1 **Title**

2 Quantitative time-course proteome analysis of *Mesorhizobium loti* during nodule maturation

3

4 **Running title**

5 Time-course proteome analysis of *Mesorhizobium loti*

6

7 **Authors and affiliations**

8 Mami Nambu^{a, 1}, Yohei Tatsukami^{a, b, 1}, Hironobu Morisaka^{a, c}, Kouichi Kuroda^a, Mitsuyoshi Ueda^{a, c, *}

9 ^aDivision of Applied Life Sciences, Graduate School of Agriculture, Kyoto University, Sakyo-ku, Kyoto

10 606-8502, Japan

11 ^bJapan Society for the Promotion of Science, Sakyo-ku, Kyoto 606-8502, Japan

12 ^cKyoto Industrial Science and Technology Innovation Center, Shimogyo-ku, Kyoto 600-8813, Japan

13 *Tel: +81-(0)75-753-6110; Fax: +81-(0)75-753-6112; E-mail: miueda@kais.kyoto-u.ac.jp

14 ¹These authors contributed equally to this work.

Abstract

Rhizobia are nitrogen-fixing bacteria that establish a symbiotic relationship with leguminous plants. To understand the mechanism by which rhizobia alter their metabolism to establish successful nitrogen-fixing symbiotic relationship with hosts, *Lotus japonicus* were inoculated with *Mesorhizobium loti*. Bacteroids were isolated from nodules harvested at 2 weeks (the early stage of nodule development), and at 3 and 4 weeks (the intermediate stage of nodule development) post-inoculation. Using a quantitative time-course proteome analysis, we quantified the variations in the expression of 537 proteins in *M. loti* bacteroids during the course of nodule maturation. The results revealed significant changes in the carbon and amino acid metabolisms by *M. loti* upon differentiating into bacteroids. Furthermore, our findings suggested that *M. loti* enters a nitrogen-deficient condition during the early stages of nodule development, and then a nitrogen-rich condition during the intermediate stages of nodule development. In addition, our data indicated that *M. loti* assimilated ammonia during the intermediate stages of nodule development. Our results provide new insights into course of physiological transitions undergone by *M. loti* during nodule maturation.

Key words

Mesorhizobium loti, *Lotus japonicus*, Symbiosis, Quantitative proteome analysis, Plant-microbe interaction

32 **Introduction**

33 Rhizobia are soil bacteria that establish a symbiotic relationship with leguminous plants. Upon
34 interaction with a specific leguminous host, these bacteria induce the formation of symbiotic host organ
35 referred to as the root nodule. After colonization of the root nodule, rhizobia differentiate into a
36 nitrogen-fixing form, known as a bacteroid. These bacteroids then provide the host plant with fixed
37 nitrogen in exchange for carbon and other essential nutrients. Nitrogen fixation by rhizobia requires
38 elaborate integration of plant and bacterial metabolism [1]. For effective nitrogen fixation, rhizobia
39 construct an optimal intracellular environment during nodule maturation; however, the specific
40 mechanism by which this process occurs is unclear. As such, analyses that are more comprehensive are
41 necessary to understand the sophisticated metabolism in bacteroids.

42 Omics analyses are important to characterize dynamic variations in complex biological systems. In
43 particular, proteome analysis is an ideal tool for the characterization of intracellular conditions, as
44 proteins play critical roles in nearly all cellular events [2]. Conventional studies involve the use of
45 2-dimensional polyacrylamide gel electrophoresis (2D-PAGE)-based analyses combined with
46 matrix-assisted laser desorption and ionization time-of-flight mass spectrometry. However,
47 characterizing the bacteroid differentiation process by 2D-PAGE-based analysis would be impractical
48 due to the low identification rates and low-throughput associated with this method. To address these

limitations, liquid chromatography-tandem mass spectrometry (LC-MS/MS)-based technology has been developed as a shotgun approach to identify large numbers of proteins in a high-throughput manner with high resolution [3]. In this study, we used a nano LC-MS/MS system equipped with a long monolithic silica capillary column, which previously exhibited high separation efficiency and low ionization suppression [4]. Indeed, a one shot proteome analysis of *Escherichia coli* using a monolithic column successfully identified 2602 proteins [5]. Furthermore, in previous studies, we performed quantitative proteome analyses of *Candida albicans* and *Clostridium cellulovorans* using a monolithic column and the isobaric tag method [6, 7]. In these studies, 1024 and 639 proteins were identified and quantified, respectively.

Mesorhizobium loti is a model rhizobium that induces formation of root nodules in *Lotus japonicus*. The genome sequences of *M. loti* MAFF303099 and *L. japonicus* MG-20 have been reported [8, 9], and the genome database is maintained by the Kazusa DNA Research Institute (Rhizobase; <http://genome.microbedb.jp/rhizobase>). We previously reported a qualitative proteome analysis of *Mesorhizobium loti* under both free-living and symbiotic conditions [10]. These analyses revealed distinct protein profiles of *M. loti* and the expression of distinct cell surface structures under each condition. Previous studies utilized the 2-D PAGE or LC-MS/MS approaches to analyze the proteome of the rhizobia, such as *Bradyrhizobium japonicum* and *Sinorhizobium meliloti* [11-19]. Although these

studies described the protein profiles of these organisms after bacteroid differentiation, the specific physiological changes that occur during the transition to the bacteroid state have yet to be fully elucidated.

