

Developmental fates and N₂-fixing efficiency of terminally-differentiated versus undifferentiated bacteroids from legume nodules

Carmen Sánchez-Cañizares^{1*}, Raphael Ledermann^{1*}, Joseph McKenna^{2*}, Thomas J. Underwood¹, Marcela Mendoza-Suárez¹, Rob Green³, Karunakaran Ramakrishnan³, Alison K. East¹, Isabel Webb¹, Charlotte Kirchhelle¹, Beatriz Jorrín¹, Gerhard Saalbach⁴, Euan K. James⁵, Flavia Moreira-Leite⁶, Jason Terpolilli⁷⁺, Philip S. Poole¹⁺

*These authors contributed equally

+ For correspondence contact J.Terpolilli@murdoch.edu.au or

philip.poole@biology.ox.ac.uk

Short title: Bacteroid fates and fixation efficiency in legume nodules

Author affiliations

¹ Molecular Plant Sciences Section, Department of Biology, University of Oxford, Oxford, OX1 3RB, UK

² School of Life Sciences, University of Warwick, Coventry, CV4 7AL, UK.

³ Department of Molecular Microbiology, John Innes Centre, Norwich, NR4 7UH, UK

⁴ Proteomics Facility, Department of Biochemistry and Metabolism, John Innes Centre, Norwich, NR4 7UH, UK

⁵ James Hutton Institute, Invergowrie, Dundee DD2 5DA, Scotland, UK

⁶ Department of Biological and Medical Sciences, Oxford Brookes University, Oxford, OX3 0BP, UK

⁷ Legume Rhizobium Sciences, Food Futures Institute, Murdoch University, 90 South St, Murdoch 6150, Western Australia, Australia

The author responsible for distribution of materials integral to the findings presented in this article in accordance with the policy described in the Instructions for Authors (<https://academic.oup.com/plphys/pages/General-Instructions>) are Jason Terpolilli and Philip S. Poole.

Abstract

Within legume root nodules, rhizobia differentiate into bacteroids, which reduce N₂ into NH₃ for secretion to the plant. Bacteroids may be swollen and terminally differentiated or non-swollen and can regenerate outside nodules. It is unclear why these different endosymbiotic lifestyles exist and whether they differ in symbiotic efficiency. Here, we compared N₂ fixing bacteroids of the near isogenic strains *Rhizobium leguminosarum* bv. phaseoli 4292 (Rlp4292) and *R. leguminosarum* bv. viciae A34 (RlvA34), nodulating *Phaseolus vulgaris* (common bean) and *Pisum sativum* (pea), respectively. The larger bean plants fixed more N₂, but peas fixed 1.6-3-fold more per unit nodule mass. Values per unit volume were similar between bean and pea because bean nodules are 2.7-fold denser (i.e., mass per unit volume). Bean nodules have higher numbers of smaller (~1/5 the volume) bacteroids than peas. Bean bacteroids are denser (i.e., 2.5-fold protein per unit volume) although less closely packed than pea bacteroids (i.e. more space between bean bacteroids). Critically, pea bacteroids, fix N₂ at higher rates versus bean per unit bacteroid protein, as protein expression is skewed towards N₂ fixation and TCA-cycle enzymes. Pea bacteroids infect 1.6 times the percentage of nodule volume of beans (i.e.,

14.2% versus 9.1%). Overall, the increased packing density of pea bacteroids, as well as the bias of their proteome to nitrogenase, associated N₂ fixation processes, and dicarboxylate metabolism, contributes to their greater symbiotic efficiency, which is likely driven by plant antimicrobial peptides.

7 Introduction

8 Rhizobia are Alpha- or Beta-proteobacteria that reduce atmospheric N₂ to ammonia in
9 symbiosis with a wide variety of leguminous plants (González et al., 2019), with the
10 bacteria accommodated inside root nodules. N₂ fixation by legumes is a key component
11 of the biosphere's biologically available nitrogen, with crop legume - rhizobia symbioses
12 providing some 21 Tg annually of nitrogen, while forage and fodder legume – rhizobia
13 symbioses contribute approximately 12–25 Tg (Herridge et al., 2008). Establishment of
14 the symbiosis involves signal exchange, where the bacteria detect flavonoids released
15 from legume roots via the receptor protein NodD, which is the transcriptional activator of
16 a large group of *nod* genes. The encoded Nod proteins (e.g. NodOTNMLEFDABCIJ in
17 the *Pisum sativum* or pea symbiont *Rhizobium leguminosarum* biovar *viciae*) synthesize
18 short oligomers of four or five amino sugars, decorated with a long chain fatty acid,
19 forming lipochitooligosaccharides (LCOs) (Spaink et al., 1991). LCOs are perceived by
20 LysM receptors in the plant, which initiate the common symbiosis (Sym) pathway, the first
21 part of which is activated in the plant by both nodulating rhizobia and mycorrhizal fungi
22 (Gleason et al., 2006; Bozsoki et al., 2020). Although best known for nodulating legumes,
23 rhizobia undergo a complex variation in lifecycle from free-living bacteria in soil, to root-
24 colonizing members of the microbiota through to infection of a legume host, typically via
25 root hairs (Wheatley et al., 2020). The rhizobia grow down root hairs via an infection

1 thread and are eventually engulfed by host cells and surrounded by a plant-derived
 2 symbiosome (also called peribacteroid) membrane (Poole et al., 2018). Plant cells
 3 become densely packed with bacteria, with their endosymbiotic N₂-fixing form known as
 4 bacteroids.

5
 6 Inside legume root nodules, bacteroids shut down expression of many genes,
 7 concentrating on expression of N₂ fixation (*nif* and *fix*) genes (Karunakaran et al., 2009;
 8 Green et al., 2019) and secreting the ammonia they produce to the host, where it is
 9 incorporated into amino acids (Allaway et al., 2000). In return, bacteroids receive a carbon
 10 source in the form of C₄ dicarboxylic acids (fumarate, succinate and malate) to generate
 11 ATP and reductant for N₂ fixation (Lodwig and Poole, 2003). Modelling and ¹³C metabolic
 12 flux analysis indicate that O₂ limitation restricts the decarboxylating arm of the TCA cycle
 13 in bacteroids, which limits ammonia assimilation into glutamate, forcing bacteroids to
 14 secrete this fixed N₂ to the plant (Schulte et al., 2021). Consequently, bacteroids become
 15 symbiotic auxotrophs, dependent on the plant for branched-chain amino acids, even
 16 though they synthesize them as free-living bacteria (Lodwig et al., 2003; Prell et al., 2009).

17
 18 Nodule morphology can vary between legume tribes, and based on differences in size,
 19 shape and infection by rhizobia, several distinct nodule types have been described
 20 (Sprent et al., 2017). Two major groups include the determinate and indeterminate type
 21 nodules, which differ to each other in their morphology and development. Determinate
 22 nodules (e.g. *Phaseolus vulgaris* or common bean and *Glycine max* or soybean) are
 23 typically spherical, lack a persistent meristem (a region of actively dividing cells), and do

1 not show a clear developmental gradient. Indeterminate nodules (e.g. pea and *Medicago*
2 *truncatula*) have an elongated, branched or lobed shape, and possess a persistent apical
3 meristem, with the nodule divided into distinct zones which delineate different
4 physiological states of both legume cells and rhizobia (meristematic, infection, N₂ fixation
5 and senescence zones).

6
7 Inside root nodules at the bacteroid level, legumes of the Invert-Repeat Lacking Clade
8 (IRLC), including pea and *M. truncatula*, can produce hundreds of nodule cysteine rich
9 (NCR) antimicrobial peptides in root nodule cells, causing bacteroids to form multiple
10 copies of their chromosome by endoreduplication. These polyploid endosymbiotic cells
11 are terminally differentiated because they lose their ability to return to free-living growth
12 and regenerate outside of the nodule, as they cannot undergo cell division with multiple
13 copies of their chromosomes (Mergaert et al., 2006; Van de Velde et al., 2010; Wang et
14 al., 2010). Terminally differentiated bacteroids are usually elongated or irregular in shape,
15 forming Y-shaped cells with leaky membranes, that are swollen in size compared to free-
16 living rhizobia, with symbiosomes within infected plant cells occupied by a single
17 bacteroid. In contrast, non-IRLC legumes such as common bean and soybean do not
18 produce NCR peptides, so bacteroids retain a chromosome number of 1-2, have a simple
19 rod shape reminiscent in size to free-living cells, and retain the capacity to resume growth
20 and regenerate outside the nodule, with multiple bacteroids occupying each symbiosome
21 (Maróti and Kondorosi, 2014; Poole et al., 2018).

1 While the evolutionary driver for the development of terminally differentiated and swollen
2 bacteroids is not clear, it seems the attack of antimicrobial peptides in IRLC legumes
3 reprograms bacteroids as ammonia secreting organelle-like organisms (also known as
4 'ammonioplasts'), which may increase the efficiency of N₂ fixation (Oono and Denison,
5 2010; Farkas et al., 2014; Penterman et al., 2014; Montiel et al., 2017; Sankari et al.,
6 2022). In their study of pea (producing terminally differentiated swollen bacteroids) with
7 common bean (non-swollen and undifferentiated bacteroids), Oono and Denison (2010)
8 compared the symbiotic efficiency of both legumes inoculated separately with *Rhizobium*
9 *leguminosarum* bv. *viciae* A34 (RlvA34). However, RlvA34 is a transgenic strain,
10 produced by curing the common bean nodulating strain *R. leguminosarum* bv. *phaseoli*
11 4292 (Rlp4292) of its symbiotic (Sym) plasmid (pRP2) and replacing it with Sym plasmid
12 pRL1J from pea-nodulating strain *R. leguminosarum* 248. RlvA34 therefore carries a
13 symbiosis plasmid compatible for N₂ fixation with peas, but not common bean (Hirsch,
14 1979; Lamb et al., 1982), making measurements of symbiotic efficiency between pea and
15 common bean inoculated with RlvA34 alone, difficult to compare. Therefore, to
16 quantitatively compare efficiency between the distinct bacteroid types in these two
17 legume hosts, we selected the strains Rlp4292 and RlvA34, that efficiently nodulate and
18 fix N₂ in common bean or pea nodules, respectively. The near-isogenic nature of these
19 two strains, differing only in their Sym plasmid, but otherwise sharing a core genome,
20 facilitates direct comparison of nodule and bacteroid N₂ fixation efficiency between
21 common bean and pea. Transcriptomic analysis of the determinate and indeterminate
22 nodules formed by Rlp4292 and RlvA34, respectively, provides detailed bacteroid gene
23 expression data (Green et al., 2019). We also sought to map the genetic, proteomic and

developmental changes from the whole plant to nodule and cellular levels to understand how these different developmental fates alter the symbiotic interaction between legumes and rhizobia and to provide a solid benchmark against which quantitative parameters that distinguish the two bacteroid types and their inherent efficiency could be compared.

Results

Modelling nodulation properties of bean and pea

To develop a coherent model of N₂ fixation in which bacteroid types with different developmental fates are compared, we considered it essential to assess their efficiency at all physiological levels, starting at the whole plant level, then moving to the nodule level, and finishing at the cellular level. By using Rlp4292 and RlvA34 as our model strains, we enabled the closest comparison on nodulation and N₂ fixation between terminally differentiated versus undifferentiated bacteroids, the two primary bacteroid types induced by rhizobia and associated in this study with determinate nodule-forming common bean (*Phaseolus vulgaris*) and indeterminate nodule-forming peas (*Pisum sativum*), respectively (Green et al., 2019). These strains constitute a unique resource because of their genetic similarity, differing only in a small region of their genome – their Sym plasmids. Their high degree of genetic homology provides the best possible comparative system to perform detailed physiological, biochemical and microscopy analyses, that were not achieved in previously published studies to assess N₂ fixation effectiveness. These two strains are effective relative to other characterized rhizobia that infect *P. vulgaris* and *P. sativum* (Table S1) (Mwenda et al., 2018). Likewise, the hosts *P. vulgaris* cv. Tendergreen and *P. sativum* cv. Avola performed similarly relative to other

tested cultivars (Table S1) (Mwenda et al., 2018). Initial measurements of dry weight, nodule number, nodule mass and nodule volume were made on plants at flowering, when nitrogenase activity is at a maximum, and hence a logical point to compare peak efficiency. Bean plants were larger than peas with a 1.9-fold higher shoot dry weight (Table 1). Reflecting their larger size, beans produced on average, a 1.4-fold increase in number of nodules. Although bean nodules were spherical in shape, compared to the cylindrically shaped pea nodules, their nodules were similar in volume to pea nodules (6.07 ± 0.48 v 4.61 ± 0.22 mm³ respectively), so the higher nodule number was reflected in a similar 1.8-fold greater total nodule volume per bean plant. This led to an expectation of a similar increase in total nodule mass in beans relative to peas. However, the nodule dry biomass of beans was 4.7-fold greater than in peas, rather than the expected 1.4-1.8-fold. Thus, the biomass per nodule in beans is 3.4-fold greater than peas. Similarly, the biomass per unit volume is 2.7-fold higher in beans than peas. The startling conclusion is that the density of bean nodules is almost three-times that of pea nodules.

