

# SCIENTIFIC REPORTS



OPEN

## Transcriptome Profiles of Nod Factor-independent Symbiosis in the Tropical Legume *Aeschynomene evenia*

Djamel Gully<sup>1</sup>, Pierre Czernic<sup>2</sup>, Stéphane Cruveiller<sup>3</sup>, Frédéric Mahé<sup>1</sup>, Cyrille Longin<sup>3</sup>, David Vallenet<sup>3</sup>, Philippe François<sup>1</sup>, Sabine Nidelet<sup>4</sup>, Stéphanie Rialle<sup>4</sup>, Eric Giraud<sup>1</sup>, Jean-François Arrighi<sup>1</sup>, Maitrayee DasGupta<sup>5</sup> & Fabienne Cartieaux<sup>1</sup>

Nod factors (NF) were assumed to be indispensable for the establishment of a rhizobium-legume symbiosis until the discovery that certain *Bradyrhizobium* strains interacting with certain *Aeschynomene* species lack the canonical *nodABC* genes required for their synthesis. So far, the molecular dialogue between *Aeschynomene* and its symbionts remains an open question. Here we report a time course transcriptional analysis of *Aeschynomene evenia* in response to inoculation with *Bradyrhizobium* ORS278. The NF-independent symbiotic process was monitored at five time points between bacterial infection and nodule maturity. The five time points correspond to three specific events, root infection by crack entry, nodule organogenesis, and the establishment of the nitrogen fixing process. During the third stage, about 80 NCR-like genes and eight symbiotic genes known to be involved in signaling, bacterial infection or nodulation regulation were highly expressed. Comparative gene expression analyses at the five time points also enabled the selection of genes with an expression profile that makes them promising markers to monitor early plant responses to bacteria. Such markers could be used in bioassays to identify the nature of the bacterial signal(s). Our data represent valuable resources for investigation of this Nod factor-independent symbiosis.

Legumes and cereals are among the earliest plants domesticated by human beings. Beans have been cultivated since the Neolithic in South America and lentils were already well known and consumed in Antiquity in Europe<sup>1</sup>. In addition to their use as human food, legumes such as alfalfa, clover and soybean have been widely used as forage since humans started breeding animals. Legumes have a beneficial effect on human health and nutrition thanks to their high protein content. Legumes also help improve soil fertility in crop rotations or intercropping systems. These two traits rely on the ability of legumes to biologically fix nitrogen. Legumes, in symbiosis with bacteria, collectively known as rhizobia, are able to convert atmospheric dinitrogen into ammonia, which is then assimilated by the host plant. Rhizobium-legume symbiosis is a win-win association in which bacteria fix nitrogen for the benefit of the plant in exchange for dicarboxylic acids resulting from photosynthesis. Nitrogen fixation takes place in a dedicated symbiotic organ, a nodule, which develops on the roots of the host plant and hosts the bacteria.

Two model legumes emerged in studies of biological nitrogen fixation in the 1990s. The need for two models was because the process of nodule organogenesis differs in *Medicago truncatula* and *Lotus japonicus*: in the former, nodules possess an active meristem throughout their life, which leads to an elongated organ, while in *L. japonicus*, meristematic activity is transient and gives rise to a spherical nodule. Apart from their meristematic nodule activity, *M. truncatula* and *L. japonicus* share a common bacterial infection process and a very similar genetic program recruited from the more ancient and widespread mycorrhizal symbiosis<sup>2,3</sup>.

<sup>1</sup>LSTM, Univ. Montpellier, CIRAD, INRA, IRD, SupAgro, Montpellier, France. <sup>2</sup>Université de Montpellier, Place Eugène Bataillon, F-34095, Montpellier Cedex 5, France. <sup>3</sup>LABGeM, Génomique Métabolique, Genoscope, Institut François Jacob, CEA, CNRS, Univ Evry, Université Paris-Saclay, F-91057, Evry, France. <sup>4</sup>MGX, Univ. Montpellier, CNRS, INSERM, BioCampus, Montpellier, France. <sup>5</sup>Department of Biochemistry, University of Calcutta, Kolkata, 700019, India. Correspondence and requests for materials should be addressed to F.C. (email: [fabienne.cartieaux@ird.fr](mailto:fabienne.cartieaux@ird.fr))

The interaction between legumes and rhizobia relies on a sophisticated recognition system. Legumes first produce flavonoids that are perceived by rhizobia and induce the expression of *nod* genes involved in the synthesis and secretion of lipochitooligosaccharide signals (called Nod factors, NF). In turn, the perception and recognition of NFs by the host plant trigger a series of responses involved in nodule inception and infection. In most legumes, including model legumes, rhizobia invade plant tissues in an intracellular way through infection threads (IT) consisting in an invagination of the plasma membrane initiated at the tip of a deformed root hair<sup>2,3</sup>.

However, in about 25% of legumes found mainly in tropical and warm temperate areas, rhizobial invasion does not involve the formation of infection threads but an intercellular process<sup>4</sup>. In a few rare cases, the symbiotic interaction may even take place despite the absence of Nod factors. This is the case of some *Aeschynomene* species that are nodulated by photosynthetic *Bradyrhizobium* strains lacking the canonical *nodABC* genes required for NF synthesis<sup>5</sup>. This NF independent infection process is initiated when the root is penetrated through epidermal cracks at the lateral root emergence sites<sup>6</sup>. Subsequently, sub-epidermal cell death occurs that enables the propagation of the bacteria within the root endodermis. Finally, the infection of one or several cortical cells by direct endocytosis of rhizobia, triggers their division and the development of a nodule that becomes functional (i.e. able to fix atmospheric nitrogen) five days after inoculation<sup>6</sup>. Thanks to studies of the genetic diversity of *Aeschynomene* species that use an NF-independent process, *Aeschynomene evenia* has become a new model legume for the purpose of deciphering the molecular mechanisms of the NF-independent symbiosis<sup>7</sup>. Using this new model species, we recently demonstrated that two common symbiotic genes<sup>8,9</sup> (*SYMRK* and *CCaMK*) and a cytokinin receptor involved in nodule organogenesis<sup>10</sup> (*HKT1*) were recruited for efficient nodulation in *A. evenia*<sup>11</sup>. Unexpectedly, the mechanisms also appeared to be conserved at the nodule maturation level. Indeed, although *Aeschynomene* are not close to the IRLC clade, some species - including *A. evenia* - use cysteine rich peptides similar to NCR to control bacterial differentiation in nitrogen fixing bacteroids<sup>12</sup>.

Therefore, the most enigmatic characteristic of *Aeschynomene* model system remains the absence of the need for NFs to induce the symbiosis. The plant and bacterial signal molecule(s) that control the establishment of the symbiosis between *A. evenia* and its symbiont remain to be identified. To advance our understanding of the initial steps of the interaction, we performed RNAseq analysis during a complete time course of the infection process starting as early as 6 hours post inoculation to 6 days post infection when the nodules have just become functional. The RNAseq data were confirmed by RTqPCR analyses for selected genes. Identification of differentially expressed genes (DEGs) at the different time points, including genes known to be involved in the symbiotic process of model legumes, was consistent with the general process of nodule organogenesis, development and functioning. Special focus was also placed on genes whose early stable induction following inoculation makes them suitable for a reporter gene strategy.

