bioRxiv preprint doi: https://doi.org/10.1101/2021.07.28.454262; this version posted July 29, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

Genome-scale metabolic modelling of lifestyle changes in *Rhizobium leguminosarum*

Carolin C. M. Schulte[1,2], Vinoy K. Ramachandran[1], Antonis Papachristodoulou[2]#, Philip S. Poole[1]#

[1]Department of Plant Sciences, University of Oxford, Oxford, UK

[2]Department of Engineering Science, University of Oxford, Oxford, UK

Running title: Metabolic modelling of rhizobial lifestyles

#correspondence to antonis@eng.ox.ac.uk or philip.poole@plants.ox.ac.uk

bioRxiv preprint doi: https://doi.org/10.1101/2021.07.28.454262; this version posted July 29, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

Abstract

Biological nitrogen fixation in rhizobium-legume symbioses is of major importance for sustainable agricultural practices. To establish a mutualistic relationship with their plant host, rhizobia transition from free-living bacteria in soil to growth down infection threads inside plant roots and finally differentiate into nitrogen-fixing bacteroids. We reconstructed a genome-scale metabolic model for *Rhizobium leguminosarum* and integrated the model with transcriptome, proteome, metabolome and gene essentiality data to investigate nutrient uptake and metabolic fluxes characteristic of these different lifestyles. Synthesis of leucine, polyphosphate and AICAR is predicted to be important in the rhizosphere, while *myo*-inositol catabolism is active in undifferentiated nodule bacteria in agreement with experimental evidence. The model indicates that bacteroids utilize xylose and glycolate in addition to dicarboxylates, which could explain previously described gene expression patterns. Histidine is predicted to be actively synthesized in bacteroids, consistent with transcriptome and proteome data for several rhizobial species. These results provide the basis for targeted experimental investigation of metabolic processes specific to the different stages of the rhizobium-legume symbioses.

bioRxiv preprint doi: https://doi.org/10.1101/2021.07.28.454262; this version posted July 29, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

Importance

Rhizobia are soil bacteria that induce nodule formation on plant roots and differentiate into nitrogen-fixing bacteroids. A detailed understanding of this complex symbiosis is essential for advancing ongoing efforts to engineer novel symbioses with cereal crops for sustainable agriculture. Here, we reconstruct and validate a genome-scale metabolic model for *Rhizobium leguminosarum* bv. *viciae* 3841. By integrating the model with various experimental datasets specific to different stages of symbiosis formation, we elucidate the metabolic characteristics of rhizosphere bacteria, undifferentiated bacteria inside root nodules, and nitrogen-fixing bacteroids. Our model predicts metabolic flux patterns for these three distinct lifestyles, thus providing a framework for the interpretation of genome-scale experimental datasets and identifying targets for future experimental studies.

bioRxiv preprint doi: https://doi.org/10.1101/2021.07.28.454262; this version posted July 29, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

Introduction

Nitrogen is commonly the main limiting nutrient in agriculture because plants are unable to assimilate atmospheric N_2 (1). Some legumes, such as peas, beans and lentils, circumvent this problem by entering into complex symbiotic relationships with soil bacteria called rhizobia. Legumes secrete signaling molecules (flavonoids) that are recognized by compatible rhizobia, which produce their own signaling molecules (Nod factors) in response. As a result of this signal exchange, rhizobia are typically entrapped by root hairs and grow down so-called infection threads until they are endocytosed by plant cells in the developing nodule. The bacteria then undergo further cell division and eventually differentiate into bacteroids converting atmospheric N_2 into ammonia, which is secreted to the plant host in exchange for carbon sources, mainly dicarboxylates (2–4).

Symbiosis formation is a multi-stage process, requiring distinct metabolic capabilities at each stage. The ability of rhizobia to adapt to various environmental conditions is reflected in their large genomes, which often comprise several replicons (5–7), and in the importance of different genomic regions for each lifestyle (8, 9). While significant research efforts have focused on understanding bacteroid metabolism in rhizobium-legume symbioses, several recent studies have begun to unravel the plant-bacteria interactions preceding the formation of differentiated nitrogen-fixing bacteroids. For example, transcriptomic changes in response to root exudates of different plants have been investigated (10, 11) and biosensors have been developed to elucidate nutrient availability in the rhizosphere (12). Importantly, a study using transposon-based insertion sequencing (INSeq) assessed gene essentiality in *Rhizobium leguminosarum* for rhizosphere bacteria, root-attached

bioRxiv preprint doi: https://doi.org/10.1101/2021.07.28.454262; this version posted July 29, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

bacteria, undifferentiated nodule bacteria and nitrogen-fixing bacteroids (13). It was found that 603 genetic regions were essential for a successful transition from free-living bacteria to bacteroids, highlighting the complexity of development during formation of a successful symbiosis. Understanding the metabolic features at different stages of symbiosis is required for developing effective rhizobial inocula for agricultural applications. Rhizobia that efficiently fix nitrogen are not necessarily adapted to persistence in the rhizosphere as wells as nodulating a plant host in the presence of genetically different bacterial strains, a characteristic described as competitiveness (13–15). Knowledge of the nutrient exchanges between plants and rhizosphere bacteria is thus required for the design of microbial inocula that are competitive and stably persist when applied in the field (14, 16). Once rhizobia have successfully entered the plant root, elucidating the metabolism of undifferentiated rhizobia inside the nodule is also important to avoid delays in the onset of nitrogen fixation.

Due to the complexity of nutrient exchanges in symbioses, metabolic modelling has become a popular tool for investigating rhizobium-legume interactions (3, 17). Metabolic models describe the reactions that are catalyzed by the enzymes annotated in an organism's genome (18, 19). By defining nutrient availability as well as an objective function reflecting the metabolic strategy of the organism, flux distributions at steady state can be calculated using flux balance analysis (20). Due to the gene-protein-reaction associations contained in metabolic models, they also provide a convenient framework for contextualizing genome-scale data obtained by omics technologies, such as transcriptomics or proteomics (21). Most metabolic models of rhizobial species so far have focused on fully differentiated bacteroids

bioRxiv preprint doi: https://doi.org/10.1101/2021.07.28.454262; this version posted July 29, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

(22–25). One *in silico* study of *Sinorhizobium meliloti* has addressed the differences in metabolism for free-living growth in the bulk soil, growth the rhizosphere, and symbiotic nitrogen fixation during the bacteroid stage (9). However, this study focused on the contributions of the different replicons to fitness in the different environments rather than specifics of changes in metabolic flux distributions and did not integrate experimental data. Another study compared the metabolism of free-living *Bradyrhizobium japonicum* with bacteroids (26). While transcriptome and proteome datasets were used to generate models for rhizosphere bacteria and bacteroids, the data for the rhizosphere model were obtained for bacteria grown in a laboratory culture. Only one recent study has addressed metabolic differences in the different nodule zones for the symbiosis between *S. meliloti* and *Medicago truncatula* (27).

In this study, we reconstruct and extensively curate a genome-scale metabolic model (GSM) for *R. leguminosarum* bv. *viciae* 3841 (Rlv3841). Various experimental datasets exist for this strain at different stages of symbiosis with its native host pea. By integrating transcriptome, proteome and gene essentiality data with the GSM, we perform a detailed investigation of nutrient uptake and metabolic pathway usage of Rlv3841 in the rhizosphere, as nodule bacteria and as nitrogen-fixing bacteroids. This genome-scale approach for data integration reproduced experimentally observed phenotypes and particularly highlighted the role of different carbon sources and amino acids throughout the different stages of symbiosis. The metabolic model developed herein provides a valuable resource for targeted investigation of metabolic requirements of the different rhizobial lifestyles and may enable the

bioRxiv preprint doi: https://doi.org/10.1101/2021.07.28.454262; this version posted July 29, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

identification of strategies for engineering strains that are metabolically advantaged at all stages of symbiosis formation.

Results

Reconstruction of a genome-scale metabolic model for *Rhizobium leguminosarum*

Most published metabolic models for rhizobia focus on bacteroids and are therefore limited to metabolic pathways active during nitrogen fixation. Curated genome-scale reconstructions are so far only available for *B. japonicum* (26) and *S. meliloti* (27). With the aim of investigating metabolism in the rhizosphere and during different stages of bacteroid development, we developed a GSM for Rlv3841 using multiple sources of information. As shown in Fig. 1A, automated reconstructions based on the KEGG (28) and MetaCyc (29) databases were combined with a homology-based reconstruction using a GSM for *S. meliloti* as a template and reactions from our previously reconstructed bacteroid model of Rlv3841 (23) (Fig. 1B). Extensive curation was then performed based on literature evidence, gene essentiality data (13, 30) and enzymatic functions predicted by DeepEC (31). Comparison with iML1515, a high-quality model for *Escherichia coli* (32) as well as the CarveMe template for Gram-negative bacteria (33) was further used to correct reaction stoichiometry and reversibility if required. We next defined a biomass function based on evidence from the literature (Table S1). Since our previous work showed the dependence of carbon polymer synthesis on environmental conditions (23), demand reactions for polymers such as glycogen, polyhydroxybutyrate (PHB) and exopolysaccharides were included in the model to allow for their flexible

bioRxiv preprint doi: https://doi.org/10.1101/2021.07.28.454262; this version posted July 29, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

accumulation. The final model (Supplementary Data 1 and 2) contained 1224 genes, 1257 reactions and 984 metabolites (Table 1), and was named iCS1224 according to standard naming conventions. The largest groups of metabolic reactions were associated with amino acid and lipid metabolism (14.1% and 13.5% of model reactions, respectively), followed by cofactor metabolism (10.7%) and purine/pyrimidine metabolism (9.0%) (Fig. 1C). Cluster of orthologous genes (COG) (34) analysis of the model genes showed that all COG categories associated with metabolic reactions were represented in iCS1224 (Fig. S1). The quality of the reconstruction was evaluated using MEMOTE (35), where iCS1224 achieved an overall score of 89% (Supplementary Data 3).

Model validation

We validated our model for free-living Rlv3841 growing in minimal media using various experimental datasets. First, we experimentally assessed growth on 190 different carbon sources using phenotype microarrays (36) (Supplementary Data 4). For the 109 carbon sources that were present as metabolites in iCS1224, an overall predictive accuracy of 89.9% with 90.9% precision and 96.4% recall was achieved (Fig. 2A), which is similar to the performance of curated GSMs for well-investigated bacteria, such as *Pseudomonas aeruginosa* (37) or *E. coli* (38). In addition, we evaluated the quality of gene essentiality predictions by comparing *in silico* gene essentiality with the results of an INSeq gene essentiality screen of Rlv3841 performed in minimal media supplemented with succinate and ammonia (30). Since the classification of genes based on transposon mutagenesis screen is subject to some variability (39), the list of essential genes was further curated by comparison

bioRxiv preprint doi: https://doi.org/10.1101/2021.07.28.454262; this version posted July 29, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

with INSeq data for growth on complex media (13). Predictions by iCS1224 achieved an accuracy, precision and recall of 91.0%, 89.6% and 87.8%, respectively (Fig. 2B), thus showing good agreement with the INSeq data and indicating high quality of the gene-protein-reaction associations as well as suitability of the biomass objective function.

Finally, we performed quantitative validation of our model by comparing the predicted flux values for 17 reactions involved in central carbon metabolism with published values measured by ^{13}C metabolic flux analysis of Rlv3841 grown in minimal media with succinate and ammonia (40). As shown in Fig. 2C, we observed excellent agreement between predicted and experimentally measured flux values. In all cases, the measured flux was within the range determined by flux variability analysis. iCS1224 thus appears to be an accurate representation of the metabolism of Rlv3841, both qualitatively and quantitatively.

Metabolism of rhizosphere bacteria

Having validated the predictive capabilities of iCS1224, we sought to extract condition-specific models for metabolism of Rlv3841 (i) in the rhizosphere, (ii) as undifferentiated nodule bacteria and (iii) as nitrogen-fixing bacteroids (Fig. 3). We chose the recently developed RIPTiDe algorithm (41) to obtain condition-specific metabolic models. Based on gene expression data, RIPTiDe assigns weights to all gene-associated reactions, assuming that higher transcript abundance makes it more likely that the corresponding reaction is used in a certain environmental condition. The overall flux through the network is then minimized and inactive reactions are removed. Finally, flux sampling of the solution space is performed,

bioRxiv preprint doi: https://doi.org/10.1101/2021.07.28.454262; this version posted July 29, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

where flux through reactions associated with highly expressed genes is favored. In contrast to other methods for transcriptome data integration, RIPTiDe does not impose arbitrary thresholds on the gene expression data, it produces functional models with flux through the objective reaction, and takes flux parsimony into account, i.e., the overall flux is minimized to find cost-efficient solutions (41). Generation of a rhizosphere-specific model thus required information about available nutrients as well as gene expression data. Nutrient availability in the rhizosphere is mainly determined by plant root exudates, and plants modulate the composition of their root exudates to select for specific soil microbes (42, 43). However, only a subset of metabolites is used by the soil microbiota (44, 45), and elucidation of nutrient uptake by rhizosphere bacteria usually requires extensive metabolomics profiling (46, 47). Taking a top-down approach for defining the rhizosphere environment, we first compiled a list of compounds present in pea root exudates based on published experimental data (10, 12, 48, 49) (Table S2 and S3, Supplementary Text). For those compounds that could be matched to model metabolites, exchange reactions were added to the model with reaction bounds set to allow for unlimited uptake. RNA-Seq data for Rlv3841 in the rhizosphere of pea plants 7 days post inoculation (Supplementary Data 5) was used as an input dataset for model contextualization. In addition, a list of genes that were classified as essential or defective in the rhizosphere in an INSeq screen (13) was provided to prevent removal of reactions associated with these genes from the rhizosphere-specific model. Biomass production was set as the objective and additional positive lower bounds were placed on exopolysaccharide, lipopolysaccharide and Nod factor synthesis, all of which are known to be important in the rhizosphere (4). During data

bioRxiv preprint doi: https://doi.org/10.1101/2021.07.28.454262; this version posted July 29, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

integration, constraining flux through the objective function to values between 50% and 95% of its maximum was tested to identify the scenario that gave the best match with the transcriptome data. Within the range of objective values tested, the highest correlation (Spearman's Rho=0.237, $P<0.001$) between metabolic fluxes and transcript abundances was obtained with the biomass reaction constrained to carry at least 77.5% of its maximum flux, and the rhizosphere-specific model contained 606 reactions and 576 metabolites (Supplementary Data 6). Remarkably, out of the 134 nutrients available for uptake before data integration, only 51 were present in the rhizosphere-specific model.

