

Rapid report

Expressed sequence-tag analysis in *Casuarina glauca* actinorhizal nodule and root

Author for correspondence:

Valérie Hocher

Tel: +33-467-41-61-96

Fax: +33-467-41-62-22

Email: hocher@mpl.ird.fr

Received: 14 October 2005

Accepted: 24 November 2005

Valérie Hocher, Florence Auguy, Xavier Argout*, Laurent Laplaze, Claudine Franche and Didier Bogusz

UMR 1098, Institut de Recherche pour le Développement (IRD), BP 64501, 911 avenue Agropolis,

34394 Montpellier cedex 5, France; *Current address: Centre de Coopération Internationale en

Recherche Agronomique pour le Développement (CIRAD), Avenue Agropolis, 34398 Montpellier, France

Summary

Key words: actinorhizal symbiosis, *Casuarina glauca*, *Frankia*, expressed sequence tags (ESTs), quantitative real time RT-PCR (qRT-PCR).

- The present study aimed to identify and assess the frequency and tissue specificity of plant genes in the actinorhizal *Casuarina glauca*-*Frankia* symbiosis through expressed sequence tag (EST) analysis.
- Using a custom analysis pipeline for raw sequences of *C. glauca* uninfected roots and nodules, we obtained an EST databank web interface. Gene expression was studied in nodules vs roots using comparative quantitative real-time reverse transcription-polymerase chain reaction (qRT-PCR).
- From roots and nodules, 2028 ESTs were created and clustered in 242 contigs and 1429 singletons, giving a total of 1616 unique genes. Half the nodule transcripts showed no similarity to previously identified genes. Genes of primary metabolism, protein synthesis, cell division and defence were highly represented in the nodule library. Differential expression was observed between roots and nodules for several genes linked to primary metabolism and flavonoid biosynthesis.
- This comparative EST-based study provides the first picture of the set of genes expressed during actinorhizal symbiosis. We consider our database to be a flexible tool that can be used for the management of EST data from other actinorhizal symbioses.

New Phytologist (2006) **169**: 681–688

© *New Phytologist* (2006) No claim to French government works

doi: 10.1111/j.1469-8137.2005.01644.x

Introduction

Two nitrogen-fixing root-nodule endosymbioses between soil bacteria and plants have been described: one between *Rhizobium* and legumes; the other between *Frankia* and actinorhizal plants. The *Rhizobium*-legume symbiosis involves 20 000 plant species of the *Leguminosae* (*Fabeaceae*) family, while actinorhizal plants comprise approx. 260 species belonging to eight angiosperm families. Unlike legume nodules, actinorhizal nodules are

structurally and developmentally similar to lateral roots, with a central vascular bundle and peripheral infected cortical tissue (Pawlowski & Bisseling, 1996; Franche *et al.*, 1998). Mature actinorhizal nodules are indeterminate multilobed structures. Phylogenetic analysis using *rbcl* chloroplast gene sequences showed that legumes and actinorhizal plants belong to the Rosid I clade, suggesting that a genetic predisposition to form root-nodule symbioses originated in a common ancestor (Soltis *et al.*, 1995; Doyle, 1998). Furthermore, legume-rhizobial

and –actinorhizal symbioses are suggested to have originated from the more ancient arbuscular mycorrhizal symbioses (Duc *et al.*, 1989; Kistner & Parniske, 2002).

The mechanisms of the symbiotic association between *Frankia* and actinorhizal plants are still poorly understood. The molecular understanding of regulatory events in actinorhizal symbiosis is mainly limited by the lack of a genetic transformation system for the microsymbiont *Frankia*. However, the current sequencing of three *Frankia* genomes should allow the identification of genes homologous to *Rhizobium* symbiotic genes and of global changes in *Frankia* gene expression within the context of actinorhizal nodule development (P. Normand, pers. comm.). In the past decade, some progress has been made in understanding the plant genes that are expressed at different stages of actinorhizal nodule differentiation. Differential screening of nodule cDNA libraries with root and nodule cDNAs has resulted in the isolation of a number of nodule-specific or nodule-enhanced plant genes in several actinorhizal plants including *Alnus*, *Datisca*, *Eleagnus* and *Casuarina* (Laplaze *et al.*, 2006).

While the construction of large-scale expressed sequence tag (EST) databases led to the description of the global gene expression patterns of the host plant in *Rhizobium*–legume and mycorrhizal symbioses (Shoemaker *et al.*, 2001; Asamizu *et al.*, 2004; Grunwald *et al.*, 2004; Cannon *et al.*, 2005; Duplessis *et al.*, 2005), no EST resource has been reported for actinorhizal plants.

Our group is currently working on the *Frankia*–*Casuarinaceae* symbiosis as a model system to study actinorhizal nodule development, as *Casuarina glauca* and its close relative *Allocasuarina verticillata* can be transformed genetically using *Agrobacterium tumefaciens* (Franche *et al.*, 1998).

In this study we report the *in silico* analyses and comparison of ESTs from nonnodulated roots and nodules of *C. glauca*. In addition, several ESTs were selected for quantitative real-time reverse transcription–polymerase chain reaction (qRT–PCR) analysis to examine changes in gene expression induced by the symbiotic interaction with *Frankia*.

