Proteome reference maps of the *Lotus japonicus* nodule and root

Svend Dam[1,2], *Thomas F. Dyrlund*[1], *Anna Ussatjuk*[1,2], *Bjarne Jochimsen*[1,2], *Kasper Nielsen*[3], *Nicolas Goffard*[4], *Miguel Ventosa*[1,2], *Andrea Lorentzen*[5], *Vikas Gupta*[1,2], *Stig U. Andersen*[1,2], *Jan J. Enghild*[1], *Clive W. Ronson*[2,6], *Peter Roepstorff*[5], *and Jens Stougaard*[1,2*]

[1]Department of Molecular Biology and Genetics, Aarhus University, DK–8000 Aarhus, Denmark

[2]Centre for Carbohydrate Recognition and Signalling

[3]Center for Biological Sequence Analysis, Technical University of Denmark, DK–2800 Kgs Lyngby, Denmark

[4]Computational Biology and Bioinformatics, IP & Science, Thomson Reuters, United Kingdom

[5]Department of Biochemistry and Molecular Biology, University of Southern Denmark, DK-5230 Odense M, Denmark

[6]Department of Microbiology and Immunology, University of Otago, Dunedin, New Zealand.

[*]Corresponding author. Tel.: +45 8715 5504; fax: +45 8612 3178. E-mail address: stougaard@mb.au.dk (J. Stougaard)

Received: August 13, 2013; Revised: October 11, 2013; Accepted: November 11, 2013

This article has been accepted for publication and undergone full peer review but has not been through the copyediting, typesetting, pagination and proofreading process, which may lead to differences between this version and the Version of Record. Please cite this article as doi: 10.1002/pmic.201300353.

List of abbreviations:

APX, ascorbate peroxidase; **DAG**, days after germination; **DAI**, days after inoculation; **DHAR**, dehydroascorbate reductase; **GME**, GDP-D-mannose 3',5'-epimerase; **L-GalDH**, L-galactose dehydrogenase; ***Lotus***, *Lotus japonicus*; ***M. loti***, *Mesorhizobium loti* strain MAFF303099; **PR**, pathogen related; **PMM**, phosphomannomutase; **PTM,** post translational modification

Keywords: comparative analysis, *Lotus japonicus*, proteomics, root, root nodule

Total number of words, including references as well as figure and table legends, is 7069.
Three figures and two tables

Abstract

Legume symbiosis with rhizobia results in the formation of a specialized organ, the root nodule, where atmospheric dinitrogen is reduced to ammonia. In *Lotus japonicus* (*Lotus*), several genes involved in nodule development or nodule function have been defined using biochemistry, genetic approaches, and high throughput transcriptomics. We have employed proteomics to further understand nodule development. Two developmental stages representing nodules prior to nitrogen fixation (white) and mature nitrogen fixing nodules (red) were compared with roots. In addition, the proteome of a spontaneous nodule formation mutant (*snf1*) was determined. From nodules and roots, 780 and 790 protein spots from 2D gels were identified and approximately 45% of the corresponding unique gene accessions were common. Including a previous proteomics set from *Lotus* pod and seed, the common gene accessions were decreased to 7%. Interestingly, an indication of more pronounced post translational modifications in nodules than in roots was determined. Between the two nodule developmental stages, higher levels of pathogen related 10 proteins, HSP's, and proteins

involved in redox processes were found in white nodules, suggesting a higher stress level at this developmental stage. In contrast, protein spots corresponding to nodulins such as leghemoglobin, asparagine synthetase, sucrose synthase, and glutamine synthetase were prevalent in red nodules. The distinct biochemical state of nodules was further highlighted by the conspicuous presence of several nitrilases, ascorbate metabolic enzymes and putative rhizobial effectors.

Introduction

Legumes are the third largest plant family with more than 20 000 species [1] and most of these are able to interact with rhizobia and develop symbiotic root nodules when available soil nitrogen is low. The bacteroids, differentiated rhizobia inside nodule cells, contain nitrogenase that converts atmospheric dinitrogen into ammonia/ammonium, which is used by the legume. In return, bacteroids are supplied with carbohydrates derived from photosynthesis. Within the nodule cells bacteroids are enclosed in a specialized plant membrane termed the symbiosome membrane separating bacteroids from the plant cytosol. A distinctive feature of a functional nodule is the red color originating from plant leghemoglobin, which is involved in buffering the oxygen level to avoid inactivation of the oxygen sensitive nitrogenase [2,3]. Formation of the nodule organ housing the bacteroids is an inducible process initiated when rhizobial lipochitin-oligosaccharide signal molecules are recognized by plant Nod factor receptors [4-7]. Plant genes such as *SymRK* and *CCaMK* including spontaneous nodule formation 1 (*snf1*), pivotal for developing a functional nodule, have been identified using forward genetics [8-11]. As a result, pathways involved in nodule development and resources for global transcriptome analysis [12,13] have provided a solid foundation for understanding symbiosis.

Biochemical characterization and analysis of the nodule proteome has also been initiated although at present this line of investigation is less well developed. In soybean, a total of 69 plant and bacteroid proteins were identified from the plant cytosolic fraction of soybean nodules [14]. In *Medicago*, nodules were separated into the bacteroid and plant fraction with 97 and 377 protein identifications, respectively [15]. The symbiosome membrane, important for transport of nutrients between the legume host and bacteroid, was analyzed in *Lotus*, soybean, and pea yielding 94, 17, and 46 protein identifications, respectively [16-18]. More recently, comparative proteomic studies of soybean nodules from a supernodulation mutant and wild type identified regulated proteins and early response proteins [19,20]. Comparative root proteomics is of interest because the nodule develops from root cells. In addition, data are available from roots subjected to various stress conditions such as flooding, temperature, and high salt stress [21-25].

In this study, we have analyzed six *Lotus* nodule and root plant cytosolic protein fractions to generate nodule and root proteome reference maps, each containing approximately 800 proteins identified from spots resolved in 2D gels. Furthermore, we used the PathExpress methodology [26] for pathway analysis and the ascorbate pathway was used to exemplify the potential of our data set. Finally, comparative proteomic analyses of four *Lotus* tissues and comparisons of proteins identified in nodule/symbiosome studies elucidate similarities and differences within *Lotus* tissues and between nodules from different legume species.

Materials and Methods

Plant and bacteria

For germination, *Lotus japonicus* Gifu B-129 and *snf1* (EMS 467.2; BC1; M5) seeds were treated for 15 min with concentrated sulfuric acid and rinsed in sterile H_2O. The seeds were surface sterilized with 1% bleach for 20 min, rinsed, and imbibed in sterile H_2O overnight. Seeds were sown in trays with sterilized Leca granulate clay watered with ¼ B&D medium [27] and grown in the greenhouse under conditions previously described [28]. A subset of the plants was inoculated with the *Mesorhizobium loti* strain MAFF303099 (*M. loti*) seven days after germination (DAG). Three nodule samples were prepared; wild type white nodules (18 days after inoculation (DAI)), wild type red nodules (25 DAI), and *snf1* red nodules (28 DAI), and these are collective termed the nodule subset. For each biological replicate, proteins from approximately 15 red nodules or 30 white nodules were separated using 2D gels. Three root samples were prepared; wild type roots (17 DAG), *snf1* roots (17 DAG), and *snf1* roots with spontaneous nodules (72 DAG), and these are collective termed the root subset. This subset was not inoculated and proteins were extracted from the whole root of all three samples.