In this study, we aimed to enhance our understanding of metabolic changes *M. loti* undergoes during the establishment of nitrogen-fixing symbiosis with the host plant *L. japonicus*. To this end, we performed a quantitative time-course proteome analysis of *M. loti* during nodule development to characterize the bacteroid differentiation process. Statistical analyses of metabolic pathways suggested that *M. loti* experienced a nitrogen-deficient condition during the early stage of nodule development, and then assimilated ammonia during the intermediate stage of nodule development. These results offer attractive insights into the mechanism of *L. japonicus* – *M. loti* symbiosis.

Material and methods

Strains and growth conditions. *M. loti* MAFF303099 was cultured in tryptone-yeast extract (TY) medium [20] at 28°C. Cultures were grown to mid log phase (OD₆₀₀ of approximately 0.9–1.1) for 36 h. Cells from this free-living condition were then collected and sample preparation was performed [10].

To establish symbiotic conditions, *L. japonicus* MG-20 Miyakojima [21] seeds were sterilized, germinated, grown in MM1 (minimal medium 1) [22] medium at 25°C with a 16 h light/8 h dark cycle,

and inoculated with *M. loti*. Root nodules from several plants were harvested at 2, 3, and 4 weeks post-inoculation (WPI), frozen with liquid nitrogen, and stored at -80°C. Nodules from 3 independently grown pools of plants were collected and processed in parallel.

Bacteroids were isolated and purified from harvested nodules, as previously reported [23]. In brief, nodules were homogenized using a mortar and pestle in 5 mL of 100 mM MOPS-KOH buffer (pH 7.2) containing 375 mM mannitol, 5 mM MgSO₄, 10 mM EGTA, 3 mM 4-aminobenzamide, 10 mM dithiothreitol, 1% (w/v) polyvinylpyrrolidone-40, 4% (w/v) Dextran T40, and 1% bovine serum albumin. Homogenates were filtered through two layers of Miracloth (Millipore, Billerica, MA, USA), and centrifuged at 4 °C, 100 g for 10 min using a Tomy MX-305 refrigerated centrifuge (Tomy Seiko, Tokyo, Japan). The supernatants were harvested and then centrifuged at 4 °C, 700 g for 15 min. Pellets were suspended in 200 µL of PBS buffer solution (pH 7.4) (Wakenyaku, Kyoto, Japan), placed on a stepwise gradient of 45%, 60%, and 80% (v/v) Percoll (GE Healthcare, Waukesha, WI, USA), and centrifuged at 4 °C, 4,200 g for 40 min. The pellets were collected from the interface of 60% and 80% (v/v) Percoll and washed 5 times in 1 mL of PBS buffer solution (pH 7.4). Isolated bacteroids were used for sample preparation.

Sample preparation. Proteins were extracted and digested, as previously reported [10]. The tryptic digests were labeled using the tandem mass tag (TMT) 6-plex labeling kit (Thermo Fisher

Scientific), according to the manufacturer's protocol. The TMT labeling reagents were dissolved in 41 μL of acetonitrile and mixed with 30 μL of each tryptic digest, as illustrated in Fig. 1. In brief, the tryptic digests of samples from free-living *M. loti* and bacteroids at 2, 3, and 4 WPI were mixed with TMT 126, 127, 128, and 129, respectively. In addition, a mixture of tryptic digests of samples from free-living and symbiotic organisms was treated with TMT-130 as an internal control for quantification. The reactants were quenched by the addition of 8 μL of 5% hydroxylamine and then lyophilized after mixing all reactants. The dried samples were reconstituted in 30 μL of 0.1% formic acid.

LC-MS/MS analysis. Proteome analyses were performed using a liquid chromatography (LC) (UltiMate3000, Thermo Fisher Scientific)/mass spectrometry (LTQ Orbitrap Velos mass spectrometer, Thermo Fisher Scientific) system equipped with a long monolithic silica capillary column (500 cm long, 0.1 mm ID) [7]. Tryptic digests were injected and separated by reversed-phase chromatography at a flow rate of 500 $\text{nL} \cdot \text{min}^{-1}$. The gradient was generated by changing the mixing ratio of the 2 eluents: A, 0.1% (v/v) formic acid and B, 80% (v/v) acetonitrile containing 0.1% (v/v) formic acid. The gradient was started with 5% B, increased to 45% B for 600 min, further increased to 95% B to wash the column, then returned to the initial condition, and held for re-equilibration. The separated analytes were detected on a mass spectrometer with a full scan range of 350–1,500 m/z . For data-dependent acquisition, the

apparatus was set to automatically analyze the top 10 most intense ions observed in the MS scan. A normalized collision energy of 40% in higher-energy C-trap dissociation with a 0.1 ms activation time was used. An electrospray ionization voltage of 2.4 kV was applied directly to the buffer end of the LC column using a MicroTee (Upchurch Scientific, Oak Harbor, WA, USA). The ion transfer tube temperature was set to 280°C. Triplicate analyses were performed for each sample of three biological replicates, and blank runs were inserted between different samples.

