Genome-scale metabolic reconstruction of the symbiosis between a leguminous plant and a nitrogen-fixing bacterium

George C. diCenzo1,2, Michelangelo Tesi1, Thomas Pfau3, Alessio Mengoni1✉ & Marco Fondi1✉

The mutualistic association between leguminous plants and endosymbiotic rhizobial bacteria is a paradigmatic example of a symbiosis driven by metabolic exchanges. Here, we report the reconstruction and modelling of a genome-scale metabolic network of *Medicago truncatula* (plant) nodulated by *Sinorhizobium meliloti* (bacterium). The reconstructed nodule tissue contains five spatially distinct developmental zones and encompasses the metabolism of both the plant and the bacterium. Flux balance analysis (FBA) suggests that the metabolic costs associated with symbiotic nitrogen fixation are primarily related to supporting nitrogenase activity, and increasing N₂-fixation efficiency is associated with diminishing returns in terms of plant growth. Our analyses support that differentiating bacteroids have access to sugars as major carbon sources, ammonium is the main nitrogen export product of N₂-fixing bacteria, and N₂ fixation depends on proton transfer from the plant cytoplasm to the bacteria through acidification of the peribacteroid space. We expect that our model, called ‘Virtual Nodule Environment’ (ViNE), will contribute to a better understanding of the functioning of legume nodules, and may guide experimental studies and engineering of symbiotic nitrogen fixation.

1Department of Biology, University of Florence, Sesto Fiorentino, Italy. 2Department of Biology, Queen’s University, Kingston, ON, Canada. 3Life Sciences Research Unit, University of Luxembourg, Belvaux, Luxembourg. ✉email: alessio.mengoni@unifi.it; marco.fondi@unifi.it
Macororganisms are colonized by a staggering diversity of microorganisms, collectively referred to as a ‘holobiont’. The intimate association between organisms is often driven by metabolic exchanges; many insects obtain essential nutrients from obligate bacterial symbionts, most plants can obtain phosphorus from arbuscular mycorrhiza in exchange for carbon, and the gut microbiota is thought to contribute to animal nutrition. Complex global patterns often emerge during these intimate biological associations, especially when nutritional inter-dependencies are involved. The communication between the two metabolic networks of the interacting organisms may give rise to unpredictable phenotypic traits and unexpected emergent properties. Metabolic relationships can span over a large taxonomic range and have profound biological relevance. For example, the interactions between bacteria and multicellular organisms have been suggested to be key drivers of evolutionary transitions, leading to eukaryotic diversification and to the occupancy of novel niches. The study of the association of two biological entities is mainly challenged by the size of the system and by the unpredictability of their metabolic interactions. Theoretical, systems-level models are required to unravel the intimate functioning of metabolic associations and to eventually exploit their potential in biotechnological applications.

Symbiotic nitrogen fixation (SNF) is a paradigmatic example of the importance and the complexity of natural biological associations. SNF is a mutualistic relationship between a group of plant families, including the Fabaceae, and a polyphyletic group of alpha- and beta-proteobacteria known as rhizobia, or a taxa of Actinobacteria (Frankia spp.), in which the plants provide a niche and carbon to the bacteria in exchange for fixed nitrogen. SNF involves constant metabolic cross-talk between the plant and the bacteria, and it is a paradigmatic example of bacterial cellular differentiation and sociomicrobiological interactions. The rhizobia intra-cellularly colonize plant cells of a specialized organ known as a root (or stem) nodule. The intra-cellular rhizobia (referred to as bacteroids) are surrounded by a plant-derived membrane, and the term symbiosome is used in reference to the structure consisting of the bacteroid, the plant-derived membrane (i.e., the peribacteroid membrane), and the intervening space (i.e., the peribacteroid space). Nodules with an indeterminate structure, such as those formed by the plant Medicago truncatula, are divided into spatially distinct developmental zones with a distal apical meristem and a proximal nitrogen fixation zone.

SNF plays a key role in the global nitrogen cycle and is central to sustainable agricultural practices by reducing the usage of synthetic nitrogen fertilizers whose application results in a multitude of adverse environmental consequences. Unfortunately, our ability to maximize the benefit of SNF is limited since rhizobial inoculants are often poorly effective due to low competitiveness and because rhizobium symbioses are specific to leguminous plants. Manipulating the rhizobium–legume interaction for biotechnological purposes will require an in-depth understanding of the symbiotic interaction, as well as an ability to predict the consequences of genetic changes and environmental perturbations.

From a metabolic perspective, genome-scale metabolic reconstruction (GENREs) and constraint-based modeling has great potential to fulfill these roles. A GENRE also serves as a comprehensive knowledgebase of an organism’s metabolism, containing hundreds to thousands of metabolic and transport reactions, most of which are linked to the corresponding gene(s) whose gene product(s) catalyzes the reaction. With the aid of mathematical approaches such as flux balance analysis (FBA), GENREs can be used to identify emergent system-level properties, to predict active reactions, and to identify essential genes. Compared to simple enrichment analyses that are typical inomics studies, GENRE-based methods allow for the interpretation of data in a connected manner based on network topology and to infer the effects of changes in remote pathways on the overall cell physiology. When considering interacting entities, for example, this approach can predict the consequence of mutations in one organism on the metabolism of the other. However, multi-organism metabolic reconstructions are still in their infancy, and very few examples of combined models exist compared to single strain GENREs.

Despite the importance of metabolism to SNF, there has been limited use of metabolic modeling in the study of rhizobia and SNF. To date, GENREs of varying quality have been reported for only three rhizobia: Sinorhizobium meliloti, Rhizobium etli, and Bradyrhizobium diazoefficiens. Currently, M. truncatula and Glycine max are the only legumes with published GENREs. With the exception of the G. max GENRE, these GENREs have been used in preliminary analyses of SNF, providing results generally consistent with expectations. However, all analyses to date suffer from two major limitations. Simulations with the rhizobium models ignore plant metabolism, while simulations with the M. truncatula GENRE (based on the genome sequence published in 2011), which has since been updated in 2014 and again in 2018, involved a very limited draft S. meliloti metabolic reconstruction (whose genome sequence was published in 2001). Furthermore, all simulations have focused on the final stage of SNF and have not considered the different steps of the preceding developmental progression where metabolism remains poorly understood.

Here, we report a holistic in silico representation of the integrated metabolism of the holobiont consisting of a M. truncatula plant nodulated by S. meliloti, which we refer to as a Virtual Nodule Environment (ViNE). Our combined, multi-compartment reconstruction accounts for the metabolic activity of shoot and root tissues together with a nodule consisting of five developmental zones. We report initial characterizations of ViNE using FBA, including zone-specific metabolic properties, trade-offs between nitrogen fixation and plant growth, and the usage of dicarboxylates as a carbon source by bacteroids. Going forward, we expect ViNE will provide a powerful platform for hypothesis generation aimed at understanding and quantitatively evaluating SNF, as well as guiding attempts at engineering SNF for increased symbiotic efficiency.

