Bradyrhizobium diazoefficiens USDA110 nodulation of *Aeschynomene afraspera* is associated with atypical terminal bacteroid differentiation and suboptimal symbiotic efficiency

Running title: *Bradyrhizobium* differentiation in *Aeschynomene*

Quentin Nicoud[1*], Florian Lamouche[1*], Anaïs Chaumeret[1], Thierry Balliau[2], Romain Le Bars[1], Mickaël Bourge[1], Fabienne Pierre[1], Florence Guérard[3], Erika Sallet[4], Solenn Tuffigo[1], Olivier Pierre[1], Yves Dessaux[1], Françoise Gilard[3], Bertrand Gakière[3], Istvan Nagy[5,6], Attila Kereszt[5,6], Michel Zivy[2], Peter Mergaert[1], Benjamin Gourion[4], Benoit Alunni[1#]

[1] Université Paris-Saclay, CEA, CNRS, Institute for Integrative Biology of the Cell (I2BC), Gif-sur-Yvette, 91198, France

[2] PAPPSO, GQE-Le Moulon, INRAE, CNRS, AgroParisTech, Paris-Saclay University, 91190 Gif-sur-Yvette, France

[3] SPOmics platform, Institute of Plant Sciences Paris-Saclay (IPS2), CNRS, INRAE, Universities Paris-Saclay, Evry and de Paris, Batiment 630, 91405 Orsay, France

[4] LIPM, Université de Toulouse, INRAE, CNRS, Castanet-Tolosan, France

[5] Institute of Biochemistry, Hungarian Academy of Sciences, Biological Research Centre, 6726 Szeged, Hungary

[6] Seqomics Biotechnology Ltd., 6782 Mórahalom, Hungary

* co-first authors

Author to whom correspondence should be addressed: benoit.alunni@i2bc.paris-saclay.fr

Abstract (max. 250 words)

Legume plants can form root organs called nodules where they house intracellular symbiotic rhizobium bacteria. Within nodule cells, rhizobia differentiate into bacteroids, which fix nitrogen for the benefit of the plant. Depending on the combination of host plants and rhizobial strains, the output of rhizobium-legume interactions is varying from non-fixing associations to symbioses that are highly beneficial for the plant. *Bradyrhizobium diazoefficiens* USDA110 was isolated as a soybean symbiont but it can also establish a functional symbiotic interaction with *Aeschynomene afraspera*. In contrast to soybean, *A. afraspera* triggers terminal bacteroid differentiation, a process involving bacterial cell elongation, polyploidy and membrane permeability leading to loss of bacterial viability while plants increase their symbiotic benefit. A combination of plant metabolomics, bacterial proteomics and transcriptomics along with cytological analyses was used to study the physiology of USDA110 bacteroids in these two host plants. We show that USDA110 establish a poorly efficient symbiosis with *A. afraspera*, despite the full activation of the bacterial symbiotic program. We found molecular signatures of high level of stress in *A. afraspera* bacteroids whereas those of terminal bacteroid differentiation were only partially activated. Finally, we show that in *A. afraspera*, USDA110 bacteroids undergo an atypical terminal differentiation hallmarked by the disconnection of the canonical features of this process. This study pinpoints how a rhizobium strain can adapt its physiology to a new host and cope with terminal differentiation when it did not co-evolve with such a host.

Importance (max 150 words)

Legume-rhizobium symbiosis is a major ecological process in the nitrogen cycle, responsible for the main input of fixed nitrogen in the biosphere. The efficiency of this symbiosis relies on the coevolution of the partners. Some legume plants, but not all, optimize

their return-on-investment in the symbiosis by imposing on their microsymbionts a terminal differentiation program that increases their symbiotic efficiency but imposes a high level of stress and drastically reduce their viability. We combined multi-omics with physiological analyses to show that the non-natural symbiotic couple formed by *Bradyrhizobium diazoefficiens* USDA110 and *Aeschynomene afraspera* is functional but displays a low symbiotic efficiency associated to a disconnection of terminal bacteroid differentiation features.

Introduction

Nitrogen availability is a major limitation for plant development in many environments, including agricultural settings. To overcome this problem and thrive on substrates presenting a low nitrogen content, crops are heavily fertilized, causing important environmental damage and financial drawbacks[1,2]. Plants of the legume family acquired the capacity to form symbiotic associations with soil bacteria, the rhizobia, which fix atmospheric nitrogen for the plants' benefit. These symbiotic associations lead to the development of rhizobia-housing root organs called nodules. In these nodules, the rhizobia adopt an intracellular lifestyle and differentiate into bacteroids that convert atmospheric dinitrogen into ammonia and transfer it to the plant. Critical recognition steps occur all along the symbiotic process and define the compatibility of the plant and bacterial partners[3]. While the mechanisms involved at the early stages of the symbiosis are well described, those of the later stages are much less clear and might affect not only the ability to interact but also the efficiency of the symbiosis (ie. the plant benefit).

Nodule-specific Cysteine-Rich (NCR) antimicrobial peptides produced by legumes of the Dalbergioids and Inverted Repeat Lacking Clade (IRLC) were proposed to play a crucial role in the control of host-symbiont specificity at the intracellular stage of the symbiosis[4]. NCR peptides are targeted to the bacteroids where they govern the bacteroid differentiation[5-9]. In these legumes, the differentiation process entails such profound changes that they suppress the bacteroids' capacity to resume growth and is therefore referred to as terminal bacteroid differentiation (TBD). TBD contrasts with bacteroid formation in legumes that lack NCR genes (eg. soybean), where bacteroids are in a reversible state and can resume growth when released from nodules[10]. Specifically, TBD is associated with cell elongation, an increase in bacteroid DNA content through a cell cycle switch toward endoreduplication[6,9,11]. Furthermore, an increased permeability of the bacteroid envelope also occurs during TBD, most probably due to the interaction of NCR peptides with bacterial membranes [6,7,10,12]. Together, these alterations

of bacteroid physiology are associated to a strongly decreased viability of the differentiated bacteria, that fail to recover growth when extracted from nodules[6].

While many rhizobia have a narrow host range, some species can nodulate a large array of plant species. One of them, *Bradyrhizobium diazoefficiens* USDA110, can trigger functional nodules without TBD on soybean (*Glycine max*), cowpea (*Vigna unguiculata*) or siratro (*Macroptilium atropurpureum*) (Fig 1A-B)[13]. In addition to these species, USDA110 induces also functional nodules on the TBD-inducing legume *Aeschynomene afraspera* (Fig. 1A-C)[14,15]. However, in *A. afraspera*, USDA110 shows only very limited features that are usually associated with TBD, suggesting that the bacterium might be resistant to the TBD process[16].

Herein, we further characterized the bacteroid differentiation in the symbiosis between USDA110 and *A. afraspera*. Our observations, supported by whole-nodule metabolome analysis, indicate that USDA110 is poorly matched for nitrogen fixation with *A. afraspera*. To understand better the adaptation of *B. diazoefficiens* physiology to the *G. max* and *A. afraspera* nodule environment, we used a combination of transcriptomics (RNA-seq) and shotgun proteomics (LC-MS/MS) approaches. Finally, we find that USDA110 undergoes a terminal but atypical bacteroid differentiation in *A. afraspera* with a reduced cell viability and an increased membrane permeability, while cell size and ploidy level remain unchanged.

Results

B. diazoefficiens USDA110 is poorly matched with *A. afraspera* for nitrogen fixation

Previous reports indicate that *B. diazoefficiens* USDA110, the model symbiont of soybean, is able to establish a functional nitrogen-fixing symbiosis with *A. afraspera*, a phylogenetically distant host belonging to the Dalbergioid clade, which naturally interacts with photosynthetic rhizobia such as *Bradyrhizobium* sp. ORS285 (Fig. 1A-C)[14-18]. To evaluate the efficiency of this symbiosis, nitrogenase activity of USDA110- and ORS285-infected plants and their

115 nitrogen content were determined. Although nitrogenase activity was detected in both types of
116 nodules, it was significantly lower in USDA110-nodulated plants (Fig. 1D). A similar trend is
117 observed for mass gain per nodule mass although this difference is not significant (Fig. 1E).
118 Nitrogen and carbon contents seemedalso reduced in USDA110-nodulated plants as compared
119 to ORS285-nodulated plants, reaching levels similar to those of non-inoculated plants (Fig. S1).
120 Accordingly, ORS285-nodulated *A. afraspera* plants display darker green leaves than those
121 interacting with USDA110.

122 Moreover, their shoot/root mass ratio, a metrics that reflects the nutritional status of the
123 plant, is reduced in USDA110-nodulated *A. afraspera* plants as compared to ORS285-
124 nodulated plants, indicating that the plant nutritional needs are not fulfilled (Fig. S2)[19]. To
125 characterize further this suboptimal symbiosis, we analyzed the whole-nodule metabolome.
126 Soybean nodules infected with USDA110 were used as a reference (Fig. S3). Allantoin, which
127 is known to be the major nitrogen form exported from soybean nodules, is specifically detected
128 in them (Fig. 1F)[20]. On the contrary, asparagine and glutamine are the principal exported
129 nitrogen compounds in *A. afraspera* nodules and their amount is lower in USDA110-infected
130 nodules as compared to ORS285-infected nodules, indicating a reduced nitrogen fixation by the
131 bacteroids (Fig. 1F)[18].

132 In addition, we find specifically in USDA110-infected *A. afraspera* nodules the
133 accumulation of sucrose, phosphoric acid and ascorbate, and oppositely, a strong reduction in
134 the trehalose content (Fig. 1F, Fig. S3). Sucrose derived from phloem sap is metabolized in
135 nodules to fuel the bacteroids with carbon substrates, usually dicarboxylates. The accumulation
136 of sucrose in nodules indicates a symbiotic dysfunction. Also, the accumulation of phosphoric
137 acid in nodules suggests that nitrogen fixation is not reaching its optimal rate. Ascorbate has
138 been shown to increase nitrogen fixation activity by modulating the redox status of
139 leghemoglobin[21,22]. Thus, its accumulation in nodules with reduced nitrogen fixation capacity

could be a stress response to rescue nitrogen fixation in nodules that do not fix nitrogen efficiently. A trehalose biosynthesis gene is upregulated in ORS285 bacteroids in *A. afraspera*, suggesting that TBD is accompanied by the synthesis of this osmo-protectant disaccharide[17]. The lower synthesis in USDA110 bacteroids suggests an altered TBD. Together these data indicate a metabolic disorder in the USDA110-infected nodules, in agreement with USDA110 being a suboptimal symbiont of *A. afraspera*.

Overview of the USDA110 bacteroid proteomes and transcriptomes

In order to better understand the poor interaction between USDA110 and *A. afraspera*, the bacteroid physiology was analyzed through transcriptome and proteome analysis. The efficient soybean bacteroids and the free-living USDA110 cells cultivated in rich medium (exponential growth phase in aerobic condition) were used as references (Fig. 2A).

Prior to quantification of transcript abundance or identification and quantification of protein accumulation, the transcriptome dataset was used to re-annotate the USDA110 genome with the EugenePP pipeline[23]. This allowed the definition of 876 new CDS, ranging from 92 to 1091bp (median size = 215bp or 71.6 aa) with 11.5% of them having a predicted function or at least a match using InterProScan (IPR). This extends the total number of CDS in the USDA110 genome to 9171. Moreover, we also identified 246 ncRNAs, ranging from 49 to 765 bp (median = 76 bp), which were not annotated before. Proteomic evidence could be found for 28 new CDS (3.2% of the new CDS, median size = 97.6 aa). The complete reannotation of the genome is described in Table S1.

In the proteome dataset, 1808 USDA110 proteins were identified. Principal component analysis (PCA) of all the replicate samples and sample types revealed their partitioning according to the tested conditions. The first axis of the PCA (40.2% of the observed variance) separates bacteroid profiles from the exponential culture, whereas the second axis separates *G.*

max bacteroids from *A. afraspera* bacteroids (14.9% of the observed variance; Fig. 2B). The samples of the transcriptome datasets are similarly distributed on the PCA plot, with a first axis explaining 42.6% of the observed variance and a second axis explaining 23.5% of the observed variance (Fig. 2B).

Although differences are less pronounced in the proteome dataset than in the transcriptome dataset, COG analysis shows similar profiles across functional categories, except for membrane proteins that are less well identified in proteomics than transcriptomics (Fig. S4). In the transcriptomic dataset, 3150 genes are differentially expressed in at least one condition (differentially expressed genes or DEGs). Among the 1808 proteins identified, 815 show differential accumulation (differentially accumulated proteins or DAPs) and 438 of the cognate genes are also differentially expressed in the transcriptome datasets, whereas 175 DEGs are not DAPs (Fig. 2C).

We analyzed the Pearson correlation between transcriptomic and proteomic profiles and found that ~66% of the bacterial functions that show significant differences in both approaches display a high correlation coefficient ($r>0.9$) whereas less than 1% of the functions show strong negative correlation ($r<-0.9$; Fig. 2D). This observation suggests that the transcriptome (which provides a more exhaustive view than the proteome) and the proteome provide a complementary picture of bacterial physiology, and they tend to show a congruent information if we restrict our analysis to the genes with differential accumulation/expression (Fig. 2E). However, there is still around 66% of the DEGs, which were detected by the proteomic analysis, that are not DAPs. Our description of the bacterial functions will be primarily based on the functions that are both DEGs and DAPs, as there is stronger evidence of their modulation in the tested conditions. The transcriptome alone will be used only when proteomics is not informative, for example to analyze regulons and stimulons that have been described previously in USDA110.

Symbiotic functions common to both types of USDA110 bacteroids

Among the 815 DAPs, 705 and 699 proteins are significantly differentially accumulated in *G. max* and *A. afraspera* respectively compared to the bacterial culture control. Strikingly, 646 proteins are commonly differentially accumulated in both plant nodules (Table S1).

In the transcriptomic dataset, 1999 DEGs, representing ~21% of the genome, were identified between the bacterial culture and the bacteroids, regardless of the host. Among them, 1076 genes displayed higher expression in nodules (including seven newly annotated ncRNAs and one newly annotated CDS among the 20 differentially expressed genes with highest fold change) and 923 genes were repressed *in planta* (including two newly annotated ncRNAs and two newly annotated CDS among the 20 DEG with highest fold change, Table S1).

Restricting the analysis to the bacterial functions that are both differentially expressed (DEG) and differentially accumulated (DAP) *in planta* in both hosts as compared to the bacterial culture identified 222 genes/proteins, 150 being upregulated and 72 being repressed *in planta* respectively (Fig. 3A). Notably, six newly annotated genes are in this gene list including one putative regulator (Bd110_01119) that is induced during symbiosis. Among the functions commonly DEG and DAP *in planta*, only four functions showed opposite trends in proteomics and transcriptomics.