To measure nitrogenase activity, acetylene reduction was conducted on beans and peas at flowering (3.31 ± 0.39 versus 1.92 ± 0.20 μ moles C₂H₂ reduced/h/plant, respectively). The total acetylene reduction per plant was 1.7-fold greater in beans than peas (Table 1), which agrees very well with the 1.8-fold greater nodule volume of beans and the overall 1.9-fold greater biomass of bean plants compared to peas. Thus, after correction for volume, the rates were very similar in the beans versus peas (4.67 ± 0.56 versus 4.81 ± 0.50 μ moles C₂H₂ reduced/h/cm³ nodule, respectively). However, when the rates of acetylene reduction are calculated per unit mass of nodule, the rate in pea is 3-fold higher

(124 ± 16.7 $\mu\text{moles C}_2\text{H}_2$ reduced/h/ μg nodules in pea versus 40.7 ± 2.60 $\mu\text{moles C}_2\text{H}_2$ reduced/h/ μg nodules in bean). As reported above, this is because the density of bean nodules per unit volume is 2.7-fold higher than in peas. If we consider that the nodule density and total mass per nodule reflects investment in resources, then pea nodules are more efficient at fixing N_2 than beans. This raises fundamental questions about why bean nodules are much denser, which was next considered by dissecting the various components of the nodule, particularly bacteroid form, packing and density (Table 1).

Spatial packing of bacteroids within bean and pea nodules

To establish the packing properties of bacteroids in bean and pea nodules, Rlp4292 and RlvA34 were fluorescently tagged with *sfGFP* expressed via a NifA-dependent promoter, to delineate the regions within nodules harboring N_2 -fixing bacteroids. Importantly, by using image analysis of *sfGFP* fluorescence, it was possible to separate the sizeable vacuoles present in pea cells from bacteroids. Longitudinal sections were then used to calculate the total nodule volume and the infected volume (i.e., the total volume of plant tissue colonized by N_2 -fixing bacteroids) (Figure 1, Table 1). The average total volume from eight nodules was 5.40 ± 1.45 mm^3 for beans and 5.16 ± 0.59 mm^3 for peas, which is in good agreement with the volumes 6.07 ± 0.48 mm^3 and 4.61 ± 0.22 mm^3 , determined by volume displacement of nodules (Table 1). The average volume of infected plant tissue was further calculated as 1.36 ± 0.40 mm^3 or 24% of the total nodule volume for beans, and 1.41 ± 0.23 mm^3 or 26% of the total nodule volume for peas, indicating that both types of nodules contain a similar volume of infected tissue despite their different morphology and organization. Having calculated the volume of a nodule occupied by

infected plant tissue, the next step was to determine how much of the infected nodule volume was composed of bacteroids rather than plant cytoplasm. Our approach below was to use serial sectioning of nodules.

Serial sectioning reveals the three-dimensional morphology and fine scale packing of bacteroids

To determine the 3D-shape of bean and pea bacteroids in legume root nodules, we used Serial Block Face-Scanning Electron Microscopy (SBF-SEM) which has previously been used in plant cell biology to determine the organization of the plant vacuole during auxin induced growth repression (Scheuring et al., 2016), to confirm that protein storage vacuoles originate from a remodeled embryonic vacuole (Feeney et al., 2018), and to demonstrate the density of plasmodesmata (Yan et al., 2019). Sequentially captured Z-sections of osmium-labelled bean and pea nodule samples had sufficient contrast to allow SBF-SEM imaging and reconstruction of individual and clustered bacteroids (Figure 2, Video 1), clearly showing the expected fundamental difference in bacteroid morphology between the samples with rod-shaped bean bacteroids, compared to Y-shaped pea bacteroids (Figure 2A & B).

Further analysis and segmentation of the data allowed 3D volume and 2D surface area determinations, showing that bean and pea bacteroids differ substantially in volume ($0.93 \pm 0.02 \mu\text{m}^3$ in bean versus $4.50 \pm 0.09 \mu\text{m}^3$ in pea) and surface area ($6.73 \pm 0.14 \mu\text{m}^2$ in bean versus $21.1 \pm 0.36 \mu\text{m}^2$ in pea). As would be expected, the smaller volume and surface area of bean bacteroids compared to pea bacteroids led to a greater surface

area:volume ratio in bean compared to pea bacteroids (7.44 versus 4.78) (Table 1). To ensure there was no prior intrinsic difference in cell morphology between Rlp4292 and RlvA34 that could explain this variation in bacteroid size, we also measured volume and surface area of free-living cells grown on UMS minimal media with confocal imaging (Table 1). There was no significant difference in either average cell surface area ($8.34 \pm 0.32 \mu\text{m}^2$ for Rlp4292 and $8.68 \pm 0.26 \mu\text{m}^2$ for RlvA34 cells), or volume ($1.66 \pm 0.10 \mu\text{m}^3$ for Rlp4292 and $1.77 \pm 0.44 \mu\text{m}^3$ for RlvA34) between the two strains. Therefore, pea bacteroid cells undergo significant enlargement inside pea nodules, differentiating into Y-shaped bacteroids with volumes almost five-fold greater than rod-shaped bean bacteroids.

While segmenting multiple individual bacteroids in the two datasets (Figure 2), we observed that bean bacteroids appeared to be much less densely packed than bacteroids in pea (Figure 2). This correlates with the observations made when imaging nodules with light microscopy (Figure 1). To quantify the packing density of bacteroids, we segmented every bacteroid in a defined volume of the plant cell (Figure 2C), allowing us to determine the ratio of plant cell volume occupied by bacteroids (Table 1). This showed that bean bacteroids occupy $36 \pm 0.9\%$ and pea bacteroids occupy $52 \pm 2.3\%$ of the volume of an infected cell. Applying the average volume of an infected plant cell determined above, a bean nodule has $0.49 \pm 0.16 \text{ mm}^3$, or 9.1% of the total nodule volume occupied by bacteroids, while the value for a pea nodule is almost 1.5-fold higher at $0.73 \pm 0.15 \text{ mm}^3$, or 14.2% of the total nodule volume occupied by bacteroids. However, as was shown above, bean bacteroids are almost five-times smaller in volume than pea bacteroids.

Taking these values into account allowed the calculation of the number of bacteroids per nodule for each nodule type. This showed that bean nodules are occupied by more than three times the number of bacteroids than pea nodules, with $(5.26 \pm 1.56) \times 10^8$ bean bacteroids versus $(1.63 \pm 0.26) \times 10^8$ pea bacteroids (Table 1). So, although bean and pea nodules are equivalent in volume, bean nodules carry a greater number of smaller bacteroids that are less closely packed within the nodule than in pea, a difference which is supported by the segmentation data in Figure 2C.

Nodule and bacteroid protein density

To obtain an accurate measurement of the protein content of a bacteroid, it was imperative to obtain a pure sample of bacteroids devoid of plant tissue. This was achieved by Percoll gradient centrifugation and protein measurement of purified bacteroids. At this stage, we realized that bean bacteroids are much denser than pea bacteroids, as they migrated further into Percoll gradients (Figure S1). To normalize the protein contents, bacteroid counts from the same purified bacteroid samples were obtained by flow cytometry. This showed that bean bacteroids contained almost half the protein content of pea bacteroids (0.32 ± 0.02 pg protein/bacteroid versus 0.62 ± 0.03 pg protein/bacteroid). However, with almost 5-fold difference in volume between the two bacteroid types, the protein density within bean bacteroids is almost 2.5-fold higher than pea bacteroids (0.34 ± 0.02 versus 0.14 ± 0.01 pg bacteroid protein/ μm^3 bacteroid). Furthermore, as the volume of the infected zone containing bacteroids in bean nodules is smaller than that of pea nodules, the density of bacteroid protein per nodule is also almost 2.5-fold higher in beans than in peas (340 ± 20.5 versus 137 ± 6.53 μg bacteroid protein/ mm^3 nodule, Table

1) Thus, relative to pea bacteroids, bean bacteroids have a higher protein density, and by extension, a greater protein density per infected nodule volume. The key question we next sought to address was whether these differences in bacteroid volume, packing and protein densities have any impact on the efficiency of symbiotic N₂ fixation.

N₂ fixation at all scales

While ARA provides an indirect estimate of nitrogenase activity by measuring ethylene production, ¹⁵N₂ labelling methods offer a direct and more accurate quantification of nitrogen fixation by tracking the incorporation of labeled nitrogen into plant tissues. Due to its precision, the ¹⁵N₂ approach is considered more reliable for estimating nitrogen fixation across different environmental and physiological conditions. Therefore, to compare N₂ fixation rates between Rlp4292-inoculated beans and RlvA34-inoculated peas, ¹⁵N₂ labelling assays were conducted on beans and peas at flowering time, as per the earlier ARA to assess nitrogenase activity. On a whole plant basis, the larger bean plants clearly fixed more N₂ than the smaller pea plants, with a 3.26-fold higher ¹⁵N₂ fixation rate (Table 1), consistent with the increased nitrogenase activity in beans measured earlier by ARA. Similarly, when expressed on a per nodule volume basis, this difference was nullified (1.22 ± 0.14 versus 1.12 ± 0.09 $\mu\text{mol N}_2\text{fixed/h/cm}^3$ nodule, Table 1), and the similar volumes of nodule tissue infected by bacteroids in both beans and peas also resulted in a non-significant difference in these rates (5.15 ± 0.66 versus 4.10 ± 0.30 nmol/h/mm^3 infected nodule tissue). This indicates that bean and pea nodules are equivalent in symbiotic N₂ efficiency. Bean nodules are occupied by a much greater number of bacteroids than pea nodules, so individual bean bacteroids on average fix 2.6-

fold less N_2 than RlvA34 bacteroids (13.3 ± 1.70 versus 35.4 ± 2.57 amol/h/bacteroid, Table 1) but when the large difference in volumes between these bacteroid types is considered, bean bacteroids fix almost twice as much N_2 per unit bacteroid volume than pea bacteroids (14.4 ± 1.83 versus 7.87 ± 0.57 amol/h/ μm^3 bacteroid). Crucially though, the higher protein content of bean bacteroids mean they fix less N_2 per unit protein than pea bacteroids (42.2 ± 5.4 versus 56.8 ± 4.1 amol/h/pg protein, $P = 0.043$). Indeed, when rates of N_2 fixation are expressed per nodule dry weight ($\mu mol/h/\mu g$ nodule), bean nodules are 1.6 times significantly less efficient at fixing N_2 than pea nodules (9.76 ± 0.76 versus 15.6 ± 1.26 , $P = 0.001$). Therefore, because protein density per bacteroid and per nodule is higher in bean nodules, pea bacteroids are more efficient at N_2 fixation per unit protein, suggesting a fundamental difference in protein expression under N_2 -fixing conditions between the two bacteroid types. We therefore sort to explore the cause of this difference at the protein level.

Mapping the proteome

Quantitative proteomics was employed to compare protein abundance in bean and pea bacteroids, with a total of 184 and 197 proteins detected in bean and pea bacteroids, respectively. Comparison of the overall abundance between bacteroid types as a percentage of the mass of the proteome (Figure 3A) showed a significant difference between protein populations in bacteroid types ($P = 0.016$, Kolmogorov-Smirnov test). Pea bacteroids expressed their most abundant proteins to a much higher level, with 50% of the proteome made up of the top 18 proteins detected, in comparison to bean, where 30 proteins made up the first 50% of the proteome.

