (Fig. 4K).

Although the *agNod-GHRPs* are expressed in infected cortical cells of *A. glutinosa* nodules before the onset of nitrogen fixation (Pawlowski et al. 1997), the expression of *dgGHRP1* (*dg200*) in *Datisca* nodules was confined mostly to cells with suberized walls that form the nodule periderm. Lower levels of *dgGHRP1* expression could be detected in cortical cells at a stage before infection (Fig. 4D and L). The absence of microsymbiont in these cells was proven by hybridizing a parallel section with a *Frankia* rRNA probe (Fig. 4L and M). In *D. glomerata* nodules, during the infection of cortical cells, *Frankia* hyphae start propagating at the periphery rather than from the center of the cell (Hafeez et al. 1984); therefore, infection is not as easily visible as in other actinorhizal nodules. Cortical cells at the same distance from the lobe meristem, but on the uninfected side of the lobe, do not show *dgGHRP1* expression (*data not shown*). Thus, *dgGHRP1* is expressed in cortical cells in advance of infection by *Frankia*.

The expression of *dg dgGHRP1* also was examined in roots and in "nodule roots", which are agraviotropically growing roots with large air spaces in the cortex that are formed by nodules grown hydroponically. In both types of roots, *dgGHRP1* expression was confined to the cells of the multilayered pericycle (Fig. 4N). The cell walls of the nodule periderm and of the (nodule) root pericycle share the same fluorescence when

viewed under UV light, which is somewhat different than the color of fluorescence of xylem vessels (Fig. 4O), indicating that, like the periderm cell walls, the root pericycle cell walls are suberized. Hence, *dgGHRP1* is expressed in cells with suberized cell walls. This is particularly obvious in the cross section of an unusual root that showed suberization in the pericycle, and in the root cortex and a superficial periderm layer at one side. All cells with suberized walls expressed high levels of *dgGHRP1* (Fig. 4O).

Comparison of metabolic specialization of nodule tissues in *Datisca* spp., *Alnus* spp., and legumes.

To compare tissue specialization for nitrogen metabolism in the different nodule types, the expression of the key enzyme in ammonium assimilation, GS, was studied. The expression patterns of legume and *A. glutinosa* GS already have been examined (Guan et al. 1996; Temple et al. 1995). GS genes from different plants are highly conserved, allowing analysis of expression by in situ hybridization using a heterologous probe that should detect all members of the GS gene family. Expression of GS in *Datisca* nodules was analyzed by in situ hybridization using *A. glutinosa* GS antisense RNA (Guan et al. 1996) as a probe.

In actinorhizal nodules of *A. glutinosa*, GS was expressed in the infected cells and in the vascular system (Guan et al. 1996). In legume nodules, GS expression also could be found in the uninfected cells (Miao et al. 1991), but never exclusively so, or at higher levels than in infected cells (Forde et al. 1989; Temple et al. 1995). However, in actinorhizal nodules of *D. glomerata*, high levels of GS expression were detected in the uninfected cortical cells surrounding the infected cells, as well as in senescent infected cells, while expression levels in the infected cells were below the detection limit of the heterologous GS probe (Fig. 4E and F). No signal was detected in hybridizations with GS sense probes (Fig. 4G) or other sense probes (*data not shown*).

DISCUSSION

PR10 protein gene family members in nodules and roots of legumes and actinorhizal plants.

Genes encoding members of the so-called PR10 family (also termed 18-kDa protein family) have been found to be expressed developmentally, as with pollen allergens (Breiteneder et al. 1989) or in the root epidermis (Mylona et al. 1994), or in response to various environmental stimuli, such as pathogens (Walter et al. 1990), wounding (Warner et al. 1992), or abscisic acid (Itturriaga et al. 1994). The synthesis of several other pathogen-inducible proteins also can be activated under various stress conditions (Keen 1992; Lamb et al. 1989). One member of the PR10 protein family has been shown to have ribonuclease activity in vitro (Bufe et al. 1996; Swoboda et al. 1996). Similar RNase activity has been shown in a group of grass pollen allergens that are encoded by members of a gene family which also contains genes induced by salicylic acid (Bufe et al. 1995). Intracellular RNases seem to be involved in the defense against pathogens in several systems.

Several PR10 protein genes have been characterized in legumes. In *Medicago truncatula*, two PR10 protein genes were characterized (Gamas et al. 1998). *MtPR10-1* was expressed in *M. truncatula* roots and induced by pathogen infection, while *MtN13* was expressed in the nodule cortex similarly to pea *RH2*. Two PR10 protein genes have been characterized from lupine, whose expression pattern in nodules resembled those of *RH2* and *MtN13* (Sikorski et al. 1999). Both *RH2* and the lupine PR10 protein genes are expressed constitutively in roots, whereas *MtN13* is not. Apparently, PR10 proteins are active in roots as well as nodules of legumes, and different members of

Frankia hyphae and vesicles are degraded by the host cell. Whatever the function of *dgPR10A* in the symbiosis may be, this expression pattern differs from that of the legume *PR10* genes and, in all likelihood, with that of the *A. glutinosa* *agPR10A* gene. However, a gene encoding a member of a different subfamily of the PR10 family, *dgMLP1*, shows an expression pattern in *Datisca* nodules comparable to that of pea *RH2*, *MtN13*, and lupine *PR10.1A/B*. Thus, with regard to the recruitment of different PR10 protein genes for conserved functions, *A. glutinosa* seems to fall into the same group as legumes, whereas *D. glomerata* stand out.

Genes encoding small glycine- and histidine-rich proteins in actinorhizal nodules.

The expression of the *agNod-GHRPs* is induced in *A. glutinosa* nodules immediately upon infection of cortical cells by *Frankia* (Pawlowski et al. 1997). The *D. glomerata* homolog *dgGHRP1* (*dg200*) was not expressed in a nodule-specific manner, but in all cells whose cell walls contain a specific type of fluorescent aromatic compound. Additionally, *dgGHRP1* was expressed transiently in young cortical cells localized between the apical meristem of the nodule lobe and the early infection zone—a cell type without detectable secondary cell wall modifications. Altogether, this expression pattern suggests that *dgGHRP1* may have a function in the differentiation process that makes nodule cortical cells infectible by *Frankia*, and that this process may have something in common with the suberization of plant cell walls.

In contrast, the expression of the *agNod-GHRPs* is confined to cortical cells which are already in the process of infection by *Frankia* (Pawlowski et al. 1997). The expression of *dgGHRP1* in the periderm of roots and nodules is not expected in *A. glutinosa* because the periderm cell walls of *A. glutinosa* plants contain different aromatic compounds as judged by observation under fluorescent light (data not shown). The expression pattern of the *agNod-GHRPs*, while also related to the infection of cortical cells, is different from that of *dgGHRP1*. This is consistent with cytological observations indicating differences in the infection process between *A. glutinosa* and *D. glomerata* (Burgess and Peterson 1987; Hafeez et al. 1984). For instance, *Alnus* nodule cortical cells do not become multinucleate before infection, whereas *Datisca* cells do. Also, *Datisca* nodule cells are filled with branching *Frankia* hyphae from the periphery inward, whereas *Alnus* nodule cells are filled from the center outward.