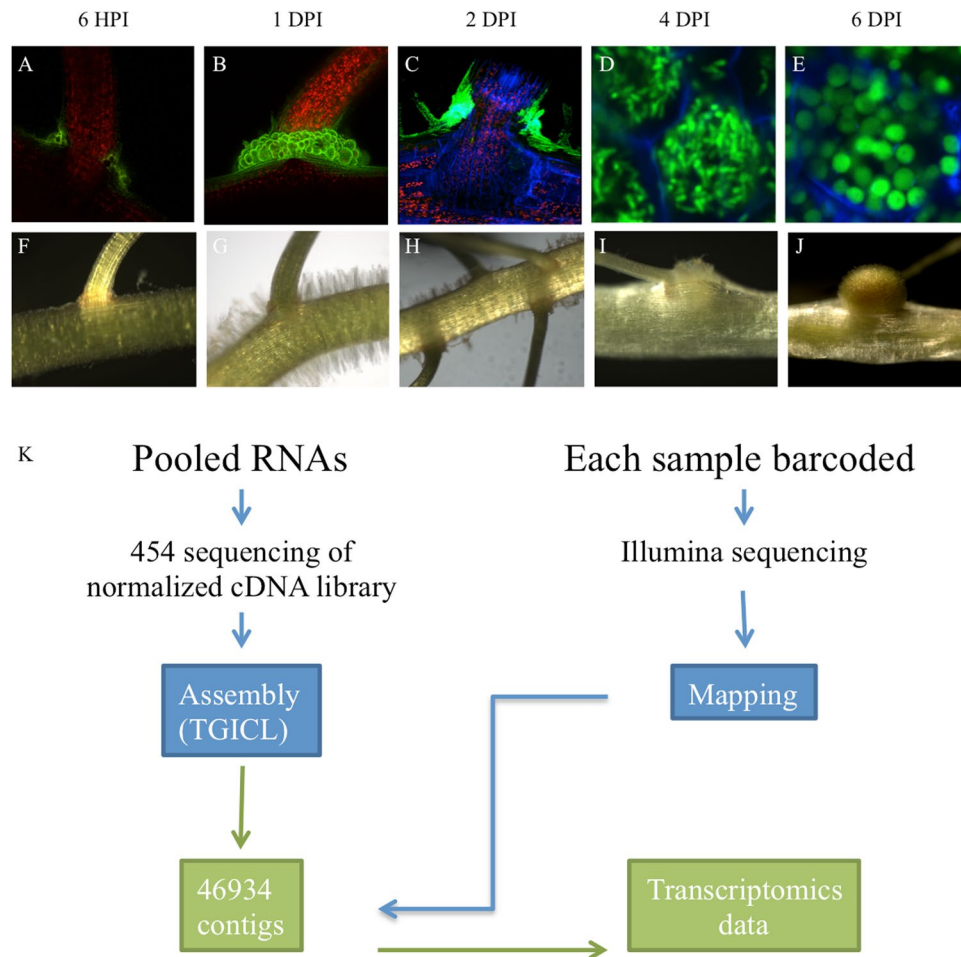
## Results and Discussion

**RNA sampling according to the symbiotic infection process.** To decipher the molecular mechanisms underlying the NF-independent symbiotic process in *A. evenia* spp. *serrulata*, the accession used for the functional characterization of four symbiotic genes in *Aeschynomene*<sup>11,12</sup>, we prepared RNA samples from roots inoculated or not with *Bradyrhizobium* sp. ORS278. Roots were collected before inoculation (T0), 6 hours post-inoculation (hpi), 1 day post-inoculation (dpi), 2, 4 and 6 dpi. These time points correspond to the different steps of infection and nodule development already described in *A. evenia*<sup>7</sup>: we collected roots at 6 hpi to analyze early responses to infection before any macroscopic changes occurred, except bacterial colonization of lateral roots (Fig. 1A,F); as early as 1 dpi, dense colonization of the axillary root hairs surrounding the lateral roots was observed (Fig. 1B,G); at 2 dpi, a few cortical cells had been invaded and began to divide (Fig. 1C,H); at 4 dpi the nodule primordium was clearly visible (Fig. 1I) and cells infected with undifferentiated rod-shaped bacteria were visible (Fig. 1D); at 6 dpi, bacteria were differentiated into spherical bacteroids (Fig. 1E), the nodule was mature, as evidenced by the pink color due to the accumulation of leghemoglobin (Fig. 1J) and nitrogenase activity was detectable<sup>6</sup>.

**EST sequencing and clustering, sequence annotation and database creation.** A normalized cDNA library from *A. evenia* tissues corresponding to non-inoculated roots, roots inoculated with *Bradyrhizobium* sp. ORS278 and collected 6 hpi, 2 dpi, 4 dpi, 7 dpi, mature nodules and leaves harvested 21 dpi, was pyrosequenced to obtain a reference set of transcripts. A total of 1,488,532 raw reads with an average length of 297 nucleotides corresponding to a total of 442 Mb were obtained (Table 1). The sequences were analyzed and clustered using a modified version of a pipeline called Expressed Sequence Tag treatment and investigation kit<sup>13</sup> (ESTtik). A total of 46,934 *A. evenia* spp. *serrulata* EST contigs were assembled, of which 71% were blast annotated. The annotation results revealed that 29,120 contigs (61.3%) had significant matches in the TrEMBL database, 1,719 (3.61%) in the Nt database, 1,682 (3.54%) in the Nr database and 1,205 (2.53%) in the SWISS-PROT database. These *A. evenia* EST sequences are accessible on the SESAM web site: <https://www.genoscope.cns.fr/agc/sesam>.

In parallel, the six different non-normalized cDNA libraries corresponding to the infection kinetic time points (T0, 6 hpi, 1 dpi, 2 dpi, 4 dpi, 6 dpi) were individually sequenced in triplicate using HiSeq2000 technology (Illumina Inc., USA). A total of 327,152,229 reads with a length of 100 nucleotides corresponding to 32,715 MB were obtained and mapped onto the contig clusters generated from the 454 sequencing data using the ssaha2 package (Table 1; Fig. 1K). The expression level of each assembled transcript sequence in different samples was measured using DESeq. 1 statistical analysis.

**Identification of differentially expressed genes.** Differentially expressed genes (DEG) in one sample compared to the non-inoculated root were defined using the following parameters: adjusted p-value < 0.05 and log<sub>2</sub> fold change > 2. In total, 1276 genes were found to be up-regulated during the infection kinetics and 712



**Figure 1.** Experimental design and description of root samples collected at the different time points during the nodulation kinetics from non-inoculated roots (T0, control roots) to 6 days post inoculation (dpi) with *Bradyrhizobium* ORS278. (A–E) Confocal microscopy images of the root system inoculated with a GFP-tagged *Bradyrhizobium* strain to monitor the infection and differentiation processes. (F–J) Bright-field macroscopic pictures of the roots. (A,F) Roots 6 hours post inoculation: the initial steps of perception. (B,G) Roots 1 dpi: colonization of an emerging lateral root; (C,H) roots 2 dpi: sub-epidermal infection; (D,I) roots 4 dpi: immature nodule colonized by undifferentiated bacteria; (E,J) roots 6 dpi: fixing nodule colonized by differentiated rhizobia, adapted from Arrighi *et al.*<sup>7</sup>. (K) RNAseq strategy: 454 sequences were clustered and assembled using TGICL to generate a transcriptome reference library. Illumina reads were mapped to 46,934 concatenated contigs of the transcriptome reference before statistical analyses were performed and differentially expressed gene patterns were obtained.

genes were found to be down regulated. Focusing on DEGs specific to each stage of the infection process (Fig. 2) led to the identification of 1,032 up-regulated genes (548 functionally annotated, Supplementary Table 1) and 604 down-regulated genes (448 functionally annotated, Supplementary Table 1).

Among the stage-specific DEGs, most were identified in nitrogen-fixing nodules (6 dpi), where 954 and 470 genes were found to be up- and down-regulated, respectively; and to a lesser extent in the initial step of the colonization (6 hpi, 59 and 113 genes up- and down-regulated, respectively). Only a few number of genes appeared to be specific to the intermediate time points (47 up- and 50 down-regulated genes for 1, 2 and 4 dpi combined; Fig. 2A). These observations were corroborated through hierarchical clustering analysis of the different time points. The two first (6 hpi and 1 dpi) and the two intermediate (2 and 4 dpi) time points appeared to cluster together, respectively, whereas the late point (6 dpi) corresponding to mature nodule behaved independently (Supplementary Fig. 1). We therefore focused our analyses on the early and late time points and performed a GO annotation for the biological processes involved (Fig. 2B,C,E and F).

**Early responses.** At 6 hpi, the majority of DEGs were involved in metabolic, biosynthetic and single organism processes (metabolic, cellular, signaling). These processes probably correspond to the induction of the signaling pathways leading to the shift from a developing root to nodule organogenesis.