Data analyses. The mass spectrometry data from each sample were used for protein identification and quantification. It was performed using Mascot (Matrix Science, London, UK) working on Proteome Discoverer 1.4. (Thermo Fisher Scientific) against the Rhizobase database, containing 7,283 sequences, with a precursor mass tolerance of 50 ppm, a fragment ion mass tolerance of 50 mmu, and strict specificity allowing for up to 1 missed cleavage. For trypsin digestion, cysteine carbamidomethylation (+57.021 Da), TMT 6-plex of N-term (+229.1629 Da), and TMT 6-plex of lysine (+229.1629 Da) were set as variable modifications. The data was then filtered at a q-value ≤ 0.01 , corresponding to a 1% false discovery rate (FDR) on the spectral level. Protein quantification was performed using the Reported Ions Quantifier with the TMT 6-plex method on Proteome Discoverer software. Three independent biological experiments for each sample were performed, and proteins identified in every replicate with a number of

used peptides per protein ≥ 3 were accepted. Proteins that were identified in every replicate were accepted. Global median normalization was carried out to normalize the amount of tryptic digest injected into the LC-MS/MS. The adjusted relative abundances of proteins identified in this study are listed in Table S1. Our experimental information was summarized in Table S2 following to the Minimum Information About a Proteomics Experiment (MIAPE) guideline (<http://www.psdev.info/node/91>).

Non-hierarchical *K*-means clustering, using the Hartigan-Wong algorithm, was performed to identify clustered proteome changes. Principal component analysis (PCA) was used to investigate similarities in the production profiles between the three biological replicates. Both analyses were performed using R.

Functional classification of proteins was performed using the Rhizobase at Kazusa DNA Research Institute (<http://www.kazusa.or.jp/e/index.html>). The pathway analysis of identified proteins was performed using the KEGG pathway-mapping tool (<http://www.genome.jp/kegg/>).

To identify proteins that were differentially expressed at distinct sampling points, an empirical Bayes moderated *t*-test was performed. To address the issue of multiple comparisons and to control the family-wise error rate, the threshold was set at $P < 0.05$ after adjustment of FDRs, using the Benjamini-Hochberg method.

Results

Protein identification and quantification. To quantify variations in the proteome of *M. loti*

bacteroids during nodule maturation, colonized *L. japonicus* nodules were harvested at 2, 3, and 4 WPI.

While nodules harvested at 2 WPI, corresponding to the early stage of nodule development, were

approximately 1 mm in diameter and were white in color, nodules harvested at 3 and 4 WPI,

corresponding to the intermediate stage of nodule development, had grown to approximately 2–3 mm in

diameter and were red in color. Bacteroid samples were isolated from harvested nodules, and proteins

were extracted, tryptic digested, and TMT-labeled. Mixtures of each sample (free-living, 2, 3, and 4 WPI,

and control) (Fig. 1) were then subjected to LC-MS/MS analysis, and 537 proteins were successfully

identified and quantified (Table S1). This number of protein identification and quantification is

comparable to other reports focused on legume-rhizobium symbiosis [11-19]. Correlation analysis and

principal component analysis (PCA) of the three biological replicates of each sample group were used to

evaluate experimental reproducibility. Pearson's correlation analysis (represented as "R") confirmed the

reproducibility (Fig. S1), and plots of PCA scores indicated high levels of similarity between the

biological replicates (Fig. 2a). These results indicate that we successfully quantified proteome variance

with high precision and accuracy.

Clustering analysis.

The proteins quantified by LC-MS/MS were classified into five clusters, according to changes in relative abundance to free-living *M. loti*, by non-hierarchical *K*-means clustering (Fig. 2c). Relative abundances were calculated by dividing the average abundance of the three biological replicates at each sampling point by the average abundance in free-living *M. loti*. Three protein clusters were up-regulated at 4 WPI during nodule maturation: Group 1 (12-fold), Group 2 (7-fold), and Group 4 (2-fold). Conversely, the Group 5 cluster was down-regulated at 4 WPI (0.8-fold). Notably, Group 3, which contained only one protein, mll4247, was up-regulated at 2 WPI (16-fold) and then down-regulated at 3 and 4 WPI (5-fold and 4-fold, respectively). Group 1, 2, 4, and 5 included 12, 35, 104, and 385 proteins, respectively.

To assess the regulation of metabolic processes during bacteroid differentiation, we classified the proteins in Group 1, 2, 4 and 5 according to their physiological functions, as listed in the rhizobase database (Fig. 2d). This classification revealed that up-regulated proteins (Group 1, 2, and 4) were generally involved in four functions: central intermediary metabolism, energy metabolism, cellular processes, and transport and binding proteins. The majority of the proteins in Group 1, including the *nif* and *fix* gene products, which are required for nitrogen fixation, were involved in central intermediary metabolism. The expression levels of these proteins markedly increased over the course of nodule maturation (Fig. 2c). These findings are consistent with previous reports that nitrogen fixation by

rhizobia occurs only during symbiosis [24]. The energy metabolism category contained various enzymes involved in central metabolic pathways, including glycolysis and the tricarboxylic acid (TCA) cycle. These findings likely reflect the fact that these processes are crucial for supplying the large amount of ATP and reducing power required for nitrogen fixation [25, 26]. Meanwhile, proteins included in the cellular processes category included enzymes involved in bacterial antioxidant defense. The increase in ATP levels needed for nitrogen fixation by bacteroids requires a concurrent increase in respiration rates, which results in the production of reactive oxygen species (ROS) within the nodules [27, 28]. Antioxidant defense enzymes would be required to scavenge ROS. Therefore, the observed up-regulation of proteins involved in antioxidant defense is likely a response to prevent oxidative damage during this life stage. Lastly, the transport and binding proteins category included proteins related to iron, sulfur, and phosphorus transport. Iron and sulfur are essential for the proteins involved in nitrogen fixation. Ferredoxins, which transfer electrons from the TCA cycle to the nitrogenase complex, contain [2Fe-2S] clusters, while the nitrogenase complex composed of NifHDK (mlr5905, mlr5906, and mlr5907, respectively) contains [4Fe-4S] clusters, [8Fe-7S] clusters (P-clusters), and [Mo-7Fe-6S] clusters (FeMo cofactors) [1]. Phosphorus is also required for the production of ATP as an energy source during mutualism. Therefore, these elements likely play an important role in nitrogen fixation.