Materials and Methods

Plant and bacterial growth conditions

Casuarina glauca Sieb. ex Spreng seeds were provided by the Desert Development Center (Cairo, Egypt). Plants were grown and inoculated with *Frankia* Thr strains (Girgis *et al.*, 1990) as previously described (Gherbi *et al.*, 1997). Nodules and non-inoculated roots (controls) were harvested 3 wk after inoculation.

cDNA library construction and sequencing

Total RNA was extracted from *C. glauca* noninoculated roots and nodules using the Invisorb Spin Plant-RNA Mini kit (Invitek, San Carlos, CA, USA) according to the manufacturer's

instructions. Poly(A⁺) RNA was obtained from total RNA using the mRNA purification kit (Amersham Pharmacia, Freiburg, Germany) for roots and the Oligotex mRNA Spin Column (Qiagen, Hilden, Germany) for nodules. Root cDNA was prepared using the Zap–cDNA synthesis and Zap–cDNA gigapackIII gold cloning kit (Stratagene, La Jolla, CA, USA); nodule cDNA was generated using the Smart cDNA library construction kit (Clontech, Palo Alto, CA, USA).

2878 clones were randomly selected from both libraries and processed through the robotic and genomic Languedoc-Roussillon Génopole platform (<http://www.genopole-montpellier-lr.org/PF/seq/index.htm>). Single-pass sequencing from the 5' end was done using the universal T3 primer for roots and the 5' Lambda Triplex2 sequencing primer (5'-CTCCGAGATCT-GGACGAGC-3') for nodules.

Sequence processing and EST database construction

Each set of sequence data (root and nodule) was first processed individually using a multimodule custom pipeline which linked sequence backup, base calling, elimination of sequences shorter than 50 bp and low-quality sequences, vector trimming and sequence assembling, as described by Jouannic *et al.* (2005). To assign functions, the valid ESTs and the assembled consensus sequences were locally aligned using BLASTALL (NCBI, <ftp://ftp.ncbi.nlm.nih.gov/blast>) to accessions in a local nonredundant protein sequence database with entries from GenPept, Swissprot, PIR, PDF, PDB and NCBI RefSeq, using the BLASTX algorithm with an *E* value cut-off at 10^{-5} . If the EST sequences did not match any database sequences, the BLASTN algorithm was used in conjunction with a nucleotide sequence database, with entries from all traditional divisions of GenBank, EMBL and DDBJ. Identified bacterial sequences were removed from the database.

Sequences were classified into three categories. 'Annotated' corresponds to sequences showing significant matches with protein sequences with an identified function in databanks. 'Unknown function' corresponds to sequences showing significant matches ($E < 10^{-5}$) and homology to a protein with no identified function. 'No homology' groups sequences for which the *E* value was $> 10^{-5}$, or for which no match was observed in databanks.

Finally, ESTs/clusters were grouped in functional categories according to the classification developed for the *Medicago truncatula* EST databank (Covitz *et al.*, 1998; Journet *et al.*, 2002). All resulting data (sequences, clustering results and BLAST results) were automatically integrated in a relational database (ESTDB), searchable via a local web browser-based interface. The nucleotide sequence data reported here are available in the DDBJ/EMBL/GenBank databases under accession numbers CO 036851–CO 038878. Contig sets and singletons are available to the public at <http://www.mpl.ird.fr/rhizo/> (see Resources). Custom-made pipeline requests should be submitted to bogusz@mpl.ird.fr.

In a second set of analyses, root and nodule ESTs were then grouped and submitted to STACKPACK (<http://www.sanbi.ac.za/Dbases.html>) to compare root and nodule ESTs and to sort nodule-specific sequences. Results were stored in the database.

Quantitative real-time RT-PCR

Gene expression was performed using a two-step qRT-PCR procedure. Poly(A)⁺ RNA was (1) purified from root and nodule samples (each from 40 plants) using the Oligotex mRNA Spin-Column kit (Qiagen) to eliminate bacterial RNA present in nodules; (2) quantified with Quant-iT Ribogreen RNA Reagent (Molecular Probes, Invitrogen, Eugene, OR, USA); and (3) reverse-transcribed (49.5 ng per reaction) using the Reverse Transcription System (Promega, Madison, WI, USA). To minimize potential heterogeneity in RT reaction yield, each cDNA sample was derived from 5 independent RT reactions. cDNAs were used as templates in qRT-PCR reactions with gene-specific primers designed using BEACON DESIGNER software (Premier Biosoft International, Palo Alto, CA, USA) (see Table 2). qRT-PCR was conducted using the FullVelocity SYBR Green QPCR Master Mix in a total volume of 25 µl and the FullVelocity cycling PCR program on an MX 3005P (Stratagene). A melting curve was recorded at the end of every run to assess product specificity. For each target gene, PCR conditions (primer concentrations, cDNA quantity) were optimized and PCR efficiency was determined. qRT-PCR reactions were run in three replicates/plates, and experiments were repeated four times. Quantification using the comparative threshold-cycle method was performed. Target gene expression was normalized to ubiquitin and corrected according to the PCR

efficiency value (Pfaffl, 2001). Data were processed by two-way ANOVA to determine the significance of gene-expression differences between roots and nodules. The products of qRT-PCR were run on agarose gel electrophoresis and showed an equal sized band as predicted in the template sequence.