Plant cytosolic protein extraction and 2D PAGE

Nodules and roots were harvested into pre-chilled mortars (-20°C) and homogenized in ice-cold extraction buffer (0.1 M Tris HCl (pH 8.0); 30% sucrose; 10 mM DTT; 1/100 (v/v) phosphatase inhibitor cocktail 1 and 2 from Sigma) and centrifuged at 12 000 x *g* for 15 min at 4°C. The protein concentration was measured according to Bradford [29] and subsequently SDS was added to 2%.

Sample and phenol (AppliChem, pH 8.0) was mixed 1:1, vortexed 30 sec, and centrifuged at 10 000 x *g* for 5 min at 4°C. The upper phase was transferred to a new tube and proteins were precipitated using 5 volumes of pre-chilled (-20°C) 100 mM ammonium acetate in methanol for 30 min and centrifuged at 10 000 x *g* for 5 min. The protein pellet was washed twice in 100 mM ice-cold ammonium acetate in methanol and twice in 80% acetone. If necessary, sonication was performed between the washes.

The protein pellet was air-dried and 500 µg protein was dissolved in 450 µL IEF buffer (7 M urea, 2 M thiourea, 4% (w/v) CHAPS, 1% (w/v) IPG buffer pH 4-7 (GE Healthcare), 10 mM Tris-HCl (pH 8.5), 10 mM DTT, and 0.002% bromphenol blue), centrifuged at 33 000 x *g*, and rehydrated using the 24 cm pH 4-7 strips (GE Healthcare) as previously described [30]. Three biological replicates were performed for each of the six tissue types. The 2D PAGE separation, staining, and scanning were performed as previously described [31].

Protein spot quantification

PDQuest Advanced software (version 8.0.1, BIO-RAD) was used for protein spot quantification. The nine 2D gels obtained from either the nodule subset or root subset was analyzed separately. The PDQuest software aligns all nine 2D gel images from the nodule or root subset and suggests the correlation of spots between the 2D gel images, then a master 2D gel was created that represents all spots detected in at least one tissue type. To perform a more accurate alignment, a manual inspection was performed. The three biological replicates for each tissue type were initially inspected followed by an inspection of all nine 2D gel images for each subset. For quantification, the spots were normalized using the method of local regression and data were exported to an Excel sheet.

To determine quantitative differences in protein spot levels between white and red nodules, two different calculation methods were used. To calculate the volume difference, the average volume of wild type and *snf1* red nodules was calculated and subtracted from the corresponding volume of the protein spot from white nodules. To calculate fold changes, the average volume of wild type and *snf1* red nodules was divided by the volume from the corresponding white nodule spot.

The p values between samples were calculated using the Prism (version 6) software and the multiple t tests with the following settings: 0.05 for statistical significant and the Holm-Sidak correction method for multiple comparisons.

Mass spectrometry and protein identifications

Spots were excised from 2D gels, digested with trypsin, and MS was performed using a 4800 *Plus* MALDI TOF/TOF Analyzer (AB SCIEX) as previously described [31]. Briefly, the MS was in positive reflector mode and MS profiles were collected in the *m/z* range 838-3500 Da. The most intense ions not corresponding to trypsin and human keratin were subjected to MS/MS analysis. The MS and MS/MS spectra were combined into one Mascot generic file and searched using Mascot version 2.3.01 (Matrix science) against an in-house *Lotus* database version 2.5 (37971 sequences; 9782974 residues) containing predicted genes available from http://www.kazusa.or.jp/lotus/. The following search settings were employed: carbamidomethyl (C) as fixed modification; oxidation (M) as variable modification; peptide mass tolerance at 70 ppm; trypsin as cleavage enzyme; one missed cleavage as maximum; and MS/MS tolerance at 0.5 Da. The significance threshold was 0.05. If no significant hit was obtained, a search against the nonredundant proteins at the National Center for Biotechnology Information [NCBInr, version 20111212 (6937173 sequences; 2395285404 residues) was

performed. Mascot results were parsed using MS Data Miner v. 1.1 [32] and all significant hits were manually inspected and only accepted if a peptide sequence tag of at least three consecutive b and/or y ions were observed. Finally, the excised protein spot numbers were manually annotated to the master gel from the PDQuest analysis to link identification and quantification data (see Supporting Information Tables 1 and 2 for the nodule and root data set, respectively).

Pathway analysis and blastp of identified symbiosome/nodule proteins between studies

Functional classes for the identified proteins were determined based on homology relationships with reviewed proteins in UniProtKB (Release-2012_02) [33], using the BLAST program [34]. For all proteins, the best match, with an expected threshold e-value less than 1E-05, was selected and functional annotations (UniPathways [35], keywords, and Gene Ontology terms [36]) were extracted from the entry content and assigned to the protein (Supporting Information Table 3).

To determine whether a particular functional category was statistically overrepresented in a nodule or root proteins compared to all identified proteins, the p-value for all terms in each functional ontology was calculated using the hypergeometric distribution, with multiple testing correction (FDR), as described in PathExpress system [26]. This p-value represents the probability that the intersection of the group of proteins belonging to the given class occurs by chance. Results from nodule and root proteins were compared, to identify for each functional category the proteins common and unique to each subset (Supporting Information Table 4).

To determine the overlap of homologous proteins between studies, the identified gene accessions from five previous symbiosome/nodule proteomics studies were extracted [14-18].

Gene accessions from the three symbiosome studies were combined into one dataset. Furthermore, only the plant gene accessions from *Medicago* were included together with *Lotus* nodule and root gene accessions identified in this study. All non-protein sequences were translated into the longest coding sequence. Within all three datasets together with the *Lotus* nodule and root datasets, the gene accession duplicates were subtracted (between *Lotus* nodule and seed, identical gene accessions occur) and a total of 1811 different gene accessions from the five datasets were obtained. A blastp between gene accessions from the five datasets was performed and the threshold for homologous protein identification between datasets was less than 1E-05 (Supporting Information Table 5).

Website for the proteomics data

All nodule and root proteomics data are available at the homepage https://www.cbs.dtu.dk/cgi-bin/lotus2_5/db.cgi. The structure together with the navigation on the homepage has been previously described [28,31]. NCBI blast v. 2.2.18 [37] was used to align and link protein sequences (program tblastn, expectation value 1E-08) and DNA sequences (using blastn, expectation value 1E-08) to genome sequences.

Results and Discussion

***Lotus* nodule and root proteome reference maps**

The founding steps in a proteome analysis are identification and quantification of proteins present in cells, tissues, or organs. For this purpose, 2D gel reference maps based on master gels give a useful overview together with an indication of the level of post translational modifications occurring. In this study, we have established proteome reference maps of the *Lotus* nodule in two developmental stages and the root. In total, six different tissue types

from wild type and *snf1* plants were harvested and divided into two subsets for the subsequent 2D gel analysis (see Supporting Information Fig. 1 for schematic overview of experimental set-up). The nodule subset encompassed wild type white nodules, wild type red nodules, and inoculated *snf1* red nodules. The root subset encompassed wild type roots, *snf1* roots, and *snf1* roots with spontaneous nodules (hereafter termed *snf1* root and nodules). To establish a comparable analysis avoiding the complication of rhizobial proteins a procedure preferentially extracting plant cytosolic proteins was used. The proteome reference maps (master gels) of nodule and root subsets display overlapping patterns of protein spots (see Supporting Information Fig. 1 and 2). Protein spots were thus excised from both the nodule and root gels to obtain positive identifications and avoid erroneous correlation between protein identification and quantification between the two subsets. In total, 834 and 895 protein spots were excised for MS analysis from the nodule and root subset respectively. With significance threshold at 0.05 (lowest protein score and coverage were 28 and 0.8%, respectively) and manual inspection of all MS/MS spectra to contain a peptide sequence of at least three consecutive b and/or y ions, protein identification was obtained for 780 nodule spots and 790 root spots. Gene accession(s) and quantitative data for each identified spot are shown in the Supporting Information Tables 1 and 2. Furthermore, at the homepage https://www.cbs.dtu.dk/cgi-bin/lotus2_5/db.cgi, the master gels with corresponding spot numbers and experimental data are stored and can be examined.