Meanwhile, Group 5, which was down-regulated in bacteroids, included multiple proteins

involved in transcription and translation. These results indicate that proteins related to gene transcription and mRNA translation, such as RNA polymerases and ribosomal proteins, were repressed during symbiotic growth. The repression of these proteins reflects the reduced multiplication rates of bacteroids during the intermediate stage of nodule development [29, 30].

Central metabolism. There was a marked change in the central metabolic profile of *M. loti* during the differentiation into bacteroids (Fig. 3). Bacteroids receive carbon and energy sources from hosts in the form of C₄-dicarboxylates, particularly malate and succinate [31]. Because C₄-dicarboxylates are directly incorporated into the TCA cycle, we expected that the expression of proteins involved in this cycle would fluctuate during the course of nodule maturation. Indeed, our proteome analysis indicated that the majority of enzymes involved in the TCA cycle were up-regulated during this period (Fig. 3). The lone exception was α -ketoglutarate dehydrogenase (mll4301), which was down-regulated at 2 and 3 WPI, but restored to normal levels at 4 WPI (Fig. 3 box (a)). α -ketoglutarate dehydrogenase catalyzes the conversion of α -ketoglutarate to S-succinyl-dihydrolipoamide-E. Furthermore, α -ketoglutarate is converted to glutamate by glutamate synthase (GOGAT; mll1646) and glutamate dehydrogenase (GDH; mll4104), which were up-regulated at 2 and 4 WPI, respectively (Fig. 4b). Our data strongly indicate that the majority of α -ketoglutarate was converted to glutamate by GOGAT or GDH, and then incorporated

into amino acid metabolism during symbiotic growth (Fig. 4b). We therefore suggest that biosynthesis of α -ketoglutarate dehydrogenase was down-regulated due to this decrease in α -ketoglutarate.

Phosphoenolpyruvate carboxykinase (PEPCK; mlr5096), which catalyzes the first step in gluconeogenesis, was down-regulated during symbiosis (Fig. 3 box (b)). Reduced PEPCK activity was previously observed in the bacteroids of *Rhizobium leguminosarum*, *S. meliloti*, and *Rhizobium* sp. NGR234 [32-34]. However, gluconeogenesis appears to be important for symbiotic nitrogen fixation in many rhizobia, including *S. meliloti*, *Rhizobium* sp. NGR234, and *Rhizobium etli*, based on mutation in PEPCK leading to reduced acetylene reduction [33, 35, 36]. These results suggest that gluconeogenesis was less activated in bacteroids than in free-living *M. loti*, although it was essential for symbiotic nitrogen fixation.

The expression level of branched-chain amino acid transferase (mll2664), which catalyzes the synthesis of the branched-chain amino acids valine, leucine, and isoleucine, decreased during nodule maturation (Fig. 3 box (c)). It has been reported that transporters for branched-chain amino acids were required for effective nitrogen fixation in *R. leguminosarum* [37, 38]. These findings therefore suggest that rhizobia must obtain the branched-chain amino acids from the host.

A previous report demonstrated that bacteroids provide ammonium as well as the amino acids alanine and/or aspartate for the host [39]. The expression levels of two aspartate aminotransferases,

which produce aspartate from oxaloacetate, were quantified (Fig. 3 box (d)). One of these aminotransferases, mlr5883, is encoded within the *M. loti* symbiosis island, and was up-regulated at 3 and 4 WPI (the intermediate stage of nodule development) [8]. The other, mlr2541, is encoded in a region outside the symbiosis island and was constitutively expressed. It was reported that expression of the symbiosis island is highly up-regulated during symbiotic growth [24]. While bacteroids likely used both of the aspartate aminotransferases, aspartate aminotransferase mlr5883 was specifically produced at the intermediate stage of nodule development. Conversely, the biosynthesis of alanine dehydrogenase (mll9089), which produces alanine from pyruvate, decreased at the early stage of nodule development. However, expression of this protein was restored to levels similar to that observed in free-living *M. loti* at the intermediate stage of nodule development (Fig. 3 box (e)). The fold change (symbiosis/free-living) of aspartate aminotransferase (mlr5883) was higher than that of alanine dehydrogenase (mll9089) (Fig. 3 boxes (d, e)). As a result, bacteroids likely biosynthesized excessive amounts of aspartate during symbiosis, and may therefore secrete larger amounts of aspartate than alanine in the host. This idea is supported by previous reports that aspartate aminotransferase mutants of *S. meliloti* did not fix nitrogen in nodules [37, 40].

Nitrogen-deficiency at the early stage of nodule development. The loading scatter plot displays

the principal components of the loadings, which are the correlation between the original variables and the factors, and therefore the key to understand the underlying nature of a particular factor (Fig. 2b).

Analysis of this figure indicated that the nitrogen regulatory protein P-II; GlnK (mll4247) had the largest effect on bacteroid development at 2 WPI (Fig. 4c). GlnK is uridylylated under nitrogen-deficiency

resulting in the activation of nitrogen stress response [41-44]. Previous studies in *E. coli* demonstrated that GlnK expression was enhanced under nitrogen-deficient conditions [44, 45]. In *Corynebacterium*

glutamicum, GOGAT expression was enhanced under nitrogen-deficient conditions. Conversely, GOGAT and GDH were down- and up-regulated under nitrogen-rich conditions, respectively [46].

