DNA Research **11**, 263–274 (2004)

# Large-Scale Analysis of Gene Expression Profiles during Early Stages of Root Nodule Formation in a Model Legume, *Lotus japonicus*

Hiroshi Kouchi,[1,*,†] Kenshiro Shimomura,[2,†] Shingo Hata,[3] Atsuko Hirota,[2] Guo-Jiang Wu,[1] Hirotaka Kumagai,[1] Shigeyuki Tajima,[2] Norio Suganuma,[4] Akihiro Suzuki,[5] Toshio Aoki,[6] Makoto Hayashi,[7] Tadashi Yokoyama,[8] Takuji Ohyama,[9] Erika Asamizu,[10] Chikara Kuwata,[10] Daisuke Shibata,[10] and Satoshi Tabata[10]

*National Institute of Agrobiological Sciences, Tsukuba, Ibaraki 305-8602, Japan,[1] Kagawa University, Miki-cho, Kagawa 761-0795, Japan,[2] Kyoto University, Kyoto 606-8502, Japan,[3] Aichi University of Education, Kariya, Aichi 448-8542, Japan,[4] Kagoshima University, Kagoshima 890-0065, Japan,[5] Nihon University, Fujisawa, Kanagawa 252-8510, Japan,[6] Osaka University, Suita, Osaka 565-0871, Japan,[7] Tokyo University of Agriculture and Technology, Fuchu, Tokyo 183-8538, Japan,[8] Niigata University, Niigata 950-2181, Japan,[9] and Kazusa DNA Research Institute, Kisarazu, Chiba 292-0812, Japan[10]*

(Received 13 May 2004; revised 12 July 2004)

**Abstract**

Gene expression profiles during early stages of formation of symbiotic nitrogen-fixing nodules in a model legume *Lotus japonicus* were analyzed by means of a cDNA array of 18,144 non-redundant expressed sequence tags (ESTs) isolated from *L. japonicus*. Expression of a total of 1,076 genes was significantly accelerated during the successive stages that represent infection of *Mesorhizobium loti*, nodule primordium initiation, nodule organogenesis, and the onset of nitrogen fixation. These include 32 nodulin and nodulin-homolog genes as well as a number of genes involved in the catabolism of photosynthates and assimilation of fixed nitrogen that were previously known to be abundantly expressed in root nodules of many legumes. We also identified a large number of novel nodule-specific or enhanced genes, which include genes involved in many cellular processes such as membrane transport, defense responses, phytohormone synthesis and responses, signal transduction, cell wall synthesis, and transcriptional regulation. Notably, our data indicate that the gene expression profile in early steps of *Rhizobium*-legume interactions is considerably different from that in subsequent stages of nodule development. A number of genes involved in the defense responses to pathogens and other stresses were induced abundantly in the infection process, but their expression was suppressed during subsequent nodule formation. The results provide a comprehensive data source for investigation of molecular mechanisms underlying nodulation and symbiotic nitrogen fixation.

**Key words:** cDNA array; nitrogen fixation; nodulin; *Rhizobium*-plant interactions

## 1. Introduction

The formation of symbiotic nitrogen-fixing nodules is established by complex interactions between *Rhizobium* bacteria and legume plants. These interactions start from the attachment of rhizobia to legume root hairs, followed by invasion of rhizobia into root cortical cells through infection threads, induction of root cortical cell division to form nodule primordia, and development of a highly organized symbiotic organ, the root nodule, in which rhizobia differentiate to bacteroids and fixation of atmospheric nitrogen takes place.[1] Progress in the molecular genetics of *Rhizobium* bacteria during the past decade has uncovered the mechanisms of early signaling from the microsymbionts that enable host recognition and induce specific host responses. Rhizobial chitooligosaccharide signal molecules, Nod factors, can elicit both infection and nodule formation processes with strict host specificity.[2] However, the molecular mechanisms programmed in the host legumes, such as mechanisms of perception of Nod factors, and of subsequent signal transduction and nodule organogenesis are still largely

Communicated by Kazuki Saito

* To whom correspondence should be addressed. Tel. +81-29-838-8377, Fax. +81-29-838-8347, E-mail: kouchih@nias.affrc.go.jp

† These authors contributed equally to this work.

unknown.

*Lotus japonicus* has been proposed as a model legume for molecular genetic studies of symbiotic nitrogen fixation.[3] This legume species, as well as another model legume *Medicago truncatula*, has a relatively small genome size and short generation time, is self-fertile and capable of molecular transfection. Thus its use overcomes difficulties in molecular genetic studies with major crop legumes like soybean, pea, and common bean. A large number of symbiotic mutant lines of *L. japonicus* have been established,[4,5,6] and the resources required for molecular genetics are being prepared systematically, *e.g.* accumulation of expressed sequence tags,[7,8] construction of high-density genetic maps,[9] and whole genome sequencing (http://www.kazusa.or.jp/lotus/). On the basis of the genomic information, the plant genes that are essential in early steps of Nod factor perception have been cloned recently.[10,11]

The nodulation process involves a drastic alteration of the expression of the host legume genes that is due to interactions with rhizobia and nodule organogenesis. In particular, a unique set of plant genes that are exclusively induced in the nodulation process are termed "nodulin genes" and have been isolated from many legume species.[12,13] Although their exact functions are mostly unknown at present, they are believed to play essential roles in nodule formation. Comprehensive analysis of gene expression profiles during the nodulation process has critical importance in understanding *Rhizobium*-legume interactions and subsequent nodule formation.

We describe here a systematic effort for such analysis by means of a cDNA macroarray consisting of *ca.* 18,000 non-redundant ESTs from the model legume *L. japonicus*. Application of a cDNA array to legume nodules was first described by Colebatch et al.[13] for *L. japonicus* on a much smaller scale. Gene expression profiling for nodules has also been reported using *in silico* anaysis of an EST database for *M. truncatula*.[14] In this paper, we present detailed analysis of variation of transcript levels through the crucial steps of the nodulation process, that is the infection of *Mesorhizobium loti*, nodule primordium initiation, nodule development, and the onset of nitrogen fixation (fully mature nodules). These results provide the most comprehensive data source at present for investigation of molecular mechanisms underlying nodulation and symbiotic nitrogen fixation.

## 2. Materials and Methods

### *2.1. Plant materials and RNA isolation*

Seeds of *Lotus japonicus* accession B129 'Gifu'[3] were surface-sterilized and germinated on agar plates. Six-day-old seedlings were transplanted to the pots with 'pillows'.[6] In brief, the pillows were made of tea bags (8×5 cm) filled with moist vermiculite and placed in a rectangular plastic pot. After autoclaving the pots, the roots of the seedlings were inserted between each pillow followed by inoculation with *Mesorhizobium loti* strain *Tono*[4] 2 days later. The plants were grown in an artificially lit growth cabinet controlled at 24°C (day, 16 hr) and 22°C (night, 8 hr) and were supplemented with nitrogen-free medium.[15] After 2 days of bacterial inoculation, root segments of the susceptible zone (*ca.* 5 mm in length from 3 mm above the root tip) were harvested. After 4, 7, and 12 days, root segments with nodule primordia or nodules were excised under a binocular microscope. Uninfected roots were harvested from 6- and 9-day-old seedlings excluding the root tip meristematic region. The harvested samples were frozen and ground to a fine powder under liquid nitrogen followed by total RNA isolation with the RNeasy Plant Mini-Kit (Qiagen GmbH, Hilden, Germany) according to the manufacturer's instructions.

### *2.2. EST resources*

Expressed sequence tags (ESTs) were generated from cDNA libraries of 2-week-old seedlings[7] and various organs of *L. japonicus* accessions B-129 'Gifu' or MG-20 'Miyakojima'[16] (see Table 1). These ESTs were all cloned directionally in a plasmid Bluescript II (Stratagene, La Jolla, CA) at *Eco*RI (5′) and *Xho* I (3′) sites and stored as bacterial suspensions of *E. coli* XL1-Blue (Stratagene).[7] Among the total 67,916 ESTs generated, 19,127 non-redundant clones were selected by their sequence comparison under the criteria described previously[7] for the array construction.