Assessing nodulin gene homology.

When comparing gene expression patterns in different organisms, it is important to know whether the genes involved are orthologs. The nodulin genes under comparison here, the PR10 nodulins, and the glycine- and histidine-rich protein

genes, have been shown to be members of small multigene families (Fig. 3C and D) (Osmark et al. 1998). Although it is not known whether there are other uncharacterized members of these gene families, we have addressed the question of orthology by phylogenetic reconstruction of the known nodulins, comparing the placement of proteins from nodulating hosts of interest.

When the amino acid sequences of known PR10 proteins from legumes and actinorhizal plants were analyzed phylogenetically (Fig. 3C), the results suggested that there are three broad groupings of PR10 proteins: i) a *Pisum/Medicago* group; ii) a legume PR10 group consisting mainly of proteins from species that form determinate nodules (*Lupinus luteus*, *Phaseolus vulgaris*, and *Glycine max*), but also two PR10 proteins from pea; and iii) a group of PR10 family pollen allergens from the plant order Fagales that includes the PR10 proteins from *D. glomerata* and *A. glutinosa* discussed here. The PR10 proteins expressed in the outer cortex of legume nodules (*Pisum*AAB32504 RH2 and *Lupinus*AAC12791PR10.1B) are found in groups 1 and 2, respectively, whereas the N13 protein from *M. truncatula* (*Medicago*CAA71481N13) is not a member of any of these three groups, instead appearing closer to the outgroup represented by the major latex proteins (Fig. 3C). Based on these data, it seems likely that each grouping represents the products of different, paralogous genes. Hence, the genes recruited as nodulins in actinorhizal plants and legumes are likely to represent nonorthologous members of the PR10, and this may be reflected in our comparisons of gene expression.

All known members of glycine- and histidine-rich protein family are shown in Fig. 3A. These genes appear to be absent in some plants (e.g., *Arabidopsis thaliana*). The glycine- and histidine-rich proteins described here form a subgroup of a family of stress-, drought-, or pathogen-induced glycine-rich plant proteins (Fig. 3D) (Park et al. 2000). These glycine-rich proteins also can contain histidine residues in the glycine-rich domain, but in *DgGHRP1* (*Datisca*GHRP1), *Ag164*, *AgNt84*, and *QrPCR1* (*Quercus*CAB72442), the histidine residues alternate with glycine residues (Fig. 3A). Alternating glycine-histidine stretches have been implicated in the binding of metal ions (Conklin et al. 1992) and, for truncated versions of *AgNt84*, metal binding has been shown experimentally (Gupta et al. 2002). The phylogenetic analysis suggests that, similar to PR10 proteins, the glycine-rich proteins consist of several groups possibly representing different paralogous genes (Fig. 3D). The *agNod-GHRPs* and *dgGHRP1* (*Alnus*Ag164, *Alnus*Nt84, and *Datisca*GHRP1) belong to the same subgroup, indicating that they might be orthologs. This subgroup also includes the other known protein with alternating glycine and histidine residues, *Quercus*CAB72442 (Fig. 3A and D). Interestingly, one glycine-rich protein that does not contain GHGH stretches (*Daucus* P37704) also is included in this subgroup.

Fig. 3. Protein families. **A** and **B**, Amino acid sequence comparisons. Identical amino acids at conserved positions are labeled by inverse print, chemically similar amino acids at conserved positions are shaded in gray. Gaps to optimize the alignment were introduced using the Program ClustalW (EMBL); editor GeneDoc was used to present the alignment (Nicholas et al. 1997). **A**, *DgGHRP1* (*Dg200*) is a homolog of the small glycine- and histidine-rich metal-binding actinorhizal nodulins *Ag164* and *AgNt84* from genus *Alnus* (Pawlowski et al. 1997). Another member of this family is a phase-change related protein precursor from *Quercus robur* (GenBank accession no. AJ271778). **B**, *DgPR10A* (*Dg241*), *AgPR10A* (*Ag241*), and *DgMLP1* are members of the PR10 protein family which includes pollen allergens (*Cor-a-1* from hazelnut [Luttkopf et al. 2002], *Aln-g-1* from *Alnus glutinosa* [Breiteneder et al. 1992]), proteins from legume roots and nodules (*Pisum sativum* RH2 [Mylona et al. 1994], *Medicago truncatula* PR10-1 [Gamas et al. 1998]), and major latex proteins (MLPs [Osmark et al. 1998]; *CmMLP*, major latex protein I homolog from *Cucumis melo* [Aggelis et al. 1997]). **C** and **D**, Phylogenetic analysis of protein families. Numbers above branches are branch lengths (ACCTRAN optimization); numbers below branches are bootstrap percentages. Protein names include genus and GenBank accession number. Proteins discussed in this study are in bold and underlined, and the protein name is added after the accession number. **C**, PR10 protein family from Rosid I plants. One of two equally parsimonious trees obtained via parsimony analysis of 38 PR10 protein sequences. Tree length = 1085; consistency index (CI) = 0.626 (excluding autapomorphies); rescaled consistency index (RC) = 0.455. The dotted line in the outgroup indicates a branch that does not appear in both trees. **D**, Glycine-rich protein family. Single tree obtained via parsimony analysis of 19 glycine-rich proteins. Tree length = 1042; CI = 0.728 (excluding autapomorphies); RC = 0.408.

The glycine-rich proteins from solanaceous plants (*Nicotiana* and *Petunia* spp.) form a subgroup; however, glycine-rich proteins from legumes (CAA07551, O07202, T06256, Q09134, and AAL15649) appear in five separate locations in the tree, providing little clue as to their orthology relationships.

The glycine- and histidine-rich proteins are predicted to be exported to the apoplast (PSORT program) (Nakai and Horton

1999). It is interesting that all known glycine- and histidine-rich proteins have been found in woody plant organs, whereas *dgGHRP* (*dg200*) expression in *D. glomerata* roots is linked to secondary cell wall modifications. Thus, it is tempting to speculate that, in woody plants, a member of the family of glycine-rich proteins evolved for a specific function in modified cell walls, and that these proteins later were recruited for a function in symbiosis.

GS expression in actinorhizal and legume nodules.

In both actinorhizal and legume nodules, the microsymbiont exports ammonium to the cytoplasm of the infected cells where it is assimilated by the plant via the GS/glutamate synthase cycle (Schubert 1986). GS expression in the vascular system is thought to play a role in nitrogen transport (Edwards et al. 1990). Accordingly, in actinorhizal nodules of *A. glutinosa*, GS expression was found in the infected cortical cells and in the vascular system (Guan et al. 1996). GS expression in nodules of another actinorhizal member of group B, *C. glauca*, followed the same pattern (L. Laplace, C. Guan, and K. Pawlowski, unpublished data). This resembles the GS expression pattern in legume nodules, except that, in *C. glauca*, expression takes place in both infected and uninfected cells of the central tissue (Forde et al. 1989; Miao et al. 1991; Temple et al. 1995).