For the analysis of the contextualized model, we focused on those metabolic pathways that are either not universally essential or that are retained in the model despite their end product being available for uptake from the environment. Pathways such as membrane lipid or PHB synthesis, for instance, will always be retained in the model, since they are required to maintain flux through the biomass objective function and uptake of lipids and PHB is not possible. In addition, we limit our discussion to reactions that had a non-zero median flux value based on the flux sampling results, since those reactions are most likely to be active in the rhizosphere. The TCA cycle was predicted to be a central catabolic pathway (Fig. 4), which is consistent with previous reports of organic acids being the predominant carbon sources for rhizobia in the rhizosphere (9, 10). In particular, the model predicted high uptake of glycolate in agreement with the induction of C2 metabolism observed in previous gene expression studies (10). Glycolate was converted into pyruvate via glycolate oxidase and an aminotransferase. The model also showed high uptake rates for aspartate, which could explain the induction of a *dctA* biosensor

bioRxiv preprint doi: https://doi.org/10.1101/2021.07.28.454262; this version posted July 29, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

in Rlv3841 in the pea rhizosphere (12). Aspartate and 2-oxoglutarate were transaminated to produce glutamate and oxaloacetate, which is a TCA cycle intermediate.

In addition to organic acids, amylotriose, which is hydrolyzed into glucose, was partly metabolized via the Entner-Doudoroff pathway and glycolysis in the model and entered the pentose phosphate pathway to enable production of nucleotides required for the synthesis of various polysaccharides and Nod factors. The solute-binding protein of a carbohydrate uptake transporter-1 (CUT1) family transporter (RL3840) was 2.6-fold upregulated in the pea rhizosphere compared to free-living cells (10), which supports the predicted uptake of a di- or oligosaccharide. Ribulose, a monosaccharide metabolized via the pentose phosphate pathway was also predicted to be taken up. Catabolism of a monosaccharide in the rhizosphere is highly probable considering the strong signals of a fructose and a xylose biosensor in the pea rhizosphere (12). The fructose biosensor is based on the solute-binding protein of the CUT2 family *frcABC* transporter, which has been shown to transport ribose in addition to fructose in *S. meliloti* (50) and may therefore also contribute to pentose uptake in the rhizosphere. The model further contained reactions for glycerol uptake and catabolism, which could explain the decreased competitiveness observed for a glycerol catabolism mutant of *R. leguminosarum* bv. *viciae* VF39 (51).

With regard to amino acids, all of which are present in root exudates, biosynthetic pathways were generally retained in the rhizosphere model due to the essentiality of the associated genes. Low levels of uptake were however predicted for most amino acids, mainly to support protein synthesis. Notably, the biosynthetic pathway for leucine was predicted to be active, which was partly supported by uptake of 2-

bioRxiv preprint doi: https://doi.org/10.1101/2021.07.28.454262; this version posted July 29, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

isopropylmalate, an intermediate of branched-chain amino acid synthesis. The need for leucine synthesis in the rhizosphere agrees with a *leuD* mutant of Rlv3841 requiring the addition of 1 mM leucine to nodulate pea (52). Mutation of the isopropylmalate synthase gene in *S. meliloti* impaired nodulation even in the presence of leucine, and it was shown that either isopropylmalate synthase or intermediates of the leucine biosynthetic pathway are required for the activation of *nod* gene expression (53). It is therefore possible that the predicted leucine synthesis is at least partly related to the synthesis of Nod factors in the rhizosphere. High uptake rates were further predicted for glutamine, which is consistent with its high concentration in pea root exudates (49). Glutamine was converted into glutamate, which was mostly used to sustain leucine synthesis. The model also contained active uptake reactions for several nucleotides. This agrees with the reported uptake of nucleosides and nucleotides by rhizosphere bacteria (44, 47, 54) and agrees with the gene essentiality predictions for rhizosphere bacteria obtained by INSeq, where purine auxotrophs appear to be rescued by plant root exudates (13).

Among biomass components that are present in root exudate but not predicted to be taken up, biosynthesis of the polyamine putrescine was retained in the model, attesting to the ability of the RIPTiDe algorithm to choose metabolic reactions that agree with gene expression and/or essentiality rather than choosing the least resource-intensive solution. Putrescine and related polyamines are important for survival under stress conditions and their synthesis has been suggested to play an important role during root colonization (55). As part of the model reconstruction, several demand reactions were included for compounds such as carbon polymers whose accumulation can vary with environmental conditions. The only non-essential

bioRxiv preprint doi: https://doi.org/10.1101/2021.07.28.454262; this version posted July 29, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

demand reactions that were not removed during the pruning process were those for glutathione and polyphosphate, where polyphosphate synthesis in particular had a non-zero median flux. Glutathione is important to deal with stress conditions, such as osmotic and oxidative stress, encountered in the rhizosphere, and mutants in glutathione biosynthesis are severely affected in rhizosphere colonization (56). The predicted catabolism of glycolate via glyoxylate produces the reactive oxygen species hydrogen peroxide, which could contribute to the need for glutathione synthesis. Polyphosphate has recently been suggested to play a role in the global carbon regulatory system (57), but its function remains to be investigated in detail. It is interesting to note that an exopolyphosphatase gene (RL1600) was classified as essential for persistence in the rhizosphere (13), indicating an important role for phosphate homeostasis in rhizosphere colonization and/or competition.

Catabolism of several other compounds, such as erythritol, *myo*-inositol and homoserine, has been described to be important for competitiveness (10, 58, 59), however, these compounds were not included in the rhizosphere model. This could be due to the catabolism of these compounds being important at later stages of the symbiosis, e.g. for growth in infection threads rather than in the rhizosphere. Alternatively, uptake of these compounds could be masked in the model due to catabolic routes that are shared with other metabolites. For example, erythritol is metabolized via the pentose phosphate pathway (60), hence the predicted uptake and metabolism of ribulose could partly be due to erythritol catabolism.

bioRxiv preprint doi: https://doi.org/10.1101/2021.07.28.454262; this version posted July 29, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

Reporter metabolites highlight plant-specific rhizosphere metabolism

As an independent validation and extension of our analysis of metabolic changes, we identified reporter metabolites using previously published microarray data comparing Rlv3841 in the rhizosphere of pea plants with free-living cells grown on minimal media with glucose and ammonium chloride (10). Based on the network topology defined by a metabolic model, the reporter metabolite algorithm identifies those compounds around which significant changes in gene expression occur (61). This method is therefore independent of specifying nutrient uptake from the environment. Reporter metabolites associated with upregulated genes matched several observations from the RNA-Seq data integration described in the previous section. In particular, several intermediates of branched-chain amino acid synthesis, such as acetolactate, 2-hydroxyethyl-thiamin diphosphate and 2-aceto-2-hydroxybutanoate, were identified as reporter metabolites (Fig. 5A). Significant transcriptional changes were also observed around various nucleobase derivatives. This may be related to their predicted uptake from plant root exudates but could also indicate an increased need for nucleotide synthesis for the production of polysaccharides and Nod factor. Phosphoribosyl-AMP and phosphoribosylformiminoaicar-phosphate are intermediates of histidine biosynthesis and direct precursors of AICAR, which is involved in purine metabolism. Seeing as no additional metabolites of the histidine biosynthetic pathway were identified as reporter metabolites, this analysis indicates an increase in AICAR synthesis, which seems to be required for successful legume infection by various *Rhizobium* species (62).

bioRxiv preprint doi: https://doi.org/10.1101/2021.07.28.454262; this version posted July 29, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

Comparison of the rhizosphere reporter metabolites for pea (host legume for Rlv3841) with those for alfalfa (non-host legume) (Fig. 5B) and sugar beet (non-legume) (Fig. 5C) highlighted several plant-specific features. For alfalfa, phosphoribosyl-AMP was identified as a reporter metabolite similar to pea. In addition, phenylalanine and tyrosine support the role of aromatic amino acid metabolism in colonization competitiveness (10). Significant transcriptional changes also occurred around the carbon polymer beta-glucan and the diamine putrescine. While beta-glucan generally appears to be important for persistence in the rhizosphere (10, 13), its identification as a reporter metabolite together with putrescine indicates increased osmotic stress in the alfalfa rhizosphere compared to pea. For sugar beet, the identification of several compounds involved in nitrogen metabolism (ammonia, urea, urate) agrees with the suggested nitrogen limitation in the sugar beet rhizosphere, but nitrogen sufficiency in legume rhizospheres (10). This could also explain why the carbon polymer glycogen was a reporter metabolite specifically in the sugar beet rhizosphere, since glycogen synthesis is probably linked to nitrogen limitation (63). Notably, multiple mono- and disaccharides and their derivatives indicate an increased importance of sugar metabolism compared to legume rhizospheres. However, many genes involved in saccharide metabolism are associated with multiple reactions (e.g. unspecific glucoside hydrolases), and therefore the identity of the metabolized sugar cannot be derived from this analysis. Finally, the reporter metabolites 3-dehydrocarnitine and betainyl-CoA indicate either accumulation of amines for osmoprotection or catabolism of carnitine or related amines. These findings present interesting targets for future investigations using gene essentiality screens on different plant hosts.

bioRxiv preprint doi: https://doi.org/10.1101/2021.07.28.454262; this version posted July 29, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

Overall, both the context-specific model obtained by transcriptome data integration and the reporter metabolite analysis were in good agreement with experimental data for rhizobial metabolism in the rhizosphere without forcing the uptake of any compound through arbitrary constraints. Instead, insights into nutrient uptake were facilitated by the integration of gene expression and gene essentiality data with iCS1224. If biomass production were simply maximized with unlimited availability of all root exudate compounds, this would result in uptake of all available compounds that are required for biomass formation, which would not reflect a biologically meaningful scenario.

Metabolism of undifferentiated nodule bacteria

We next sought to develop models for Rlv3841 inside the nodule environment. For this purpose, it is important to differentiate between nodule bacteria at the tip of the nodule, which are dividing and undergoing differentiation, and bacteroids in the central nitrogen fixation zone of the nodule (64). While nodule bacteria are still dividing, bacteroids are growth-arrested and mainly catabolize plant-provided dicarboxylates to fix atmospheric N_2 into ammonia. The distinction between these developmental stages is required in the context of gene essentiality analyses since genes required for the differentiation process may not be essential for nitrogen fixation and vice versa. Similar to the approach for the rhizosphere model, we used RIPTiDe to obtain models for nodule bacteria and bacteroids and performed flux sampling to identify those reactions that are most likely to be active in each contextualized model.

bioRxiv preprint doi: https://doi.org/10.1101/2021.07.28.454262; this version posted July 29, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

The model for nodule bacteria was obtained using published dRNA-Seq data for RNA extracted from nodule tips (65), as well as a list of genes that were predicted to be specifically essential for nodule bacteria (13). Nutrient availability was defined based on a study using biosensors to detect metabolites inside nodules (12) and our direct measurement of metabolites in pea root exudate, in pea bacteroids and in the nodule cytosol as described previously (40) (Table S4 and S5, Supplementary Data 7). The biomass objective function was used to account for the cell division occurring as rhizobia grow down infection threads and differentiate into bacteroids and positive lower bounds were placed on demand reactions for exopolysaccharides and lipopolysaccharides. The nodule bacteria model contained 510 reactions and 502 metabolites and achieved highest correlation with the transcriptome data (Spearman's Rho=0.335, P<0.001) when the objective value was constrained to 65% of its maximum (Supplementary Data 8). The observation that higher correlation was obtained for lower flux through the objective reaction (compared to the rhizosphere) indicates that the metabolism of nodule bacteria is not oriented towards maximum growth. This agrees with experimental data showing that growth of infection threads proceeds at highly variable rates controlled by the plant host (66). The improved correlation of flux predictions and gene expression data compared to the rhizosphere model can be explained by the lower number of essential genes, which places fewer constraints on the reactions included in the contextualized model.

Malate, fructose, xylose, *myo*-inositol and γ-aminobutyrate (GABA) were all predicted to be taken up by nodule bacteria (Figs. 6A and S2). Biosensors for these carbon sources were strongly induced in young nodules, whereas biosensors for the carbon

bioRxiv preprint doi: https://doi.org/10.1101/2021.07.28.454262; this version posted July 29, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

sources that were removed during the data integration process (erythritol, mannitol, formate, malonate, tartrate) only showed weak induction (12). Malate and GABA are both catabolized in the TCA cycle, indicating that it is an important catabolic route in differentiating nodule bacteria despite transport of dicarboxylates being non-essential for differentiation into bacteroids (67). Enzymes involved in GABA metabolism are highly induced in bacteroids, although GABA catabolism is not essential for effective nitrogen fixation (68). The predicted catabolism of fructose is consistent with the strong induction of a fructose-specific biosensor in nodules (12) as well as a previous modelling study of *S. meliloti* suggesting the use of sucrose-derived sugars as a carbon source by differentiating nodule bacteria (27). Sucrose uptake was removed from the nodule bacteria model, which may be due to the inability of the model to accurately distinguish between sucrose and fructose uptake based on the gene expression data. It is interesting to note that Rlv3841 bacteroids mutated in a subunit of succinyl-CoA synthetase, which had severely reduced nitrogen fixation capacity, had 168-fold higher levels of fructose than wild-type bacteroids and 151-fold elevated levels of sucrose (Supplementary Data 7), which may be the result of carbon source build-up in the developmentally impaired nodule bacteria and bacteroids. *Myo*-inositol is present in the rhizosphere (12) and abundant in pea nodules (69), and mutants in *myo*-inositol catabolism have strongly reduced competitiveness compared to wild-type Rlv3841 (58). However, the activity of enzymes involved in *myo*-inositol catabolism is very low in mature bacteroids (69), and mutants in *myo*-inositol catabolism were not disadvantaged during growth in the rhizosphere compared to wild-type Rlv3841 (58). In addition, it has been proposed that catabolism of rhizopines, which are inositol derivatives, by undifferentiated

bioRxiv preprint doi: https://doi.org/10.1101/2021.07.28.454262; this version posted July 29, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

nodule bacteria may be important as a kin selection strategy (70). Catabolism of *myo*-inositol is therefore most likely to play a role during infection and in undifferentiated nodule bacteria, which is correctly predicted by the model. Xylose enters the pentose phosphate pathway, and its predicted uptake could be related to the importance of nucleotide synthesis, both for DNA endoreduplication and synthesis of exopolysaccharides and lipopolysaccharides. Similarly, uptake reactions for the nucleoside guanosine and uridine as well as the nucleobase adenine were present in the model.