A smaller number of DEGs were linked to biological processes corresponding to “Response to Stress” and “Abiotic Stimuli”. More detailed analysis of the genes specifically induced at 6 hpi highlighted two genes encoding

	Samples	Number of reads	Number of bases	Mapped reads	Non-ambiguous mapping rate (%)
Illumina raw data	T0	52 300 676	5 230 067 567	48 067 215	91,91
	6HPI	51 082 419	5 108 241 867	46 906 291	91,82
	24HPI	48 800 573	4 880 057 300	44 793 759	91,79
	2DPI	55 930 717	5 593 071 700	51 898 744	92,79
	4DPI	61 361 611	6 136 161 133	51 898 744	84,58
	6DPI	57 676 234	5 767 623 367	53 456 518	92,68
	Total	327 152 229	32 715 222 933	297 021 271	90,79
Raw 454 data		1 488 532	442 226 657		

**Table 1.** Raw data from Illumina and 454 sequencing.

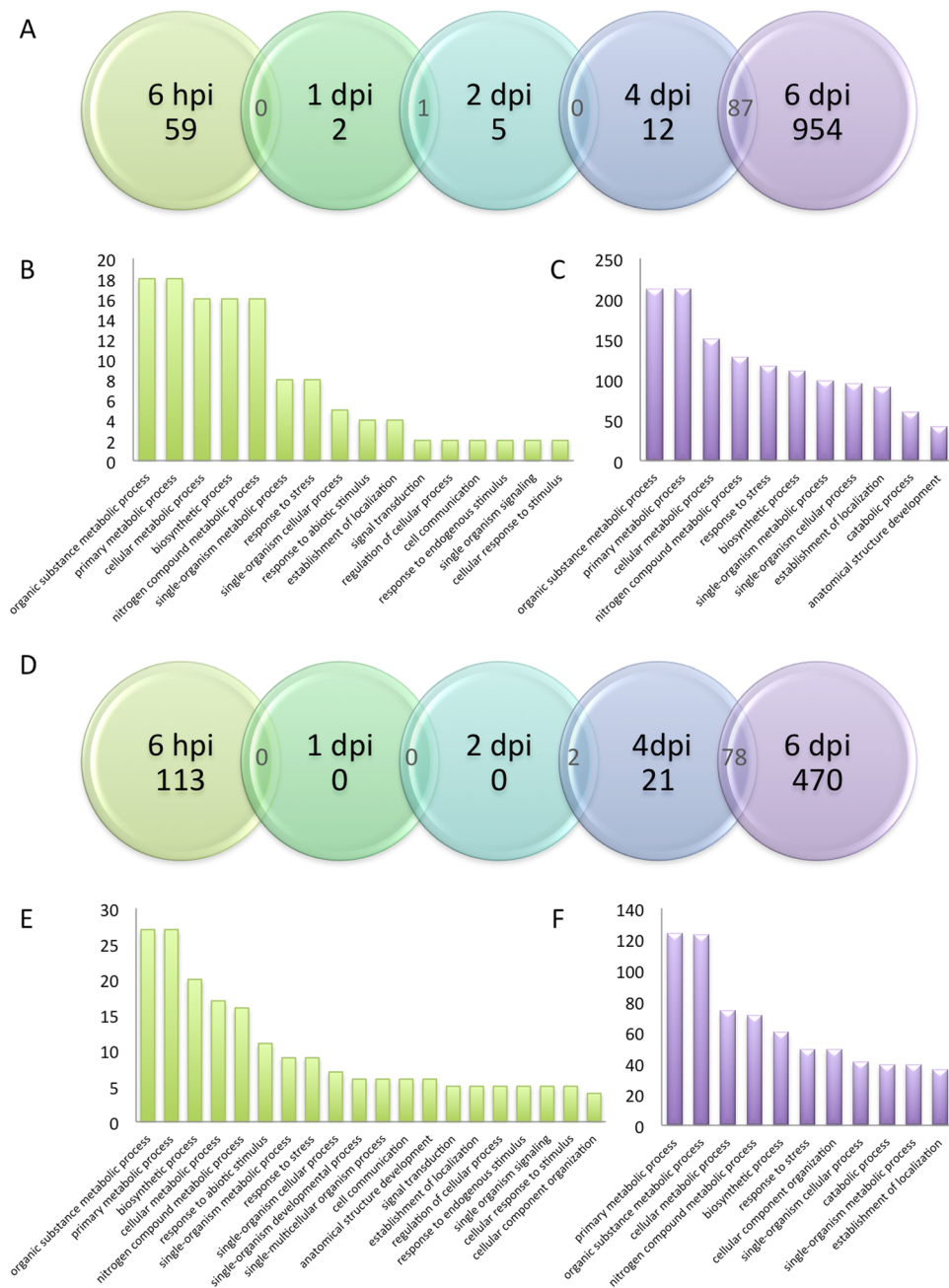
isochorismatases (CL6312Contig1 & 2; Supplementary Table 1; Fig. 3C). Isochorismatases are able to deplete isochorismate, the precursor in the synthesis pathway of the main plant defense hormone, salicylic acid<sup>14</sup>. It was recently reported that such enzyme activity was used by both an oomycete and a fungal pathogen to interfere with salicylic acid signaling and to reduce defense responses<sup>15</sup>. In order to confirm this observation, we focused on several defense-related genes previously shown to be transiently induced in model legumes such as medicago<sup>16,17</sup>, lotus<sup>18</sup> or soybean<sup>19</sup>. First, we identified the corresponding *Aeschynomene* orthologs (Fig. 3A) and then analyzed their expression profiles. As can be seen in Fig. 3B, none of them appeared significantly induced throughout the nodulation kinetics (adjusted p-value < 0,05). In agreement with this general absence in defense response, many other putative defense signaling genes were found to be down regulated (Fig. 3C) e.g. LRR-containing proteins (2 genes), receptor-like-kinases (7 genes), TIR-NBS disease resistance protein (2 genes) and a LysM domain containing protein (2 genes). These results suggest that the plant defense mechanisms are muted, at least in the early stages of the interaction. Such trade-offs between the symbiotic infection process and defense mechanisms are well documented<sup>20</sup>.

In sharp contrast, other candidate genes for signal transduction or gene regulation were induced as early as 6 hpi. These included 1 serine/threonine kinase, 1 MYB and 4 other transcription factors, as well as 3 genes encoding two component response regulators. These latter have been shown to be associated with cytokinin signaling, the hormone associated with nodule organogenesis in legumes<sup>21</sup>. These observations suggest that they all play an important role in the shift from root development to nodule organogenesis. The analysis of the genes specifically involved in the symbiotic process is reported in the “Behavior of symbiotic genes in the *A. evenia* RNA-seq database” section below.

**Late responses.** As mentioned in the “Early responses” section above, among the 1,276 up-regulated genes, 954 were specific to the 6 dpi time point (Fig. 2A) at which time the nodule is functional and able to fix nitrogen. A functional categorization based on gene ontology (GO) was conducted on the 593 contigs that were annotated. As expected, and as shown in Fig. 2C, six out of the 11 GO groups were related to metabolic or catabolic processes. A fine analysis of the “Response to Stress” GO group enabled the identification of 84 nodule-cysteine-rich (NCR) peptides encoding genes as those with the highest induction levels at 6 dpi (as manually annotated in Supplementary Table 1). NCR-like peptides have already been shown to be implicated in the differentiation of bacteroids in *Aeschynomene* nodules, and some have already been shown to be highly expressed in this organ<sup>12</sup>. Among the previously identified 82 NCR-like genes<sup>12</sup>, 73 were highly expressed at 6 dpi. A more detailed analysis enabled the identification of an additional 11 NCR-like encoding genes (Supplementary Table 1). Interestingly, in the “Response to Stress” GO group, several defense-related encoding genes, including 1 PAL, 2 PR proteins; 11 hypersensitive-induced response proteins and 9 MLO-like proteins, were induced in the 6 dpi time point (Fig. 3C). Taken together, these data suggest that *Aeschynomene* not only uses NCRs but also several other defense genes in mature nodules to keep its symbionts under control. This situation contrasts with the early time points where defense reactions appeared to be attenuated. Interestingly, one of the most highly induced genes at the 6 dpi time point corresponds to a NPR1 homolog (CL1720Contig3). NPR1 is a transcriptional regulator required for defense gene induction in response to salicylic acid<sup>22</sup>. Although we were able to identify 5 NPR1 annotated genes in our database, only the CL1720Contig3 gene appeared specifically induced in nodules (Fig. 3C). It is therefore tempting to speculate that many of the defense-related genes, and especially the NCR genes, that are highly induced in nodules could be regulated through this regulator. A reverse genetic approach such as the one we already performed for DNF1<sup>12</sup> would help to test this hypothesis.