### *2.3. cDNA array construction*

Inserts of the selected non-redundant EST clones were amplified by polymerase-chain-reaction (PCR) directly from aliquots of bacterial suspension cultures as templates. PCR was performed with the Ex-Taq system (Takara Bio, Otsu, Japan) according to the manufacturer's instructions. The primers were designed outside the multi-cloning site of Bluescript plasmid and were 5′-TCATTAGGCACCCCAGGCTTTACAC-3′ (reverse) and 5′-GTAATACGACTCACTATAGGGC-3′ (forward). The amplification conditions were 94°C for 2 min; followed by 40 cycles of 94°C for 45 sec, 55°C for 45 sec, and 72°C for 2 min; and 72°C for 10 min. Alternatively, a condition of 94°C for 2 min; 40 cycles of 96°C for 10 sec and 68°C for 6 min; and 70°C for 7 min was used for long inserts. The PCR products were qualified by agarose gel electrophoresis, and finally 18,144 clones were spotted onto nylon membranes (8×12 cm, Biodyne A, Pall, East Hills, NY) using a BioMek 2000 Laboratory Automation Workstation (Beckman Instruments, Fullerton, CA) equipped with a 384 gridding tool with a 4×4 secondary grid pattern. Thus, one membrane held 6,144 spots including 96 spots of control $\lambda$ DNA, and one array set

**Table 1**. Resources of the expressed sequence tags (ESTs) of *Lotus japonicus* used for the cDNA array construction. Non-redundant ESTs were selected by sequence comparisons under criteria described previously.[7] *) N, normalized library; L, size-selected library. Identifiers of gene I.D. for each library are shown in parentheses. **) The 3′-ESTs from MG-20 seedlings are all included in the 5′-ESTs. ***) Non-redundant cDNAs for MG-20 seedlings were selected from the 5′-ESTs.

| Source | Library*) | 3'-ESTs | 5'-ESTs | Non-redundant |
|---|---|---|---|---|
| MG-20 seedlings (2 weeks) | N (MWM) | 3,921**) | 18,280 | 5,468***) |
| | L (MWL) | 1,468 | 4,703 | 2,208 |
| MG-20 Pods | N (MPD) | 6,787 | | 2,209 |
| | L (MPDL) | 4,811 | | 1,875 |
| MG-20 Roots | N (MR) | 7,751 | | 1,720 |
| | L (MRL) | 2,357 | | 675 |
| Gifu Nodule Primordia | N (GN) | 7,459 | | 1,038 |
| (har1; 4 d after inoculation) | L (GNL) | 1,158 | | 1,450 |
| Gifu Mature Nodules (23 d) | N (GEN) | 4,091 | | 2,239 |
| | L (GENL) | 5,130 | | 445 |
| Total | | 44,933 | 22,983 | 19,127 |

consisted of three separate membranes. DNA on the membranes was denatured by standard procedures[17] and fixed by UV irradiation.

### *2.4. Hybridization*

Total RNA (5 $\mu$g) prepared from each plant material was reverse-transcribed using Superscript II reverse transcriptase (Invitrogen, Groningen, The Netherlands) in the presence of [$\alpha$-$^{33}$P]dCTP (Amersham Biosciences, Piscataway, NJ) with oligo-dT$_{14-18}$(C/G/A)(T/C/G/A) as a primer. The array filters were pre-hybridized in Church buffer (0.5 M sodium phosphate, pH 7.2, 7% [w/v] SDS, 1 mM EDTA)[18] containing 1 $\mu$g/ml poly-dA (Amersham Biosciences) at 62°C for more than 2 hr. Target cDNA labeled with $^{33}$P was injected into the hybridization mixture and the hybridization reaction was carried out at 62°C for more than 16 hr. After hybridization, the filters were washed successively in 2× SSC (1× SSC is 0.15 M NaCl, 0.015 M sodium citrate, pH 7.0) containing 0.5% SDS at room temperature for 15 min, and in 0.2× SSC, 0.1% SDS at 63°C for 15 min twice. They were exposed to a phosphor imaging plate (Fuji, Tokyo, Japan) for 12 to 48 hr, followed by detection of radioactive images by a high-resolution imaging scanner (Storm 840, Amersham Biosciences). All the experiments were carried out using four replicate membrane sets with two independent target RNA preparations.

### *2.5. Primary data analysis*

Primary data collection was performed using Array Vision software (Amersham Biosciences), and the subsequent data analysis generated a mean pixel intensity within a defined area around each spot. Global normalization was adopted for normalizing the differences of signal intensities between the membranes.[18] Signal intensities of individual spots were normalized by adjusting the sum total of signal intensities of all the spots on a given membrane to $1\times10^6$.

### *2.6. Statistical analysis*

Positive and negative significant genes in each time point sample relative to uninfected roots were detected by the SAM (Significance Analysis for Microarray) software.[19] Positive significant genes detected by this procedure were further subjected to hierarchical and K-means cluster analyses using J-Express ver. 2.0 software (Molmine, Bergen, Norway).

### *2.7. Northern blot hybridization and reverse transcription-PCR analyses*

Total RNA (5 $\mu$g) was subjected to denatured agarose gel electrophoresis and blotted on a nylon membrane (Biodyne A, Pall) by standard procedures.[17] Primers for reverse transcription-polymerase chain reaction (RT-PCR) were synthesized according to 3′-EST sequences to amplify DNA fragments of 300 to 400 bp in length. Total RNA was reverse-transcribed with anchored oligo-dT$_{18}$(C/G/A)(T/C/G/A) (Promega, Madison, WI, USA) as a primer, and the resulting single-stranded cDNAs equivalent to 2 to 5 ng of total RNA were subjected to PCR with the anchor- and gene-specific primers. The PCR condition was 94°C for 4 min; followed by 20 cycles of 94°C for 30 sec, 55°C for 30 sec, and 72°C for 1 min; and 72°C for 7 min. The reaction products were electrophoresed on an agarose gel, blotted onto nylon membrane filters, and hybridized with $^{32}$P-

labelled probes prepared from each cDNA insert using a random prime labeling kit (Takara). Radioactive images were obtained as described above, and quantified with ImageQant software (Amersham Biosciences). Ubiquitin cDNA from *L. japonicus* was used as controls in both experiments.

### *2.8. DNA sequencing*

The cDNA inserts were sequenced by the dye-termination method with an automatic sequencer (ABI 3700, Applied Biosystems, Foster City, CA). Their annotation to a public database was performed using the BLASTX program with nr database that represents all non-redundant GenBank CDS translations, PDB, SwissProt and PIR.

## 3. Results and Discussion

### *3.1. Resources of expressed sequence tags (ESTs)*

The EST resource from *L. japonicus* is summarized in Table 1 (see also the Web database; http://www.kazusa.or.jp/en/plant/lotus/EST/index.html). They were collected from normalized and size-selected cDNA libraries of seedlings (2 weeks old), uninfected roots, and developing pods of *L. japonicus* MG-20 Miyakojima, nodulated roots of hyper-nodulating mutant *har1*[20] 4 days after bacterial inoculation, and mature nodules (23 days old) of B-129 Gifu. They were sequenced from the 3′-ends and/or from the 5′-ends, generating a total 67,916 ESTs. Non-redundant ESTs were selected by sequence comparisons under the criteria that the clones with more than 95% identity over at least a 100-bp stretch were included in the same cluster.[7] Accordingly, 19,127 ESTs were picked up for cDNA array construction. A few hundred clones did not give sufficient amplification in PCR so that a total of 18,144 ESTs were used for spotting onto membranes.

### *3.2. Identification of positive and negative significant genes during nodulation process*

Under our growth conditions, nodules become visible under a binocular microscope 4 to 5 days after *M. loti* inoculation, and nitrogen fixation (acetylene reduction) activity was first detected at around 10 to 12 days (data not shown). Based on these preliminary observations together with microscopy on the earlier stages of bacterial infection, we chose 2, 4, 7, and 12 days after inoculation as the material for the target RNA preparation. These time points represent stages of the most abundant infection, nodule primordium initiation, nodule organ development, and full developement (onset of nitrogen fixation) of nodules, respectively.

The RNA preparations were reverse-transcribed and subjected to hybridization with the cDNA array. The data of hybridization signal intensities were normalized for each membrane and the spots of which maximum signal intensity values throughout all experiments were less than three times the background (an average of all $\lambda$ DNA spots) were excluded. As a consequence, data sets for a total of 14,971 ESTs were subjected for further analyses. Normalized signal intensity of $\lambda$ DNA was 13.7 as an average of overall experiments, and this background value was not subtracted from the data for ESTs in the analyses that followed. Global view of transcript levels in comparison with uninfected roots is shown in Fig. 1. While the variations of individual signal intensities due to RNA preparation from the same source organs and/or between each sets of membranes were very small (Fig. 1A), the gene expression profile after *M. loti* infection was markedly different from that in uninfected roots (Fig. 1B–D). The detection range of gene expression with the cDNA macroarray was judged to be around $10^4$.

To detect statistically significant up- and down-regulated genes, we adopted the Significance Analysis of Microarray (SAM) program.[19] In brief, the significance of the changes of expression was evaluated by scores assigned to each gene on the basis of "gene-specific" fluctuations in expression relative to the standard deviation of repeated measurements. The significant genes were selected as the differences in their expression levels from uninfected roots exceed the adjustable '$\delta$-values,' which was determined to minimize the expected false positive frequencies (the false discovery rate; FDR)[19] in each given comparison. In addition to this criterion, genes with signal intensities more than threefold and less than 1/3-fold compared to uninfected roots were listed as up-regulated (positive significant) and down-regulated (negative significant) genes, respectively. The results of the SAM analysis are summarized in Fig. 2. As a consequence, we listed 1,076 up-regulated and 277 down-regulated genes. It is noteworthy that the numbers of up-regulated genes at 2 and 4 days after inoculation were much larger than those in later stages of the nodulation process, indicating that very dynamic global changes in gene expression occur in early stages of *Rhizobium*-legume interactions. The total number of down-regulated genes was smaller than that of up-regulated genes, mainly because the majority of down-regulated genes exhibited low expression levels with relatively large fluctuations of the signal (see online supplemental data; URL, http://www.dnares.kazusa.or.jp/11/4/04/supplement_table/supplement.html).