Proteins identified in nodule and root: differences and similarities

The separation of nodules and roots enabled the elucidation of specific and common proteins in each tissue type. Supporting Information Fig. 1 shows the total number of identified gene accessions, unique gene accessions, and intersecting gene accessions. For both subsets, a total

of approximately 1100 gene accessions were identified and, unexpectedly, approximately 20% additional unique gene accessions were identified in the root. This indicates a higher post translational modification (PTM) level in the nodule, and comparing the 2D gels, more horizontal and vertical "string of pearls", are visible in nodule 2D gels than in root 2D gels. Fig. 1A shows an example where the pattern of eleven spots (7607 to 7617) indicates PTMs affecting pI and molecular mass of the identified protein (gene accession chr3.CM0996.1250.r2.d) corresponding to a putative aspartic proteinase, called nodulin 41 in common bean [38]. Further enrichment for PTM peptides was not performed in this study and as expected it was therefore not possible to assign peptides with different PTMs using the error tolerance search feature (data not shown). In conclusion, the lower number of proteins identified and the spot pattern for the nodules indicates a higher level of PTM. This observation may reflect prompt adaptation under the different development stages of the fast growing nodule compared to the more steadily growing root.

Approximately 45% (363) of the unique gene accessions were identified in both nodule and root. As expected, leghemoglobin, the hallmark of mature nodules, was only identified in nodules while no obviously root specific proteins could be determined for the 275 corresponding gene accessions identified in root alone. For a more systematic analysis, all predicted *Lotus* genes were searched against the RefSeq database and annotated (Supporting Information Table 3). In Supporting Information Table 4, the grouping of identified nodule and root gene accessions into pathways is shown. The majority of enzymes/proteins involved in ascorbate metabolism were identified in nodule and/or root, as shown in Fig. 2. Ascorbate is an important antioxidant and an increased ascorbate level in *Lotus* nodule compared to root has been measured previously [39]. Our protein identifications support this observation. The putative GDP-D-mannose 3',5'-epimerase (GME) which synthesizes GDP-L-galactose and

GDP-L-gulose using GDP-D-mannose was solely identified in several spots from the *Lotus* nodule and was previously suggested to be positively correlated with the ascorbate level [40]. The GDP-L-galactose pathway is a major pathway for ascorbate synthesis, whereas the GDP-L-gulose pathway was suggested as an alternative pathway [41]. In *Lotus*, a putative phosphomannomutase (PMM) and putative L-galactose dehydrogenase (L-GalDH) were identified in nodule/root and root, respectively. PMM is suggested as a regulator for the ascorbate level whereas L-GalDH is suggested to be a feedback inhibitor of ascorbate in pea seedlings while no regulatory function was found in tobacco [42-46]. A third pathway for synthesis of ascorbate is a salvage pathway (not illustrated in Fig. 2) and several *Lotus* D-galacturonate reductase proteins were identified in nodule/root or only in root. However, with our current knowledge, the *Lotus* D-galacturonate reductase used for ascorbate synthesis could not be identified. Surprisingly, more protein spots from root than nodule were identified to be involved in redox processes of ascorbate i.e. cytosolic ascorbate peroxidase (APX), monodehydroascorbate reductase, and dehydroascorbate reductase (DHAR). This was unexpected as leghemoglobins are thought to contribute significantly to ROS together with a high respiration rate in nodules, which is probably a significant factor for the higher ascorbate level in the nodule than in the root [47].

Altogether, the differences in the ascorbate pathway between nodule and root, especially, the identification of GME in several spots from the nodule, predominantly support the finding of a higher ascorbate level in the nodule than in the root.

Protein regulation during nodule maturation

White and red wild type nodules were analyzed together with inoculated *snf1* red nodules to identify regulated proteins/pathways expressed in the fast growing nodule (see Supporting

Information Fig. 1). Protein spot volumes for corresponding spots of wild type and *snf1* red nodules were similar and, thus, the average volume was used to determine differences between white and red nodules. To determine the difference between white and red nodules, either the volume difference or fold change can be used. In Fig 1B to 1E the most differential volumes and fold changes between white and red nodule spots are shown and in Table 1 the ten most regulated protein spots, as determined by either volume difference or fold change, are shown. By using volume difference, the protein spots that were not quantified in either white or red nodules can potentially be included and, thus, these identifications are discussed further. Interestingly, protein spots corresponding to pathogen related (PR) 10 proteins had higher volume in white nodules (see Table 1). Spots 7281, 7456, 7286, and 7294 were identified with two gene accessions; LjSGA_063085.1 and chr6.CM0573.230.r2.m, which were all in the top seven of spots with the highest difference in spot volume (highest white nodules). Given that the two *Lotus* PR10 proteins were identified in both nodule and root (see below), it is likely that they participate in a more general defense mechanism rather than a specific symbiotic defense mechanism. In *Medicago*, mRNA for a closely related *PR10* gene (*MtN13*) is nodule specific [48], whereas the PR10 *Medicago* gene *MsPR10-1* is primarily expressed in uninoculated roots. Following infection by the pathogenic *Pseudomonas syringae*, *MsPR10-1* is expressed in both roots and leaves [48]. In comparison, the two *Lotus PR10* genes are more similar to *MsPR10-1* gene than to *MtN13* (data not shown). Furthermore, the volume of putative HSP's was higher (spots 7936 and 7938) in white *Lotus* nodules. HSP's are known to be important for cell survival under stress conditions [49] and together with the higher level of PR10 proteins this indicates a partial defense reaction occurs in white nodules than in red nodules. As expected, the highest volume/fold change differences corresponding to leghemoglobins, sucrose synthases, asparagine synthetase, and

glutamine synthetase were found in the red *Lotus* nodules.

Mesorhizobium loti proteins identified from the plant nodule cytosolic fraction

The *M. loti* proteins identified in the nodule subset were excluded from the comparison of nodule and root presented above. In total, 31 protein spots corresponding to 24 unique *M. loti* accessions (Table 2) were found suggesting that the method used for extraction of plant cytosolic proteins was adequate. Quantified bacteroid protein spots were present in red nodules, both wild type and *snf1*, with similar volumes. In white nodules, 50% of these spots were not visible and 50% had lower/similar volumes probably reflecting the difference in bacteroid content between the two nodule developmental stages. Several of the identified *M. loti* proteins are annotated as outer membrane or periplasmic proteins. This indicates that some symbiosome membranes and bacterial outer membranes might have been disrupted during extraction of plant cytosolic proteins. Compatible observations were made in an analysis of soybean nodules. Here seven out of 69 identified protein spots from the plant cytosolic fraction corresponded to bacteroid proteins [14]. Four of these, 1-aminocyclopropane-1-carboxylate-deaminase, nitrogenase β-chain, and two GroEL chaperonins, are homologous to *M. loti* proteins identified in our analysis, whereas the rhizobial ABC transporter identified in both studies shows no similarity. Furthermore, a comparison of bacteroid proteins identified from the *Medicago truncatula* symbiosome membrane and proteins identified in this study shows three homologous proteins; GroEL chaperonin, 30S ribosomal protein, and sulfate binding protein of ABC transporter [50].