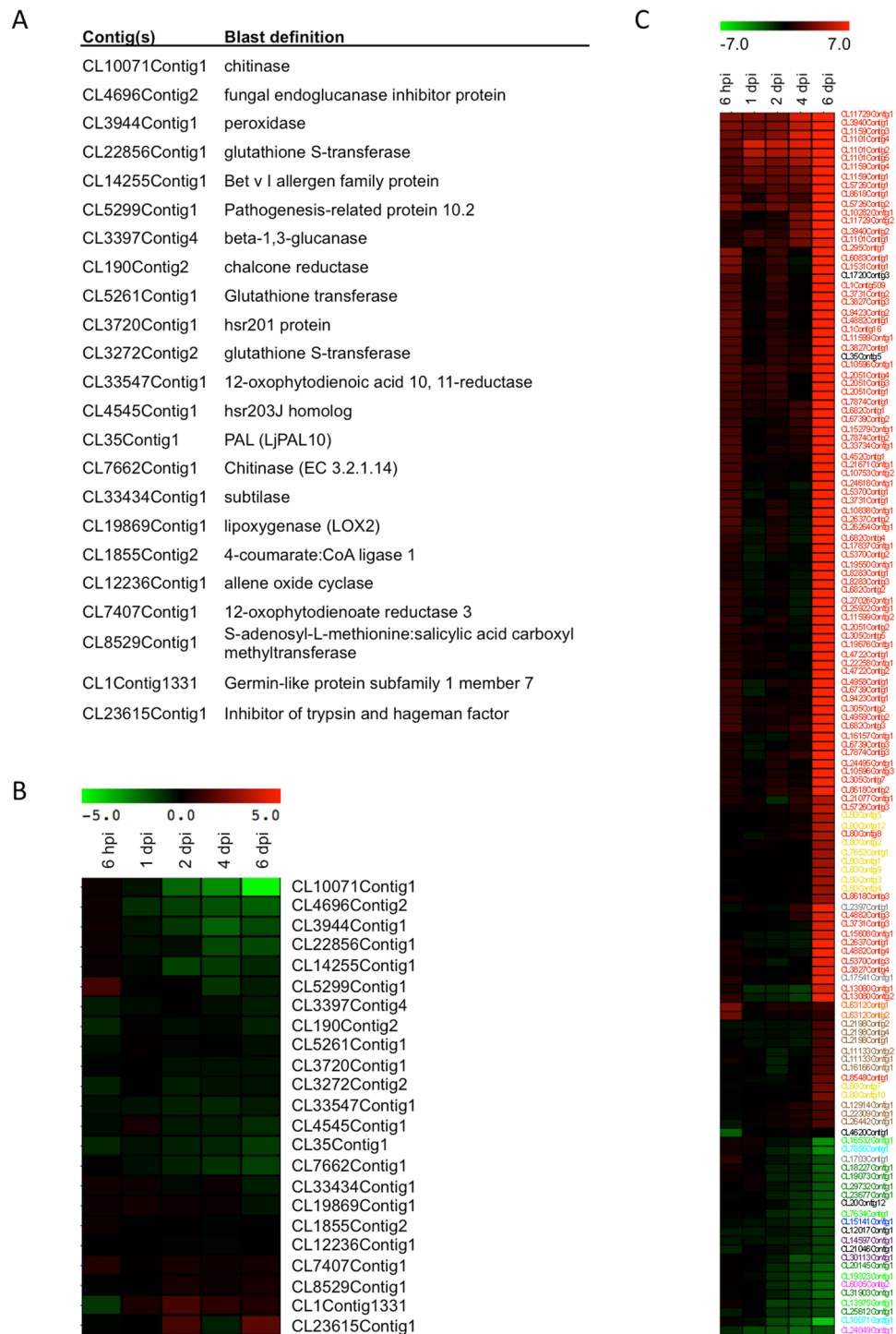
The mature nodules also had the highest number of down-regulated genes. Among the 470 genes, 283 were annotated and almost 50% of them (201) were involved in cell wall dynamics (Supplementary Table 1) including structural proteins (hydroxyproline-rich glycoprotein also known as extensin, 82 genes) or modeling enzymes (polygalacturonase, 6 genes; pectate lyase, 6 genes; pectin esterase, 13 genes). It should be noted that the time point 6 dpi corresponds to the isolated mature module stage (with limited growth and cell divisions) compared to non-inoculated roots (young tissue under strong elongation driven by cell division and cell wall extension). Therefore, it seems logical that all these genes linked with cell-wall dynamics are down regulated (Supplementary Fig. 2).

**Validation of differentially expressed genes using qRT-PCR.** As a first step in validating our RNAseq data, we focused on 17 genes among the most up- and down-regulated in *A. evenia* during the first step of the



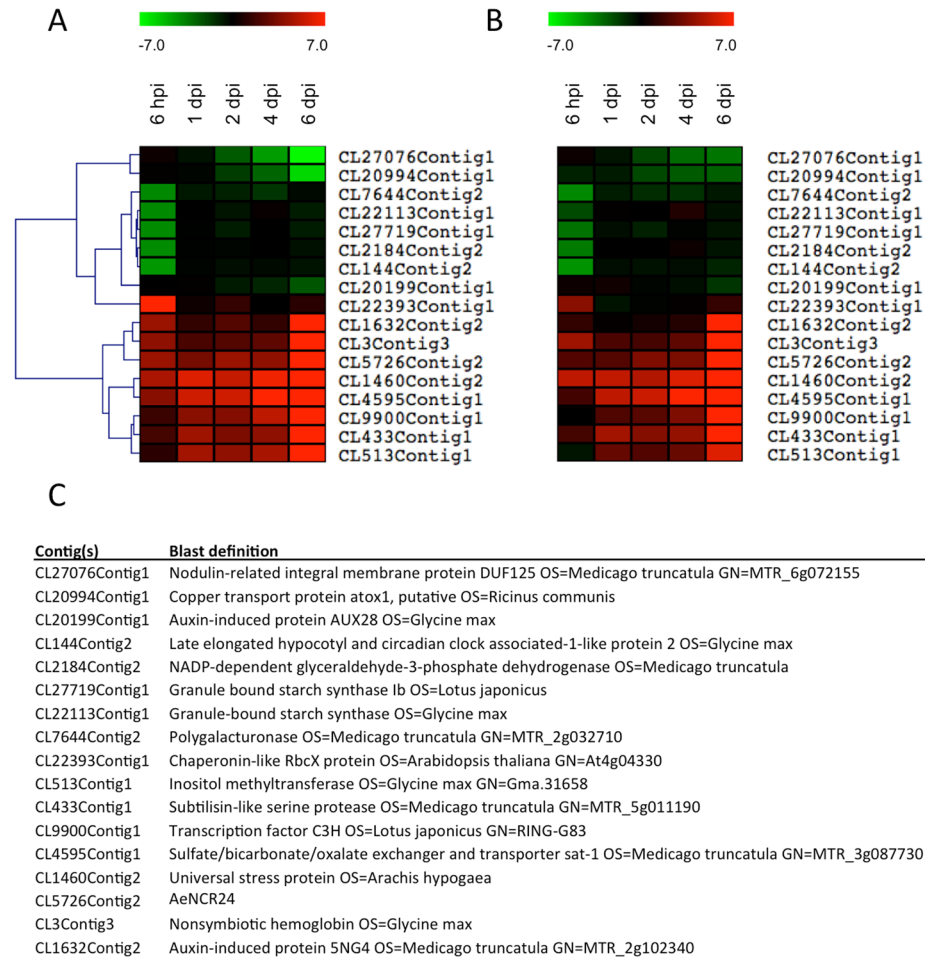
**Figure 2.** Differentially expressed genes. Venn diagram showing the genes found to be up-regulated (A) or down-regulated (D) during the time course of *Aeschynomene evenia* spp. *serrulata* inoculation with *Bradyrhizobium* ORS278. In both situations, DEGs were classified by gene ontology annotation for biological processes at 6 hpi (B,E) or 6 dpi (C,F) respectively. Overlaps between 2 adjacent time point conditions are also indicated. The details of the kinetics are the same as those described in Fig. 1.

interaction (6 hpi) or in the nodule tissue (6 dpi). Their corresponding messenger-RNA accumulation was examined in three biological replicates using qRT-PCR to confirm the reliability of the RNA-seq analysis (Fig. 4A–C). It is noteworthy that among the most up-regulated genes at 6 dpi, we found genes encoding a nodulin ortholog, leghemoglobin and a nodule rich cysteine peptide<sup>12</sup> (AeNCR24) known to be expressed in nodules of model legumes. All the orthologs of the candidate genes selected for validation of RNA-seq data had a similar expression profile in *M. truncatula* and in *Glycine max*<sup>23</sup>. The fold changes calculated from qRT-PCR data are expressed in log<sub>2</sub> to facilitate comparison with RNA-seq data. For the same reason, expression profiles, visualized as heatmaps, were subjected to hierarchical clustering using MEV software<sup>24</sup> (MeV 4.8.0 software). The results showed that the gene expression patterns evaluated by qRT-PCR (Fig. 4B) were consistent with the RNAseq data (Fig. 4A), evidence that our RNA-Seq analysis was robust.