**Results**

**Validation of iGD1348, an updated S. meliloti reconstruction.** Prior to constructing the integrated plant–bacterium metabolic model, an updated metabolic reconstruction of S. meliloti Rm1021 was prepared as described in Methods and Supplementary Note 1. Briefly, the highly refined core metabolic network iGD726 was combined with the comprehensive accessory metabolism of the genome-scale metabolic network iGD1575. Most of the reactions were manually curated through comparison against the literature, referenced where possible, and mass and charge balanced. The updated model consists of 1348 genes (Supplementary Table 1), and incorporates information from 240 literature sources (Supplementary Data 1) that includes transposon-sequencing (Tn-seq) data and Phenotype MicroArray data for wild-type and mutant strains. The reduced size of iGD1348 relative to the older iGD1575 model is a consequence of deleting poorly characterized genes that appeared to be incorrectly associated with reactions, and the removal of pathways that produced dead-end metabolites.

Several tests were performed to validate the quality of the newly prepared S. meliloti reconstruction. Flux balance analysis (FBA) was used to simulate growth using glucose or succinate as the sole source of carbon, with or without the inclusion of an NGAM reaction. Inclusion of an NGAM reaction resulted in a specific growth rate reduction of −0.043 h⁻¹ and 0.030 h⁻¹ for...
growth with glucose and succinate, respectively. This result confirmed the absence of energy leaks in iGD1348 that would allow for spontaneous energy production.

Using FBA, the ability of S. meliloti to catabolize 163 carbon sources to support growth was predicted with the iGD1348 and iGD1575 models (Supplementary Data 2). As previously reported33, simulations with the iGD1575 model correctly predicted growth with 67 of the 85 (79%) substrates experimentally shown to support growth of S. meliloti. Nicely, simulations with the iGD1348 model correctly predicted growth with 76 of these 85 (89%) substrates, including all 67 that supported growth of the iGD1575 model. This result confirmed that iGD1348 incorporates the majority of the accessory metabolism of S. meliloti, and that it is a better representation of total cellular S. meliloti metabolism than the previous genome-scale model.

Context-specific core metabolic models, containing a minimal set of reactions for biomass production in a given environment, were extracted from the iGD1348 and iGD1575 metabolic models through the integration of Tn-seq data34 using Tn-Core48. The accuracy of the resulting core metabolic models was determined through comparison with iGD726, a manually prepared core metabolic model of S. meliloti34. As summarized in Fig. 1, the iGD1348 core model displayed greater overlap with the iGD726 model than did the core model generated from iGD1575 (Fig. 1a).

In particular, of the genes essential in the iGD1348 core model, 96% were essential in the iGD1348 core model, whereas only 62% were essential in the iGD1575 core model (Fig. 1b). This result confirmed that the newly prepared iGD1348 reconstruction better represents the core metabolic network of S. meliloti than does the iGD1575 reconstruction. Overall, these tests confirmed that iGD1348 is a high-quality representation of S. meliloti metabolism, and that it is an improvement over previous iterations of S. meliloti GENREs.

**Construction and validation of a model of a nodulated legume.**

Obtaining high-quality reconstructions of M. truncatula and S. meliloti metabolism was a prerequisite to generating an in silico genome-scale metabolic network of an entire nodulated legume (referred to as ViNE). For S. meliloti, we used the updated model described in the previous section. In the case of M. truncatula, we used a recently published reconstruction that was updated to match the most recent version of the M. truncatula genome annotation (see Supplementary Note 2).

Integrating the S. meliloti and M. truncatula metabolic models resulted in a model encompassing shoot, root, and nodule tissues as summarized in Fig. 2 and Table 1. Memote49 was used to evaluate the overall quality of the integrated reconstruction. This analysis revealed an overall quality score of 72%, which falls well in the range of the scores of the original two models (78% and 65% for the iGD1348 and the M. truncatula reconstructions, respectively) and is consistent with ViNE being a high-quality metabolic reconstruction. The detailed reports are provided in Supplementary Data 3.

In total, ViNE includes 746 unique S. meliloti genes and 1,327 unique M. truncatula genes. The nodule was subdivided into five zones to match the spatially, and transcriptionally, distinct developmental zones that are simultaneously present in indeterminate legume nodules such as those formed by M. truncatula21,22,24. Several simulations were performed to evaluate the reliability of the model. Using FBA, the maximal rate of plant (shoot + root) growth of the nodulated system was predicted to be ~0.044 g day⁻¹ (g plant dry weight)⁻¹, while the N₂-fixation rate was predicted to be ~3 µmol h⁻¹ (g plant dry weight)⁻¹. Both values are reasonable predictions; Medicago sativa plants have an experimentally determined growth rate of ~0.1 g day⁻¹ (g plant dry weight)⁻¹31 and N₂-fixation rates around 7 µmol h⁻¹ (g plant dry weight)⁻¹ based on acetylene reduction assays52. Moreover, the predicted carbon cost of supporting N₂-fixing nodules (nodule CO₂ export per N₂ fixed) was ~4.2 g C g⁻¹ N, falling within the normal range of 3–5 g C g⁻¹ N based on experimental measurements53.

Next, FBA was used to examine the effects of adding exogenous ammonium to the soil on plant growth considering two situations: (i) the N₂-fixation efficiency (defined as the rate of N₂-fixation per gram nodule) could vary while the rate of nodulation (defined as the percent of total plant biomass being nodule biomass) was constant at 2%, and (ii) the rate of nodulation could vary while the N₂-fixation efficiency was constant at 150 µmol h⁻¹ (g nodule dry weight)⁻¹. As expected, increasing the availability of exogenous ammonium increased the rate of plant growth, with the effect more pronounced when the rate of nodulation was allowed to decrease with increasing ammonium since the plant no longer had to invest in nodule maintenance (Fig. 3).

We then simulated the effects of individual bacteria gene deletions on plant biomass production (Supplementary Data 4).
Table 1 Summary properties of ViNE.

| Model feature | Shoot | Root | Zone I | Zone IId | Zone IIp | Zone IZ | Zone III |
|---------------|-------|------|--------|----------|----------|--------|---------|
| Genes         |       |      |        |          |          |        |         |
| M. truncatula | 1295  | 1292 | 236    | 265      | 243      | 228    | 116     |
| S. meliloti   | 0     | 0    | 0      | 640      | 629      | 638    | 246     |
| Reactions     | 937   | 944  | 494    | 543      | 559      | 530    | 318     |
| M. truncatula | 0     | 0    | 0      | 662      | 654      | 670    | 201     |
| S. meliloti   | 831   | 825  | 490    | 568      | 597      | 581    | 286     |
| Metabolites   | 0     | 0    | 0      | 751      | 747      | 764    | 222     |

*Excludes reactions for transfer of metabolites between tissues or between M. truncatula and the peribacteroid space.