Interestingly, our data demonstrated that GlnK and GOGAT were up-regulated in bacteroids at 2 WPI (the early stage of nodule development). Then GlnK expression returned to the levels observed in free-living *M. loti* at 3 and 4 WPI (the intermediate stage of nodule development) (Fig. 4b, c).

Meanwhile, the expression levels of GOGAT and GDH were low and high in bacteroids at 4 WPI compared to the levels present in free-living *M. loti*, respectively (Fig. 4b). Accordingly, our results

strongly indicate that *M. loti* may enter a nitrogen-deficient condition at the early stage of nodule development, and then change to a nitrogen-rich condition at the intermediate stage of nodule development.

Discussion

This study aimed to characterize variations in the physiological conditions of bacteroids during nodule maturation. To the best of our knowledge, this is the first report describing a time-course of proteome variations in *M. loti* bacteroids. We quantitatively identified 537 proteins, which comprise only 32.3% of a total of identified proteins in a qualitative proteome analysis of *M. loti* [10]. However, our data provide novel insights into molecular mechanisms of *L. japonicus* – *M. loti* symbiosis because we described the intracellular variations of *M. loti* bacteroids during nodule maturation, which have never been revealed.

Our quantitative proteome analysis demonstrated that GlnK and GOGAT, which have been reported to be expressed at high levels under nitrogen-deficiency [44-46], were up-regulated in bacteroids at 2 WPI (Fig 4b, c). These results suggest that *M. loti* undergoes nitrogen-deficiency during the early stage of nodule development. While nodules harvested at 3 and 4 WPI were red, the nodules harvested at 2 WPI were white. This white phenotype is a strong indicator that at 2 WPI, nodules contained significantly less leghemoglobin than at 3 and 4 WPI. In the nitrogen-fixing nodules, leghemoglobin is induced by host plants to avoid inactivation of oxygen-labile nitrogenase and to maintain a sufficient oxygen flux for respiration [47, 48]. In addition, nitrogenase complex core subunits (NifHDK) were slightly up-regulated (less than 3-fold) at 2 WPI, but highly up-regulated (more than

7-fold) at 3 and 4 WPI (Fig. 4a). These results indicated that bacteroids were associated with lower levels of nitrogenase activity during the early stage of nodule development than during the intermediate stage of nodule development. Therefore, nitrogen-deficiency would likely occur in the nodule at the early stage of development. Meanwhile, nitrogen fixation resulted in nitrogen-rich conditions in bacteroids at the intermediate stage of nodule development. Furthermore, GlnK and GOGAT levels were similar to those in free-living *M. loti* at this stage of nodule development (Fig. 4 (b, c)).

Although it was generally accepted that bacteroids shut down ammonium assimilation during the symbiotic growth stage [31, 41], we observed high expression levels of enzymes involved in ammonium assimilation during the intermediate stage of nodule development, indicating that bacteroids assimilated ammonia during this period. Ammonium assimilation occurs through the GS/GOGAT or GDH/GS pathways [41]. GS plays an important role in ammonium assimilation by catalyzing the condensation of glutamate and ammonia to yield glutamine. GSI (mll0343) deficiency in *M. loti* has been reported to reduce nitrogen-fixing activity per nodule and induce early nodule senescence [49]. Our data indicated up-regulations of GSI (mll0343) and GDH (mll4104) production in bacteroids at 4 WPI (Fig. 4b). These results suggest that *M. loti* assimilates ammonia through the GDH/GS pathway during the intermediate stage of nodule development.

In conclusion, quantitative time-course proteome analysis of *M. loti* bacteroids revealed that the

protein expression profiles of *M. loti* varied greatly over the course of bacteroid differentiation, and provides a comprehensive characterization of the intracellular conditions of bacteroids (Fig. 5). Based on these data, we propose three characteristics regarding the cellular conditions of *M. loti* during mutualism. First, *M. loti* experiences nitrogen-deficiency at the early stage of nodule development resulting from low levels of nitrogenase expression. Second, high levels of nitrogenase expression induce nitrogen-rich conditions in *M. loti* at the intermediate stage of nodule development. Third, *M. loti* assimilates ammonia produced by nitrogen fixation at the intermediate stage of nodule development. These characteristics were hard to find by previous individual researches focused on several genes or proteins. We find some important phenomena by using proteomics approach, and next we have to proof these phenomena at gene or protein level analysis.

Our results provide novel insights into the mutualistic relationship of *L. japonicus* and *M. loti*, and helpful in carrying out detailed studies on the phenotypic variations of bacteroids during nodule maturation.

The mass spectrometry proteomics data have been deposited to the ProteomeXchange Consortium [50] via the PRIDE partner repository with the dataset identifier PXD001964 and 10.6019/PXD001964.

Acknowledgements

This work was supported by the 12th Innovations Inspired by Nature Research Support of Sekisui Chemical Co., Ltd., by a Grant-in-Aid for Scientific Research (B) (No. 25289293), JST-CREST, and by the Japan Society for the Promotion of Science (grant 26-1156).

We thank the National BioResource Project (Legume Base), Japan, for kindly providing *Lotus japonicus* seeds. We also thank the National Institute of Technology and Evaluation, Japan, for kindly providing *Mesorhizobium loti*.