Several of the identified *M. loti* proteins are encoded on the symbiosis island [51,52]. Many of the corresponding genes are directly regulated by the master regulator NifA and are strongly expressed in bacteroids, including *mll5873* (*msi334* in *M. loti* strain R7A), *mlr5906*

(*nifD*), *mlr5907* (*nifK*), *mlr5932* (*acdS*), *mlr5943* (*msi260*), *mlr5944* (*msi259*) and *mll6127* (*msi158*) [53].

Interestingly, the outer membrane protein (gi|13475117) encoded by *mll6127* was also identified in a previous study of the *Lotus* symbiosome membrane [18] and is essential for efficient nitrogen fixation [53]. Several of the other identified proteins are also encoded on the symbiosis island but the regulation of the corresponding genes (*mlr5787*, *mll5810*, *mll5998*, *mll6062*, *mlr6296*) has not been reported. The identification of the proteins here suggests that the genes are also strongly expressed in bacteroids.

Lotus compatible *Mesorhizobium* strains possess either a type III secretion system, including the *M. loti* MAFF303099 strain used in this study, or a type IV secretion system capable of secreting effector proteins into the plant cytoplasm [54]. Using the effector predicting software, EffectiveT3 [51,55], approximately 9% (661 of 7272) of predicted *M. loti* proteins are putative effectors (data not shown). This seems a large fraction but is similar to the prediction for *Bradyrhizobium japonicum* USDA110, a soybean symbiont [55]. Two of the identified *M. loti* proteins, a hypothetical protein (gi|13474949) and a 30S ribosomal protein (gi|13474344), were predicted as putative effectors. However it is still an open question whether the hypothetical protein is a true effector as the corresponding gene *mlr5944* is part of a NifA-regulated operon located outside of the cluster of genes encoding the Type III secretion system.

In conclusion, the level of *M. loti* protein contamination appeared to be low when using the described cytosolic plant protein extraction protocol. However, further *in vivo* studies are needed to verify the *M. loti* protein localization.

The proteome of root (wild type and *snf1*) and *snf1* root nodules

The root subset (see Supporting Information Fig. 1), wild type roots and *snf1* roots, were similar in volumes for the majority of corresponding protein spots. For the *snf1* root nodules, some spot volumes were different and higher volumes were for example found for two putative PR10 proteins; LjSGA_063085.1 and chr6.CM0573.230.r2.m. These were also found regulated in the nodule set. In roots (wild type and *snf1*), the spot volumes of several putative HSP's (LjSGA_048339.1, chr2.CM0210.60.r2.m, LjSGA_038906.1 and chr3.CM0786.370.r2.a) and detoxification enzymes such as APX1, GST, and DHAR2, (chr3.CM0616.30.r2.d, LjSGA_072403.1, chr5.CM0180.360.r2.d and LjSGA_044448.2) had higher volumes. This indicates a higher level of control of different stress factors in root (wild type and *snf1*) compared to *snf1* root nodules. Since there was only a small intersection between *snf1* roots carrying spontaneous nodules, we infer that the differences observed most likely correspond to a later developmental stage of the *snf1* root nodules. In conclusion, to obtain more detailed information about proteins present and regulated in *snf1* spontaneous nodules, this tissue has to be harvested separately.

Comparative proteomic analysis within *Lotus* tissues and other nodule/symbiosome studies

To obtain more knowledge of proteins represented in different *Lotus* tissues, which can be the initial step to identify crucial proteins in, for example, nitrogen transport, a comparative analysis was performed using data from this study together with a proteomic data set from *Lotus* pod and seed [31]. All identified gene accessions for pod and seed were updated and renamed according to the current version of the *Lotus* genome (version 2.5). In total, 1327 unique gene accessions were compared (see Supporting Information Table 5) and the intersection between nodule, root, pod, and seed data was approximately 7% (96 gene

accessions) whilst more than 56% (748 gene accessions) were identified in only one tissue (Fig. 3A). Four putative nitrilase enzymes, chr1.CM0133.900.r2.m, chr3.CM0164.140.r2.d, chr3.CM0164.160.r2.d, and chr5.CM0180.950.r2.d, were only identified in *Lotus* nodule, while chr3.CM0164.130.r2.d was identified in both nodule and seed. In Arabidopsis, nitrilase genes are divided into four subgroups *NIT 1-4* [56]. It is suggested that nitrilases are important in plant-bacteria symbiosis for nitrogen assimilation, hormone synthesis, and nitrile detoxification [57]. The function of identified *Lotus* nitrilases is unknown; however, the majority of identified *Lotus* spots are Arabidopsis *NIT 4* homologs, which are likely to be important for the detoxification pathway where cyanide production may be linked to defense pathways [58,59]. Alternatively, nitrilase could be a part of a nitrogen transport mechanism in the *Lotus* nodule, similar to plants colonized with bacteria using the ammonia released from bacterial nitrilase activity as a nitrogen source [60].

A total of 18 putative GST's including DHAR (Fig. 2) were identified in all four tissues. Interestingly, GST identifications were predominantly correlated to the root and nodule with 16 and nine gene accessions identified, whereas, only four and three putative GST are identified for the pod and seed. The primary role of GST is to detoxify cellular environments by conjugating toxic compounds to GSH and the data indicates a higher level of xenobiotics in root/nodule than pod/seed.

Proteomics studies of the nodule/symbiosome membrane have been performed in *Lotus* and other legumes. To compare the protein identifications in this study and five other nodule/symbiosome proteomics studies [14-18], blastp analysis of all identified gene accessions (see Supporting Information Table 5) was performed and data are shown in Fig. 3B and 3C. As expected, the intersection is higher when using *Lotus* nodules (Fig. 3B) than

Lotus roots (Fig. 3C). This type of analysis may help to identify the conserved proteins and pathways that are important for root nodule function.

Concluding remarks

The current comparative proteomic analysis of *Lotus* nodule and root tissues has identified proteins corresponding to more than 800 gene accessions from nodule and root. Post translational modifications appeared to be more frequent in the nodule than in the root. One of the *M. loti* proteins identified is encoded on the symbiosis island and may be an effector; however, further studies are needed to verify this. The intersection of identified gene accessions between nodule/root and pod/seed is low (Fig. 3) considering that a 2D gel approach was used for both studies. One bias could be the different protein extraction methods, where a plant cytosolic protein extraction and total protein extraction for nodule/root and pod/seed were used, respectively [31]. To elucidate the function of the putative nitrilases, predominantly identified in the nodule, plant lines from the large LORE1 retrotransposon insert mutant population [61,62] that have a LORE1 insert in the identified nitrilase gene accessions can be used.

Acknowledgements

We thank Fleur C. Dolman for suggestions on the article. This work was supported by the Danish National Research Foundation grant no. DNRF79 and the Danish Council for Independent Research | Technology and Production Sciences (FTP).

Conflict of interest

The authors have declared no conflicts of interest.

Supporting Information Available

Two supporting information figure and five supporting information tables

Supporting Information Fig. S1: Experimental setup and proteins identified.

Supporting Information Fig. S2: 2D gels for all six tissues together with nodule and root master gels.

Supporting Information Table 1: Identification and quantification of protein spots from the nodule.

Supporting Information Table 2: Identification and quantification of protein spots from the root.