Fig. 3 Effect of exogenous ammonium on M. truncatula growth. The effects of increasing the availability of soil ammonium on the growth rate of nodulated M. truncatula was examined. Simulations were run allowing either the N2-fixation efficiency to vary while the rate of nodulation remained constant at 2% (blue) or allowing the rate of nodulation to vary while the N2-fixation efficiency remained constant at 150 µmol h−1 (g nodule dry weight)−1 (red). The dashed lines indicate the maximal rate of plant growth with exogenous ammonium (upper) and the maximal rate of plant growth when relying on N2-fixation (lower).

and compared the results to published experimental data. The model was able to accurately predict the phenotypes of many S. meliloti mutants. For example, S. meliloti genes such as nifH (nitrogenase), dctA (succinate transport), ilvI (branched-chain amino acid biosynthesis), aatA (aspartate transaminase), pgk (phosphoglycerate kinase), and mrdI (ribonucleotide reductase) were correctly predicted to be essential, while pyc (pyruvate carboxylase), glnA (glutamine synthetase), pckA (phosphoenolpyruvate carboxykinase), and leuB (leucine biosynthesis) were correctly predicted to be non-essential. Similarly, the removal of plant-encoded nodule sucrose synthase, phospho- enolpyruvate carboxylase, and homocitrate synthase reactions abolished nitrogen fixation, as expected. However, it is important to note that the predictions were not perfect. For example, deleting argG (arginine biosynthesis) or carA (carbon-methyl phosphate synthase) did not result in the expected phenotypes, while the incorrect malic enzyme (time instead of dme) was predicted to be essential. Taken together, these analyses provide support for the general reliability of ViNE as a representation of nodule metabolism.

Metabolic progression and nutrient exchange during nodule development. The presence of five nodule zones in ViNE provided an opportunity to examine the metabolic changes associated with the development of an effective nodule. To accomplish this, FBA was used to predict the flux distribution through the integrated metabolic networks of each nodule zone, and to simulate the effects of individually deleting each gene, or removing each reaction, specifically in a single nodule zone.

Additionally, a robustness analysis was performed to evaluate how perturbations in the flux of individual bacteroid reactions influence the predicted rate of plant growth. The outputs of these analyses are provided as Supplementary Data 5 and 6, and they are summarized in Fig. 4 and Supplementary Figs. S1 and S2. For simplicity, here we focus on the reaction-level analyses, and we split the nodule into only three sections: uninfected (zone I), differentiating (zones IId, IIp, and IZ), and nitrogen-fixing (zone III). Highlighting the overall similarity of zones IId, IIp, and IZ, and thus supporting their grouping, the robustness analysis indicated that roughly 90% of the bacteroid reactions that had to carry flux in one of these zones had to carry flux in all three zones to maximize plant growth.

The most notable difference comparing the uninfected and differentiation zones was an increase in the number of active reactions related to energy production, including carbon and nucleotide metabolism. This result suggests that the accommodation of differentiating bacteroids may place additional energy demands on the plant cell, and that few additional metabolic functions are required. In contrast, the transition from the differentiation zone to the N2-fixing zone was associated with a marked decrease in the number of active reactions in both the plant and bacterial cells, consistent with published transcriptomic and proteomic datasets. Highlighting this result, ~ 560 bacteroid reactions had to carry flux in the differentiation zones to optimize plant growth, whereas only 167 bacteroid reactions had to carry flux in the N2-fixation zone for maximal plant growth.

The lack of biomass production in the N2-fixation zone meant that most biomass biosynthetic pathways were predicted to be inactive and non-essential. However, bacterial pathways related to the production of cofactors for nitrogenase or energy production remained essential; this included FMN, heme, cobalamin, pyridoxine phosphate, and glutathione biosynthesis, as well as the pentose phosphate pathway. Similarly, the TCA cycle, oxidative phosphorylation, and purine biosynthesis in bacteroids were predicted to be essential in the N2-fixation zone, presumably to supply the massive amounts of energy required by nitrogenase. Biosynthesis of methionine and SAM were also predicted to be essential. Few other notable bacterial reactions were required in the N2-fixation zone (Supplementary Data 5). In the plant compartment, the majority of the active reactions were related to central carbon metabolism for the production of energy and C4-dicarboxylates for use by the bacteroids, while other active reactions were involved in the assimilation of ammonium through the formation of glutamine. Consistent with experimental works, the FBA results indicated that the plant nodule cells are provided sucrose as a carbon/energy source; in fact, ~ 80% of all carbon fixed by the plant leaves was sent to nodule zone III. The sucrose was then hydrolyzed and metabolized to phosphoenolpyruvate, of which ~ 80% was...
ammonia75,76; however, some studies have suggested that L-alanine could be a major nitrogen export product77,78. Our FBA simulations suggested that rhizobial biosynthesis of approximately half of the amino acids was essential for the symbiosis.

Sociobiology of symbiotic nitrogen fixation. By containing a representation of an entire nodule, ViNE allowed for an evaluation of the metabolic costs associated with SNF. Using FBA, the maximal plant growth rate of the nodulated system (without exogenous ammonium) was estimated to be ~72% of the maximal growth rate of a nodule-free system supplied with non-limiting amounts of exogenous ammonium (Fig. 3, Table 2). The largest factor contributing to the difference in growth was the direct energetic cost of supporting N2-fixation (~67% of the difference; Table 2). The remaining third of the difference was explained by the cost of synthesizing (~11% of the difference) and maintaining (~22% of the difference) the nodule and bacteroid tissue (Table 2).