Our nodule bacteria model predicted uptake of most amino acids, which agrees with the severe symbiotic defect of a *gltB* mutant unable to transport amino acids (71) but may also be a result of a beginning general downregulation of biosynthetic functions as rhizobia transition into growth-arrested bacteroids. Similar to the rhizosphere bacteria, leucine was predicted to be synthesized from 2-isopropylmalate. Expression of *nod* genes is elevated in nodule bacteria at 7 days post inoculation (72), which could explain the predicted leucine synthesis as discussed for the rhizosphere model.

Metabolism of bacteroids

To extract a model specific for nitrogen-fixing bacteroids, we used dRNA-Seq data derived from the middle of nodules (65), which contains fully differentiated bacteroids performing nitrogen fixation (64). In addition, a list of 38 genes that were present in the model and encoded proteins significantly upregulated in bacteroids compared to free-living bacteria (23) and the *dct* genes (73) were specified to ensure inclusion of those genes in the bacteroid model. Nitrogenase activity was set as the objective

bioRxiv preprint doi: https://doi.org/10.1101/2021.07.28.454262; this version posted July 29, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

function while low levels of protein and fatty acid production were enforced through demand reactions. Nutrient availability was specified similar to the considerations for nodule bacteria (Table S6). Gene essentiality data from the INSeq screen were not included for model contextualization due to the aforementioned difficulty of determining the developmental stage where a gene is essential inside the nodule environment. The bacteroid model contained 307 reactions and 308 metabolites (Supplementary Data 9) and achieved significant correlation with the transcriptome data (Spearman's Rho=0.348, P<0.001) when nitrogenase activity was constrained to 65% of its maximum. The reduced model size compared to both the rhizosphere and the nodule bacteria model is in agreement with the reduced physiological complexity of the non-dividing bacteroids (3, 72).

The bacteroid model contained the C4 dicarboxylates malate, succinate and fumarate as the main carbon sources in agreement with experimental evidence (67, 74) (Figs. 6B and S3), and only low levels of GABA uptake were predicted. Ammonia was the only nitrogenous export product. Consistent with our previous modelling study of Rlv3841 bacteroids (23), constraining the oxygen uptake prior to data integration resulted in nitrogen partly being secreted as alanine. With the metabolites provided in initial simulations, the glyoxylate cycle comprising isocitrate lyase and malate synthase was contained in the model, which is consistent with the high induction of malate synthase (72) but disagrees with the lack of isocitrate lyase activity in pea bacteroids (75). The source of glyoxylate for the malate synthase reaction has so far not been elucidated. Because the metabolomics data showed that glycolate is present in the nodule cytosol and glycolate concentrations in bacteroids are 2-fold elevated compared to free-living cells (Supplementary Data 7),

bioRxiv preprint doi: https://doi.org/10.1101/2021.07.28.454262; this version posted July 29, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

we allowed for glycolate uptake by the bacteroid model and inactivated the isocitrate lyase reaction. This resulted in substantial uptake of glycolate, which was converted into glyoxylate that was used in the malate synthase reaction. Glycolate provision by the plant may therefore explain the increase in malate synthase expression in the absence of isocitrate lyase activity.

The model also contained uptake of xylose, which was metabolized in the pentose phosphate pathway and supported synthesis of nucleotides. Dicarboxylate catabolism generally requires gluconeogenesis to provide precursors for the synthesis of nucleotides and some amino acids. Due to the predicted xylose uptake, only minor flux through the reactions involved in gluconeogenesis was predicted, highlighting the importance of this pathway in bacteroids as an interesting question to explore using targeted mutant studies. Proton uptake by bacteroids was further required as previously predicted for *S. meliloti* bacteroids (27) and a demand reaction for PHB was retained in the model. PHB synthesis was highly variable across flux samples, which is in agreement with its previously suggested role for carbon and redox balancing (23, 40).

Low levels of uptake were predicted for most amino acids to support the required synthesis of protein, but no significant catabolism of any amino acid was observed. Mutant studies have shown a requirement for branched-chain amino acid supply to bacteroids (52), and isoleucine was predicted to be supplied by the plant. Interestingly, histidine was predicted to be synthesized rather than taken up by bacteroids. Several proteins involved in histidine synthesis were upregulated or unchanged in abundance in the bacteroid proteome compared to free-living Rlv3841 (23), in contrast to the general downregulation of amino acid biosynthesis (3). Similar

bioRxiv preprint doi: https://doi.org/10.1101/2021.07.28.454262; this version posted July 29, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

493 results were obtained in a proteome study of *Rhizobium etli* (24) and RNA-Seq data
494 for bacteroids of *R. leguminosarum* bv. *viciae* A34 and *R. leguminosarum* bv.
495 *phaseoli* 4292 (76). In addition, mutants of *R. leguminosarum* bv. *trifolii* lacking
496 histidinol dehydrogenase activity formed ineffective nodules on clover (77). To
497 investigate the requirement for histidine biosynthesis, we compared the amino acid
498 composition of the Nif and Fix proteins, which are highly expressed in bacteroids,
499 with the overall amino acid composition of the Rlv3841 proteome (Table S7). We
500 found a significant (P=0.042) enrichment of histidine in the Nif and Fix proteins,
501 which could at least partly explain why histidine biosynthesis is required in
502 bacteroids.
503

504 **Discussion**

505 In this study, we present the first curated GSM for Rlv3841, a model strain for
506 investigating rhizobium-legume interactions and a natural symbiont of the
507 agriculturally important crop pea. GSMs have emerged as promising tools for
508 informing experimental design, addressing fundamental research questions, and
509 contextualizing experimental data (78). In order to obtain a high-quality model,
510 integration of experimental data during model curation and validation is essential
511 (79). We therefore evaluated our model using carbon source utilization, gene
512 essentiality data and flux data obtained by ^{13}C labelling and observed high
513 agreement between model predictions and experimental data.
514 We further used the GSM to elucidate metabolic changes in Rlv3841 as it transitions
515 from a free-living soil bacterium in the rhizosphere to an undifferentiated nodule
516 bacterium and finally to a nitrogen-fixing bacteroid. While significant advances in

bioRxiv preprint doi: https://doi.org/10.1101/2021.07.28.454262; this version posted July 29, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

517 determining metabolic requirements for successful symbiosis formation have been
518 made using transcriptome data (10) and gene essentiality screens (13), genome-
519 scale datasets are often difficult to interpret without the framework of a
520 comprehensive model, especially when information about nutrient uptake is missing.
521 To this end, we employed approaches integrating gene expression and metabolome
522 data as well as gene essentiality predicted by INSeq to obtain condition-specific
523 models. This allowed us to contextualize our model based on experimental data
524 without assuming uptake rates for any nutrient. In addition, during the process of
525 data integration, different fractions of the optimum objective value were tested as
526 constraints to find a solution with the highest correlation between gene expression
527 and associated reaction fluxes. Especially for nodule bacteria and bacteroids, using
528 sub-optimal fluxes through the objective function as constraints during model
529 contextualization was found to produce better agreement with experimental data.
530 Objective functions can be difficult to define outside of defined growth in a laboratory
531 culture, and our results highlight the need to adopt strategies beyond maximization
532 of a biomass objective function to accurately capture metabolic behavior in complex
533 environmental settings. A clear limitation of our approach is the imperfect correlation
534 of gene expression and protein abundance, as well as protein abundance and
535 enzyme activity (80, 81). Catabolic pathways common to multiple different
536 compounds can further make it difficult to specifically determine which nutrient is
537 taken up. Nevertheless, our model predictions are in good agreement with known
538 metabolic characteristics of the different lifestyles of Rlv3841, attesting to the
539 biological relevance of our findings.

bioRxiv preprint doi: https://doi.org/10.1101/2021.07.28.454262; this version posted July 29, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

The rhizosphere model showed substantial uptake of glycolate, aspartate and glutamine as well as mono- and oligosaccharides. These predictions are consistent with previous studies about gene expression (10) as well as nutrient uptake of a *Rhizobium* sp. from root exudates of *Arabidopsis* (44). We further identified a requirement for leucine synthesis in the rhizosphere, as well as a potentially important role for polyphosphate synthesis. However, the predicted nutrient uptake was not supported by gene essentiality predictions in all cases. While both INSeq gene essentiality assignments and our metabolic model generate predictions that warrant detailed investigation using isolated mutant strains, there are other possible explanations for this observation. First, root exudates might not contain sufficient quantities of a compound to complement an auxotrophy. In addition, the composition of plant root exudates has been shown to change over time (47), and compounds present at the time of RNA extraction may not be present from the time of inoculation, causing the loss of some mutants. Finally, for genes that are essential on complex media, the corresponding mutants may already be lost from or underrepresented in the bacterial population that is inoculated onto plants.

The model for nodule bacteria confirmed previous results suggesting supply of nutrients other than dicarboxylates, in particular sucrose-derived sugars, during the differentiation process (27). Interestingly, we found that *myo*-inositol catabolism was only predicted for nodule bacteria, but not in the rhizosphere or in bacteroids. While the importance of *myo*-inositol catabolism for competitiveness has been established (58), our results suggest that it may be particularly important for differentiating bacteria rather than those in the rhizosphere. For bacteroids, biosynthesis of histidine was found to be important in contrast to the general uptake predicted for

bioRxiv preprint doi: https://doi.org/10.1101/2021.07.28.454262; this version posted July 29, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

most other amino acids. In addition, low levels of xylose uptake were predicted to support nucleotide synthesis in bacteroids. This result indicates that a carbon source metabolized in the pentose phosphate pathway may be provided to bacteroids, which presents an interesting area to explore experimentally using mutants affected in gluconeogenesis. Initial predictions of isocitrate lyase activity, which disagree with measured enzyme activities in bacteroids, led us to hypothesize that glycolate is provided to bacteroids. This is supported by metabolomics data and could explain the induction of malate synthase in bacteroids without concomitant expression of isocitrate lyase.

In summary, our results provide insights into rhizobial metabolism in the rhizosphere, which can inform the design of more competitive rhizobial inocula as well as plants that secrete metabolites to specifically enrich beneficial bacterial strains. Our understanding of the nutrient exchanges between plants and rhizobia at different developmental stages inside nodules remains incomplete (3, 82), and the predictions presented herein provide a foundation for targeted investigation of amino acid and central carbon metabolism in particular. We anticipate that the highly curated metabolic model for Rlv3841 presented in this paper will provide a valuable resource for the reconstruction of GSMs for related species.

bioRxiv preprint doi: https://doi.org/10.1101/2021.07.28.454262; this version posted July 29, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

Materials and Methods

Model reconstruction

To reconstruct a GSM for Rlv3841, we combined information from multiple databases, which has been shown to significantly improve the scope of metabolic network reconstructions (83). All reconstructions were performed based on RefSeq assembly GCF_000009265.1. We used the RAVEN Toolbox 2.0 (84) to create draft models from KEGG (28) and MetaCyc (29) using the functions `getKEGGModelForOrganism` and `getMetaCycModelForOrganism`, respectively. In addition, template-based reconstruction based on BLAST bidirectional hits was performed using a curated GSM for *S. meliloti* 1021 (iGD1348 (27)) as a template for the function `getModelFromHomology`. All models were merged into one reaction list and reaction and compound identifiers were unified based on the reaction database provided by the ModelSEED (85), followed by removal of duplicate reactions. Starting from this database of reactions compiled from different sources, the reconstruction was curated. First, reactions without gene association were removed. Reactions involving non-specific compounds such as “acceptor” or “protein” were also deleted, as well as reactions involved in the biosynthesis and catabolism of secondary metabolites and non-metabolic processes, such as DNA and RNA modification because those were outside of the scope of our model. Extensive curation was then performed by evaluating metabolic pathways guided by the literature and the KEGG database. Pathways for catabolism of small carbon sources in particular were reconstructed based on predictions obtained from GapMind (86). Gene-protein-reaction associations were curated based on published

bioRxiv preprint doi: https://doi.org/10.1101/2021.07.28.454262; this version posted July 29, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

gene essentiality data for growth in minimal (30) and complete (13) media as well as enzyme commission (EC) number predictions obtained from DeepEC (31). Transport reactions were annotated based on literature evidence, in particular homology to experimentally characterized transporters in *S. meliloti* (87) and the annotation obtained from TransportDB 2.0 (88). We manually reconstructed pathways for organism-specific biomass components, such as lipopolysaccharides and exopolysaccharides, as well as pathways which were not present in any of the databases used for reconstruction, such as carnitine metabolism. To improve information on reaction directionality, upper and lower bounds were adjusted according to the information in a highly curated model for *E. coli* (iML1515 (32)) and the CarveMe template model for Gram-negative bacteria (33).

The biomass objective function was defined as follows: The composition of DNA was determined from the RefSeq genome sequence. Similarly, RNA and protein composition were determined by counting nucleotides or amino acids in the annotated RNAs and protein coding sequences, respectively. Since the lipid composition of Rlv3841 has not been investigated so far, we adopted the values reported *R. leguminosarum* bv. *trifolii* ANU843 (89). *R. leguminosarum* produces predominantly C18 fatty acids, as well as smaller quantities of C16 fatty acids (40, 90), and representative phospholipids in our model included fatty acids with these chain lengths. Lipopolysaccharides and exopolysaccharides were included with the fractions previously reported for *S. meliloti* (91). Cyclic beta-glucans have so far not been considered in metabolic models for rhizobia, however they can make up a significant fraction of the cellular dry weight (92) and were therefore also included as a biomass component. Apart from the main cell components, trace amounts of

bioRxiv preprint doi: https://doi.org/10.1101/2021.07.28.454262; this version posted July 29, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

cofactors identified as universally essential in prokaryotes (93) were included in the biomass objective function. Phytoene was also added to the biomass reaction due to the essentiality of the genes associated with its biosynthetic pathway. Carbon polymers such as glycogen, PHB and fatty acids, as well as polyamines such as homospermidine and putrescine are known to be produced by Rlv3841, however, the quantities in which they are produced vary depending on nutrient availability. Similar to our previous work (23), we therefore added demand reactions for these compounds to allow for variable accumulation. Glycogen and PHB were also included in the biomass objective function as they are commonly synthesized by free-living *R. leguminosarum* (94). A complete description of the biomass composition used in this study is given in Table S1.

General modelling procedures

Standard metabolic modelling computations were performed in MATLAB R2020b (Mathworks) using scripts from the COBRA Toolbox v3.0 (95) and the Gurobi 9.1.1 solver (www.gurobi.com). When using the `optimizeCbModel` function, the Taxicab norm was minimized to avoid loops in the calculated flux distributions. All scripts are available on Github (https://github.com/CarolinSchulte/Rlv3481-lifestyles).