Results

Casuarina glauca root and nodule EST analysis

Sequencing produced 2878 ESTs with an average length of 447 and 357 bp for *C. glauca* root and nodule libraries, respectively (Table 1). Of these sequences, 70% were considered to be of high quality; 948 of these derived from nodules and 1080 from roots. Then 30% of the sequences were eliminated because of low quality, small size (<50 bp) or insert problems. A total of 1616 different genes in the form of 187 clusters (242 contigs), assembled from two or more ESTs and 1429 singletons, were formed from root and nodule data sets (Table 1). Many ESTs that produced identical BLAST hits were grouped in the same contigs. We also found few clusters containing ESTs with similar but not identical DNA sequences that may encode different isoforms. Sequence redundancy (ESTs in clusters/total ESTs) reached 26% for roots and 34% for nodules, suggesting that continued sequencing has the potential to uncover other novel sequences from the constructed libraries.

Functional annotation of sequences and clustering

Examination of the initial BLAST matches resulted in classification of the nodule and root sequences (singletons

Table 1 *Casuarina glauca* root and nodule expressed sequence tags (ESTs) and cluster collection statistics

	Root	Nodule	Total
Number of cDNAs sequenced	1534	1344	2878
EST summary:			
Number of high-quality ESTs	1080 (70%)	948 (71%)	2028 (70%)
Mean EST length (bp)	447	357	402
EST size range (bp)	50–882	50–825	50–880
Nonvalid sequences	454 (30%)	396 (29%)	850 (30%)
Small size sequence (< 50 bp)	213 (47%)	84 (21%)	297 (35%)
Low-quality sequences	137 (30%)	274 (69%)	411 (48%)
Sequences with multi-inserts	7 (2%)	1 (0.2%)	8 (0.9%)
Sequences with no insert	97 (21%)	37 (9%)	134 (16%)
Cluster summary:			
Number of ESTs assembled	1080	948	2028
Number of clusters	103	84	187
Number of singletons	800	629	1429
Number of contig sequences	126	116	242
Mean contig length (bp)	620	500	
Contig size range (bp)	50–1374	83–919	
Redundancy (%)	26	34	30

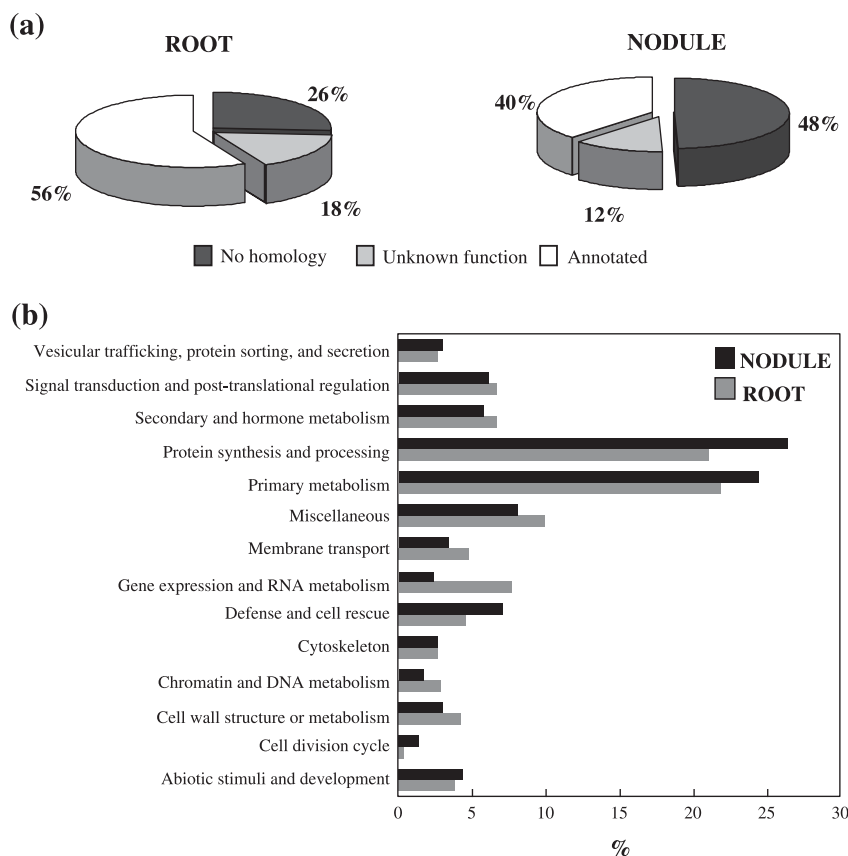


Fig. 1 Classification of *Casuarina glauca* clusters and singletons. (a) Distribution of clusters and singletons based on *E* value and top 10 results of BLAST. (b) Distribution of annotated *C. glauca* clusters and singletons according to the functional categories developed for annotation of *Medicago truncatula* expressed sequence tags (ESTs).