In contrast, GS expression in *D. glomerata* nodules seems to be confined to the uninfected cortical cells surrounding the patches of infected cells. Furthermore, in *A. glutinosa*/*C. glauca* and legume nodules, GS expression is not detectable in senescent infected cells (K. Pawlowski, C. Guan, L. Laplace, and T. Bisseling, unpublished observation), but in *Datisca* nodules, GS is expressed at high levels in infected cortical cells where the microsymbiont is degraded. Thus, the GS expression pattern in actinorhizal nodules of *Datisca* is strikingly different from that of actinorhizal nodules of *Alnus*/*Casuarina* and that of legume nodules. Although the physiological basis underlying this novel expression pattern is the subject of another study (A. M. Berry, K. Pawlowski, T. M. Murphy, P. A. Okubara, and K. R. Jacobsen, unpublished data), the most immediate conclusion regarding the tissue specialization for nitrogen metabolism is that actinorhizal nodules from different phylogenetic groups can show greater differences than between actinorhizal and legume nodules.

Comparison of root nodule-forming plants from different groups of Eurosid I.

The results outlined above show that, regarding the recruitment of PR10 protein gene family member for particular functions, expression patterns of homologous genes during the differentiation of nodule cortical cells, and tissue specialization in nitrogen metabolism, differences between actinorhizal nodules of groups B and E exceed the differences between actinorhizal nodules of group B and legume nodules (group A). Actinorhizal nodules of *A. glutinosa* (group B)

and *D. glomerata* (group E) were different in all aspects examined. In combination with the known morphological differences between nodules formed by these plants (Berry and Sunell 1990; Hafeez et al. 1984), the data presented here support the hypothesis of separate independent origins of the symbiosis in intracellularly infected actinorhizal plants (group B; *Alnus*) and *Datisca*/*Coriaria* (group E). These analyses also suggest that the actinorhizal plants of group C/D may be more closely related to legumes than was assumed previously, or than is supported by phylogenetic analysis. No differences between gene expression patterns were observed in nodules of actinorhizal plants from the same phylogenetic group, either by comparison of GS expression patterns between *A. glutinosa* and *C. glauca*, or in previous studies (*ag12/cg12*) (Laplace et al. 2000; Ribeiro et al. 1995).

The phylogenetic position of genus *Parasponia* that interacts with rhizobia, among actinorhizal members of families Rhamnaceae and Elaeagnaceae (Swensen 1996), indicates that microsymbiont specificity may not have been part of the early evolution of the symbiotic syndrome. Studies on hemoglobin genes expressed in nodules (Hunt et al. 2001; Jacobsen-Lyon et al. 1995; Landsmann et al. 1986) showed that, in *P. andersonii* (group C/D), a hemoglobin gene was recruited for infected nodule cells from a different family than in legumes (group A) and *C. glauca* (group B), and that, although symbiotic hemoglobin/leghemoglobin gene regulation mechanisms are conserved among legumes and *C. glauca*, they are not conserved among legumes and *P. andersonii*. These results also suggest that legumes and intracellularly infected actinorhizal plants are more closely related than legumes and actinorhizal plants (including *Parasponia* spp.) from group C/D, respectively. The PR10 gene expression patterns presented in this study, as well as the comparison of GS expression patterns, suggest that actinorhizal plants of group B are more closely related to legumes than to actinorhizal plants of group D. Based on these analyses, the following series in the evolution of nitrogen-fixing root nodule symbioses appears plausible: first legumes (group A), then intracellularly infected actinorhizal plants (group B), then intercellularly infected species of Elaeagnaceae, Rhamnaceae, *Parasponia*, and Rosaceae (group C/D; phylogenetic analysis indicates that there were three independent origins of the symbiosis in this branch, but does not yield informa-

Fig. 4. Gene expression patterns in nodules detected by in situ hybridization. **A, B, C, E, H, K, N, and O,** Bright-field photos; silver grains denoting hybridization are visible as black dots. **D, F, G, I, J, L, and M,** Dark-field photos; silver grains appear as white dots. **O,** Photographed under fluorescent light. The aromatic compounds in secondary cell wall modifications show blue fluorescence. **A–D,** Series of adjacent longitudinal sections of *Datisca* nodules were hybridized with **A,** *Frankia nifH* to identify the nitrogen-fixing infected cells, and with **B,** *dgPR10A* (*dg241*) and **C and D,** *dgGHRP1* (*dg200*). **A,** Open arrow points at the nodule meristem; v, vascular system. Bar = 250 µm. **A and B,** *dgPR10A* (*dg241*) is expressed in all infected cells. Expression starts upon infection (broad arrows; as confirmed by hybridization of parallel sections with a *Frankia* rRNA probe; data not shown), declines as the infection process reaches completion, and reaches a high level again when nitrogen-fixation starts in the infected cells (arrow). **C and D,** *dgGHRP1* (*dg200*) is expressed in the nodule periderm (open arrow) and in certain cortical cells (arrows) located between the meristem and the infected cortical cells that express *dgPR10A* (*dg241*) in (B). **E–G,** Expression of GS in *D. glomerata* nodules: a heterologous probe (*Alnus glutinosa* GS) was used in order to detect the expression of all members of the *D. glomerata* GS gene family. Hybridization is detected mainly in the uninfected cortical cells. The infected cortical cells (**E**, stained red) show very little, if any, hybridization, except for senescent infected cells which are characterized by the appearance of small lytic vacuoles (**E**, arrow). **G,** Negative control for E and F: An adjacent section was hybridized with a sense RNA probe. Bar = 250 µm. **H and I,** Expression of *RH2* in a pea nodule (15 days after infection). Signal is confined to the nodule cortex (arrows). No signal is found in the infected cells. Bar = 250 µm. **J and K,** Expression of *dgMLP1* in sections of *D. glomerata* nodules. Expression is found in the cortical cell layers directly below the nodule periderm, at the infected side of the nodule lobe (**J**, black arrow). Some uninfected cortical cells of the nodule also express *dgMLP1* (**K**, arrows). **J,** Bar = 250 µm; **K,** bar = 50 µm. **L and M,** Adjacent longitudinal sections of a *D. glomerata* nodule lobe were hybridized with **L,** *dgGHRP1* (*dg200*) and **M,** a *Frankia* rRNA probe. White arrows point at corresponding positions in both sections between **L,** cortical *dgGHRP1* expression and **M,** *Frankia* rRNA hybridization. Cortical cells expressing *dgGHRP1* are not infected by *Frankia*. Bar = 250 µm. **N and O,** Expression of *dgGHRP1* (*dg200*) in cross sections of roots and nodule roots from *D. glomerata*. Expression is found in the suberized pericycle of roots (thin arrows) and in the nodule periderm (thick white arrow), but where it is interrupted by lenticels (le). **O,** One root was found where both pericycle (pc) and periderm (pd) and even part of the cortical cells were suberized. In this root, all suberized cells expressed *dgGHRP1*. Thin white arrows point at the xylem element of the vascular tissue which show blue cell wall fluorescence due to lignification and do not express *dgGHRP1*. **N,** Bar = 250 µm; **O,** bar = 50 µm.

tion as to which group was basal), and last, species of families Datisceae and Coriariaceae (group E).