References

- [1] Seefeldt LC, Hoffman BM, Dean DR. Mechanism of Mo-Dependent Nitrogenase. *Annu Rev Biochem.* 2009;78:701-22.
- [2] Tyers M, Mann M. From genomics to proteomics. *Nature.* 2003;422:193-7.
- [3] Yates JR, Ruse CI, Nakorchevsky A. Proteomics by mass spectrometry: approaches, advances, and applications. *Annu Rev Biomed Eng.* 2009;11:49-79.
- [4] Motokawa M, Kobayashi H, Ishizuka N, Minakuchi H, Nakanishi K, Jinnai H, *et al.* Monolithic silica columns with various skeleton sizes and through-pore sizes for capillary liquid chromatography. *J Chromatogr A.* 2002;961:53-63.

337 [5] Iwasaki M, Miwa S, Ikegami T, Tomita M, Tanaka N, Ishihama Y. One-dimensional capillary liquid
 338 chromatographic separation coupled with tandem mass spectrometry unveils the *Escherichia coli*
 339 proteome on a microarray scale. *Anal Chem.* 2010;82:2616-20.

340 [6] Aoki W, Tatsukami Y, Kitahara N, Matsui K, Morisaka H, Kuroda K, *et al.* Elucidation of potentially
 341 virulent factors of *Candida albicans* during serum adaptation by using quantitative time-course
 342 proteomics. *J Proteomics.* 2013;91:417-29.

343 [7] Matsui K, Bae J, Esaka K, Morisaka H, Kuroda K, Ueda M. Exoproteome profiles of *Clostridium*
 344 *cellulovorans* grown on various carbon sources. *Appl Environ Microbiol.* 2013;79:6576-84.

345 [8] Kaneko T, Nakamura Y, Sato S, Asamizu E, Kato T, Sasamoto S, *et al.* Complete genome structure of
 346 the nitrogen-fixing symbiotic bacterium *Mesorhizobium loti*. *DNA Res.* 2000;7:331-8.

347 [9] Sato S, Nakamura Y, Kaneko T, Asamizu E, Kato T, Nakao M, *et al.* Genome structure of the legume,
 348 *Lotus japonicus*. *DNA Res.* 2008;15:227-39.

349 [10] Tatsukami Y, Nambu M, Morisaka H, Kuroda K, Ueda M. Disclosure of the differences of
 350 *Mesorhizobium loti* under the free-living and symbiotic conditions by comparative proteome analysis
 351 without bacteroid isolation. *BMC Microbiol.* 2013;13:180-8.

352 [11] Barra-Bily L, Fontenelle C, Jan G, Flechard M, Trautwetter A, Pandey SP, *et al.* Proteomic
 353 alterations explain phenotypic changes in *Sinorhizobium meliloti* lacking the RNA chaperone Hfq. *J*

354 Bacteriol. 2010;192:1719-29.

355 [12] Chen H, Teplitski M, Robinson JB, Rolfe BG, Bauer WD. Proteomic analysis of wild-type
356 *Sinorhizobium meliloti* responses to N-acyl homoserine lactone quorum-sensing signals and the
357 transition to stationary phase. J Bacteriol. 2003;185:5029-36.

358 [13] Delmotte N, Ahrens CH, Knief C, Qeli E, Koch M, Fischer HM, *et al.* An integrated proteomics and
359 transcriptomics reference data set provides new insights into the *Bradyrhizobium japonicum* bacteroid
360 metabolism in soybean root nodules. Proteomics. 2010;10:1391-400.

361 [14] Djordjevic MA. *Sinorhizobium meliloti* metabolism in the root nodule: a proteomic perspective.
362 Proteomics. 2004;4:1859-72.

363 [15] Gao M, Chen H, Eberhard A, Gronquist MR, Robinson JB, Connolly M, *et al.* Effects of
364 AiiA-mediated quorum quenching in *Sinorhizobium meliloti* on quorum-sensing signals, proteome
365 patterns, and symbiotic interactions. Mol Plant Microbe Interact. 2007;20:843-56.

366 [16] Hempel J, Zehner S, Goettfert M, Patschkowski T. Analysis of the secretome of the soybean
367 symbiont *Bradyrhizobium japonicum*. J Biotechnol. 2009;140:51-8.

368 [17] Larrainzar E, Wienkoop S, Weckwerth W, Ladrera R, Arrese-Igor C, Gonzalez EM. *Medicago*
369 *truncatula* root nodule proteome analysis reveals differential plant and bacteroid responses to drought
370 stress. Plant Physiol. 2007;144:1495-507.

371 [18] Sarma AD, Emerich DW. Global protein expression pattern of *Bradyrhizobium japonicum*
372 bacteroids: a prelude to functional proteomics. *Proteomics*. 2005;5:4170-84.

373 [19] Sarma AD, Emerich DW. A comparative proteomic evaluation of culture grown vs nodule isolated
374 *Bradyrhizobium japonicum*. *Proteomics*. 2006;6:3008-28.

375 [20] Beringer JE. R factor transfer in *Rhizobium leguminosarum*. *J Gen Microbiol*. 1974;84:188-98.

376 [21] Kawaguchi M. *Lotus japonicus* 'Miyakojima' MG-20: An early-flowering accession suitable for
377 indoor handling. *J Plant Res*. 2000;113:507-9.

378 [22] Becard G, Fortin JA. Early Events of Vesicular Arbuscular Mycorrhiza Formation on Ri T-DNA
379 Transformed Roots. *New Phytologist*. 1988;108:211-8.