We next evaluated the relationship between N2-fixation efficiency (while keeping the rate of nodulation at 2%) and the rate of plant growth. When the N2-fixation efficiency was below the optimum, there was a linear relationship between N2-fixation and biomass production (Fig. 5a). However, excessive N2-fixation quickly resulted in impaired plant growth, with 10% excess of N2-fixation collapsing the symbiosis (Fig. 5a). We hypothesized that this result was due to insufficient energy to support both the excess N2-fixation and the ATP maintenance costs. Consistent with this hypothesis, removing the upper limit on the rate of zone III oxygen uptake resulted in a gradual decrease in plant growth as the rate of N2-fixation was increased above the optimal (Fig. 5a). In this case, excessive N2-fixation was less detrimental than insufficient N2-fixation; the effect of increasing the N2-fixation efficiency by 50 μmol h−1 (g nodule dry weight)−1 increased or decreased the rate of plant growth by 14.7 or 3.4 mg day−1 (g plant dry weight)−1 when below or above the optimum, respectively. We next examined the consequences of varying the rate of nodulation (i.e., the ratio between plant and nodule biomass) while maintaining a constant N2-fixation efficiency. The simulations demonstrated linear relationships between the rate of nodulation and plant growth both above and below the optimum (Fig. 5b), with increasing the percent nodulation resulting in a 3-fold greater impact when below the optimum compared to above the optimum. Overall, these

### Table 2 Contributions of N2-fixation and nodulation to the metabolic costs of SNF.

| Nitrogen source | Nodulation state | Relative plant growth rate |
|-----------------|------------------|---------------------------|
| N2-fixation     | Nodulated        | 0.717                     |
| N2-fixation     | Nodulated but without maintenance costs | 0.781 |
| N2-fixation     | Non-nodulated    | 0.812                     |
| Exogenous ammonium | Non-nodulated  | 1                         |

*Non-nodulated indicates that nodule biomass was excluded from the overall biomass reaction and nodule maintenance costs were removed. As N2-fixation in ViNE is unrelated to the production or maintenance of nodule biomass (unless stated otherwise), N2-fixation can occur without nodulation.*
The previous simulations represented simple scenarios where only a single variable differed. In reality, a change in the \( \text{N}_2 \)-fixation efficiency should be accompanied by a change in the rate of nodulation as a result of legume autoregulation of nodulation. We therefore ran simulations where the \( \text{N}_2 \)-fixation efficiency was varied and the rate of nodulation was optimized to maximize plant growth. Strikingly, the simulations suggested a pattern of diminishing returns associated with increasing the \( \text{N}_2 \)-fixation efficiency (Fig. 5c and Supplementary Fig. 3); decreasing the \( \text{N}_2 \)-fixation efficiency 50% from the maximum tested value resulted in a mere 10% decrease in plant growth. The half-maximal growth rate was achieved with a \( \text{N}_2 \)-fixation efficiency of just 10% the maximal, although this required that the nodule accounted for almost 13% of the total biomass. If we assume an upper limit of nodulation at 10% or 5% of the total biomass, the benefits of low \( \text{N}_2 \)-fixation efficiencies are decreased although the pattern of diminishing returns remains (Fig. 5c and Supplementary Fig. 3). In these cases, half-maximal plant growth rate is achieved at 12% or 21%, respectively, of the highest tested \( \text{N}_2 \)-fixation efficiency. Overall, these simulations support that even a poor symbiosis is likely to provide a noticeable benefit to the plant.
DICARBOXYLATES OF N2-FIXING BACTERIA

**Fig. 6 Effects of limiting the mitochondrial terminal oxidase of the N2-fixing zone.** FBA simulations were run with a modified version of ViNE in which bacteroid metabolism could be supported by sucrose in addition to C4-dicarboxylates. No overall limit on oxygen usage of the nodule was set during the simulations, but a limit was set on the activity of the mitochondrial terminal oxidase of the zone III nodule tissue. **a** The effect on plant growth rate of varying the mitochondrial terminal oxidase of the zone III nodule tissue. **b** The effect on specified flux rates of varying the mitochondrial terminal oxidase of the zone III nodule tissue. Red - the sucrose uptake rate of N2-fixing bacteroids; blue - the uptake rate of C4-dicarboxylates of N2-fixing bacteroids; purple - the flux rate of the terminal oxidase of N2-fixing bacteroids. The dashed line indicates the mitochondrial terminal oxidase flux rate below which no sucrose is used by the bacteroids.

Experimentally determined uptake rates by *S. meliloti* bacteroids (1.1 to 1.3 mmol h\(^{-1}\) [g bacteroid dry weight]\(^{-1}\))\(^{35,91}\). Considered together, these simulations provide evidence that C4-dicarboxylates can support optimal plant growth under physiologically relevant conditions.

**Discussion**

Models of the integrated metabolism of various holobionts (consisting of a host and its symbiotic microorganisms) would be valuable tools to understand the emergent properties of these systems\(^ {32,93}\). However, to date there are few examples of constraint-based metabolic modeling being used to study metabolic interactions [e.g.\(^ {8,31}\)], with this approach most commonly used to study the human gut microbiome\(^ {84}\). Here, we developed a broadly adaptable pipeline for modeling the metabolism of interacting organisms across physiologically distinct tissue (sub) sections. Using metabolic network reconstruction and constraint-based modeling, we studied the metabolism of a legume root nodule and SNF, a well-established model of inter-organismal metabolic exchange and cellular differentiation.

Our model (ViNE) accounts for plant shoot, root, and nodule tissues, with the nodule encompassing the metabolism of both the plant and bacterial partners and subdivided into five developmental zones. This is an advance over previous attempts at modeling SNF\(^ {33,35,40}\), most of which focused solely on bacterial metabolism while treating the plant as a black box. The increased complexity of ViNE allows for: i) more accurate simulations of the nutrient exchange, ii) analysis of the metabolic differentiation associated with nodule development, iii) examination of unexpected emergent properties of the symbiosis resulting from inter-organism interactions, and iv) the possibility to perturb the network at the single reaction level. Initial simulations with ViNE supported that this model does a good job at capturing the metabolism of a legume nodule. Nevertheless, as with all models, ViNE predictions were imperfect; of the 38 genes both present in ViNE and listed as being involved in symbiosis according to Additional File 5 of Galardini et al.\(^ {35}\) (excluding genes involved in early nodulation, a process absent in ViNE), deletion of 84% were correctly predicted to have a symbiotic phenotype in ViNE.

However, as we often compared simulated phenotypes for *M. truncatula* with experimental data for *S. sativa*, and given that rhizobium mutant phenotypes are often plant specific (e.g.\(^ {96-98}\)), we cannot rule out that some of the inconsistencies are the result of plant-specific phenotypes. Going forward, we intend to continue to manually refine and update ViNE to maximize consistency with experimental observations.

FBA simulations with ViNE revealed a pattern of diminishing returns in terms of plant growth (as a proxy for fitness) as the N2-fixation efficiency increased, assuming that the rate of nodulation could also vary (Fig. 5c). This observation has potential implications for engineering SNF for biotechnological applications. It suggests that when developing rhizobial inoculants, maximizing competition for nodule occupancy may have a greater impact than maximizing the rate of N2-fixation. This result also supports efforts aimed at engineering N2-fixing symbiosis with cereals\(^ {99}\) by highlighting how even a low-efficiency symbiosis has the potential to have a noticeable benefit on crop yield.