The proteome and transcriptome data provided a coherent view of the nitrogen fixation metabolism of *B. diazoefficiens* in the tested conditions. Key enzymes involved in microoxic respiration and nitrogen fixation were detected amongst the proteins having the highest spectra number in the nodule samples (Fig. 3A, Table S1) and the corresponding genes are among the most strongly expressed ones in bacteroids, while almost undetectable in the free-living condition. This includes for instance, the nitrogenase and the nitrogenase reductase subunits, which constitute the nitrogenase enzyme complex responsible for nitrogen conversion into

ammonia. They belong to a locus of 21 genes from *blr1743* (*nifD*) to *bll1778* (*ahpC_2*), including the genes involved in nitrogenase cofactor biosynthesis, in electron transport to nitrogenase, and in microaerobic respiration, that are among the highest expressed ones in bacteroids of both host plants, both at the gene and protein expression level. The slightly higher level of the dinitrogenase reducatse NifH detected in proteomics was not supported by western blot analysis, which showed apparent similar protein level in both bacteroid conditions (Fig. S5). Strikingly, the two bacteroid types did not show a notable difference in the expression of these genes and proteins, suggesting that the activation of the nitrogen fixation machinery is not a limiting factor underlying the suboptimal efficiency of strain USDA110 in *A. afraspera* nodules.

In addition to these expected bacteroid functions, many other proteins were identified that specifically and strongly accumulated in both nodule types. This is the case of the chaperonins GroEL1/GroES1, which are strongly upregulated and reach high gene expression and protein levels in both bacteroids. The upregulation of these chaperonins is remarkable because other GroEL/GroES (4, 5 and 7) proteins are also very strongly accumulated in a constitutive manner. This indicates that bacteroids have a high demand for protein folding, possibly requiring specific GroEL isoforms, a situation reminding the requirement of one out of five GroEL isoforms for symbiosis in *Sinorhizobium meliloti*, the symbiont of *Medicago sativa*[12,24]. Another example of a bacteroid-specific function is the hydrogenase uptake system, whose gene expression was induced in both bacteroid types from nearly no expression in culture. Hydrogenase subunit HupL_2 (*bll6941*) was found amongst the proteins displaying the highest spectra number in the nodule samples suggesting important electron recycling in bacteroids of the two hosts. Another one is the 1-aminocyclopropane-1-carboxylic acid (ACC) deaminase (*blr0241*), which was also amongst the most strongly accumulated proteins in nodules and was significantly less abundant in free-living USDA110. An outer membrane protein (*bll1872*)

belonging to the NifA regulon[25] was also strongly induced *in planta*, with a transcript level among the top 10 genes in *A. afraspera*. Additionally, a locus of seven genes (*blr7916-blr7922*) encoding an amidase enzyme and a putative peptide transporter composed of two transmembrane domain proteins, two ATPases and two solute-binding proteins was strongly upregulated in the two bacteroid types, with three protein being also over-accumulated *in planta* (Fig. 3A; Table S1).

Oppositely, motility genes encoding flagella subunits (*bll5844-bll5846*), metabolic enzymes and transporter subunits are strongly downregulated during symbiosis and hardly detectable at the protein level *in planta* (Fig. 3A).

Taken together, these data show that both bacteroid types display a typical nitrogen fixation-oriented metabolism, with a partial shutdown of housekeeping functions. This indicates that despite the apparent reduced symbiotic efficiency of USDA110 in *A. afraspera* nodules, the bacterium fully expresses its symbiotic program within this non-native host as it does in soybean, its original host. Thus, the sub-optimal functioning of the *A. afraspera* nodules does not seem to come from a bacterial defect to express the symbiotic program, but possibly from an unfavorable host microenvironment or from a lack of metabolic integration of these maladapted partners.

Host-specific functions

Comparison of the *A. afraspera* and *G. max* bacteroids revealed also significant differences in the proteomes and transcriptomes. At the transcriptomic level, 935 DEGs could be identified between the two bacteroid types (509 *A. afraspera* > *G. max* and 426 *G. max* > *A. afraspera*). One notable feature of the transcriptome is the identification of four newly annotated ncRNA and one new CDS amongst the 20 most induced DEGs in *A. afraspera* nodules and the presence of five newly annotated CDS amongst the 20 most induced DEGs in *G. max* nodules (Table

S1). However, when considering only the functions that display congruent and significant differences in terms of transcripts and protein levels between plant hosts, we fall down to 63 genes/proteins, 33 being induced in *A. afraspera* nodules and 30 being induced in *G. max* nodules (Fig. 3B).

Interestingly, the phenylacetic acid degradation pathway (PaaABCDEIK, *blr2891-blr2897*) was highly expressed in *A. afraspera* nodules (although only PaaABCD and PaaK have been detected by proteomics), as well as a yet uncharacterized cluster of genes putatively involved in toluene degradation (*blr3675-blr3680*). The chaperone GroEL2 is also specifically induced in *A. afraspera*. Similarly, three S1 peptidases (Dop: *blr2591*, *blr3130* and *blr7274*) are highly expressed in the nodules of this latter host together with a RND efflux pump (*bll3903*) and a LTXXQ motif protein (*bll6433*), a motif also found in the periplasmic stress response CpxP[26]. The over-accumulation of these proteins suggests that bacteroids are facing stressful conditions during this interaction with *A. afraspera*. An uncharacterized ABC transporter solute binding protein (*blr7922*) was also overexpressed in *A. afraspera*.

One αβ hydrolase (*blr6576*) and a TonB-dependent receptor-like protein (*bll2460*) were over-accumulated in a *G. max*-specific manner. Similarly, an uncharacterized metabolic cluster including transketolases (*blr2167-blr2170*), the heme biosynthetic enzyme HemN1 (*bll2007*) and to a lesser extent an anthranilate phosphoribosyl-transferase (TrpD encoded by *bll2049*) are overexpressed in soybean nodules.

USDA110 transcriptomics data in the perspective of previously described regulons and stimulons

USDA110 is one of the best-characterized rhizobial strains in terms of transcriptomic responses to various stimuli as well as the definition of regulons[27]. We analyzed the behavior of these previously defined gene networks in USDA110 in our dataset (Table S2). To initiate

the molecular dialog that leads to nodule formation, plants secrete flavonoids like genistein in their root exudates, which are perceived by the rhizobia and trigger Nod factor production. At 14 dpi, when the nodule is formed and functioning, the genistein stimulon, which comprises the NodD1, NodVW, TtsI and LafR regulons, is not anymore activated in bacteroids. The symbiotic regulons controlled by NifA, FixK1, FixK2, FixLJ and sigma54 (RpoN) were activated *in planta*, indicating that nitrogen fixation was going on in both hosts. Accordingly, the nitrogen metabolism genes controlled by NtrC were activated *in planta*. Additionally, the PhyR/EcfG regulon involved in general stress response is not activated in bacteroids. Differences between hosts were however not observed for any of these regulons/stimulons. The only stimulon that showed differential expression between hosts is the one involved in aromatic compound degradation, which was highly expressed in *A. afraspera* nodules. Similar upregulation of the vanillate degradation pathway was observed in the transcriptome of *Bradyrhizobium* sp. ORS285 in *A. afraspera* and *A. indica* nodules[17], suggesting that Dalbergioid hosts display a higher aromatic compound content in nodules than *G. max*. In line with this hypothesis, some of the most differentially accumulated sets of proteins (*A. afraspera* > *G. max*) are involved in the degradation of phenylacetic acid (PaaABCDK and *bll0339*) suggesting that the bacterium converts phenylalanine (or other aromatic compounds) ultimately to fumarate through this route (Fig. 3B)[28]. Similarly, enzymes of another pathway involved in phenolic compound degradation (*blr3675-blr3680*) are accumulated in *A. afraspera* nodules (Fig. 3B, Table S1).

Expression pattern of orthologous genes between ORS285 and USDA110 in *A. afraspera* nodules

In a previous study[17], a transcriptome analysis was performed on *Bradyrhizobium sp.* ORS285 in interaction with *A. afraspera* and in culture. *Bradyrhizobium* sp. ORS285 is a strain that co-evolved with *A. afraspera*, leading to an efficient symbiosis hallmarked by TBD, *id est*

315 cell elongation and polyploidization of the bacteroids. In order to compare gene expression of
316 these two nodule-forming rhizobia in culture and *in planta*, we determined the set of
317 orthologous genes between the two strains using the Phyloprofile tool of Mage Microscope
318 website. This analysis yielded a total of 3725 genes (Table S3). The heatmap on Figure 4A
319 presents the modulation of gene expression (LFC) between *A. afraspera* nodules and the
320 bacterial culture for the orthologous genes in each bacterium, regardless of their statistical
321 significance. When taking FDR < 0.01 in account, we identified sets of genes that are
322 differentially expressed *in planta* in either bacterium or in both (Fig. 4B).

323 Only 343 genes displayed differential expression (FDR < 0.01 and |LFC| > 1.58) *in planta*
324 in both bacteria as compared to their respective culture control (Fig. 4C). A majority of these
325 genes (86.8%) exhibited congruent expression patterns. First, the *nif*, *fix* and *hup* genes are
326 commonly and highly induced in both strains during their symbiotic life with *A. afraspera*, a
327 hallmark of a functional symbiosis. However, there are differences in their expression level,
328 with a higher expression of the symbiotic genes in ORS285 (*nifHDK* represent 12.5% of all
329 reads in *A. afraspera* nodules)[17] than in USDA110 (*nifHDK* represent only 2.5% of all reads in
330 *A. afraspera* nodules), consistently with a more efficient interaction occurring between ORS285
331 and *A. afraspera*. Additionally, the Kdp high affinity transport system, the phosphate (*pstCAB,*
332 *phoU, phoE, phoC*) and phosphonate metabolism (*phnHIJKL*) are activated *in planta* in both
333 bacteria (Fig. 4B-C). The stress-marker *dop* protease gene is also induced in both bacteria in *A.*
334 *afraspera* nodules (Fig. 4C).

335 Additionally, 1026 genes were differentially expressed solely in ORS285, and similarly
336 there was 604 DEG specific to USDA110 (Fig. 4B). For example, the general secretory pathway
337 seems to be specifically induced in ORS285[17]. Oppositely, USDA110 displays an induction of
338 the *rhcJQRU* genes which are involved in the injection of type three effector proteins that can
339 be important for the establishment of the symbiosis whereas they are not induced or even

repressed in ORS285 (Fig. 4B). This is also the case of the nitrite reductase encoding gene *nirK* (*blr7089*/BRAD285_v2_0763; Fig. 4C). In addition, USDA110 induces the expression of an ACC deaminase (*blr0241*), while its ortholog is repressed in ORS285 (BRAD285_v2_3570) during symbiosis (Fig. 4C). Bacterial ACC deaminases can degrade ACC, a precursor of ethylene, and thereby modulate ethylene levels in the plant host and promote the nodulation process[29].

Bradyrhizobium diazoefficiens USDA110 bacteroids undergo _bona fide_ TBD in _Aeschynomene afraspera_ nodules despite very weak morphological and ploidy modifications

In a previous description of the *A. afraspera* - *B. diazoefficiens* USDA110 interaction, the typical TBD features were not observed and the bacteroids were very similar to those in *G. max* where no TBD occurs[16]. At the molecular level, accumulation of the replication initiation factor DnaA is higher in soybean than in *A. afraspera* (Table S1). Similarly, the MurA peptidoglycan synthesis enzyme (encoded by *bll0822*) that may play a role in cell elongation during TBD was detected to similar levels in both bacteroids (Table S1). Taken together, the molecular data do not clearly indicate whether USDA110 bacteroids undergo TBD in *A. afraspera*. Therefore, we investigated the features of the USDA110 bacteroids in *A. afraspera* nodules in more detail.

We analyzed bacteroid differentiation features in USDA110 bacteroids extracted from soybean and *A. afraspera* nodules. The interaction between *A. afraspera* and *Bradyrhizobium* sp. ORS285 was used as a positive control for TBD features[9,30,31]. TBD is characterized by cell elongation. We quantified cell length, width, area and shape of purified bacteroids and culture controls. Whereas ORS285 bacteroids were enlarged within *A. afraspera* nodules as compared to their free-living conterparts, USDA110 bacteroids were similar to free-living bacteria in both soybean and *A. afraspera* (Fig. 5A; Fig. S6). Another feature of TBD is endoreduplication.

Analysis of the bacterial DNA content of ORS285 bacteroids in *A. afraspera* by flow cytometry shows peaks at 6C and more[9]. As expected, USDA110 bacteroids in *G. max* yields only two peaks, at 1C and 2C, similarly to the cycling cells in the bacterial culture sample (Fig. 5B)[16]. Strikingly, similar results were obtained for USDA110 in *A. afraspera*. Thus, with respect to the DNA content and cell size, the USDA110 bacteroids do not display the typical TBD features in *A. afraspera* nodules. Loss of membrane integrity is a third hallmark of TBD that likely strongly contributes to the loss of viability of bacteroids. Time-course analysis of propidium iodide (PI) uptake by bacteroids and the corresponding culture controls were performed to assess bacteroid permeability (Fig. S7). Twenty minutes after PI application, USDA110 bacteroids from *A. afraspera* display an increased permeability that is much closer to ORS285 bacteroids in interaction with *A. afraspera* than to the low permeability of USDA110 bacteroids from *G. max* nodules (Fig. 5C). Also the free-living counterparts exhibit a very low permeability. Taken together, this suggests that the envelope of USDA110 bacteroids is more permeable in the NCR-producing *A. afraspera* nodules, even if it does not reach the permeability level of the ORS285 strain. To analyze bacterial viability, bacteroids extracted from nodules were plated and the colony forming units (cfu) were determined (Fig. 5D). In *G. max*, USDA110 formed $1.46x10^{10}$ colonies/mg nodule (~100% survival). Oppositely, ORS285 formed only $5.42x10^{7}$ colonies/mg nodule in *A. afraspera* (~0.5% survival). Interestingly, USDA110 formed $1.13x10^{8}$ colonies/mg nodule in *A. afraspera* (~1% survival), indicating that, despite the absence of cell enlargment and endoreduplication USDA110 bacteroids lose their viability and undergo a *bona fide* terminal differentiation in *A. afraspera*. Thus, in the NCR-producing plant *A. afraspera*, USDA110 bacteroids display a disconnection of the four canonical TBD features (ie. cell size, ploidy level, membrane permeability and cell viability).

Discussion

Aeschynomene afraspera triggers atypical but terminal differentiation of USDA110 bacteroids

In a previous study, we noticed that, in *A. afraspera*, USDA110 forms a functional symbiosis although bacteroids do not display features that are usually associated with TBD[16]. Here we show that no endoreduplication and cell elongation of USDA110 occur in terminally differentiated bacteroids that fix nitrogen in a suboptimal way. Accordingly, the protein level of DnaA, the genome replication initiator, was higher in soybean than in *A. afraspera* bacteroids and the MurA level was not different between bacteroid conditions, confirming that polyploidization and cell elongation did not occur in this host. Such unusual terminal bacteroid differentiation is reminiscent of the bacteroids in *Glycyrrhiza uralensis*. This plant of the IRLC expresses NCR peptides[11]. However, one of its compatible symbionts, *Sinorhizobium fredii* strain HH103, does not undergo any loss of viability, no change in DNA content and no cell elongation[32], while another symbiont, *Mesorhizobium tianshanense* strain HAMBI 3372 showed all TBD features[33]. The influence of the bacterial genotype on terminal/non-terminal differentiation of bacteroids was also suggested in *Medicago truncatula* in which, the gene *hrrP* might confer to some *Sinorhizobium* strains a resistance against the differentiation process triggered by some *M. truncatula* ecotypes[34]. In these two IRLC plants (ie. *M. truncatula* and *G. uralensis*), bacteria undergo a complete TBD or no TBD at all in a strain-dependent manner, but there is no clear uncoupling of the features of TBD (cell elongation/endoreduplication/altered viability) as shown here in the case of *B. diazoefficiens* USDA110-*A. afraspera*.