### *3.3. Cluster analysis of expression profiles of up-regulated genes*

A total of 1,076 genes up-regulated during nodulation was further analyzed by hierarchical clustering on the basis of their expression levels relative to those in uninfected roots. Taken together with the results of the SAM analysis, we classified these genes into 5 clusters accord-

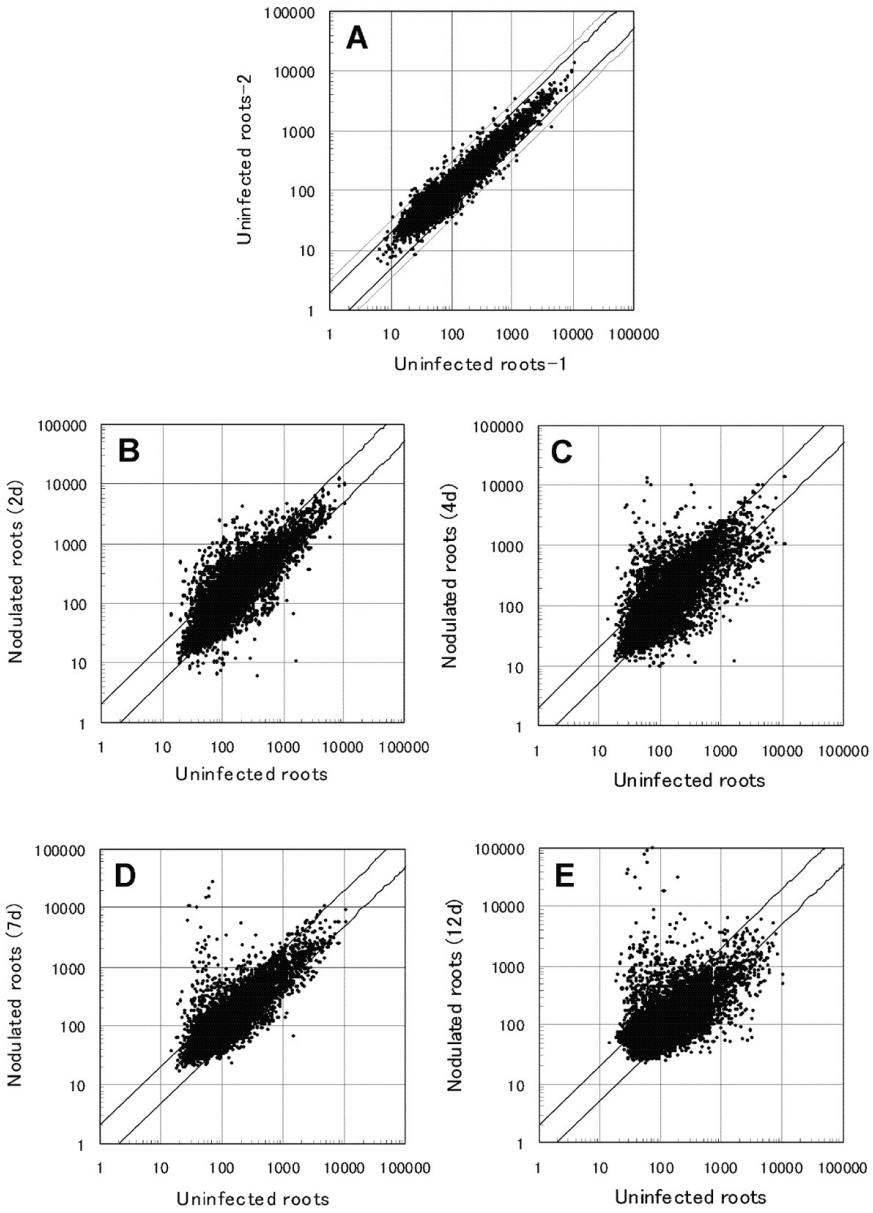



**Figure 1**. Scatter plots of relative gene expression levels in uninfected roots and root nodules of *Lotus japonicus*. The axes represent normalized expression levels in logarithmic scale. (A), Comparison of expression profiles with target RNAs from two independent uninfected roots (12 and 15 days after germination). The data were calculated as the averages from duplicate membrane hybridizations. Inner and outer lines represent twofold and threefold expression ratios, respectively. (B, C, D, and E), Comparisons of relative transcript levels of nodulated roots or nodules at day 2, 4, 7, and 12 after *M. loti* inoculation, respectively, with those of uninfected roots. The data were calculated as the averages of four replicate experiments. Lines indicate the twofold expression ratio. All data points in each figure represent 14,971 ESTs that showed signal intensities larger than three times the background intensity.

ing to their temporal expression patterns (Fig. 3). Genes in clusters 1 and 2 were induced transiently during early stages of infection and nodule initiation, and comprised 56% of total up-regulated genes. Their expression was significantly up-regulated 2 and/or 4 days after inoculation, but either recovered to the levels seen in uninfected roots or were down-regulated in the later stages of nodule development. Cluster 3 represents genes that were induced at a very early stage of infection and their expression levels further increased or remained high in the later stages. A well-characterized early nodulin gene, ENOD40 belongs to this cluster. Expression of genes in cluster 4 appeared to accompany nodule organogenesis because their expression was first induced at the time of nodule primordium formation and increased during further nodule development. Many known nodulin genes such as leghemoglobin were grouped in this cluster, indicating that expression of these nodule-specific genes are

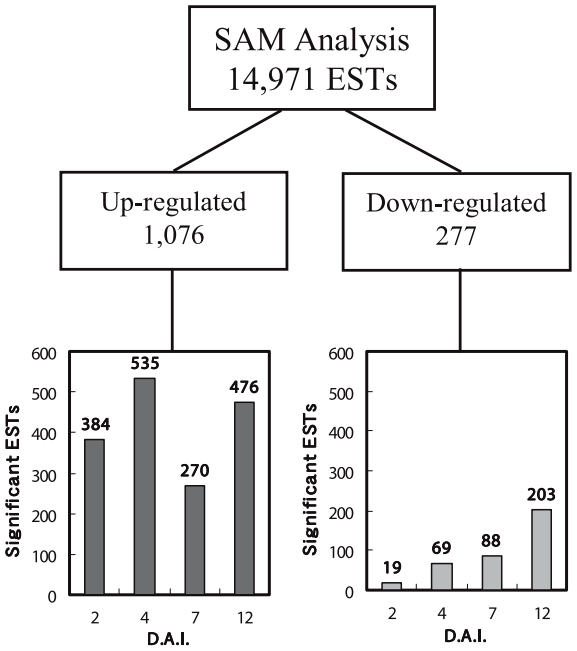



**Figure 2**. Summary of the significance analyses of genes differentially expressed during the nodulation process. The number of significant genes at each time point (D.A.I.=days after inoculation) as compared with uninfected roots is depicted. Details of selection criteria of up-regulated and down-regulated genes by the SAM software are described in the text. Figure 3.

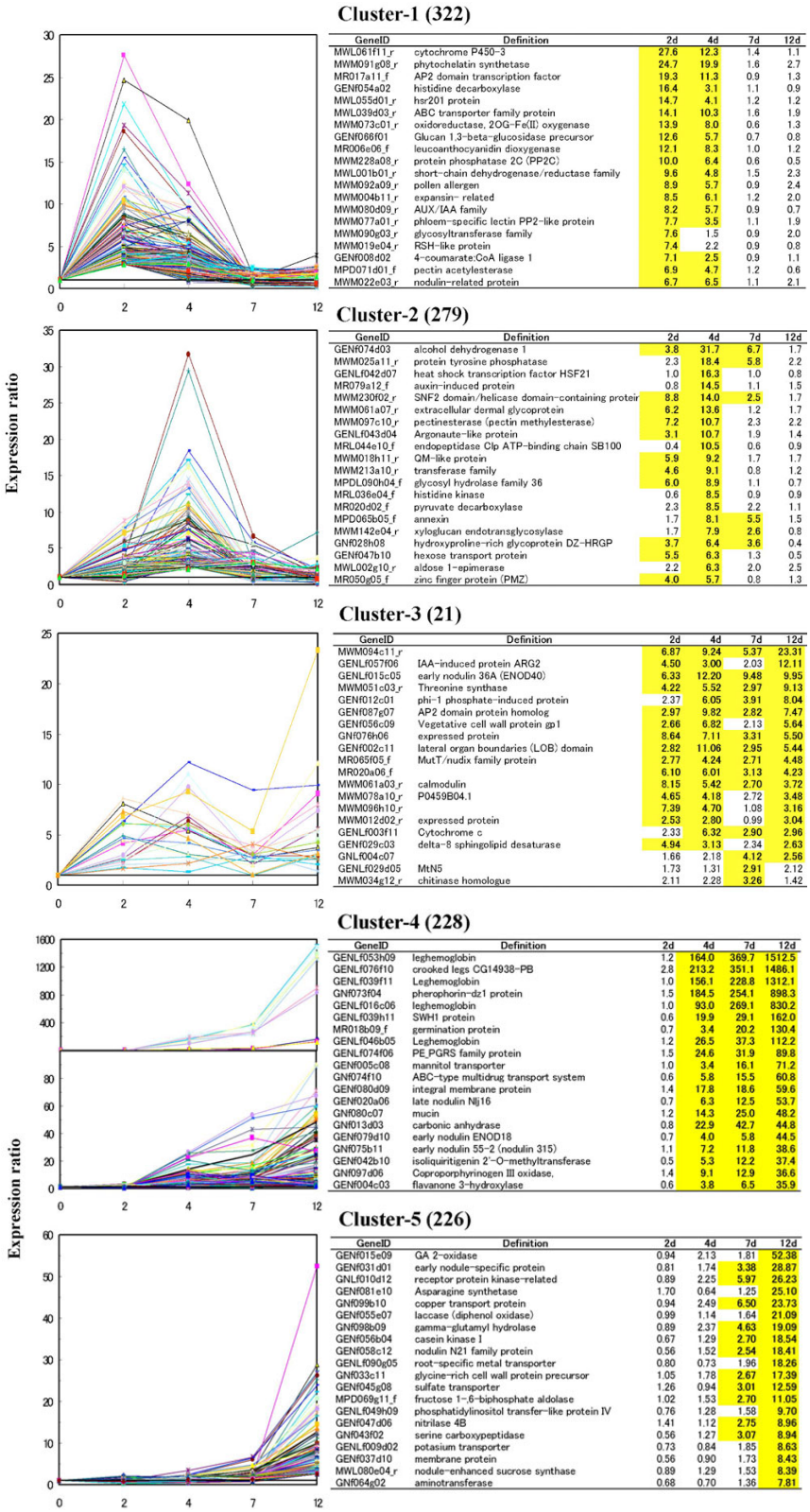
developmentally regulated. Finally, cluster 5 represents genes induced almost concomitantly with the onset of nitrogen fixation in fully developed nodules. Many genes encoding enzymes involved in primary carbon and nitrogen metabolism were found in cluster 5. The lists of the twenty most highly significant genes are also given for each cluster in Fig. 3 with their database annotations. The complete list of significant ESTs in each cluster with their relative expression ratios is available in the online supplemental data as well as in the web database at http://www.kazusa.or.jp/en/plant/lotus/EST/index.html.