and clusters) in three categories (Fig. 1a). A large proportion of nodule sequences (48% compared with 26% for root sequences) showed no significant match with any protein sequence in the public databases, and 18% of the root and 12% of the nodule sequences match to protein sequences classified as 'unknown function'. Around 60% of root and 40% of nodule sequences were annotated, and the sequences identified were subsequently assigned to 14 functional categories on the basis of the classification developed for the *M. truncatula* EST databank (Covitz *et al.*, 1998; Journet *et al.*, 2002) (Fig. 1b). All categories were represented, and for both nodule and root ESTs the major categories were 'protein synthesis' and 'primary metabolism'. These two categories, plus 'cell-division cycle' and 'defence and cell rescue', were more common in nodules than in roots.

A number of nodule EST/cluster sequences showed similarity to previously described proteins such as actinorhizal nodulins: haemoglobin (Gherbi *et al.*, 1997); metallothionein (Laplaze *et al.*, 2002); subtilisin (Laplaze *et al.*, 2000); rubisco activase (Okubara *et al.*, 1999); saccharose synthase (Van Ghelue *et al.*, 1996); glycine and histidine-rich proteins (Pawlowski *et al.*, 1997). Furthermore, several ESTs showed some similarity to nodulin genes identified in legume–*Rhizobium* symbioses, such as carbonic anhydrase (Table 2), which is thought to be involved in oxygen regulation in some legume

nodules (Galvez *et al.*, 2000; Flegmetakis *et al.*, 2003). However, EST homologues to early nodulin genes in legumes such as *ENOD2*, *ENOD12* and *ENOD11* (Vijn *et al.*, 1995; Journet *et al.*, 2001) were not detected.

Comparison of *C. glauca* root and nodule sequences

To compare root and nodule sequences, an assembling method using STACKPACK software was used for the complete set of ESTs. This analysis enabled sorting of nodule- and root-specific sequences, and revealed a number of sequences that were present in both organs. This provided insights into the level of expression of genes in roots and nodules (data not shown). The majority of nodule-specific sequences (70.5%) were composed of unique sequences or groups of two ESTs, which were considered to be relatively low-copy gene transcripts. Twenty per cent of the sequences grouped three to four ESTs. The largest and unique nodule-specific group of sequences consisted of a total of 33 ESTs with homology to haemoglobin. It should be noted that the largest group of sequences (82 ESTs, named CL1) encoded a class of protein that did not match any known protein sequence. Of these 82 ESTs, 78 occurred in nodules and only four in roots, suggesting that CL1 expression is induced on nodule formation.

Table 2 *Casuarina glauca* genes selected for qRT–PCR analysis on the basis of their putative involvement in nodule development and/or functioning

Sequence name in EST database	Accession number	Description	Primers	
			Forward	Reverse
<i>Casuarina glauca</i> genes selected for qRT–PCR analysis				
CG-N01_013_A07	CO038723	Putative peptide transporter (<i>CgPPT</i>)	GGAGTCGGAAACTTTATGAGCAG	AAGATGGGAAGCATTGAGTTG
CG-N01_006_A05	CO038323	Carbonic anhydrase (<i>CgCA</i>)	AAAACCGGGTTCGCTCACTT	GCAAACACGAGAGTCTGAGCAT
CG-N01_008_F06	CO038106	Sucrose synthase (<i>CgSS</i>)	CTGGAAGCACGTGTCCAACC	GACGGCAAGAGGAACGGACT
CG-N01_006_C04	CO038303	Dihydroflavonol 4-reductase (<i>CgDFR</i>)	CTGCGTCACTGGTGCTTCTG	CCTTCTTCATATTCCTGGGTCTCTC
CG-N01_002_A10	CO038460	Flavonol synthase (<i>CgFS</i>)	CCACAGGAGACCATCATCGG	CAGCAGTTTACCACCCTTTTCG
CG-N01_002_B10	CO038451	Flavonoid 3',5'-hydroxylase (<i>CgF3'5'H</i>)	CAACCACACATGCGAAGTG	ACATCAGGGTCTCGCCCAAT
CG-N01_012_A03	CO038801	Laccase (<i>CgLAC</i>)	GCAATCAGATTCAAGGCAGATAACC	GGTGGTAGTAGTCTCATTCAAGC
CG-N01_011_B09*	CO038855	No hit found (<i>N11B9</i>)	GGGCGGGGGCCCTGCCTCAGTG	GAGGGGAGGGACGGGATTTCG
Genes already described in <i>C. glauca</i>				
CG-N01_012_E09	CO038761	Chalcone synthase (<i>CgCHS1</i>) (Laplaze <i>et al.</i> , 1999)	CTTCGCCCATCCGTCAAAG	TCTCCGAGCACACGACAAGC
CG-N01_002_C11	CO038438	Subtilase (<i>Cg12</i>) (Laplaze <i>et al.</i> , 2000)	ATGCCACGCTTGATACCAC	CCGACTTGACAAATTCCTTTCC
CgENOD40**		Homologous to leguminous ENOD40 (Santi <i>et al.</i> , 2003)	GATTCTTAACTCTGCTGATGC	TTGCCGTTCTGTGACTTG
CG-N01_002_E06	CO038421	Haemoglobin (<i>CgHb</i>) (Gherbi <i>et al.</i> , 1997)	TGGGAGGTACTGAAGCAAAC	AAGGAGAACACATACTTGGATTC
CG-N01_002_F01	CO038414	Metallothionein (<i>CgMT1</i>) (Laplaze <i>et al.</i> , 2002)	TGTCTTCTGTGGCTGTG	TCCTCCTTCAACGCTCATC
Genes used for normalization of real-time PCR results				
CG-R01_010_F8	CO037049	Poly ubiquitin (<i>CgUBI</i>)	CCAGGATAAGGAGGGCATTTC	GGAGACGGAGCACAAAGATG

