Transcriptomic dissection of *Bradyrhizobium* sp. strain ORS285 in symbiosis with *Aeschynomene* spp. inducing different bacteroid morphotypes with contrasted symbiotic efficiency

Florian Lamouche[a‡], Djamel Gully[b‡], Anaïs Chaumeret[a+], Nico Nouwen[b], Camille Verly[a+], Olivier Pierre[a+], Coline Sciallano[b], Joël Fardoux[b], Christian Jeudy[c], Attila Szücs[d], Samuel Mondy[a+], Christophe Salon[c], István Nagy[d,e], Attila Kereszt[d,e], Yves Dessaux[a], Eric Giraud[b], Peter Mergaert[a,*], Benoit Alunni[a,*]

[a]Institute for Integrative Biology of the Cell, UMR 9198, CNRS/Université Paris-Sud/CEA, 91198 Gif-sur-Yvette, France. [b]Laboratoire des Symbioses Tropicales et Méditerranéennes, Institut pour la Recherche et le Développement, UMR IRD/SupAgro/INRA/UM2/CIRAD, Campus International de Baillarguet, TA A-82/J, 34398 Montpellier, France. [c]Agroécologie, AgroSup Dijon, INRA, Université Bourgogne Franche-Comté, 21065 Dijon, France. [d]Biological Research Centre, Hungarian Academy of Sciences, 6726 Szeged, Hungary. [e]Seqomics Biotechnology Ltd., 6782 Mórahalom, Hungary.

[+]present address: Laboratoire de Reproduction et Développement des plantes, Lyon, France (AC); Institut Jean-Pierre Bourgin, Versailles, France (CV); Institut Sophia Agrobiotech, Sophia Antipolis, France (OP); Agroécologie, AgroSup Dijon, INRA, Université Bourgogne Franche-Comté, Dijon, France (SM); Department of Medical Biology. Szeged University (AS).

*Correspondance: benoit.alunni@i2bc.paris-saclay.fr; Tel +33 1 69 82 34 92

peter.mergaert@i2bc.paris-saclay.fr; Tel +33 1 69 82 34 81

[‡]equal contribution

This article has been accepted for publication and undergone full peer review but has not been through the copyediting, typesetting, pagination and proofreading process which may lead to differences between this version and the Version of Record. Please cite this article as an 'Accepted Article', doi: 10.1111/1462-2920.14292

Significance statement

The legume-rhizobium symbiosis is of major ecological importance in agriculture and more broadly in the global nitrogen cycle. The symbiosis takes place in root nodules where large rhizobium populations are hosted to fix sufficient nitrogen to sustain plant growth. The rhizobium bacteria inside the nodules are differentiated into their nitrogen-fixing form called bacteroids. In recent years, the fate of the bacteroids has drawn lots of attention in the research on this symbiosis and this has revealed that some legume hosts can induce a terminal differentiation in their microsymbionts resulting in polyploid and strongly enlarged bacteria that become elongated or spherical. This peculiar process relies on the massive production of nodule-specific antimicrobial peptides that are targeted to the bacteroids. Phylogenetic evidence suggest that terminal differentiation improves the symbiotic benefits to the host. We show here that the level of differentiation indeed correlates with the efficiency of the symbiosis as highly polyploid spherical bacteroids are more efficient than the low polyploidy elongated ones. To reach this conclusion, we used a single bacterial strain that nodulates two closely related plant species that induce elongated and spherical bacteria, respectively. Further investigations on the transcriptomes of these two differentiated bacteroid types revealed differences in their metabolism as well as in their stress status. As a whole, this study provides insights on the molecular basis of this increased symbiotic efficiency that accompanies bacteroid differentiation and more generally on the molecular consequences of cellular differentiation on bacterial metabolism.

Summary

To circumvent the paucity of nitrogen sources in the soil legume plants establish a symbiotic interaction with nitrogen-fixing soil bacteria called rhizobia. During symbiosis, the plants form root organs called nodules, where bacteria are housed intracellularly and become active nitrogen fixers known as bacteroids. Depending on their host plant, bacteroids can adopt different morphotypes, being either unmodified (U), elongated (E) or spherical (S). E- and S-type bacteroids undergo a terminal differentiation leading to irreversible morphological changes and DNA endoreduplication. Previous studies suggest that differentiated bacteroids display an increased symbiotic efficiency (E>U and S>U). In this study, we used a combination of *Aeschynomene* species inducing E- or S-type bacteroids in symbiosis with *Bradyrhizobium* sp. ORS285 to show that S-type bacteroids present a better symbiotic efficiency than E-type bacteroids. We performed a transcriptomic analysis on E- and S-type bacteroids formed by *Aeschynomene afraspera* and *Aeschynomene indica* nodules and identified the bacterial functions activated in bacteroids and specific to each bacteroid type. Extending the expression analysis in E- and S-type bacteroids in other *Aeschynomene* species by qRT-PCR on selected genes from the transcriptome analysis narrowed down the set of bacteroid morphotype-specific genes. Functional analysis of a selected subset of 31 bacteroid-induced or morphotype-specific genes revealed no symbiotic phenotypes in the mutants. This highlights the robustness of the symbiotic program but could also indicate that the bacterial response to the plant environment is partially anticipatory or even maladaptive. Our analysis confirms the correlation between differentiation and efficiency of the bacteroids and provides a framework for the identification of bacterial functions that affect the efficiency of bacteroids.

Keywords: *Aeschynomene*; *Bradyrhizobium*; functional analysis; legume nodule; nitrogen

fixation; symbiotic efficiency; terminal bacteroid differentiation; transcriptome

Introduction

Legume plants circumvent the paucity of soil nitrogen sources by developing symbiotic organs called nodules in which bacteria collectively called rhizobia fix atmospheric nitrogen and transfer ammonia to the plant (Udvardi and Poole, 2013). During nodule formation, rhizobia are released inside the host plant cells by an endocytosis-like process and transform into nitrogen-fixing bacteroids. In several legume lineages, a terminal bacteroid differentiation (TBD) process occurs after their intracellular uptake, producing elongated polyploid bacterial cells that switched their cell cycle towards endoreduplication (Mergaert *et al.*, 2006; Kondorosi *et al.*, 2013; Alunni and Gourion, 2016). TBD is controlled by the host plant, which produces defensin-like peptides called Nodule-specific Cysteine Rich (NCR) peptides (Mergaert *et al.*, 2003; Van de Velde *et al.*, 2010; Czernic *et al.*, 2015). NCR peptides are encoded by large gene families. *Medicago truncatula* for example has more than 600 NCR genes (Mergaert *et al.*, 2003; Montiel *et al.*, 2017). Moreover, the numbers and types of NCR peptides produced by other Inverted-Repeat Lacking Clade (IRLC) legumes correlate with the degree of bacteroid elongation (Montiel *et al.*, 2017). In the *Aeschynomene* genus, belonging to the distinct legume clade of the Dalbergioids, the TBD process is also driven by a family of several dozens of NCR peptides, which are however unrelated to the NCR peptides of IRLC and thus arose by a convergent evolution process (Czernic *et al.*, 2015).

Legumes of the *Aeschynomene* genus are nodulated by members of the *Bradyrhizobium* genus such as the photosynthetic *Bradyrhizobium* sp. strain ORS285. Depending on the host, bacteroids exhibit different morphotypes, being either elongated (E-type) as in *Aeschynomene afraspera* or *Aeschynomene nilotica*, where bacteroid ploidy raises to 8C (1C being one haploid genome complement), or spherical (S-type) as in *Aeschynomene indica* or *Aeschynomene evenia* where bacteroid ploidy raises to higher levels, up to 16C (Czernic *et*

al., 2015; Guefrachi *et al.*, 2015). The spherical bacteroid differentiation in *A. indica* comprises a transient cell elongation step (Czernic *et al.*, 2015), suggesting that E- and S-type bacteroids can be regarded as a differentiation gradient from less differentiated E-type to more differentiated S-type bacteroids.

The TBD process is not the default state of bacteroids in legumes and several plants do not produce NCR peptides and hence do not induce TBD. For example, soybean induces unmodified bacteroids (U-type) that remain similar to free-living bacteria. Ancestral state reconstruction indicates that TBD is a derived state that evolved independently in different legume clades (Oono *et al.*, 2011) in agreement with the independent acquisition of NCR genes in the IRLC and *Aeschynomene*. This convergence toward the evolution of TBD raises questions regarding the selective advantage provided by this process. The biological cost of TBD seems high, as hundreds of peptides are massively produced to induce bacterial differentiation. Studies comparing symbiotic performance of U- *vs.* E-type bacteroids and U- *vs.* S-type bacteroids showed that both types of TBD (E- or S-type nodules) confer a higher symbiotic efficiency than that observed in the absence of TBD (U-type bacteroids) (Sen and Weaver, 1984; Oono and Denison, 2010). These results suggest that TBD increases the efficiency of bacteroids. On the other hand, a comparison of the symbiotic efficiency of E- *vs.* S-type bacteroids has not been performed. Since S-type bacteroids are an enhanced form of differentiation, it is possible that those bacteroids are more efficient than the less differentiated E-type. Moreover, the molecular basis of the increased symbiotic efficiency in E- and S-type bacteroids remains unknown.