The surprising differentiation of USDA110 in *A. afraspera* nodules raises questions about the molecular mechanisms supporting this phenomenon. We consider two possible hypotheses: strain USDA110 might be more sensitive to the differentiation factors of the host than strain

ORS285 and be rapidly "terminally" differentiated, before the other differentiation features, that are potentially important for symbiotic efficiency, can take place. Alternatively, USDA110 might be resistant to the plant effectors that trigger the elongation and polyploidization features.

In agreement with the latter possibility, the application of NCR peptides has very limited effect on strain USDA110 as compared to *S. meliloti* and to other plant-associated bacteria[16,35]. NCR insensitivity may be due to the thick hopanoid layer that is present in the outer membrane of strain USDA110, as the hopanoid biosynthesis mutant *hpnH* is more sensitive to NCR peptides and shows symbiotic defects in *A. afraspera* but not in *G. max*[36]. Moreover, the altered peptidoglycan structure in the strain USDA110 DD-carboxypeptidase mutant resulted in an increased TBD process with endoreduplicated and elongated bacteroids in *A. afraspera*[16]. This suggests that the envelope of strain USDA110 prevents a canonical TBD to occur. Possibly, NCR peptides are not able to reach their intracellular targets required to induce endoreduplication and cell division arrest, while their effect on cell viability through pore formation and membrane destabilization is still effective.

A survey of TBD in the legumes has identified multiple occurrences of the process in several subclades of the legumes but found that the majority of legumes do not have TBD[37]. The classification in this study was based on a morphological analysis of the bacteroids. Ancestral state reconstruction based on this classification suggested that the non-differentiated bacteroids are ancestral and that TBD evolved at least five times independently in legumes[37]. The discovery of bacteroids that are terminally differentiated without any obvious morphological changes opens the possibility that the occurrence of TBD might be underestimated in the legume family. Similarly, in the IRLC clade, the extent of morphological bacteroid differentiation was correlated to the size of the cationic NCR peptides repertoire and in legumes with few NCR peptides, the morphological modification of bacteroids can be minor[11,33]. In addition, at the molecular level, TBD is originally ascribed to the production of symbiotic

antimicrobial peptides, the NCRs, by nodules[7], but more recently, other types of antimicrobial peptides such as the NCR-like, GRP, MBP1 and CAPE peptides specifically produced in nodules of different plants were proposed to contribute to bacteroid differentiation[9,38-40]. Thus, if TBD would indeed be more widespread than currently estimated on the basis of morphological bacteroid features, the currently porposed evolutionary scenario of bacteroid formation might require revision.

Terminal differentiation is associated with specific stress response

The TBD of strain USDA110 in *A. afraspera* is associated with a higher accumulation of stress markers compared to the *G. max* bacteroids. These markers include four proteases (Dop, Lon_2, *blr3130* and *blr7274*) and one chaperonin (GroEL_2). Similar induction of proteases and chaperonins have been reported in NCR-treated *S. meliloti* cultures[35], indicating that this response may be linked to the perception of *A. afraspera* NCR-like peptides in USDA110.

The genes encoding these stress related proteins are not part of the well-characterized general stress response (GSR) controlled by the PhyR/EcfG signaling cascade in *B. diazoefficiens* USDA110[41]. On the other hand, we found that the PhyR/EcfG regulon in USDA110 is not activated in the bacteroids of both host plants (Table S2). This observation contrasts with our previous study of *Bradyrhizobium* sp. ORS285 transcriptome during symbiosis with *Aeschynomene* plants, which showed that the PhyR/EcfG cascade was upregulated *in planta*[17]. Nevertheless, the expression of the Dop protease was induced in *A. afraspera* in both bacteria (Fig. 4C). Together, the omics data suggest that bacteroids of *Bradyrhizobium* spp. activate stress-related genes in the TBD-inducing *A. afraspera* host but differences exist in the activation of specific stress responses at the strain level.

Correlation between bacteroid differentiation features and symbiotic efficiency for the plant

TBD is associated with the massive production of symbiotic antimicrobial peptides such as NCR, NCR-like and CAPE peptides in different plants[5,9,38,40]. They represent ~10% of the nodule transcriptomes in *M. truncatula* (analysis of the data from ref 42) and their production is thus potentially a strong energetic cost for the plant, raising questions about the benefits of the TBD process. TBD appeared independently in different legume clades[9,37], suggesting that plants imposing this process obtain an advantage which might be a higher symbiotic benefit. Increased symbiotic efficiency has indeed been observed in hosts imposing TBD[17,43,44]. The findings reported here, comparing bacteroids and symbiotic efficiency in *A. afraspera* infected with strain ORS285 and strain USDA110, are in agreement with this hypothesis. Also in the symbiosis of *M. truncatula* in interaction with different *S. meliloti* strains, a similar correlation was observed between the level of bacteroid differentiation and the plant growth stimulation[45]. However, the simultaneous analysis of the bacteroid differentiation and symbiotic performance of an extended set of *Aeschynomene-Bradyrhizobium* interactions has shown that, perhaps not unexpectedly, the symbiotic efficiency of the plant-bacterium couple is not solely correlated with bacteroid differentiation and that other factors can interfere with the symbiotic efficiency as well[46].

Conclusion

Bradyrhizobium diazoefficiens USDA110 is a major model in the legume-rhizobium symbiosis, mainly thanks to its interaction with *G. max*, the worldwide most cultivated legume. Although omic studies have been conducted in this strain in symbiosis with various hosts[13,25], this is the first time that this bacterium is studied at the molecular level in symbiosis with a NCR-producing plant that normally trigger a typical terminal bacteroid differentiation in its

symbionts. The symbiosis between USDA110 and *A. afraspera* is functional even if nitrogen fixation and plant benefits are sub-optimal.

Terminal bacteroid differentiation is taking place in the NCR-producing host *A. afraspera*, as bacterial viability is impaired in USDA110 bacteroids, whereas morphological changes and the cell cycle switch to endoreduplication are not observed. We also show by combining proteomics and transcriptomics that the bacterial symbiotic program is expressed in *A. afraspera* nodules in a similar way as in *G. max*, although host-specific patterns were also identified. However, the bacterium is under stressful conditions in the *A. afraspera* host, possibly due to the production of NCR-like peptides in this plant. Integration of datasets from different bacteria in symbiosis with a single host, like ORS285 and USDA110 in symbiosis with *A. afraspera*, shed light on the differences in the stress responses activated in *A. afraspera* and confirmed that the symbiosis is functional but suboptimal in this interaction. The molecular data presented here provide a set of candidate functions that could be analyzed for their involvement in the adaptation to a new host and to the TBD process.

Material and Methods

Bacterial cultures and bacteroid extraction

B. diazoefficiens USDA110[47] and *Bradyrhizobium* sp. ORS285 were cultivated in YM medium at 30°C in a rotary shaker[48]. For transcriptomic analysis, culture samples (OD_{600nm} = 0.5) were collected and treated as in Chapelle et al. (2015)[49].

G. max ecotype Williams 82 and *A. afraspera* seeds were surface-sterilized and the plants were cultivated and infected with rhizobia for nodule formation as described in Barrière et al. (2017)[16]. Nodules were collected at 14 days post inoculation (dpi), immediately immersed in liquid nitrogen and stored at -80°C until use. Each tested condition (in culture and *in planta*) was produced in biological triplicates.

514 **Phylogeny analysis**

515 Nucleotide sequence of *matK* genes were collected on NCBI using accession numbers
516 described in references 50 and 51 and analyzed on phylogeny.fr (www.phylogeny.fr). They
517 were aligned using ClustalW with manual corrections, before running a phyML (GTR - Gamma
518 model) analysis with 500 bootstraps. A Bayesian inference tree was also generated (GTR + G
519 + I) and provided similar topology as the maximum likelihood tree (data not shown). Trees
520 were visualized and customized using TreeDyn.

521

522 **Genome annotation and RNA-seq analysis**

523 Nodule and bacterial culture total RNA was extracted and treated as previously described
524 in [17]. Oriented (strand-specific) libraries were produced using the SOLiD Total RNA-seq kit
525 (Life Technologies) and were sequenced on a SOLiD 3 station yielding ~40 million 50bp single
526 reads. Trimming and normalization of the reads were performed using the CLC workbench
527 software. Subsequently, the reads were used to annotate the genome using EugenePP[23], and the
528 mapping was performed using this new genome annotation. Analysis of the transcriptome using
529 DE-seq2 and data representation were performed as previously described[17]. Differentially
530 expressed genes (DEG) showed an asbolute $\log_2$ fold change ($|LFC| > 1.58$ (ie. fold change >
531 3) with a false discovery rate (FDR) < 0.01.

532

533 **Proteomic analysis**

534 Bacteroids were extracted from 14 dpi frozen nodules[6], while bacterial culture samples were
535 collected as above, and the bacterial pellets were resuspended in -20°C acetone and lysed by
536 sonication. Protein solubilization, dosage, digestion (trypsin 2% w/w) and solid phase
537 extraction (using Phenomenex polymeric C18 column) were performed as described before[52].
538 Peptides from 800 ng of proteins were analyzed by LC-MS/MS with a Q Exactive mass

spectrometer (Thermo Electron) coupled to a nanoLC Ultra 2D (Eksigent) using a nanoelectrospray interface (non-coated capillary probe, 10 µ i.d.; New Objective). Peptides were loaded on a Biosphere C18 trap-column (particle size: 5 µm, pore size: 12 nm, inner/outer diameters: 360/100 µm, length: 20 mm; NanoSeparations) and rinsed for 3 min at 7,5µl minute of 2% Acetonitrile (ACN), 0,1% Formic acid (FA) in water. Peptides were then separated on a Biosphere C18 column (particle size: 3 µm, pore size: 12 nm, inner/outer diameters: 360/75 µm, length: 300 mm; NanoSeparations) with a linear gradient from 5% of 0,1% FA in ACN (buffer B) and 95% of 0,1% FA in Water (buffer A) to 35% of buffer B and 65% of buffer A in 80 min at 300nl/min followed by a rinsing step at 95% of buffer B and 5% of buffer A for 6 min and a regeneration step with parameters of the start of the gradient for 8 min. peptide ions were analyzed using Xcalibur 2.1 software in data dependent mode with the following parameters: (I) full ms was acquire for the 400-1400 mz range at a resolution of 70000 with an AGC target of 3.10^6; (ii) MS^2 scan was acquired at a resolution of 17500 with an agc target of 5.10^4, a maximum injection time of 120 ms and an isolation window of 3 m/z. The normalized collision energy was set to 27. MS^2 scan was performed for the eight most intense ions in previous full MS scan with an intensity threshold of 1.10^3 and a charge between 2 and 4. Dynamic exclusion was set to 50s. After conversion to mzXML format using msconvert (3.0.3706)[53], data were searched using X!tandem (version 2015.04.01.1)[54] against the USDA110 reannotated protein database and a homemade database containing current contaminants. In a first pass trypsin was set to strict mode and cysteine carbamidomethylation as a fixed modification and methionine oxidation, protein N-terminal acetylation with or without protein N-terminal methionine excision, N-terminal glutamine and carbamidomethylated cysteine deamidation, N-terminal glutamic dehydration as potential modifications. In a refine pass, Semi enzymatic peptides were allowed. Proteins inference was performed using X!TandemPipeline (version 3,4,3)[55]. A protein was validated with an E-value

564 $< 10^{-5}$ and 2 different peptides with an E-value < 0.05. Protein from contaminant database
565 (*Glycine max* proteins and unpublished *Aeschynomene* Expressed Sequence Tags) were
566 removed after inference. Proteins were quantified using the spectral counting method[56]. To
567 discriminate differentially accumulated proteins (DAPs), ANOVA analysis was performed on
568 the spectral counts and proteins were considered DAP when p-value < 0.05.
569
570 **Metabolomic analysis**
571 Metabolites and cofactors were extracted from lyophilized nodules and analyzed by GC-
572 MS and LC-MS respectively according to Su et al. (2016)[57].
573
574 **Plant biomass and nitrogen fixation analysis**
575 Dry mass of shoot, root and nodules was measured, and shoot-root mass ratio was
576 calculated. The mass gain per g of dry nodule was calculated as the difference between total
577 mean masses of the plants of interest and of the non-inoculated plants, divided by the mean
578 mass of nodules. Thirty plants were used per condition. Nitrogenase activity was assessed by
579 Acetylene Reduction Assay (ARA) on ten plants per condition as previously described[31]. The
580 elemental analysis of leaf carbon and nitrogen content was performed as described in reference
581 18.
582
583 **Analysis of *B. diazoefficiens* USDA110 regulons and stimulons**
584 Gene sets defined as regulons and stimulons were collected form the literature and the
585 regulons/stimulons were considered as activated/repressed when ≥ 40% of the corresponding
586 genes were DEG in a host plant as compared to the culture condition.
587
588 **Comparison of orthologous gene expression between *B. diazoefficiens* USDA110 and**
589 ***Bradyhrizobium* sp. ORS285**

The list of orthologous genes between USDA110 and ORS285 was determined using the Phyloprofile tool of the MicroScope-MAGE platform[58], with identity threshold of 60%, maxLrap > 0 and minLrap > 0.8. The RNA-seq data from reference 17 and those of this study were used to produce heatmaps, for the genes displaying FDR < 0.01 (*A. afraspera* vs. YM), using R (v3.6.3) and drawn using pheatmap (v1.0.12) coupled with kohonen (v3.0.10) for gene clustering using the Self Organizing Maps (SOM) method. The DEG in both organisms (*A. afraspera* vs YM) were plotted for USDA110 and ORS285.

Analysis of TBD features

Bacteroids were extracted from 14 dpi nodules and analyzed using a CytoFLEX S (Beckman-Coulter)[31]. For ploidy and live/dead analyses, samples were stained with propidium iodide (PI, ThermoFisher, 50 µg.mL^{-1} final) and Syto9 (ThermoFisher, 1.67 µM final). PI permeability was assessed over time on live bacteria. *Bradyrhizobium* sp. ORS285.pMG103-*nptII-GFP*[30] and *B. diazoefficiens* USDA110 sYFP2-1[59] strains were used to distinguish bacteroid from debris during flow cytometry analysis. For each time point, the suspension was diluted 50 times for measurement in the flow-cytometer. The percentage of bacteroids permeable to PI was estimated as the ratio of PI-positive over total bacteroids (GFP/YFP positive). Heat-killed bacteroids were used as positive control to identify the PI-stained bacteroid population.