Supporting Information Table 3: Putative annotation of predicted *Lotus* genes.

Supporting Information Table 4: Pathway similarities and differences between the nodule and root.

Supporting Information Table 5: Unique proteins identified in different *Lotus* tissues together with nodule/symbiosome membrane proteins identified from *Lotus*, soybean, and pea.

References

[1] Doyle JJ, Luckow MA. The rest of the iceberg. Legume diversity and evolution in a phylogenetic context. *Plant Physiol* 2003, 131, 900-10.

[2] Madsen O, Sandal L, Sandal NN, Marcker KA. A soybean coproporphyrinogen oxidase gene is highly expressed in root nodules. *Plant Mol Biol* 1993, 23, 35-43.

[3] Ott T, van Dongen JT, Günther C, Krusell L, Desbrosses G, Vigeolas H, Bock V, Czechowski T, Geigenberger P, Udvardi MK. Symbiotic leghemoglobins are crucial for nitrogen fixation in legume root nodules but not for general plant growth and development. *Curr Biol* 2005, 15, 531-5.

[4] Radutoiu S, Madsen LH, Madsen EB, Felle HH, Umehara Y, Gronlund M, Sato S, Nakamura Y, Tabata S, Sandal N, Stougaard J. Plant recognition of symbiotic bacteria requires two LysM receptor-like kinases. *Nature* 2003, 425, 585-92.

[5] Madsen EB, Madsen LH, Radutoiu S, Olbryt M, Rakwalska M, Szczyglowski K, Sato S, Kaneko T, Tabata S, Sandal N, Stougaard J. A receptor kinase gene of the LysM type is involved in legume perception of rhizobial signals. *Nature* 2003, 425, 637-40.

[6] Limpens E, Franken C, Smit P, Willemse J, Bisseling T, Geurts R. LysM domain receptor kinases regulating rhizobial Nod factor-induced infection. *Science* 2003, 302, 630-3.

[7] Broghammer A, Krusell L, Blaise M, Sauer J, Sullivan JT, Maolanon N, Vinther M, Lorentzen A, Madsen EB, Jensen KJ, Roepstorff P, Thirup S, Ronson CW, Thygesen MB, Stougaard J. Legume receptors perceive the rhizobial lipochitin oligosaccharide signal molecules by direct binding. *Proc Natl Acad Sci U S A* 2012, 109, 13859-64.

[8] Madsen LH, Tirichine L, Jurkiewicz A, Sullivan JT, Heckmann AB, Bek AS, Ronson CW, James EK, Stougaard J. The molecular network governing nodule organogenesis and infection in the model legume *Lotus japonicus*. *Nat Commun* 2010, 1, 10.

[9] Tirichine L, Imaizumi-Anraku H, Yoshida S, Murakami Y, Madsen LH, Miwa H, Nakagawa T, Sandal N, Albrektsen AS, Kawaguchi M, Downie A, Sato S, Tabata S, Kouchi H, Parniske M, Kawasaki S, Stougaard J. Deregulation of a Ca2+/calmodulin-dependent kinase leads to spontaneous nodule development. *Nature* 2006, 441, 1153-6.

[10] Stracke S, Kistner C, Yoshida S, Mulder L, Sato S, Kaneko T, Tabata S, Sandal N, Stougaard J, Szczyglowski K, Parniske M. A plant receptor-like kinase required for both bacterial and fungal symbiosis. *Nature* 2002, 417, 959-62.

[11] Tirichine L, Sandal N, Madsen LH, Radutoiu S, Albrektsen AS, Sato S, Asamizu E, Tabata S, Stougaard J. A gain-of-function mutation in a cytokinin receptor triggers spontaneous root nodule organogenesis. *Science* 2007, 315, 104-7.

[12] Høgslund N, Radutoiu S, Krusell L, Voroshilova V, Hannah MA, Goffard N, Sanchez DH, Lippold F, Ott T, Sato S, Tabata S, Liboriussen P, Lohmann GV, Schauser L, Weiller GF, Udvardi MK, Stougaard J. Dissection of symbiosis and organ development by integrated transcriptome analysis of *Lotus japonicus* mutant and wild-type plants. *PLoS One* 2009, 4, e6556.

[13] Verdier J, Torres-Jerez I, Wang M, Andriankaja A, Allen SN, He J, Tang Y, Murray JD, Udvardi MK. Establishment of the *Lotus japonicus* Gene Expression Atlas (LjGEA) and its use to explore legume seed maturation. *Plant J* 2013, 74, 351-62.

[14] Oehrle NW, Sarma AD, Waters JK, Emerich DW. Proteomic analysis of soybean nodule cytosol. *Phytochemistry* 2008, 69, 2426-38.

[15] Larrainzar E, Wienkoop S, Weckwerth W, Ladrera R, Arrese-Igor C, Gonzalez EM. *Medicago truncatula* root nodule proteome analysis reveals differential plant and bacteroid responses to drought stress. *Plant Physiol* 2007, 144, 1495-507.

[16] Panter S, Thomson R, de Bruxelles G, Laver D, Trevaskis B, Udvardi M. Identification with proteomics of novel proteins associated with the peribacteroid membrane of soybean root nodules. *Mol Plant Microbe Interact* 2000, 13, 325-33.

[17] Saalbach G, Erik P, Wienkoop S. Characterisation by proteomics of peribacteroid space and peribacteroid membrane preparations from pea (*Pisum sativum*) symbiosomes. *Proteomics* 2002, 2, 325-37.

[18] Wienkoop S, Saalbach G. Proteome analysis. Novel proteins identified at the peribacteroid membrane from *Lotus japonicus* root nodules. *Plant Physiol* 2003, 131, 1080-90.

[19] Lim CW, Park JY, Lee SH, Hwang CH. Comparative proteomic analysis of soybean nodulation using a supernodulation mutant, SS2-2. *Biosci Biotechnol Biochem* 2010, 74, 2396-404.

[20] Salavati A, Bushehri AA, Taleei A, Hiraga S, Komatsu S. A comparative proteomic analysis of the early response to compatible symbiotic bacteria in the roots of a supernodulating soybean variety. *J Proteomics* 2012, 75, 819-32.

[21] Nanjo Y, Skultety L, Uvackova L, Klubicova K, Hajduch M, Komatsu S. Mass spectrometry-based analysis of proteomic changes in the root tips of flooded soybean seedlings. *J Proteome Res* 2012, 11, 372-85.

[22] Salavati A, Khatoon A, Nanjo Y, Komatsu S. Analysis of proteomic changes in roots of soybean seedlings during recovery after flooding. *J Proteomics* 2012, 75, 878-93.

[23] Dumont E, Bahrman N, Goulas E, Valot B, Sellier H, Hilbert JL, Vuylsteker C, Lejeune-Hénaut I, Delbreil B. A proteomic approach to decipher chilling response from cold acclimation in pea (*Pisum sativum* L.). *Plant Sci* 2011, 180, 86-98.

[24] Ahsan N, Donnart T, Nouri MZ, Komatsu S. Tissue-specific defense and thermo-adaptive mechanisms of soybean seedlings under heat stress revealed by proteomic approach. *J Proteome Res* 2010, 9, 4189-204.

[25] Rodriguez-Celma J, Lattanzio G, Grusak MA, Abadia A, Abadia J, Lopez-Millan AF. Root responses of *Medicago truncatula* plants grown in two different iron deficiency conditions, changes in root protein profile and riboflavin biosynthesis. *J Proteome Res* 2011, 10, 2590-601.