At the same time, the pattern of diminishing returns is interesting from an evolutionary perspective\(^ {100}\). In particular, the evolution of N2-fixation efficiency may be influenced by the rhizobium community diversity, assuming that nodule infection increases the fitness of rhizobia\(^ {101}\). In an environment dominated by a single rhizobium, kin selection may favor the evolution of a poorly efficient symbiosis as it would increase nodule number and thus the size of the niche for colonization by the rhizobia. On the other hand, in a highly diverse environment, evolution of strains capable of entering into a highly efficient symbiosis may be favored, as this would lead to fewer nodules and thus less plant resources being allocated to competing rhizobium strains, thereby limiting the spread of less mutualist (viz. cheater) strains\(^ {20,102}\).

Of particular interest to us were the metabolic exchanges between the plant and rhizobia, both during N2-fixation and during differentiation. The carbon source(s) of differentiating rhizobia remain poorly understood. Results with ViNE suggested that sucrose may be a major carbon source for the differentiating bacteroids. However, *S. meliloti* mutants unable to transport sucrose are not impaired in nodule formation\(^ {103}\), suggesting that differentiating bacteroids have access to at least one other carbon source. Interestingly, a *S. meliloti* pyc mutant unable to grow with glycolytic carbon sources was not impaired in differentiation\(^ {60}\). Similarly, *S. meliloti* pckA\(^ {58}\) and tpI\(^ {104}\) mutants unable to grow with gluconeogenic carbon sources remained capable of differentiating. Thus, it seems likely that differentiating bacteroids have access to a variety of glycolytic and gluconeogenic carbon substrates, with sugars possibly serving as the main carbon source in wild type nodules. If so, the restriction of carbon flow to N2-fixing bacteroids to just C4-dicarboxylates may be the result of active remodeling of the peribacteroid membrane during differentiation.

In attempting to identify conditions favoring the use of C4-dicarboxylates as a carbon source by N2-fixing bacteroids, ViNE also provided insights into the metabolic exchange in the N2-fixation zone. The peribacteroid space of N2-fixing bacteroids is known to be acidic due to the activity of H\(^+\)-ATPases on the peribacteroid membrane\(^ {81-83}\). This acidification contributes to the import of C4-dicarboxylates and the export of ammonium from/to the plant cytosol and the peribacteroid space\(^ {105}\), and it
may contribute to the lysis of non-functional symbiosomes. Our FBA simulations suggest that the plant-derived protons of the peribacteroid space may also be actively used by the bacteroid to support its metabolism.

Although it is generally accepted that nodules are low oxygen environments, the site of O₂-limitation has been debated. Based on the average concentration of free oxygen in the nodule, enzyme kinetics data are consistent with the mitochondria being O₂-limited and the bacteroids being O₂-sufficient. Measurements of the adenylate pools of the plant and bacterial nodule fractions support this conclusion. However, others have argued that nodule adenylate measurements suggest that bacteroids, not the plant, are O₂-limited. Similarly, it was suggested that mitochondria cluster near the periphery of the cell near air pockets, resulting in elevated local O₂ concentrations. The FBA results presented here predicted that C₄-dicarboxylates are the optimal carbon source for N₂-fixing bacteroids only when the plant mitochondria are O₂-limited while the bacteroids are O₂-sufficient. This result supports the hypothesis that mitochondria, and not bacteroids, are O₂-limited in wild type nodules.

The use of metabolic modeling to study SNF is not without limitations. In particular, the accuracy of predictions is restricted by the quality of the imposed flux constraints, and unfortunately, experimental kinetic data for key enzymes and nutrient exchange reactions in the nodule is generally lacking. FBA also fails to actively incorporate regulatory enzymes and nutrient exchange reactions in the nodule is generally lacking. Therefore, the lack of finished legume genomes, an incomplete ability to ensure correct subcellular compartmentalization in eukaryotic cells, and difficulty in experimentally validating the functions of plant genes can limit the quality of models of SNF.

In sum, this work presented a complex metabolic model representing the full metabolism of a rhizobium-nodulated legume, as well as a series of simulations demonstrating the potential for this model to help address genetic, evolutionary, metabolic, and sociobiological questions. Future work will be aimed at continuing to refine and improve the quality of the model, and to using it to generate hypotheses to guide experimental studies and to assist in the interpretation of experimental datasets.

**Methods**

**Construction of a virtual nodule environment.** Before starting the reconstruction of the integrated metabolic model, we refined/updated the existing reconstructions for *S. meliloti* and *M. truncatula*. Concerning *S. meliloti*, we started from the existing core metabolic reconstruction (iGD276) which is a highly curated model built with 726 genes and 681 reactions covering the pathways required to produce all biomass components when provided glucose, succinate, or sucrose as a carbon source, and made the following changes: (i) updated the biomass composition (Supplementary Data 7), (ii) manually added and curated accessory metabolic pathways, (iii) performed an automated expansion based on a nearly modified version of the iGD1575 model reported in, (iv) mass and charge balanced all reactions, and (v) added an ATP hydrolysis reaction to account for non-growth associated maintenance (NGAM) costs. The final *S. meliloti* reconstruction, termed iGD1348, is available in Supplementary Data 1 in SBML, XLS, and MATLAB COBRA format. In terms of the *M. truncatula* reconstruction, a recently published reconstruction was used in this work. Since that version was based on the *M. truncatula* genome annotation version M13.5v51, the gene names in the reconstruction were updated to match the most recent annotation (version 5.0.44 see Supplementary Note 2 for the details of this procedure, and Supplementary Data 8 for the model in SBML format). In addition, we included an energetic cost (0.25 mol of ATP per mol of transported compound) for all single-metabolite diffusion reactions to limit inappropriate transport between compartments. Next, a tissue-specific *M. truncatula* model containing shoot and root tissues was generated using the ‘BuildTissueModel’ function of Plato et al., all genes were relabeled according to the different model compartments (e.g., ‘Root’), and a few ion transport reactions were added. See Supplementary Notes 1 and 2 for additional details on preparation of the *S. meliloti* and *M. truncatula* GENRES, respectively.