### *3.4. Functional classification of up-regulated genes*

Among a total of 1,076 up-regulated genes, 710 (66.0%) ESTs displayed more or less similarity to sequences with known or putative functions that were deposited in a public database, 221 (20.5%) had similarities to 'hypothetical,' 'expressed,' or 'putative' genes with no defined function, and 145 (13.5%) had no similarity to deposited sequences. Table 2 presents a summary of the classification of the significant ESTs by their predicted biochemical functions according to MIPS (http://mips.gsf.de/proj/thal/db/index.html) with some modifications or by biological processes that they are involved in or respond to. In cluster 1, which represents genes transiently induced in the stage of infection, a category of cell rescue, defense, cell death and aging was most prominent, as comprising 20.7% of the ESTs with predicted functions. A total of 93 genes were assigned to this category over all time points and 74 (79.6%) of them were found in clusters 1 and 2, indicating that activation of genes in this functional category is mostly restricted in the early phase of *Rhizobium*-legume interactions. A number of genes involved in cell growth, cell division and DNA synthesis, as well as those of cell wall components or related to their synthesis, were found in clusters 2 through 4, reflecting the activity of cell division and new organ development during the stages of nodule primordium initiation and subsequent nodule organogenesis. A number of genes involved in phytohormone biosynthesis and responses were induced through the nodulation process. Among them, a gene encoding gibberellin-2-oxydase, that is responsible in inactivation of gibberellins, was strongly expressed in mature nodules in nearly a nodule-specific manner. However, gibberellin 2$\beta$-hydorolase and gibberellin-20-oxidase, which are involved in the final steps of the conversion of gibberellins to their biologically active forms, were also highly induced in the mature nodules (see the online supplemental data). Further analyses of spatial expression patterns of these genes in nodule tissues may provide more clues to solve these conflicting results.

Genes of enzymes involved in primary metabolism were found predominantly throughout all clusters. A number of nodule-specific or nodule-enhanced enzymes involved in photosynthate breakdown and assimilation of fixed nitrogen in the nodules was found in cluster 5, such as sucrose synthase (nodulin-100), [21] phosphoenolpyruvate carboxylase and its kinase,[22] nodule-enhanced malate dehydrogenase,[23] asparagine synthase,[24] glutamine synthase and aminotransferases. Genes with transport functions also comprised a prominent class over all stages of the nodulation process. This class contained previously identified nodule-specific or nodule-enhanced transporters, such as sulfate transporter[13] and sugar transporter.[25] In addition, a number of transporter genes with strong similarities to metal and ion transporters, peptide transporters, and mannitol transporters were found to be almost specific or significantly enhanced in nodules. Most of these nodule-specific and nodule-enhanced transporters are presumed to localize on symbiosome membranes in infected nodule cells. It would be interesting to determine the exact localization of these gene products and their physiological roles in functional symbiosis. More details of the significant gene list and their functional annotation are available in the online supplemental data.

The majority of genes included in the class of cell rescue, defense, cell death and aging was related to the defense response to pathogen attack. Representative defense-related genes are presented in Table 3. These genes included enzymes involved in phytoalexin biosynthesis (phenylalanine ammonia-lyases, 4-coumarate:CoA




| GeneID | Definition | 2d | 4d | 7d | 12d |
|---|---|---|---|---|---|
| MWL061f11_r | cytochrome P450-3 | 27.6 | 12.3 | 1.4 | 1.1 |
| MWM091g08_r | phytochelatin synthetase | 24.7 | 19.9 | 1.6 | 2.7 |
| MR017a11_f | AP2 domain transcription factor | 19.3 | 11.3 | 0.9 | 1.3 |
| GENf054a02 | histidine decarboxylase | 16.4 | 3.1 | 1.1 | 0.9 |
| MWL055d01_r | hsr201 protein | 14.7 | 4.1 | 1.2 | 1.2 |
| MWL039d03_r | ABC transporter family protein | 14.1 | 10.3 | 1.6 | 1.9 |
| MWM073c01_r | oxidoreductase, 2OG-Fe(II) oxygenase | 13.9 | 8.0 | 0.6 | 1.3 |
| GENf086f01 | Glucan 1,3-beta-glucosidase precursor | 12.6 | 5.7 | 0.7 | 0.8 |
| MR006e06_f | leucoanthocyanidin dioxygenase | 12.1 | 8.3 | 1.0 | 1.2 |
| MWM228a08_r | protein phosphatase 2C (PP2C) | 10.0 | 6.4 | 0.6 | 0.5 |
| MWL001b01_r | short-chain dehydrogenase/reductase family | 9.6 | 4.8 | 1.5 | 2.3 |
| MWM092a09_r | pollen allergen | 8.9 | 5.7 | 0.9 | 2.4 |
| MWM004b11_r | expansin- related | 8.5 | 6.1 | 1.2 | 2.0 |
| MWM080d09_r | AUX/IAA family | 8.2 | 5.7 | 0.9 | 0.7 |
| MWM077a01_r | phloem-specific lectin PP2-like protein | 7.7 | 3.5 | 1.1 | 1.9 |
| MWM090g03_r | glycosyltransferase family | 7.6 | 1.5 | 0.9 | 2.0 |
| MWM019e04_r | RSH-like protein | 7.4 | 2.2 | 0.9 | 0.8 |
| GENf008d02 | 4-coumarate:CoA ligase 1 | 7.1 | 2.5 | 0.9 | 1.1 |
| MPD071d01_f | pectin acetylesterase | 6.9 | 4.7 | 1.2 | 0.6 |
| MWM022e03_r | nodulin-related protein | 6.7 | 6.5 | 1.1 | 2.1 |

| GeneID | Definition | 2d | 4d | 7d | 12d |
|---|---|---|---|---|---|
| GENf074d03 | alcohol dehydrogenase 1 | 3.8 | 31.7 | 6.7 | 1.7 |
| MWM025a11_r | protein tyrosine phosphatase | 2.3 | 18.4 | 5.8 | 2.2 |
| GENLf042d07 | heat shock transcription factor HSF21 | 1.0 | 16.3 | 1.0 | 0.8 |
| MR079a12_f | auxin-induced protein | 0.8 | 14.5 | 1.1 | 1.5 |
| MWM230f02_r | SNF2 domain/helicase domain-containing protein | 8.8 | 14.0 | 2.5 | 1.7 |
| MWM061a07_r | extracellular dermal glycoprotein | 6.2 | 13.6 | 1.2 | 1.7 |
| MWM097c10_r | pectinesterase (pectin methylesterase) | 7.2 | 10.7 | 2.3 | 2.2 |
| GENLf043d04 | Argonaute-like protein | 3.1 | 10.7 | 1.9 | 1.4 |
| MRL044e10_f | endopeptidase Clp ATP-binding chain SB100 | 0.4 | 10.5 | 0.6 | 0.9 |
| MWM018h11_r | QM-like protein | 5.9 | 9.2 | 1.7 | 1.7 |
| MWM213a10_r | transferase family | 4.6 | 9.1 | 0.8 | 1.2 |
| MPDL090h04_f | glycosyl hydrolase family 36 | 6.0 | 8.9 | 1.1 | 0.7 |
| MRL036e04_f | histidine kinase | 0.6 | 8.5 | 0.9 | 0.9 |
| MR020d02_f | pyruvate decarboxylase | 2.3 | 8.5 | 2.2 | 1.1 |
| MPD065b05_f | annexin | 1.7 | 8.1 | 5.5 | 1.5 |
| MWM142e04_r | xyloglucan endotransglycosylase | 1.7 | 7.9 | 2.6 | 0.8 |
| GNf028h08 | hydroxyproline-rich glycoprotein DZ-HRGP | 3.7 | 6.4 | 3.6 | 0.4 |
| GENf047b10 | hexose transport protein | 5.5 | 6.3 | 1.3 | 0.5 |
| MWL002g10_r | aldose 1-epimerase | 2.2 | 6.3 | 2.0 | 2.5 |
| MR050g05_f | zinc finger protein (PMZ) | 4.0 | 5.7 | 0.8 | 1.3 |