For bacteroid viability assays, nodules were collected and surface sterilized (1 min NaClO 0.4%, 1 min 70% ethanol, two washes in sterile water). Bacteroids were subsequently prepared as previously described[31] and serially diluted and plated (five µl per spot) in triplicate on YM medium containing 50 µg.mL^{-1} carbenicillin. Colony-forming units (cfu) were counted five days post plating and divided by the total nodule mass.

Bacterial cell shape, length and width were determined using confocal microscopy image analysis. Bacteroid extracts and stationary phase bacteria cultures we stained with 2.5 nM Syto9 for 10 minutes at 37°C and mounted between slide and coverslip. Bacteria imaging was performed on a SP8 laser scanning confocal microscope (Leica microsystems) equipped with hybrid detectors and a 63x oil immersion objective (Plan Apo, NA: 1.4, Leica). For each condition, multiple z-stacks (2.7µm width, 0.7 µm step) were automatically acquired (excitation: 488 nm; collection of fluorescence: 520-580 nm).

Prior to image processing, each stack was transformed as a maximum intensity projection using ImageJ software (https://imagej.nih.gov/ij/). Bacteria detection was performed with MicrobeJ (https://www.microbej.com/)[60]. First, bacteria were automatically detected on every image using an intensity based thresholding method with a combination of morphological filters (area: 1-20 µm²; length: 1 µm-∞; width: 0.5-1.3 µm) and every object was fitted with a "Rod-shaped" bacteria model. To ensure high data quality every image was manually checked to remove false positive (mainly plant residues) and include rejected objects (mainly fused bacteria). Then the morphology measurements and figures were directly extracted from MicrobeJ. ORS285 in culture and in symbiosis with *A. afraspera* were used as references for the analysis of TBD features.

Western blot analysis

Detection of NifH by western blotting was performed using a commercial polyclonal antibody against a NifH peptide (Agrisera) respectively. The western blotting was carried out as previously described[61] using bacterial exponential (OD_{600nm} = 0.5) and stationary (OD_{600nm} > 2.5) phase cultures as well as 14 dpi nodule-extracted bacteroids.

Acknowledgments

The authors would like to thank Dora Latinovics for production and sequencing of the RNA-seq libraries and Mélisande Blein-Nicolas for her advices regarding the statistical analysis of the proteomic dataset. Q.N. and F.L. were supported by a PhD fellowship from the Université Paris-Sud. The present work has benefited from the core facilities of Imagerie-Gif (http://www.i2bc.paris-saclay.fr), member of IBiSA (http://www.ibisa.net), supported by 'France-BioImaging' (ANR-10-INBS-04–01), and the Labex 'Saclay Plant Sciences' (ANR-11-IDEX-0003-02). This work was funded by the Agence Nationale de la Recherche, grants n° ANR-13-BSV7-0013 and ANR-17-CE20-0011 and used resources from the National Office for Research, Development and Innovation of Hungary, grant n° 120120 to A.K.

Authors' contribution

QN, FL, PM, BGo and BA designed the work. QN, FL, AC, TB, MB, FGu, ES, ST and OP performed the experiments. QN, FL, MB, ES, YD, BGa, FGi, IN, AK, MZ, PM, BGo and BA analyzed the data. QN, FL, PM, BGo and BA wrote the paper.

References

1. Erisman, J.W., Galloway, J.N., Seitzinger, S., Bleeker, A., Dise, N.B., Petrescu, A.M.R., et al. (2013) Consequences of human modification of the global nitrogen cycle. *Philos Trans R Soc B Biol Sci* **368**: 20130116–20130116.
2. Zhao, C., Liu, B., Piao, S., Wang, X., Lobell, D.B., Huang, Y., et al. (2017) Temperature increase reduces global yields of major crops in four independent estimates. *Proc Natl Acad Sci U S A* **114**: 9326–9331.
3. Oldroyd, G. (2013) Speak, friend, and enter: signalling systems that promote beneficial symbiotic associations in plants. *Nat Rev Microbiol* **11**: 252–263.

4. Gourion, B. and Alunni, B. (2018) Strain-specific symbiotic genes: A new level of control in the intracellular accommodation of rhizobia within legume nodule cells. *Mol Plant-Microbe Interact* **31**: 287–288.
5. Mergaert, P., Nikovics, K., Kelemen, Z., Maunoury, N., Vaubert, D., Kondorosi, A., et al. (2003) A novel family in *Medicago truncatula* consisting of more than 300 nodule-specific genes coding for small, secreted polypeptides with conserved cysteine motifs. *Plant Physiol* **132**: 161–73.
6. Mergaert, P., Uchiumi, T., Alunni, B., Evanno, G., Cheron, A., Catrice, O., et al. (2006) Eukaryotic control on bacterial cell cycle and differentiation in the rhizobium-legume symbiosis. *Proc Natl Acad Sci U S A* **103**: 5230–5235.
7. Van de Velde, W., Zehirov, G., Szatmari, A., Debreczeny, M., Ishihara, H., Kevei, Z., et al. (2010) Plant peptides govern terminal differentiation of bacteria in symbiosis. *Science* **327**: 1122–1126.
8. Guefrachi, I., Nagymihaly, M., Pislariu, C.I., Van de Velde, W., Ratet, P., Mars, M., et al. (2014) Extreme specificity of NCR gene expression in *Medicago truncatula*. *BMC Genomics* **15**(1):712.
9. Czernic, P., Gully, D., Cartieaux, F., Moulin, L., Guefrachi, I., Patrel, D., et al. (2015) Convergent evolution of rndosymbiont differentiation in Dalbergioid and Inverted Repeat-Lacking Clade legumes mediated by nodule-specific cysteine-rich peptides. *Plant Physiol* **169**: 1254–1265.
10. Alunni, B., and Gourion, B. (2016) Terminal bacteroid differentiation in the legume-rhizobium symbiosis: nodule-specific cysteine-rich peptides and beyond. *New Phytol* **211**: 411–417.

686 11. Montiel, J., Downie, J.A., Farkas, A., Bihari, P., Herczeg, R., Bálint, B., et al. (2017)
687 Morphotype of bacteroids in different legumes correlates with the number and type of
688 symbiotic NCR peptides. *Proc Natl Acad Sci U S A* **114**: 5041–5046.
689 12. Farkas, A., Maróti, G., Dürgo, H., Györgypal, Z., Lima, R.M., Folkl-Medzihradsky, K.,
690 Kereszt, A., Mergaert, P., and Kondorosi, E. (2014) *Medicago truncatula* symbiotic
691 peptide NCR247 contributes to bacteroid differentiation through multiple mechanisms.
692 *Proc Natl Acad Sci U S A* **111**: 5183-5188.
693 13. Koch, M., Delmotte, N., Rehrauer, H., Vorholt, J.A., Pessi, G., and Hennecke, H. (2010)
694 Rhizobial adaptation to hosts, a new facet in the legume root-nodule symbiosis. *Mol*
695 *Plant-Microbe Interact* **23**: 784–790.
696 14. Renier, A., Maillet, F., Fardoux, J., Poinsot, V., Giraud, E., and Nouwen, N. (2011)
697 Photosynthetic *Bradyrhizobium* Sp. strain ORS285 synthesizes 2-O-methylfucosylated
698 lipochitooligosaccharides for nod gene-dependent interaction with *Aeschynomene*
699 plants. *Mol Plant Microbe Interact* **24**: 1440–7.
700 15. Ledermann, R., Bartsch, I., Müller, B., Wülser, J., and Fischer, H.M. (2018) A
701 functional general stress response of *Bradyrhizobium diazoefficiens* is required for early
702 stages of host plant infection. *Mol Plant-Microbe Interact* **31**: 537–547.
703 16. Barrière, Q., Guefrachi, I., Gully, D., Lamouche, F., Pierre, O., Fardoux, J., et al. (2017)
704 Integrated roles of BclA and DD-carboxypeptidase 1 in *Bradyrhizobium* differentiation
705 within NCR-producing and NCR-lacking root nodules. *Sci Rep* **7**: 1–13.
706 17. Lamouche, F., Gully, D., Chaumeret, A., Nouwen, N., Verly, C., Pierre, O., et al. (2019)
707 Transcriptomic dissection of *Bradyrhizobium* sp. strain ORS285 in symbiosis with
708 *Aeschynomene* spp. inducing different bacteroid morphotypes with contrasted symbiotic
709 efficiency. *Environ Microbiol.* **21**: 3244–3258.

710 18. Lamouche, F., Chaumeret, A., Guefrachi, I., Barrière, Q., Pierre, O., Guérard, F., et al.
711 (2019) From intracellular bacteria to differentiated bacteroids: transcriptome and
712 metabolome analysis in *Aeschynomene* nodules using the *Bradyrhizobium* sp. ORS285
713 *bclA* mutant. *J. Bacteriol* **201**: e00191-19.
714 19. Andrews, M., Raven, J.A., Lea, P.J., and Sprent, J.I. (2006) A role for shoot protein in
715 shoot-root dry matter allocation in higher plants. *Ann Bot* **97**: 3-10.
716 20. Collier, R., and Tegeder, M. (2012) Soybean ureide transporters play a critical role in
717 nodule development, function and nitrogen export. *Plant J*. **72**: 355-67.
718 21. Ross, E.J.H.H., Kramer, S.B., and Dalton, D.A. (1999) Effectiveness of ascorbate and
719 ascorbate peroxidase in promoting nitrogen fixation in model systems. *Phytochemistry*
720 **52**: 1203–1210.
721 22. Bashor, C.J. and Dalton, D.A. (1999) Effects of exogenous application and stem
722 infusion of ascorbate on soybean (*Glycine max*) root nodules. *New Phytol* **142**: 19–26.
723 23. Sallet, E., Gouzy, J., and Schiex, T. (2014) EuGene-PP: a next-generation automated
724 annotation pipeline for prokaryotic genomes. *Bioinformatics* **30**: 2659–61.
725 24. Bittner, A.N., Foltz, A., and Oke, V. (2007) Only one of five *groEL* genes is required
726 for viability and successful symbiosis in *Sinorhizobium meliloti*. *J Bacteriol* **189**: 1884–
727 9.
728 25. Lardi, M., Murset, V., Fischer, H. M., Mesa, S., Ahrens, C. H., Zamboni, N. et al. (2016)
729 Metabolomic Profiling of *Bradyrhizobium diazoefficiens*-Induced Root Nodules
730 Reveals Both Host Plant-Specific and Developmental Signatures. *Int J Mol Sci* **17**(6) :
731 815.
732 26. Thede, G. L., Arthur, D. C., Edwards, R. A., Buelow, D. R., Wong, J. L., Raivio, T. L.,
733 & Glover, J. N. (2011) Structure of the periplasmic stress response protein CpxP.
734 *Journal Bacteriol* **193**(9): 2149–57.

27. diCenzo, G.C., Zamani, M., Checcucci, A., Fondi, M., Griffitts, J.S., Finan, T.M., and Mengoni, A. (2018) Multi-disciplinary approaches for studying rhizobium – legume symbioses. *Peer J* cjm-2018-0377.

28. Teufel, R., Mascaraque, V., Ismail, W., Voss, M., Perera, J., Eisenreich, W., et al. (2010) Bacterial phenylalanine and phenylacetate catabolic pathway revealed. *Proc Natl Acad Sci U S A* **107**: 14390–5.

29. Okazaki, S., Nukui, N., Sugawara, M., and Minamisawa, K. (2004) Rhizobial strategies to enhance symbiotic interactions: rhizobitoxine and 1-aminocyclopropane-1-carboxylate deaminase. *Microb. Environ* **19**: 99–111.

30. Bonaldi, K., Gargani, D., Prin, Y., Fardoux, J., Gully, D., Nouwen, N., et al. (2011) Nodulation of *Aeschynomene afraspera* and *A. indica* by photosynthetic *Bradyrhizobium* sp. strain ORS285: the Nod-dependent versus the Nod-independent symbiotic interaction. *Mol Plant-Microbe Interact* **24**: 1359–1371.

31. Guefrachi, I., Pierre, O., Timchenko, T., Alunni, B., Barrière, Q., Czernic, P., et al. (2015) *Bradyrhizobium* BclA is a peptide transporter required for bacterial differentiation in symbiosis with *Aeschynomene* legumes. *Mol Plant-Microbe Interact* **28**: 1155–1166.

32. Crespo-Rivas, J.C., Guefrachi, I., Mok, K.C., Villaécija-Aguilar, J.A., Acosta-Jurado, S., Pierre, O., et al. (2016) *Sinorhizobium fredii* HH103 bacteroids are not terminally differentiated and show altered O-antigen in nodules of the Inverted Repeat-Lacking Clade legume *Glycyrrhiza uralensis*. *Environ Microbiol* **18**: 2392–2404.

33. Montiel, J., Szűcs, A., Boboescu, I.Z., Gherman, V.D., Kondorosi**,** E., and Kereszt, A. (2016) Terminal bacteroid differentiation is associated with variable morphological changes in legume species belonging to the Inverted Repeat-Lacking Clade. *Mol Plant Microbe Interact* **29**(3): 210-9.

760 34. Price, P.A., Tanner, H.R., Dillon, B.A., Shabab, M., Walker, G.C., and Griffitts, J.S.
761 (2015) Rhizobial peptidase HrrP cleaves host-encoded signaling peptides and mediates
762 symbiotic compatibility. *Proc Natl Acad Sci U S A* **112**: 15244–15249.
763 35. Tiricz, H., Szücs, A., Farkas, A., Pap, B., Lima, R.M., Maróti, G., et al. (2013)
764 Antimicrobial nodule-specific cysteine-rich peptides induce membrane depolarization-
765 associated changes in the transcriptome of *Sinorhizobium meliloti*. *Appl Environ*
766 *Microbiol* **79**: 6737–6746.
767 36. Kulkarni, G., Busset, N., Molinaro, A., Gargani, D., Chaintreuil, C., Silipo, A., et al.
768 (2015) Specific hopanoid classes differentially affect free-living and symbiotic states of
769 *Bradyrhizobium diazoefficiens*. *MBio* **6**: 1–9.
770 37. Oono, R., Schmitt, I., Sprent, J.I., and Denison, R.F. (2010) Multiple evolutionary
771 origins of legume traits leading to extreme rhizobial differentiation. *New Phytol* **187**(2):
772 508-20
773 38. Karmakar, K., Kundu, A., Rizvi, A.Z., Dubois, E., Severac, D., Czernic, P., et al. (2019)
774 Transcriptomic analysis with the progress of symbiosis in 'crack-entry' legume *Arachis*
775 *hypogaea* highlights its contrast with 'infection thread' adapted legumes. *Mol Plant-*
776 *Microbe Interact* **32**(3):271-285.
777 39. Kereszt, A., Mergaert, P., Montiel, J., Endre, G., and Kondorosi, E. (2018) Impact of
778 plant peptides on symbiotic nodule development and functioning. *Front Plant Sci*.
779 **9**:1026
780 40. Trujillo, D.I., Silverstein, K.A.T., and Young, N.D. (2019) Nodule-specific PLAT
781 domain proteins are expanded in the *Medicago* lineage and required for nodulation. *New*
782 *Phytol* **222**: 1538–1550.