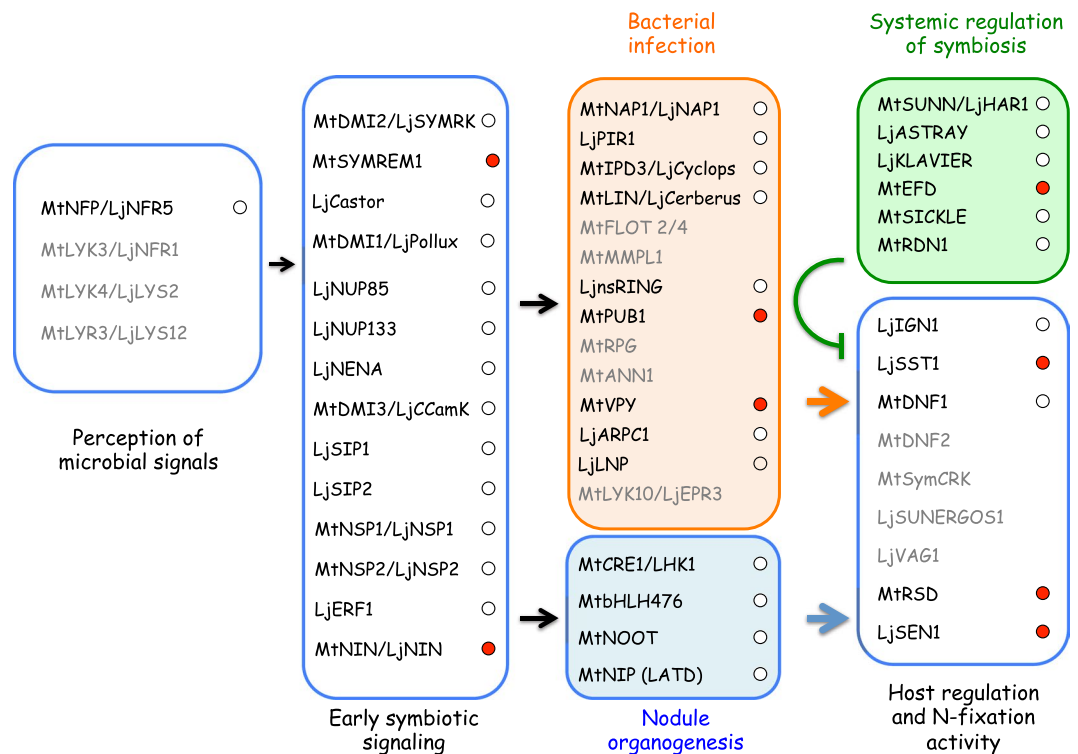
**Figure 3.** Heatmap of defense-related genes throughout nodulation kinetics. **(A)** The contig identification number and the BLAST definition of 23 genes known to be transiently up-regulated in response to symbionts in model legumes are presented in the table. **(B)** Behavior of the 23 corresponding orthologs in *Aeschynomene*. **(C)** Heatmap of differentially expressed defense-related genes throughout nodulation kinetics. Contigs corresponding to NCR are red, hypersensitive-induced response protein yellow, pathogenesis-related protein grey, MLO-like protein brown, isochorismatase orange, RLK protein dark green, disease resistance protein light green, LRR-containing protein purple, TIR-NBS disease resistance protein in pink, LysM domain containing protein in blue and chitinase in cyan. Red indicates up-regulation and green indicates down-regulation. Chroma color from green to red indicates  $\text{Log}_2$  (fold change) from less to more.

**Behavior of symbiotic genes in the *A. evenia* RNA-seq database.** We searched our database for genes reported to be involved in all the different steps in the symbiotic process in model legumes. Among the 52 symbiotic genes characterized in *M. truncatula* and *L. japonicus* we checked (Fig. 5; Supplementary Table 2),



**Figure 4.** Heat maps of 17 differentially expressed genes with contrasted behavior during the infection process obtained from RNAseq data (A) and from real time quantitative RT-PCR (qPCR) (B). The contig identification number and the BLAST definition of these 17 genes are presented in the table (C). Three biological replicates were performed for each RT-qPCR validation; the results are expressed as means. The elongation factor one alpha (EF1 $\alpha$ ) gene was used as internal control. All results are expressed as Log<sub>2</sub> fold change. The color scale corresponding to the expression fold change is presented above panels (A and B). Clustering was performed with the MultiExperiment Viewer 4.8 by hierarchical clustering of all genes and samples using Manhattan distance calculations.

we were able to identify 39 orthologs in *A. evenia*. The list of present or absent symbiotic genes is consistent with previous Illumina data used to develop molecular markers in another *A. evenia* accession of the spp. *evenia*<sup>25</sup>. The symbiotic genes for which orthologs were identified in both *A. evenia* accessions are required in all the different steps of the nodulation process: from the perception of the rhizobium to the induction of the signaling pathways that control bacterial infection and nodule organogenesis and finally in local and systemic regulation of nitrogen fixation (Fig. 5). The distinctive feature and advantage of the present study is monitoring the expression profiles of these symbiotic genes throughout nodulation kinetics. We found strong induction of expression in eight of them (*MtSYMREM1*, *MtNIN/LjNIN*, *MtPUB1*, *MtVPY*, *MtEFD*, *LjSST1*, *MtRSD*, *LjSEN1*) throughout nodule development, comparable to the behavior reported in *M. truncatula* and *L. japonicus* (for a review see<sup>2,3</sup>). *MtSYMREM1* is a symbiotic receptor-binding protein that interacts with the *bona fide* receptors MtLYK3/LjNFR1, MtNFP/LjNFR5 and MtDMI2/LjSYMRK<sup>21</sup>. Expression pattern of *MtSYMREM1* and the phenotypes of loss-of-function mutants suggest a role for this gene in bacterial signal perception during the initial stages of infection and throughout nodule development<sup>26</sup>. The expression profile of the *SYMREM1* ortholog in *A. evenia* differs slightly from the expression profile of *MtSYMREM1*<sup>25</sup>: no induction was observed during the very first steps of bacterial perception by the host plant; the induction of expression *AeSYMREM1* started at 4 dpi and peaked at 6 dpi when nodules were fully functional (Supplementary Fig. 3). We observed the same time shift in induction for *NIN*, encoding a transcription regulator that is essential for nodulation<sup>27</sup>, *MtPUB1*, encoding an E3 ubiquitin ligase that interacts with the NF receptor MtLYK3<sup>28</sup> and *MtVPY*, a common symbiotic gene<sup>29</sup>. The fact that the crack-entry infection process bypasses epidermal responses to rhizobia could explain why the induction of *SYMREM1*, *NIN*, *PUB1* and *VPY* orthologs during the early steps of nodulation is not observed in *A. evenia*, whereas the high level of transcription in nodules for those genes is conserved between model legumes and *A.*



**Figure 5.** Schematic representation of the gene cascades in symbiotic signaling. Symbiotic genes identified in the *Medicago truncatula* and *Lotus japonicus* genomes were classified according to their putative roles in root nodule symbiosis. The genes identified in *A. evenia* spp. *serrulata* are black and the symbiotic genes for which no transcripts were detected are grey. Regulation of transcription during the nodulation kinetics is in the circle: in white for constitutive expression; in red for up-regulation; in green for down-regulation. Adapted from Chaintreuil *et al.*<sup>25</sup>.

*evenia*. Nodule specific induction of expression was also conserved in the ortholog of *MtEFD* and in the orthologs of *LjSST1*, *MtRSD* and *LjSEN1*, the former being involved in systemic regulation of symbiosis and the latter in host regulation and nitrogen-fixation activity (Fig. 5).

In addition to this panel of symbiotic genes whose expression profile is at least partly conserved between model legumes and *A. evenia*, we also identified nine orthologs of symbiotic genes (*LjNENA*, *MtNSP1/LjNSP1*, *MtNSP2/LjNSP2*, *LjSIP1*, *LjJERF1*, *MtLIN/LjCERBERUS*, *MtNOOT*, *LjnsRING*, *MtDNF1*) whose expression was not significantly affected during the symbiotic process contrary to what has been reported in *M. truncatula* or *L. japonicus*<sup>30–40</sup>. In the corresponding model legumes, five of the genes (*MtNSP1/LjNSP1*, *MtNSP2/LjNSP2*, *LjSIP1*, *LjJERF1*, *MtNOOT*) display a moderate, and sometimes transitory (*LjSIP1*, *LjJERF1*), increase in their transcription level under symbiotic conditions, which could explain why we failed to detect the modulation of transcription of these genes in our dataset. There were still four symbiotic genes (*LjNENA*, *MtLIN/LjCERBERUS*, *LjnsRING* and *MtDNF1*) whose expression significantly increased during nodule development in model legumes but not in *A. evenia*. *DNF1* encodes the SPC22 subunit of a nodule-specific signal peptidase complex required for maturation of NCR peptides. *DNF1* is the only one of the subset of four nodule-expressed genes that have undergone RNAi experiments and whose role in nodulation was demonstrated in *A. evenia*<sup>12</sup>. It is even more surprising not to find an expression profile similar to that observed in *M. truncatula*. Still, no straightforward hypothesis can be proposed concerning the absence of *DNF1* up-regulation in *A. evenia* nodules. In the case of *LjNENA*, *LjnsRING* and *MtLIN/LjCERBERUS*, three genes required for the initiation or growth of infection threads, the intercellular infection process of *Aeschynomene* species could once again explain the differences in transcriptional regulation observed between *A. evenia* and temperate model legumes in which rhizobia use root hair infection threads to invade plant tissues in an intracellular way<sup>6,30,37–39</sup>. The phenotype of *nena* mutants is particularly relevant. No root hair curling or root hair infection threads were observed in *nena* mutants, but a few infected nodules appeared 21 days post infection<sup>30</sup>. Detailed microscopic analyses of the mutants revealed that the rhizobial infection mode in *nena* mutants carried the hallmarks of crack entry, the infection process at work in *A. evenia*. This is additional evidence that some, if not all, epidermal responses to rhizobia are bypassed during the crack entry infection process.