Bacteroid N₂ fixation is dominated by C₄-dicarboxylate metabolism, which bacteroids receive as a carbon source from the host plant. At the individual cell level, pea bacteroids showed greater abundance in dicarboxylate metabolizing enzymes NAD⁺-dependent malic enzyme (Dme, 3.24-fold), malate dehydrogenase (Mdh, 9.44-fold) and phosphoenolpyruvate carboxykinase (PckA, 1.42-fold) enzymes compared to bean (Table 2). Similarly, pyruvate dehydrogenase (PdhABC) and the enzymes of the decarboxylating arm of the TCA cycle (citrate synthase, GltA; aconitase, AcnA; isocitrate dehydrogenase; Icd, 2-oxoglutarate dehydrogenase; SucAB; succinyl CoA synthetase, SucCD), which are known to be highly active in N₂-fixing bacteroids, were all increased in abundance in pea bacteroids relative to bean. Crucially, a much greater number of proteins constituting the nitrogenase enzyme complex (Fe protein NifH and Mo-Fe proteins NifD and NifK) were present in pea bacteroids, as well as a greater than 2.6-fold increase in essential electron transfer proteins FixA and FixC compared to Rlp4292, and alpha and beta subunits of the ATP synthase (Table 2). Therefore, pea bacteroids support higher rates of N₂ fixation, as they express a greater abundance of nitrogenase enzyme and supporting proteins, on an individual cell basis, compared to bean bacteroids. Expressing the absolute protein abundance values per unit bacteroid volume, dramatically nullifies differences between pea and bean bacteroids, with the numbers of proteins (including NifH, NifD, FixA) very similar, while others such as Dme, FixC and PdhABC, are now more abundant in bean bacteroids. This change reflects the very large difference (4.5-fold) in volume between pea and bean bacteroids. Crucially, when these proteins are expressed per unit bacteroid protein, the abundance of NifHDK and TCA

cycle enzymes is significantly higher in pea relative to bean bacteroids (Table 2). Thus, the pea bacteroid proteome, relative to that of bean, is biased towards production of nitrogenase and supporting enzymes. This may be a key outcome of endoreduplication of pea bacteroids by host plant NCR peptides. Absolute levels of individual proteins in bacteroids were also compared to their transcript levels measured by RNA-Seq (Green et al., 2019). For bean bacteroids, normalized RNA counts show a moderate positive correlation with mass of protein per bacteroid, with a Pearson correlation coefficient of $r = 0.66$ and $R^2 = 0.44$ (Fig. 3B). In contrast, for pea bacteroids, the correlation is stronger, with $r = 0.84$ and $R^2 = 0.71$ (Fig. 3C). Thus, there is a reasonable correlation between transcription and protein abundance in both types of bacteroids.

Discussion

In this work we address the question of why legumes have evolved to produce terminally differentiated bacteroids. Previous studies had already shown that terminal differentiation enhances N_2 fixation efficiency (Oono and Denison, 2010; Kazmierczak et al., 2017; Montiel et al., 2017; Lamouche et al., 2019; Chen Wen et al., 2023). However, these earlier investigations had some significant limitations, including the use of suboptimal bacterial strains, non-physiological growth conditions, or comparisons across plant species with inherently different physiologies. Therefore, we revisit this evolutionary question with an optimized comparative system using two extremely close strains of rhizobia that nodulate bean and pea, supported by comprehensive physiological, biochemical, and microscopic analyses not previously applied in this context. This comparative study highlights how the differentiation of rhizobia into bacteroids in root

1 nodules leads to substantial changes in their physiology and metabolism, altering their
2 symbiotic efficiency. Overall, we show that pea bacteroids relative to bean fix more N_2 on
3 an individual basis. This is accompanied by a much larger bacteroid size, which partly
4 explains their increased N_2 fixation. The nitrogenase complex and supporting enzymes
5 are also more abundant in pea bacteroids (normalized by protein), with the proteome
6 biased towards these proteins. This was also observed in a comparative study between
7 pea (*P. sativum* cv. Frisson) and lentil (*Lens culinaris* cv. Magda) nodules from plants
8 inoculated with RlvUPM791 (Durán et al., 2021), thus confirming the specialization of
9 these endosymbiotic cells for the conversion of atmospheric nitrogen into ammonia. From
10 transcriptomic analysis of pSym encoded genes from Rlp4292 bean and RlvA34 pea
11 bacteroids, there is stronger upregulation of a smaller number of genes on pRL1JI (pea)
12 than on pRP2 (bean) (Green et al., 2019). While the *nif* and *fix* gene clusters are > 10-
13 fold upregulated in both, in pea almost every other pRL1JI gene is downregulated > 2-
14 fold (and some > 10-fold). This is not the case for beans where expression of most pRP2
15 genes is unaffected, with only a very small proportion down-regulated > 2-fold. In
16 comparing the whole genomes, a similar picture emerges, more genes are down-
17 regulated and fewer genes are up-regulated in pea than in bean bacteroids (Green et al.,
18 2019). This transcriptomic data is in good agreement with the proteomic data presented
19 in this study; in RlvA34 fewer genes are up-regulated and more genes are down-regulated
20 which would lead to a lower number of different proteins in pea bacteroids.

21 Pea bacteroids are more densely packed within infected plant cells than bean bacteroids,
22 although bean nodules harbor three-times more bacteroids than pea nodules. Bean
23 nodules are 2.7-fold the density of pea nodules, partly because bean bacteroids are much

denser than pea bacteroids. However, as bean and pea bacteroids respectively only make up 9.1 % and 14.2 % of the nodule volume, much of the increased density of bean nodules must be due to the inherent density of plant tissue. The striking change in pea nodules is not only the much larger cell volume of individual bacteroids but also that the packing and distribution of bacteroids leads to a 1.6-fold greater volume of infected nodule volume (i.e. 14.2% versus 9.1%). When the greater density of bean nodules is combined with the bias of the pea bacteroid proteome to nitrogenase and associated N₂ fixation proteins, as well as dicarboxylate metabolism, the greater efficiency of pea versus bean nodules can be explained.

Based on a substantially robust and integrated dataset, our findings align with earlier conclusions. For example, in a characterization of symbiotic root nodules performed with *Bradyrhizobium arachidis* strain CCBAU051107, a strain which can form both determinate dalbergoid-type nodules on peanut (*Arachis hypogaea*) differentiating into swollen bacteroids, and indeterminate nodules on the woody shrub *Sophora flavescens* (non-swollen bacteroids), the authors observed the N₂ fixation activity per unit mass of peanut nodules was 3-fold higher than that of *S. flavescens* (12.06 ± 2.04 versus 3.81 ± 0.58 nmol/h/mg, respectively) (Chen Wen et al., 2023). The expression levels of N₂-fixing genes measured by RNA-Seq, were 2- to 5-fold greater in bacteroids of peanut nodules than those of *S. flavescens*, suggesting that differentiated bacteroids are more efficient (Chen Wen et al., 2023). A caveat is that this comparison is between nodules on an annual crop (peanut) with those on a woody shrub (*Sophora flavescens*), which are plants with radically different physiologies. Although similar conclusions for increased symbiotic efficiency were also reached comparing peas and common bean in the study

by Oono and Denison (2010), RlvA34 was chosen as the only infecting strain for both pea and bean (Oono and Denison, 2010). This is highly unusual as the RlvA34 Sym plasmid (pRLJ1) is specific for peas and lacks the specific *nod* genes that would allow bean nodulation. In the same study, Oono and Denison (2010) also compared the symbiotic efficiency of strain *Bradyrhizobium* sp. 32H1 inoculated on peanut (*Arachis hypogaea*) and cowpea (*Vigna unguiculata*), both symbiotic hosts forming determinate nodules, but hosting swollen and terminally differentiated versus non-swollen and undifferentiated bacteroids, respectively. The bacteroids of peanut nodules conferred more net host benefit in terms of nodule resource investment (plant growth per gram of nodule growth) and efficiency of N₂ fixation (Oono and Denison, 2010). However, while strain 32H1 is effective at fixing N₂ with peanut, it is not highly effective on cowpea (Sen and Weaver, 1981). Nevertheless, while some aspects of these previous studies have been limited by slightly indirect comparisons, they are suggestive that differentiated bacteroids are more efficient at N₂ fixation than undifferentiated bacteroids.

The use in this work of near isogenic strains of *R. leguminosarum* (Rlp4292 and RlvA34), both of which are fully effective on their respective hosts of bean and pea (Table S1; Mwenda et al., 2018), confers confidence that we are measuring real differences in efficiency, related to fundamentally different developmental fates of bacteroids. They broadly agree with studies on other legumes (Oono and Denison, 2010; Chen Wen et al., 2023) and we have no reason to think that these quantitative measurements and differences would not be applicable to a wide variety of symbioses.

It is known that the endoreduplication and variations in ploidy level of pea bacteroids by NCR-like peptides encoded within the plant genome leads to enlarged

1 bacteroids (Montiel et al., 2017). This difference in efficiency is linked to variations in
2 ploidy level, with highly polyploid (containing multiple sets of chromosomes) spherical
3 bacteroids being more efficient than the elongated ones with lower ploidy levels.
4 Comparisons between terminally differentiated bacteroids with varying genome contents
5 (6C in *Pisum* and 18C in *Vicia*, as reported by Mergaert et al., 2006) raises the question
6 of whether higher ploidy levels correlate with enhanced metabolic capacity or symbiotic
7 performance. These ploidy differences suggest that host-specific modulation of bacteroid
8 differentiation could influence symbiotic efficiency and may represent an adaptive trait
9 shaped by co-evolution. Indeed, studies of the determinate nodules of *Aeschynomene*
10 demonstrate that plant host NCR peptides can induce formation of spherical bacteroids,
11 which are associated with higher N₂ fixation efficiency compared to elongated bacteroids,
12 even though both are terminally differentiated (Lamouche et al., 2019). A protein atlas of
13 *M. truncatula* contained entries for 252 detected and quantified NCR peptides showed
14 their expression patterns changed during the stages of nodule development (Marx et al.,
15 2016). NCR peptides seem to interfere with the bacterial cell cycle regulatory network,
16 with bacteroids displaying a major drop in expression of the CtrA master regulator of the
17 cell cycle and of its cognate regulon (Penterman et al., 2014; Pini et al., 2015). Addition
18 of *Medicago truncatula* NCR peptide NCR247 to *S. meliloti* cultures slowed down or fully
19 inhibited protein synthesis in a concentration-dependent manner (Tiricz et al., 2013). In
20 bacteroids, protein synthesis continues, although the complexity of the proteome is
21 reduced; with a large portion of proteins of <30 kDa reduced and others like nitrogenase
22 and GroEL chaperone produced in high quantities (Farkas et al., 2014). In considering
23 different legume species and their microsymbionts, such as *M. truncatula* and

Sinorhizobium strains, there is often large strain variation in bacterial ploidy and efficiency of N₂-fixation (Kazmierczak et al., 2017). In our work we sort to avoid these differences by choosing highly efficient pairs of legume host and microsymbiont, but of course compatibility between a rhizobial strain and a legume is a major determinant of N₂-fixation efficiency.

In RNA-Seq and microarray experiments examining bacteroids at 28 dpi (pea) or 35 dpi (bean), there is general down-regulation of translation compared to free-living bacteria (Karunakaran et al., 2009; Green et al., 2019). However, there is specific down-regulation of genes associated with cell division only in peas, consistent with endoreduplication (without cell division), a consequence of NCR peptide exposure (Green et al., 2019).

Overall, this work provides insights into the functional physiology of different bacteroid types and offers compelling evidence for the adaptive advantage conferred by terminal differentiation. Terminal bacteroid differentiation seems to be a plant-driven strategy to control the endosymbiont. By understanding why NCR peptide-producing legumes program bacteroids for greater N₂-fixing efficiency we can better inform efforts to improve symbiotic performance in crop species, with significant implications for sustainable agriculture and reduced reliance on synthetic fertilizers.