Model validation

To evaluate the agreement between model predictions and experimentally determined carbon source utilization, we limited our analysis to those compounds that were either present in the model or showed a positive growth phenotype in the phenotype microarray experiment. The lower bounds for the exchange reactions

bioRxiv preprint doi: https://doi.org/10.1101/2021.07.28.454262; this version posted July 29, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

were then adjusted according to the composition of universal minimal salts (UMS) media (30) with ammonium as a nitrogen source, and flux through the biomass reaction was evaluated for each carbon source individually added to the model. Accuracy, precision and recall for carbon source utilization and gene essentiality analysis were calculated according to the following equations:

$$accuracy = \frac{TP + TN}{TP + TN + FP + FN}$$

$$precision = \frac{TP}{TP + FP}$$

$$recall = \frac{TP}{FP + FN}$$

TP: true positives FP: false positives

TN: true negatives FN: false negatives

Gene essentiality analysis was performed using the function `singleGeneDeletion` with the MOMA (minimization of metabolic adjustment) option in the COBRA Toolbox, while all components of UMS media with succinate and ammonia were available without constraints on their uptake rate. The predictions were compared with gene essentiality data for Rlv3841 determined by INSeq (13, 30). Genes that were experimentally classified as essential or defective were considered essential in our analysis. The threshold for a gene to be considered essential *in silico* was set to 50% of the wild-type growth rate since all mutant strains are grown in a single culture for an INSeq experiment, and a slower growth rate will

bioRxiv preprint doi: https://doi.org/10.1101/2021.07.28.454262; this version posted July 29, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

675 therefore decrease the abundance of a mutant even if the gene carrying the insertion
676 is not absolutely essential.
677 For comparison with ^{13}C labelling data, boundary conditions were set to allow for
678 unlimited uptake of UMS media components. The succinate uptake rate was
679 constrained to 1 flux unit and flux balance analysis was performed maximizing the
680 biomass objective function. In addition, loopless flux variability analysis was
681 performed where the objective fraction was set to 95% of the optimum value.
682
683 **Data integration for model contextualization**
684 The Python implementation of RIPTiDe (https://github.com/mjenior/riptide) (41) was
685 used to generate condition-specific models of iCS1224. Max fit RIPTiDe was run for
686 objective flux fractions between 0.5 and 0.95 with 0.05 increments, and the context-
687 specific models with the highest correlation between flux values and transcriptome
688 data were used in further analyses.
689 In addition to the nutrient availability determined based on experimental data, trace
690 elements and vitamins required for flux through the objective function were added to
691 the *in silico* representation of each environment. If an exchange and transport
692 reaction for a compound already existed in the model, the lower bound of the
693 exchange reaction was set to -1000. If a compound was only present as an
694 intracellular metabolite, a sink reaction for this metabolite with lower bound set to -
695 1000 and upper bound set to 0 was added. This was done to avoid erroneous
696 exclusion of metabolites which are present in the environment, but for which
697 transporters have not been identified. Some cofactors and central intermediates of
698 carbon metabolism, such as glyceraldehyde 3-phsophate, were excluded from

bioRxiv preprint doi: https://doi.org/10.1101/2021.07.28.454262; this version posted July 29, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

699 environmental representations since their uptake would result in unspecific
700 predictions for metabolic pathway activity (Supplementary Text).
701
702 **Data integration for rhizosphere model**
703 For the rhizosphere model, compounds that have been detected in pea root
704 exudates (10, 12, 48, 49) and that could be matched to model metabolites were
705 specified with unlimited availability (Table S2 and S3). Flux through the biomass
706 reaction as described in the previous section was set as the objective function and in
707 addition, a lower bound of one flux unit was set for demand reactions for Nod factor,
708 lipopolysaccharides and exopolysaccharides, since these compounds are known to
709 be produced as part of the root colonization process (4). RPKM values for RNA-Seq
710 data obtained from Rlv3841 in the pea rhizosphere 7 days post inoculation were
711 provided as an input, and all genes that are present in the model and were classified
712 as essential or defective in the rhizosphere (13) were specified as model tasks to
713 prevent removal of the associated reactions during the pruning process.
714
715 **Data integration for nodule bacteria model**
716 For nodule bacteria, all metabolites that were detected by rhizobial biosensors in pea
717 nodules (12) were allowed to be taken up without limitation, as well as all amino
718 acids and metabolites whose abundance was at least ten-fold higher in the nodule
719 cytosol compared to root exudates (Table S4 and S5). A lower bound of one flux unit
720 was set for lipopolysaccharide and exopolysaccharide demand reactions and
721 biomass production was used as the objective function. RPKM values for dRNA-Seq
722 data obtained from the tip of pea nodules were provided as an input, and all genes

bioRxiv preprint doi: https://doi.org/10.1101/2021.07.28.454262; this version posted July 29, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

723 that were classified as essential or defective for nodule bacteria (13) were specified
724 as model tasks to prevent removal of the associated reactions during the pruning
725 process.
726
727 **Data integration for bacteroids**
728 Similar to the nodule bacteria model, metabolites detected in nodules by rhizobial
729 biosensors, and all amino acids were made available to the bacteroid model (Table
730 S6). However, fructose and sucrose were not included since they are known to be
731 poorly oxidized by bacteroids (96). Inclusion of the metabolites increased in the
732 nodule cytosol compared to root exudates led to a decrease in correlation of flux
733 predictions and gene expression data, and those metabolites were therefore omitted
734 from the nutrients available to bacteroids. A lower bound of one flux unit was set for
735 the synthesis of fatty acids and proteins and flux through the nitrogenase reaction
736 was used as the objective function. RPKM values for dRNA-Seq data obtained from
737 the middle of pea nodules were provided as an input, and all genes associated with
738 proteins significantly upregulated in bacteroids compared to free-living cells (23)
739 were specified as model tasks to prevent removal of the associated reactions during
740 the pruning process.
741
742 **Phenotype MicroArray™ analysis**
743 Carbon source utilization of Rlv3841 was assessed using the phenotype microarray
744 technology (Biolog, Hayward, USA). A liquid culture of Rlv3841 was grown at 28 ºC
745 in UMS media supplemented with 10 mM glucose, 10 mM ammonium chloride and a
746 vitamin solution as previously described (30). Cells were spun down and washed

bioRxiv preprint doi: https://doi.org/10.1101/2021.07.28.454262; this version posted July 29, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

three times in UMS without addition of a carbon or nitrogen source. The optical density at 600 nM was then adjusted to 0.1 with UMS supplemented with 10 mM ammonium chloride and vitamins, and 100 ml of cell suspension were added to each well of the phenotype microarray plate. After overnight incubation without shaking at 28 ºC, 10 ml of a 0.1 % (w/v) stock solution of 2,3,5-triphenyltetrazolium chloride in water were added to each well. Plates were then incubated in an Omega FluoStar plate reader with double orbital shaking at 500 rpm and the absorbance at 505 nm was measured every 15 min. Absorbance values were analyzed using the DuctApe software (97), and all carbon sources with an activity value higher than the water control were considered to support growth. For activity values close to the growth threshold, curves were manually inspected, and literature searches were performed to determine if the carbon source supports growth of *R. leguminosarum*. The full DuctApe output for the phenotype microarray analysis is available in Supplementary Data 4.

Metabolomics data

Metabolomics data were obtained in a previous study (40), where only values for metabolites relevant to the investigated metabolic pathways were published . The full metabolomics dataset is included as Supplementary Data 7.

Sample preparation for RNA-Seq of rhizosphere bacteria

For total RNA extraction from Rlv3841 in the pea rhizosphere, *Pisum sativum* cv. Avola seeds were surface sterilized and sown in sterilized boiling tubes with fine vermiculite and nitrogen-free rooting solution. Pea seeds were grown in the dark for

bioRxiv preprint doi: https://doi.org/10.1101/2021.07.28.454262; this version posted July 29, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

771 3 days and then transferred to a controlled environment room, where they were
772 grown at 25 °C with a 16:8 h photoperiod for another 4 days. On day 7, 1 ml (108
773 CFU/ml) of washed Rlv3841 cells was added near the root. At 7 days post
774 inoculation, rhizobial cells were harvested from the rhizosphere as previously
775 described (10). RNA was extracted for three biological replicates where the total
776 RNA extracted from the pea rhizosphere of 16 boiling tubes was pooled for each
777 replicate. Quality and quantity of the total RNA was assessed using Experion
778 StdSens (Standard Sensitivity) and HighSens (High Sensitivity) analysis kits. Total
779 RNA (3 mg per sample) was treated with the TURBO DNA-free kit (Invitrogen
780 AM1907) as previously described (10). Depletion of genomic DNA was confirmed by
781 performing a Qubit fluorometer double-stranded DNA broad range assay. Finally, the
782 ribosomal RNA was depleted from the RNA sample using the Illumina Ribo-Zero
783 rRNA removal kit - Gram-negative (MRZGN126) according to the manufacturer's
784 instructions. The rRNA-depleted mRNA was purified using the ZymoResearch RNA
785 Clean & Concentrator 50. mRNA samples were used to generate barcoded cDNA
786 libraries for multiplexing during sequencing using the Ion Total RNA-Seq kit v2
787 (Thermo Fisher Scientific). Each barcoded cDNA library was quantified using the
788 Agilent Bioanalyzer High Sensitivity DNA kit and diluted to a final concentration of 70
789 pM. Equal volumes of the diluted cDNA libraries were pooled before loading on the
790 IonChef for template preparation and chip loading. Finally, the chips were sequenced
791 in an Ion Proton Semiconductor based sequencing platform (Thermo Fisher
792 Scientific). The full dataset is available on the NCBI SRA database, BioProject
793 number PRJNA748006.
794

bioRxiv preprint doi: https://doi.org/10.1101/2021.07.28.454262; this version posted July 29, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

795 **Data analysis for RNA-Seq of rhizosphere bacteria**

796 RNA-Seq data was de-multiplexed based on valid barcodes and data for each library
797 was downloaded in fastq format. The overall quality of the sequencing and the data
798 was assessed based on the Torrent Browser suite sequencing run report summary.
799 Data from each library was assessed using FastQC (Babraham Institute;
800 https://www.bioinformatics.babraham.ac.uk/projects/fastqc/) and any remaining
801 adapters and low-quality reads were filtered using cutadapt (98). The data for each
802 library was mapped against the Rlv3841 genome using EDGE-pro (99) developed
803 specifically for bacterial RNA-Seq data. EDGE-pro uses Bowtie2 to map the reads to
804 the genome and calculates the frequencies per nucleotide. EDGE-pro calculates the
805 number of reads and RPKM value for each gene feature in the genome including
806 mRNA, rRNA, tRNA. The mapped reads from each library were visualized with the
807 Integrative Genomics Viewer (100) for further analysis.
808

809 **dRNA-Seq data for nodule bacteria and bacteroids**

810 The dRNA-Seq data used for creation of the nodule bacteria and the bacteroid
811 model were described previously (65) and are available on the NCBI SRA database,
812 BioProject number PRJNA667846.

bioRxiv preprint doi: https://doi.org/10.1101/2021.07.28.454262; this version posted July 29, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

References

1. Ågren GI, Wetterstedt JÅM, Billberger MFK. 2012. Nutrient limitation on terrestrial plant growth – modeling the interaction between nitrogen and phosphorus. New Phytol 194:953–960.
2. Udvardi M, Poole PS. 2013. Transport and metabolism in legume-rhizobia symbioses. Annu Rev Plant Biol 64:781–805.
3. Ledermann R, Schulte CCM, Poole PS. 2021. How rhizobia adapt to the nodule environment. J Bacteriol 203:e00539-20.
4. Poole P, Ramachandran V, Terpolilli J. 2018. Rhizobia: from saprophytes to endosymbionts. Nat Rev Microbiol 16:291–303.
5. Young JPW, Crossman LC, Johnston AWB, Thomson NR, Ghazoui ZF, Hull KH, Wexler M, Curson ARJ, Todd JD, Poole PS, Mauchline TH, East AK, Quail MA, Churcher C, Arrowsmith C, Cherevach I, Chillingworth T, Clarke K, Cronin A, Davis P, Fraser A, Hance Z, Hauser H, Jagels K, Moule S, Mungall K, Norbertczak H, Rabbinowitsch E, Sanders M, Simmonds M, Whitehead S, Parkhill J. 2006. The genome of *Rhizobium leguminosarum* has recognizable core and accessory components. Genome Biol 7:R34.
6. Galibert F, Finan TM, Long SR, Puhler A, Abola P, Ampe F, Barloy-Hubler F, Barnett MJ, Becker A, Boistard P, Bothe G, Boutry M, Bowser L, Buhrmester J, Cadieu E, Capela D, Chain P, Cowie A, Davis RW, Dreano S, Federspiel NA, Fisher RF, Gloux S, Godrie T, Goffeau A, Golding B, Gouzy J, Gurjal M, Hernandez-Lucas I, Hong A, Huizar L, Hyman RW, Jones T, Kahn D, Kahn ML, Kalman S, Keating DH, Kiss E, Komp C, Lelaure V, Masuy D, Palm C, Peck MC, Pohl TM, Portetelle D, Purnelle B, Ramsperger U, Surzycki R,

bioRxiv preprint doi: https://doi.org/10.1101/2021.07.28.454262; this version posted July 29, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

838 Thebault P, Vandenbol M, Vorholter FJ, Weidner S, Wells DH, Wong K, Yeh
839 KC, Batut J. 2001. The composite genome of the legume symbiont
840 *Sinorhizobium meliloti*. Science 293:668–672.
841 7. González V, Santamaría RI, Bustos P, Hernández-González I, Medrano-Soto
842 A, Moreno-Hagelsieb G, Janga SC, Ramírez MA, Jiménez-Jacinto V, Collado-
843 Vides J, Dávila G. 2006. The partitioned *Rhizobium etli* genome: Genetic and
844 metabolic redundancy in seven interacting replicons. Proc Natl Acad Sci
845 103:3834–3839.
846 8. Jiao J, Ni M, Zhang B, Zhang Z, Young JPW, Chan T-F, Chen WX, Lam H-M,
847 Tian CF. 2018. Coordinated regulation of core and accessory genes in the
848 multipartite genome of *Sinorhizobium fredii*. PLOS Genet 14:e1007428.
849 9. DiCenzo GC, Checcucci A, Bazzicalupo M, Mengoni A, Viti C, Dziewit L, Finan
850 TM, Galardini M, Fondi M. 2016. Metabolic modelling reveals the specialization
851 of secondary replicons for niche adaptation in *Sinorhizobium meliloti*. Nat
852 Commun 7:1–10.
853 10. Ramachandran VK, East AK, Karunakaran R, Downie JA, Poole PS. 2011.
854 Adaptation of *Rhizobium leguminosarum* to pea, alfalfa and sugar beet
855 rhizospheres investigated by comparative transcriptomics. Genome Biol
856 12:R106.
857 11. Fagorzi C, Bacci G, Huang R, Cangioli L, Checcucci A, Fini M, Perrin E, Natali
858 C, diCenzo GC, Mengoni A. 2021. Nonadditive transcriptomic signatures of
859 genotype-by-genotype interactions during the initiation of plant-rhizobium
860 symbiosis. mSystems 6:e00974-20.
861 12. Pini F, East AK, Appia-Ayme C, Tomek J, Karunakaran R, Mendoza-Suárez

bioRxiv preprint doi: https://doi.org/10.1101/2021.07.28.454262; this version posted July 29, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

M, Edwards A, Terpolilli JJ, Roworth J, Downie JA, Poole PS. 2017. Bacterial biosensors for *in vivo* spatiotemporal mapping of root secretion. Plant Physiol 174:1289–1306.