Phylogenetic analyses presented here, as well as that of Soltis and associates (2000), indicate that, within the Eurosid I clade, legumes are basal to the actinorhizal plants and *Parasponia* spp. (Fig. 1). However, in both cases, this conclusion lacked support from parsimony jackknife analysis. It was shown clearly that groups C and D as well as *Parasponia* spp. are more closely related to each other than to other actinorhizal groups or legumes, but the analysis could not provide information on whether there is a closer relationship between groups A (legumes) and B (intracellularly infected actinorhizal plants, including *Alnus* spp.) than between groups A and C/D (Rhamnaceae, Elaeagnaceae, and Rosaceae), as would be predicted based on the hemoglobin data. It also is unclear whether actinorhizal plants of group E (Datisceae and Coriariaceae) are less closely related to legumes than to actinorhizal plants of group B (e.g., *Alnus* spp.). Even more extensive phylogenetic analyses of Eurosid I species (Soltis et al. 2000) are unable to provide well-supported ideas of relationships among major groups. Thus, data from the comparative molecular analysis of different nodule types helps inform our view of the evolution of nitrogen-fixing symbioses.

Obviously, significant differences can exist between symbioses from different branches of the Eurosid I clade, involving variation in nodule tissue specialization and gene expression during the infection process. Thus, plants seem to have multiple strategies to meet the requirements imposed by an intracellular nitrogen-fixing symbiont. A comparison of these strategies appears to be useful in understanding the evolutionary history of plant root symbioses. Further studies will help to characterize and compare developmental events from different taxonomic groups and, possibly, to identify features that are necessary to allow the symbiosis to exist.

MATERIALS AND METHODS

Phylogenetic methods: *rbcL* analysis.

New phylogenetic analyses were performed based on published *rbcL* sequences of Eurosid I (Angiosperm Phylogeny Group 1998) species. This analytical approach was similar to that conducted by Soltis and associates (1995), except that additional Eurosid I species were included here in an effort to find more resolution among nitrogen-fixing host plants. In this phylogenetic analysis of 107 *rbcL* sequences, trees were found using three successive rounds of parsimony searches implemented in PAUP*4.0 (beta test versions d53–d55) (Swofford 2002). The first round was used only to generate a minimum tree length to be used as a limit on the subsequent search (no trees were saved from this search). This initial search consisted of 10 replications of a heuristic search using NNI branch swapping. The second search consisted of 780 replications of a heuristic search using tree bisection-reconnection (TBR) branch swapping and random addition taxon entry and saving only two trees greater than or equal to the minimum-length tree specified from search 1. By saving only two trees per replication, the search sampled more possible topologies without being detained by swapping on significantly suboptimal trees. In all, 719 trees were obtained in this second search and subsequently were used as starting trees for the third search, which consisted of TBR branch swapping on the starting trees. During the third search, all most-parsimonious trees were saved. The consensus trees were constructed in PAUP*4.0 by summarizing all the branching patterns found in the 72 trees obtained through this search strategy.

For this analysis, we designated 12 taxa as outgroup species based on results from previous analyses (Soltis et al. 1995).

The outgroup comprised the genera *Geranium*, *Francoa*, *Crossosoma*, *Dillenia*, *Vitis*, *Lythrum*, *Clarkia*, *Ailanthus*, *Acer*, *Qualea*, *Gossypium*, and *Capparis*. Due to the length of time required by this initial analysis (550 h), typical estimates of internal branch support (e.g., via bootstrap or decay analysis) were not possible. Instead, the *rbcL* data set was analyzed by parsimony jackknife analysis (Farris et al. 1996). This analysis was conducted by S. Farris using the *Jac* program. In all, 1,000 jackknife replications were conducted; each replication included TBR branch swapping on trees from each of 10 random taxon addition sequences. Parsimony jackknifing was selected specifically for estimating branch support because of the significant decrease in computational time associated with this program. Branches with support $\geq 50\%$ were identified by this analysis; these support values were transferred onto the strict consensus tree obtained with PAUP*4.0.

Phylogenetic methods: protein families.

Phylogenetic analyses of PR10 and glycine- and histidine-rich proteins was carried out using PAUP*4.0 following alignment with ClustalX1.81 (a graphical interface for ClustalW) (Thompson et al. 1994). Alignments were carried out using the Gonnet 250 protein weight matrix; pairwise alignments used gap opening and gap extension penalties of 10.0 and 0.1, respectively; multiple alignments used gap opening and gap extension penalties of 10.0 and 0.2, respectively. Phylogenetic trees were obtained in PAUP*4.0 using 1,000 repetitions of a heuristic search with random-taxon entry and TBR branch swapping. To assess branch support, bootstrap analyses (Felsenstein 1985) were carried out by analyzing 100 bootstrapped datasets using the same approach as with the initial analyses. Thirty-six PR10 protein sequences, obtained from GenBank, were aligned with the *D. glomerata* and *A. glutinosa* PR10 protein sequences. *L. japonicus*, *M. truncatula*, and *D. glomerata* MLP1 sequences were designated as the outgroup. Sixteen glycine- and histidine-rich protein sequences, obtained from GenBank, were aligned with the *A. glutinosa* and *D. glomerata* glycine- and histidine-rich protein sequences. A glycine-rich RNA-binding protein from *Brassica napus* (GenBank accession number Q05966) was designated the outgroup because it belongs to another protein family.

Plant material.

D. glomerata seed were collected at a local source near Lake Berryessa, CA, USA. Plants were grown in a peat:fir bark:sand (1:1:1) mix and watered with one-quarter-strength Hoagland solution (Hoagland and Arnon 1938). Leaves, flowers, and developing fruit were harvested from uninfected plants. For nodulation, plants were watered with $1/4$ Hoagland solution without nitrogen starting 2 weeks before infection. Crushed nodules from aeroponically cultivated *Ceanothus griseus* stock plants were used for infection. *Datisca* nodules were harvested 4 to 8 weeks after inoculation. Root tips (1 to 2 cm) were harvested from *D. glomerata* plants grown hydroponically in one-quarter-strength Hoagland solution. Growth conditions for *A. glutinosa* have been described elsewhere (Ribeiro et al. 1995).

Molecular biological methods.