Expression was studied using a comparative qRT–PCR method. The primers used for each gene are indicated.

*This sequence is the most representative of the CL1 group, and was used for primer design.

**This gene was not found in the EST database, but was used as a control in qRT–PCR experiments.

Comparison of expression pattern in *C. glauca* roots and nodules for selected ESTs by qRT-PCR

In order to test the usefulness of our EST database to identify symbiotic genes, 13 genes were selected to verify nodule-specific and/or nodule-enhanced expression by qRT-PCR on the basis of their putative involvement in nodule development and/or functioning (Table 2). CL1, which groups 82 ESTs, was also selected, as *in silico* analysis had shown it to be overexpressed in nodules.

Using the ubiquitin (*CgUBI*) gene, which is equally expressed in roots and nodules (Fig. 2a), as internal control, we first analysed the level of expression of previously described actinorhizal symbiotic genes that were present in the database. Figure 2b shows that metallothionein (*CgMT1*) was equally expressed in roots and nodules. We also evaluated transcript abundance of the *C. glauca* *ENOD40* gene (Santi *et al.*,

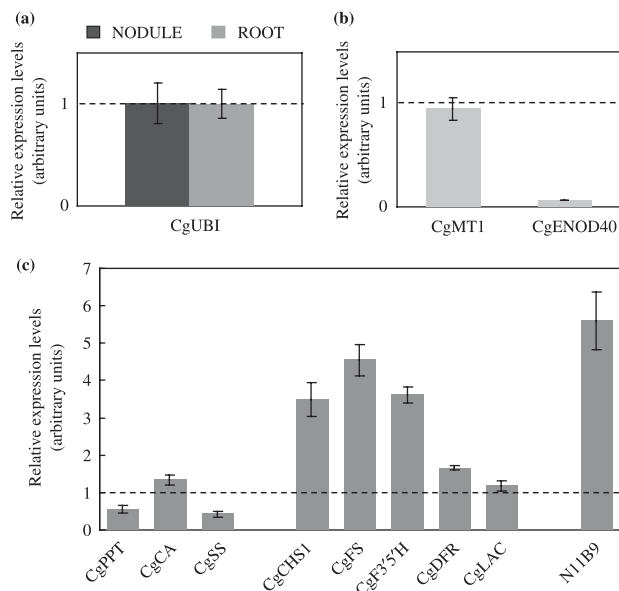


Fig. 2 Analysis by qRT-PCR of the expression of selected genes in nodules of *Casuarina glauca*. Relative quantification was calculated by comparing levels of gene expression in nodule samples with root samples (control sample) after gene normalization using ubiquitin (*CgUBI*) as reference gene. Data are expressed as fold-difference of gene expression with respect to the control sample, expression of which is set at 1 (dashed line). Values > 1 correspond to overexpression in nodules, while values < 1 correspond to underexpression. Values are the mean of four experiments. (a) Expression of *CgUBI* in roots and nodules. Expression was similar in both organs and this gene was used for qRT-PCR normalization. (b) Expression of *CgMT1* and *CgENOD40* previously characterized in *C. glauca*. These were used as controls for qRT-PCR validation. (c) Expression of target genes. *CgUBI*, ubiquitin; *CgMT1*, metallothionein; *CgENOD40*, ENOD 40; *CgPPT*, putative peptide transporter; *CgCA*, carbonic anhydrase; *CgSS*, sucrose synthase; *CgLAC*, laccase; *CgCHS1*, chalcone synthase; *CgFS*, flavonoid synthase; *CgF3'5'H*, flavonoid 3',5'-hydroxylase; *CgDFR*, dihydroflavonol 4-reductase; *N11B9*, unknown gene.

2003), which was not expressed in nodules. Conversely, haemoglobin (*CgHb*) and *Cg12*, encoding a subtilase, appeared to be expressed only in nodules. The transcripts of a putative peptide transporter (*CgPPT*) and of sucrose synthase (*CgSS*) were less abundant in nodules, while carbonic anhydrase (*CgCA*) was more expressed in roots and nodules (Fig. 2c).