[26] Goffard N, Weiller G. PathExpress, a web-based tool to identify relevant pathways in gene expression data. *Nucleic Acids Res* 2007, 35, W176-81.

[27] Broughton WJ, Dilworth MJ. Control of leghaemoglobin synthesis in snake beans. *Biochem J* 1971, 125, 1075-80.

[28] Dam S, Laursen BS, Ørnfelt JH, Jochimsen B, Staerfeldt HH, Friis C, Nielsen K, Goffard N, Besenbacher S, Krusell L, Sato S, Tabata S, Thøgersen IB, Enghild JJ, Stougaard

J. The proteome of seed development in the model legume *Lotus japonicus*. *Plant Physiol* 2009, 149, 1325-40.

[29] Bradford M. A rapid and sensitive method for the quantitation of microgram quantities of protein utilizing the principle of protein-dye binding. *Anal Biochem* 1976, 72, 248-54.

[30] Zheng X, Hong L, Shi L, Guo J, Sun Z, Zhou J. Proteomics analysis of host cells infected with infectious bursal disease virus. *Mol Cell Proteomics* 2008, 7, 612-25.

[31] Nautrup-Pedersen G, Dam S, Laursen BS, Siegumfeldt AL, Nielsen K, Goffard N, Staerfeldt HH, Friis C, Sato S, Tabata S, Lorentzen A, Roepstorff P, Stougaard J. Proteome analysis of pod and seed development in the model legume *Lotus japonicus*. *J Proteome Res* 2010, 9, 5715-26.

[32] Dyrlund TF, Poulsen ET, Scavenius C, Sanggaard KW, Enghild JJ. MS Data Miner, a web-based software tool to analyze, compare, and share mass spectrometry protein identifications. *Proteomics* 2012, 12, 2792-6.

[33] UniProt Consortium. Update on activities at the Universal Protein Resource (UniProt) in 2013. *Nucleic Acids Res* 2013, 41, D43-7.

[34] Altschul SF, Gish W, Miller W, Myers EW, Lipman DJ. Basic local alignment search tool. *J Mol Biol* 1990, 215, 403-10.

[35] Morgat A, Coissac E, Coudert E, Axelsen KB, Keller G, Bairoch A, Bridge A, Bougueleret L, Xenarios I, Viari A. UniPathway: a resource for the exploration and annotation of metabolic pathways. *Nucleic Acids Res* 2012, 40, D761-9.

[36] Ashburner M, Ball CA, Blake JA, Botstein D, Butler H, Cherry JM, Davis AP, Dolinski K, Dwight SS, Eppig JT, Harris MA, Hill DP, Issel-Tarver L, Kasarskis A, Lewis S, Matese JC, Richardson JE, Ringwald M, Rubin GM, Sherlock G. Gene ontology: tool for the unification of biology. The gene ontology consortium. *Nat Genet* 2000, 25, 25-9.

[37] Altschul SF, Madden TL, Schaffer AA, Zhang J, Zhang Z, Miller W, Lipman DJ. Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. *Nucleic Acids Res* 1997, 25, 3389-402.

[38] Olivares JE, Díaz-Camino C, Estrada-Navarrete G, Alvarado-Affantranger X, Rodriguez-Kessler M, Zamudio FZ, Olamendi-Portugal T, Márquez Y, Servín LE, Sánchez F. Nodulin 41, a novel late nodulin of common bean with peptidase activity. *BMC Plant Biol* 2011, 11, 2229-11-134.

[39] Matamoros MA, Loscos J, Coronado MJ, Ramos J, Sato S, Testillano PS, Tabata S, Becana M. Biosynthesis of ascorbic acid in legume root nodules. *Plant Physiol* 2006, 141, 1068-77.

[40] Wolucka BA, Goossens A, Inze D. Methyl jasmonate stimulates the *de novo* biosynthesis of vitamin C in plant cell suspensions. *J Exp Bot* 2005, 56, 2527-38.

[41] Wolucka BA, Van Montagu M. GDP-mannose 3',5'-epimerase forms GDP-L-gulose, a putative intermediate for the *de novo* biosynthesis of vitamin C in plants. *J Biol Chem* 2003, 278, 47483-90.

[42] Dowdle J, Ishikawa T, Gatzek S, Rolinski S, Smirnoff N. Two genes in *Arabidopsis thaliana* encoding GDP-L-galactose phosphorylase are required for ascorbate biosynthesis and seedling viability. *Plant J* 2007, 52, 673-89.

[43] Qian W, Yu C, Qin H, Liu X, Zhang A, Johansen IE, Wang D. Molecular and functional analysis of phosphomannomutase (PMM) from higher plants and genetic evidence for the involvement of PMM in ascorbic acid biosynthesis in *Arabidopsis* and *Nicotiana benthamiana*. *Plant J* 2007, 49, 399-413.

[44] Hoeberichts FA, Vaeck E, Kiddle G, Coppens E, van de Cotte B, Adamantidis A, Ormenese S, Foyer CH, Zabeau M, Inze D, Perilleux C, Van Breusegem F, Vuylsteke M. A temperature-sensitive mutation in the *Arabidopsis thaliana* phosphomannomutase gene disrupts protein glycosylation and triggers cell death. *J Biol Chem* 2008, 283, 5708-18.

[45] Pallanca JE, Smirnoff N. The control of ascorbic acid synthesis and turnover in pea seedlings. *J Exp Bo*t 2000, 51, 669-74.

[46] Gatzek S, Wheeler GL, Smirnoff N. Antisense suppression of L-galactose dehydrogenase in *Arabidopsis thaliana* provides evidence for its role in ascorbate synthesis and reveals light modulated L-galactose synthesis. *Plant J* 2002, 30, 541-53.

[47] Günther C, Schlereth A, Udvardi M, Ott T. Metabolism of reactive oxygen species is attenuated in leghemoglobin-deficient nodules of *Lotus japonicus*. *Mol Plant Microbe Interact* 2007, 20, 1596-603.

[48] Gamas P, de Billy F, Truchet G. Symbiosis-specific expression of two *Medicago truncatula* nodulin genes, *MtN1* and *MtN13*, encoding products homologous to plant defense proteins. *Mol Plant Microbe Interact* 1998, 11, 393-403.

[49] Rajan VB, D'Silva P. *Arabidopsis thaliana* J-class heat shock proteins: cellular stress sensors. *Funct Integr Genomics* 2009, 9, 433-46.

[50] Catalano CM, Lane WS, Sherrier DJ. Biochemical characterization of symbiosome membrane proteins from *Medicago truncatula* root nodules. *Electrophoresis* 2004, 25, 519-31.

[51] Kaneko T, Nakamura Y, Sato S, Asamizu E, Kato T, Sasamoto S, Watanabe A, Idesawa K, Ishikawa A, Kawashima K, Kimura T, Kishida Y, Kiyokawa C, Kohara M, Matsumoto M, Matsuno A, Mochizuki Y, Nakayama S, Nakazaki N, Shimpo S, Sugimoto M, Takeuchi C, Yamada M, Tabata S. Complete genome structure of the nitrogen-fixing symbiotic bacterium *Mesorhizobium loti*. *DNA Res* 2000, 7, 331-8.

[52] Sullivan JT, Trzebiatowski JR, Cruickshank RW, Gouzy J, Brown SD, Elliot RM, Fleetwood DJ, McCallum NG, Rossbach U, Stuart GS, Weaver JE, Webby RJ, De Bruijn FJ, Ronson CW. Comparative sequence analysis of the symbiosis island of *Mesorhizobium loti* strain R7A. *J Bacteriol* 2002, 184, 3086-95.