Isolation of RNA from *A. glutinosa* plants has been described previously (Ribeiro et al. 1995). RNA was isolated from *D. glomerata* material as described by Pawlowski and associates (1994) for legume nodules. A directional cDNA library was constructed from *D. glomerata* nodule mRNA using the kit from Stratagene (La Jolla, CA, U.S.A.) according to the instructions of the manufacturer, and 480 in vivo-excised clones were randomly selected for DNA isolation and DNA gel blot hybridization with ^{32}P -labeled nodule and

root cDNA, respectively (Okubara et al. 1999). Clones whose inserts were not or only weakly hybridizing with root but strongly hybridizing with nodule cDNA were selected for further analysis. The initial cDNA fragment of *dgMLP1* was amplified from nodule first-strand cDNA prepared using 3 µg of total RNA and primer 5'-CTCGAGGATCCGCGGCCGCTTTTTTTTTTTTTTTTTT-3'. Primers 5'-CCGGCAAGT-CACIAAAIGGCAATG-3' and 5'-CTCGAGGATCCGCGGCCG-3' were used for the amplification. The product was cloned in pGEM-T (Promega Corp., Madison, WI, U.S.A.). To obtain the full-size cDNA sequence, 5'-RACE was performed. The Thermoscript RT-PCR System Kit (Life Technologies, Gaithersburg, MD, U.S.A.) was used for reverse transcription of approximately 5 µg of total nodule RNA with the first gene-specific primer (5'-TTCACACACAGCCA-CAGC-3') at 55°C for 60 min. Nucleotides were removed using spin columns from the Concert TM Rapid PCR Purification System (Life Technologies). DNA was eluted from the column in 50 µl of terminal transferase buffer (Promega Corp.) and C-tailing was performed using 15 U of terminal transferase (Promega Corp.). A 1/10 volume of the reaction mixture was used for PCR with primer 5'-GGCCACGCGTCTGACTAGTACGGGIGGGIIG-3' and the second gene-specific primer 5'-GAGCTCATAACAGTCATGAGC-3'. The PCR product was cloned in pGEM-T Easy (Promega Corp.) and sequenced. RT-PCR was used to obtain the full-length cDNA. First-strand cDNA was reverse transcribed from 3 µg of nodule mRNA using the primer 5'-CTCGAGGATCCGCGGCCGCTTTTTTTTTTTTTTTTTT-3'. Amplification of the first-strand cDNA was carried out using the primers 5'-CACAAGAGAATTTCTTAGTGA-3' and 5'-CTCGAGGATCCGCGGCCG-3'. The PCR product was cloned in pGEM-T Easy and sequenced.

Sequencing was performed using an automatic sequencer (Model 373A; Applied Biosystems, Foster City, MA, U.S.A.). Sequence data were analyzed using the programs of the Wisconsin Genetics Computer Group (Devereux et al. 1984). Database searches were performed using the BLAST algorithm (Altschul et al. 1990) in the nucleotide sequence data bases of the National Center of Biotechnology Information, National Library of Medicine, National Institute of Health, in Bethesda, MD, U.S.A.

The preparation and hybridization of RNA gel blots has been described previously (Ribeiro et al. 1995). For *dgGHRP1* (*dg200*) RNA probes for in situ hybridization, the cDNA clone was linearized using *Xba*I and transcribed with T7 RNA polymerase (antisense), or linearized using *Xho*I and transcribed with T3 RNA polymerase (sense), respectively. For the *dgPRI0A* (*dg241*) RNA probes, a partial cDNA clone containing nucleotides 213 to 763 was linearized using *Xba*I and transcribed with T7 RNA polymerase (antisense), or linearized using *Xho*I and transcribed with T3 RNA polymerase (sense), respectively. The pea *RH2* RNA probes were prepared as described by Mylona and associates (1994), the *Alnus* GS RNA probes as described by Guan and associates (1996), and the *Frankia nifH* antisense RNA probe as described by Ribeiro and associates (1995). For the *dgMLP1* probes, a partial cDNA clone containing nucleotides 346 to 763 was linearized using *Sal*I and transcribed with T7 RNA polymerase (antisense), or linearized using *Sst*II and transcribed with SP6 RNA polymerase (sense), respectively. For the *Frankia* antisense 23S rRNA probe, clone pLL27-1, representing base pairs 3,588 to 3,942 subcloned in pGEM-T (Promega Corp.), was linearized using *Nco*I and transcribed with SP6 RNA polymerase. Fixation, embedding, sectioning of plant material, and the hybridization procedure were as described by Van de Wiel and associates (1990).

ACKNOWLEDGMENTS

Parts of the research were supported by a Katherine Esau Fellowship from the University of California, Davis, to K. Pawlowski and by California Experimental Station Project California Agricultural Experiment Station Project CA-AES-6258 (AMB). K. Pawlowski and C. Guan acknowledge support by the Netherlands' Foundation for Scientific Research (NWO). S. Swenson acknowledges support from DEB-9408347 and DEB-9815816. We thank T. van Kampen for DNA sequencing, L. Laplace (IRD-GeneTrop, Montpellier, France) for kindly providing the *Frankia* rRNA clone pLL27-1, D. Neale (United States Forest Service, PSW, Albany, CA) for facilities and advice during cDNA library constructions, and S. Farris for conducting parsimony jackknife analyses.