Concerning the 13 symbiotic genes for which transcripts were not detected in *A. evenia* transcriptomes, sequencing of *A. evenia* genome is currently underway and will enable us to clarify if these genes do exist in the *A. evenia* genome but are not, or only very weakly, expressed in our conditions. Nevertheless, as previously observed in *A. evenia* spp. *evenia*<sup>7</sup>, it is significant that nine out of thirteen non-expressed genes in *A. evenia* (*MtLYK3/LjNFR1*, *MtLYK4/LjLYS2*, *MtLYR3/LjLYS12*, *MtLYK10/LjEPR3*, *MtFLOT2*, *MtFLOT4*, *MtMMPL1*, *MtRPG*,



*MtANN1*) are either involved in NF perception or IT formation and growth, two processes that are not involved in the *A. evenia-Bradyrhizobium* ORS278 symbiosis<sup>41–48</sup>. The four other genes missing in our data set are required for nodule functioning<sup>49,50</sup> (*LjSUNERGOS1*, *LjVAG1*) or control of immunity<sup>51,52</sup> (*MtDNF2*, *MtSymCRK*).

In conclusion, the BLAST search performed using the sequence of symbiotic genes characterized in *M. truncatula* and *L. japonicus* enabled the identification of 39 orthologs in *A. evenia* among which 13 display different expression profiles from those observed in the two model legumes. These observations demonstrate that specific mechanisms occur during the NF-independent symbiosis even if the common signaling pathway was shown to be partly recruited in *A. evenia*<sup>11</sup>.

**Toward the identification of molecular markers for the search for *Bradyrhizobium* signal molecule(s).** Since the discovery of an NF-independent symbiosis in *Aeschynomene* species, the major challenge has been the elucidation of the first steps of the molecular dialogue between the plant and its symbionts. On the bacterial side, the signal molecule(s) used by *Bradyrhizobium* ORS278 to interact with its host plant in a NF-independent way remain(s) to be identified. The main obstacle to achieving this goal is the absence in *Aeschynomene* of an early easy to monitor symbiotic response to inoculation, like the root hair curling observed in model legumes in response to NFs. The use of molecular markers of the early stages of interaction as a bioassay appeared to be a good alternative to get round this problem. In fact such strategy was originally set up to purify Myc signals using a transgenic line of *M. truncatula* carrying a fusion between the promoter region of the early nodulin gene *MtENOD11* and the *GUS* reporter gene<sup>53</sup>. We attempted to use this bioassay but our results demonstrated that *pMtENOD11* is not activated during the early stages of the interaction of *Aeschynomene* with photosynthetic *bradyrhizobia* strains<sup>54</sup>.

RNAseq data will probably be able to provide better markers of early symbiotic stages. Initially, we were particularly concerned about the expression profile of the ortholog of well-characterized symbiotic genes such as *NIN*. *NIN* is a transcription regulator that is essential for nodulation but not required for arbuscular mycorrhizal (AM) symbiosis<sup>27</sup>. Its loss of function inhibits nodule organogenesis and bacterial infection and its expression is dramatically induced in the root epidermis and in the cortex during the infection process in both *M. truncatula* and *L. japonicus*<sup>27,55</sup>. The rapid induction of *NIN* expression in response to rhizobia inoculation in the two model legumes is conserved in actinorhizal plants that form root nodule symbiosis with *Frankia*<sup>56</sup>. Unfortunately, in *A. evenia*, the induction of the putative ortholog of *NIN* significantly starts with nodule inception (Supplementary Fig. 3). This does not preclude conserved symbiotic function but the gene is apparently not a suitable candidate for a promoter::reporter gene construct capable of monitoring early plant responses to its symbionts.

Other candidates emerged from the RNAseq data according to their expression profiles, like those illustrated in Fig. 4. We focused on genes that displayed early and stable induction during the complete nodulation kinetics as observed in RNAseq data (Fig. 4A) and confirmed by RT-qPCR (Fig. 4B). Blast search revealed three contigs that appeared not to be related to symbiotic signaling processes: CL4595Contig1, CL1460Contig2 and CL5726Contig2 (Fig. 4C). Three candidates were retained: CL9900Contig1, which encodes a transcription factor; CL513Contig1, which shares homology with an inositol-O-methyltransferase, a protein involved in abiotic stress tolerance in *Glycine max*<sup>57</sup> and CL433Contig1, which encodes a subtilase whose putative homolog in *L. japonicus*, *SbtM1*, is an early molecular marker of arbuscular mycorrhiza<sup>58</sup>. Silencing of *SbtM1* confirmed the role of subtilase during the fungal infection process<sup>57</sup>. Amongst other modes of action, plant subtilases are believed to weaken cell-cell connections by degrading structural protein in the apoplastic space and may thus be required for elongation of the fungal hyphae in the intercellular space. This putative role of subtilase in cell wall loosening could also be at play during the intercellular infection process in *Aeschynomene*. Taken together, these lines of evidence strongly support the choice of CL433Contig1 for the development of a bioassay. We will take advantage of genomics data currently being produced for *A. evenia* to identify the promoter sequences of CL513Contig1 and CL433Contig1 to develop a bioassay consisting in a fusion between their promoter regions with a *GUS::GFP* cassette. In parallel, it would be interesting to address the role of these new symbiotic genes, especially CL433Contig1, in silencing experiments.

In conclusion, the data generated by this RNAseq analysis represent an important advance in the understanding and decryption of the NF-independent symbiotic process. They also contain promising new genes to use as reporter markers to better characterize the nodulation process and to enable the identification of the signal molecule(s) used by *Bradyrhizobium* ORS278 to interact with *A. evenia*.

## Methods

**Plant and Bacterial Strains.** *A. evenia* spp. *serrulata* species used in this study corresponds to line IREFL6945<sup>7</sup> (USDA). Rhizobial inoculations and growth conditions for plants are detailed in<sup>11</sup>. Briefly, seeds were scarified with concentrated (95%) sulfuric acid and surface sterilized with 3% sodium hypochlorite. Germination was induced on 0.8% water agar plates for one day in a 34 °C chamber in the dark. One day-old seedlings were transferred to test tubes filled with BNM medium. Plants were grown in a 28 °C growth chamber with a 16h/8h light regime and 70% humidity. One week later, the plants were inoculated with 1 ml of a 5 day-old *Bradyrhizobium* sp. ORS278 strain<sup>59</sup> washed and with the optical density at 600 nm adjusted to 1. Root tissues were collected from non-inoculated samples just before inoculation. Six hours post-inoculation (hpi), and 1, 2, 4, 6 and 7 days post-inoculation (dpi), root sections corresponding to the nodulation inception zone (1 to 2 cm long section below the collet) were harvested. Young developing leaves and mature nodules were collected from 3-week-old *A. evenia* plants.

**Total RNA isolation.** Total RNA was extracted from roots using the SV total RNA Isolation System (Promega) and quantified using a NanoDrop ND-1000 spectrophotometer. We used a pool of 24 plants for each

time point. The integrity of the RNA samples was checked using an Agilent 2100 Bioanalyzer according to the manufacturer's instructions.

**454 and Illumina sequencing.** (GATC Biotech, Mulhouse, France) and then used for 454 sequencing (Roche). The sequences obtained were analyzed using a pipeline called Expressed Sequence Tag treatment and investigation kit<sup>13</sup> (ESTtik). This pipeline enabled editing and assembling of the input cDNA sequences in a non-redundant data set.

Six mRNA libraries corresponding to non-inoculated roots and roots inoculated at 6 hpi, 1, 2, 4 and 6 dpi, were built and sequenced at the MGX platform (Montpellier Genomix, *Institut de Genomique Fonctionnelle*, Montpellier France). The RNA libraries were constructed using the TruSeq stranded mRNA library construction kit (Illumina Inc., USA). The quantitative and qualitative analyses of the library were carried on Agilent\_DNA 1000 chip and qPCR (Applied Biosystems 7500, SYBR Green). RNA was sequenced using the Illumina SBS (sequence by synthesis) technique on a Hiseq2000 in single read 100 nt mode. Image analysis, base calling and quality filtering were performed using Illumina software.