Afterward, the *S. meliloti* and *M. truncatula* reconstructions were integrated following the procedure described in detail in Supplementary Notes 4–6, and all reactions for producing dead-end metabolites were removed (Supplementary Data 9). This pipeline generated an integrated, multi-compartment metabolic reconstruction embedding seven overall compartments: shoot tissue, root tissue, the nodal apical meristem (nodule zone I) and four nodule developmental zones representing the combined metabolism of the plant and the bacterium (zone II proximal, interzone II-III, and the nitrogen-fixing zone III). Additionally: (i) the nodule compartments were metabolically connected to the root compartment using appropriate nutrient exchange reactions, (ii) the reaction space of each nodule zone was constrained based on the *M. truncatula* – *S. meliloti* model of specific RNA-seq data in SBML, XLS, and scripts from the COBRA Toolbox commit 9610af110, the TIGER Toolbox version 1.2.0-beta111, FASTCORE version 1.0.112, and the Trn-Core Toolbox version 2.2.4. The iLOG CPLEX Studio 12.7.1 solver (ibm.com) was used for nearly all FBA simulations; the exception was for the preparation of iGD1348, during which the Gurobi version 7.0.1 solver (gurobi.com) was used. The switch to CPLEX was prompted by numerical issues that were solved by switching the solver. All custom scripts used in this study are available through a GitHub repository (github.com/diCenzo-GC/VINE_Reconstruction).

Each gene found in multiple tissues or nodule zones was distinguished by a unique gene name to facilitate tissue-specific gene deletion analyses. When performing global single or double gene deletion analyses, all versions of the gene were simultaneously deleted followed by the removal of all constrained reactions. In contrast, zone- or tissue-specific gene deletion analyses involved deleting just the gene version specific to the zone or tissue of interest. Flux variability analyses were performed with the requirement that flux through the objective function was at least 99% the optimal flux. The reaction data were prepared for each nodule zone first identifying the approximate flux range for each reaction in which the plant growth rate was non-zero. Then, for each reaction, the flux rate of the reaction was set to various values within the previously identified flux range, and the objective value was maximized. For simulations in which the rate of nodulation could vary, nodule biomass was removed from the objective reaction and instead forced through a nodule biomass sink reaction at the appropriate rate; maintenance costs and oxygen availability were modified accordingly (see Supplementary Note 8, for details). For simulations comparing the effect of providing zone III bacteroids sucrose versus C₄-dicarboxylates as the carbon source a modified version of VINE was prepared as described in Supplementary Note 3.

Except when stated otherwise, the objective function for all simulations was: 0.65333 Shoot biomass + 0.32666 Root biomass + 0.00100 Nodule Zone I biomass + 0.00675 Nodule Zone IIId biomass + 0.00225 Bacteroid Zone IIIId biomass + 0.00675 Nodule Zone IIIlp biomass + 0.00075 Nodule Zone IZ biomass + 0.00005 Bacteroid Zone IZ biomass — 1.000 biomass

Except when stated otherwise, the following constraints were placed on VINE during the simulations: (i) when N₂-fixation was the sole source of nitrogen, the uptake of nitrate and ammonia by the roots was set to 0; (ii) the rate of oxygen uptake by zone III tissue was 8.9847 µmol h⁻¹ (g plant dry weight)⁻¹; (iii) the rate of light uptake by the leaves was 1000 µmol h⁻¹ (g plant dry weight)⁻¹; (iv) the rate of CO₂ uptake by the leaves was 1000 µmol h⁻¹ (g plant dry weight)⁻¹; (v) maintenance costs (ATP hydrolysis) for shoot and root tissue were 74 and 35 µmol h⁻¹ (mg plant dry weight)⁻¹, respectively; (vi) maintenance costs (ATP hydrolysis) for zone I, Id, Ilp, IZ, and III plant nodules were set to 90; (vii) maintenance costs (ATP hydrolysis) for zone IId, IIId, Ilp, IIIlp, and IIIId bacteroids were set to 2; (viii) the only carbon sources available to bacteroids in the zone IIIId nodule tissue was succinate, malate, and fumarate at a maximal rate of 1000 µmol h⁻¹ (g plant dry weight)⁻¹ each. No limit on the carbon sources available to the bacteroids of other nodule zones was set.

**Reporting summary.** Further information on research design is available in the Nature Research Reporting Summary linked to this article.

**Data availability**

The authors declare that the data supporting the findings of this study are available within the paper (in Supplementary Data 1–10). No restrictions apply to data availability.
References