| GeneID | Definition | 2d | 4d | 7d | 12d |
|---|---|---|---|---|---|
| MWM094c11_r | | 6.87 | 9.24 | 5.37 | 23.31 |
| GENLf057f06 | IAA-induced protein ARG2 | 4.50 | 3.00 | 2.03 | 12.11 |
| GENLf015c05 | early nodulin 36A (ENOD40) | 6.33 | 12.20 | 9.48 | 9.95 |
| MWM051c03_r | Threonine synthase | 4.22 | 5.52 | 2.97 | 9.13 |
| GENf012c01 | phi-1 phosphate-induced protein | 2.37 | 6.05 | 3.91 | 8.04 |
| GENf087g07 | AP2 domain protein homolog | 2.97 | 9.82 | 2.82 | 7.47 |
| GENf056c09 | Vegetative cell wall protein gp1 | 2.66 | 6.82 | 2.13 | 5.64 |
| GNf076h06 | expressed protein | 8.64 | 7.11 | 3.31 | 5.50 |
| GENf002c11 | lateral organ boundaries (LOB) domain | 2.82 | 11.06 | 2.95 | 5.44 |
| MR065f05_f | MutT/nudix family protein | 2.77 | 4.24 | 2.71 | 4.48 |
| MR020a06_f | | 6.10 | 6.01 | 3.13 | 4.23 |
| MWM061a03_r | calmodulin | 8.15 | 5.42 | 2.70 | 3.72 |
| MWM078a10_r | P0459B04.1 | 4.65 | 4.18 | 2.72 | 3.48 |
| MWM096h10_r | | 7.39 | 4.70 | 1.08 | 3.16 |
| MWM012d02_r | expressed protein | 2.53 | 2.80 | 0.99 | 3.04 |
| GENLf003f11 | Cytochrome c | 2.33 | 6.32 | 2.90 | 2.96 |
| GENf029c03 | delta-8 sphingolipid desaturase | 4.94 | 3.13 | 2.34 | 2.63 |
| GNLf004c07 | | 1.66 | 2.18 | 4.12 | 2.56 |
| GENLf029d05 | MtN5 | 1.73 | 1.31 | 2.91 | 2.12 |
| MWM034g12_r | chitinase homologue | 2.11 | 2.28 | 3.26 | 1.42 |

| GeneID | Definition | 2d | 4d | 7d | 12d |
|---|---|---|---|---|---|
| GENLf053h09 | leghemoglobin | 1.2 | 164.0 | 369.7 | 1512.5 |
| GENLf076f10 | crooked legs CG14938-PB | 2.8 | 213.2 | 351.1 | 1486.1 |
| GENLf039f11 | Leghemoglobin | 1.0 | 156.1 | 228.8 | 1312.1 |
| GNf073f04 | pherophorin-dz1 protein | 1.5 | 184.5 | 254.1 | 898.3 |
| GENLf016c06 | leghemoglobin | 1.0 | 93.0 | 269.1 | 830.2 |
| GENLf039h11 | SWH1 protein | 0.6 | 19.9 | 29.1 | 162.0 |
| MR018b09_f | germination protein | 0.7 | 3.4 | 20.2 | 130.4 |
| GENLf046b05 | Leghemoglobin | 1.2 | 26.5 | 37.3 | 112.2 |
| GENLf074f06 | PE_PGRS family protein | 1.5 | 24.6 | 31.9 | 89.8 |
| GENf005c08 | mannitol transporter | 1.0 | 3.4 | 16.1 | 71.2 |
| GNf074f10 | ABC-type multidrug transport system | 0.6 | 5.8 | 15.5 | 60.8 |
| GENf080d09 | integral membrane protein | 1.4 | 17.8 | 18.6 | 59.6 |
| GENf020a06 | late nodulin Nlj16 | 0.7 | 6.3 | 12.5 | 53.7 |
| GNf080c07 | mucin | 1.2 | 14.3 | 25.0 | 48.2 |
| GNf013d03 | carbonic anhydrase | 0.8 | 22.9 | 42.7 | 44.8 |
| GENf079d10 | early nodulin ENOD18 | 0.7 | 4.0 | 5.8 | 44.5 |
| GNf075b11 | early nodulin 55-2 (nodulin 315) | 1.1 | 7.2 | 11.8 | 38.6 |
| GENf042b10 | isoliquiritigenin 2'-O-methyltransferase | 0.5 | 5.3 | 12.2 | 37.4 |
| GNf097d06 | Coproporphyrinogen III oxidase, | 1.4 | 9.1 | 12.9 | 36.6 |
| GENf004c03 | flavanone 3-hydroxylase | 0.6 | 3.8 | 6.5 | 35.9 |

| GeneID | Definition | 2d | 4d | 7d | 12d |
|---|---|---|---|---|---|
| GENf015e09 | GA 2-oxidase | 0.94 | 2.13 | 1.81 | 52.38 |
| GENf031d01 | early nodule-specific protein | 0.81 | 1.74 | 3.38 | 28.87 |
| GNLf010d12 | receptor protein kinase-related | 0.89 | 2.25 | 5.97 | 26.23 |
| GENf081e10 | Asparagine synthetase | 1.70 | 0.64 | 1.25 | 25.10 |
| GNf099b10 | copper transport protein | 0.94 | 2.49 | 6.50 | 23.73 |
| GENf055e07 | laccase (diphenol oxidase) | 0.99 | 1.14 | 1.64 | 21.09 |
| GNf098b09 | gamma-glutamyl hydrolase | 0.89 | 2.37 | 4.63 | 19.09 |
| GENf056b04 | casein kinase I | 0.67 | 1.29 | 2.70 | 18.54 |
| GENf058c12 | nodulin N21 family protein | 0.56 | 1.52 | 2.54 | 18.41 |
| GENLf090g05 | root-specific metal transporter | 0.80 | 0.73 | 1.96 | 18.26 |
| GNf033c11 | glycine-rich cell wall protein precursor | 1.05 | 1.78 | 2.67 | 17.39 |
| GENf045g08 | sulfate transporter | 1.26 | 0.94 | 3.01 | 12.59 |
| MPD069g11_f | fructose 1-,6-biphosphate aldolase | 1.02 | 1.53 | 2.70 | 11.05 |
| GENLf049h09 | phosphatidylinositol transfer-like protein IV | 0.76 | 1.28 | 1.58 | 9.70 |
| GENf047d06 | nitrilase 4B | 1.41 | 1.12 | 2.75 | 8.96 |
| GNf043f02 | serine carboxypeptidase | 0.56 | 1.27 | 3.07 | 8.94 |
| GENLf009d02 | potasium transporter | 0.73 | 0.84 | 1.85 | 8.63 |
| GENf037d10 | membrane protein | 0.56 | 0.90 | 1.73 | 8.43 |
| MWL080e04_r | nodule-enhanced sucrose synthase | 0.89 | 1.29 | 1.53 | 8.39 |
| GNf064g02 | aminotransferase | 0.68 | 0.70 | 1.36 | 7.81 |

**Figure 3**. Expression profiles of genes differentially expressed during the nodulation process. The graphs show the expression patterns of significant genes for each cluster. The data are expressed as expression ratios (fold increase) relative to the uninfected roots. The total number of genes in each cluster is indicated in parenthesis. The twenty most highly significant genes are listed for each cluster. The values in yellow-shaded boldface were judged significant by SAM analysis. The complete list of genes in each cluster is available in the online supplemental tables.