In this study, we compared the symbiotic efficiency of E- *vs.* S-type bacteroids formed by *Aeschynomene* spp. and found an increased symbiotic performance for S-type bacteroids. To identify the underlying basis, a transcriptomic study was conducted. Bacterial functions specifically activated in E- or S-type bacteroids highlighted important metabolic differences

between the two bacteroid types. A functional analysis of a selected set of genes identified by transcriptomics, suggested that the bacterial symbiotic program shows robustness and redundancy. This study provides novel insights on the metabolism of differentiated bacteroids and shows to which extent E- and S-morphotype bacteroids behave differently.

Results and discussion

S-morphotype bacteroids have an increased symbiotic efficiency compared to E-morphotype bacteroids

To analyze symbiotic efficiency of E-morphotype and S-morphotype bacteroids, we first determined the plant biomass gained by the symbiosis and per investment in the symbiosis, estimated by the total mass of nodules (see Supporting Information Appendix S1). We find that the S-morphotype inducing plants *A. indica* and *A. evenia*, infected with strain ORS285 obtain a higher benefit than the E-morphotype inducing plants *A. afraspera* and *A. nilotica*, infected with the same strain ORS285 (Fig. 1A). In order to further and more specifically compare the gain in the N-content of the total plant biomass through symbiosis, we analyzed N and C elemental composition of *A. afraspera* and *A. indica* plants infected or not by ORS285 strain (see Supporting Information Appendix S1). In agreement with the previous result, the S-morphotype inducing plant *A. indica* obtained a higher amount of nitrogen per nodule mass than the E-morphotype inducing plant *A. afraspera* and this was paralleled with a higher increase in carbon as well (Fig. 1B), confirming the different performance of the two bacteroid types. This elemental composition analysis determines the N and C accumulation during the whole lifespan of the plant. Therefore, we analyzed also the N and C fluxes in a 24-hour time window at 14dpi, a stage of plant growth when nodules are fully developed and have acquired full nitrogen fixation activity (Bonaldi *et al.*, 2011). Roots and shoots of *A. evenia* and *A. afraspera* plants at 14dpi were labelled with $^{15}N_2$ and $^{13}CO_2$ isotopes,

respectively, for one day followed by a 48h chase and EA-IRMS analysis, to calculate photosynthesis and nitrogen fixation specific activities. Although the photosynthesis specific activity remained comparable for both *A. evenia* and *A. afraspera*, the *Bradyrhizobium* sp. ORS285 provided more nitrogen per nodule mass to *A. evenia* than to *A. afraspera* (Fig. 1C). Together, these results highlight the higher efficiency of S-type bacteroids over E-type bacteroids in *Aeschynomene* spp. plants at 14dpi. To investigate the metabolic processes that could account for this apparent increased symbiotic efficiency we conducted a transcriptomic study by RNA-seq of *Bradyrhizobium* sp. ORS285 grown as free-living cells and in interaction with *A. indica* and *A. afraspera* at 14dpi.

Overview of the transcriptomes

The transcriptomic dataset (Supporting Information Table S1) contained two bacteroid conditions and one culture condition, each in triplicate (Fig. 2A). Principal component analyses revealed a clear partitioning of the samples, the first axis (55% of the observed variance) separating culture from bacteroid conditions and the second axis separating E-morphotype (*A. afraspera*) from S-morphotype (*A. indica*) (20% of the observed variance; Fig. 2B). To assess the validity of the transcriptomic analysis, qRT-PCR was performed on a subset of 23 genes differentially regulated in one of the three conditions and including examples for each of the three conditions (Supporting Information Table S2). The $\log_2$-fold changes (LFC) obtained by both techniques were compared and appeared to be consistent (linear regression R^2 value = 0.77; Fig. 2C).

A first, general analysis of the transcriptomes was obtained by comparing the representation of the different Cluster of Orthologous Genes (COG) under each condition (Fig. 2D). This COG analysis revealed that bacteroids - regardless of their morphotype - undergo a metabolic transition hallmarked by the upregulation of genes involved in energy production/conversion

(COG class C) and inorganic ion transport/metabolism (COG class P) on the one hand and by the drastic downregulation of genes involved in translation (COG class J), signal transduction (COG class T), secondary metabolites biosynthesis (COG class Q) and cell motility (COG class N) on the other hand (Fig. 2D). These data corroborate the common view of bacteroid physiology, which is rewired towards nitrogen fixation and export of the fixed nitrogen, whereas several housekeeping functions are shut down, indicating that bacteroids behave as nitrogen-fixing organelles or ammoniaplasts within plant cells (Udvardi and Poole, 2013).

At the gene level, our differential expression analysis, using parameters of False Discovery Rate (FDR) < 0.01 and LFC > 1.58 (fold-change > 3), allowed the identification of 1680 (24.9% of total CDS) and 2123 (31.5%) differentially expressed genes (DEG) for E-bacteroids and S-bacteroids, respectively, compared to the YM culture, and 515 (7.6%) DEGs in the E-bacteroid *vs.* S-bacteroid comparison. This number of DEGs between E- and S-type bacteroids is substantially higher as compared to those found by a previous study in *Bradyrhizobium diazoefficiens* on three different phaseoloid host plants, soybean, cowpea and siratro, all forming U-type bacteroids (35, 38 and 129 DEGs per comparison) (Koch *et al.*, 2010). This high number of DEGs between E- and S-type bacteroids is in agreement with a significantly different functioning of these morphotypes.

Common bacteroid-specific genes

The transcriptome dataset contained 1342 DEGs, which are commonly up- or down-regulated in the two bacteroid types (Fig. 3; Supporting Information Table S3). As expected, the genes encoding the nitrogen fixation functions were amongst the most upregulated ones (Supporting Information Fig. S1). Together, the transcripts of this gene cluster represent about 22% of all transcripts in bacteroids. This amazing number illustrates in how far bacteroids are truly ammoniaplasts, converging all their efforts towards nitrogen fixation.

A distinct, strongly upregulated locus carries the genes encoding an uptake hydrogenase and accessory functions (LFC=2-7; Supporting Information Fig. S2). About one third of the energy consumed by nitrogenase is lost through hydrogen production, which is an obligate by-product of nitrogen fixation. Rhizobia that possess an uptake hydrogenase can recycle this hydrogen and thereby recover some of the lost energy by oxidizing hydrogen. This process is thus a benefit for the plant. Interestingly, the cluster is expressed about twice as high in *A. indica* bacteroids as compared to *A. afraspera* bacteroids. Although this difference does not match the threefold induction (LFC > 1.58) criterion for qualifying these genes as DEGs between the two bacteroid types, this observation suggests a more efficient hydrogen recycling in the spherical bacteroids which could be a contributing factor to the different symbiotic performance of these bacteroids.

The upregulation of phosphate (Supporting Information Fig. S2) and phosphonate (Supporting Information Fig. S1) transport and utilization coupled to the upregulation of a set of genes implicated in phospholipid metabolism (Supporting Information Fig. S3) indicate that phosphate and phospholipid homeostasis needs to be modulated in both the E-morphotype and S-morphotype bacteroids (Supporting Information Appendix S2). Remarkably, the locus carrying the phosphonate utilization genes is immediately adjacent to the above-mentioned nitrogen fixation genes (Supporting Information Fig. S1) while the phosphate transport locus is adjacent to the hydrogenase uptake locus (Supporting Information Fig. S2), reinforcing the view that phosphate and phosphonate utilization are important for symbiotic nitrogen fixation. Besides these phosphate and phosphonate transport systems, several other genes involved in metabolite transport were upregulated in the two bacteroid types, reflecting the adaptation of the metabolism of the bacteroids to the intracellular environment in the symbiotic nodule cells (Supporting Information Appendix S2).

Some members of the *Aeschynomene* genus are only effectively nodulated in nature by a nod factor independent mechanism. They belong to the *Aeschynomene* inoculation group III (Chaintreuil *et al*., 2013). This group is only nodulated by photosynthetic bradyrhizobia and a specific group of non-photosynthetic bradyrhizobia including the strain STM3843 (Okazaki *et al.*, 2016; Miché *et al.,* 2010). We reasoned that these group of strains may have specific functions that allow them to do so. We used the comparative genomics tool of the MicroScope web-based application (https://www.genoscope.cns.fr; Vallenet *et al.*, 2017) to identify gene sets only present in photosynthetic *Aeschynomene* symbionts and absent in other bradyrhizobia. This resulted in the identification of 116 genes (Supporting Information Table S4). Interestingly, we identified several gene regions that were upregulated in bacteroids (Supporting Information Fig. S4). Among them is a large gene cluster of 12 genes (Supporting Information Fig. S5), two other clusters of two or three contiguous genes and several individual genes. Unfortunately, for none of these genes a specific function based on annotation can be proposed (Supporting Information Appendix S2).