Materials and Methods

Bacterial strains and growth conditions

Escherichia coli was grown in LB (Sambrook et al., 1989) at 37°C with antibiotics added at the following concentrations if appropriate: Gentamicin 10 µg/ml, Streptomycin 50

1 $\mu\text{g/ml}$, Spectinomycin 50 $\mu\text{g/ml}$, Ampicillin 100 $\mu\text{g/ml}$. For *E. coli* ST18, 50 $\mu\text{g/ml}$ of
 2 δ -aminolevulinic acid was supplemented (Thoma and Schobert, 2009) *R. leguminosarum*
 3 strains (Table 3) were grown at 28°C in TY medium (Johnston and Beringer, 1975) or
 4 UMS minimal medium (Terpolilli et al., 2016) with appropriate carbon and nitrogen
 5 sources at 10 mM unless otherwise stated. Antibiotics were added at the following
 6 concentrations where needed: Rifampicin 50 $\mu\text{g/ml}$, Streptomycin 500 $\mu\text{g/ml}$, Gentamicin
 7 20 $\mu\text{g/ml}$.

9 **Plasmid and strain construction**

10 To tag both RlvA34 and Rlp4292 wild type strains with a constitutive *mCherry* and a NifA-
 11 dependent *sfGFP*, we first assembled pOPS1699 via Golden Gate assembly from the
 12 following plasmids: pOGG037, pOGG043, pOGG157, and pUC18R6K-miniTn7T. This
 13 plasmid contains a mTn7 module with *sfGFP* expressed by a synthetic *nifH* promoter
 14 (P_{snifH}). Subsequently, pOPS1699 was linearized with *NsiI* and a fragment containing
 15 P_{J23104} -*mCherry*, previously amplified from pOPS1015 using primers *oxp5931* and
 16 *oxp5932*, was inserted via HiFi assembly, so that is expressed in opposite direction to
 17 P_{snifH} -*sfGFP*. The resulting plasmid pOPS2031 was mobilized into both recipient strains
 18 in triparental matings using pTNS3 as a helper plasmid for mTn7 transposition. For
 19 confocal imaging of free-living cells, we built plasmid pOPS0875 using Golden Gate
 20 cloning to express the *mCherry* module under the constitutive P_{neo} promoter in the
 21 medium-copy, broad-host range pOGG024 BEVA vector (Geddes et al., 2019). This
 22 plasmid was introduced into RlvA34 and Rlp4292 by conjugation.

Plant inoculation and growth

Pisum sativum cv. Avola and *Phaseolus vulgaris* cv. Tendergreen seeds were sterilized in 70% ethanol and 2% sodium hypochlorite followed by extensive washing in distilled water. Seeds were pregerminated on 0.8% w/v agar plates at 28°C in the dark for two days. Seeds were then placed into 1 l beakers filled with washed medium grade vermiculite for peas, and 2 l beakers with fine grade vermiculite for beans. For peas, 400 ml of N-free rooting solution (Poole et al., 1994) was added to each pot prior to autoclaving and for beans, 400 ml of Broughton and Dilworth's nutrient solution (1971). Strains for inoculation were grown on TY slopes with appropriate antibiotics added. Cells were then resuspended in UMS and washed once and OD₆₀₀ was adjusted to 0.01, corresponding to 1×10^7 cfu/ml. A 1 ml suspension was used to inoculate each pot, with one plant per pot, and 11 plants per inoculated treatment. Plants were grown for 28 (*P. sativum*) or 35 (*P. vulgaris*) days at 21°C with a 16 h photoperiod. *P. vulgaris* were watered weekly with 100 ml of sterile Broughton and Dilworth's nutrient solution.

Acetylene reduction and ¹⁵N₂ assays

Acetylene reduction was determined on plants incubated in 95% air: 5% acetylene for 1 h in 250 ml Schott bottles, as previously described (Allaway et al., 2000). Nodules were then excised, counted and weighed fresh. Finally, root and shoot were separated, dried at 70°C in a dry-heat incubator for 3 days then weighed. For ¹⁵N₂ fixation assays were conducted as previously published (Lodwig et al., 2003). Briefly, plants were incubated in 80% air: 20% ¹⁵N₂ gas (98%+, Cambridge Isotope Laboratories, Inc. USA) for 2 h in 250 ml Schott bottles. The incorporation of ¹⁵N₂ was determined separately for roots,

shoots, and nodules. Tissue samples were then dried overnight at 70°C, and dry weights were determined for each fraction separately. For each sample, 4 mg of dried tissue was then crushed and sent for combustion and mass spectrometric analysis. Total nitrogen and $^{15}\text{N}_2$ enrichment were measured in a continuous flow isotope ratio mass spectrometer (James Hutton Institute, UK). From the $\delta^{15}\text{N}$ (‰) values obtained separately for each plant fraction, BNF enrichment or δ APE (atom% enrichment) was calculated for each sample by subtracting the ^{15}N natural abundance (0.36782) from the amount of atmospheric N_2 fixed by each plant fraction (^{15}N Atom %). Then, the N_2 fixed/h for each fraction was calculated as $((\delta \text{ APE})/100) \times \text{total N} \times 5/4$, to account for the 20% $^{15}\text{N}_2$ gas injected and for the two nitrogen atoms and the two hours of $^{15}\text{N}_2$ incubation. These shoot, root and nodule values were added together to give the final $\mu\text{g N}_2$ fixed/h/plant. The Atom % values for both the shoot and root fractions in the water controls were 0.368, matching the $^{15}\text{N}_2$ natural abundance and confirming that the assay worked well.

Bacteroid isolation and protein measurement

Harvested nodules were crushed with a pestle in a mortar with 2 ml of mulching buffer (100 mM potassium phosphate buffer pH 7.4, 300 mM sucrose, 2 mM MgCl_2). Debris was filtered through two layers of gauze. The filtrate was loaded onto a 30 ml preformed Percoll gradient (43% Percoll, 57% Percoll buffer consisting of 232.4 mM potassium phosphate buffer pH 7.4, 697 mM sucrose, 4.65 mM MgCl_2 , centrifuged at 35 000 x g , 30 min, 4°C in a swing-out rotor) and centrifuged at 35 000 x g , 4°C for 20 min. To estimate bacteroid densities, identical gradients were loaded with colored beads of defined densities (Cospheric LLC, Somis, CA, USA). Bacteroid fractions were aspirated and

loaded on a fresh Percoll gradient and processed as above. Aspirated pure bacteroid fractions were washed twice in resuspension buffer (50 mM potassium phosphate buffer pH 7.4, 2 mM $MgCl_2$) before further use. Bacteroids were subsequently lysed using BugBuster (Sigma Aldrich) according to the manufacturer's recommendations, and protein was measured using the Bradford assay (Sigma Aldrich).

Flow cytometry analysis

An Amnis® Cellstream® (Luminex Ltd.) with autosampler and equipped with a 488 nm laser to excite sYFP2 was used for flow cytometer. Flow rates were set to low speed/high sensitivity (3.66 μ l/min) and the denser population in FSC-SSC was defined as "nodule". Pea and Bean bacteroid samples were diluted 100-fold and 1,000-fold respectively. A total of 20,000 events in the nodule population were counted for each sample. Using Cellstream® Analysis 1.3.384 software the nodule population was afterwards gated based on FSC (threshold > 0) and the aspect-ratio of SSC (threshold >0.4) defining the "singlets" population. Then singlets events were gated based on their fluorescence emission of sYFP2 emission was detected at 528/46 and the events above 3000 fluorescent units (FI) units were defined as the "yellow" population. Three technical replicates of six independent bacteroid isolations were processed and the number of events $\cdot$ mL⁻¹ recorded.

Confocal imaging

Five independent free-living cultures with cells expressing constitutive mCherry on pOPS0875 were grown in UMS with 0.1% agarose, so they could be fixed for confocal

imaging, and normalized to the same OD_{600nm}. Nodules were placed in a molten 8% w/v solution of agarose. After the agar solidified, cubes with one nodule were sectioned longitudinally to 100 µm thickness using a Leica VT1200S vibratome. Nodule sections were imaged with a Zeiss LSM 880 Airy Scan confocal microscope and the ZEN Black software. mCherry was visualized using a 561 nm laser and emission detected between 598 and 649 nm, sfGFP was visualized using a 488 nm wavelength laser and emission detected between 498 and 562 nm.

Image analysis and volume calculations

Fluorescent nodules were analyzed using Fiji/ImageJ 2.9.0/1.53t. To establish the total infected area for each nodule we used the NifA-dependent sfGFP signal (green channel). To enable comparable thresholding, we processed each image as follows: To subtract background noise, first, we subtracted 20 from each pixel intensity and then subtracted background using a sliding paraboloid of a 200 pixel radius. Then we enhanced local contrast (CLAHE; blocksize 500, histogram bins 256, max. slope 6.00) to account for uneven signal intensity. Lastly, we used a median filter (3.00 px) to smooth edges before thresholding (otsu with manual correction where needed) the infected area. To determine the infected zone (i.e. the inner part of each nodule, excluding epidermis and cortex, where infected cells are found), we used the above determined infected areas to manually establish the borders. For the whole nodule area, the white light channel was used with inverse thresholding and manual correction as needed.

For bean nodules, the following parameters were obtained: The total area of each nodule section (A_{nod}) and both maximum and minimum Feret's diameter (F_{dmax} and F_{dmin}); the

total infected area as determined by green-fluorescent areas containing *sfGFP*-expressing bacteroids (A_i); the area of the infected zone (A_{iz}) and both maximum and minimum Feret's diameter. Being approximately spherical, we calculated the volumes of both the whole bean nodule and the infected zone as follows:

$$V = \frac{4}{3}AF_d$$

Both maximum and minimum volumes (using F_{dmax} and F_{dmin}) were calculated and the mean volume was taken as our final volume. The fractional colonization density (f_c) of infection in the infected zone was calculated using the ratio A_i/A_{iz} which was then used to calculate the total infected volume per nodule.

For pea nodules, we assumed that both the infection zone and the nodule as a whole took a cylindrical form, with an uninfected spherical cap at the distal end of the nodule. The infected area, infection zone, and total nodule area were determined as described for bean nodules. The cylindrical body of the nodule and the cap were then dealt with separately. For the cap, we treated the image as a 2D segment of the spherical cap. The chord length (c) and the sagitta (h) of the circular segment were measured. The volume of the uninfected spherical cap (V_c) was calculated as follows:

$$V_c = \frac{1}{6}\pi h(3c^2 + h^2)$$

For the cylindrical body of the nodule and the infected zone within we calculated the length of the cylinders (l_c) by using the skeletonize function of Fiji to create a mid-line of the nodule and then measuring this from end to end. Using the determined areas (in case of

the whole nodule body excluding the cap area) the average width (d) of each was calculated ($d=A/l_c$). The volumes of both the nodule body cylinder (V_b) and the infected zone (V_{iz}) were calculated then as follows:

$$V = \frac{A d \pi}{4}$$

Total nodule volume (V_n) was then calculated ($V_n=V_c+V_b$). The total infected volume (V_i) was again calculated using the ratio of infected tissue within the infected zone ($V_i=V_{iz}*A_i/A_{iz}$).

Serial Block Face Scanning Electron Microscopy (SBF-SEM) imaging and reconstruction

Selected nodules were immersed in 8 % w/v agar and these agar blocks containing fixed and embedded nodules were mounted on aluminum pins using conductive epoxy resin, sputter coated with gold and imaged on a Zeiss Merlin VP compact coupled with a 3view2XP chamber and an OnPoint backscattered electron detector (both from Gatan/Ametek, Pleasanton, CA). Samples were imaged by SBF-SEM, using the following imaging conditions: 1.8kV accelerating voltage, 100% focal charge compensation, 20 μ m aperture, 7 nm pixel size, 3-4 μ s pixel time, 100 nm section thickness. Data were assembled into 3D volumes and aligned using ETomo (part of the IMOD software package, v4.12.47; (Kremer et al., 1996)), and three-dimensional reconstruction was performed using 3dmod (also part of IMOD). Bacteroids were segmented manually by tracing structures in each slice, and surface meshing and capping were performed in 3dmod. For packing data, individual volume datasets were trimmed from main data sets

1 at defined volumes (X, 8 μ m, Y 8 μ m by Z 2 μ m). Within these individual packing datasets
2 every bacteroid was segmented to determine the percentage of plant cell volume
3 occupied by bacteroids. Graphs were generated in Graphpad Prism (v10.0.1) and the
4 figure generated using Inkscape (v1.0.2).