1. Theis, K. R. et al. Getting the holonome concept right: an eco-evolutionary framework for hosts and their microorganisms. *mSystems* **1**, e002226 (2016).
2. Vandenkooijhuyse, P., Quaiser, A., Duhamel, M., Le Van, A. & Dufresne, A. The importance of the microbiome of the plant holobiont. *N. Phytol.** **206**, 1196–1206 (2015).
3. McCutcheon, J. P., McDonald, B. R. & Moran, N. A. Convergent evolution of metabolic roles in bacterial co-symbionts of insects. *Proc. Natl Acad. Sci. USA* **106**, 15394–15399 (2009).
4. Smith, S. E., Jakobsen, I., Grenlund, M. & Smith, F. A. Roles of arbuscular mycorrhizas in plant phosphorus nutrition: interactions between pathways of phosphorus uptake in arbuscular mycorrhizal roots have important implications for understanding and manipulating plant phosphorus acquisition. *Plant Physiol.* **156**, 1090–1097 (2011).
5. den Besten, G. et al. The role of short-chain fatty acids in the interplay between diet, gut microbiota, and host energy metabolism. *J. Lipid Res.* **54**, 2325–2340 (2013).
6. Heinken, A. et al. Systematic assessment of secondary bile acid metabolism in gut microbes reveals distinct metabolic capabilities in inflammatory bowel disease. *Microbiome* **7**, 75–78 (2019).
7. Lu, M. & Hedlin, L. O. Global plant-symbiont organization and emergence of biogeochemical cycles resolved by evolution-based trait modelling. *Nat. Ecol. Evol.* **3**, 239–250 (2019).
8. Fondi, M. & Di Patti, F. A synthetic ecosystem for the multi-level modelling of host-microbe interactions. *J. Lipid Res.* **59**, 13–22 (2019).
9. Patopovski, J. et al. Modeling trophic dependencies and exchanges among insects’ bacterial symbionts in a host-simulated environment. *BMGCMicrobiomes* **19**, 402 (2018).
10. Fabbiano, S., Suarez-Zamarano, N. & Trajkovic, M. Host-microbiota mutualism in metabolic diseases. *Front Endocrinol.* **8**, 267 (2017).
11. Klein, C. C. et al. Exploration of the core metabolism of symbiotic bacteria. *BMGC Microbiomes* **13**, 438 (2012).
12. Bertrand, E. M. et al. Phytoplankton-bacterial interactions mediate microeukaryote colonisation at the coastal Antarctic sea ice edge. *Proc. Natl Acad. Sci. USA* **112**, 9938–9943 (2015).
13. Amin, S. A., Parker, M. S. & Armbrust, E. V. Interactions between diatoms and bacteria. *Microbiol. Mol. Biol. Rev.* **76**, 667–684 (2012).
14. Ankrah, N. Y. D., Chouaia, B. & Douglas, A. E. The cost of metabolic interactions in symbioses between insects and bacteria with reduced genomes. *mbio* **9**, 141 (2018).
15. Zibber-Rosenberg, I. & Rosenberg, E. Role of microorganisms in the evolution of animals and plants: the holobiont theory of evolution. *FEMS Microbiol. Rev.* **32**, 723–735 (2008).
16. Moya, A., Pereto, J., Gil, R. & Latorre, A. Learning how to live together: genomic insights into prokaryote-animal symbioses. *Nat. Rev. Genet.* **9**, 218–228 (2008).
17. Mus, F. et al. Symbiotic nitrogen fixation and the challenges to its extension to nonlegumes. *Appl. Environ. Microbiol.* **82**, 3698–3710 (2016).
18. Udvardy, M. & Poole, P. S. Transport and metabolism in legume-rhizobia symbioses. *Annu. Rev. Plant Biol.* **64**, 781–805 (2013).
19. Gibson, K. E., Kobayashi, H. & Walker, G. C. Molecular determinants of a symbiotic chronic infection. *Annu. Rev. Genet.* **42**, 413–441 (2008).
20. Checcucci, A., diCenzo, G. C., Bazzicalupo, M. & Mengoni, A. Trade, diplomacy, and warfare: the quest for elite rhizobia inoculant strains. *Front Microbiol.* **8**, 2207 (2017).
21. Vasse, J., De Billy, F., Camut, S. & Truchet, G. Correlation between ultrastructural differentiation of bacteroids and nitrogen fixation in alfalfa nodules. *J. Bacteriol.* **172**, 4295–4306 (1990).
22. Erisman, J. W., Sutton, M. A., Fleming, R. M. & Thiele, I. Systems-level characterization of a host-microbe symbiosis in the mammalian gut. *Nat. Microbiol.* **4**, 28–40 (2013).
23. Ankrah, N. Y. D., Luan, J. & Douglas, A. E. Cooperative metabolism in a three-partner insect-bacterial symbiosis revealed by metabolite modeling. *J. Bacteriol.* **199**, 17 (2017).
24. Rodenbarg, S. Y. et al. Metabolic model of the *Rhizobium leguminosarum* biovar *lucerne*–tomato interaction reveals metabolic switches during host colonization. *mbio* **10**, 96 (2019).
25. diCenzo, G. C. et al. Metabolic modelling reveals the specialization of secondary replicons for niche adaptation in *Sinorhizobium meliloti*. *Nat. Commun.* **7**, 12219 (2016).
26. diCenzo, G. C. et al. Robustness encoded across essential and accessory replicons of the ecologically versatile bacterium *Sinorhizobium meliloti*. *PLOS Genet.* **14**, e1007357 (2018).
27. Zhao, H., Li, M., Fang, K., Chen, W. & Wang, J. In silico insights into the symbiotic nitrogen fixation in *Sinorhizobium meliloti* via metabolic reconstruction. *PLOS One* **7**, e31287 (2012).
28. Resendis-Antonio, O., Encarnación, S., Collado-Vides, J. & Palsson, B. O. Metabolic reconstruction and modeling of nitrogen fixation in *Rhizobium etli* *eti*. *PLOS Comput Biol.* **3**, 1887–1895 (2007).
29. Resendis-Antonio, O., Hernández, M., Mora, Y. & Encarnación, S. Functional modules, structural topology, and optimal activity in metabolic networks. *PLOS Comput Biol.* **8**, e1002720 (2012).
30. Resendis-Antonio, O. et al. Systems biology of bacterial nitrogen fixation: high-throughput technology and its integrative description with constraint-based modeling. *BMCSyst. Biol.* **5**, 120 (2011).
31. Yang, Y., Yu, X.-P. & Ma, B.-G. Construction and simulation of the *Bradyrhizobium diazoefficiens* USDA110 metabolic network: a comparison between free-living and symbiotic states. *Mol. Biolyst.* **13**, 607–620 (2017).
32. Pfaus, T. et al. The intertwined metabolism during symbiotic nitrogen fixation elucidated by metabolic modelling. *Sci. Rep.* **8**, 12504 (2018).
33. Moreira, T. B. et al. A genome-scale metabolic model of soybean (*Glycine max*) highlights metabolic fluxes in seedlings. *Plant Physiol.* **180**, 1912–1929 (2019).
34. Young, N. D. et al. The *Medicago* genome provides insights into the evolution of rhizobial symbioses. *Nat. Biotechnol.* **30**, 520–524 (2011).
35. Tang, H. et al. An improved genome release (version Mt4.0) for the model legume *Medicago truncatula*. *BMGC Genomics* **15**, 312 (2014).
36. Pecrix, Y. et al. Whole-genome landscape of *Medicago truncatula* symbiotic genes. *Nat. Plants* **4**, 1017–1025 (2018).
37. Gallibert, F. et al. The composite genome of the legume symbiont *Sinorhizobium meliloti*. *Science* **293**, 668–672 (2001).
38. Biondi, E. G. et al. Metabolic capacity of *Sinorhizobium* (*Ensifer*) *meliloti* strains as determined by Phenotype MicroArray analysis. *Appl. Environ. Microbiol.* **75**, 5396–5404 (2009).
39. diCenzo, G. C., MacLean, A. M., Milunovic, B., Golding, G. B. & Finan, T. M. Examination of prokaryotic multipartite genome evolution through experimental genome reduction. *PLOS Genet.* **10**, e1004742 (2014).
40. diCenzo, G. C., Mengoni, A. & Fondi, M. Tn-Core: a toolbox for integrating Tn-seq gene essentiality data and constraint-based metabolic modelling. *ACS Synth. Biol.* **8**, 158–169 (2019).
41. Lieven, C. et al. MEMOTE for standardized genome-scale metabolic model construction. *Nat. Biotechnol.* **38**, 272–276 (2020).
42. Roux, B. et al. An integrated analysis of plant and bacterial gene expression in symbiotic root nodules using laser-capture microdissection coupled to RNA sequencing. *Plant J.* **77**, 817–837 (2014).
43. Lötschler, M., Klumpp, K. & Phytoplast, H. S. N. Growth and maintenance respiration for individual plants in hierarchically structured canopies of *Medicago sativa* and *Helianthus annuus*: the contribution of current and old assimilates. *N. PhytoL* **164**, 305–316 (2004).
44. Zamani, M., diCenzo, G. C., Milunovic, B. & Finan, T. M. A putative 3-hydroxyisobutyrly-CoA hydrolase is required for efficient symbiotic nitrogen fixation in *Sinorhizobium meliloti* and *Sinorhizobium fredii* NGR234. *Environ. Microbiol.* **19**, 218–230 (2017).
45. Minchin, F. R. & Witty, J. F. Respiratory/carbon costs of symbiotic nitrogen fixation in legumes. in *Plant Respiration: From Cell to Ecosystem* (eds Lambers,
54. Durek, R., Koudriakova, C., Kondorosi, A., Krishnapillai, V. & Merrick, M. Complementation analysis of Klebsiella pneumoniae mutants defective in nitrogen fixation. *Mol. Gen. Genet.* 157, 189–198 (1977).