**RNaseq: short read counting method.** The complete transcriptomic high-throughput sequencing data were analyzed with the TAMARA bioinformatic pipeline (Transcriptome Analyses based on Massive sequencing of RNAs; Cruveiller S., unpublished) currently implemented on the Microscope platform<sup>60</sup>. The pipeline is a "Master" shell script that launches the different parts of the analysis and checks that all tasks are completed without error. We first assessed RNA-Seq data quality including options such as read trimming or the use of merging/split paired-end reads. Reads were then mapped onto the EST contigs of *A. evenia* spp. *serrulata* with the SSAHA2 package<sup>61</sup>. We minimized the false positive discovery rate by selecting only reliable alignments from SAM-formatted files using SAMtools<sup>62</sup> (v.0.1.8). The number of reads matching each EST contig was then calculated with an enhanced version of the coverageBed software from the BEDtools suite<sup>63</sup>. Reads matching several genomic objects were weighted by the number of objects matched so as to keep the total number of reads in the library constant. Finally, the Bioconductor-DESeq package<sup>64</sup> (v1.4.1) was used with default parameters for the analysis of raw count data and to determine whether expression levels differed significantly between conditions. The R script described in Supplementary Method 1 was used to select the up- or down-regulated gene sets specific for each time lapse point.

**qPCR analysis.** One hundred nanograms of total RNA per sample were reverse transcribed using SuperScriptII reverse transcriptase (Invitrogen) and oligo(dT)12–18. Real time qPCR was performed using the Brilliant II SYBR<sup>®</sup> Green QPCR Master Mix (Agilent Technologies). Primer sequences are listed in Supplementary Table 3. The Real Time SYBR<sup>®</sup> Green cycling PCR program on a Stratagene MX3005P (Agilent Technologies) was as follows: one cycle at 95 °C for 10 min, 40 cycles at 95 °C for 10 s, 60 °C for 30 s and 72 °C for 30 s, ending with one cycle at 95 °C for one min, 55 °C for 30 s, and 95 °C for 30 s. Expression levels were normalized with the *AeEF1α* reference gene. Three independent biological replicates were performed for each experiment and two qPCR technical replicates were performed for each biological replicate. The efficiency-corrected comparative quantification method was used to quantify the expression of genes of interest<sup>65</sup>.

**Data Availability.** The reference transcriptome and the raw FASTQfiles for the 18 libraries were respectively deposited on NCBI under the bioprojects PRJNA396532 and PRJNA399681 with the accession numbers: SRR5983062, SRR5983061, SRR5983060, SRR5983059, SRR5983066, SRR5983065, SRR5983064, SRR5983063, SRR5983069, SRR5983067, SRR5983070, SRR5983068, SRR5983072, SRR5983071, SRR5983073, SRR5983074, SRR5983076, SRR5983075.

## References

- Graham, P. H. & Vance, C. P. Update on Legume Utilization Legumes: Importance and Constraints to Greater Use. *Plant Physiol.* **131**, 872–877 (2014).
- Oldroyd, G. E. D., Murray, J. D., Poole, P. S. & Downie, J. A. The rules of engagement in the legume-rhizobial symbiosis. *Annu. Rev. Genet.* **45**, 119–44 (2011).
- Kouchi, H. *et al.* How many peas in a pod? Legume genes responsible for mutualistic symbioses underground. *Plant Cell Physiol.* **51**, 1381–97 (2010).
- Sprent, J. I. Evolving ideas of legume evolution and diversity: a taxonomic perspective on the occurrence of nodulation. *New Phytol.* **174**, 11–25 (2007).
- Giraud, E. *et al.* Legumes symbioses: absence of Nod genes in photosynthetic bradyrhizobia. *Science* **316**, 1307–12 (2007).
- Bonaldi, K. *et al.* Nodulation of *Aeschynomene afraspera* and *A. indica* by photosynthetic Bradyrhizobium Sp. strain ORS285: the nod-dependent versus the nod-independent symbiotic interaction. *Mol. Plant. Microbe. Interact.* **24**, 1359–71 (2011).
- Arrighi, J.-F. *et al.* *Aeschynomene evenia*, a model plant for studying the molecular genetics of the nod-independent rhizobium-legume symbiosis. *Mol. Plant. Microbe. Interact.* **25**, 851–61 (2012).
- Mitra, R. M. *et al.* A Ca<sup>2+</sup>/calmodulin-dependent protein kinase required for symbiotic nodule development: Gene identification by transcript-based cloning. *Proc. Natl. Acad. Sci. USA* **101**, 4701–5 (2004).
- Stracke, S. *et al.* A plant receptor-like kinase required for both bacterial and fungal symbiosis. *Nature* **417**, 959–62 (2002).
- Gonzalez-Rizzo, S., Crespi, M. & Frugier, F. The Medicago truncatula CRE1 cytokinin receptor regulates lateral root development and early symbiotic interaction with Sinorhizobium meliloti. *Plant Cell* **18**, 2680–93 (2006).
- Fabre, S. *et al.* The Nod factor-independent nodulation in *Aeschynomene evenia* required the common plant-microbe symbiotic 'toolkit'. *Plant Physiol.* **169**, 2654–2664 (2015).
- Czerniec, P. *et al.* Convergent Evolution of Endosymbiont Differentiation in Dalbergioid and Inverted Repeat-Lacking Clade Legumes Mediated by Nodule-Specific Cysteine-Rich Peptides. *Plant Physiol.* **169**, 1254–1265 (2015).
- Argout, X. *et al.* Towards the understanding of the cocoa transcriptome: Production and analysis of an exhaustive dataset of ESTs of Theobroma cacao L. generated from various tissues and under various conditions. *BMC Genomics* **9**, 512 (2008).