6 **Quantitative proteomic analysis**

7 Bacteroids were obtained from nodules of bean and pea plants inoculated with Rlp4292
8 and RlvA34, respectively, as described above. Bacteroids were isolated from nodules of
9 three plants each of bean and pea and processed as three independent replicates.
10 Bacteroid soluble proteins were extracted as described (Schulte et al., 2021), precipitated
11 in acetone in the presence of NaCl and the resulting pellets dissolved in 50 μ l of 8 M
12 urea/0.1 M Tris (pH 8). A 5 μ l aliquot was used to quantify protein by Bradford assay. The
13 remainder of all samples was adjusted to 4.6 mM dithiothreitol (DTT) by adding 0.9 μ l of
14 235 mM DTT and the samples were incubated at 37°C for 30 min. Then 0.8 μ l of 545 mM
15 iodoacetamide was added to final concentration of 9 mM and samples incubated for 30
16 min at room temperature. Two μ l (1 μ g) of trypsin/Lys-C mix (Promega) was added and
17 the samples incubated at 37°C for 5 h, diluted one in six with 0.1 M Tris (pH 8) and
18 digested at 37°C for a further 16 h. The reaction was stopped by adding 20% (v/v)
19 trifluoroacetic acid to a final concentration of 1%.

20 For loading to the LCMS analysis, aliquots were prepared from each sample containing
21 380 ng of digested bacteroid protein. Each aliquot was spiked with 12.5 fmole Hi3 Phos
22 B standard (Waters, 186006011) for absolute quantification. Each biological replicate was
23 run three times on a nanoLCMS system comprising an LTQ-Orbitrap (Thermo Fisher

Scientific, Hemel-Hempstead, UK) and a nanoACQUITY UPLC system (Waters, Manchester, UK as described previously (Schulte et al., 2021) except a longer gradient was used from 3% to 40% acetonitrile over 100 min.

All raw files were processed individually in MaxQuant (1.5.0.0) with protein sequence databases for RlvA34 and Rlp4292 (Green et al., 2019). The PHS2_RABIT (P00489) sequence, from which the Hi3 standard peptides are derived, was added to the databases, along with the MaxQuant contaminants database. MaxQuant parameters were set to: trypsin as enzyme with no missed cleavages; precursor and fragment tolerances 6 ppm/0.5 Da; carbamidomethyl on C as fixed; and oxidation on M and acetylation of protein N-terminus as variable modification. The MaxQuant output tables modificationSpecificPeptides.txt were merged for RlvA34 and Rlp4292 and used to calculate absolute protein abundances using the top3 approach and the Hi3 standard (see above). An in-house script (top3) was applied to the merged modificationSpecificPeptides.txt table which calculates the average of the top three unique peptides for each protein. No missed cleavage was allowed, and only unmodified peptides and peptides with oxidized methionine were used. If the same peptide was found unmodified and oxidized, the sum of both was used. Data were normalized using the RegrRun procedure (Kultima et al., 2009). The mass spectrometry proteomics data have been deposited to the ProteomeXchange Consortium via the PRIDE (Perez-Riverol et al., 2022) partner repository with the dataset identifier PXD058609.

To calculate the absolute abundance of each protein as fmole in the loaded sample, each protein abundance was divided by the Hi3 standard abundance and multiplied by 12.5 resulting in the protein abundance value (A_b). The results were transformed into protein

copy numbers (C_p) based on the Avogadro constant (N_A) using 6.0221413×10^8 copies per fmole. Each 380 ng loaded sample of bacteroid protein was calculated to have been derived (B_D) from 1,202,531 bean and 608,974 pea bacteroids, based on protein contents of 0.32 pg protein/bacteroid for bean and 0.62 pg protein/bacteroid for pea (Table 1). Copies per bacteroid were subsequently determined with the formula:

$$C_p (\text{bacteroid}) = \frac{C_p}{B_D}$$

Copies per bacteroid volume were calculated by:

$$C_p (\text{bacteroid vol}) = \frac{C_p (\text{bacteroid})}{B_{vol}}$$

Where B_{vol} = bacteroid volumes of 0.93 and $4.50 \mu\text{m}^3$ /bacteroid for bean and pea (Table 1). Copies per nodule were calculated by:

$$C_p (\text{nodule}) = \frac{C_p (\text{bacteroid})}{B_{nod}}$$

Where B_{nod} = bacteroids per nodule of 5.28×10^8 and 1.63×10^8 bacteroids for bean and pea (Table 1). Copies per nodule volume were calculated by:

$$C_p (\text{nodule volume}) = \frac{C_p (\text{nodule})}{Nod_{vol}}$$

Where Nod_{vol} = Mean nodule volume of 6.07 and 4.61 mm^3 /nodule for bean and pea (Table 1).

Copies per unit protein were calculated by:

$$C_p (\text{unit protein}) = \frac{C_p (\text{bacteroid})}{B_p}$$

Where B_P = bacteroid protein content of 0.32 and 0.62 pg/bacteroid for bean and pea (Table 1). Welch's two-sample T test was used to identify proteins common to both RlvA34 and Rlp4292 bacteroids whose expression was significantly different between each bacteroid type ($P < 0.05$). The mass of each protein detected in Rlp4292 and RlvA34 bacteroids was calculated and used to rank proteins in RlvA34 and Rlp4292 bacteroids and compare the percentage proteome by mass to proteins ranked from most to least abundant.

RNA-Seq data from bacteroid of *R. leguminosarum* bv. *phaseoli* 4292 and *R. leguminosarum* bv. *viciae* A34, isolated from bean and pea plants respectively (Green et al., 2019), were used to calculate normalized counts using the EDGE-pro (Magoc et al., 2013) and DESeq2 (Love et al., 2014) software packages. Only genes present in the proteome were further analyzed. Pearson correlation between proteome and transcriptome for bean and pea bacteroid was performed using Prism (version 10.4.0).

Accession Numbers

Sequence data from this article can be found in the GenBank data library under accession numbers NZ_AQZR000000000.1 for *R. leguminosarum* bv. *phaseoli* 4292, and ARRT000000000.1 for the symbiotic plasmid of pRLJ1 in *Rhizobium leguminosarum* bv. *viciae* 248, with locus tags for protein sequence database used in this study listed in Green et al., (2019).

Acknowledgments

We would like to thank the center for Bioimaging at Oxford Brookes University for access to and use of the SBF-SEM including support during image analysis. We would also like to thank the late Professor Chris Hawes for advice at the initial stages of the SBF-SEM aspects of this project.

Funding

This work was supported by the Biotechnology and Biological Sciences Research Council [grant numbers BB/J014524/1, BB/J007749/1, BB/T001801/1, BB/T006722/1, BB/W006219/1] and The Leverhulme Trust [grant RPG-2019-246] awarded to Philip Poole; the Grains Research and Development Corporation of Australia Grants [UMU2407 and UoA2312] to Jason Terpolilli; The Royal Society [URF\R1\221030] to Carmen Sánchez Cañizares; the Swiss National Science Foundation Postdoc.Mobility Fellowship [number 183 901] to Raphael Ledermann; a BBSRC Discovery fellowship to Joseph McKenna [BB/X01651/1]; BBSRC PhD studentship [BB/T008784/1] to Thomas Underwood.

Data Availability

Proteomic Data are available via ProteomeXchange with identifier PXD058609. Flow cytometer data is available at <http://flowrepository.org>, experiment codes are FR-FCM-Z623 and FR-FCM_Z624 for pea and bean bacteroids respectively.

Author Contributions

P.S.P and J.T. conceived and designed the research work; C.S.C., R.L., T.J.U., R.G., K.R., I.W., B.J., J.T. carried out the research experiments and analyzed the data; J.M. and F.M.L. performed the SBF-SEM imaging and reconstruction; M.M.S. and C.K. did the confocal free-living experiments and imaging analysis; G.S. carried out the proteomic computational analysis; C.S.C., R.L., J.M., A.E., J.T. and P.S.P wrote the paper.

Tables

Table 1. Comparison of *Phaseolus vulgaris* (bean) inoculated with Rlp4292 versus *Pisum sativum* (pea) inoculated with RlvA34.

Parameter				Bean with Rlp4292	Pea with RlvA34	Fold-change ^a	P value ^b
Plant and nodule comparison							
Plant shoot dry mass (mg/plant)				610 ± 50.2	321 ± 33.2	1.90	<0.001
Plant nodule fresh mass (mg nodules/plant)				1,445.3 ± 94.8	251.40 ± 13.35	5.82	<0.001
Plant nodule dry mass (mg nodules/plant)				81.8 ± 9.4	17.3 ± 1.85	4.72	<0.001
Nodule Number (nodules/plant)				125 ± 10.3	88.6 ± 4.18	1.41	0.006
Nodule fresh mass (mg/nodule)				5.54 ± 0.70	2.56 ± 0.19	2.16	<0.001
Nodule dry mass (µg/nodule)				642 ± 25.1	195 ± 18.4	3.29	<0.001
Plant nodule volume (mm ³ nodule/plant)				709 ± 43	400 ± 6.9	1.77	<0.001
Mean nodule volume (mm ³ /nodule)				6.07 ± 0.48	4.61 ± 0.22	1.32	0.013

Nodule density (mg/cm ³ nodule)	115 ± 13.3	43.2 ± 4.64	2.67	<0.001
--	------------	-------------	------	--------

C₂H₂ Reduction per

Plant (μmol/h/plant)	3.31 ± 0.39	1.92 ± 0.20	1.72	0.007
Plant (μmol/h/g SDW)	5.28 ± 0.25	6.00 ± 0.13	0.88	0.022
Nodule volume (μmol/h/cm ³ nodule)	4.67 ± 0.56	4.81 ± 0.50	0.97	0.862
Nodule dry mass (nmol/h/μg nodules)	40.7 ± 2.60	124 ± 16.7	0.33	<0.001

Nodule infection and packing

Calculated nodule volume (mm ³ /nodule)	5.40 ± 1.45	5.16 ± 0.59	1.05	0.881
Infected nodule volume (mm ³ /nodule)	1.36 ± 0.40	1.41 ± 0.23	0.97	0.917
Proportion infected nodule volume (%)	24.2 ± 1.52	26.2 ± 2.19	0.92	0.473

Bacteroid volume and surface area

Volume (μm ³ /bacteroid)	0.93 ± 0.02	4.50 ± 0.09	0.21	<0.001
Surface area (μm ³ /bacteroid)	6.73 ± 0.14	21.1 ± 0.36	0.32	<0.001
Surface area:volume ratio	7.44 ± 0.06	4.78 ± 0.03	1.56	<0.001
Proportion bacteroids/infected cell (%)	36 ± 0.9	52 ± 2.3	0.69	<0.001
Total bacteroid volume (mm ³ /nodule)	0.49 ± 0.14	0.73 ± 0.12	0.67	0.235

Bacteroid Protein

Protein (pg/bacteroid)	0.32 ± 0.02	0.62 ± 0.03	0.52	<0.001
Number (bacteroids x 10 ⁸ /nodule)	5.28 ± 1.56	1.63 ± 0.26	3.24	0.036
Protein density per bacteroid (pg/μm ³ bacteroid)	0.34 ± 0.02	0.14 ± 0.01	2.42	<0.001
Protein density per nodule (μg/μm ³ nodule)	340 ± 20.5	137 ± 6.53	2.48	<0.001

Free-living cells

Volume ($\mu\text{m}^3/\text{cell}$)	1.66 $\pm$ 0.10	1.77 $\pm$ 0.44	0.94	0.389
Surface area ($\mu\text{m}^2/\text{cell}$)	8.34 $\pm$ 0.32	8.68 $\pm$ 0.26	0.96	0.420
Surface area:volume ratio	5.82 $\pm$ 0.13	5.10 $\pm$ 0.58	1.14	<0.001
$^{15}\text{N}_2$ Fixation rate per				
Plant ($\mu\text{mol/h/plant}$)	1.89 $\pm$ 0.22	0.58 $\pm$ 0.05	3.26	<0.001
Plant ($\mu\text{mol/h/g SDW}$)	1.33 $\pm$ 0.11	1.07 $\pm$ 0.12	1.24	0.143
Nodule dry mass ($\mu\text{mol/h}/\mu\text{g}$ nodule)	9.76 $\pm$ 0.76	15.6 $\pm$ 1.26	0.63	0.001
Nodule volume ($\mu\text{mol/h}/\text{cm}^3$ nodule)	1.22 $\pm$ 0.14	1.12 $\pm$ 0.09	1.09	0.552
Volume infected nodule ($\text{nmol/l}/\text{mm}^3$)	5.15 $\pm$ 0.66	4.10 $\pm$ 0.30	1.26	0.147
Bacteroid ($\text{amol/h}/\text{bacteroid}$)	13.3 $\pm$ 1.70	35.4 $\pm$ 2.57	0.38	<0.001
Bacteroid protein ($\text{amol/h}/\text{pg}$ protein)	42.2 $\pm$ 5.4	56.8 $\pm$ 4.1	0.74	0.043
Bacteroid volume ($\text{amol/h}/\mu\text{m}^3$ bacteroid)	14.4 $\pm$ 1.83	7.87 $\pm$ 0.57	1.83	0.003

Plants harvested at flowering (35 days post inoculation for beans and 28 days post inoculation for pea) at peak nitrogenase activity. Values represent the average plus/minus standard error, with sample sizes per bean and pea treatment, respectively of: 11 each for plant and nodule comparison and C_2H_2 reduction; eight each for nodule infection and packing; 240 and 246 for bacteroid volume, surface area; six for bacteroid protein; 75 and 132 for free-living cells for bean and pea; nine and ten for $^{15}\text{N}_2$ fixation rate for bean and pea.