**Table 2**. Functional classification of significantly up-regulated genes during the nodulation process. A complete list of genes in each functional group is available in the online supplemental data.

| Category | Infection (Cluster-1) | Nodule Primordia Initiation (Cluster-2) | Nodule Development (Cluster-3 & 4) | Fully Developed Nodules (Cluster-5) | Total |
|---|---|---|---|---|---|
| ***Primary Metabolism*** | **26** | **35** | **26** | **35** | **122** |
| ***Secondary Metabolism*** | **19** | **7** | **13** | **16** | **55** |
| ***Fatty Acid & Phospholipid Metabolism*** | **15** | **6** | **8** | **6** | **35** |
| ***Cell Wall*** | **9** | **17** | **6** | **4** | **36** |
| ***Cell Growth, Cell Division and DNA Synthesis*** | **3** | **12** | **10** | **4** | **29** |
| ***Nodulins*** | **4** | **3** | **15** | **10** | **32** |
| ***Phytohormone Synthesis and Response*** | **13** | **5** | **8** | **12** | **38** |
| ***Protein Fate*** | **9** | **8** | **8** | **6** | **31** |
| ***Cell Rescue, Defense, Cell Death and Ageing*** | **43** | **31** | **9** | **10** | **93** |
| ***Signal Transduction*** | **17** | **10** | **11** | **6** | **44** |
| ***Transport/Membrane*** | **18** | **12** | **18** | **17** | **65** |
| ***Transcription/Translation*** | **17** | **18** | **18** | **9** | **62** |
| ***Unclassified*** | **15** | **21** | **18** | **14** | **68** |
| ***Unknown Function*** | **59** | **55** | **54** | **53** | **221** |
| ***No Match*** | **54** | **40** | **27** | **24** | **145** |
| ***Total*** | **321** | **280** | **249** | **226** | **1,076** |

**Table 3**. Transient induction of the defense genes in the *M. loti* infection and nodule initiation stages. The data indicate the expression level relative to uninfected control roots (level=1). MaxRE is the greatest of the normalized expression levels during the nodulation process.

| GeneID | Definition | 2d | 4d | 7d | 12d | Max RE | E-value |
|---|---|---|---|---|---|---|---|
| ***Cluster-1*** | | | | | | | |
| MWM142a05_r | fungal endoglucanase inhibitor protein | **4.9** | 8.1 | 1.1 | 0.8 | 337.2 | 2E-036 |
| MWM103b11_r | proline-rich protein F26K10.180 | **2.9** | 1.0 | 0.8 | 0.4 | 332.8 | 2E-009 |
| GENf046f08 | glycine-rich protein | **3.5** | 1.3 | 0.9 | 1.5 | 454.6 | 1E-011 |
| MWM207h11_r | proline-rich cell wall protein | **2.7** | 1.4 | 0.9 | 1.5 | 210.7 | 8E-026 |
| GENf066f01 | glucan 1,3-beta-glucosidase precursor | **12.6** | **5.7** | 0.7 | 0.8 | 1311.3 | 2E-026 |
| MR050h02_f | 1-aminocyclopropane-1-carboxylate synthase 4 | 3.1 | **2.2** | **0.9** | 0.9 | 334.9 | 1E-068 |
| MR021d06_f | 12-oxophytodienoic acid 10, 11-reductase | **2.8** | 1.1 | 0.5 | 0.3 | 1001.7 | 1E-119 |
| MWM056d02_r | phenylalanine ammonia-lyase | **5.0** | 1.7 | 0.5 | 0.4 | 2945.8 | 4E-095 |
| MWL052f09_r | phenylalanine ammonia-lyase class II | **5.0** | 1.9 | 0.7 | 0.4 | 2970.1 | 1E-065 |
| MPD051d06_f | hsr203J homolog | **4.9** | **4.1** | 0.9 | 1.3 | 237.8 | 2E-099 |
| MR073a02_f | elicitor inducible gene product | **2.8** | **2.9** | 1.1 | 0.6 | 1079.3 | 4E-054 |
| GENLf020h11 | respiratory burst oxidase homolog | **4.0** | **3.6** | 1.2 | 1.2 | 545.7 | 4E-057 |
| MR092c12_f | syringolide-induced protein 1-3-1A | **5.3** | **4.1** | 0.7 | 0.8 | 640.7 | 2E-071 |
| MWM085h01_r | syringolide-induced protein 1-3-1B | **5.8** | **4.0** | 0.8 | 1.3 | 413.4 | 2E-088 |
| MWM143b11_r | allene oxide cyclase | **5.4** | **3.4** | 1.6 | 1.1 | 1080.4 | 2E-023 |
| MWM004e12_r | 12-oxophytodienoate reductase 3 | **6.0** | **3.1** | 0.8 | 1.4 | 303.2 | 4E-092 |
| GNf053e10 | beta-1,3-glucanase | **5.4** | **2.7** | 0.6 | 0.4 | 692.0 | 1E-034 |
| MWM060f08_r | chitinase class II | **5.8** | **3.8** | 0.9 | 1.9 | 500.7 | 1E-134 |
| MWL055d01_r | hsr201 protein | **14.7** | **4.1** | 1.2 | 1.2 | 777.3 | 9E-051 |
| GNf072h03 | disease resistance protein (CC-NBS-LRR class) | **4.3** | **5.2** | 1.1 | 0.8 | 518.2 | 5E-013 |
| GNf090d05 | chalcone reductase | **2.7** | 2.1 | 0.5 | 0.3 | 2717.0 | 2E-075 |
| MR006e06_f | leucoanthocyanidin dioxygenase | **12.1** | **8.3** | 1.0 | 1.2 | 564.0 | 1E-041 |
| GENf008d02 | 4-coumarate:CoA ligase 1 | **7.1** | **2.5** | 0.9 | 1.1 | 364.8 | 8E-090 |
| ***Cluster-2*** | | | | | | | |
| MWM012b07_r | polygalacturonase-inhibitor protein | 2.0 | **4.9** | 2.4 | 2.2 | 219.3 | 4E-068 |
| MWM061c04_r | proline-rich protein | 1.0 | **3.9** | 1.6 | 1.2 | 271.5 | 2E-035 |
| MWM225d01_r | S-adenosylmethionine synthetase 2 | **8.3** | **16.2** | 2.3 | **3.7** | 3030.1 | 1E-109 |
| MWM002h01_r | bacterial-induced peroxidase precursor | 1.6 | **4.1** | 2.0 | 1.3 | 323.1 | 2E-055 |
| MWM033e05_r | harpin-induced protein 1 | 2.4 | **3.4** | 1.0 | 1.9 | 1630.8 | 7E-014 |
| MWM074d06_r | patatin-related | 1.4 | **2.9** | 1.4 | 2.9 | 103.3 | 4E-041 |
| MR049h04_f | cationic peroxidase 2 precursor (PNPC2) | 1.5 | **3.4** | 1.3 | 1.2 | 940.6 | 6E-016 |
| GNf025c02 | class 5 chitinase | 0.4 | **3.4** | **3.6** | 2.0 | 406.2 | 3E-017 |
| MPD006d05_f | elicitor-responsive Dof protein ERDP | 1.1 | 1.8 | **2.7** | 0.6 | 3409.0 | 2E-013 |
| MPD008f11_f | elicitor inducible beta-1,3-glucanase | 0.8 | 2.0 | **2.6** | 1.7 | 127.5 | 1E-037 |

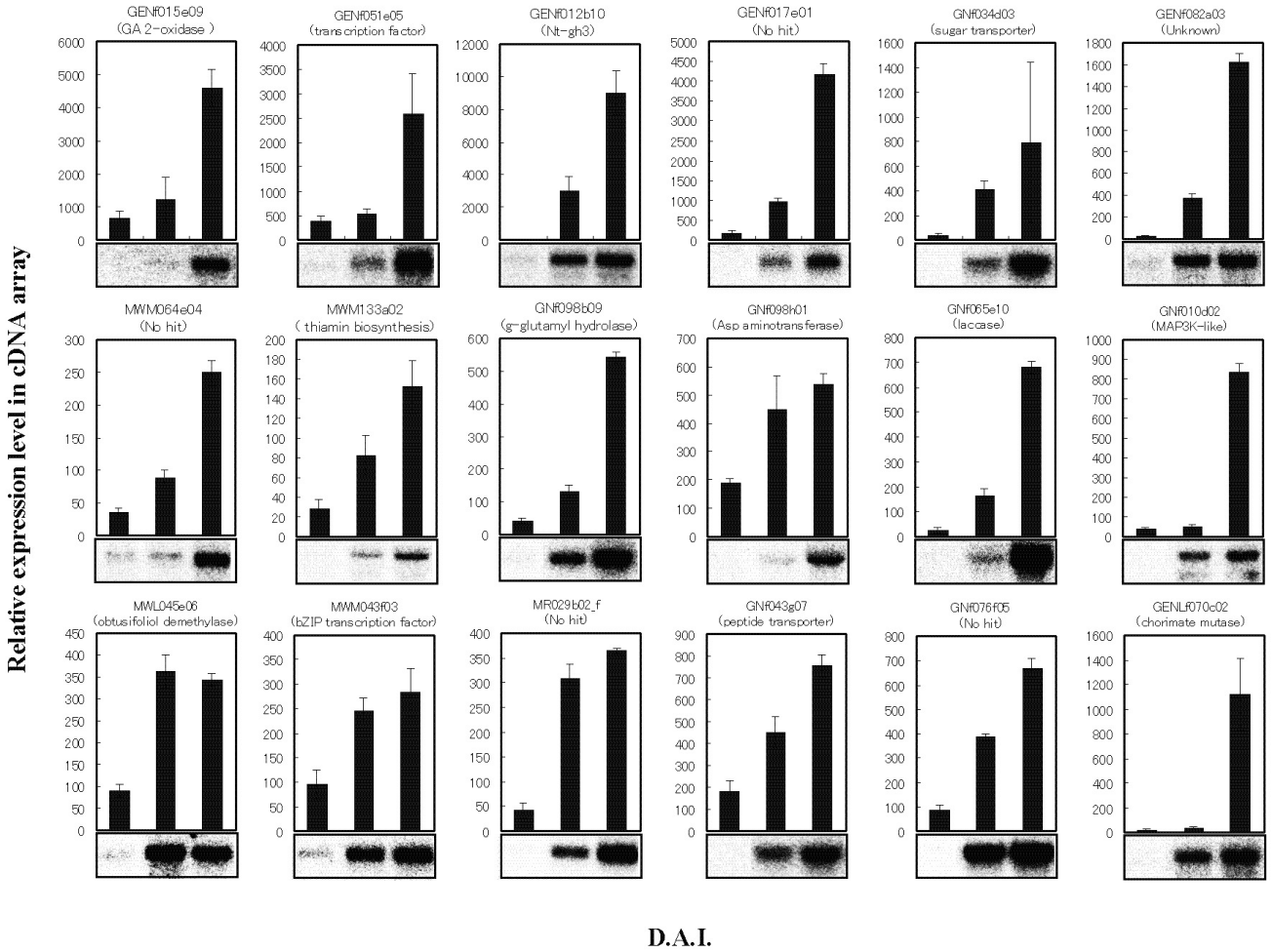



**Figure 4**. Comparison of relative expression levels in the cDNA array analysis with RNA gel blot analysis for representative novel nodule-enhanced genes. The y-axis indicates the normalized expression levels.