ACCESSIONS

Species and corresponding accession numbers are as follows. Outgroup species: Aceraceae (*Acer saccharum* L., L01881), Capparaceae (*Capparis hastata* Jacq., M95755), Crossosomataceae (*Crossosoma californicum* Nutt., L11179), Dilleniaceae (*Dillenia indica* L., L01903), Francoaceae (*Francoa sonchifolia* Cav., L11184), Geraniaceae (*Geranium cinerium* Cav., L14695), Lythraceae (*Lythrum hyssopifolia* L., L10218), Malvaceae (*Gossypium robinsonii* F. Muell., L13186), Onagraceae (*Clarkia xantiana* A. Gray, L10896), Simaroubaceae (*Ailanthus altissima* L., L12566), Vitaceae (*Vitis aestivalis* Michx., L01960), Vochysiaceae (*Qualea* sp., U02730). Ingroup species: Ingroup "A": Ochnaceae (*Ochna serrulata* Walp., Z75273), Passifloraceae (*Passiflora quadrangularis* L., L01940), Humeriaceae (*Humiria balsamifera* Aubl., L01926), Trigonaceae (*Trigonostemon nivea* Cambess., AF206830), Chrysobalanaceae (*Licania tomentosa* Fritsch, L11193), Violaceae (*Viola sororia* Willd., L11674), Euphorbiaceae (*Drypetes roxburghii* (Wall.) Hurus., M95757), Saxifragaceae (*Lepuropetalon spathulatum* (Muhl.) Elliot, L11192). Ingroup "B": Eucryphiaceae (*Eucryphia lucida* Druce, L01918), Cephalotaceae (*Cephalotus follicularis* Labill., L01894), Cunoniaceae (*Bauera rubioides* Andrews, L11174), Cunoniaceae (*Ceratopetalum gumiferum* Small, L01895), Begoniaceae (4) (*Hillebrandia sandwicensis* Oliver, U59822; *Begonia herbacea* Vell., U59816; *Begonia oxyloba* Welw. ex Hook. f., U59815; *Begonia ulmifolia* Willd., U59822), Betulaceae (5) (*Carpinus caroliniana* Walt., X56621; *Alnus incana* Moench, X56618; *Betula nigra* L., L01889; *Corylus cornuta* Marsh., X56619; *Ostrya virginiana* C.Koch, X56620), Cucurbitaceae (5) (*Neolalsomitra sarcophylla* (Wall.) Hutchinson, U59823; *Marah macrocarpa* Greene, U59825; *Xerosicyos decaryi* Guillaumin & Keraudr., U59824; *Cucumis sativus* L., L21937; *Cucurbita pepo* L., L21938), Coriariaceae (2) (*Coriaria arborea* Lindsay, U59829; *Coriaria myrtifolia* L., L01897), Cannabaceae (2) (*Cannabis sativa* L., AJ390068; *Humulus lupulus* L., U2729), Corynocarpaceae (*Corynocarpus laevigatus* Forst., X69731), Datisceae (2) (*Datisca glomerata* Baill., L21940; *Datisca cannabina* L., L21939), Elaeagnaceae (3) (*Elaeagnus angustifolia* Blanco, U01738; *Hippophae salicifolia* D.Don, U59821; *Shepherdia canadensis* (Pursh.) Nutt., U17039), Fabaceae (12) (*Parkia roxburghii* G. Don, U74209; *Erythrophleum ivorense* A. Chev., U74205; *Glycine tabacina* (Labill.) Benth., U74197; *Cadja purpurea* (G. Piccoli) Aiton, U74192; *Macrolobium acaciifolium* (Benth.) Benth., U74191; *Albizia julibrissin* Durazz., AF206730; *Caesalpinia pulcherrima* (L.) Sw., U74190; *Xanthocercis zambeziaca* (Baker) Dumas-le-Grand, U74189; *Cercis canadensis* L., U74188; *Medicago sativa* L., X04975; *Pisum sativum* L., X03853; *Bauhinia* sp., AF206739), Fagaceae (6) (*Castanea sativa* Mill., M94936; *Quercus rubra* L., M58931; *Trigonobalanus verticellata* Forman, AJ235812; *Chrysopsis sempervirens* (Kellogg) Hjelmq., U02727; *Fagus sylvatica* L., L13340; *Nothofagus dombeya* Blume, L13350), Juglandaceae (2) (*Juglans nigra* L., U00437; *Carya glabra* Sweet, L12637), Moraceae (2) (*Ficus trigonata* L., AF206767; *Morus alba* L., L01933), Myricaceae (2) (*Myrica cerifera* L., U01934; *Comptonia peregrina* Coulter, X69529), Polygalaceae (*Securidaca diversifolia* (L.) S.F.Blake, L01955), Rhamnaceae (5) (*Colletia ulicina* Gill. & Hook., U59819; *Discaria chacaye* (G. Don) Tortosa, U59826; *Ceanothus thyrsiflorus* Esch., U59827; *Ziziphus glabrata* Heyne ex. Roth, U60313; *Rhamnus cathartica* L., L13189; *Trevoa trinervia* Gill. & Hook., U59828), Rosaceae (12) (Rosoidae 1 (3): *Geum chiloense* Balb. ex Ser., L01921; *Rosa woodsii* A.Gray, U06824; *Rubus idaeus* L., U06825; Rosoidae 2 (4): *Cowaniana stansburiana* Torr., U59817; *Cercocarpus ledifolius* Nutt., U06796; *Purshia tridentata* DC., U06821; *Dryas drummondii* Richards, U59818; Maloideae (2): *Photinia fraseri* Dress, L11200; *Crataegus columbiana* Howell, U06799; Other Rosaceae (3): *Prunus virginiana* L., L01947; *Neillia sinensis* Oliver, U06813; *Spiraea vanhouttei* (Briot) Zabel, L11206), Suraniaceae (*Cadellia pentastylis* F. Muell., L29491), Tetrameleaceae (2) (*Tetrameles nudiflora* R.Br., L21943; *Octomeles sumatrana*

Miq., L21942), Tiodendraceae (*Ticodendron incognitum* J. Gomez-Lauri & L. D. Gomez P., AF061197), Ulmaceae (5) (*Celtis yunnanensis* C. K. Schneid., L12638; *Zelkova serrata* Makino, AF6835, *Ulmus alata* Michx., U00441; *Barbeya oleoides* Schweinf., U60314; *Parasponia rigida* Merrill & Perry, U59820), Urticaceae (2) (*Pilea pumila* (L.) Gray, U00438; *Boehmeria nivea* Gaudich., AF2062005), Casuarinaceae (3) (*Casuarina litorea* Salisb., 01893; *Allocasuarina japonicum*, X69527; *Gymnostoma webbianum* (Miq.) L.A.S.Johnson; X69531).