Interestingly, a large gene cluster encoding a type II secretion system (T2SS) is strongly upregulated in both bacteroid types (LFC=2.8-9.6; Supporting Information Fig. S6). T2SSs are widespread in rhizobia but their role in symbiosis has not been analyzed. This machinery secretes proteins to the extracellular environment, ranging from one to several tens of protein substrates depending on the species (Cianciotto and White, 2017). An intriguing possibility is that these T2SS substrates function as (extracellular) effectors in symbiosis in analogy to the effectors secreted by the well-known rhizobial type III secretion systems (T3SS) which suppress host immune responses or induce nodulation (Miwa and Okazaki, 2017). The T2SS locus of strain ORS285 carries in addition to the core components of the apparatus, a few genes that potentially encode secretion substrates. These genes are also strongly or very strongly induced in bacteroids (LFC=6-9). However, they encode proteins that are only

vaguely annotated (see Supporting Information Appendix S2 for a more detailed description of the T2SS in *Bradyrhizobium* sp. ORS285).

Many rhizobia carry their major specific symbiotic functions, including the nitrogenase encoding genes and the nodulation genes for Nod factor biosynthesis, on a plasmid or a symbiotic island (Poole *et al.*, 2018). *Bradyrhizobium* sp. strain ORS285 has a symbiotic island carrying the nodulation genes and a T3SS but the above described nitrogenase and hydrogenase genes are not located on it, suggesting that this bacterium may have been a free nitrogen fixer before it became a plant symbiont (see Supporting Information Appendix S2 for a full description of the symbiosis island). We specifically analyzed the expression pattern of the genes located on the island of ORS285 (Supporting Information Fig. S7). None of them showed an upregulation in bacteroids, suggesting that contrary to other rhizobia, the symbiosis island of strain ORS285 has only a function at early stages of nodulation for Nod factor production encoded by the nodulation genes (Nouwen *et al.*, 2016). In addition, the T3SS of the symbiosis island in ORS285 does not seem to have a role in bacteroids, in agreement with a previously reported mutant analysis (Okazaki *et al.*, 2016).

The two hallmarks of the TBD process are the morphological changes of the bacteria, either elongation to a long rod or inflation to a large sphere, and the polyploidization of their genome (Czernic *et al.*, 2015). The nature of these two processes suggests the involvement of the peptidoglycan wall and the cell cycle regulatory machinery in bacteroid differentiation. Accordingly, many of the cell cycle (Supporting Information Fig. S8) and peptidoglycan biosynthesis (Supporting Information Fig. S9) genes show a shift in expression, either up- or down-regulated, in bacteroids confirming the importance of the cell wall dynamics and the cell cycle regulation during bacteroid differentiation (Supporting Information Appendix S2). Other important features involved in the chronic establishment of the bacteroids are the lipopolysaccharide (LPS) and hopanoid lipids of the membrane since mutants that are affected

in the production of these lipids form defective bacteroids (Silipo *et al.*, 2014; Kulkarni *et al.*, 2015; Busset *et al.*, 2016, 2017). However, the genes involved in their biosynthesis show no or only slight differences in expression (Supporting Information Fig. S10, S11 and S12). Only three hopanoid biosynthesis genes *hpnO*, *hpnP* and *hpnG* are slightly but significantly upregulated in bacteroids suggesting that the bacteroids use hopanoids to adjust the properties of the bacteroid membranes to tune their rigidity in accordance to the local stress condition (see Supporting Information Appendix S2).

Finally, we noticed the strong downregulation of several bacterial functions in bacteroids including pilus formation, motility and chemotaxis (Supporting Information Fig. S13), protein synthesis (Supporting Information Fig. S14), photosynthesis (Supporting Information Fig. S15) and secondary metabolite synthesis (Supporting Information Fig. S16). We refer to Supporting Information Appendix S2 for a further description of these functions.

Identification of morphotype-specific genes

To identify the physiological bases of the differential efficiency of the E- and S-morphotype bacteroids, we next focused on the 515 genes which are differentially expressed between these bacteroid types in the *A. afraspera* and *A. indica* nodules. Hierarchical clustering, including the culture condition, grouped the 515 DEGs into five distinct classes (Fig. 4, Supporting Information Table S5). The first class comprises 142 genes induced only in *A. indica* bacteroids, another class of 136 genes are induced only in *A. afraspera* bacteroids, 35 genes are repressed in *A. afraspera* bacteroids and 35 genes in *A. indica* bacteroids and finally, 167 genes are repressed in both bacteroid types but more strongly in the *A. indica*.

Although a large portion of the DEGs between *A. afraspera* and *A. indica* bacteroids have no or an imprecise annotation (42% are conserved proteins of unknown function and 15% proteins with putative functions), the annotated ones hint at significant difference in the

metabolism of both bacteroid types. The most striking difference is the citrate synthase encoding gene *citA* (BRAD285_v2_0083) which is almost 200 times induced in the *A. indica* bacteroids while it is only weakly induced in the *A. afraspera* bacteroids. This high induction makes *citA* having the 11th rank of most highly expressed genes in *A. indica* bacteroids underscoring its importance in these bacteroids. It suggests that in these bacteroids the tricarboxylic acid (TCA) cycle is particularly active. On the other hand, *A. afraspera* bacteroids have a much higher expression of the isocitrate lyase gene *aceA* suggesting that these bacteroids have an enhanced use of the glyoxylate shunt of the TCA cycle as compared to the *A. indica* bacteroids. The higher expression of biotin (*bioADFB*, LFC>2.1; Supporting Information Fig. S17) and cobalamin (vitamin B_{12}) (16 genes, LFC=1.3-2.2; Supporting Information Fig. S18) biosynthesis genes in the *A. indica* bacteroids compared to the *A. afraspera* bacteroids is another observation that could hint at a difference in the TCA cycle between these two bacteroid types. These two cofactors are used by the enzymes propionyl-CoA carboxylase and methylmalonyl-CoA mutase (MutB, BRAD285_v2_4653) respectively to form the TCA metabolite succinyl-CoA from propionyl-CoA. We identified in *Bradyrhizobium* sp. strain ORS285 two other enzymes using the cobalamin cofactor. They are methionine synthase (MetH, BRAD285_v2_0497) and the class II ribonucleotide reductase (RNR, BRAD285_v2_3862, *nrd* gene). However while these genes are strongly downregulated in both bacteroid types, the *mutB* gene as well as *mutA* encoding the small subunit of the methylmalonyl-CoA mutase are slightly upregulated, suggesting that this enzyme is active in bacteroids. The two genes encoding the alpha- and beta-subunits of the propionyl-CoA carboxylase are also highly expressed in both bacteroid-types. Thus, the biotin and vitamin B12 production could be increased in the *A. indica* bacteroids in response to a higher demand of the TCA metabolite succinyl-CoA.

In addition, the inferred higher biotin production in the *A. indica* bacteroids is correlated with

the higher expression of several other biotin-dependent carboxylases (Supporting Information Fig. S17). This includes three carboxylase subunits encoded by a cluster of seven genes organized in one or two operons (from BRAD285_v2_5128 to BRAD285_v2_5135). The cluster is only vaguely annotated but carries besides the carboxylase genes, genes encoding an ABC transporter as well as enzymes involved in the metabolism of amines suggesting an implication in nitrogen metabolism.

Besides these examples, several other genes differentially expressed between the two bacteroid types hint at metabolic differences. They include for example several amino acid biosynthesis genes. However, pinpointing precise pathways is in many cases difficult because of incomplete annotations. In addition, the many transporter-encoding genes (ABC transporters, outer membrane proteins, RND efflux transporters) among these DEGs also confirm metabolic differences.

A set of DEGs encoding proteins with functions related to bacterial stress responses underlies in another way the differences between the two bacteroid types. Among the most strongly expressed genes in the *A. indica* bacteroids, 18th rank, is *csbD* (BRAD285_v2_5411). It has an almost 20-fold induction while in the *A. afraspera* it is less than 3-fold induced. A second *csbD* gene, BRAD285_v2_0601, shows a similar pattern but is less strongly expressed. CsbD is widely conserved in bacteria and is a general stress response protein although its precise role remains to be determined (Prágai and Harwood, 2002). Adjacent to the *csbD* gene BRAD285_v2_5411 is the gene BRAD285_v2_5408, also very strongly and specifically induced (about 130 fold) and highly expressed (43th rank) in the *A. indica* bacteroids. This gene encodes a GTPase with unknown function but it is also widely conserved in bacteria and contributes to bacterial fitness in various stress conditions in different bacteria (Khil and Camerini-otero, 2002; http://fit.genomics.lbl.gov/). We further identified the PhyR-σ^{EcfG} signaling cascade (BRAD285_v2_0553 to BRAD285_v2_0555) as upregulated in bacteroids

but two to three times higher in the *A. indica* bacteroids. This cascade is a master regulator of the general stress response mediating resistance to various stresses in Alphaproteobacteria (Gourion *et al.*, 2009). Interestingly, a Tn5 mutant in the σ^{EcfG} gene of *Bradyrhizobium* sp. ORS278, a strain closely related to ORS285 formed non-functional nodules on *A. indica* (Bonaldi *et al.*, 2010) and corresponding mutants in *Bradyrhizobium japonicum* strain USDA110 similarly formed defective nodules on soybean and mungbean (Gourion *et al.*, 2009). The *otsA* gene (BRAD285_v2_7091) encodes the trehalose-phosphate synthase, which catalyzes the first step in the synthesis of the osmo-protectant disaccharide trehalose. The *otsA* gene is strongly induced in both bacteroid types but to a much higher level in the *A. indica* bacteroids. In addition, the *glgX* gene (BRAD285_v2_2141) has a similar expression pattern as *otsA*. GlgX is a glycogen debranching enzyme providing glucose units that is coupled to trehalose synthesis in some bacteria (Schneider *et al.*, 2000; Seibold and Eikmanns, 2007). Interestingly, it was reported that overexpression of *otsA* in *Rhizobium etli* resulted in higher nitrogen fixation of *Phaseolus vulgaris* nodules which was correlated with an increased plant biomass, grain-yield and drought tolerance (Suárez *et al.*, 2008).