[53] Sullivan JT, Brown SD, Ronson CW. The NifA-RpoN regulon of *Mesorhizobium loti* strain R7A and its symbiotic activation by a novel LacI/GalR-family regulator. *PLoS One* 2013, 8, e53762.

[54] Sánchez C, Iannino F, Deakin WJ, Ugalde RA, Lepek VC. Characterization of the *Mesorhizobium loti* MAFF303099 type-three protein secretion system. *Mol Plant Microbe Interact* 2009, 22, 519-28.

[55] Arnold R, Brandmaier S, Kleine F, Tischler P, Heinz E, Behrens S, Niinikoski A, Mewes HW, Horn M, Rattei T. Sequence-based prediction of type III secreted proteins. *PLoS Pathog* 2009, 5, e1000376.

[56] Bartel B, Fink GR. Differential regulation of an auxin-producing nitrilase gene family in *Arabidopsis thaliana*. *Proc Natl Acad Sci U S A* 1994, 91, 6649-53.

[57] Howden AJ, Preston GM. Nitrilase enzymes and their role in plant-microbe interactions. *Microb Biotechnol* 2009, 2, 441-51.

[58] Piotrowski M. Primary or secondary? Versatile nitrilases in plant metabolism. *Phytochemistry* 2008, 69, 2655-67.

[59] Piotrowski M, Schönfelder S, Weiler EW. The *Arabidopsis thaliana* isogene NIT4 and its orthologs in tobacco encode beta-cyano-L-alanine hydratase/nitrilase. *J Biol Chem* 2001, 276, 2616-21.

[60] Howden AJ, Harrison CJ, Preston GM. A conserved mechanism for nitrile metabolism in bacteria and plants. *Plant J* 2009, 57, 243-53.

[61] Urbanski D, Malolepszy A, Stougaard J, Andersen S. Genome-wide *LORE1* retrotransposon mutagenesis and high-throughput insertion detection in *Lotus japonicus*. *Plant J* 2012, 69, 731-41.

[62] Fukai E, Soyano T, Umehara Y, Nakayama S, Hirakawa H, Tabata S, Sato S, Hayashi M. Establishment of a *Lotus japonicus* gene tagging population using the exon-targeting endogenous retrotransposon *LORE1*. *Plant J* 2012, 69, 720-30.

Figure 1: Close up of 2D gel sections. A), horizontal and vertical spots (7607 to 7617) from nodules identified with identical gene accession (chr3.CM0996.1250.r2.d) indicating PTMs. B) and C), shows the protein spots (7423 and 7304) with the highest volume differences between white and red nodules. D) and E), shows the protein spots (7269 and 7504) with the highest fold change differences between white and red nodules. A is a close up view from a 2D gel corresponding to red nodules whereas B to E the initial image corresponding to a 2D gel from white nodules and the second image is from red nodules.

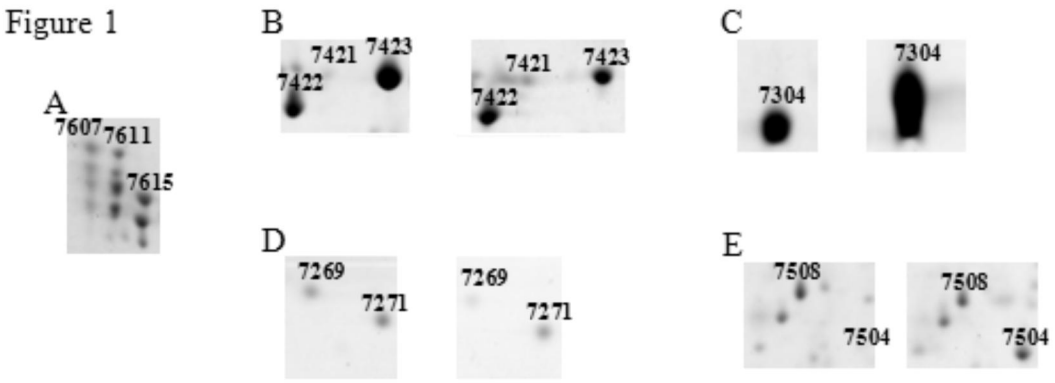


Figure 2: Enzymes involved in the ascorbate pathway identified in nodules and roots. The gene accessions in *Arabidopsis* coding for enzymes involved in the ascorbate pathway were searched against the *Lotus* genome. Several *Lotus* homologous were identified in the nodule and/or root. Identifications from the nodule subset are marked with (1) and identifications from the root subset are marked with (2). APX: cytosolic ascorbate peroxidase (nodules and roots: chr3.CM0616.30.r2.d); DHAR: dehydroascorbate reductase (roots: chr5.CM0180.360.r2.d); GLDH: L-galactono-1,4-lactone dehydrogenase; GME: GPD-D-mannose 3',5'-epimerase (nodules: LjSGA_138374.1, LjSGA_049053.1, LjSGA_117719.1); GR: glutathione reductase; L-GalDH: L-galactose dehydrogenase (roots: chr4.CM1616.90.r2.d); MDAR: monodehydroascorbate reductase (nodules: chr4.CM0004.930.r2.d, nodules and roots: LjSGA_039892.1); PMM: phosphomannomutase (nodules and roots: chr1.CM0284.440.r2.a); VTC: vitamin C; VTC4 (nodules and roots: chr6.CM0367.220.r2.a). For VTC2/VTC5 no homologous were found in *Lotus*. For further details, see text.

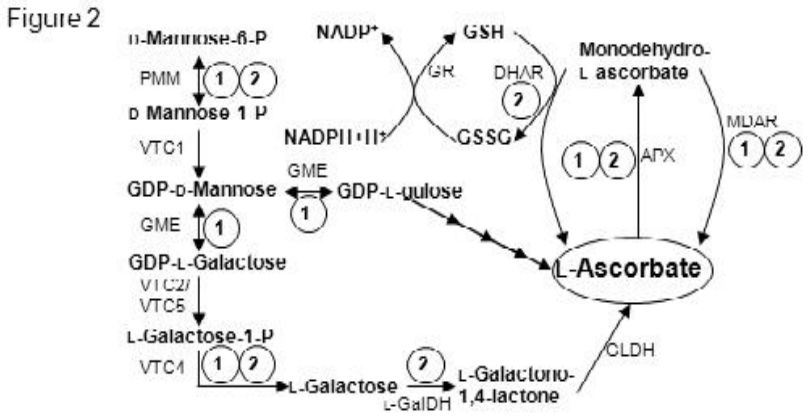
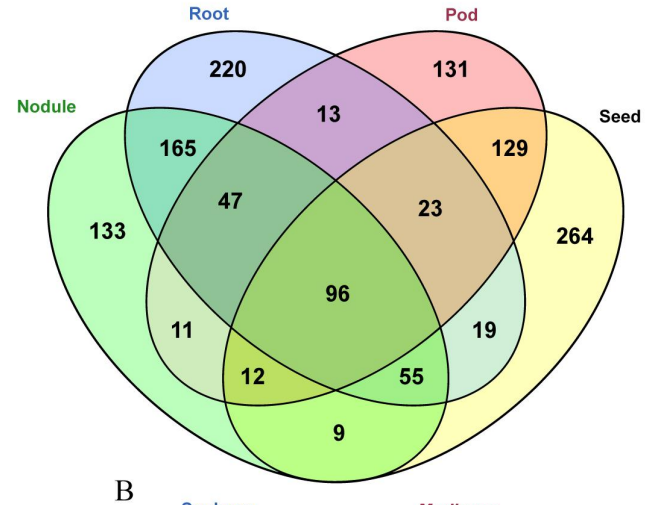
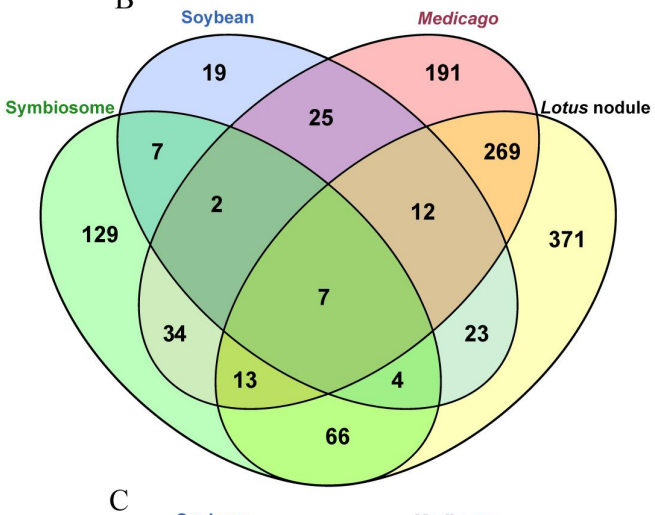