LITERATURE CITED

- Aggelis, A., John, I., Karvouni, Z., and Grierson, D. 1997. Characterization of two cDNA clones for mRNAs expressed during ripening of melon (*Cucumis melo* L.) fruits. *Plant Mol. Biol.* 33:313-322.
- Altschul, S. F., Gish, W., Miller, W., Myers, E. W., and Lipman, D. J. 1990. Basic local alignment search tool. *J. Mol. Biol.* 215:403-410.
- Andersson, C. R., Llewellyn, D. J., Peacock, W. J., and Dennis, E. S. 1997. Cell-specific expression of the promoters of two nonlegume hemoglobin genes in a transgenic legume, *Lotus corniculatus*. *Plant Physiol.* 113:45-57.
- Angiosperm Phylogeny Group 1998. An ordinal classification for the families of flowering plants. *Ann. Mol. Bot. Gard.* 85:531-553.
- Benson, D. R., and Silvester, W. B. 1993. Biology of *Frankia* strains, actinomycete symbionts of actinorhizal plants. *Microbiol. Rev.* 57:293-319.
- Berry, A. M., and Sunell, L. A. 1990. The infection process and nodule development. Pages 61-81 in: *The Biology of Frankia and Actinorhizal Plants*. C. R. Schwintzer and J. D. Tjepkema, eds. Academic Press, New York.
- Berry, A. M., Pawlowski, K., Murphy, T. M., Okubara, P. A., and Jacobsen, K. R. Expression of cytosolic glutamine synthetase in root nodules of the actinorhizal host, *Datisca glomerata*. Submitted for publication.
- Breiteneder, H., Ferreira, F., Reikerstorfer, A., Duchene, M., Valenta, R., Hoffmann-Sommergruber, K., Ebner, C., Breitenbach, M., Kraft, D., and Scheiner, O. 1992. Complementary DNA cloning and expression in *Escherichia coli* of Aln I, the major allergen in pollen of alder (*Alnus glutinosa*). *J. Allergy Clin. Immunol.* 90:909-917.
- Breiteneder, H., Pettenburger, K., Bito, A., Valenta, R., Kraft, D., Rumpold, H., Scheiner, O., and Breitenbach, M. 1989. The gene coding for the major birch pollen allergen Bet v 1 is highly homologous to a pea disease resistance response gene. *EMBO (Eur. Mol. Biol. Organ.) J.* 8:1935-1938.
- Bufe, A., Schramm, G., Keown, M. B., Schlaak, M., and Becker, W. M. 1995. Major allergen Phl p Vb in timothy grass is a novel pollen RNase. *FEBS (Fed. Eur. Biochem. Soc.) Lett.* 363:6-12.
- Bufe, A., Spangfort, M. D., Kahlert, H., Schlaak, M., and Becker, W. M. 1996. The major birch pollen allergen, Bet v 1, shows ribonuclease activity. *Planta* 199:413-415.
- Burgess, D., and Peterson, R. L. 1987. Development of *Alnus japonica* root nodules after inoculation with *Frankia* strain HFPAr13. *Can. J. Bot.* 65:1647-1657.
- Conklin, D. S., McMaster, J. A., Culbertson, M. R., and Kung, C. 1992. *COT1*, a gene involved in cobalt accumulation in *Saccharomyces cerevisiae*. *Mol. Cell Biol.* 12:3678-3788.
- Devereux, J., Haerberli, P., and Smithies, O. 1984. A comprehensive set of sequence analysis programs for the VAX. *Nucleic Acids Res.* 12:387-195.
- Doyle, J. J., Doyle, J. L., Ballenger, J. A., Dickson, E. E., Kajita, T., and Ohashi, H. 1997. A phylogeny of the chloroplast gene *rbcl* in the leguminosae—taxonomic correlations and insights into the evolution of modulation. *Am. J. Bot.* 84:541-554.
- Edwards, J. W., Walker, E. L., and Coruzzi, G. M. 1990. Cell-specific expression in transgenic plants reveals nonoverlapping roles for chloroplast and cytosolic glutamine synthetase. *Proc. Natl. Acad. Sci. USA* 87:3459-3463.
- Farris, J. S., Albert, V. A., Kallersjo, M., Lipscomb, D., and Kluge, A. G. 1996. Parsimony jackknifing outperforms neighbor-joining. *Cladistics* 12:99-124.
- Felsenstein, J. 1985. Confidence limits on phylogenies: an approach using the bootstrap. *Evolution* 39:783-791.
- Forde, B. G., Day, H. M., Turton, J. F., Shen, W. J., Cullimore, J. V., and Oliver, J. E. 1989. Two glutamine synthetase genes from *Phaseolus vulgaris* L. display contrasting developmental and spatial patterns of expression in transgenic *Lotus corniculatus* plants. *Plant Cell* 1:391-401.
- Gamas, P., de Billy, F., and Truchet, G. 1998. Symbiosis-specific expression of two *Medicago truncatula* nodulin genes, *MtN1* and *MtN13*, encoding products homologous to plant defense proteins. *Mol. Plant-Microbe Interact.* 11:393-403.
- Guan, C., Ribeiro, A., Akkermans, A. D. L., Jing, Y., van Kammen, A., Bisseling, T., and Pawlowski, K. 1997. Nitrogen metabolism in actinorhizal nodules of *Alnus glutinosa*: Expression of glutamine synthetase and acetylornithine transaminase. *Plant Mol. Biol.* 32:1177-1184.
- Gupta, R. K., Dobritsa, S. V., Stiles, C. A., Essington, M. E., Liu, Z. Y., Chen, C. H., Serpersu, E. H., and Mullin, B. C. 2002. Metallothioneins: A new class of plant metal-binding proteins. *J. Prot. Chem.* 21:529-536.
- Hafeez, F., Akkermans, A. D. L., and Chaudhary, A. H. 1984. Observations on the ultrastructure of *Frankia* sp. in root nodules of *Datisca cannabina* L. *Plant Soil* 79:383-402.
- Hoagland, D. R., and Arnon, D. T. 1938. The water-culture method for growing plants without soil. *Calif. Agric. Exp. Stn. Circ.* 347. Univ. Calif. Berkeley.
- Hunt, P. W., Watts, R. A., Trevaskis, B., Llewellyn, D. J., Burnell, J., Dennis, E. S., and Peacock, W. J. 2001. Expression and evolution of functionally distinct hemoglobin genes in plants. *Plant Mol. Biol.* 47:677-692.
- Itturiaga, E. A., Leech, M. J., Barratt, D. H., and Wang, T. L. 1994. Two ABA-responsive proteins from pea (*Pisum sativum* L.) are closely related to intracellular pathogenesis-related proteins. *Plant Mol. Biol.* 24:235-240.
- Jacobsen-Lyon, K., Østergaard-Jensen, E., Jørgensen, J.-E., Marcker, K. A., Peacock, J., and Dennis, E. 1995. Symbiotic and nonsymbiotic hemoglobin genes of *Casuarina glauca*. *Plant Cell* 7:213-223.
- Keen, N. T. 1992. The molecular biology of disease resistance. *Plant Mol. Biol.* 19:109-122.
- Kouchi, H., and Hata, S. 1993. Isolation and characterization of novel nodulin cDNAs representing genes expressed at early stages of soybean nodule development. *Mol. Gen. Genet.* 238:106-119.
- Lamb, C. J., Lawton, M. A., Dron, M., and Dixon, R. A. 1989. Signals and transduction mechanisms for activation of plant defenses against microbial attack. *Cell* 56:215-224.
- Landsmann, J., Dennis, E. S., Higgins, T. J. V., Appleby, C. A., Kortt, A. A., and Peacock, W. J. 1986. Common evolutionary origin of legume and non-legume plant hemoglobins. *Nature* 324:166-168.
- Laplace, L., Ribeiro, A., Franche, C., Duhoux, E., Auguy, F., Bogusz, D., and Pawlowski, K. 2000. Characterization of a *Casuarina glauca* nodule-specific subtilisin-like protease gene, a homolog of *Alnus glutinosa agl2*. *Mol. Plant-Microbe Interact.* 13:113-117.
- Luttkopf, D., Müller, U., Skov, P. S., Ballmer-Weber, B. K., Wuthrich, B., Skamstrup Hansen, K., Poulsen, L. K., Kastner, M., Haustein, D., and Vieths, S. 2002. Comparison of four variants of a major allergen in hazelnut (*Corylus avellana*) Cor a 1.04 with the major hazel pollen allergen Cor a 1.01. *Mol. Immunol.* 38:515-525.
- Miao, G. H., Hirel, B., Marsolier, M. C., Ridge, R. W., and Verma, D. P. 1991. Ammonia-regulated expression of a soybean gene encoding cytosolic glutamine synthetase in transgenic *Lotus corniculatus*. *Plant Cell* 3:11-22.
- Mylona, P., Moerman, M., Yang, W. C., Gloudemans, T., van de Kerckhove, J., van Kammen, A., Bisseling, T., and Franssen, H. J. 1994. The root epidermis-specific pea gene *RH2* is homologous to a pathogenesis-related gene. *Plant Mol. Biol.* 26:39-50.
- Nakai, K., and Horton, P. 1999. PSORT: a program for detecting sorting signals in proteins and predicting their subcellular localization. *Trends Biochem. Sci.* 24: 34-35.
- Newcomb, W., and Pankhurst, C. E. 1982. Fine structure of actinorhizal nodules of *Coriaria arborea* (Coriariaceae). *N. Z. J. Bot.* 20:93-103.
- Nicholas, K. B., Nicholas, H. B., Jr., and Deerfield, D. W., II. 1997. GeneDoc: Analysis and visualization of genetic variation. *EMBNETNEWS* 4:14. Online publication.
- Okubara, P., Pawlowski, K., Murphy, T. M., and Berry, A. M. 1999. Symbiotic root nodules of the actinorhizal plant *Datisca glomerata* express Rubisco activase mRNA. *Plant Physiol.* 120:411-420.
- Osmark, P., Boyle, B., and Brisson, N. 1998. Sequential and structural homology between intracellular pathogenesis-related proteins and a group of latex proteins. *Plant Mol. Biol.* 38:1243-1246.
- Park, C. J., Park, C. B., Hong, S. S., Lee, H. S., Lee, S. Y., and Kim, S. C. 2000. Characterization and cDNA cloning of two glycine- and histidine-rich antimicrobial peptides from the roots of shepherd's purse, *Capsella bursa-pastoris*. *Plant Mol. Biol.* 44:187-197.
- Parsons, R., Silvester, W. B., Harris, S., Gruijters, W. T. M., and Bullivant, S. 1987. *Frankia* vesicles provide inducible and absolute oxygen protection for nitrogenase. *Plant Physiol.* 83:728-731.
- Pawlowski, K., and Bisseling, T. 1996. Rhizobial and actinorhizal symbioses: What are the shared features? *Plant Cell* 8:1899-1913.
- Pawlowski, K., Kunze, R., De Vries, S., and Bisseling, T. 1994. Isolation of total, poly(A) and polysomal RNA from plant tissues. Pages 1-13 in: *Plant Molecular Biology Manual D5*, 2nd ed. S. B. Gelvin and R. A. Schilperoort, eds. Kluwer Academic Publishers, Dordrecht, The Netherlands.