An additional osmo-protectant function specifically activated in the *A. indica* bacteroids is provided by the high-affinity potassium uptake transporter encoded by the *kdpABC* operon (Huang *et al.*, 2017). The role of this transporter is to maintain cell shape and turgor. The operon is regulated by the downstream two component regulator genes *kdpDE*. KdpD is a turgor- and osmosensor that is activated by high osmolarity (Jung and Altendorf, 2002). Upon activation of the sensor, the response regulator KdpE is phosphorylated and activates transcription of the *kdpABC* genes. Upstream of *kdpA* are two small open reading frames, encoding peptides of 26 and 29 amino acids, and the gene BRAD285_v2_1469 which are co-transcribed with *kdpABC*. These genes are conserved in *Bradyrhizobium* spp. and could be new components or regulators of this well characterized potassium transporter.

Furthermore, we identified DEGs likely involved in stress regulation that are specifically upregulated in the *A. afraspera* bacteroids. This includes the gene BRAD285_v2_4589, encoding the protease DegP (protease Do). The DegP protease and chaperone is part of the envelope stress response in bacteria and permits the recognition and relief of periplasmic protein-folding stress (Grabowicz and Silhavy, 2017). The genes displaying the highest differential expression in the *A. afraspera* bacteroids compared to the *A. indica* bacteroids, were the *czcABC* genes (LFC>5) which encode a cation efflux RND transporter that is potentially involved in the regulation of the cation homeostasis. We further identified an additional gene cluster involved in metal homeostasis. Two *suf* clusters are present in the *Bradyrhizobium* sp. ORS285 genome. One of them, *sufBCDSE_1* is part of the nitrogenase locus described above and is induced in both bacteroid types (LFC>8.8), whereas the second *sufBCDSTA* cluster (BRAD285_v2_3776 to BRAD285_v2_3783), located on another genome locus, was specifically expressed in *A. afraspera* bacteroids (LFC>3.1). The SufABCDSE proteins are involved in iron metabolism by contributing to iron–sulfur cluster biogenesis and repair and they play a critical role in oxidative stress response (Nachin *et al.*, 2003; Jang and Imlay, 2010). It remains to be determined whether the role of the second *sufBCDS* cluster is redundant with that of the first one or specifically necessary for a functional symbiosis in *A. afraspera*.

Taken together our observations suggest that the *A. indica* and *A. afraspera* bacteroids display, besides a number of common responses, also substantial transcriptome differences, notably in metabolism including energy metabolism (hydrogenase, TCA cycle), amino acid and amine metabolism, as well as in the management of stresses imposed by the nodule environment. It will be of importance to determine whether these specificities are related to the different symbiotic efficiencies of the E- and S-morphotype bacteroids in these host plants.

Are differences between A. afraspera and A. indica bacteroids generalizable to E- and S-morphotype bacteroids in other Aeschynomene hosts?

We next addressed the questions whether the observed differential gene expressions between culture and nodule conditions are generalizable to other host plants and whether the observed differential gene expressions between *A. afraspera* and *A. indica* bacteroids are merely related to the host plant or rather reflect fundamental features of the different bacteroid differentiation states. To help answering these two questions, we performed qRT-PCR analyses on cultured bacteria and bacteroids from *A. afraspera* and *A. nilotica* generating E-type bacteroids and *A. evenia* and *A. indica* generating S-type bacteroids. Based on RNA-seq dataset, 23 genes were selected, representing examples of the observed expression profiles. Interestingly, the overall expression profiles of these genes were more similar between hosts inducing the same morphotypes than between hosts inducing contrasted morphotypes, even if the profiles were more variable between the two E-type bacteroids than between the two S-type bacteroids (Fig. 5A). The phylogenetic distance between *A. nilotica* and *A. afraspera* is higher than between *A. evenia* and *A. indica* (Chaintreuil *et al.*, 2013). This could translate into more similar microenvironments for the bacteroids residing within these latter two plants. The tested gene set included 14 genes equally expressed in the *A. afraspera* and *A. indica* transcriptomes, including seven genes repressed in the two bacteroid types and seven upregulated genes. The expression patterns of all these genes were validated by qRT-PCR (Fig. 5B), although three of the bacteroid-enhanced genes seemed to be slightly higher expressed in E-type bacteroids than S-type bacteroids (Fig. 5C). We then tested nine genes specifically induced in *A. indica* or *A. afraspera* respectively, to discriminate if they are host-specific or morphotype-specific. Four genes appeared to be specific of the host *A. afraspera* as they were not upregulated in *A. nilotica* bacteroids. Finally, among the five putative S-morphotype-specific genes tested, one

was specific for the *A. indica* host, while the four others appeared to be S-morphotype-specific with similar level of expression in the *A. indica* and *A. evenia* bacteroids and no or weak expression in both E-morphotype bacteroids (Fig. 5C).

Thus overall, the expression patterns revealed by the transcriptome analysis of *A. indica* and *A. afraspera* bacteroids are partially generalizable to S- and E-morphotype bacteroids although it must be recognized that some of the expression profiles are specific for a particular host and are thus not inherent features of particular bacteroid types.

Functional analysis of candidate genes reveals strong robustness and functional redundancy in the symbiotic transcriptome

We selected 31 genes or loci from our transcriptomic dataset that were previously uncharacterized in symbiosis (thus excluding canonical symbiotic genes such as *nif* and *fix* genes) for functional analysis (Supporting Information Table S6). Because of the intriguing upregulation of the T2SS and its complex organization (local duplication of the locus), different deletion mutants were constructed in this locus. Both copies of the T2SS were individually deleted and a double mutant carrying both deletions was also constructed. Furthermore, mutants were constructed in which the single prepilin signal peptidase gene was removed as well as a mutant in which the complete region was deleted. An insertional mutant in one of the T2SS substrate proteins, the putative S-layer protein was also constructed. In addition, insertional mutants in 27 genes were produced (Supporting Information Table S6). The selection criteria were based on their expression profile and level. Indeed, genes chosen are either bacteroid-induced and/or morphotype-specific. Twenty of them have a high expression level in at least one bacteroid condition (first decile of highest expression). Two genes have been selected for their relevance in plant microbe-interactions (BRAD285_v2_5219, putatively involved in salicylic acid metabolism) and nodule

development (BRAD285_v2_5527, putatively involved in auxin metabolism).

Mutants were tested for their symbiotic performance. We assessed the capacity of the mutant bacteria to form nodules and sustain plant growth in hosts inducing E-type and S-type bacteroids. Unexpectedly, in all cases, plants displayed pink nodules and green leaves similarly to the controls inoculated with the wild type strain (data not shown). For ten of those mutants and the T2SS deletion mutants, we determined the nodule number per plant and measured nitrogenase activity by Acetylene Reduction Assays. No difference in nodule formation and its kinetics, nodule tissue organization and nitrogen fixation could be identified between plants inoculated by wild-type and mutant bacteria (Supporting Information Fig. S19 and S20). This result indicates that none of the candidate genes is essential for symbiosis under the tested conditions, which were identical to the conditions used for the transcriptome analysis. This observation is consistent with a previous report in another symbiotic system, the interaction of *Rhizobium leguminosarum* with pea, where the mutation of 37 bacteroid-induced genes resulted in no detectable symbiotic phenotype (Karunakaran *et al.*, 2009).

To further analyze these surprising results, we compiled the expression patterns of all reported genes with symbiotic phenotypes in the corresponding mutants of *Bradyrhizobium* sp. ORS285 or related photosynthetic bradyrhizobia (Supporting Information Table S7). For a total of 131 mutants, the corresponding genes in 93 (71%) cases were not differentially expressed in bacteroids. Only 22 genes (17%) with symbiotic phenotypes were upregulated in bacteroids, seven of them encoding components or regulators of the nitrogenase. Counterintuitively, 16 (12%) of the genes displaying a nodulation phenotype when mutated were even downregulated in the bacteroids. Since the large majority of these mutants were obtained by unbiased forward genetic screens (Giraud *et al.*, 2007; Bonaldi *et al.*, 2010), we can conclude that genes that result in a symbiotic phenotype when mutated, are not necessarily differentially expressed.