^aFold-change determined as the ratio of the value for bean divided by pea

^bStudent's unpaired t-Test, with significance set at 0.05

Shoot Dry Weight abbreviated as SDW

Table 2. Selected bacteroid proteins and their fold change relative to Rlp4292 (i.e. A34/4292)

			Fold-Change* (RlvA34/Rlp4292) per:				
Locus Tag	Protein	Product	Bacteroid	Bacteroid volume	Nodule	Nodule volume	Bacteroid protein
Nitrogenase, electron transfer and ATP							
#RHLv8088	NifH	Nitrogenase iron protein	5.49	1.13	1.63	2.15	2.78
#RHLv8087	NifD	Nitrogenase molybdenum-iron protein alpha chain	6.42	1.33	1.92	2.53	3.25
#RHLv8086	NifK	Nitrogenase molybdenum-iron protein beta chain	16.75	3.46	5.18	6.83	8.48
#RHLv8128	FixA	Electron transfer flavoprotein, beta subunit	4.99	1.03	1.47	1.93	2.53
#RHLv8126	FixC	Electron transfer flavoprotein oxidoreductase	2.64	0.55	0.82	1.08	1.34
RHL3232	AtpA	Proton translocating ATP synthase, F1 alpha subunit	3.88	0.80	1.16	1.52	1.96
RHL3230	AtpD	ATP synthase, F1 beta subunit	3.28	0.68	0.98	1.29	1.66
RHL2934	-	Putative GTPases (G3E family)	4.56	0.94	1.34	1.77	2.31
Dicarboxylates and TCA cycle							
RHL1472	Dme	Malic enzyme, NAD ⁺ -dependent	3.24	0.67	0.96	1.26	1.64
RHL3264	Mdh	Malate dehydrogenase, NAD ⁺ -dependent	9.44	1.95	2.82	3.71	4.78
RHL3615	PckA	Phosphoenolpyruvate carboxykinase	1.43	0.29	0.42	0.56	0.72
RHL1082	PdhA	PDH E1 component, alpha subunit	4.05	0.84	1.19	1.57	2.05
RHL1083	PdhB	PDH E1 component, beta subunit	2.48	0.51	0.74	0.97	1.25
RHL1084	PdhC	PDH E2 component, dihydrolipoamide acetyltransferase	2.03	0.42	0.61	0.80	1.03
RHL1087	LpdA	PDH E3 component^, dihydrolipoamide dehydrogenase	3.46	0.71	1.05	1.39	1.75
RHL1075	GltA	Citrate synthase I	21.19	4.38	6.32	8.32	10.73
RHL3358	AcnA	Aconitate hydratase I	16.86	3.48	5.03	6.62	8.54
RHL1435	Icd	Isocitrate dehydrogenase, NADP-dependent	6.42	1.33	1.92	2.52	3.25
RHL3260	SucA	2-OGDH E1 component	13.39	2.77	3.99	5.25	6.78
RHL3258	SucB	2-ODDH E2 component	10.28	2.12	3.13	4.12	5.21
RHL3263	SucC	Succinyl-CoA synthetase, beta subunit	8.93	1.85	2.67	3.51	4.52

- 1 #Locus tag for genes encoded on plasmid pRL1 in RlvA34. Corresponding genes on pRP2 in Rlp4292 are: *nifH*
2 (RHLp7062 and RHLp7131), *nifD* (RHLp7061), *nifD* (RHLp7060), *fixA* (RHLp7092) and *fixC* (RHLp7094).
3 *Values in bold are significantly different ($P<0.05$) according to Welch's t-Test, while those in italics are not.
4 ^Also the E3 Component of 2-OGDH
5 Abbreviations: 2-OGDH, 2-oxoglutarate dehydrogenase; PDH, pyruvate dehydrogenase

1 **Table 3: Strains and plasmids used in this study**

Strain or Plasmid	Relevant genotype or phenotype	Reference
Escherichia coli		
DH5α	supE44 ΔlacU169 (φ80 lacZΔM15) hsdR17 recA1 gyrA96 thi-1 relA2	BRL, Gaithersburg, USA
ST18	Sm ^R , pro thi hsdR ⁺ ; chromosome::RP4-2 Tc::Mu- Kan::Tn7/λpir ΔhemA	(Thoma and Schobert, 2009)
Rhizobium leguminosarum		
4292	Rf ^R , biovar phaseoli wild type	(Lamb et al., 1982)
A34	Sm ^R , biovar viciae wild type	(Downie et al., 1983)
OPS2220	Sm ^R Gm ^R , (A34) pOPS0875	This work
OPS2221	Rf ^R Gm ^R , (4292) pOPS0875	This work
OPS3405	Sm ^R Gm ^R , (A34) Tn7(P _{J23104} - mCherry; P _{snifH} -sfGFP; aacC1)	This work
OPS3411	Rf ^R Gm ^R , (4292) Tn7(P _{J23104} - mCherry; P _{snifH} -sfGFP; aacC1)	This work
Plasmids		

pTNS3	Ap ^R , transposase delivery plasmid for integration of mini-Tn7 cassettes	(Choi et al., 2008)
pOGG037	Sp ^R , pL0M-SC-sfGFP	This work
pOGG043	Sp ^R , pL0M-PU-P _{snifH} (synthetic P _{nifH})	(Mendoza-Suarez et al., 2020)
pOGG157	Sp ^R , pL0M-T-DT16	(Geddes et al., 2019)
pOPS0875	Gm ^R , pneo-mCherry-pOGG024	This work
pUC18R6KminiTn7T	Ap ^R Gm ^R , pUC18R6KminiTn7T	(Choi et al., 2008)
pOPS1015	Ap ^R Gm ^R , pUC18T-miniTn7T-Gm-T1-J ₁₂₃₁₀₄ -RBstd-mCherry-DT16-T0	(Westhoek et al., 2021)
pOPS1699	Ap ^R Gm ^R , pUC18R6K-miniTn7T P _{snifH} -sfGFP-DT16	This work
pOPS2031	Ap ^R Gm ^R , (pOPS1699) P _{J23104} -mCherry, P _{snifH} -sfGFP-DT16	This work

1

2

Figure legends

Figure 1. Volumes of bean and pea nodules. For spherical bean nodules (A), whole nodule volumes were calculated (B) from the areas of cross sections (light purple outline). The volume of the infected zone (blue outline) was calculated analogously and the fractional colonization density (f_c) was calculated as area (A_i ; represented by green autofluorescence of sfGFP-expressing bacteroids, outlined in dark purple in A) per infected zone (A_{iz}). From these, the total infected volume (green) was calculated. For pea nodules (C and D), each nodule was separated into an uninfected cap and a cylindrical body. Total nodule volumes were calculated from the cap using its chord length and sagitta (orange) and the cylindrical body using its area (light purple) and length (teal). The volume of the infected zone (blue and teal) was calculated analogously. Fractional colonization density (f_c) and final infected volume (green) was calculated as for bean nodules from the infected area (dark purple). See Materials and Methods for details and formulae. Images in panels C and D are from the same composite image of a whole nodule with all tiles taken in a single continuous scanning process.

Figure 2. Volume Electron Microscopy imaging of individual bacteroids in pea and bean. A) SBF-SEM imaging of representative bacteroid in bean (Rlp4292, red) and pea (RlvA34, green) segmented across example Z slices and their subsequent 3D reconstruction. B) Cluster of 30 bacteroids in bean (Rlv4292) and pea (RlvA34) showing individual bacteroids segmented in different colors across a range of Z slices and 3D reconstruction of these bacteroids. C) Packing density of Rlp4292 in bean, and RlvA34

in pea were determined by segmenting all bacteroid occupied volume in a defined cellular volume. All scale bars = 2 μ m.

Figure 3. Relative protein and RNA abundance. (A) Comparison of the overall abundance of proteins detected from quantitative proteomics of Rlp4292 bean (red dots) and RlvA34 pea (blue dots) bacteroids, expressed as a percentage of the mass of the proteome. Dotted line represents number of proteins constituting 50% of the proteome. (B) Correlation between normalized counts of RNA (x-axis) and fg of protein per bean bacteroid (red dots). (C) Correlation between normalized counts of RNA (x-axis) and fg of protein per pea bacteroid (blue dots). RNA data used to construct panel B and C from Green et al., 2019. Abbreviations for selected proteins in panels B and C: AcnA, Aconitate hydratase I; Dme, NAD⁺-dependent malic enzyme; FixA, Electron transfer flavoprotein, beta subunit; FixC, Electron transfer flavoprotein oxidoreductase; GltA, Citrate synthase I; Icd, Isocitrate dehydrogenase; Mdh, Malate dehydrogenase; NifD, Nitrogenase molybdenum-iron protein alpha chain; NifH, Nitrogenase iron protein; NifK, Nitrogenase molybdenum-iron protein beta chain; PdhA, Pyruvate dehydrogenase (PDH) E1 component alpha subunit; PdhB, PDH E1 component, beta subunit; PdhC, PDH E2 component dihydrolipoamide acetyltransferase; SucA, 2-oxoglutarate dehydrogenase (OGDH) E1 component; SucC, 2-ODDH E2 component

Video 1. <https://youtu.be/tSkER3KRBfM>

Volume electron microscopy of pea and bean nodules showing segmented bacteroids. Sequential Z-sections are displayed for nodules containing Rlv4292 (bean) and RlvA34

bacteroids (pea). The movie first presents consecutive Z-slices progressing through nodule cells and then returning to the initial Z-plane. A representative bacteroid cluster is subsequently enlarged and segmented in the Z-series. Finally, segmented bacteroid clusters are rendered in 3D without the underlying EM images to illustrate bacteroid morphology and spatial packing. Identical scaling is used for both pea and bean samples. Scale bar: 2 μ m.