In previous studies we showed that *C. glauca* produces large amounts of polyphenols in response to *Frankia* infection, and in mature nodules (Laplaze *et al.*, 1999). We therefore studied the expression of five genes present in the database encoding enzymes involved in polyphenol metabolism. Chalcone synthase (*CgCHS1*); flavonoid 3',5'-hydroxylase (*CgF3'5'H*); dihydroflavonol 4-reductase (*CgDFR*); and flavonol synthase (*CgFS*) were more highly expressed in nodules than in roots (3.49, 3.6, 1.6 and 4.5 times higher, respectively). A laccase (*CgLAC*) was also shown to be expressed slightly more in nodules.

The transcript annotated as having no homology to any known genes in available databases (CL1) was studied on the basis of the expression of its most representative sequence (*N11B9*; Table 2): it was found to be up to 5.6 times more abundant in nodules, and may be a new, symbiotic, enhanced *C. glauca* gene (Fig. 2c). For each gene, ANOVA revealed a significant difference in expression between root and nodule ($P = 0.000$ at the 5% level) excepted for *CgUBI* ($P = 0.155$) and for *CgMT1* ($P = 0.432$).

Discussion

Here we provide the first genomic platform for the study of plant gene expression in actinorhizal symbiosis. Our *C. glauca* root and nodule EST database contains 1616 different unique transcripts (clusters and singletons) assembled from over 2000 valid sequences. A low level of redundancy was observed (26 and 34% for roots and nodules, respectively), indicating that the two libraries still have considerable potential to uncover novel sequences. Database similarity searches with our ESTs showed that 48% of the total nodule ESTs did not match any protein sequence, suggesting that these transcripts represent important genes that are specific to actinorhizal symbiosis. A majority of annotated root and nodule sequences were classified as 'primary metabolism' or 'protein synthesis and processing'; the percentages of these subgroups are similar to those noted for *M. truncatula* and *Lotus japonicus* (Journet *et al.*, 2002; Poulsen & Podenphant, 2002). Together with the two latter functional subgroups, transcripts involved in the 'cell-division cycle' are more abundant in nodules than in roots, a finding consistent with the increase in protein metabolism, cell division and cell growth observed in developing organs. We also found that, similarly to *M. truncatula* reported data, the 'defence and cell rescue' functional category is represented more often in nodules than in roots, suggesting that responses induced by symbionts and pathogens overlap (Journet *et al.*, 2002).

We used qRT-PCR to assess differential expression of several genes in nodules vs roots. The validity of this approach

was confirmed because, as previously shown by Gherbi *et al.* (1997) and Laplace *et al.* (2000), we observed nodule-specific expression of haemoglobin (*CgHB*) and serine protease (*Cg12*). The expression profiles of *CgENOD40* and *CgMT1* were also in agreement with earlier studies by Santi *et al.* (2003) and Laplace *et al.* (2002). Surprisingly, all selected nodule genes linked to primary metabolism (*CgSS*, *CgCA*, *CgPPT*), and predicted to be upregulated during nodule development as shown in legumes (Journet *et al.*, 2002; Colebatch *et al.*, 2004; El Yahyaoui *et al.*, 2004), appeared to be under- or only slightly overexpressed in *C. glauca* nodules. Similar observations were reported for *M. truncatula*, and could be linked to the fact that these genes are members of multigene families, with alternative expression during nodule development (Journet *et al.*, 2002; El Yahyaoui *et al.*, 2004).

Enhanced transcript expression in nodules of a laccase (*CgLAC*), a polyphenoloxidase involved in lignification process and plant–pathogen responses (Mayer & Staples, 2002), and of several enzymes related to flavonoid synthesis (Winkel-Shirley, 2001) support previous data showing that the amount of phenolic compounds increases dramatically in *C. glauca* nodules compared with uninfected roots (Laplace *et al.*, 1999). The involvement of flavonoids in the establishment of legume–*Rhizobium* interactions is well documented (Perret *et al.*, 2000), and they could also be involved in the morphogenesis of legume nodules (Stafford, 1997; Mathesius, 2001). The role of flavonoids in actinorhizal species is still unknown but, together with previous work (Laplace *et al.*, 1999; 2006), our results suggest that, similarly to legume–*Rhizobium* symbioses, polyphenols play a significant role in actinorhizal symbiosis.

In conclusion, a collection of 2028 ESTs of an actinorhizal species has been generated for the first time. Obviously this EST database needs to be amplified, but it already represents a resource for further functional analysis of specific genes and for comparative genomics with legume–*Rhizobium* and mycorrhizal symbioses. This should provide clues to the identification of genetic factors, thus enabling successful root–nodule symbioses to be established.

Acknowledgements

We thank B. Piegu and R. Cook for discussion on the ESTs analysis pipeline design, C. Montavon for helpful discussion on qRT–PCR, and S. Dussert for statistical analysis of our results.

References

Asamizu E, Nakamura Y, Sato S, Tabata S. 2004. Characteristics of the *Lotus japonicus* gene repertoire deduced from large-scale expressed sequence tag (EST) analysis. *Plant Molecular Biology* 54: 405–414.

Cannon SB, Crow JA, Heuer ML, Wang X, Cannon EK, Dwan C, Lamblin AF, Vasdevani J, Mudge J, Cook A, Gish J, Cheung F, Kenton S, Kunau TM, Brown D, May GD, Kim D, Cook DR, Roe BA, Town CD, Young ND, Retzel EF. 2005. Databases and information integration for the *Medicago truncatula* genome and transcriptome. *Plant Physiology* 138: 38–46.