The absence of symbiotic phenotypes in bacteroid-upregulated genes in this and previous work

requires further explanation because induced and high expression is in general taken as an adaptive response of the organism or cell to a specific condition so that genes are expressed when they are required. First, it may suggest a strong robustness and a functional redundancy within the bacterial symbiotic program. Robustness implies that each of the induced gene functions have a slight additive contribution to the symbiotic process but that the removal of an individual function has no or an imperceptible effect, not enough to let the edifice crumble, especially under controlled laboratory conditions. For example, the putative T2SS determined S-layer could together with the LPS and the hopanoid lipids constitute a multilayered envelope which provides sufficient strength to the bacteria to withstand the NCR peptides or other stresses in the symbiotic nodule cells. Removing the individual components of this envelope could have an imperceptible or moderate effect on the bacteroids (Silipo *et al.*, 2014; Kulkarni *et al.*, 2015; Busset *et al.*, 2016, 2017; this work) but, possibly, removing multiple parts simultaneously would be much more deleterious for the bacteroids. Redundancy on the other hand can be achieved by homologous genes or through a functional complementation by an unrelated pathway. For some of the insertional mutants, we cannot exclude that redundancy with a homologous gene is responsible for the lack of symbiotic phenotype although none of the selected genes has a homolog of more than 55% (Supporting Information Table S6). The robustness and redundancy could be experimentally verified by the construction of mutants carrying multiple mutations.

It is also possible that the mutated genes have a role in symbiosis under particular conditions that are not captured by our laboratory nodulation and nitrogen fixation assays but that might exist in natural growth conditions. In that case, one has to assume that these genes are not finely regulated but activated to a standby mode even if not needed in the present condition but in such a way that their gene products are available in the advent of a changing condition. In the same vein, the genes could be activated to function in a future stage of the nodule lifecycle

such as for example in nodule senescence. Nodule senescence is the last stage of the nodule lifetime and leads to the digestion of the host tissues as well as a large part of the bacteroids. This process is mediated by the massive activation of macromolecule-degrading enzymes by the host (Van de Velde *et al.*, 2006). It is possible that bacteroids activate functions that could help them in their attempt to survive the future attack of the host to digest them. Still another possibility is that these genes are only required for symbiosis with certain *Aeschynomene* hosts but not with others while their regulation does not distinguish between differences in host environments. A striking example is the *nifV* gene encoding a homocitrate synthase. We find that this gene is induced to very high and equal levels in both nodule types (LFC>11; Supporting Information Table S1) but mutagenesis has shown that the gene is only required for symbiosis with *A. indica* and not with *A. afraspera* (Nouwen *et al.*, 2017). Homocitrate is an essential intermediate in the production of the iron-molybdenum cofactor (FeMo-co) of nitrogenase. Thus, nitrogen fixation in bacteroids requires a functional *nifV* gene except if homocitrate is supplied by the host plant through the nodule-specific expression of the *FEN1* gene, a plant homocitrate synthase gene (Hakoyama *et al.*, 2009; Nouwen *et al.*, 2017). *A. afraspera* has such a nodule-specific *FEN1* gene but *A. indica* not (Nouwen *et al.*, 2017). Thus, *Bradyrhizobium* sp. ORS285 activates *nifV* expression in nodules irrespectively of the supply of homocitrate by the host plant.

Finally, we cannot exclude that part of the transcriptome changes that we have observed are non-adaptive responses. From recent findings, the picture emerges that transcriptional responses in bacteria or yeast are in large part suboptimal. Systematic comparisons of mutant fitness and gene expression have shown that these two parameters are poorly correlated (Giaever *et al.*, 2002; Deutschbauer *et al.*, 2011; Price *et al.*, 2013). Thus, genes may be upregulated in a specific context only as a side effect without providing a functional contribution. In the case of rhizobia, the bacteria have a dual lifestyle as free-living soil and

rhizosphere bacteria and as nodule symbionts. It is possible that these bacteria have certain adaptations enhancing their fitness during free-living growth in the proximity of plants and that they therefore recognize certain plant-derived signals that induce the genes encoding these adaptations. It is further possible that the same or similar signaling molecules are also present inside nodules thereby enhancing the expression of the same gene set which make however no specific contribution to the symbiosis.

Conclusions

Bacteroids in legume nodules come in different flavors, characterized by their typical size, morphology, degree of polyploidy and membrane permeabilization. The functional implications on these differences are still mostly unclear. Since the differentiation process is imposed by the plant through the production or not of the NCR peptides, it was predicted that the differentiation process provides advantage to the host and thus different efficiencies of these bacteroids were suggested (Oono and Denison, 2010; Kereszt *et al.*, 2011). Our analysis, a case study comparing E- and S-morphotype bacteroids formed by two closely related host plants, revealed that S-morphotype bacteroids tended to be more efficient nitrogen fixers than elongated ones at 14dpi. At the molecular level, our transcriptomic survey of differentiated bacteroids displaying contrasted morphotypes allowed the identification of a bacteroid gene set that included determinants of the main metabolic pathways required for nitrogen fixation and symbiotic life. When comparing E- and S-morphotype bacteroids, we identified large gene sets that were specific for one or the other bacteroid type and that could provide hints on candidate functions associated with the increased symbiotic efficiency of spherical bacteroids. Nevertheless, as the two host plants are different species, we have to acknowledge that the detected differences in symbiotic efficiency and in gene expression can only be correlated with bacteroid types and that at this stage, no causal relationship can be concluded. Indeed, as our

qRT-PCR data demonstrates, some of the expression differences are not related to the bacteroid type but probably to some other, yet unknown, differences between the hosts. Extending the analyses to a larger number of hosts inducing E- and S-morphotype bacteroids would increase our capacity to distinguish between these confounding factors.

Disappointingly, mutants in a selected set of upregulated genes did not reveal any phenotype even if they were tested under identical conditions as the expression analysis. To go further in the characterization of these genes, possible strategies are enlarging the set of conditions in which phenotypes of mutants are tested or the creation of double, triple… mutants.

Finally, the present analysis could be usefully complemented with a metabolome and proteome comparison of the bacteroids. Our analysis hints at a different metabolism in the E- and S-morphotype bacteroids based on differential gene expression. However, cellular metabolic fluxes are determined more by substrate concentrations than protein abundances (Hackett *et al.*, 2016) and protein abundance does not always correlate with gene expression. Thus, metabolome and proteome studies have the potential to further deepen our understanding of the differences between the bacteroid types.

Experimental procedures

Plant growth and inoculation

Bradyrhizobium sp. strain ORS285 was grown at 30°C in YM medium supplemented with carbenicillin (50 µg/mL) (Guefrachi *et al.*, 2015). *A. indica, A. nilotica, A. evenia* and *A. afraspera* plants were grown in transparent test tubes filled with BNM medium and inoculated with strain ORS285 as described before (Guefrachi *et al.*, 2015). Plants were grown at 28°C, 80% humidity and under a 16h : 8 h of light : dark regimen.

Symbiotic efficiency analyses

The symbiotic efficiencies of the *A. indica, A. nilotica, A. evenia* and *A. afraspera* interactions with *Bradyrhizobium* sp. strain ORS285 were determined at 14 dpi using three independent methods: 1) plant biomass determination, 2) elemental analysis to determine the nitrogen content in whole plants and 3) fluxomic analysis to determine the transfer of fixed nitrogen to the plant in 24h. Details of these methods are provided in Supporting Information Appendix S1.

Transcriptome and qRT-PCR analyses

For gene expression analyses, nodules were harvested at 14 dpi. Procedures for qRT-PCR measurements, RNA extraction, RNA-seq with the SOLiD technology, sequencing data treatment, statistical analyses and genome mining are described in detail in Supporting Information Appendix S3. The DESeq2 (Love et al., 2014) normalized RNA-seq data are represented in the heat maps as the standard score (z-score) of the normalized read numbers (Fig. 3B and Fig. 4) or as the percentage of the sum of the normalized read numbers for all the samples (Supporting Information Fig. S1 to S18). The qRT-PCR data are shown as the percentage of the sum of values for all the samples (Fig. 5B and 5C).

Generation and phenotyping of bacterial mutants

For insertion mutagenesis, internal fragments (300-600 bp) of the target gene were PCR-amplified and cloned into the pVO155-nptII-aadA-GFP vector. The resulting constructs were introduced into *Bradyrhizobium* sp. ORS285 and mutants were selected on 50 µg/mL kanamycin and 50 µg/mL spectinomycin plates (Wongdee *et al.*, 2016). For the construction of deletion mutants, flanking regions of the gene were PCR amplified, fused by overlap extension PCR and cloned in the *sacB* suicide pNPTS129 plasmid. The resulting plasmids were introduced into *Bradyrhizobium* sp. ORS285 and mutants were obtained after a double

recombination event as described (Nouwen *et al.*, 2016). The genotype of all mutant strains was verified by PCR analysis. Primers for constructions and strain verification are listed in Supporting Information Table S8. The phenotype of the mutants was analyzed by inoculating *A. indica* and *A. afraspera* plants as indicated above. Fourteen dpi plants were collected, and their leaf and nodule color and number were visually inspected. Nitrogenase activity was determined by the acetylene reduction assay as previously described (Barrière *et al.*, 2017).