REFERENCES

- Allaway D, Lodwig E, Crompton LA, Wood M, Parsons TR, Wheeler T, Poole PS (2000) Identification of alanine dehydrogenase and its role in mixed secretion of ammonium and alanine by pea bacteroids. *Mol Microbiol* **36**: 508-515
- Bozsoki Z, Gysel K, Hansen SB, Lironi D, Kronauer C, Feng F, de Jong N, Vinther M, Kamble M, Thygesen MB, Engholm E, Kofoed C, Fort S, Sullivan JT, Ronson CW, Jensen KJ, Blaise M, Oldroyd G, Stougaard J, Andersen KR, Radutoiu S (2020) Ligand-recognizing motifs in plant LysM receptors are major determinants of specificity. *Science* **369**: 663-670
- Chen Wen F, Meng Xiang F, Jiao Yin S, Tian Chang F, Sui Xin H, Jiao J, Wang En T, Ma Sheng J (2023) Bacteroid development, transcriptome, and symbiotic nitrogen-fixing comparison of *Bradyrhizobium arachidis* in nodules of peanut (*Arachis hypogaea*) and medicinal legume *Sophora flavescens*. *Microbiol Spectr* **11**: e01079-01022
- Choi K-H, Mima T, Casart Y, Rholl D, Kumar A, Beacham Ifor R, Schweizer Herbert P (2008) Genetic tools for select-agent-compliant manipulation of *Burkholderia pseudomallei*. *Appl Environ Microbiol* **74**: 1064-1075
- Downie JA, Hombrecher G, Ma QS, Knight CD, Wells B, Johnston AWB (1983) Cloned nodulation genes of *Rhizobium leguminosarum* determine host range specificity. *Mol. Genet. Genomics* **190**: 359-365
- Durán D, Albareda M, García C, Marina A-I, Ruiz-Argüeso T, Palacios J-M (2021) Proteome analysis reveals a significant host-specific response in *Rhizobium leguminosarum* bv. viciae endosymbiotic cells. *Molecular & Cellular Proteomics* **20**: 100009
- Farkas A, Maróti G, Dürögő H, Györgypál Z, Lima RM, Medzihradszky KF, Kereszt A, Mergaert P, Kondorosi É (2014) *Medicago truncatula* symbiotic peptide NCR247 contributes to bacteroid differentiation through multiple mechanisms. *Proc Natl Acad Sci USA* **111**: 5183-5188
- Feeney M, Kittelmann M, Menassa R, Hawes C, Frigerio L (2018) Protein storage vacuoles originate from remodeled preexisting vacuoles in *Arabidopsis thaliana*. *Plant Physiol* **177**: 241-254
- Geddes BA, Mendoza-Suárez MA, Poole PS (2019) A Bacterial Expression Vector Archive (BEVA) for flexible modular assembly of Golden Gate-compatible vectors. *Front Microbiol* **9**: 10.3389/fmicb.2018.03345

- Gleason C, Chaudhuri S, Yang TB, Munoz A, Poovaiah BW, Oldroyd GED (2006) Nodulation independent of rhizobia induced by a calcium-activated kinase lacking autoinhibition. *Nature* **441**: 1149-1152
- González V, Santamaría RI, Bustos P, Pérez-Carrascal OM, Vinuesa P, Juárez S, Martínez-Flores I, Cevallos MÁ, Brom S, Martínez-Romero E, Romero D (2019) Phylogenomic *Rhizobium* species are structured by a continuum of diversity and genomic clusters. *Front Microbiol* **10**
- Green RT, East AK, Karunakaran R, Downie JA, Poole PS (2019) Transcriptomic analysis of *Rhizobium leguminosarum* bacteroids in determinate and indeterminate nodules. *Micro Genetics* **5**: DOI 10.1099/mgen.1090.000254
- Herridge D, Peoples M, Boddey R (2008) Global inputs of biological nitrogen fixation in agricultural systems. *Plant and Soil* **311**: 1-18
- Hirsch PR (1979) Plasmid-determined bacteriocin production by *Rhizobium leguminosarum*. *J Gen Microbiol* **113**: 219-228
- Johnston AWB, Beringer JE (1975) Identification of the *Rhizobium* strains in pea root nodules using genetic markers. *J Gen Microbiol* **87**: 343-350
- Karunakaran R, Ramachandran VK, Seaman JC, East AK, Moushine B, Mauchline TH, Prell J, Skeffington A, Poole PS (2009) Transcriptomic analysis of *Rhizobium leguminosarum* b.v. *viciae* in symbiosis with host plants *Pisum sativum* and *Vicia cracca*. *J Bacteriol* **191**: 4002-4014
- Kazmierczak T, Nagymihaly M, Lamouche F, Barriere Q, Guefrachi I, Alunni B, Ouadghiri M, Ibijbijen J, Kondorosi E, Mergaert P, Gruber V (2017) Specific host-responsive associations between *Medicago truncatula* accessions and *Sinorhizobium* strains. *Mol Plant Microbe Interact* **30**: 399-409
- Kremer JR, Mastronarde DN, McIntosh JR (1996) Computer visualization of three-dimensional image data using IMOD. *J Struct Biol* **116**: 71-76
- Kultima K, Nilsson A, Scholz B, Rossbach UL, Falth M, Andren PE (2009) Development and evaluation of normalization methods for label-free relative quantification of endogenous peptides. *Mol Cell Proteomics* **8**: 2285-2295
- Lamb JW, Hombrecher G, Johnston AWB (1982) Plasmid-determined nodulation and nitrogen-fixation abilities in *Rhizobium phaseoli*. *Mol Gen Genetics* **186**: 449-452
- Lamouche F, Bonadé-Bottino N, Mergaert P, Alunni B (2019) Symbiotic efficiency of spherical and elongated bacteroids in the *Aeschynomene-Bradyrhizobium* symbiosis. *Front Plant Sci* **Volume 10 - 2019**
- Lodwig E, Poole P (2003) Metabolism of *Rhizobium* bacteroids. *Crit Rev Plan Sci* **22**: 37-78
- Lodwig EM, Hosie AHF, Bourdes A, Findlay K, Allaway D, Karunakaran R, Downie JA, Poole PS (2003) Amino-acid cycling drives nitrogen fixation in the legume-*Rhizobium* symbiosis. *Nature* **422**: 722-726
- Love MI, Huber W, Anders S (2014) Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. *Genome Biol* **15**: 550
- Magoc T, Wood D, Salzberg SL (2013) EDGE-pro: Estimated degree of gene expression in prokaryotic genomes. *Evol Bioinform Online* **9**: 127-136
- Maróti G, Kondorosi É (2014) Nitrogen-fixing *Rhizobium*-legume symbiosis: are polyploidy and host peptide-governed symbiont differentiation general principles of endosymbiosis? *Front Microbiol* **Volume 5 - 2014**
- Marx H, Minogue CE, Jayaraman D, Richards AL, Kwiecien NW, Siahipirani AF, Rajasekar S, Maeda J, Garcia K, Del Valle-Echevarria AR, Volkening JD, Westphall MS, Roy S, Sussman MR, Ané J-M, Coon JJ (2016) A proteomic atlas of the legume *Medicago truncatula* and its nitrogen-fixing endosymbiont *Sinorhizobium meliloti*. *Nature Biotechnology* **34**: 1198-1205

- Mendoza-Suarez MA, Geddes BA, Sanchez-Canizares C, Ramirez-Gonzalez RH, Kirchhelle C, Jorin B, Poole PS (2020) Optimizing *Rhizobium*-legume symbioses by simultaneous measurement of rhizobial competitiveness and N₂ fixation in nodules. *Proc Natl Acad Sci USA* **117**: 9822-9831
- Mergaert P, Uchiumi T, Alunni B, Evanno G, Cheron A, Catrice O, Mausset A-E, Barloy-Hubler F, Galibert F, Kondorosi A, Kondorosi E (2006) Eukaryotic control on bacterial cell cycle and differentiation in the *Rhizobium*-legume symbiosis. *Proc Natl Acad Sci USA* **103**: 5230-5235
- Montiel J, Downie JA, Farkas A, Bihari P, Herczeg R, Balint B, Mergaert P, Kereszt A, Kondorosi E (2017) Morphotype of bacteroids in different legumes correlates with the number and type of symbiotic NCR peptides. *Proc Natl Acad Sci USA* **114**: 5041-5046
- Mwenda GM, O'Hara GW, De Meyer SE, Howieson JG, Terpolilli JJ (2018) Genetic diversity and symbiotic effectiveness of *Phaseolus vulgaris*-nodulating rhizobia in Kenya. *Syst Appl Microbiol* **41**: 291-299
- Oono R, Denison RF (2010) Comparing symbiotic efficiency between swollen versus nonswollen rhizobial bacteroids. *Plant Physiol* **154**: 1541-1548
- Penterman J, Abo RP, De Nisco NJ, Arnold MFF, Longhi R, Zanda M, Walker GC (2014) Host plant peptides elicit a transcriptional response to control the *Sinorhizobium meliloti* cell cycle during symbiosis. *Proc Natl Acad Sci USA* **111**: 3561-3566
- Perez-Riverol Y, Bai J, Bandla C, Garcia-Seisdedos D, Hewapathirana S, Kamatchinathan S, Kundu DJ, Prakash A, Frericks-Zipper A, Eisenacher M, Walzer M, Wang S, Brazma A, Vizcaino JA (2022) The PRIDE database resources in 2022: a hub for mass spectrometry-based proteomics evidences. *Nucleic Acids Res* **50**: D543-D552
- Poole PS, Ramachandran RK, Terpolilli J (2018) Rhizobia: from saprophytes to endosymbionts. *Nat. Rev. Microbiol.* **16**: 291-303
- Prell J, White JP, Bourdes A, Bunnewell S, Bongaerts RJ, Poole PS (2009) Legumes regulate *Rhizobium* bacteroid development and persistence by the supply of branched-chain amino acids. *Proc Natl Acad Sci USA* **106**: 12477-12482
- Sambrook J, Fritsch E, Maniatis T (1989) *Molecular Cloning: A Laboratory Manual*, Ed Third Edition. Cold Spring Harbor, New York
- Sankari S, Babu VMP, Bian K, Alhhazmi A, Andorfer MC, Avalos DM, Smith TA, Yoon K, Drennan CL, Yaffe MB, Lourido S, Walker GC (2022) A haem-sequestering plant peptide promotes iron uptake in symbiotic bacteria. *Nat Microbiol*: 1453-1465
- Scheuring D, Löffke C, Krüger F, Kittelmann M, Eisa A, Hughes L, Smith RS, Hawes C, Schumacher K, Kleine-Vehn J (2016) Actin-dependent vacuolar occupancy of the cell determines auxin-induced growth repression. *Proceedings of the National Academy of Sciences* **113**: 452-457
- Schulte CCM, Borah K, Wheatley RM, Terpolilli JJ, Saalbach G, Crang N, Groot DHd, Ratcliffe RG, Kruger NJ, Papachristodoulou A, Poole PS (2021) Metabolic control of nitrogen fixation in *Rhizobium*-legume symbioses. *Science Advances* **7**: eabh2433
- Sen D, Weaver RW (1981) A comparison of nitrogen-fixing ability of peanut, cowpea and siratro plants nodulated by different strains of *Rhizobium*. *Plant and Soil* **60**: 317-319
- Spaink HP, Sheeley DM, Vanbrussel AAN, Glushka J, York WS, Tak T, Geiger O, Kennedy EP, Reinhold VN, Lugtenberg BJJ (1991) A novel highly unsaturated fatty acid moiety of lipo-oligosaccharide signals determines host specificity of *Rhizobium*. *Nature* **354**: 125-130
- Sprent JI, Ardley J, James EK (2017) Biogeography of nodulated legumes and their nitrogen-fixing symbionts. *New Phytologist* **215**: 40-56
- Terpolilli JJ, Masakapalli SK, Karunakaran R, Webb IUC, Green R, Watmough NJ, Kruger NJ, Ratcliffe RG, Poole PS (2016) Lipogenesis and redox balance in nitrogen-fixing pea bacteroids. *J Bacteriol* **198**: 2864-2875