380 [23] Kumagai H, Hakoyama T, Umehara Y, Sato S, Kaneko T, Tabata S, *et al*. A novel ankyrin-repeat
381 membrane protein, IGN1, is required for persistence of nitrogen-fixing symbiosis in root nodules of
382 *Lotus japonicus*. *Plant Physiol*. 2007;143:1293-305.

383 [24] Uchiumi T, Ohwada T, Itakura M, Mitsui H, Nukui N, Dawadi P, *et al*. Expression islands clustered
384 on the symbiosis island of the *Mesorhizobium loti* genome. *J Bacteriol*. 2004;186:2439-48.

385 [25] Ronson CW, Lyttleton P, Robertson JG. C₄-dicarboxylate transport mutants of *Rhizobium trifolii*
386 form ineffective nodules on *Trifolium repens*. *Proc Natl Acad Sci U S A*. 1981;78:4284-8.

387 [26] Finan TM, Wood JM, Jordan DC. Symbiotic properties of C₄-dicarboxylic acid transport mutants of

388 *Rhizobium leguminosarum*. J Bacteriol. 1983;154:1403-13.

389 [27] Hanyu M, Fujimoto H, Tejima K, Saeki K. Functional differences of two distinct catalases in

390 *Mesorhizobium loti* MAFF303099 under free-living and symbiotic conditions. J Bacteriol.

391 2009;191:1463-71.

392 [28] Becana M, Dalton DA, Moran JF, Iturbe-Ormaetxe I, Matamoros MA, Rubio MC. Reactive oxygen

393 species and antioxidants in legume nodules. Physiol Plantarum. 2000;109:372-81.

394 [29] Ludwig RA. Rhizobium free-living nitrogen fixation occurs in specialized nongrowing cells. Proc

395 Natl Acad Sci U S A. 1984;81:1566-9.

396 [30] Verma D. Signals in Root Nodule Organogenesis and Endocytosis of *Rhizobium*. Plant Cell.

397 1992;4:373-82.

398 [31] Prell J, Poole P. Metabolic changes of rhizobia in legume nodules. Trends Microbiol.

399 2006;14:161-8.

400 [32] McKay IA, Dilworth MJ, Glenn AR. Carbon Catabolism in Continuous Cultures and Bacteroids of

401 *Rhizobium leguminosarum* Mnf-3841. Arch Microbiol. 1989;152:606-10.

402 [33] Finan TM, Mcwhinnie E, Driscoll B, Watson RJ. Complex Symbiotic Phenotypes Result from

403 Gluconeogenic Mutations in *Rhizobium meliloti*. Mol Plant Microbe Interact. 1991;4:386-92.

404 [34] Saroso S, Dilworth MJ, Glenn AR. The Use of Activities of Carbon Catabolic Enzymes as a Probe

405 for the Carbon Nutrition of Snakebean Nodule Bacteroids. J Gen Microbiol. 1986;132:243-9.

406 [35] Tate R, Ferraioli S, Filosa S, Cermola M, Riccio A, Iaccarino M, *et al.* Glutamine utilization by

407 *Rhizobium etli*. Mol Plant Microbe Interact. 2004;17:720-8.

408 [36] Osteras M, Finan TM, Stanley J. Site-directed mutagenesis and DNA sequence of *pckA* of

409 *Rhizobium* NGR234, encoding phosphoenolpyruvate carboxykinase: gluconeogenesis and

410 host-dependent symbiotic phenotype. Mol Gen Genet. 1991;230:257-69.

411 [37] Lodwig EM, Hosie AH, Bourdes A, Findlay K, Allaway D, Karunakaran R, *et al.* Amino-acid

412 cycling drives nitrogen fixation in the legume-Rhizobium symbiosis. Nature. 2003;422:722-6.

413 [38] Hosie AHF, Allaway D, Galloway CS, Dunsby HA, Poole PS. *Rhizobium leguminosarum* has a

414 second general amino acid permease with unusually broad substrate specificity and high similarity to

415 branched-chain amino acid transporters (Bra/LIV) of the ABC family. J Bacteriol. 2002;184:4071-80.

416 [39] Allaway D, Lodwig EM, Crompton LA, Wood M, Parsons R, Wheeler TR, *et al.* Identification of

417 alanine dehydrogenase and its role in mixed secretion of ammonium and alanine by pea bacteroids. Mol

418 Microbiol. 2000;36:508-15.

419 [40] Watson RJ, Rastogi VK. Cloning and nucleotide sequencing of *Rhizobium meliloti* aminotransferase

420 genes: an aspartate aminotransferase required for symbiotic nitrogen fixation is atypical. J Bacteriol.

421 1993;175:1919-28.

422 [41] Patriarca EJ, Tate R, Iaccarino M. Key role of bacterial NH_4^+ metabolism in rhizobium-plant
 423 symbiosis. Microbiol Mol Biol Rev. 2002;66:203-22.

424 [42] Coutts G, Thomas G, Blakey D, Merrick M. Membrane sequestration of the signal transduction
 425 protein GlnK by the ammonium transporter AmtB. EMBO J. 2002;21:536-45.

426 [43] Bueno R, Pahel G, Magasanik B. Role of *glnB* and *glnD* gene products in regulation of the *glnALG*
 427 operon of *Escherichia coli*. J Bacteriol. 1985;164:816-22.

428 [44] Atkinson MR, Ninfa AJ. Role of the GlnK signal transduction protein in the regulation of nitrogen
 429 assimilation in *Escherichia coli*. Mol Microbiol. 1998;29:431-47.