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

Covitz PA, Smith LS, Long SR. 1998. Expressed sequence tags from a root-hair-enriched *Medicago truncatula* cDNA library. *Plant Physiology* 117: 1325–1332.

Doyle JJ. 1998. Phylogenetic perspectives on nodulation: evolving views of plants and symbiotic bacteria. *Trends in Plant Science* 3: 473–478.

Duc G, Trouvelet A, Gianinazzi-Pearson V, Gianinazzi S. 1989. First report of non-mycorrhizal plant mutants (Myc-) obtained in pea (*Pisum sativum* L.) and faba bean (*Vicia faba* L.). *Plant Science* 60: 215–222.

Duplessis S, Courty PE, Tagu D, Martin F. 2005. Transcript patterns associated with ectomycorrhiza development in *Eucalyptus globulus* and *Pisolithus microcarpus*. *New Phytologist* 165: 599–611.

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

Flemetakis E, Dimou M, Cotzur D, Aivalakis G, Efroze RC, Kenoutis C, Udvardi M, Katinakis P. 2003. A *Lotus japonicus* beta-type carbonic anhydrase gene-expression pattern suggests distinct physiological roles during nodule development. *Biochimica Biophysica Acta* 1628: 186–194.

Franché C, Laplace L, Duhoux E, Bogusz D. 1998. Actinorhizal symbioses: recent advances in plant molecular and genetic transformation studies. *Critical Reviews in Plant Science* 17: 1–28.

Galvez S, Hirsch AM, Wycoff KL, Hunt S, Layzell DB, Kondoroski A, Crespi M. 2000. Oxygen regulation of a nodule-located carbonic anhydrase in alfalfa. *Plant Physiology* 124: 1059–1068.

Gherbi H, Duhoux E, Franché C, Pawlowski K, Nassar A, Berry A, Bogusz D. 1997. Cloning of a full-length symbiotic hemoglobin cDNA and *in situ* localization of the corresponding mRNA in *Casuarina glauca* root nodule. *Physiologia Plantarum* 99: 608–616.

Girgis ZGM, Ishac ZY, El-Haddad M, Saleh AE, Diem HG, Dommergues RY. 1990. First report on isolation and culture of effective *Casuarina*-compatible strains of *Frankia* from Egypt. In: El-Lakany MH, Turnbull JW, Bewbaker JL, eds. *Advances in Casuarina research and utilisation*. Cairo, Egypt: Desert Development Center, American University of Cairo, pp. 156–164.

Grunwald U, Nyamsuren O, Tamasloukht M, Lapopin L, Becker A, Mann P, Gianinazzi-Pearson V, Krajinski F, Franken P. 2004. Identification of mycorrhiza-regulated genes with arbuscule development-related expression profile. *Plant Molecular Biology* 55: 553–566.

Jouannic S, Argout X, Lechaue F, Fizames C, Borgel A, Morcillo F, Aberlenc-Bertossi F, Duval Y, Tregear J. 2005. Analysis of expressed sequence tags from oil palm (*Elaeis guineensis*). *FEBS Letters* 579: 2709–2714.

Journet E-P, El-Gachtouli N, Vernoud V, de Billy F, Pichon M, Dodieu A, Arnould C, Morandi D, Barker DG, Gianinazzi-Pearson V. 2001. *Medicago truncatula ENOD11*: a novel RPRP encoding early nodulin gene expressed during mycorrhization in arbuscule-containing cells. *Molecular Plant–Microbe Interactions* 14: 737–748.

Journet E-P, van Tuinen D, Gouzy J, Crespeau H, Carreau V, Farmer M-J, Niebel A, Schiex T, Jaillon O, Chatagnier O, Godiard L, Micheli F, Kahn D, Gianinazzi-Pearson V, Gamas P. 2002. Exploring root symbiotic programs in the model legume *Medicago truncatula* using EST analysis. *Nucleic Acids Research* 30: 5579–5592.

Kistner C, Parniske M. 2002. Evolution of signal transduction in intracellular symbiosis. *Trends in Plant Science* 7: 511–518.

Laplace L, Gherbi H, Frutz T, Pawlowski K, Franché C, Macheix J-J, Auguy F, Bogusz D, Duhoux E. 1999. Flavan-containing cells delimit *Frankia*-infected compartments in *Casuarina glauca* nodules. *Plant Physiology* 121: 113–122.