ligase, chalcone reductase), proteins involved in cell wall modification (glycine-rich proteins, hydroxyproline-rich proteins, polygalacturonase inhibitor proteins), PR-proteins (chitinase, $\beta$-1,3-glucanase, peroxidase), and hyper-sensitivity related (hsr) proteins. In particular, $\beta$-1,3-glucanase, hsr201 protein, leucoanthocyanidin dioxygenase, and S-adenosylmethionine synthetase were highly induced during the infection and nodule initiation stages. It is noteworthy that genes involved in jasmonate biosynthesis, such as allene oxide cyclase and 12-oxophytodienoate reductase, were highly induced in early infection stages. This finding is not surprising because the jasmonate signaling pathway is known to play crucial roles in defense gene expression in response to pathogen attack. In mature nodules, however, the expression levels of these genes were reduced, and genes encoding lipoxygenase, which catalyzes the initial step of jasmonate biosynthesis were the most significantly down-regulated genes (see the online supplemental data). Genes known to respond to other environmental stresses such as osmotic pressure, drought and dessication were also found to be significantly induced in the early stages of nodulation.

### *3.5. Quantification of gene expression by RNA gel blot and RT-PCR analyses*

To confirm the fidelity of the cDNA macroarray analysis in our study, selected genes were subjected to RNA gel blot and reverse transcription-polymerase chain reaction (RT-PCR) analyses. For RNA gel blot analysis, a total of 161 genes were selected from both up- and down-regulated genes to cover the wide range of expression levels, and RNAs for the gel blot analysis were prepared from the source plant materials independent of those used for the cDNA array analysis. The results for representative novel nodule-specific and nodule-enhanced genes are shown in Fig. 4, demonstrating that the expression profiles obtained by the two methods were fairly consistent. As for the genes induced transiently during the early stages of nodulation (clusters 1 and 2), we adopted RT-PCR analysis because of the limited amounts of RNA available. A total of 26 clones, including a number of defense-related genes, were selected from clusters 1 and

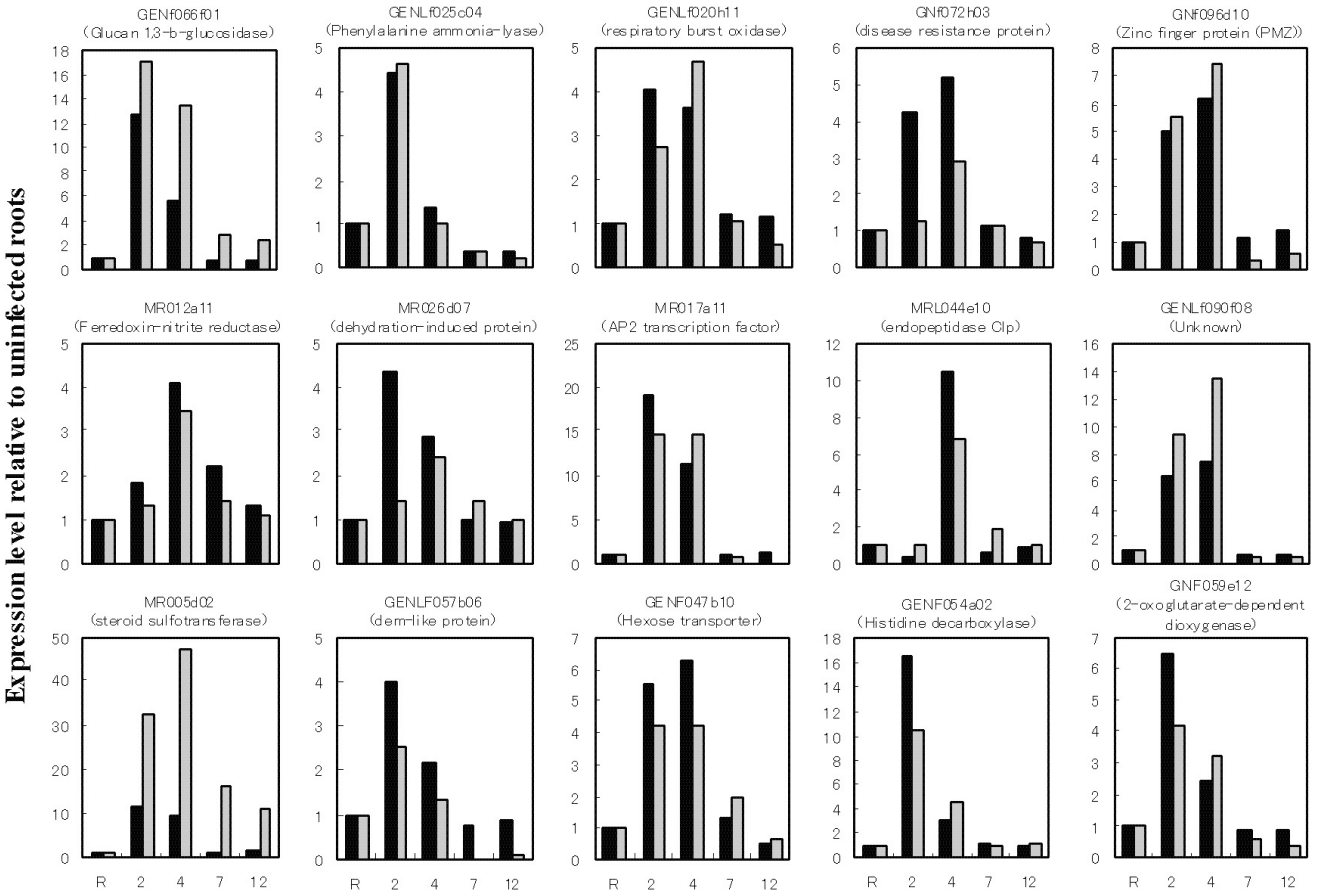



**Figure 5**. Comparison of relative expression levels in the cDNA array analysis (solid bars) with RT-PCR analysis (shaded bars) for representatives of genes transiently induced in the stages of infection (2 D.A.I.) and nodule primordium initiation (4 D.A.I.). The data are expressed as expression levels relative to uninfected roots (R).

2. The results (Fig. 5) are shown after quantification of the expression levels by Southern hybridization of the amplified DNA fragments on an agarose gel with $^{32}$P-labelled probes of each insert. Expression profiles along the time course correlated well between the two methods.

The overall data obtained by RNA gel blot and RT-PCR analyses were plotted against those of cDNA array analysis (Fig. 6). Although absolute values of the expression levels relative to uninfected roots differed markedly in some cases, the comparisons of over 400 expression data indicated a good correlation between cDNA array analysis and Northern or RT-PCR analyses (R=0.76). Detection of positive and negative significant genes by means of cDNA array was justified in about 80% of clones tested.

### *3.6. Concluding remarks and future perspectives*

Our results provide a comprehensive data source of gene expression profiles during early symbiotic process from *Rhizobium* infection to initiation of nitrogen-fixing symbiosis in developed nodules. One of the most intriguing findings from the analysis is that there was little overlap between a) genes that were markedly induced during the early stages of infection and nodule primordium initiation and b) those expressed abundantly in subsequent later stages of nodulation. Thus, the expression of the majority of positive significant genes in the initial period of the interaction was transient. Defense- and stress-related genes were a significant component in this group. It is noteworthy that this gene expression profile is in good agreement with that recently reported for arbuscular mycorrhizal (AM) symbiosis.[26] In time-course analysis of gene expression during AM-symbiosis in *Medicago trucatula*, the genes induced during interactions with AM fungi were divided into two clusters: one showed a sustained increase in transcripts as the symbiosis developed, and the other showed a transient increase in transcripts during the initial period of the interactions but a subsequent decrease as the symbiosis developed. The latter group contained a number of defense- and stress-response genes as a major component. Thus initial induction and subsequent rapid suppression of these defense- and stress-related genes are the common feature in both symbioses, and may not be involved directly in the specific interaction with rhizobia mediated by Nod factors. Compar-