430 [45] van Heeswijk WC, Hoving S, Molenaar D, Stegeman B, Kahn D, Westerhoff HV. An alternative PII
 431 protein in the regulation of glutamine synthetase in *Escherichia coli*. Mol Microbiol. 1996;21:133-46.

432 [46] Beckers G, Nolden L, Burkovski A. Glutamate synthase of *Corynebacterium glutamicum* is not
 433 essential for glutamate synthesis and is regulated by the nitrogen status. Microbiology.
 434 2001;147:2961-70.

435 [47] Appleby CA. Leghemoglobin and *Rhizobium* Respiration. Annu Rev Plant Physiol Plant Mol Biol.
 436 1984;35:443-78.

437 [48] Ott T, van Dongen JT, Gunther C, Krusell L, Desbrosses G, Vigeolas H, *et al.* Symbiotic
 438 leghemoglobins are crucial for nitrogen fixation in legume root nodules but not for general plant growth

and development. Curr Biol. 2005;15:531-5.

[49] Chungopast S, Thapanapongworakul P, Matsuura H, Van Dao T, Asahi T, Tada K, *et al.* Glutamine synthetase I-deficiency in *Mesorhizobium loti* differentially affects nodule development and activity in *Lotus japonicus*. J Plant Physiol. 2014;171:104-8.

[50] Vizcaino JA, Deutsch EW, Wang R, Csordas A, Reisinger F, Rios D, *et al.* ProteomeXchange provides globally coordinated proteomics data submission and dissemination. Nat Biotechnol. 2014;32:223-6.

Figure legends

Fig.1. Experimental workflow for proteome analysis

Samples were harvested from cultured *Mesorhizobium loti* (free-living) and bacteroids isolated from root nodules at 2, 3, and 4 weeks post-inoculation (WPI). Cells were disrupted using glass beads, and the resulting lysates were digested with trypsin. Tryptic peptides were tandem mass tag (TMT) labeled and subjected to liquid chromatography- mass spectrometry (LC-MS/MS) analysis. A mixture of tryptic digests of samples from free-living and symbiotic cells was used as a control. Details of the experimental workflow are described in the materials and methods.

Fig.2. Data analysis of protein quantification data.

(a) Principal component analysis (PCA) score plot for the data from three biological replicates in this study. Proteome data with similar expression profiles are clustered in close proximity. The cumulative contribution rate for principal component (PC) 1 to PC2 was 74.9%, indicating high reliability of the clustering results. (b) The loading scatter plots corresponding to these PCA score plots. (c) The individual protein profiles classified by *K*-means clustering. Each cluster profile indicates the change in average abundance of clustered proteins. (d) Functional classification of the cluster profiles of Groups 1, 2, 4, and 5. The relative frequencies were calculated by dividing the number of proteins in each category by the total number of proteins identified.

Fig.3. Map of central metabolic pathways

The relative abundances of proteins involved in glycolysis, the Entner-Doudoroff (ED) pathway, the pentose phosphate (PP) pathway, the tricarboxylic acid cycle, and amino acid synthesis at each sampling point were shown. Values were calculated by dividing the average abundance of the three biological replicates at each sampling point by the average abundance in free-living *Mesorhizobium loti*. The values are presented as means $\pm$ standard error of the mean (SEM), based on three independent experiments. To compare the individual protein abundances in bacteroids at each sampling point with those in free-living

M. loti, false discovery rate (FDR)-adjusted *P*-values were determined using an empirical Bayes moderated *t*-test and the Benjamini-Hochberg method: * *P* < 0.05, ** *P* < 0.01. The vertical axes indicate the relative abundances of proteins at each sampling point. The relative abundances of isozymes of each protein are depicted in thin-lined boxes. The relative abundances of specific proteins examined in this study are shown in gray boxes as follows: (a) phosphoenolpyruvate carboxykinase (mlr5096); (b) α -ketoglutarate dehydrogenase (mll4301); (c) branched-chain amino acid transferase (mll2664); (d) alanine dehydrogenase (mll9089); (e) two aspartate aminotransferases (mlr2541, and mlr5883).

Fig.4. Adaptation of nitrogen metabolism to symbiotic conditions

(a) The relative abundances of nitrogenase complex core subunits (NifHDK). (b) The relative abundances of proteins involved in ammonium assimilation. (c) The relative abundance of the nitrogen regulatory protein P-II; GlnK. Values were calculated by dividing the average abundances of three biological replicates at each sampling point by the average abundance in free-living *Mesorhizobium loti*. The values are presented as means $\pm$ SEM, based on three independent experiments. To compare individual protein abundances in bacteroids at each sampling point with those in free-living *M. loti*, false discovery rate (FDR)-adjusted *P*-values were determined using an empirical Bayes moderated *t*-test and the Benjamini-Hochberg method: * *P* < 0.05, ** *P* < 0.01. Abbreviations are as follows: GDH, glutamate

490 dehydrogenase; GOGAT, glutamate synthetase; GSI, Glutamine synthase I

491

492 **Fig.5. Overview of transport and metabolism in *Mesorhizobium loti* bacteroids**

493 The production of proteins involved in nitrogen fixation, antioxidant defense, iron, sulfur, and
494 phosphorus transporters, the tricarboxylic acid cycle, glycolysis, and nitrogen assimilation were
495 enhanced in *M. loti* bacteroids. Pink words indicate metabolic pathways focused in this study.

Figure 1

Figure 2

Figure 3

Figure 4

a

b

c

Figure 5