13. Wheatley RM, Ford BL, Li L, Aroney STN, Knights HE, Ledermann R, East AK, Ramachandran VK, Poole PS. 2020. Lifestyle adaptations of *Rhizobium* from rhizosphere to symbiosis. Proc Natl Acad Sci 117:23823–23834.
14. Checcucci A, DiCenzo GC, Bazzicalupo M, Mengoni A. 2017. Trade, diplomacy, and warfare: the quest for elite rhizobia inoculant strains. Front Microbiol 8:2207.
15. Mendoza-Suárez MA, Geddes BA, Sánchez-Cañizares C, Ramírez-González RH, Kirchhelle C, Jorrin B, Poole PS. 2020. Optimizing *Rhizobium*-legume symbioses by simultaneous measurement of rhizobial competitiveness and N_2 fixation in nodules. Proc Natl Acad Sci 117:9822–9831.
16. Busby PE, Soman C, Wagner MR, Friesen ML, Kremer J, Bennett A, Morsy M, Eisen JA, Leach JE, Dangl JL. 2017. Research priorities for harnessing plant microbiomes in sustainable agriculture. PLOS Biol 15:e2001793.
17. DiCenzo GC, Zamani M, Checcucci A, Fondi M, Griffitts JS, Finan TM, Mengoni A. 2019. Multidisciplinary approaches for studying rhizobium–legume symbioses. Can J Microbiol 65:1-33.
18. Nielsen J. 2017. Systems biology of metabolism. Annu Rev Biochem 86:245–275.
19. Thiele I, Palsson B. 2010. A protocol for generating a high-quality genome-scale metabolic reconstruction. Nat Protoc 5:93–121.
20. Orth JD, Thiele I, Palsson BO. 2010. What is flux balance analysis? Nat

bioRxiv preprint doi: https://doi.org/10.1101/2021.07.28.454262; this version posted July 29, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

886 Biotechnol 28:245-248.

887 21. Blazier AS, Papin JA. 2012. Integration of expression data in genome-scale
888 metabolic network reconstructions. Front Physiol 3:299.

889 22. Contador CA, Lo S-K, Chan SHJ, Lam H-M. 2020. Metabolic analyses of
890 nitrogen fixation in the soybean microsymbiont *Sinorhizobium fredii* using
891 constraint-based modeling. mSystems 5:e00516-19.

892 23. Schulte CCM, Borah K, Wheatley RM, Terpolilli JJ, Saalbach G, Crang N, de
893 Groot DH, Ratcliffe RG, Kruger NJ, Papachristodoulou A, Poole PS. 2021.
894 Metabolic constraints on nitrogen fixation by rhizobia in legume nodules.
895 bioRxiv https://doi.org/10.1101/2021.02.16.431433.

896 24. Resendis-Antonio O, Hernández M, Salazar E, Contreras S, Batallar GM,
897 Mora Y, Encarnación S. 2011. Systems biology of bacterial nitrogen fixation:
898 high-throughput technology and its integrative description with constraint-
899 based modeling. BMC Syst Biol 5:120.

900 25. Zhao H, Li M, Fang K, Chen W, Wang J. 2012. *In silico* insights into the
901 symbiotic nitrogen fixation in *Sinorhizobium meliloti* via metabolic
902 reconstruction. PLoS One 7:e31287.

903 26. Yang Y, Hu X-P, Ma B-G. 2017. Construction and simulation of the
904 *Bradyrhizobium diazoefficiens* USDA110 metabolic network: a comparison
905 between free-living and symbiotic states. Mol BioSyst 13:607.

906 27. diCenzo GC, Tesi M, Pfau T, Mengoni A, Fondi M. 2020. Genome-scale
907 metabolic reconstruction of the symbiosis between a leguminous plant and a
908 nitrogen-fixing bacterium. Nat Commun 11:1–11.

909 28. Kanehisa M, Furumichi M, Tanabe M, Sato Y, Morishima K. 2016. KEGG: new

bioRxiv preprint doi: https://doi.org/10.1101/2021.07.28.454262; this version posted July 29, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

910 perspectives on genomes, pathways, diseases and drugs. Nucleic Acids Res
911 45:353–361.
912 29. Caspi R, Billington R, Keseler IM, Kothari A, Krummenacker M, Midford PE,
913 Ong WK, Paley S, Subhraveti P, Karp PD. 2020. The MetaCyc database of
914 metabolic pathways and enzymes - a 2019 update. Nucleic Acids Res
915 48:D445–D453.
916 30. Wheatley RM, Ramachandran VK, Geddes BA, Perry BJ, Yost CK, Poole PS.
917 2017. Role of O2 in the growth of *Rhizobium leguminosarum* bv. *viciae* 3841
918 on glucose and succinate. J Bacteriol 199:e00572-16.
919 31. Ryu JY, Kim HU, Lee SY. 2019. Deep learning enables high-quality and high-
920 throughput prediction of enzyme commission numbers. Proc Natl Acad Sci
921 116:13996–14001.
922 32. Monk JM, Lloyd CJ, Brunk E, Mih N, Sastry A, King Z, Takeuchi R, Nomura W,
923 Zhang Z, Mori H, Feist AM, Palsson BO. 2017. iML1515, a knowledgebase
924 that computes *Escherichia coli* traits. Nat Biotechnol 35:904–908.
925 33. Machado D, Andrejev S, Tramontano M, Patil KR. 2018. Fast automated
926 reconstruction of genome-scale metabolic models for microbial species and
927 communities. Nucleic Acids Res 46:7542-7553.
928 34. Galperin MY, Wolf YI, Makarova KS, Vera Alvarez R, Landsman D, Koonin E
929 V. 2021. COG database update: focus on microbial diversity, model
930 organisms, and widespread pathogens. Nucleic Acids Res 49:D274–D281.
931 35. Lieven C, Beber ME, Olivier BG, Bergmann FT, Ataman M, Babaei P, Bartell
932 JA, Blank LM, Chauhan S, Correia K, Diener C, Dräger A, Ebert BE,
933 Edirisinghe JN, Faria JP, Feist AM, Fengos G, Fleming RMT, García-Jiménez

bioRxiv preprint doi: https://doi.org/10.1101/2021.07.28.454262; this version posted July 29, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

B, Hatzimanikatis V, van Helvoirt W, Henry CS, Hermjakob H, Herrgård MJ, Kaafarani A, Kim HU, King Z, Klamt S, Klipp E, Koehorst JJ, König M, Lakshmanan M, Lee DY, Lee SY, Lee S, Lewis NE, Liu F, Ma H, Machado D, Mahadevan R, Maia P, Mardinoglu A, Medlock GL, Monk JM, Nielsen J, Nielsen LK, Nogales J, Nookaew I, Palsson BO, Papin JA, Patil KR, Poolman M, Price ND, Resendis-Antonio O, Richelle A, Rocha I, Sánchez BJ, Schaap PJ, Malik Sheriff RS, Shoaie S, Sonnenschein N, Teusink B, Vilaça P, Vik JO, Wodke JAH, Xavier JC, Yuan Q, Zakhartsev M, Zhang C. 2020. MEMOTE for standardized genome-scale metabolic model testing. Nat Biotechnol 38:272–276.

36. Bochner BR, Gadzinski P, Panomitros E. 2001. Phenotype microarrays for high-throughput phenotypic testing and assay of gene function. Genome Res 11:1246–1255.

37. Bartell JA, Blazier AS, Yen P, Thøgersen JC, Jelsbak L, Goldberg JB, Papin JA. 2017. Reconstruction of the metabolic network of *Pseudomonas aeruginosa* to interrogate virulence factor synthesis. Nat Commun 8:14631.

38. Monk JM, Charusanti P, Aziz RK, Lerman JA, Premyodhin N, Orth JD, Feist AM, Palsson BØ. 2013. Genome-scale metabolic reconstructions of multiple *Escherichia coli* strains highlight strain-specific adaptations to nutritional environments. Proc Natl Acad Sci 110:20338–20343.

39. Blazier AS, Papin JA. 2019. Reconciling high-throughput gene essentiality data with metabolic network reconstructions. PLOS Comput Biol 15:e1006507.

40. Terpolilli JJ, Masakapalli SK, Karunakaran R, Webb IUC, Green R, Watmough NJ, Kruger NJ, Ratcliffe RG, Poole PS. 2016. Lipogenesis and redox balance

bioRxiv preprint doi: https://doi.org/10.1101/2021.07.28.454262; this version posted July 29, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

in nitrogen-fixing pea bacteroids. J Bacteriol 198:2864–2875.

41. Jenior ML, Moutinho Jr. TJ, Dougherty B V, Papin JA. 2020. Transcriptome-guided parsimonious flux analysis improves predictions with metabolic networks in complex environments. PLOS Comput Biol 16:e1007099.

42. Rudrappa T, Czymmek KJ, Paré PW, Bais HP. 2008. Root-secreted malic acid recruits beneficial soil bacteria. Plant Physiol 148:1547–1556.

43. Biedrzycki ML, Bais HP. 2009. Root secretions: from genes and molecules to microbial associations. J Exp Bot 60:1533–1534.

44. Jacoby RP, Martyn A, Kopriva S. 2018. Exometabolomic profiling of bacterial strains as cultivated using *Arabidopsis* root extract as the sole carbon source. Mol Plant Microbe Interact 31:803–813.

45. Baran R, Brodie EL, Mayberry-Lewis J, Hummel E, Da Rocha UN, Chakraborty R, Bowen BP, Karaoz U, Cadillo-Quiroz H, Garcia-Pichel F, Northen TR. 2015. Exometabolite niche partitioning among sympatric soil bacteria. Nat Commun 6:8289.

46. Jacoby RP, Kopriva S. 2019. Metabolic niches in the rhizosphere microbiome: new tools and approaches to analyse metabolic mechanisms of plant-microbe nutrient exchange. J Exp Bot 70:1087–1094.

47. Zhalnina K, Louie KB, Hao Z, Mansoori N, da Rocha UN, Shi S, Cho H, Karaoz U, Loqué D, Bowen BP, Firestone MK, Northen TR, Brodie EL. 2018. Dynamic root exudate chemistry and microbial substrate preferences drive patterns in rhizosphere microbial community assembly. Nat Microbiol 3:470–480.

48. Gaworzewska ET, Carlile MJ. 1982. Positive chemotaxis of *Rhizobium*

bioRxiv preprint doi: https://doi.org/10.1101/2021.07.28.454262; this version posted July 29, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

leguminosarum and other bacteria towards root exudates from legumes and other plants. Microbiology 128:1179–1188.

49. Knee EM, Gong F-C, Gao M, Teplitski M, Jones AR, Foxworthy A, Mort AJ, Bauer WD. 2001. Root mucilage from pea and its utilization by rhizosphere bacteria as a sole carbon source. Mol Plant-Microbe Interact 14:775–784.

50. Annie L, Magne Ø, Karine M, Marie-Christine P, Daniel LR. 2001. Fructose uptake in *Sinorhizobium meliloti* is mediated by a high-affinity ATP-binding sassette transport system. J Bacteriol 183:4709–4717.

51. Ding H, Yip CB, Geddes BA, Oresnik IJ, Hynes MF. 2012. Glycerol utilization by *Rhizobium leguminosarum* requires an ABC transporter and affects competition for nodulation. Microbiology 158:1369–1378.

52. 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 106:12477–12482.

53. Sanjuán-Pinilla JM, Muñoz S, Nogales J, Olivares J, Sanjuán J. 2002. Involvement of the *Sinorhizobium meliloti leuA* gene in activation of nodulation genes by NodD1 and luteolin. Arch Microbiol 178:36–44.

54. Cole BJ, Feltcher ME, Waters RJ, Wetmore KM, Mucyn TS, Ryan EM, Wang G, Ul-Hasan S, McDonald M, Yoshikuni Y, Malmstrom RR, Deutschbauer AM, Dangl JL, Visel A. 2017. Genome-wide identification of bacterial plant colonization genes. PLOS Biol 15:e2002860.

55. Becerra-Rivera VA, Dunn MF. 2019. Polyamine biosynthesis and biological roles in rhizobia. FEMS Microbiol Lett 366:fnz084.

56. Cheng G, Karunakaran R, East AK, Munoz-Azcarate O, Poole PS. 2017.

bioRxiv preprint doi: https://doi.org/10.1101/2021.07.28.454262; this version posted July 29, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

Glutathione affects the transport activity of *Rhizobium leguminosarum* 3841 and is essential for efficient nodulation. FEMS Microbiol Lett 364:fnx045.

57. Webb I, Xu J, Sanchez-Cañizares C, Karunakaran R, Ramachandran V, Rutten P, East A, Huang W, Watmough N, Poole P. 2021. Regulation and characterization of mutants of *fixABCX* in *Rhizobium leguminosarum*. Mol Plant-Microbe Interact https://doi.org/10.1094/MPMI-02-21-0037-R.

58. Fry J, Wood M, Poole PS. 2001. Investigation of *myo*-inositol catabolism in *Rhizobium leguminosarum* bv. *viciae* and its effect on nodulation competitiveness. Mol Plant-Microbe Interact 14:1016–1025.

59. Yost CK, Rath AM, Noel TC, Hynes MF. 2006. Characterization of genes involved in erythritol catabolism in *Rhizobium leguminosarum* bv. *viciae*. Microbiology 152:2061–2074.