783 41. Gourion, B., Sulser, S., Frunzke, J., Francez-Charlot, A., Stiefel, P., Pessi, G., et al.
784 (2009) The PhyR-σEcfG signalling cascade is involved in stress response and symbiotic
785 efficiency in *Bradyrhizobium japonicum*. *Mol Microbiol* **73**: 291–305.
786 42. Roux, B., Rodde, N., Jardinaud, M.F., Timmers, T., Sauviac, L., Cottret, L., et al. (2014)
787 An integrated analysis of plant and bacterial gene expression in symbiotic root nodules
788 using laser-capture microdissection coupled to RNA sequencing. *Plant J* **77**(6):817-37.
789 43. Sen, D. and Weaver, R.W. (1981) A comparison of nitrogen-fixing ability of peanut,
790 cowpea and siratro plants nodulated by different strains of *Rhizobium*. *Plant Soil* **60**:
791 317–319.
792 44. Oono, R. and Denison, R.F. (2010) Comparing symbiotic efficiency between swollen
793 versus nonswollen rhizobial bacteroids. *Plant Physiol* **154**: 1541–1548.
794 45. Kazmierczak, T., Nagymihaly, M., Lamouche, F., Barriere, Q., Guefrachi, I., Alunni,
795 B., et al. (2017) Specific host-responsive associations between *Medicago truncatula*
796 accessions and *Sinorhizobium* strains. Mol. Plant Microbe Interact. 30:399–409.
797 46. Lamouche, F., Bonadé-Bottino, N., Mergaert, P., and Alunni, B. (2019) Symbiotic
798 efficiency of spherical and elongated bacteroids in the *Aeschynomene-Bradyrhizobium*
799 symbiosis. *Front. Plant Sci* **10**: 377.
800 47. Regensburger, B. and Hennecke, H. (1983) RNA polymerase from *Rhizobium*
801 *japonicum*. *Arch Microbiol* **135**: 103–109.
802 48. Giraud, E., Hannibal, L., Fardoux, J., Verméglio, A., and Dreyfus, B. (2000) Effect of
803 *Bradyrhizobium* photosynthesis on stem nodulation of *Aeschynomene sensitiva*. *Proc*
804 *Natl Acad Sci U S A* **97**: 14795–14800.
805 49. Chapelle, E., Alunni B, Malfatti P, Solier L, Pedron J, Kraepiel Y, Van Gijsegem F.
806 (2015) A straightforward and reliable method for bacterial *in planta* transcriptomics:

807 application to the *Dickeya dadantii*/*Arabidopsis thaliana* pathosystem. *Plant J* **82**:352–
808 362.
809 50. Azani, N., Babineau, M., Bailey, C.D., Banks, H., Barbosa, A.R., Pinto, R.B., *et al.*
810 (2017) A new subfamily classification of the Leguminosae based on a taxonomically
811 comprehensive phylogeny – The Legume Phylogeny Working Group (LPWG). *Taxon*
812 **66**: 44–77.
813 51. Brottier, L., Chaintreuil, C., Simion, P., Scornavacca, C., Rivallan, R., Mournet, P., et
814 al. (2018) A phylogenetic framework of the legume genus *Aeschynomene* for
815 comparative genetic analysis of the Nod-dependent and Nod-independent symbioses.
816 *BMC Plant Biol* **18**: 333.
817 52. Langella O, Valot B, Jacob D, Balliau T, Flores R, Hoogland C, Joets J, Zivy M. (2013)
818 Management and dissemination of MS proteomic data with PROTICdb: example of a
819 quantitative comparison between methods of protein extraction. *Proteomics*
820 13(9):1457-66.
821 53. Kessner D, Chambers M, Burke R, Agus D, Mallick P. (2008) ProteoWizard: open
822 source software for rapid proteomics tools development. *Bioinformatics* **24**(21):2534-6.
823 54. Craig R, Beavis RC. (2004) TANDEM: matching proteins with tandem mass spectra.
824 *Bioinformatics* **20**(9):1466-7.
825 55. Langella O, Valot B, Balliau T, Blein-Nicolas M, Bonhomme L, Zivy M. (2017)
826 X!TandemPipeline: A tool to manage sequence redundancy for protein inference and
827 phosphosite identification. *J Proteome Res*. **16**(2):494-503.
828 56. Delmotte N, Mondy S, Alunni B, Fardoux J, Chaintreuil C, Vorholt JA, Giraud E,
829 Gourion B. (2014) *Int J Mol Sci*. 15(3):3660-70.
830 57. Su, F., Gilard, F., Guérard, F., Citerne, S., Clément, C., Vaillant-Gaveau, N., and
831 Dhondt-Cordelier, S. (2016) Spatio-temporal responses of *Arabidopsis* leaves in

photosynthetic performance and metabolite contents to *Burkholderia phytofirmans* PsJN. *Front Plant Sci* **7**: 1–15.

58. Médigue, C., Calteau, A., Cruveiller, S., Gachet, M., Gautreau, G., Josso, A., *et al.* (2019) MicroScope-an integrated resource for community expertise of gene functions and comparative analysis of microbial genomic and metabolic data. *Brief Bioinform* **20**: 1071-1084.
59. Ledermann, R., Bartsch, I., Remus-Emsermann, M.N., Vorholt, J.A., and Fischer, H.M. (2015) Stable fluorescent and enzymatic tagging of *Bradyrhizobium diazoefficiens* to analyze host-plant infection and colonization. *Mol Plant-Microbe Interact* **28**: 959–967.
60. Ducret, A., Quardokus, E. M., & Brun, Y. V. (2016). MicrobeJ, a tool for high throughput bacterial cell detection and quantitative analysis. *Nature Microbiology*, **1**(7), 16077.
61. Beroual, W., and Biondi, E.G. (2019) A new factor controlling cell envelope integrity in Alphaproteobacteria in the context of cell cycle, stress response and infection. *Mol Microbiol*. **111**(3):553-555.

Figure legends

Figure 1. The non-adapted symbiotic couple formed by *Bradyrhizobium diazoefficiens* USDA110 and the NCR-producing plant *Aeschynomene afraspera* displays suboptimal nitrogen fixation and nodule metabolic dysfunction. A. Phylogenetic ML tree of a selection of plant species based on *matK* nucleotide sequences. Red branches indicate clades of legumes plants inducing terminal bacteroid differentiation. Blue boxes indicate the distantly-related host plants used in this study. Bootstrap values are mentioned in green on each node of the tree. B,C. General aspect of the plants and nodule sections (inlays) displaying the red coloration of

leghaemoglobin of *G. max* (B) and *A. afraspera* (C) at 14 dpi. Scale bars: 5 cm (plants) and 1 mm (nodules). D, E. Nitrogen fixation activity determined by acetylene reduction assay (D) and gain in biomass attributable to the symbiosis (E) of 14 dpi plants. F. Whole-nodule metabolome determined by GC/MS or LC/MS at 14 dpi. Histograms show the average value of the relative metabolite concentration of four biological replicates. Letters represent significant differences after ANOVA and post hoc Tukey tests ($p < 0.05$). GM: *G. max* bacteroids, AA: *A. afraspera* bacteroids, USDA: *B. diazoefficiens* USDA110, ORS: *Bradyrhizobium sp*. ORS285.

Figure 2. **Experimental setup and general description of the trancriptomics and proteomics dataset.** A. Experimental setup displaying the three biological conditions of this study. B. Principal component analysis of the proteomics and transcriptomics datasets. C. Venn diagram representing the overlap between differentially expressed genes (DEGs, FDR < 0.01 & |LFC| > 1.58) and differentially accumulated proteins (DAPs, FDR < 0.05) in at least one comparison and among the population of detected proteins. D. Pearson correlation coefficient (r) distribution between transcriptomic and proteomic datasets based on DAPs only (red) or DAPs that are also DEGs (green). E. Heatmaps and hierarchical clustering of the 815 DAPs and the corresponding transcriptomic expression values. The heatmaps show the standard score (z-score) of assigned spectra and DESeq2 normalized read counts, respectively. The color-coded scale bars for the normalized expression and value of Pearson correlation coefficient of the genes are indicated below the heatmap. YM: Yeast-Mannitol culture, GM: *G. max* bacteroids, AA: *A. afraspera* bacteroids.

Figure 3: Symbiosis and host-specific functions that display congruency between transcriptomics and proteomics A. Heatmap with SOM clustering displaying bacterial functions that are commonly DAP and DEG *in planta* in both host plants as compared to the

882 culture reference. B. Heatmap displaying bacterial functions that are commonly DEG and DAP
883 in one host as compared to the other (upper panel: *A. afraspera* > *G. max*; lower panel: *G. max*
884 > *A. afraspera*). In panels A-B, data are presented as log 2 of DESeq2 normalized read counts
885 (RNA-seq) or spectral counting (Proteomics). YM: Yeast-Mannitol culture, GM: *G. max*
886 bacteroids, AA: *A. afraspera* bacteroids.
887
888 **Figure 4: Expression pattern of *B. diazoefficicens* USDA110 and *Bradyhrizobium* sp.**
889 **ORS285 orthologous genes *in planta* and in culture.** A. Heatmap after SOM clustering of all
890 the orthologous genes of USDA110 and ORS285 obtained with Phyloprofile. Values present
891 the *in planta* LFC calculated after the read counts of the culture control versus *A. afraspera* 14
892 dpi nodules. B. Heatmaps of the orthologous genes after filtering on the FDR (< 0.01) values.
893 Selected genes are highlighted for each class of interest. C. Dot plot of the orthologous genes
894 that are DEG (FDR < 0.01 and |LFC| > 1.58) *in planta* (ie. in *A afraspera* nodules) in both
895 strains. The red dashed line is for the linear regression and the blue envelope shows a 0.95
896 confidence interval of the linear regression.
897
898 **Figure 5. *B. diazoefficicens* USDA110 displays atypical bacteroid differentiation features**
899 **in *A. afraspera* nodules**. A. Average cell shape of free-living bacteria and bacteroids
900 determined by MicrobeJ (900 < n < 21 000). B. DNA content of USDA110 bacteroids extracted
901 from soybean and *A. afraspera* determined by flow cytometry. C. Assessment of the
902 permeability of USDA110 and ORS285 free-living cells and bacteroids 20 min after PI
903 treatment. *: wilcoxon test, p-value < 0.01. Five biological replicates were performed for
904 bacteroids and two for free-living bacteria. D. Viability of soybean and *A. afraspera* extracted
905 bacteroids at 14 dpi. Asterisks point out significant differences according to a wilcoxon test. *:
906 p-value < 0.05; **: p-value < 0.01. Data are representative of 10 independent plants.

Supplementary figure and table legends

Figure S1. Nitrogen and carbon content in aerial parts of the plants were determined by elemental analysis. GM: *G. max*, AA: *A. afraspera*, ORS: inoculated by *Bradyrhizobium* sp. ORS285, USDA: inoculated by *B. diazoefficiens* USDA110, NI: Non-inoculated plants.

Figure S2. Nutritional status of 14 dpi plants determined by the shoot/root mass ratios. AA: *A. afraspera*, ORS: inoculated by *Bradyrhizobium* sp. ORS285, USDA: inoculated by *B. diazoefficiens* USDA110, NI: Non-inoculated plants. Letters represent significant differences after t-test or ANOVA and post hoc Tukey tests ($p < 0.05$).

Figure S3. Overview of the 129 quantified metabolites in *G. max* and *A. afraspera* whole nodules elicited by *B. diazoefficiens* USDA110 or *Bradyrhizobium* sp. ORS285. Heatmap and hierarchical clustering of the 129 metabolites that were quantified either by gas- (GC-MS) or liquid-chromatography (LC-MS) coupled to mass spectrometry. Gm: *G. max*, Aa: *A. afraspera*, O: inoculated by *Bradyrhizobium* sp. ORS285, U: inoculated by *B. diazoefficiens* USDA110.

Figure S4. General overview of the datasets using COG classification. Repartitions of the assigned spectra (left panel) and normalized reads (right panel) among COG classes in the three conditions (blue: bacterial culture, ocher: *B. diazoefficiens* USDA110 in *G. max* nodules, green: *B. diazoefficiens* USDA110 in *A. afraspera* nodules).

Figure S5. Western blot analysis of selected USDA110 proteins in culture and in bacteroids. NifH protein were analyzed by western blots on purified USDA110 bacteroids extracted from soybean and *A. afraspera* nodules 14 dpi. Exponential and stationary phase cultures were used as controls.

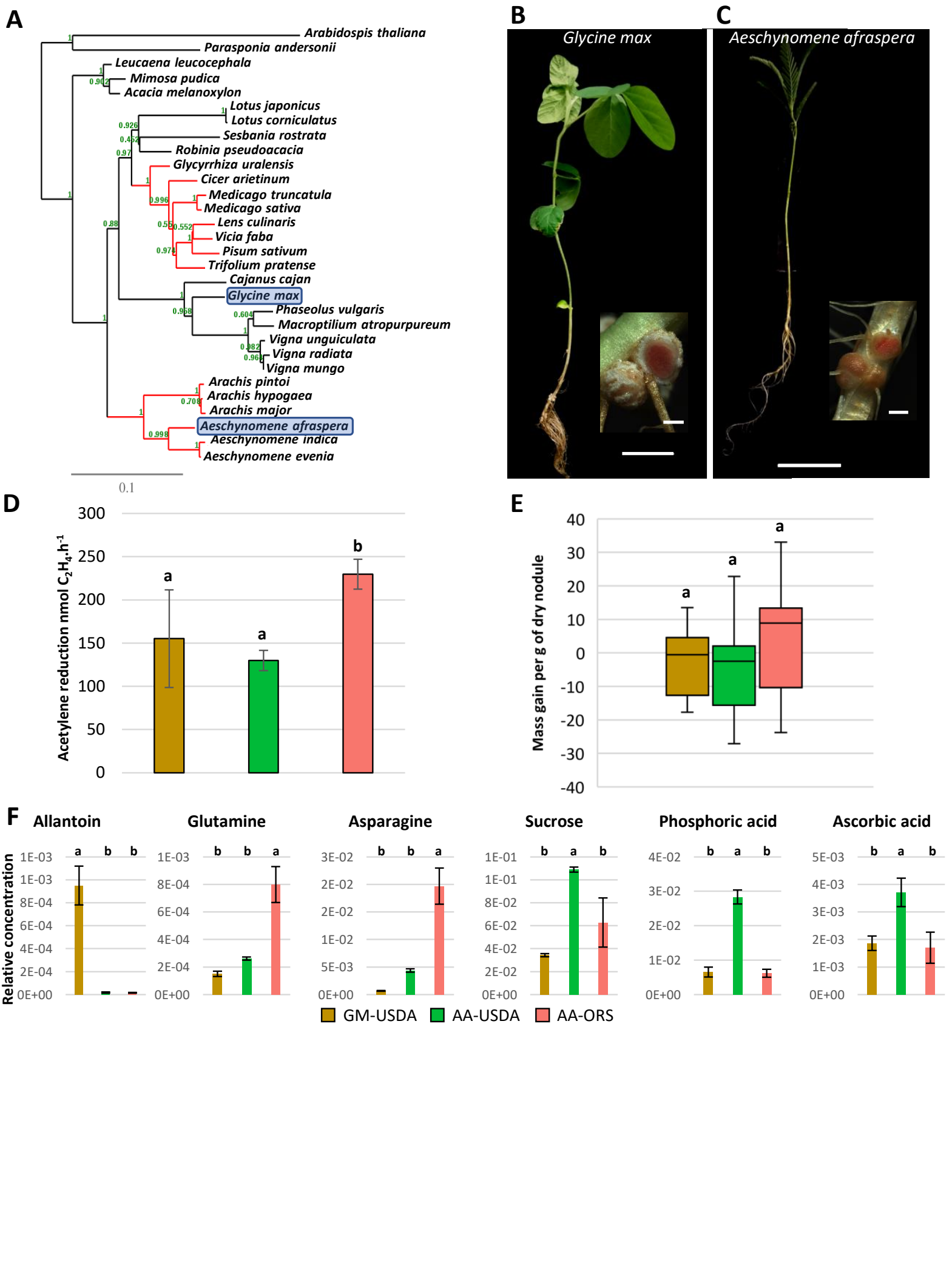
Figure S6. Analysis of cellular differentiation using automated morphometry. A, B, C & D. Parameters were quantified by image analysis of syto9 stained bacteria and bacteroids using MicrobeJ. The process from raw images (A), segmentation (B), object detection (C) and measurements (D) is depicted with these four panels. E. Cell area. F. Cell width. G. Cell length.