14. Wildermuth, M. C., Dewdney, J., Wu, G. & Ausubel, F. M. Isochorismate synthesis is required to synthesize salicylic acid for plant defense. *Nature* **414**, 562–565 (2001).
15. Liu, T. *et al.* Unconventionally secreted effectors of two filamentous pathogens target plant salicylate biosynthesis. *Nat. Commun.* **5**, (2014).
16. Lohar, D. P. *et al.* Transcript analysis of early nodulation events in *Medicago truncatula*. *Plant Physiol.* **140**, 221–34 (2006).
17. Rey, T. *et al.* NFP, a LysM protein controlling Nod factor perception, also intervenes in *Medicago truncatula* resistance to pathogens. *New Phytol.* **198**, 875–886 (2013).
18. Kouchi, H. Large-Scale Analysis of Gene Expression Profiles during Early Stages of Root Nodule Formation in a Model Legume, *Lotus japonicus*. *DNA Res.* **11**, 263–274 (2004).
19. Libault, M. *et al.* Complete Transcriptome of the Soybean Root Hair Cell, a Single-Cell Model, and Its Alteration in Response to *Bradyrhizobium japonicum* Infection. *Plant Physiol.* **152**, 541–552 (2010).
20. Zipfel, C. & Oldroyd, G. E. D. Plant signalling in symbiosis and immunity. *Nature* **543**, 328–336 (2017).
21. Gamas, P., Brault, M., Jardinaud, M. F. & Frugier, F. Cytokinins in Symbiotic Nodulation: When, Where, What For? *Trends Plant Sci.* **22**, 792–802 (2017).
22. Yan, S. & Dong, X. Perception of the plant immune signal salicylic acid. *Curr. Opin. Plant Biol.* **20**, 64–68 (2014).
23. Goodstein, D. M. *et al.* Phytozome: a comparative platform for green plant genomics. *Nucleic Acids Res.* **40**, D1178–D1186 (2012).
24. Howe, E. A., Sinha, R., Schlauch, D. & Quackenbush, J. RNA-Seq analysis in MeV. *Bioinformatics* **27**, 3209–3210 (2011).
25. Chaintreuil, C. *et al.* A gene-based map of the Nod factor-independent *Aeschynomene evenia* sheds new light on the evolution of nodulation and legume genomes. *DNA Res.* **23**, 365–76 (2016).
26. Lefebvre, B. *et al.* A remorin protein interacts with symbiotic receptors and regulates bacterial infection. *Proc. Natl. Acad. Sci. USA* **107**, 2343–8 (2010).
27. Schauser, L., Roussis, A., Stiller, J. & Stougaard, J. A plant regulator controlling development of symbiotic root nodules. *Nature* **402**, 191–5 (1999).
28. Mbengue, M. *et al.* The *Medicago truncatula* E3 Ubiquitin Ligase PUB1 Interacts with the LYK3 Symbiotic Receptor and Negatively Regulates Infection and Nodulation. *Plant Cell* **22**, 3474–3488 (2010).
29. Murray, J. D. *et al.* Vapyrin, a gene essential for intracellular progression of arbuscular mycorrhizal symbiosis, is also essential for infection by rhizobia in the nodule symbiosis of *Medicago truncatula*. *Plant J.* **65**, 244–52 (2011).
30. Groth, M. *et al.* NENA, a *Lotus japonicus* homolog of Sec. 13, is required for rhizodermal infection by arbuscular mycorrhiza fungi and rhizobia but dispensable for cortical endosymbiotic development. *Plant Cell* **22**, 2509–26 (2010).
31. Zhu, H. *et al.* A novel ARID DNA-binding protein interacts with SymRK and is expressed during early nodule development in *Lotus japonicus*. *Plant Physiol.* **148**, 337–47 (2008).
32. Heckmann, A. B. *et al.* *Lotus japonicus* nodulation requires two GRAS domain regulators, one of which is functionally conserved in a non-legume. *Plant Physiol.* **142**, 1739–50 (2006).
33. Smit, P. *et al.* NSP1 of the GRAS protein family is essential for rhizobial Nod factor-induced transcription. *Science* **308**, 1789–91 (2005).
34. Kaló, P. *et al.* Nodulation signaling in legumes requires NSP2, a member of the GRAS family of transcriptional regulators. *Science* **308**, 1786–9 (2005).
35. Asamizu, E., Shimoda, Y., Kouchi, H., Tabata, S. & Sato, S. A positive regulatory role for LjERF1 in the nodulation process is revealed by systematic analysis of nodule-associated transcription factors of *Lotus japonicus*. *Plant Physiol.* **147**, 2030–2040 (2008).
36. Couzigou, J.-M. *et al.* NODULE ROOT and COCHLEATA maintain nodule development and are legume orthologs of Arabidopsis BLADE-ON-PETIOLE genes. *Plant Cell* **24**, 4498–510 (2012).
37. Kiss, E. *et al.* LIN, a novel type of U-box/WD40 protein, controls early infection by rhizobia in legumes. *Plant Physiol.* **151**, 1239–49 (2009).
38. Yano, K. *et al.* CERBERUS, a novel U-box protein containing WD-40 repeats, is required for formation of the infection thread and nodule development in the legume-Rhizobium symbiosis. *Plant J.* **60**, 168–80 (2009).
39. Shimomura, K., Nomura, M., Tajima, S. & Kouchi, H. LjnsRING, a Novel RING Finger Protein, is Required for Symbiotic Interactions Between *Mesorhizobium loti* and *Lotus japonicus*. *Plant Cell Physiol.* **47**, 1572–1581 (2006).
40. Wang, D. *et al.* A nodule-specific protein secretory pathway required for nitrogen-fixing symbiosis. *Science* **327**, 1126–9 (2010).
41. Limpens, E. *et al.* LysM domain receptor kinases regulating rhizobial Nod factor-induced infection. *Science* **302**, 630–3 (2003).
42. Radutoiu, S. *et al.* Plant recognition of symbiotic bacteria requires two LysM receptor-like kinases. *Nature* **425**, 585–92 (2003).
43. Lohmann, G. V. *et al.* Evolution and Regulation of the *Lotus japonicus* LysM Receptor Gene Family. *Mol. Plant-Microbe Interact.* **23**, 510–521 (2010).
44. Arrighi, J.-F. *et al.* The *Medicago truncatula* lysine motif-receptor-like kinase gene family includes NFP and new nodule-expressed genes. *Plant Physiol.* **142**, 265–79 (2006).
45. Arrighi, J.-F. *et al.* The RPG gene of *Medicago truncatula* controls Rhizobium-directed polar growth during infection. *Proc. Natl. Acad. Sci. USA* **105**, 9817–22 (2008).
46. Haney, C. H. & Long, S. Plant flotillins are required for infection by nitrogen-fixing bacteria. *Proc. Natl. Acad. Sci. USA* **107**, 478–83 (2010).
47. Combier, J.-P. *et al.* The MtMMPL1 early nodulin is a novel member of the matrix metalloendoproteinase family with a role in *Medicago truncatula* infection by *Sinorhizobium meliloti*. *Plant Physiol.* **144**, 703–16 (2007).
48. Niebel, F. D. C., Lescure, N., Cullimore, J. V. & Gamas, P. The *Medicago truncatula* MtAnn1 gene encoding an annexin is induced by Nod factors and during the symbiotic interaction with *Rhizobium meliloti*. *Mol. Plant. Microbe. Interact.* **11**, 504–13 (1998).
49. Yoon, H. J. *et al.* *Lotus japonicus* SUNERGOS1 encodes a predicted subunit A of a DNA topoisomerase VI that is required for nodule differentiation and accommodation of rhizobial infection. *Plant J.* **78**, 811–21 (2014).
50. Suzaki, T. *et al.* Endoreduplication-mediated initiation of symbiotic organ development in *Lotus japonicus*. *Development* **141**, 2441–2445 (2014).
51. Bourcy, M. *et al.* *Medicago truncatula* DNF2 is a PI-PLC-XD-containing protein required for bacteroid persistence and prevention of nodule early senescence and defense-like reactions. *New Phytol.* **197**, 1250–1261 (2013).
52. Berrabah, F. *et al.* A nonRD receptor-like kinase prevents nodule early senescence and defense-like reactions during symbiosis. *New Phytol.* **203**, 1305–1314 (2014).
53. Maillat, F. *et al.* Fungal lipochitooligosaccharide symbiotic signals in arbuscular mycorrhiza. *Nature* **469**, 58–63 (2011).
54. Bonaldi, K. *et al.* The Nod factor-independent symbiotic signaling pathway: development of *Agrobacterium rhizogenes*-mediated transformation for the legume *Aeschynomene indica*. *Mol. Plant. Microbe. Interact.* **23**, 1537–44 (2010).
55. Marsh, J. F. *et al.* *Medicago truncatula* NIN is essential for rhizobial-independent nodule organogenesis induced by autoactive calcium/calmodulin-dependent protein kinase. *Plant Physiol.* **144**, 324–35 (2007).
56. Clavijo, F. *et al.* The *Casuarina* NIN gene is transcriptionally activated throughout *Frankia* root infection as well as in response to bacterial diffusible signals. *New Phytol.* **208**, 887–903 (2015).
57. Ahn, C., Park, U. & Park, P. B. Increased salt and drought tolerance by D-ononitol production in transgenic *Arabidopsis thaliana*. *Biochem. Biophys. Res. Commun.* **415**, 669–674 (2011).
58. Takeda, N., Sato, S., Asamizu, E., Tabata, S. & Parniske, M. Apoplastic plant subtilases support arbuscular mycorrhiza development in *Lotus japonicus*. *Plant J.* **58**, 766–77 (2009).

59. Molouba, F. *et al.* Photosynthetic bradyrhizobia from *Aeschynomene* spp. are specific to stem-nodulated species and form a separate 16S ribosomal DNA restriction fragment length polymorphism group. *Appl. Environ. Microbiol.* **65**, 3084–94 (1999).
60. Vallenet, D. *et al.* MicroScope—an integrated microbial resource for the curation and comparative analysis of genomic and metabolic data. *Nucleic Acids Res.* **41**, D636–D647 (2013).
61. Ning, Z., Cox, A. J. & Mullikin, J. C. SSAHA: a fast search method for large DNA databases. *Genome Res.* **11**, 1725–9 (2001).
62. Li, H. *et al.* The Sequence Alignment/Map format and SAMtools. *Bioinformatics* **25**, 2078–2079 (2009).
63. Quinlan, A. R. & Hall, I. M. BEDTools: a flexible suite of utilities for comparing genomic features. *Bioinformatics* **26**, 841–2 (2010).
64. Anders, S. & Huber, W. Differential expression analysis for sequence count data. *Genome Biol.* **11**, R106 (2010).
65. Pfaffl, M. W. A new mathematical model for relative quantification in real-time RT-PCR. *Nucleic Acids Res.* **29**, e45 (2001).

## Acknowledgements

This work was funded by a grant from the French National Research Agency (ANR-SESAM-2010-BLAN-170801) and supported by Grants from Govt. of India: IFCPAR/CEFIPRA (IFC/5103-4/2014/543).

## Author Contributions

D.G., P.C. and F.C. performed biological experiments. S.N. and S.R. produced RNAseq data. S.C., F.M., C.L., D.V. and P.F. realized bio-informatics. E.G., J.-F.A. and M.D.G. reviewed and considered the manuscript. P.C. and F.C. designed the experiments, analyzed the data and wrote the manuscript.

## Additional Information

**Supplementary information** accompanies this paper at <https://doi.org/10.1038/s41598-018-29301-0>.

**Competing Interests:** The authors declare no competing interests.

**Publisher's note:** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.



**Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this license, visit <http://creativecommons.org/licenses/by/4.0/>.

© The Author(s) 2018