Figure 3: The intersection of proteins within the *Lotus* tissues and other nodule/symbiosome studies. A) In total, 1327 proteins were identified for the four *Lotus* tissues. B and C show the intersection of proteins identified in the following studies: symbiosome membrane ([16-18], soybean nodule [14], and *Medicago* nodule (only plant gene accessions) [15] including either proteins from *Lotus* nodule or root obtained in the analysis. See Supporting Information Table 5, for the gene accessions belonging to each intersection.

Figure 3A



B



C

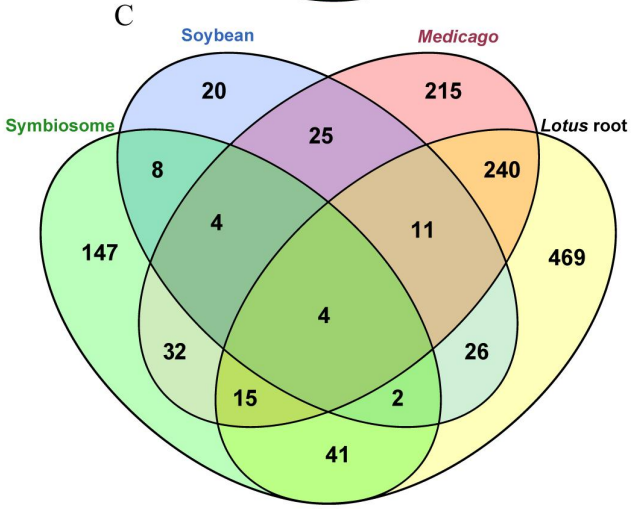


Table 1: The most regulated proteins between white and red nodules.

spot #	white>red (volume)	predicted annotation	spot #	white<red (volume)	predicted annotation
7423	95663	L-ascorbate peroxidase	7298	191804	leghemoglobin
7281	65561	PR10 protein	7304	179613	leghemoglobin
8092	61945	methionine synthase	7455	77293	14-3-3 protein
7456	57755	PR10 protein	7696	73072	glutamine synthetase
7936	52578	heat shock protein	8002	67441	glutamine synthetase
7286	51568	PR10 protein	7685	62133	glutamine synthetase
7294	45566	PR10 protein	7741	60082	phosphoglycerate kinase
8086	44149	transketolase	7693	59764	nitrilase
7938	43722	heat shock protein	7288	58135	leghemoglobin
7410	42774	lectin	7313	53640	unknown
	white>red (fold change)			**white<red (fold change)**	
7269	5,6	ribosomal protein	7504	17,2	vacuolar H+-ATPase subunit
8086	4,1	transketolase	7514	15,0	cysteine proteinase
7889	3,7	glycerol kinase	7998	13,2	asparagine synthetase
7520	3,4	caffeoyl-CoA O-methyltransferase	8098	12,1	sucrose synthase
7342	3,4	transcription factor homolog	7659	8,9	ACC oxidase
7878	3,4	helicase	7645	8,8	ACC oxidase
7309	3,2	actin depolymerizing	8005	8,8	asparagine synthetase
7410	3,2	lectin	7301	8,3	glycine-rich RNA binding protein
7358	3,1	ATP synthase subunit	8002	7,3	glutamine synthetase
7319	3,0	enolase	7997	7,3	glutamine synthetase

To the left, proteins with highest level (volume or fold change) in white nodules together with predicted annotations. To the right, proteins with highest level (volume or fold change) in red nodules together with predicted annotations. The *M. loti* proteins were excluded from this analysis.

Table 2: *M. loti* proteins identified from the nodule subset.

spot #	gene accession	annotation
7338	gi|13475117 (mll6127)	outer membrane protein
7345	gi|13475117 (mll6127)	outer membrane protein
7347	gi|13476553 (mll7908)	outer membrane protein
7348	gi|13476553 (mll7908)	outer membrane protein
7351	gi|13475117 (mll6127)	outer membrane protein
7354	gi|13474891 (mll5873)	peroxiredoxin 2 family protein
7496	gi|13473202 (mll3723)	periplasmic phosphate-binding protein
7532	gi|13474949 (mlr5944)	hypothetical protein
7534	gi|13474949 (mlr5944)	hypothetical protein
7553	gi|13472749 (mlr3148)	membrane lipoprotein
7596	gi|13474817 (mlr5787)	hypothetical protein
7601	gi|13475013 (mll5998)	hypothetical protein
7623	gi|13473431 (mll4029)	porin Omp2b
7634	gi|13470575 (mlr0325)	DNA-directed RNA polymerase subunit alpha
7698	gi|13471636 (mlr1666)	sulfate binding protein of ABC transporter
7707	gi|13475061 (mll6062)	hypothetical protein
7739	gi|13474948 (mlr5943)	diaminobutyrate--2-oxoglutarate aminotransferase
7755	gi|13474948 (mlr5943)	diaminobutyrate--2-oxoglutarate aminotransferase
7762	gi|13488218 (mlr9117)	malate dehydrogenase-like protein
7780	gi|13474939 (mlr5932)	1-aminocyclopropane-1-carboxylate deaminase
7832	gi|13474921 (mlr5907)	nitrogenase molybdenum-iron protein beta chain, nifK
7844	gi|13475266 (mlr6296)	carboxylase
7945	gi|13474595 (mlr5513)	ABC transporter, periplasmic oligopeptide-binding protein
7956	gi|13474344 (mll5210)	30S ribosomal protein S1
7959	gi|13474921 (mlr5907)	nitrogenase molybdenum-iron protein beta chain, nifK
7960	gi|13474837 (mll5810)	chaperonin GroEL
7963	gi|13474837 (mll5810)	chaperonin GroEL
7966	gi|13474837 (mll5810)	chaperonin GroEL
8004	gi|13474920 (mlr5906)	nitrogenase molybdenum-iron protein alpha chain, nifD
8056	gi|13474634 (mlr5562)	polynucleotide phosphorylase/polyadenylase
8057	gi|13470549 (mlr0286)	elongation factor G

Spot number, gene accession, and annotation for the 31 protein spots identified as *M. loti* proteins are shown.