Acknowledgments

F.L. was 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, grant n° ANR-13-BSV7-0013 and used resources from the National Office for Research, Development and Innovation of Hungary, grant n° 120120 to A.K.

Figure legends

Fig 1. Evaluation of the symbiotic efficiency of *Bradyrhizobium* sp. ORS285 in symbiosis with *Aeschynomene* spp. plants inducing E- or S-type bacteroids. (A) Mass gains at 14 dpi of inoculated plants compared to non-inoculated plants (n>30) and per nodule mass. AA: *Aeschynomene afraspera*; AN: *Aeschynomene nilotica*, AI: *Aeschynomene indica*; AE: *Aeschynomene evenia*. Letters represent significant differences after ANOVA and *post-hoc* Tukey tests ($p<0.05$). (B) Carbon and nitrogen mass gains at 14 dpi determined by elemental analysis of the C and N mass of inoculated versus non-inoculated plants and per nodule mass.

(C) Specific photosynthetic and N_2 fixation activities from fluxomic analyses at 14 days post inoculation with *Bradyrhizobium* sp. ORS285.

Fig. 2. Experimental setup and general overview of the RNA-seq dataset. (A) RNA-seq experimental setup displaying the three biological conditions and the comparisons of interest. (B) Principal component analysis. (C) qRT-PCR cross validation of RNA-seq results obtained on three independent samples, based on the log2-transformed fold change (LFC) expression values. (D) Repartition of the reads among COG classes in the three conditions.

Fig. 3. Overview of the commonly regulated bacteroid gene set. (A) Venn diagram showing the overlap of the differentially expressed genes (fdr<0.01 and |LFC|>1.58) in E- and S-type bacteroids as compared to the culture reference. (B) Heatmap and hierarchical clustering of the 1342 differentially expressed genes in both bacteroid comparisons with bacterial culture, displaying bacteroid-specific patterns. The heatmap shows the standard score (z-score) of DESeq2 normalized read numbers. The color-coded scale bars for the normalized expression and LFC of the genes are indicated below the heatmap. Gene accessions are represented by the numerical value of their BRAD285_v2_xxxx format. Asterisks indicate genes for which mutant strains were generated.

Fig. 4. Heatmap of the 515 differentially expressed genes between E-type and S-type bacteroids. The left heatmap shows the standard score (z-score) of the DESeq2 normalized expression of genes in the three experimental conditions and three biological replicates. The heatmap at the right shows the LFC values for pairwise comparisons between the mean of the conditions. The color-coded scale bars for the normalized expression and LFC of the genes are indicated below the heatmap. Gene accessions are represented by the numerical value of

their BRAD285_v2_xxxx format. Asterisks indicate genes for which mutant strains were generated.

Fig. 5. qRT-PCR expression profiles of selected genes in four plants inducing E-type or S-type bacteroids. (A) Principal component analysis of the gene expression profiles in bacterial culture (YM), hosts inducing E-type bacteroids (*A. nilotica* (AN) and *A. afraspera* (AA)) and hosts inducing S-type bacteroids (*A. indica* (AI) and *A. evenia* (AE)) determined by qRT-PCR on biological triplicates for each condition. (B,C) Heatmap of the 23 genes of interest measured by qRT-PCR, discriminating bacteroid or culture specificity (B), and morphotype or host specificity (C). Expression levels are shown as percentage of the sum of values for all the samples.

References

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.

Barrière, Q., Guefrachi, I., Gully, D., Lamouche, F., Pierre, O., Fardoux, J., *et al.* (2017) Integrated roles of BclA and DD-carboxypeptidase 1 in *Bradyrhizobium* differentiation within NCR-producing and NCR-lacking root nodules. *Sci Rep* **7**: 9063.

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.

Bonaldi, K., Gourion, B., Fardoux, J., Hannibal, L., Cartieaux, F., Boursot, M., *et al.* (2010)

Large-scale transposon mutagenesis of photosynthetic *Bradyrhizobium* sp. strain ORS278 reveals new genetic loci putatively important for nod-independent symbiosis with *Aeschynomene indica*. *Mol Plant Microbe Interact* **23**: 760–770.

Busset, N., De Felice, A., Chaintreuil, C., Gully, D., Fardoux, J., Romdhane, S., *et al.* (2016) The LPS O-antigen in photosynthetic *Bradyrhizobium* strains is dispensable for the establishment of a successful symbiosis with *Aeschynomene* legumes. *PLoS ONE* **11**: e0148884.

Busset, N., Di Lorenzo, F., Palmigiano, A., Sturiale, L., Gressent, F., Fardoux, J., *et al.* (2017) The very long chain fatty acid (C26:25OH) linked to the lipid A is important for the fitness of the photosynthetic *Bradyrhizobium* strain ORS278 and the establishment of a successful symbiosis with *Aeschynomene* legumes. *Front Microbiol* **8**: 1821.

Chaintreuil, C., Arrighi, J-F., Giraud, E., Miché, L., Moulin, L., Dreyfus, B., *et al.* (2013) Evolution of symbiosis in the legume genus *Aeschynomene*. *New Phytol* **200**: 1247–1259.

Cianciotto, N.P., and White, RC. (2017) Expanding role of type II secretion in bacterial pathogenesis and beyond. *Infect Immun* **85**: 11–16.

Czernic, P., Gully, D., Cartieaux, F., Moulin, L., Guefrachi, I., Patrel, D., *et al.* (2015) Convergent evolution of endosymbiont differentiation in dalbergioid and Inverted Repeat-Lacking Clade legumes mediated by nodule-specific cysteine-rich peptides. *Plant Physiol* **169**: 1254-1265.

Deutschbauer, A., Price, M.N., Wetmore, K.M., Shao, W., Baumohl, J.K., Xu, Z., *et al.* (2011) Evidence-based annotation of gene function in *Shewanella oneidensis* MR-1 using genome-wide fitness profiling across 121 conditions. *PLoS Genet* **7**: e1002385.

Giaever, G., Chu, A.M., Ni, L., Connelly, C., Riles, L., Véronneau, S., *et al.* (2002) Functional profiling of the *Saccharomyces cerevisiae* genome. *Nature* **418**: 387–391.

Giraud, E., Moulin, L., Vallenet, D., Barbe, V., Cytryn, E., Avarre, J-C., *et al.* (2007) Legumes symbioses: absence of nod genes in photosynthetic bradyrhizobia. *Science* **316**: 1307–1312.

Gourion, B., Sulser, S., Frunzke, J., Francez-Charlot, A., Stiefel, P., Pessi, G., *et al.* (2009) The PhyR-σ^{EcfG} signalling cascade is involved in stress response and symbiotic efficiency in *Bradyrhizobium japonicum*. *Mol Microbiol* **73**: 291–305.

Grabowicz, M., and Silhavy, T.J. (2017) Redefining the essential trafficking pathway for outer membrane lipoproteins. *Proc Natl Acad Sci U S A* **114**: 4769–4774.

Guefrachi, I., Pierre, O., Timchenko, T., Alunni, B., Barriere, 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.

Hackett, S.R., Zanotelli, V.R.T., Xu, W., Goya, J., Park, J.O., Perlman, D.H., *et al.* (2016) Systems-level analysis of mechanisms regulating yeast metabolic flux. *Science* **354**: aaf2786.

Hakoyama, T., Niimi, K., Watanabe, H., Tabata, R., Matsubara, J., Sato, S., *et al.* (2009) Host plant genome overcomes the lack of a bacterial gene for symbiotic nitrogen fixation. *Nature* **462**: 514–517.

Huang, C-S., Pedersen, B.P., and Stokes, D.L. (2017) Crystal structure of the potassium-importing KdpFABC membrane complex. *Nature* **546**: 681–685.

Jang, S., and Imlay, J.A. (2010) Hydrogen peroxide inactivates the *Escherichia coli* Isc iron-sulphur assembly system, and OxyR induces the Suf system to compensate. *Mol Microbiol* **78**: 1448–1467.

Jung, K., and Altendorf, K. (2002) Towards an understanding of the molecular mechanisms of stimulus perception and signal transduction by the KdpD/KdpE system of *Escherichia coli*. *J*

Mol Microbiol Biotechnol **4**: 223–228.

Karunakaran, R., Ramachandran, V.K., Seaman, J.C., East, A.K., Mouhsine, B., Mauchline, T.H., *et al.* (2009) Transcriptomic analysis of *Rhizobium leguminosarum* biovar *viciae* in symbiosis with host plants *Pisum sativum* and *Vicia cracca*. *J Bacteriol* **191**: 4002–4014.

Kereszt, A., Mergaert, P., and Kondorosi, E. (2011) Bacteroid development in legume nodules: evolution of mutual benefit or of sacrificial victims? *Mol Plant Microbe Interact* **24**: 1300–1309.