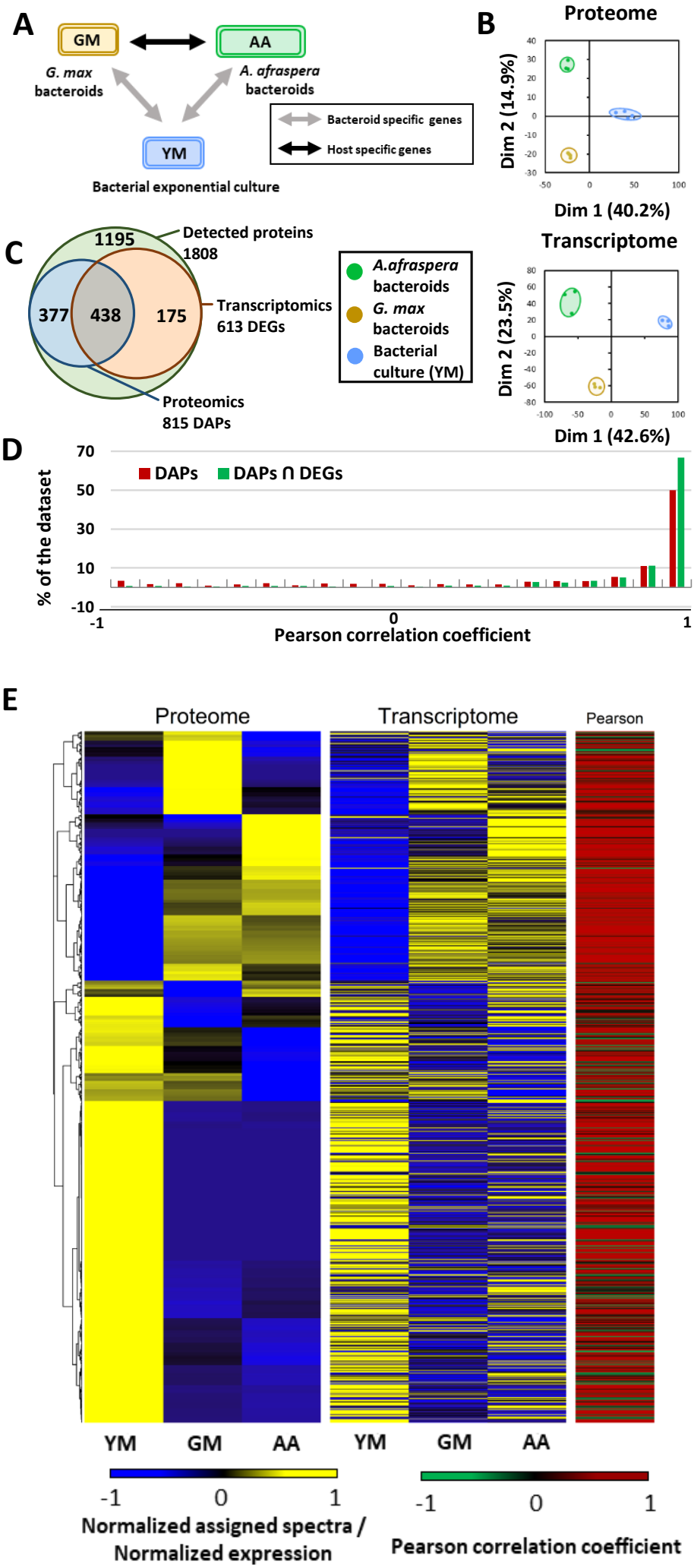
Figure S7. Kinetic analysis of bacterial membrane permeability. Kinetics of propidium iodide uptake assays (reflecting membrane permeability) from which data presented in Figure 5C were extracted. The PI permeability was measured by flow cytometry over 60 min after treatment on *A. afraspera* nodule extracted USDA110 (AaU) or ORS285 (AaO) bacteroids and *G. max* extracted USDA110 bacteroids at 14 dpi (GmU). Exponential phase bacterial culture of USDA110 and ORS285 where used as controls. Each dot represent three independent measures and error bars represent the standard deviation of the samples.

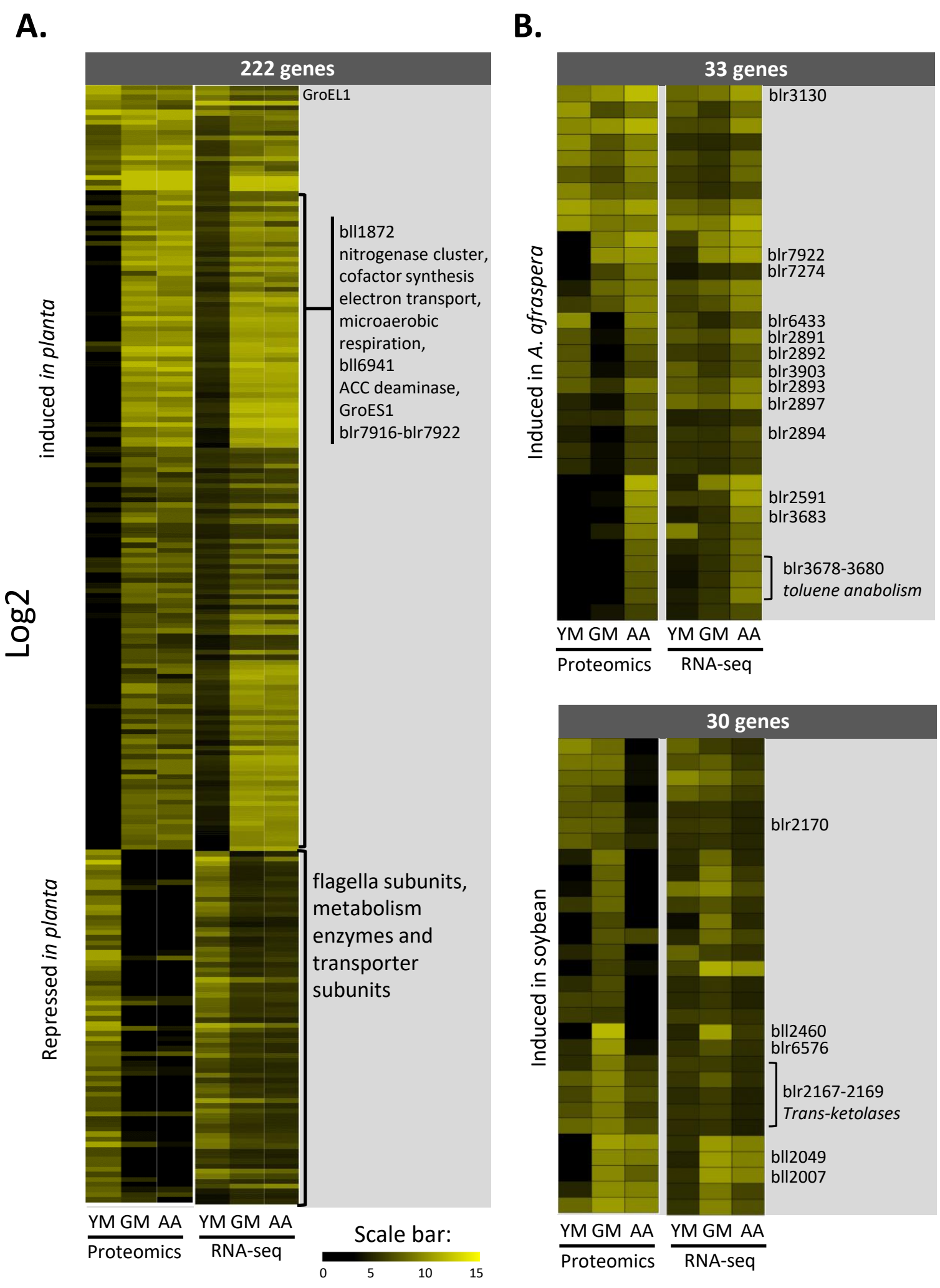
Table S1. Genome annotation, transcriptomic and proteomic data of *B. diazoefficiens* USDA110 generated in this study. Description of proteomic and transcriptomic data of USDA110 related conditions. DESeq2 normalized reads, false discovery rate (FDR) values as well as log2 fold change (LFC) are used to describe transcriptomic data. On the other side, spectral counting (SC) along with related statistical indicators, Tukey statistical test result and p-value depict the proteomic data.

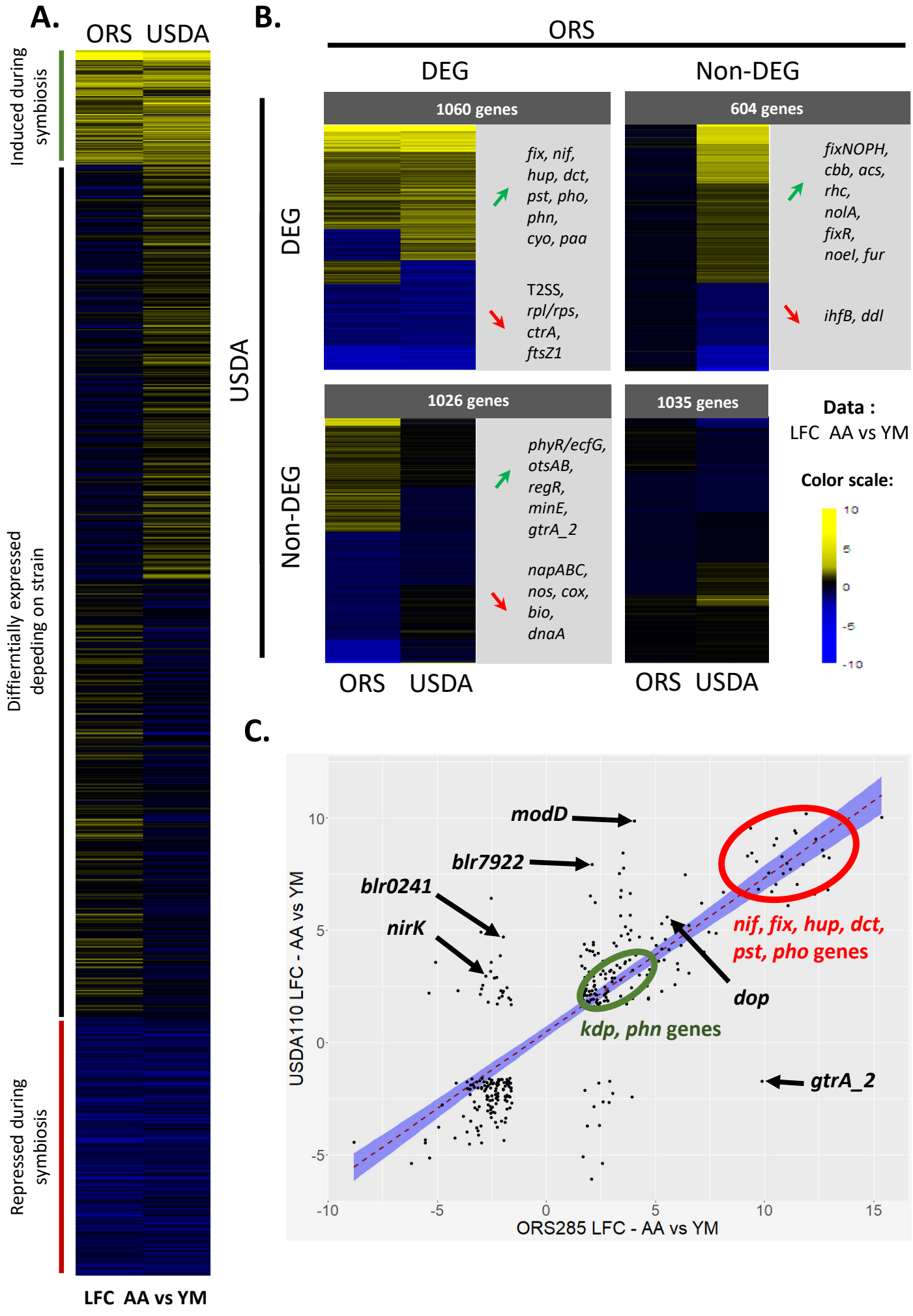
Table S2. Expression analysis of selected *B. diazoefficiens* USDA110 regulons and stimulons. Detailed analysis of the previously determined regulons and stimulons of USDA110 based on our transcriptomic data. A given regulon/stimulon was considered differentially regulated when ≥ 40% of the corresponding genes were differentially expressed in our conditions.