60. Barbier T, Collard F, Zúñiga-Ripa A, Moriyón I, Godard T, Becker J, Wittmann C, Van Schaftingen E, Letesson JJ. 2014. Erythritol feeds the pentose phosphate pathway via three new isomerases leading to D-erythrose-4-phosphate in *Brucella*. Proc Natl Acad Sci 111:17815–17820.

61. Patil KR, Nielsen J. 2005. Uncovering transcriptional regulation of metabolism by using metabolic network topology. Proc Natl Acad Sci 102:2685–2689.

62. Newman JD, Diebold RJ, Schultz BW, Noel KD. 1994. Infection of soybean and pea nodules by *Rhizobium* spp. purine auxotrophs in the presence of 5-aminoimidazole-4-carboxamide riboside. J Bacteriol 176:3286–3294.

63. Ofaim S, Sulheim S, Almaas E, Sher D, Segrè D. 2021. Dynamic allocation of carbon storage and nutrient-dependent exudation in a revised genome-scale model of *Prochlorococcus*. Front Genet 12:586293.

bioRxiv preprint doi: https://doi.org/10.1101/2021.07.28.454262; this version posted July 29, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

64. Vasse J, F de B, Camut S, Truchet G. 1990. Correlation between ultrastructural differentiation of bacteroids and nitrogen fixation in alfalfa nodules. J Bacteriol 172:4295–4306.

65. Rutten PJ, Steel H, Hood GA, Ramachandran VK, McMurtry L, Geddes B, Papachristodoulou A, Poole PS. 2021. Multiple sensors provide spatiotemporal oxygen regulation of gene expression in a *Rhizobium*-legume symbiosis. PLOS Genet 17:e1009099.

66. Fournier J, Timmers ACJ, Sieberer BJ, Jauneau A, Chabaud M, Barker DG. 2008. Mechanism of infection thread elongation in root hairs of *Medicago truncatula* and dynamic interplay with associated rhizobial colonization. Plant Physiol 148:1985–1995.

67. Ronson CW, Lyttleton P, Robertson JG. 1981. C(4)-dicarboxylate transport mutants of *Rhizobium trifolii* form ineffective nodules on *Trifolium repens*. Proc Natl Acad Sci 78:4284–8.

68. Prell J, Bourdés A, Karunakaran R, Lopez-Gomez M, Poole P. 2009. Pathway of γ-aminobutyrate metabolism in *Rhizobium leguminosarum* 3841 and its role in symbiosis. J Bacteriol 191:2177–2186.

69. Poole PS, Blyth A, Reid CJ, Walters K. 1994. *Myo*-inositol catabolism and catabolite regulation in *Rhizobium leguminosarum* bv. *viciae*. Microbiology 140:2787–2795.

70. Oono R, Denison RF, Kiers ET. 2009. Controlling the reproductive fate of rhizobia: how universal are legume sanctions? New Phytol 183:967–979.

71. Mulley G, White JP, Karunakaran R, Prell J, Bourdes A, Bunnewell S, Hill L, Poole PS. 2011. Mutation of GOGAT prevents pea bacteroid formation and N2

bioRxiv preprint doi: https://doi.org/10.1101/2021.07.28.454262; this version posted July 29, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

fixation by globally downregulating transport of organic nitrogen sources. Mol Microbiol 80:149–167.

72. Karunakaran R, Ramachandran VK, Seaman JC, East AK, Mouhsine B, Mauchline TH, Prell J, Skeffington A, Poole PS. 2009. Transcriptomic analysis of *Rhizobium leguminosarum* biovar *viciae* in symbiosis with host plants *Pisum sativum* and *Vicia cracca*. J Bacteriol 191:4002–14.

73. Finan TM, Wood JM, Jordan DC. 1983. Symbiotic properties of C4-dicarboxylic acid transport mutants of *Rhizobium leguminosarum*. J Bacteriol 154:1403–1413.

74. Mitsch MJ, DiCenzo GC, Cowie A, Finan TM. 2018. Succinate transport is not essential for symbiotic nitrogen fixation by *Sinorhizobium meliloti* or *Rhizobium leguminosarum*. Appl Environ Microbiol 84:e01561-17.

75. Johnson G V, Evans HJ, Ching T. 1966. Enzymes of the glyoxylate cycle in rhizobia and nodules of legumes. Plant Physiol 41:1330–1336.

76. Green RT, East AK, Karunakaran R, Downie JA, Poole PS. 2019. Transcriptomic analysis of *Rhizobium leguminosarum* bacteroids in determinate and indeterminate nodules. Microb Genomics 5:e000254.

77. Yadav, Vashishat, Kuykendall, Hashem. 1998. Biochemical and symbiotic properties of histidine-requiring mutants of *Rhizobium leguminosarum* biovar *trifolii*. Lett Appl Microbiol 26:22–26.

78. Gu C, Kim GB, Kim WJ, Kim HU, Lee SY. 2019. Current status and applications of genome-scale metabolic models. Genome Biol 20:121.

79. Kim WJ, Kim HU, Lee SY. 2017. Current state and applications of microbial genome-scale metabolic models. Curr Opin Syst Biol 2:10-18.

bioRxiv preprint doi: https://doi.org/10.1101/2021.07.28.454262; this version posted July 29, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

80. Machado D, Herrgård M. 2014. Systematic evaluation of methods for integration of transcriptomic data into constraint-based models of metabolism. PLoS Comput Biol 10:e1003580.

81. Liu Y, Beyer A, Aebersold R. 2016. On the dependency of cellular protein levels on mRNA Abundance. Cell 165:535–550.

82. Dunn MF. 2015. Key roles of microsymbiont amino acid metabolism in rhizobia-legume interactions. Crit Rev Microbiol 41:411–451.

83. Griesemer M, Kimbrel JA, Zhou CE, Navid A. 2018. Combining multiple functional annotation tools increases coverage of metabolic annotation. BMC Genomics 19:948.

84. Wang H, Marcišauskas S, Sánchez BJ, Domenzain I, Hermansson D, Agren R, Nielsen J, Kerkhoven EJ. 2018. RAVEN 2.0: A versatile toolbox for metabolic network reconstruction and a case study on *Streptomyces coelicolor*. PLOS Comput Biol 14:e1006541.

85. Henry CS, Dejongh M, Best AA, Frybarger PM, Linsay B, Stevens RL. 2010. High-throughput generation, optimization and analysis of genome-scale metabolic models. Nat Biotechnol 28:977–982.

86. Price MN, Deutschbauer AM, Arkin AP. 2020. GapMind: Automated annotation of amino acid biosynthesis. mSystems 5:e00291-20.

87. Mauchline TH, Fowler JE, East AK, Sartor AL, Zaheer R, Hosie AHF, Poole PS, Finan TM. 2006. Mapping the *Sinorhizobium meliloti* 1021 solute-binding protein-dependent transportome. Proc Natl Acad Sci 103:17933-17938.

88. Elbourne LDH, Tetu SG, Hassan KA, Paulsen IT. 2017. TransportDB 2.0: a database for exploring membrane transporters in sequenced genomes from all

bioRxiv preprint doi: https://doi.org/10.1101/2021.07.28.454262; this version posted July 29, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

domains of life. Nucleic Acids Res 45:D320–D324.

89. Orgambide GG, Huang Z-H, Gage DA, Dazzo FB. 1993. Phospholipid and fatty acid compositions ofRhizobium leguminosarum biovartrifolii ANU843 in relation to flavone-activated pSymnod gene expression. Lipids 28:975–979.

90. Théberge M-C, Prévost D, Chalifour F-P. 1996. The effect of different temperatures on the fatty acid composition of *Rhizobium leguminosarum* bv. *viciae* in the faba bean symbiosis. New Phytol 134:657–664.

91. diCenzo GC, Benedict AB, Fondi M, Walker GC, Finan TM, Mengoni A, Griffitts JS. 2018. Robustness encoded across essential and accessory replicons of the ecologically versatile bacterium *Sinorhizobium meliloti*. PLoS Genet 14:e1007357.

92. Breedveld MW, Miller KJ. 1994. Cyclic beta-glucans of members of the family Rhizobiaceae. Microbiol Rev 58:145-161.

93. Xavier JC, Patil KR, Rocha I. 2017. Integration of biomass formulations of genome-scale metabolic models with experimental data reveals universally essential cofactors in prokaryotes. Metab Eng 39:200-208.

94. Lodwig EM, Leonard M, Marroqui S, Wheeler TR, Findlay K, Downie JA, Poole PS. 2005. Role of polyhydroxybutyrate and glycogen as carbon storage compounds in pea and bean bacteroids. Mol Plant-Microbe Interact 18:67–74.

95. Heirendt L, Arreckx S, Pfau T, Mendoza SN, Richelle A, Heinken A, Haraldsdóttir HS, Wachowiak J, Keating SM, Vlasov V, Magnusdóttir S, Ng CY, Preciat G, Žagare A, Chan SHJ, Aurich MK, Clancy CM, Modamio J, Sauls JT, Noronha A, Bordbar A, Cousins B, El Assal DC, Valcarcel L V., Apaolaza I, Ghaderi S, Ahookhosh M, Ben Guebila M, Kostromins A,

bioRxiv preprint doi: https://doi.org/10.1101/2021.07.28.454262; this version posted July 29, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

Sompairac N, Le HM, Ma D, Sun Y, Wang L, Yurkovich JT, Oliveira MAP, Vuong PT, El Assal LP, Kuperstein I, Zinovyev A, Hinton HS, Bryant WA, Aragón Artacho FJ, Planes FJ, Stalidzans E, Maass A, Vempala S, Hucka M, Saunders MA, Maranas CD, Lewis NE, Sauter T, Palsson BØ, Thiele I, Fleming RMT. 2019. Creation and analysis of biochemical constraint-based models using the COBRA Toolbox v.3.0. Nat Protoc 14:639–702.

96. Glenn AR, Dilworth MJ. 1981. Oxidation of substrates by isolated bacteroids and free-living cells of *Rhizobium leguminosarum* 3841. Microbiology 126:243–247.

97. Galardini M, Mengoni A, Biondi EG, Semeraro R, Florio A, Bazzicalupo M, Benedetti A, Mocali S. 2014. DuctApe: A suite for the analysis and correlation of genomic and OmniLog™ Phenotype Microarray data. Genomics 103:1–10.

98. Martin M. 2011. Cutadapt removes adapter sequences from high-throughput sequencing reads. EMBnet.journal 17:10-12.

99. Magoc T, Wood D, Salzberg SL. 2013. EDGE-pro: Estimated degree of gene expression in prokaryotic genomes. Evol Bioinform Online 9:127–136.

100. Robinson JT, Thorvaldsdóttir H, Winckler W, Guttman M, Lander ES, Getz G, Mesirov JP. 2011. Integrative genomics viewer. Nat Biotechnol 29:24–26.

bioRxiv preprint doi: https://doi.org/10.1101/2021.07.28.454262; this version posted July 29, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

Data availability

All data need to evaluate the conclusions in this paper are present in the paper and/or the Supplementary Materials. RNA-Seq data for Rlv3841 in the pea rhizosphere are available on the NCBI SRA database, BioProject number PRJNA748006. All code is available on Github (https://github.com/CarolinSchulte/Rlv3481-lifestyles).

Acknowledgements

This work was supported by funding from the Biotechnology and Biological Sciences Research Council (BBSRC) [grant numbers BB/M011224/1, BB/R017859/1, BB/T001801/1 and BB/T006722/1]. C.C.M.S. is supported by the Clarendon Fund (Oxford University) and the Keble College De Breyne Scholarship. A.P. was funded in part by the Engineering and Physical Sciences Research Council (EPSRC) [grant number EP/M002454/1].

Author contributions

Conceptualization: CCMS, AP, PSP

Methodology: CCMS, AP, PSP

Formal analysis: CCMS, VKR

Investigation: CCMS, VKR, AP, PSP

Visualization: CCMS

Supervision: AP, PSP

Writing—original draft: CCMS

Writing—review & editing: CCMS, VKR, AP, PSP

bioRxiv preprint doi: https://doi.org/10.1101/2021.07.28.454262; this version posted July 29, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

Supplementary Data

Supplementary Data 1: sbml file of iCS1224

Supplementary Data 2: Excel file of iCS1224

Supplementary Data 3: MEMOTE report for iCS1224

Supplementary Data 4: Phenotype microarray data for Rlv3841

Supplementary Data 5: RNA-Seq data for Rlv3841 in the pea rhizosphere

Supplementary Data 6: Rhizosphere model of Rlv3841

Supplementary Data 7: Metabolomics data

Supplementary Data 8: Nodule bacteria model of Rlv3841

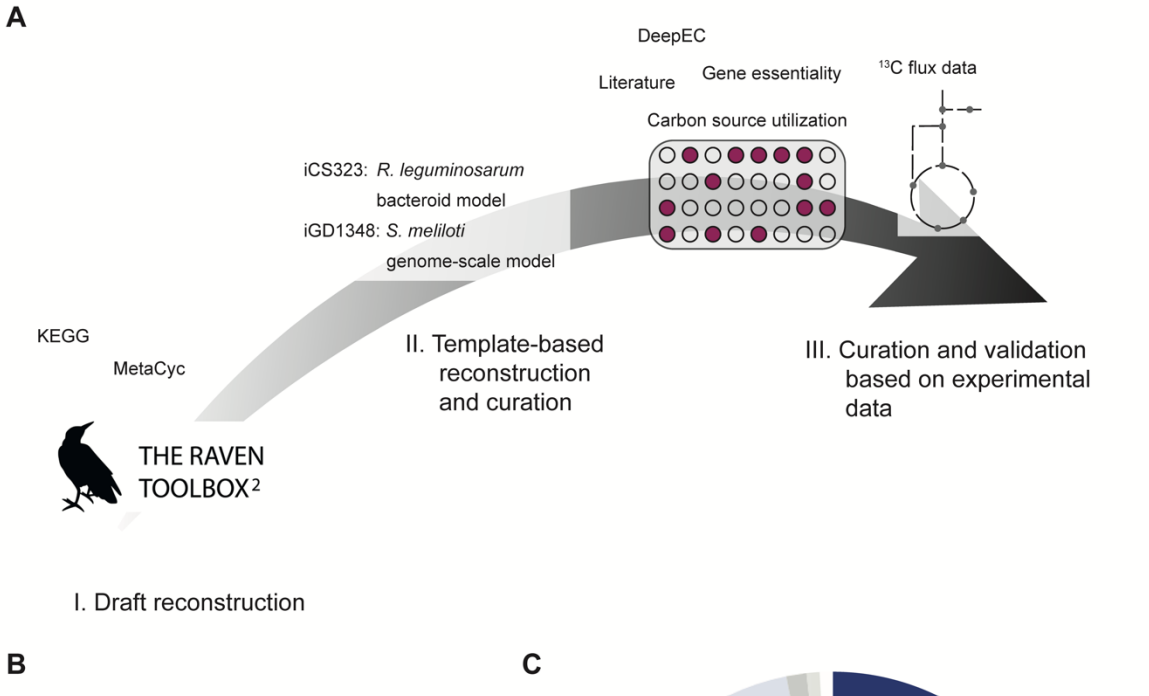
Supplementary Data 9: Bacteroid model of Rlv3841