Khil, P.P., and Camerini-Otero, R.D. (2002) Over 1000 genes are involved in the DNA damage response of *Escherichia coli*. *Mol Microbiol* **44**: 89–105.

Koch, M., Delmotte, N., Rehrauer, H., Vorholt, J., Pessi, G., and Hennecke, H. (2010) Rhizobial adaptation to hosts., a new facet in the legume root-nodule symbiosis. *Mol Plant Microbe Interact* **23**: 784–790.

Kondorosi, E., Mergaert, P., and Kereszt, A. (2013) A paradigm for endosymbiotic life: cell differentiation of Rhizobium bacteria provoked by host plant factors. *Annu Rev Microbiol* **67**: 611–628.

Kulkarni, G., Busset, N., Molinaro, A., Gargani, D., Chaintreuil, C., Silipo, A., *et al.* (2015) Specific hopanoid classes differentially affect free-living and symbiotic states of *Bradyrhizobium diazoefficiens*. *mBio* **6**: e01251-15.

Love, M.I., Huber, W., and Anders, S. (2014) Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. *Genome Biol* **15**: 550.

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**: 161–173.

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.

Miché, L., Moulin, L., Chaintreuil, C., Contreras□Jimenez, J.L., Munive□Hernández, J., Del Carmen Villegas□Hernandez, M., *et al.* (2010) Diversity analyses of *Aeschynomene* symbionts in Tropical Africa and Central America reveal that *nod*□ independent stem nodulation is not restricted to photosynthetic bradyrhizobia. *Environ Microbiol* **12**: 2152-2164.

Miwa, H., and Okazaki, S. (2017) How effectors promote beneficial interactions. *Curr Opin Plant Biol* **38**: 148–154.

Montiel, J., Downie, J.A., Farkas, A., Bihari, P., Herczeg, R., Bálint, B., *et al.* (2017) Morphotype of bacteroids in different legumes correlates with the number and type of symbiotic NCR peptides. *Proc Natl Acad Sci U S A* **114**: 5041–5046.

Nachin, L., Loiseau, L., Expert, D., and Barras, F. (2003) SufC: an unorthodox cytoplasmic ABC/ATPase required for [Fe-S] biogenesis under oxidative stress. *EMBO J* **22**: 427-437.

Nouwen, N., Arrighi, J-F., Cartieaux, F., Chaintreuil, C., Gully, D., Klopp, C., *et al.* (2017) The role of rhizobial (NifV) and plant (FEN1) homocitrate synthases in *Aeschynomene*/photosynthetic *Bradyrhizobium* symbiosis. *Sci Rep* **7**: 448.

Nouwen, N., Fardoux, J., and Giraud, E. 2016. NodD1 and NodD2 are not required for the symbiotic interaction of *Bradyrhizobium* ORS285 with nod-factor-independent *Aeschynomene* legumes. *PLoS ONE* **11**: e0157888.

Okazaki, S., Tittabutr, P., Teulet, A., Thouin, J., Fardoux, J., Chaintreuil, C., *et al.* (2016) Rhizobium-legume symbiosis in the absence of Nod factors: Two possible scenarios with or without the T3SS. *ISME J* **10**: 64–74.

Oono, R., Anderson, C.G., and Denison, R.F. (2011) Failure to fix nitrogen by non-reproductive symbiotic rhizobia triggers host sanctions that reduce fitness of their reproductive clonemates. *Proc R Soc Lond, B, Biol Sci* **278**: 2698–2703.

Oono, R., and Denison, R.F. (2010) Comparing symbiotic efficiency between swollen versus nonswollen rhizobial bacteroids. *Plant Physiol* **154**: 1541–1548.

Poole, P., Ramachandran, V., and Terpolilli, J. (2018) Rhizobia: from saprophytes to endosymbionts. *Nat Rev Microbiol* **16**: 291-303.

Prágai, Z., and Harwood, C.R. (2002) Regulatory interactions between the Pho and sigma(B)-dependent general stress regulons of *Bacillus subtilis*. *Microbiology* **148**: 1593–602.

Price, M.N., Deutschbauer, A.M., Skerker, J.M., Wetmore, K.M., Ruths, T., Mar, J.S., *et al.* (2013) Indirect and suboptimal control of gene expression is widespread in bacteria. *Mol Syst Biol* **9**: 660.

Schneider, D., Bruton, C.J., and Chater, K.F. (2000) Duplicated gene clusters suggest an interplay of glycogen and trehalose metabolism during sequential stages of aerial mycelium development in *Streptomyces coelicolor* A3(2). *Mol Gen Genet* **263**: 543–553.

Seibold, G.M., and Eikmanns, B.J. (2007) The *glgX* gene product of *Corynebacterium glutamicum* is required for glycogen degradation and for fast adaptation to hyperosmotic stress. *Microbiology* **153**: 2212–2220.

Sen, D., and Weaver, R.W. (1981) A comparison of nitrogen-fixing ability of peanut, cowpea and siratro plants nodulated by different strains of rhizobium. *Plant Soil* **60**: 317-319.

Silipo, A., Vitiello, G., Gully, D., Sturiale, L., Chaintreuil, C., Fardoux, J., *et al.* (2014) Covalently linked hopanoid-lipid A improves outer-membrane resistance of a *Bradyrhizobium* symbiont of legumes. *Nat Commun* **5**: 5106.

Suárez, R., Wong, A., Ramírez, M., Barraza, A., Del, M., Orozco, C., *et al.* (2008) Improvement of drought tolerance and grain yield in common bean by overexpressing trehalose-6-phosphate synthase in rhizobia. *Mol Plant Microbe Interact* **21**: 958–966.

Udvardi, M., and Poole, P.S. (2013) Transport and metabolism in legume-rhizobia symbioses. *Annu Rev Plant Biol* **64**: 781–805.

Vallenet, D., Calteau, A., Cruveiller, S., Gachet, M., Lajus, A., Josso, A., *et al.* (2017) MicroScope in 2017: An expanding and evolving integrated resource for community expertise of microbial genomes. *Nucleic Acids Res* **45**: D517–D528.

Van de Velde, W., Guerra, J.C.P., De Keyser, A., De Rycke, R., Rombauts, S., Maunoury, N., *et al.* (2006) Aging in legume symbiosis. A molecular view on nodule senescence in *Medicago truncatula*. *Plant Physiol* **141**: 711–720.

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.

Supporting Information

Appendix S1. Methods for symbiotic efficiency analyses.

Appendix S2. Complementary results on common bacteroid-specific genes

Appendix S3. Methods for transcriptome analysis.

Table S1. Full dataset of *Bradyrhizobium* sp. ORS285 transcriptome.

Table S2. Genes used for qRT-PCR experiments.

Table S3. Dataset of bacteroid-specific genes.

Table S4. Common genes specific for photosynthetic, Nod-independent *Aeschynomene* symbionts.

Table S5. Bacteroid morphotype-specific genes.

Table S6. List of constructed *Bradyrhizobium* sp. ORS285 mutants.

Table S7. Expression of genes with published mutant symbiotic phenotypes in strain ORS285 or related photosynthetic bradyrhizobia.

Table S8. Primers used for qRT-PCR on *Bradyrhizobium* sp. ORS285 and for mutagenesis.

Fig. S1. Nitrogenase region of *Bradyrhizobium* sp. ORS285 is highly induced in bacteroids.

Fig. S2. The *Bradyrhizobium* sp. ORS285 locus containing *pst* and *hup* genes is induced in bacteroid conditions.

Fig. S3. *Bradyrhizobium* sp. ORS285 bacteroids highly express a genomic region involved in phospholipid metabolism as well as phospholipases C.

Fig. S4. Expression of photosynthetic *Bradyrhizobium*-specific genes.

Fig. S5. Expression in *Bradyrhizobium* sp. ORS285 of the photosynthetic *Bradyrhizobium*-specific *vanB* gene cluster.

Fig. S6. *Bradyrhizobium* sp. ORS285 features a duplicated Type II Secretion System (T2SS) induced in bacteroids.

Fig. S7. Expression of the *Bradyrhizobium* sp. ORS285 symbiotic island.

Fig. S8. Expression of cell cycle regulatory genes.

Fig. S9. Expression of peptidoglycan biosynthesis genes.

Fig. S10. Expression of the hopanoid biosynthesis pathway.

Fig. S11. Expression of the lipopolysaccharide (LPS) biosynthesis pathway.

Fig. S12. Expression of polysaccharide biosynthesis genes.

Fig. S13. Expression of genes involved in chemotaxis, flagellum and pilus biosynthesis.

Fig. S14. Expression of ribosomal genes.

Fig. S15. Expression of genes involved in bacterial photosynthesis.

Fig. S16. Expression of Non-Ribosomal Peptide Synthase (NRPS) and Quorum Sensing (QS) genes.

Fig. S17. Expression of genes involved in biotin biosynthesis and utilization.

Fig. S18. Expression of genes involved in vitamin B12 synthesis and utilization.