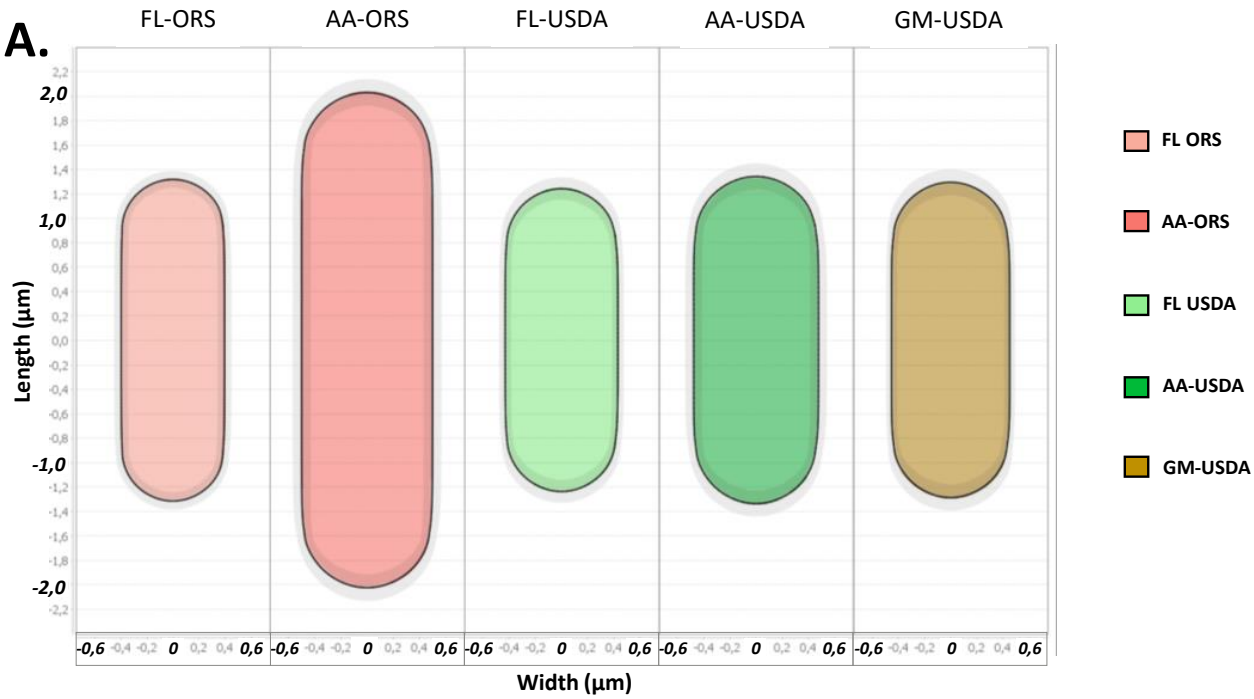
Table S3. List of the 3725 orthologous genes shared by *B. diazoefficiens* USDA110 and *Bradyrhizobium* sp. ORS285 with their corresponding expression level in rich medium and in *A. afraspera* nodules. This dataset was obtained after a Phyloprofile analysis on Mage Microscope website and was used to generate the Figure 4. Normalized read counts are shown together with the corresponding LFC and FDR as determined by DESeq2.

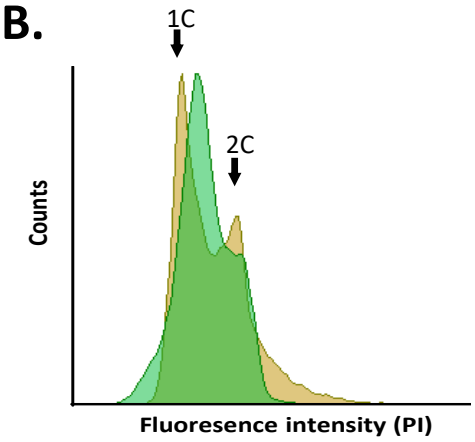

A
Arabidopsis thaliana
Parasponia andersonii
Leucaena leucocephala
Mimosa pudica
Acacia melanoxylon
Lotus japonicus
Lotus corniculatus
Sesbania rostrata
Robinia pseudoacacia
Glycyrrhiza uralensis
Cicer arietinum
Medicago truncatula
Medicago sativa
Lens culinaris
Vicia faba
Pisum sativum
Trifolium pratense
Cajanus cajan
Glycine max
Phaseolus vulgaris
Macroptilium atropurpureum
Vigna unguiculata
Vigna radiata
Vigna mungo
Arachis pintoi
Arachis hypogaea
Arachis major
Aeschynomene afraspera
Aeschynomene indica
Aeschynomene evenia
0.1
B
Glycine max
C
Aeschynomene afraspera
D
Acetylene reduction nmol $C_2H_4.h^{-1}$
300
250
200
150
100
50
0
a
a
b
E
Mass gain per g of dry nodule
40
30
20
10
0
-10
-20
-30
-40
a
a
a
F
Allantoin
Glutamine
Asparagine
Sucrose
Phosphoric acid
Ascorbic acid
Relative concentration
GM-USDA
AA-USDA
AA-ORS

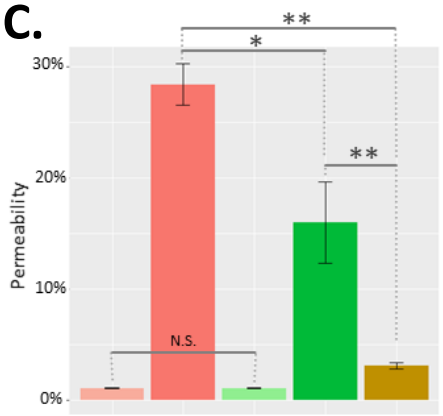

A
GM
AA
G. max bacteroids
A. afraspera bacteroids
YM
Bacteroid specific genes
Host specific genes
Bacterial exponential culture
B
Proteome
Dim 2 (14.9%)
Dim 1 (40.2%)
C
1195
1808
377
438
175
Detected proteins
Transcriptomics 613 DEGs
Proteomics 815 DAPs
A.afraspera bacteroids
G. max bacteroids
Bacterial culture (YM)
Transcriptome
Dim 2 (23.5%)
Dim 1 (42.6%)
D
DAPs
DAPs ∩ DEGs
% of the dataset
Pearson correlation coefficient
E
Proteome
Transcriptome
Pearson
YM
GM
AA
Normalized assigned spectra / Normalized expression
Pearson correlation coefficient

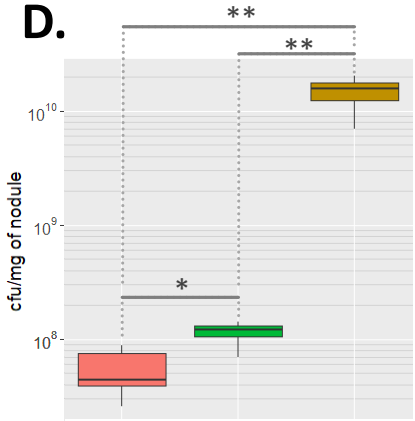

A.
222 genes
GroEL1
bll1872
nitrogenase cluster,
cofactor synthesis
electron transport,
microaerobic
respiration,
bll6941
ACC deaminase,
GroES1
blr7916-blr7922
induced in planta
Repressed in planta
Log2
flagella subunits,
metabolism
enzymes and
transporter
subunits
YM GM AA
Proteomics
YM GM AA
RNA-seq
Scale bar:
0 5 10 15
B.
33 genes
Induced in A. afraspera
blr3130
blr7922
blr7274
blr6433
blr2891
blr2892
blr3903
blr2893
blr2897
blr2894
blr2591
blr3683
blr3678-3680
toluene anabolism
YM GM AA
Proteomics
YM GM AA
RNA-seq
30 genes
Induced in soybean
blr2170
bll2460
blr6576
blr2167-2169
Trans-ketolases
bll2049
bll2007
YM GM AA
Proteomics
YM GM AA
RNA-seq


A.
ORS USDA
Induced during symbiosis
Diffierntially expressed depeding on strain
Repressed during symbiosis
LFC AA vs YM
B.
ORS
DEG
Non-DEG
USDA
DEG
Non-DEG
1060 genes
fix, nif, hup, dct, pst, pho, phn, cyo, paa
T2SS, rpl/rps, ctrA, ftsZ1
604 genes
fixNOPH, cbb, acs, rhc, nolA, fixR, noeI, fur
ihfB, ddl
1026 genes
phyR/ecfG, otsAB, regR, minE, gtrA_2
napABC, nos, cox, bio, dnaA
1035 genes
ORS USDA
ORS USDA
Data :
LFC AA vs YM
Color scale:
10
5
0
-5
-10
C.
modD
blr7922
blr0241
nirK
nif, fix, hup, dct, pst, pho genes
dop
kdp, phn genes
gtrA_2
USDA110 LFC - AA vs YM
ORS285 LFC - AA vs YM
10
5
0
-5
-10
-5
0
5
10
15


A.
FL-ORS
AA-ORS
FL-USDA
AA-USDA
GM-USDA
Length (µm)
Width (µm)
FL ORS
AA-ORS
FL USDA
AA-USDA
GM-USDA


B.
1C
2C
Counts
Fluoresence intensity (PI)


C.
Permeability
30%
20%
10%
0%
N.S.
**
*
**


D.
cfu/mg of nodule
**
**
*

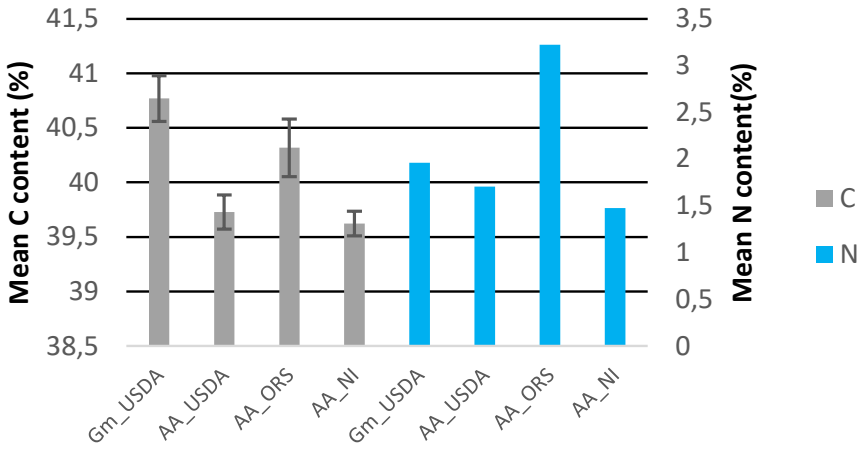

41,5
41
40,5
40
39,5
39
38,5
Mean C content (%)
3,5
3
2,5
2
1,5
1
0,5
0
Mean N content(%)
C
N
Gm_USDA
AA_USDA
AA_ORS
AA_NI
Gm_USDA
AA_USDA
AA_ORS
AA_NI

Figure S1

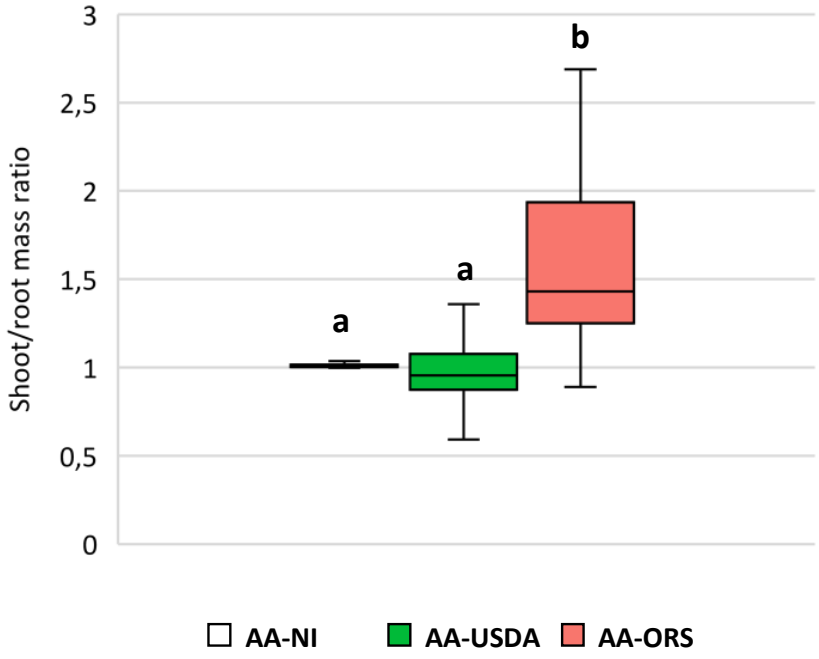
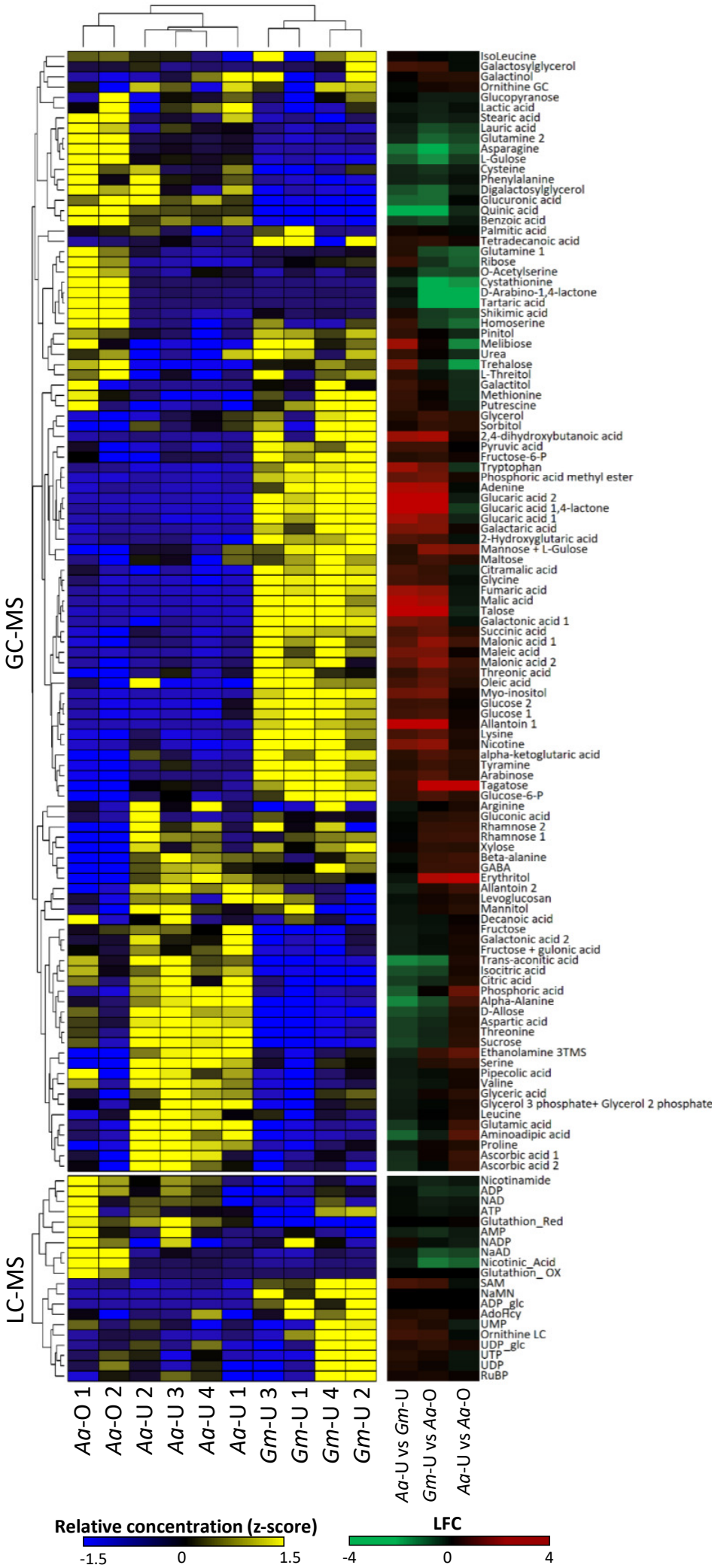