- Laplaze L, Ribeiro A, Franche C, Duhoux E, Auguy F, Bogusz D, Pawlowski K. 2000. Characterization of a *Casuarina glauca* nodule-specific subtilisin-like protease gene, a homolog of *Alnus glutinosa* Ag12. *Molecular Plant-Microbe Interactions* 13: 113–117.
- Laplaze L, Gherbi H, Duhoux E, Pawlowski K, Auguy F, Guermache F, Franche C, Bogusz D. 2002. Symbiotic and non-symbiotic expression of *cgMT1*, a metallothionein-like gene from the actinorhizal tree *Casuarina glauca*. *Plant Molecular Biology* 49: 81–92.
- Laplaze L, Svistoonoff S, Santi C, Auguy F, Franche C, Bogusz D. 2006. Molecular biology of actinorhizal symbioses. In: Newton WE, ed. *Nitrogen fixation research: origins and progress, Vol. VI: Actinorhizal symbioses*. Dordrecht, the Netherlands: Kluwer, in press.
- Mathesius U. 2001. Flavonoids induced in cells undergoing nodule organogenesis in white clover are regulators of auxin breakdown by peroxidase. *Journal of Experimental Botany* 52: 419–426.
- Mayer AM, Staples RC. 2002. Laccase: new functions for an old enzyme. *Phytochemistry* 60: 551–565.
- Okubara PA, Pawlowski K, Murphy TM, Berry AM. 1999. Symbiotic root nodules of the actinorhizal plant *Datisca glomerata* express rubisco activase mRNA. *Plant Physiology* 120: 411–420.
- Pawlowski K, Bisseling T. 1996. Rhizobial and actinorhizal symbioses: what are the shared features? *Plant Cell* 8: 1899–1913.
- Pawlowski K, Twigg P, Dobritsa S, Guan CH, Mullin B. 1997. A nodule-specific gene family from *Alnus glutinosa* encodes glycine- and histidine-rich proteins expressed in the early stages of actinorhizal nodule development. *Molecular Plant-Microbe Interactions* 10: 656–664.
- Perret X, Staehelin C, Broughton WJ. 2000. Molecular basis of symbiotic promiscuity. *Microbiology and Molecular Biology Reviews* 64: 180–201.
- Pfaffl MW. 2001. A new mathematical model for relative quantification in real-time RT-PCR. *Nucleic Acids Research* 29: 2002–2007.
- Poulsen C, Podenphant L. 2002. Expressed sequence tags from roots and nodule primordia of *Lotus japonicus* infected with *Mesorhizobium loti*. *Molecular Plant-Microbe Interactions* 15: 376–379.
- Santi C, von Groll U, Chiurazzi M, Auguy F, Bogusz D, Franche C, Pawlowski K. 2003. Comparison of nodule induction in legume and actinorhizal symbioses: the induction of actinorhizal nodules does not involve *Enod40*. *Molecular Plant-Microbe Interactions* 16: 808–816.
- Shoemaker R, Keim P, Vodkin L, Retzel E, Clifton SW, Waterston V, Smoller D, Coryell V, Khanna A, erpelding J, Gai XW, Brendel V, Raph-Schmidt C, Shoop EG, Vielweber CJ, Schmatz M, Pape D, Bowers Y, Theising B, Martin J, Dante M, Wylie T, Granger C. 2001. A compilation of soybean ESTs: generation and analysis. *Genome* 45: 329–338.
- Soltis DE, Soltis PS, Morgan DR, Swensen SM, Mullin BC, Dowd JM, Martin PG. 1995. Chloroplast gene sequence data suggest a single origin of the predisposition for symbiotic nitrogen fixation in angiosperms. *Proceedings of the National Academy of Sciences, USA* 92: 2647–2651.
- Stafford HA. 1997. Role of flavonoids in symbiotic and defense functions in legume roots. *Botanical Review* 63: 27–39.
- Van Ghelue M, Ribeiro A, Solheim B, Akkermans ADL, Bisseling T, Pawlowski K. 1996. Sucrose synthase and enolase expression in actinorhizal nodules of *Alnus glutinosa*: comparison with legume nodules. *Molecular and General Genetics* 250: 437–446.
- Vijn I, Martinez-Abarca F, Yang W-C, das Neves L, van Brussel A, van Kammen A, Bisseling T. 1995. Early nodulin gene expression during *Nod* factor-induced processes in *Vicia sativa*. *Plant Journal* 8: 111–119.
- Winkel-Shirley B. 2001. Flavonoid biosynthesis. A colorful model for genetics, biochemistry, cell biology and biotechnology. *Plant Physiology* 126: 485–493.



About New Phytologist

- *New Phytologist* is owned by a non-profit-making **charitable trust** dedicated to the promotion of plant science, facilitating projects from symposia to open access for our Tansley reviews. Complete information is available at www.newphytologist.org.
- Regular papers, Letters, Research reviews, Rapid reports and both Modelling/Theory and Methods papers are encouraged. We are committed to rapid processing, from online submission through to publication 'as-ready' via *OnlineEarly* – the 2004 average submission to decision time was just 30 days. Online-only colour is **free**, and we provide 25 offprints as well as a PDF for each article.
- For online summaries and ToC alerts, go to the website and click on 'Journal online'. You can take out a **personal subscription** to the journal for a fraction of the institutional price. Rates start at £125 in Europe/\$232 in the USA & Canada for the online edition (click on 'Subscribe' at the website).
- If you have any questions, do get in touch with Central Office (newphytol@lancaster.ac.uk; tel +44 1524 594691) or, for a local contact in North America, the US Office (newphytol@ornl.gov; tel +1 865 576 5261).