- 1 **Thoma S, Schobert M** (2009) An improved *Escherichia coli* donor strain for diparental mating. FEMS
- 2 Microbiol Lett **294**: 127-132
- 3 **Tiricz H, Szűcs A, Farkas A, Pap B, Lima Rui M, Maróti G, Kondorosi É, Kereszt A** (2013) Antimicrobial
- 4 nodule-specific cysteine-rich peptides induce membrane depolarization-associated changes in
- 5 the transcriptome of *Sinorhizobium meliloti*. Appl Environ Microbiol **79**: 6737-6746
- 6 **Van de Velde W, Zehirov G, Szatmari A, Debreczeny M, Ishihara H, Kevei Z, Farkas A, Mikulass K, Nagy**
- 7 **A, Tiricz H, Satiat-Jeunemaitre B, Alunni B, Bourge M, Kucho K, Abe M, Kereszt A, Maroti G,**
- 8 **Uchiumi T, Kondorosi E, Mergaert P** (2010) Plant peptides govern terminal differentiation of
- 9 bacteria in symbiosis. Science **327**: 1122-1126
- 10 **Wang D, Griffiths J, Starker C, Fedorova E, Limpens E, Ivanov S, Bisseling T, Long SR** (2010) A nodule-
- 11 specific protein secretory pathway required for nitrogen-fixing symbiosis. Science **327**: 1126-
- 12 1129
- 13 **Westhoek A, Clark LJ, Culbert M, Dalchau N, Griffiths M, Jorin B, Karunakaran R, Ledermann R, Tkacz**
- 14 **A, Webb I, James EK, Poole PS, Turnbull LA** (2021) Conditional sanctioning in a legume–
- 15 *Rhizobium* mutualism. Proc Natl Acad Sci USA **118**: e2025760118
- 16 **Wheatley RM, Ford BL, Li L, Aroney STN, Knights HE, Ledermann R, East AK, Ramachandran VK, Poole**
- 17 **PS** (2020) Lifestyle adaptations of *Rhizobium* from rhizosphere to symbiosis. Proc Natl Acad Sci
- 18 USA **117**: 23823-23834
- 19 **Yan D, Yadav SR, Paterlini A, Nicolas WJ, Petit JD, Brocard L, Belevich I, Grison MS, Vaten A, Karami L,**
- 20 **el-Showk S, Lee J-Y, Murawska GM, Mortimer J, Knoblauch M, Jokitalo E, Markham JE, Bayer**
- 21 **EM, Helariutta Y** (2019) Sphingolipid biosynthesis modulates plasmodesmal ultrastructure and
- 22 phloem unloading. Nature Plants **5**: 604-615

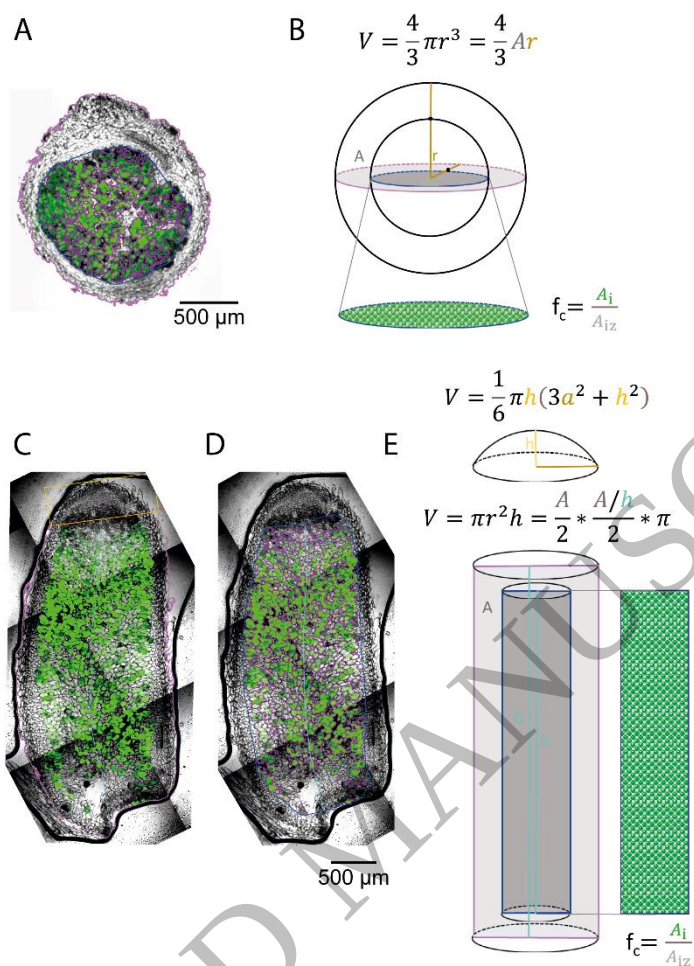


Figure 1
104x128 mm (DPI)

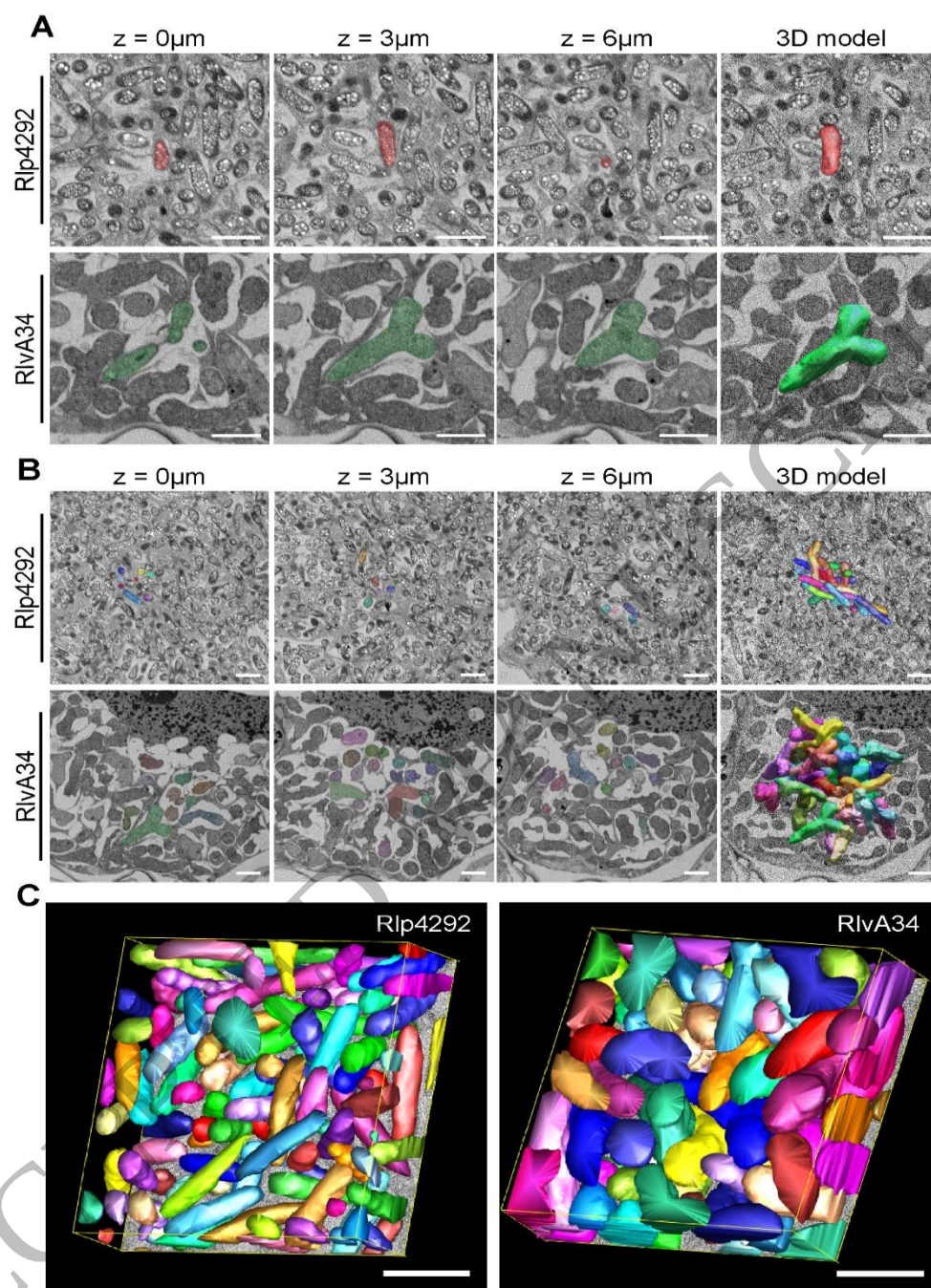


Figure 2. Volume Electron Microscopy imaging of individual bacteroids in pea and bean. **A)** SBF-SEM imaging of representative bacteroid in bean (Rlp4292, red) and pea (RlvA34, green) segmented across example Z slices and their subsequent 3D reconstruction. **B)** Cluster of 30 bacteroids in bean (Rlv4292) and pea (RlvA34) showing individual bacteroids segmented in different colours across a range of Z slices and 3D reconstruction of these bacteroids. **C)** Packing density of Rlp4292 in bean, and RlvA34 in pea were determined by segmenting all bacteroid occupied volume in a defined cellular volume. All scale bars = 2 μm.

Figure 2
135x246 mm (DPI)

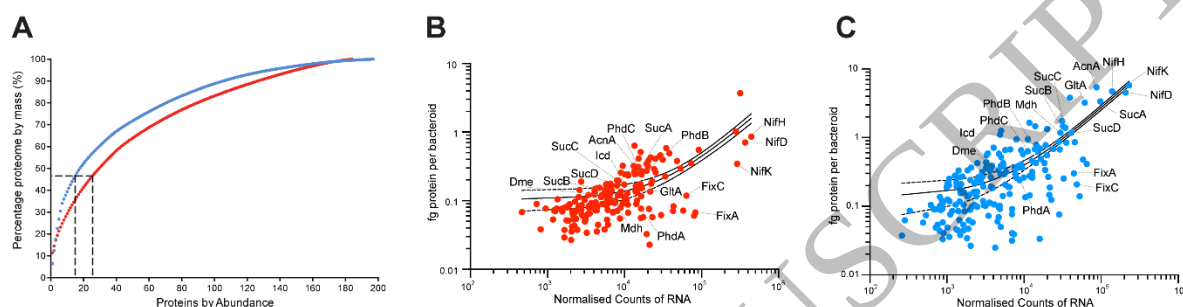
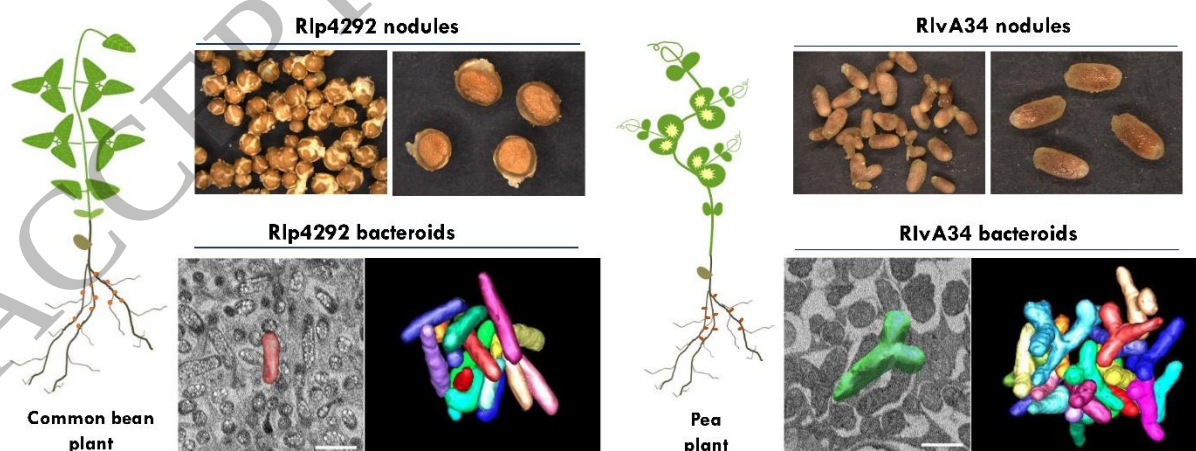


Figure 3. Relative protein and RNA abundance. (A) Comparison between protein abundance (x-axis) and percentage proteome by mass (y-axis), red dots represent bean bacteroids and blue dots represent pea bacteroids. (B) Correlation between normalised counts of RNA (x-axis) and fg of protein per bean bacteroid (red dots). (C) Correlation between normalised counts of RNA (x-axis) and fg of protein per pea bacteroid (blue dots).

Figure 3
159x112 mm (DPI)

From whole plant to single cell: Bacteroid fates and fixing efficiency in legume nodules



Physiological, microscopy, proteomics and RNASeq analyses

Graphical Abstract
159x76 mm (DPI)