- Pawlowski, K., Twigg, P., Dobritsa, S., Guan, C., and Mullin, B. C. 1997. A nodule-specific gene family from *Alnus glutinosa* encodes glycine- and histidine-rich proteins expressed in the early stages of actinorhizal nodule development. *Mol. Plant-Microbe Interact.* 10:656-664.
- Ribeiro, A., Akkermans, A. D. L., van Kammen, A., Bisseling, T., and Pawlowski, K. 1995. A nodule-specific gene encoding a subtilisin-like protease is expressed in early stages of actinorhizal nodule development. *Plant Cell* 7:785-794.
- Schubert, K. R. 1986. Products of biological nitrogen fixation in higher plants: Synthesis, transport, and metabolism. *Annu. Rev. Plant Physiol.* 37:539-574.
- Sikorski, M. M., Biesiadka, J., Kasperska, A. E., Kopcinska, J., Lotocka, B., Golinowski, W., and Legocki, A. B. 1999. Expression of genes encoding PR10 class pathogenesis-related proteins is inhibited in yellow lupine root nodules. *Plant Sci.* 149:125-137.
- Sikorski, M. M., Szlagowska, A. E., and Legocki, A. B. 1998. Structure of *Lupinus luteus* genes *Ypr10.1a* and *Ypr10.1b* encoding two homologs of pathogenesis-related proteins of PR10 class (accession nos. AF002277 and AF002278). *Plant Physiol.* 116:1192-1193.
- Soltis, D. E., Soltis, P. S., Morgan, D. R., Swensen, S., Mullin, B. C., Dowd, J. M., and Martin, P. G. 1995. Chloroplast gene sequence data suggest a single origin of the predisposition for symbiotic nitrogen fixation in angiosperms. *Proc. Natl. Acad. Sci. USA* 92:2647-2651.
- Soltis, D. E., Soltis, P. S., Chase, M. W., Mort, M. E., Albach, D. C., Zanis, M., Savlainen, V., Hahn, W. H., Hoot, S. B., Fay, M. F., Axtell, M., Swensen, S. M., Nixon, K. C., and Farris, J. S. 2000. Angiosperm phylogeny inferred from a combined data set of 18S rDNA, *rbcL*, and *atpB* sequences. *Bot. J. Linn. Soc.* 133:381-461.
- Swensen, S. M. 1996. The evolution of actinorhizal symbioses - evidence for multiple origins of the symbiotic association. *Am. J. Bot.* 83:1503-1512.
- Swoboda, I., Hoffmann-Sommergruber, K., O'Riordain, B., Scheiner, O., Heberle-Bors, E., and Vicente, O. 1996. Bet v 1 proteins, the major birch pollen allergens and members of a family of conserved pathogenesis-related proteins, show ribonuclease activity in vitro. *Physiol. Plant.* 96:433-438.
- Swoboda, I., Scheiner, O., Heberle-Bors, E., and Vicente, O. 1995. cDNA cloning and characterization of three genes in the *Bet v 1* gene family that encode pathogenesis-related proteins. *Plant Cell Environ.* 18:765-764.
- Swofford, D. L. 2002. PAUP*. Phylogenetic Analysis Using Parsimony (*and Other Methods). Version 4. Sinauer Associates, Sunderland, MA.
- Temple, S. J., Heard, J., Ganter, G., Dunn K., and Sengupta-Gopalan, C. 1995. Characterization of a nodule-enhanced glutamine synthetase from alfalfa: nucleotide sequence, in situ localization, and transcript analysis. *Mol. Plant-Microbe Interact.* 8:218-227.
- Thompson, J. D., Higgins, D. G., and Gibson, T. J. 1994. CLUSTAL W: improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position specific gap penalties and weight matrix choice. *Nucleic Acids Res.* 22:4673-80.
- Van de Wiel, C., Scheres, B., Franssen, H., van Lierop, M. J., van Lammeren, A., van Kammen, A., and Bisseling, T. 1990. The early nodulin transcript ENOD2 is located in the nodule parenchyma (inner cortex) of pea and soybean root nodules. *EMBO (Eur. Mol. Biol. Organ.) J.* 9:1-7.
- Walter, M. H., Liu, J.-W., Grand, C., Lamb, C. J., and Hess, D. 1990. Bean pathogenesis-related (PR) proteins deduced from elicitor-induced transcripts are members of a ubiquitous new class of conserved PR proteins including pollen allergens. *Mol. Gen. Genet.* 222:353-360.
- Warner, S. A. J., Scott, R., and Draper, J. 1992. Characterization of a wound-induced transcript from the monocot asparagus that shares similarity with a class of intracellular pathogenesis-related (PR) proteins. *Plant Mol. Biol.* 19:555-561.