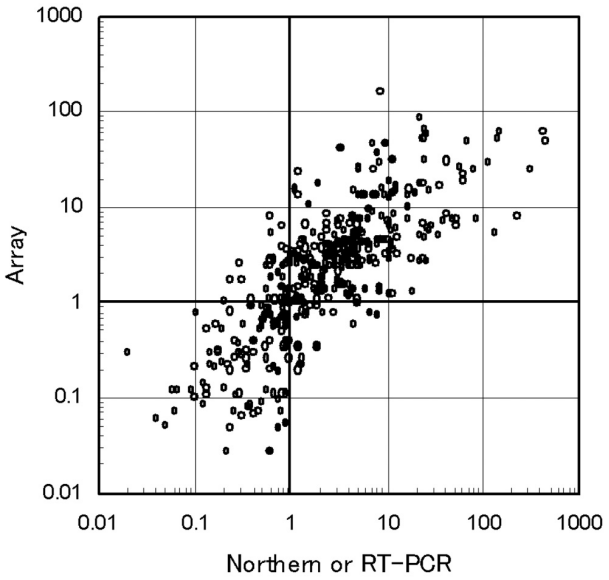



**Figure 6**. Correlation of expression profiles between the results from cDNA array and RNA gel blot (open circles) or RT-PCR (solid circles) analysis. The x- and y-axes are shown as logarithms of expression ratios to uninfected roots. The data points represent a total of 322 and 98 expression data for RNA gel blot and RT-PCR analysis, respectively. Quantification of RNA gel blot and RT-PCR results were performed by image analyzing with ubiquitin as a control (for details, see Materials and Methods).

isons with gene expression profiles after inoculation with *M. loti* that is defective in Nod factor production, as well as those in non-nodulating mutants of *L. japonicus*, may provide more insights about the possible involvement of individual transiently induced genes in clusters 1 and 2 in Nod factor signaling pathways.

Among the genes involved in clusters 4 and 5, those with more than several tenfold expression compared to uninfected roots are virtually nodule-specific, thus comprising novel nodulin genes. Their transcript levels were, in some cases, super-abundant in nodule tissues. Most of the genes represented in cluster 4 (Fig. 3) showed high expression ratios in 12-day nodules that were comparable with those of leghemoglobin genes. Analysis of the physiological functions of these novel nodule-specific genes in nodule formation and/or in functional symbiosis is an intriguing subject for future study. Systematic analysis of functions of novel nodulin genes identified from the cDNA array analysis is currently underway by means of a transient RNA silencing strategy.[27]

**Acknowledgements:** The authors express their thanks to K. Tejima and T. Hakoyama of Tokyo University of Agriculture and Technology; N. Ohtake and H. Fujikake of Niigata University; T. Nakagawa, M. Banba, K. Nakamori and Y. Ooki of Kyoto University; T. Akashi, Y. Sawada, N. Shimada and R. Imaizumi of Nihon University; T. Maekawa of Osaka University; T. Kato and T. Okada of Aichi University of Education; T. Shitaohta of Kagoshima University; and A. Kodera of Kagawa University for their technical assistance in cDNA array construction. The authors also thank Prof. Robert W. Ridge of International Christian University for critical reading of the manuscript. This work was supported by the Special Coordination Funds for promoting Science and Technology from the Ministry of Education, Culture, Sports, Science and Technology, Japan.

Online supplement data are available on a WEB site http: // www.dna-res.kazusa.or.jp/11/4/04/supplement_table/supplement.html.

## References

1. Schultze, M. and Kondorosi, A. 1998, Regulation of symbiotic root nodule development, *Annu. Rev. Genet.*, **32**, 33–57.
2. Spaink, H. P. 2000, Root nodulation and infection factors produced by rhizobial bacteria, *Annu. Rev. Microbiol.*, **54**, 257–288.
3. Handberg, K. and Stougaard, J. 1992, Lotus japonicus, an autogamous, diploid legume species for classical and molecular genetics, *Plant J.*, **2**, 487–496.
4. Kawaguchi, M., Imaizumi-Anraku, H., Koiwa, H. et al. 2002, Root, root hair, and symbiotic mutants of the model legume *Lotus japonicus*, *Mol. Plant-Microbe Interact.*, **15**, 17–26.
5. Schauser, L., Handberg, K., Sandal, N. et al. 1998, Symbiotic mutants deficient in nodule establishment identified after T-DNA transformation of *Lotus japonicus*, *Mol. Gen. Genet.*, **259**, 414–423.
6. Szczyglowski, K., Shaw, R. S., Wopereis, J. et al. 1998, Nodule organogenesis and symbiotic mutants of the model legume *Lotus japonicus.*, *Mol. Plant-Microbe Interact.*, **11**, 684–697.
7. Asamizu, E., Nakamura, Y., Sato, S., and Tabata, S. 2000, Generation of 7137 Non-redundant Expressed Sequence Tags from a Legume, *Lotus japonicus*, *DNA Res.*, **7**, 127–130.
8. Asamizu, E., Nakamura, Y., Sato, S., and Tabata, S. 2004, Characteristics of the *Lotus japonicus* Gene Repertoire Deduced from Large-Scale Expressed Sequence Tag (EST) Analysis, *Plant Mol. Biol.*, (in press)
9. Hayashi, M., Miyahara, A., Sato, S. et al. 2001, Construction of a genetic linkage map of the model legume Lotus japonicus using an intraspecific F2 population, *DNA Res.*, **8**, 301–310.
10. Radutoiu, S., Madsen, L. H., Madsen, E. B. et al. 2003, Plant recognition of symbiotic bacteria requires two LysM receptor-like kinases, *Nature*, **425**, 585–592.
11. Schauser, L., Roussis, A., Stiller, J., and Stougaard, J. 1999, A plant regulator controlling development of symbiotic root nodules, *Nature*, **402**, 191–195.
12. 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.
13. Colebatch, G., Kloska, S., Trevaskis, B., Freund, S., Alt-

mann, T., and Udvardi, M. K. 2002, Novel aspects of symbiotic nitrogen fixation uncovered by transcript profiling with cDNA arrays, *Mol. Plant-Microbe Interact.*, **15**, 411–420.
14. Fedorova, M., van de Mortel, J., Matsumoto, P. A. et al. 2002, Genome-wide identification of nodule-specific transcripts in the model legume *Medicago truncatula*, *Plant Physiol.*, **130**, 519–537.
15. Niwa, S., Kawaguchi, M., Imazumi-Anraku, H. et al. 2001, Responses of a model legume *Lotus japonicus* to lipochitin oligosaccharide nodulation factors purified from *Mesorhizobium loti* JRL501, *Mol. Plant-Microbe Interact.*, **14**, 848–856.
16. Kawaguchi, M. 2000, *Lotus japonicus* 'Miyakojima' MG-20: An early flowering accession suitable for indoor genetics, *J. Plant Res.*, **113**, 507–509.
17. Sambrook, J. and Russel, D. W. 2001, Molecular Cloning: A Laboratory Manual, 3rd Ed. Cold Spring Harbor, NY.
18. Sasaki, Y., Asamizu, E., Shibata, D. et al. 2001, Monitoring of methyl jasmonate-responsive genes in *Arabidopsis* by cDNA macroarray: self-activation of jasmonic acid biosynthesis and crosstalk with other phytohormone signaling pathways, *DNA Res.*, **8**, 153–161.
19. Tusher, V. G., Tibshirani, R., and Chu, G. 2001, Significance analysis of microarrays applied to the ionizing radiation response, *Proc. Natl. Acad. Sci. USA*, **98**, 5116–5121.
20. Nishimura, R., Hayashi, M., Wu, G.-J. et al. 2002, HAR1 mediates systemic regulation of symbiotic organ development, *Nature*, **420**, 426–429.
21. Zhang, X. Q., Lund, A. A., Sarath,G. et al. 1999, Soybean nodule sucrose synthase (Nodulin-100): Further analysis of its phosphorylation using recombinant and authentic root-nodule enzymes, *Arch. Biochem. Biophys.*, **371**, 70–82.
22. Nakagawa, T., Izumi, T., Banba, M. et al. 2003, Characterization and expression analysis of genes encoding phosphoenolpyruvate carboxylase and phosphoenolpyruvate carboxylase kinase of *Lotus japonicus*, a model legume, *Mol. Plant-Microbe Interact.*, **16**, 281–288.
23. Miller, S. S., Driscoll, B. T., Gregerson, R. G. et al. 1998, Alfalfa malate dehydrogenase (MDH): molecular cloning and characterization of five different forms reveals a unique nodule-enhanced MDH, *Plant J.*, **15**, 173–184.
24. Shi, L., Twary, S. N., Yoshioka, H. et al. 1997, Nitrogen assimilation in alfalfa: isolation and characterization of an asparagine synthetase gene showing enhanced expression in root nodules and dark-adapted leaves, *Plant Cell*, **9**, 1339–1356.
25. Flemetakis, E., Dimou, M., Cotzur, D. et al. 2003, A sucrose transporter, LjSUT4, is up-regulated during *Lotus japonicus* nodule development, *J. Exp. Bot.*, **54**, 1789–1791.
26. Liu, J., Blaylock, L. A., Endre, G. et al. 2003, Transcript profiling coupled with spatial expression analyses reveals genes involved in distinct developmental stages of an arbuscular mycorrhizal symbiosis, *Plant Cell*, **15**, 2106–2123.
27. Kumagai, H. and Kouchi, H. 2003, Gene silencing by expression of hairpin RNA in *Lotus japonicus* roots and root nodules. *Mol. Plant-Microbe Interact.*, **16**, 663–668.