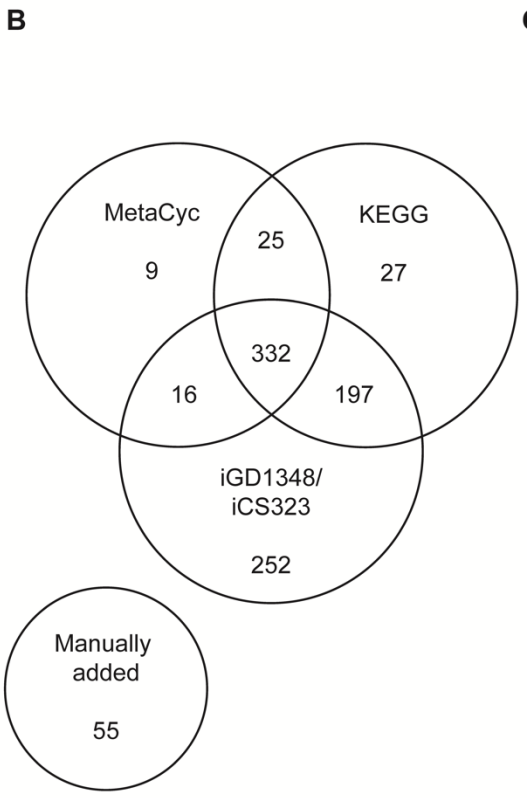
bioRxiv preprint doi: https://doi.org/10.1101/2021.07.28.454262; this version posted July 29, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

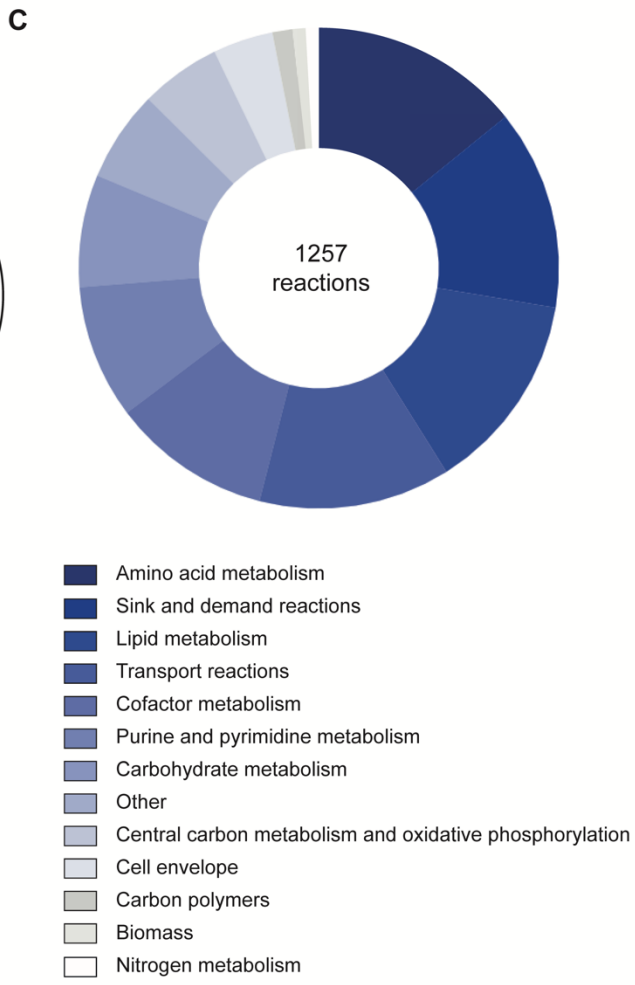
Table 1: Properties of iCS1224.

Feature	Value
Genes	1224
Metabolites	984
Unique EC identifiers	603
Reactions	1257
Metabolic reactions	913
Gene-associated metabolic reactions	897
Transport reactions	162
Gene-associated transport reactions	142
Sink reactions	155
Demand reactions	15
Other reactions (e.g. DNA synthesis, protein synthesis, biomass objective function)	12

bioRxiv preprint doi: https://doi.org/10.1101/2021.07.28.454262; this version posted July 29, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.







bioRxiv preprint doi: https://doi.org/10.1101/2021.07.28.454262; this version posted July 29, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

Fig. 1: Reconstruction of a genome-scale model for *Rhizobium leguminosarum* bv. *viciae* 3841. A: Reconstruction process for iCS1224 using automated reconstruction, template-based reconstruction, and data-based curation. B: Sources for the 913 metabolic reactions in iCS1224. Numbers indicate how many reactions from KEGG, MetaCyc or the template-based reconstruction were included in the final model, with numbers in the overlapping areas indicating reactions that were present in multiple draft reconstructions. C: Classification of the reactions in iCS1224.

bioRxiv preprint doi: https://doi.org/10.1101/2021.07.28.454262; this version posted July 29, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

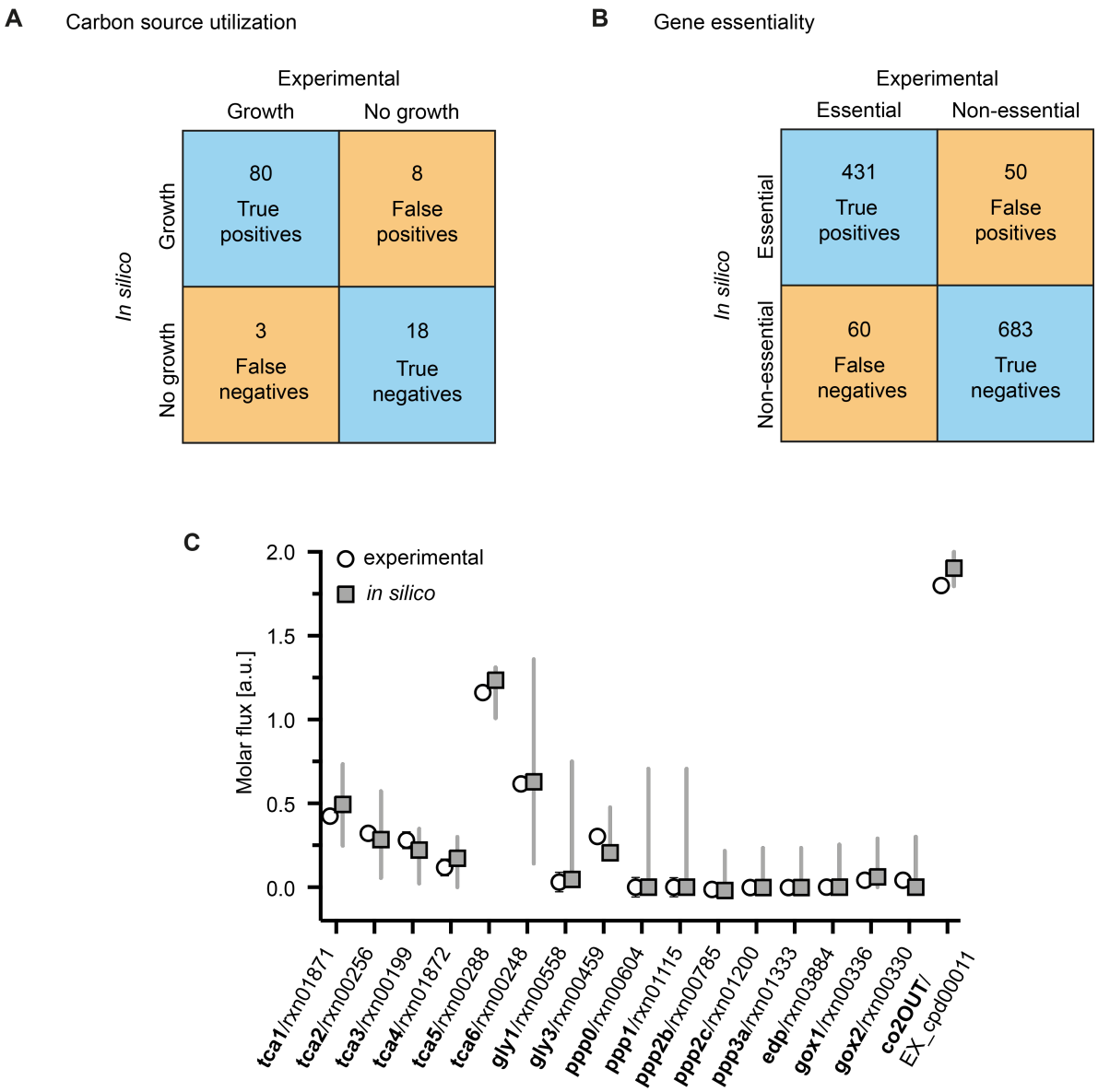


Fig 2. Validation of iCS1224. A: Table showing the agreement between carbon source utilization experimentally measured with phenotype microarrays and predicted by iCS1224. B: Table showing the agreement between gene essentiality determined by insertion sequencing (13, 30) and predicted by iCS1224. C: Comparison of metabolic fluxes determined by ^{13}C metabolic flux analysis for Rlv3841 grown on succinate (40) with flux rates predicted by iCS1224. For experimental data, symbols and bars indicate mean ± SD. Note the error bars are

bioRxiv preprint doi: https://doi.org/10.1101/2021.07.28.454262; this version posted July 29, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

too small to be visible for most data points. For *in silico* data, symbols represent the flux rate predicted by flux balance analysis, with lines indicating upper and lower bounds for each flux determined by flux variability analysis with at least 95% of the optimum flux through the biomass objective function. Labels on the x axis indicate the name of the reaction as reported in (40) (in bold), as well as the reaction identifier in the model.

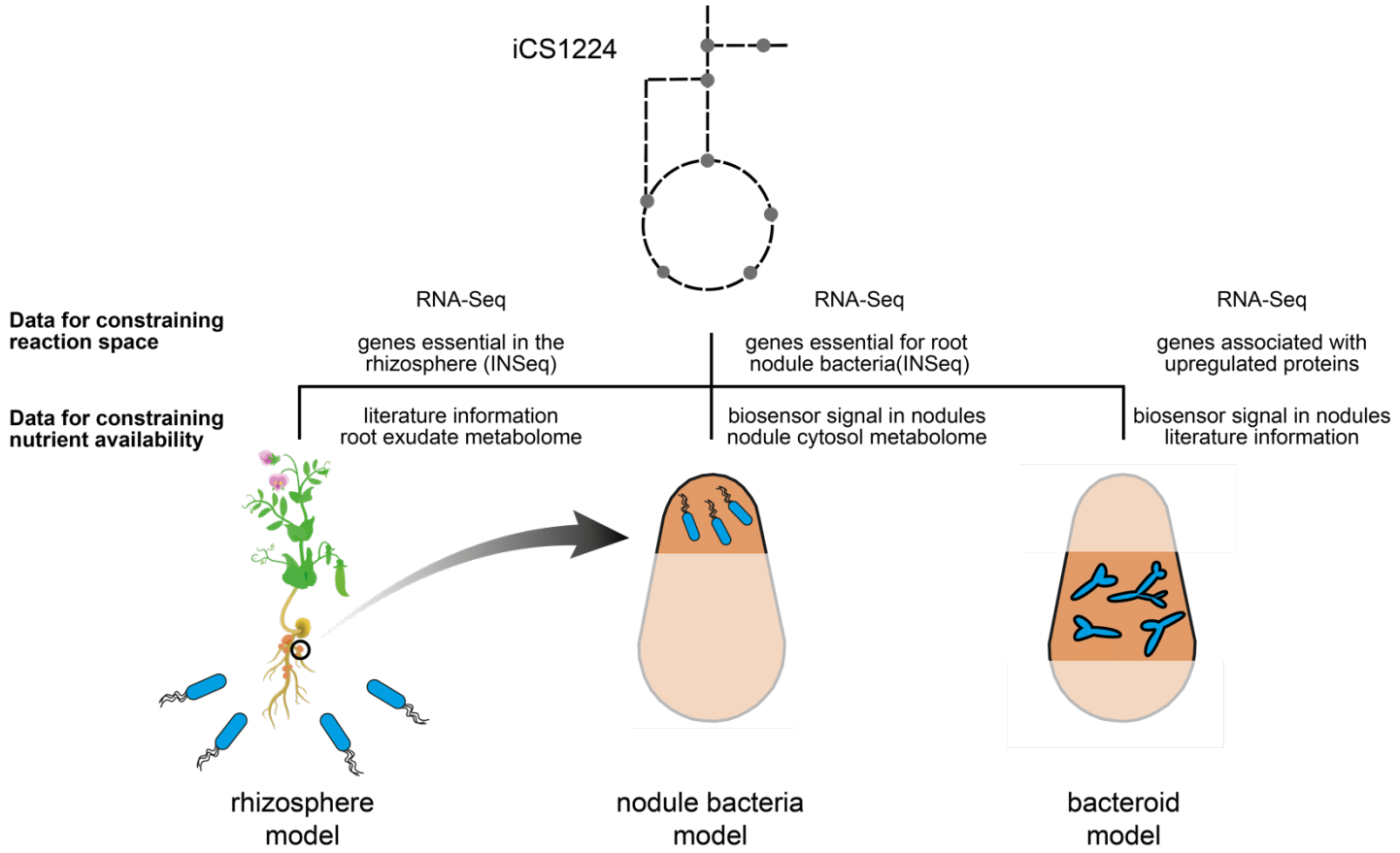


Fig. 3: Approach for generating lifestyle-specific models for *Rhizobium leguminosarum* bv. *viciae*. Based on iCS1224, transcriptome, gene essentiality and proteome data specific to a certain lifestyle were used to inform the extraction of context-specific models for the rhizosphere, nodule bacteria and nitrogen-fixing bacteroids. Boundary conditions were defined based on metabolome data and/or literature information.

bioRxiv preprint doi: https://doi.org/10.1101/2021.07.28.454262; this version posted July 29, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