Fig. S19. Nodulation phenotypes of eight mutant strains.

Fig. S20. Nodulation phenotypes of mutants in the general secretion pathway.

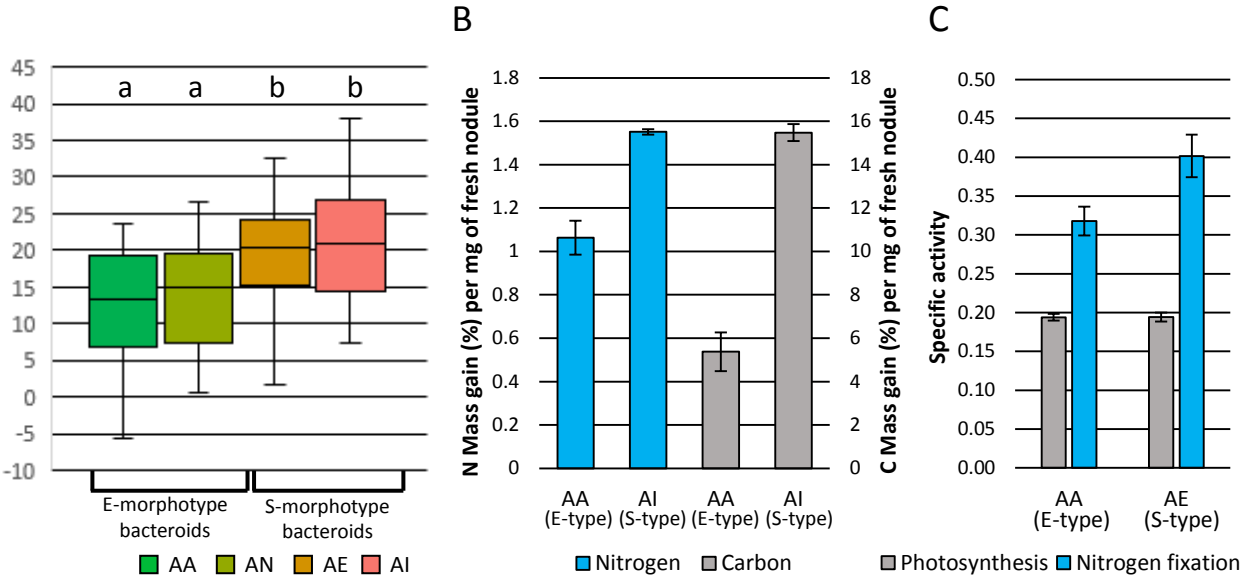
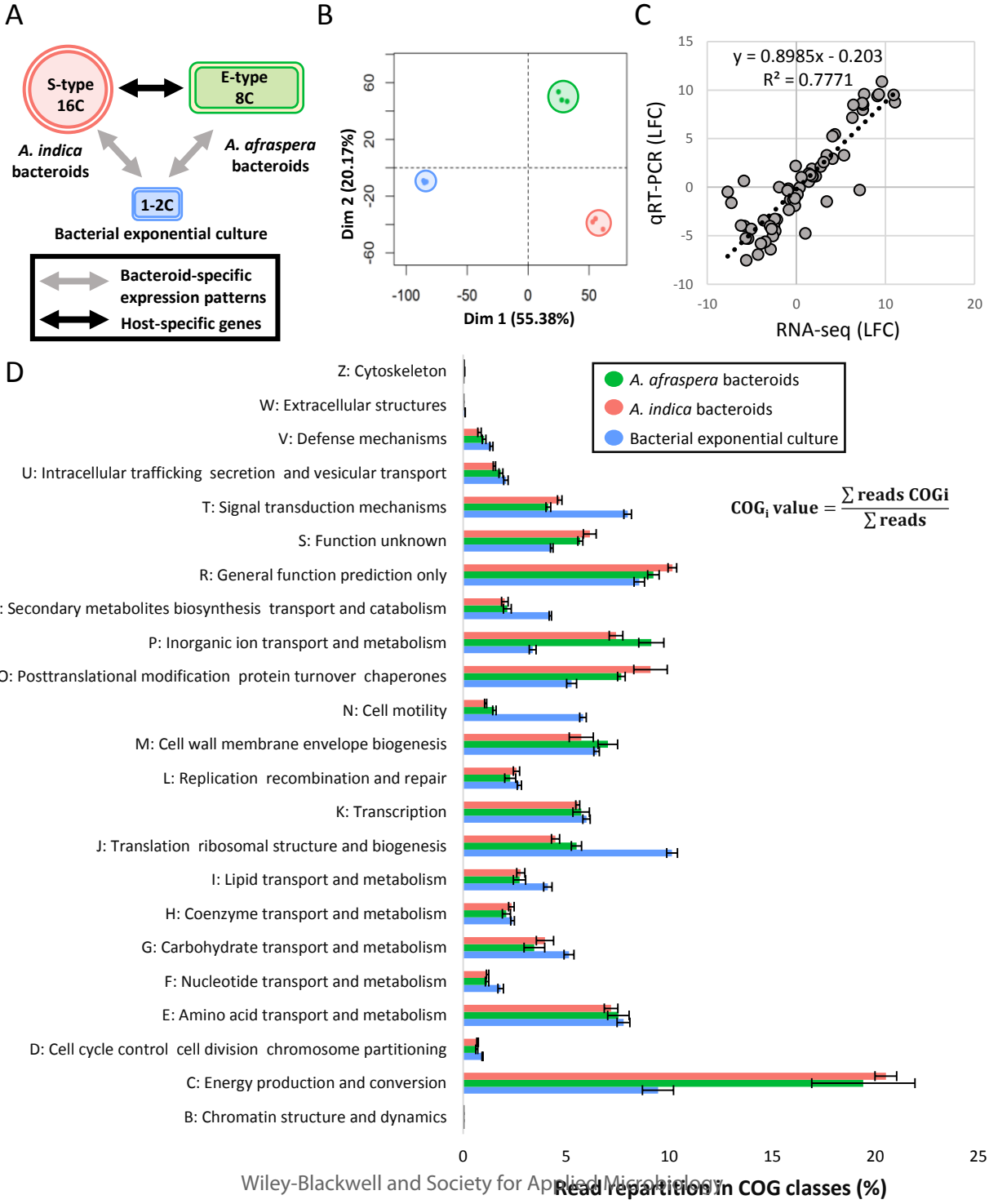


Fig. 2



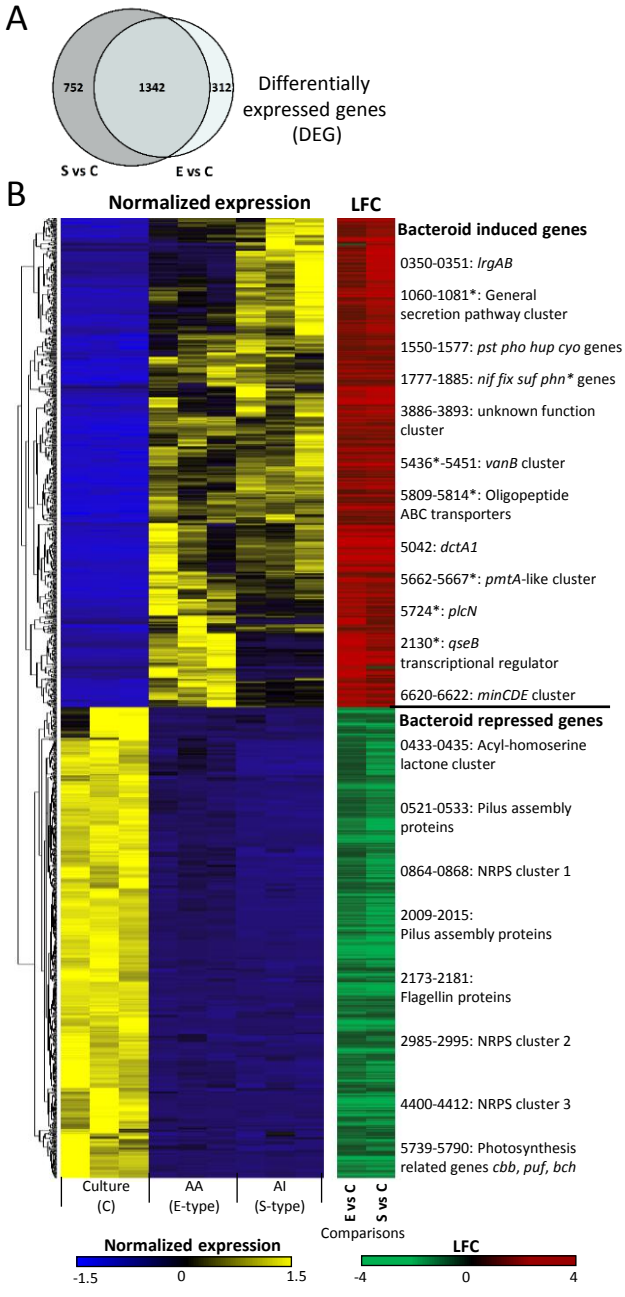


Fig. 4