Figure S2

Figure S3


GC-MS
LC-MS
IsoLeucine
Galactosylglycerol
Galactinol
Ornithine GC
Glucopyranose
Lactic acid
Stearic acid
Lauric acid
Glutamine 2
Asparagine
L-Gulose
Cysteine
Phenylalanine
Digalactosylglycerol
Glucuronic acid
Quinic acid
Benzoic acid
Palmitic acid
Tetradecanoic acid
Glutamine 1
Ribose
O-Acetylserine
Cystathionine
D-Arabino-1,4-lactone
Tartaric acid
Shikimic acid
Homoserine
Pinitol
Melibiose
Urea
Trehalose
L-Threitol
Galactitol
Methionine
Putrescine
Glycerol
Sorbitol
2,4-dihydroxybutanoic acid
Pyruvic acid
Fructose-6-P
Tryptophan
Phosphoric acid methyl ester
Adenine
Glucaric acid 2
Glucaric acid 1,4-lactone
Glucaric acid 1
Galactaric acid
2-Hydroxyglutaric acid
Mannose + L-Gulose
Maltose
Citramalic acid
Glycine
Fumaric acid
Malic acid
Talose
Galactonic acid 1
Succinic acid
Malonic acid 1
Maleic acid
Malonic acid 2
Threonic acid
Oleic acid
Myo-inositol
Glucose 2
Glucose 1
Allantoin 1
Lysine
Nicotine
alpha-ketoglutaric acid
Tyramine
Arabinose
Tagatose
Glucose-6-P
Arginine
Gluconic acid
Rhamnose 2
Rhamnose 1
Xylose
Beta-alanine
GABA
Erythritol
Allantoin 2
Levoglucosan
Mannitol
Decanoic acid
Fructose
Galactonic acid 2
Fructose + gulonic acid
Trans-aconitic acid
Isocitric acid
Citric acid
Phosphoric acid
Alpha-Alanine
D-Allose
Aspartic acid
Threonine
Sucrose
Ethanolamine 3TMS
Serine
Pipecolic acid
Valine
Glyceric acid
Glycerol 3 phosphate+ Glycerol 2 phosphate
Leucine
Glutamic acid
Aminoadipic acid
Proline
Ascorbic acid 1
Ascorbic acid 2
Nicotinamide
ADP
NAD
ATP
Glutathion_Red
AMP
NADP
NaAD
Nicotinic_Acid
Glutathion_ OX
SAM
NaMN
ADP_glc
AdoHcy
UMP
Ornithine LC
UDP_glc
UTP
UDP
RuBP
Aa-O 1
Aa-O 2
Aa-U 2
Aa-U 3
Aa-U 4
Aa-U 1
Gm-U 3
Gm-U 1
Gm-U 4
Gm-U 2
Aa-U vs Gm-U
Gm-U vs Aa-O
Aa-U vs Aa-O
Relative concentration (z-score)
-1.5
0
1.5
LFC
-4
0
4

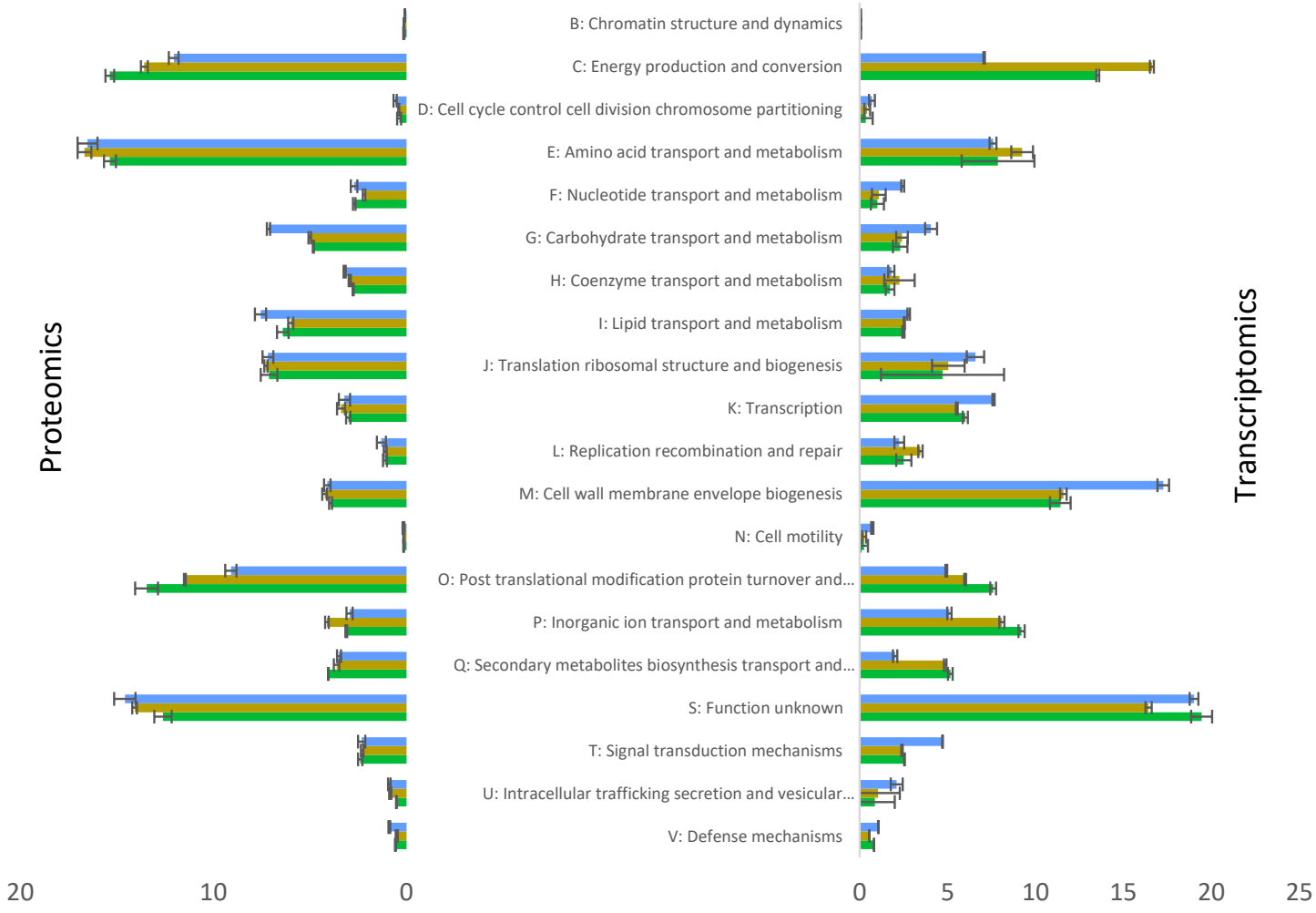

Proteomics
Transcriptomics
B: Chromatin structure and dynamics
C: Energy production and conversion
D: Cell cycle control cell division chromosome partitioning
E: Amino acid transport and metabolism
F: Nucleotide transport and metabolism
G: Carbohydrate transport and metabolism
H: Coenzyme transport and metabolism
I: Lipid transport and metabolism
J: Translation ribosomal structure and biogenesis
K: Transcription
L: Replication recombination and repair
M: Cell wall membrane envelope biogenesis
N: Cell motility
O: Post translational modification protein turnover and...
P: Inorganic ion transport and metabolism
Q: Secondary metabolites biosynthesis transport and...
S: Function unknown
T: Signal transduction mechanisms
U: Intracellular trafficking secretion and vesicular...
V: Defense mechanisms
20
10
0
0
5
10
15
20
25

Figure S4

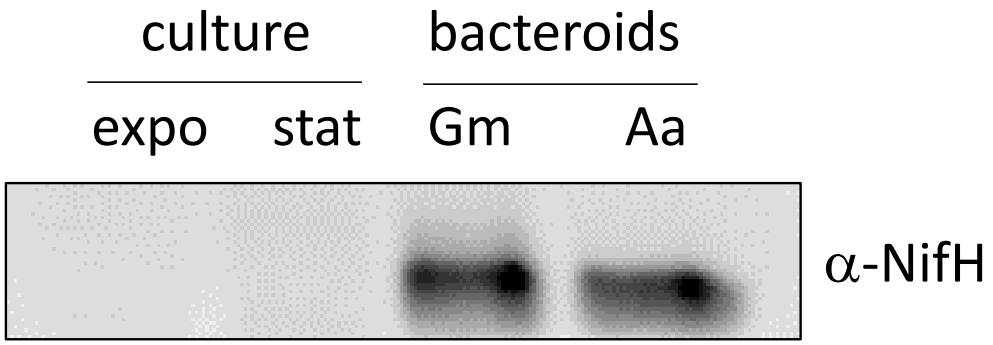


Figure S5

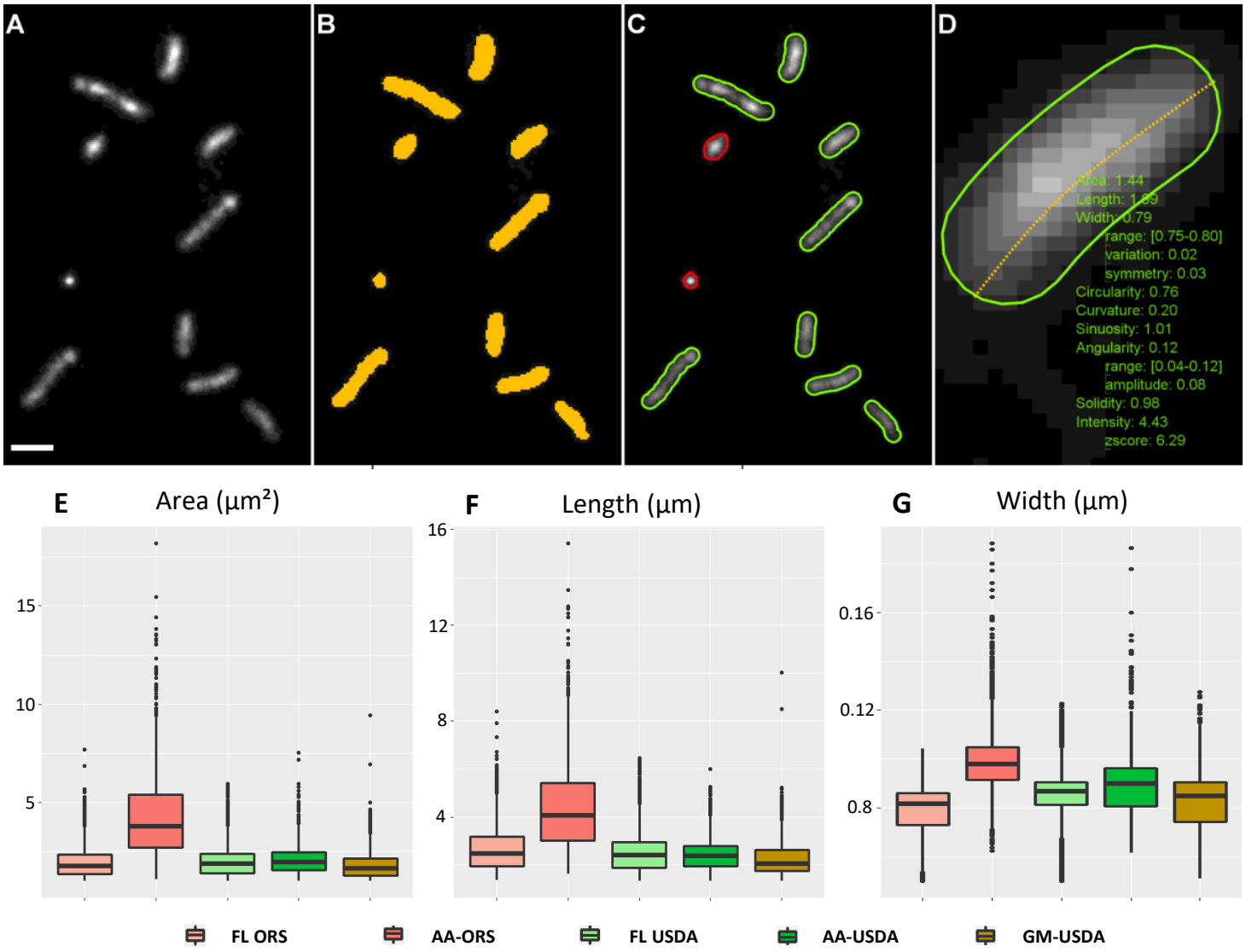

A
B
C
D
Area: 1.44
Length: 1.89
Width: 0.79
range: [0.75-0.80]
variation: 0.02
symmetry: 0.03
Circularity: 0.76
Curvature: 0.20
Sinuosity: 1.01
Angularity: 0.12
range: [0.04-0.12]
amplitude: 0.08
Solidity: 0.98
Intensity: 4.43
zscore: 6.29
E
Area (µm²)
15
10
5
F
Length (µm)
16
12
8
4
G
Width (µm)
0.16
0.12
0.8
FL ORS
AA-ORS
FL USDA
AA-USDA
GM-USDA

Figure S6

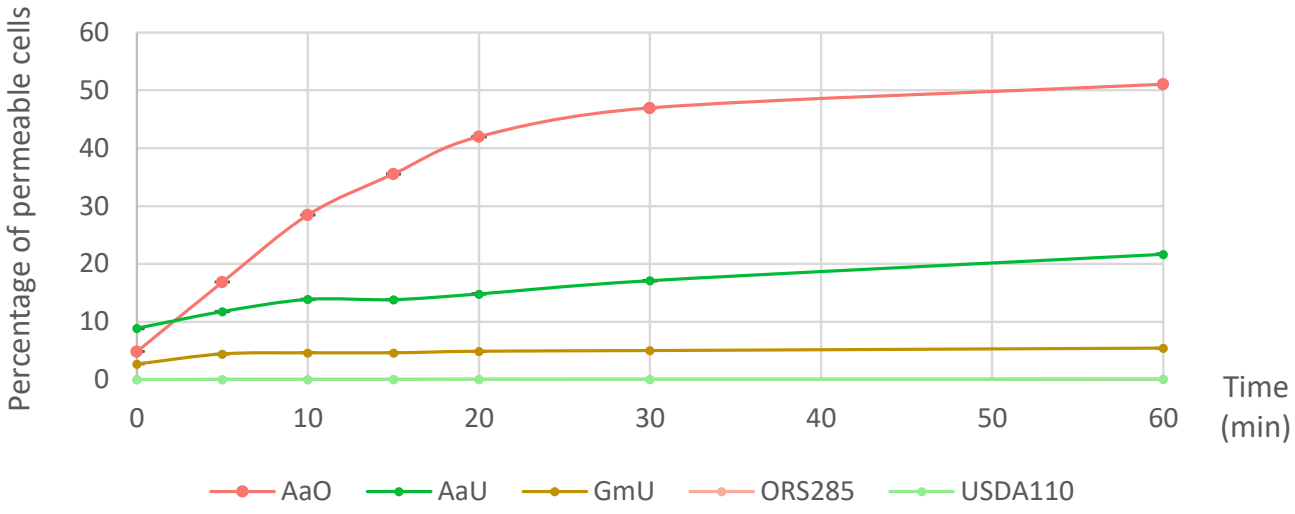

Percentage of permeable cells
60
50
40
30
20
10
0
0
10
20
30
40
50
60
Time (min)
AaO
AaU
GmU
ORS285
USDA110

Figure S7