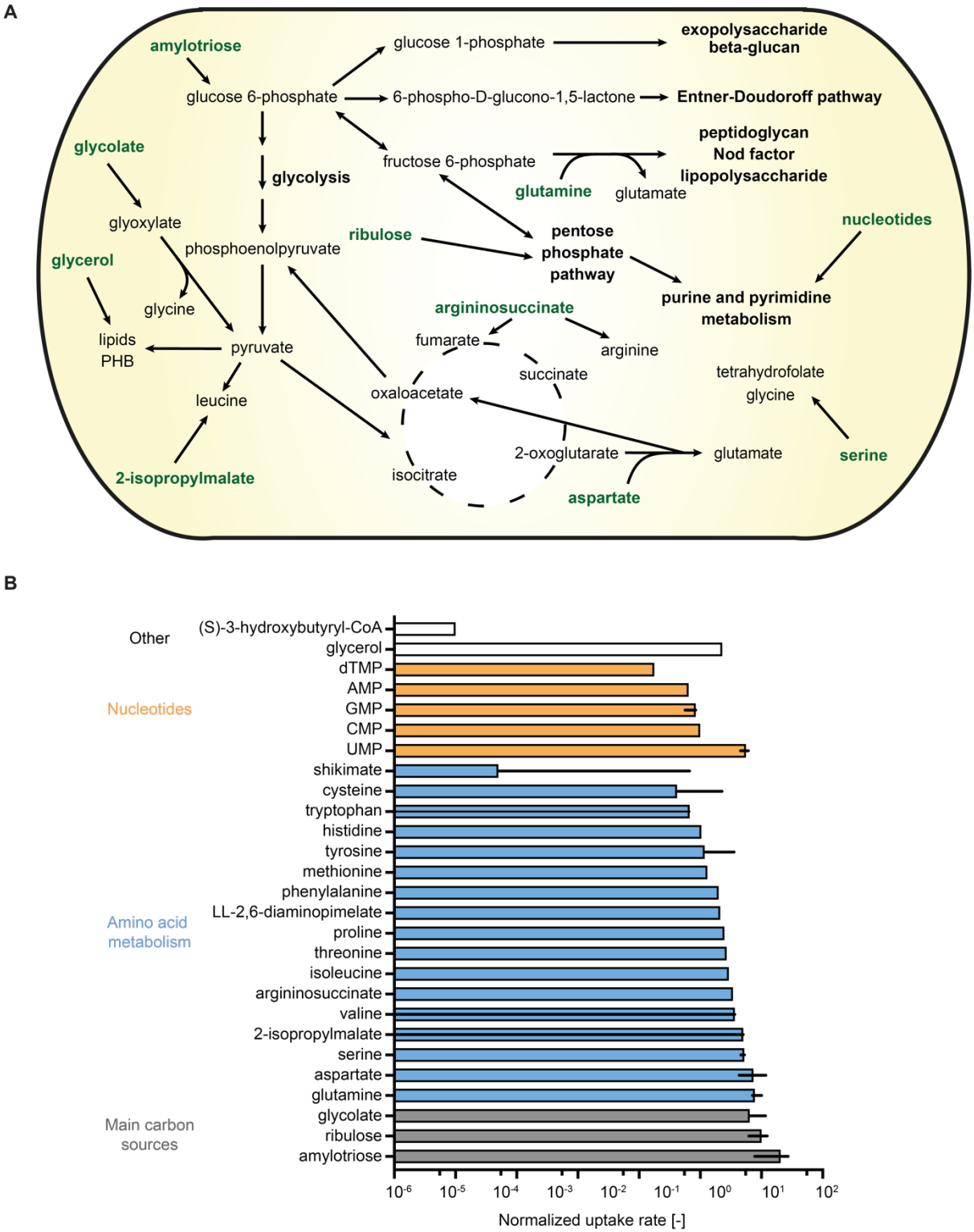


Fig. 4: Metabolism of *Rhizobium leguminosarum* in the pea rhizosphere. A rhizosphere-specific model was extracted from iCS1224 using the RIPTiDe algorithm with RNA-Seq and gene essentiality data for *R. leguminosarum* in the rhizosphere of pea plants. A: Schematic representation of the main pathways predicted to be active

bioRxiv preprint doi: https://doi.org/10.1101/2021.07.28.454262; this version posted July 29, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

in the rhizosphere-specific model. Compounds predicted to be taken up are indicated in bold green. Note that the magnitude of flux is not indicated in this summary map. B: Bar graph showing the uptake rates of metabolites predicted to be taken up from pea root exudate. Absolute flux values for the exchange reactions were normalized by flux through the biomass reaction in each sample. Only metabolites with non-zero median uptake for the 500 samples of the contextualized model are shown. Uptake of ions and cofactors has been omitted for clarity. Bars and lines indicate median and interquartile range, respectively.

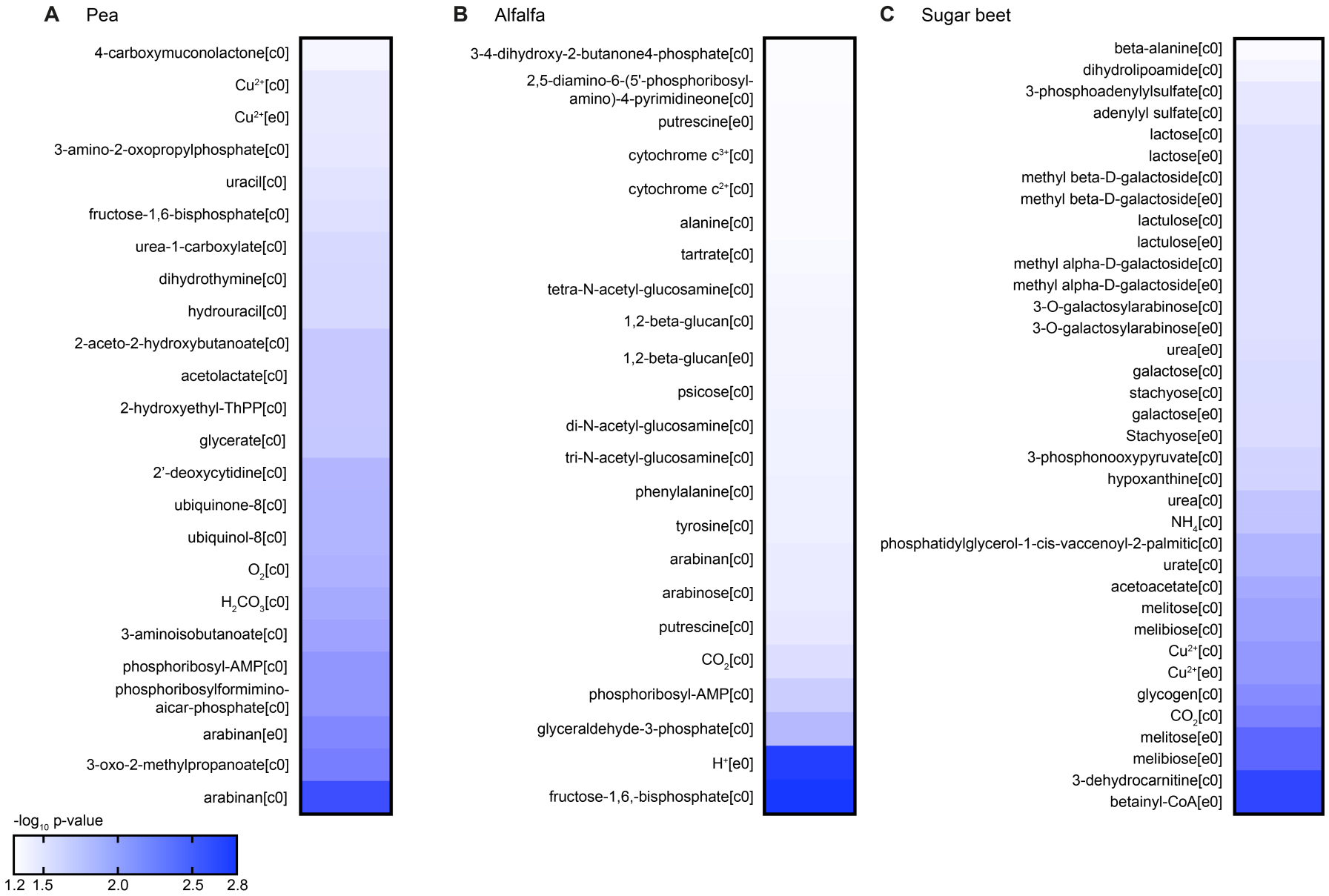

A Pea
4-carboxymuconolactone[c0]
Cu2+[c0]
Cu2+[e0]
3-amino-2-oxopropylphosphate[c0]
uracil[c0]
fructose-1,6-bisphosphate[c0]
urea-1-carboxylate[c0]
dihydrothymine[c0]
hydrouracil[c0]
2-aceto-2-hydroxybutanoate[c0]
acetolactate[c0]
2-hydroxyethyl-ThPP[c0]
glycerate[c0]
2'-deoxycytidine[c0]
ubiquinone-8[c0]
ubiquinol-8[c0]
O2[c0]
H2CO3[c0]
3-aminoisobutanoate[c0]
phosphoribosyl-AMP[c0]
phosphoribosylformimino-aicar-phosphate[c0]
arabinan[e0]
3-oxo-2-methylpropanoate[c0]
arabinan[c0]
-log10 p-value
1.2 1.5 2.0 2.5 2.8
B Alfalfa
3-4-dihydroxy-2-butanone4-phosphate[c0]
2,5-diamino-6-(5'-phosphoribosyl-amino)-4-pyrimidineone[c0]
putrescine[e0]
cytochrome c3+[c0]
cytochrome c2+[c0]
alanine[c0]
tartrate[c0]
tetra-N-acetyl-glucosamine[c0]
1,2-beta-glucan[c0]
1,2-beta-glucan[e0]
psicose[c0]
di-N-acetyl-glucosamine[c0]
tri-N-acetyl-glucosamine[c0]
phenylalanine[c0]
tyrosine[c0]
arabinan[c0]
arabinose[c0]
putrescine[c0]
CO2[c0]
phosphoribosyl-AMP[c0]
glyceraldehyde-3-phosphate[c0]
H+[e0]
fructose-1,6,-bisphosphate[c0]
C Sugar beet
beta-alanine[c0]
dihydrolipoamide[c0]
3-phosphoadenylylsulfate[c0]
adenylyl sulfate[c0]
lactose[c0]
lactose[e0]
methyl beta-D-galactoside[c0]
methyl beta-D-galactoside[e0]
lactulose[c0]
lactulose[e0]
methyl alpha-D-galactoside[c0]
methyl alpha-D-galactoside[e0]
3-O-galactosylarabinose[c0]
3-O-galactosylarabinose[e0]
urea[e0]
galactose[c0]
stachyose[c0]
galactose[e0]
Stachyose[e0]
3-phosphonooxypyruvate[c0]
hypoxanthine[c0]
urea[c0]
NH4[c0]
phosphatidylglycerol-1-cis-vaccenoyl-2-palmitic[c0]
urate[c0]
acetoacetate[c0]
melitose[c0]
melibiose[c0]
Cu2+[c0]
Cu2+[e0]
glycogen[c0]
CO2[c0]
melitose[e0]
melibiose[e0]
3-dehydrocarnitine[c0]
betainyl-CoA[e0]

Fig. 5: Reporter metabolites in different rhizospheres. Reporter metabolites were calculated using microarray data for *Rhizobium leguminosarum* bv. *viciae* 3841 in the rhizosphere of pea (A), alfalfa (B), and sugar beet (C) compared to free living cells grown in minimal media with glucose and ammonia (10). The heatmaps show the negative decimal logarithm of the *P* value for those metabolites that were associated with significant ($P<0.05$) transcriptional changes among genes upregulated in the rhizosphere. [c0] and [e0] indicate cytosolic and extracellular metabolites, respectively.

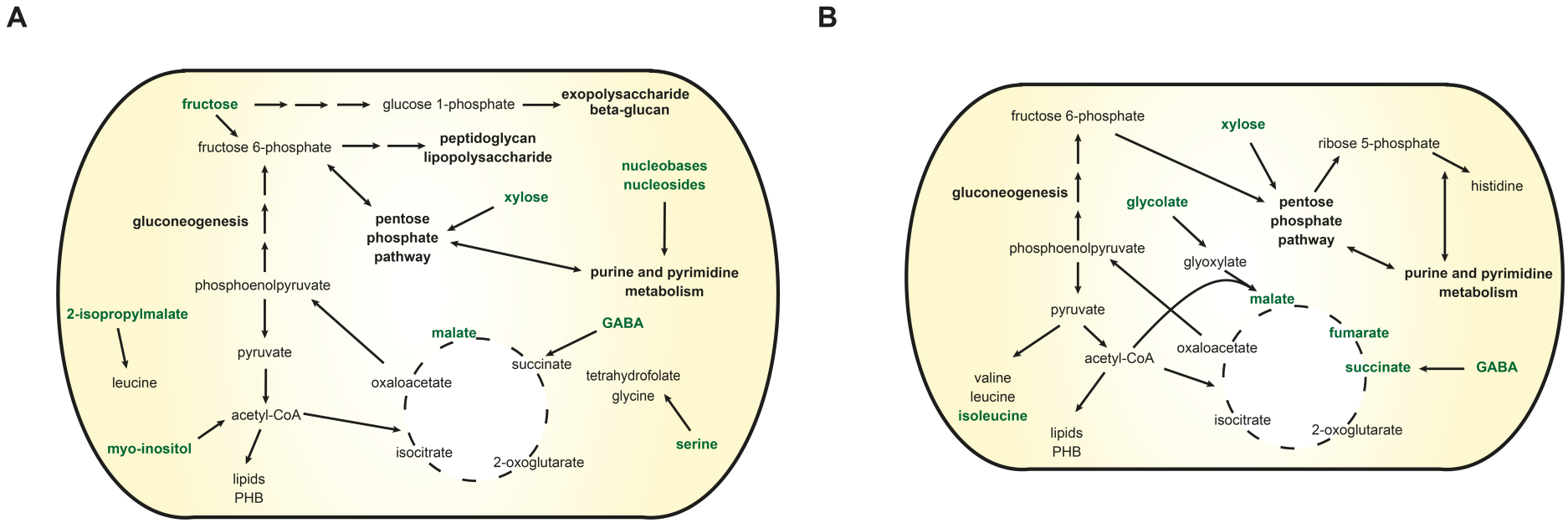


Fig. 6: Metabolism of undifferentiated nodule bacteria and nitrogen-fixing bacteroids. Maps showing schematic representations of the main pathways predicted to be active in undifferentiated nodule bacteria (A) and nitrogen-fixing bacteroids (B). Compounds predicted to be taken up are indicated in bold green. Note that the magnitude of flux is not indicated in these summary maps.