55. Engelke, T., Jording, D., Kapp, D. & Puhler, A. Identification and sequence analysis of the *Rhizobium meliloti* dca gene encoding the C4-dicarboxylate transporter. *J. Bacteriol.* 171, 2879–2887 (1989).

56. de las Nieves Peltzer, M. et al. Auxotrophy accounts for nodulation defect of most *Sinorhizobium meliloti* mutants in the branched-chain amino acid biosynthesis pathway. *Mol. Plant Microbe Interact.* 21, 1232–1241 (2008).

57. Driscoll, B. & Finan, T. M. NADP-dependent dcm transport of the *Sinorhizobium meliloti* dcmporter. *Front Physiol.* 8, 386–392 (1991).

58. Taga, M. E. & Walker, G. C. *Sinorhizobium meliloti* requires a cobalamin-dependent ribonucleotide reductase for symbiosis with its plant host. *Mol. Plant Microbe Interact.* 23, 1643–1654 (2010).

59. Dunn, M. F., Araiza, G. & Finan, T. M. Cloning and characterization of the pyruvate carboxylase from *Sinorhizobium meliloti* Rm1021. *Arch. Microbiol.* 176, 355–363 (2014).

60. De Bruijn, F. J. et al. *Rhizobium meliloti* 1021 has three differentially regulated loci involved in general amino acid biosynthesis, none of which is essential for symbiotic nitrogen fixation. *J. Bacteriol.* 171, 1673–1682 (1989).

61. Gordon, A., Minchin, F., James, C. & Komina, O. Sucrose synthase in legume nodules is essential for nitrogen fixation. *Plant Physiol.* 120, 867–878 (1999).

62. Schulze, J. et al. Inhibition of alfalfa root nodule phosphoenolpyruvate carboxylase through an antisense strategy impacts nitrogen fixation. *Proc. Natl Acad. Sci. USA* 95, 341–346 (1998).

63. Liu, A., Contador, C. A., Fan, K. & Lam, H.-M. Interaction and regulation of the ATPase of pea root nodules and its effect on the nitrogenase activity. *Plant Cell Environ.* 22, 717–728 (1999).

64. Capela, D., Filipe, C., Bobik, C., Batut, J. & Bruand, C. *Sinorhizobium meliloti* differentiation during symbiosis with alfalfa: a transcriptomic dissection. *Mol. Plant Microbe Interact.* 19, 363–372 (2006).

65. Kerpola, T. K. & Kahn, M. L. Symbiotic phenotypes of auxotrophic mutants of *Rhizobium meliloti*. *J. Bacteriol.* 171, 2879–2887 (1989).

66. Liu, A., Contador, C. A., Fan, K. & Lam, H.-M. Interaction and regulation of carbon, nitrogen, and phosphorus metabolisms in root nodules of legumes. *Front Plant Sci.* 9, 1860 (2018).

67. Ronson, C. W., Lyttleton, P. & Robertson, J. G. C4-dicarboxylate transport mutants of *Rhizobium trifolii* form ineffective nodules on *Trifolium repens*. *Proc. Natl Acad. Sci. USA* 78, 4384–4388 (1981).

68. Finan, T. M., Wood, J. M. & Jordan, D. C. Symbiotic properties of *C4* dicarboxylic acid transport mutants of *Rhizobium leguminosarum*. *J. Bacteriol.* 154, 1403–1413 (1983).

69. Mitsch, M. J., diCenzo, G. C., Cowie, A. & Finan, T. M. Saccinate transport is not essential for symbiotic nitrogen fixation by *Sinorhizobium meliloti* or *Rhizobium leguminosarum*. *Appl Environ. Microbiol.* 84, 201561–17 (2018).

70. Li, Y., Parsons, R., Day, D. A. & Bergersen, F. J. Reassessment of major products of N2 fixation by bacteria from soybean root nodules. *Microbiology* 148, 1599–1606 (2002).

71. Kumar, S., Bourdès, A. & Poole, P. De novo alanine synthesis by bacteria of *Medicago truncatula* nod is not required for nitrogen transfer in the determinate nodules of *Lotus corniculatus*. *J. Bacteriol.* 187, 5493–5495 (2005).

72. Waters, J. K. et al. Alanine, not ammonia, is excreted from N2-fixing soybean nodule bacteroids. *Proc. Natl Acad. Sci. USA* 95, 12038–12042 (1998).

73. Allaway, G. L. et al. Identification of alanine dehydrogenase and its role in mixed secretion of alanin and alanine by pea bacteroids. *Mol. Microbiol.* 36, 508–515 (2000).

74. Dunn, M. F. Key roles of microsymbiont amino acid metabolism in rhizobia-legume interactions. *Crit. Rev. Microbiol.* 41, 411–451 (2015).

75. Reid, D. E., Ferguson, B. J., Hayashi, S., Lin, Y.-H. & Gresshoff, P. M. Molecular mechanisms controlling legume autoregulation of nodulation. *Ann. Bot.* 108, 789–795 (2011).

76. Kumar, S., Bourdès, A. & Poole, P. De novo alanine synthesis by bacteria of *Medicago truncatula* nod is not required for nitrogen transfer in the determinate nodules of *Lotus corniculatus*. *J. Bacteriol.* 187, 5493–5495 (2005).

77. Rajaram, G., de Ruiter, J., Kranz, P., Simek, I. & Schütz, D. *Rhizobium meliloti* strain Rm1021 requires no key role of cobalamin-dependent ribonucleotide reductase in symbiosis with its plant host. *Mol. Plant Microbe Interact.* 23, 1643–1654 (2010).

78. Dunn, M. F. Key roles of microsymbiont amino acid metabolism in rhizobia-legume interactions. *Crit. Rev. Microbiol.* 41, 411–451 (2015).

79. Reid, D. E., Ferguson, B. J., Hayashi, S., Lin, Y.-H. & Gresshoff, P. M. Molecular mechanisms controlling legume autoregulation of nodulation. *Ann. Bot.* 108, 789–795 (2011).

80. Szafrai, M. M. & Haaker, H. Properties of the peribacteroid membrane ATPase of pea root nodules and its effect on the nitrogenase activity. *Plant Physiol.* 108, 1227–1232 (1995).

81. Pierre, O., Zuffery, R., Thöny-Meyer, L., Appleby, C. A. & Hennecke, H. A high-affinity C4-type cytochrome oxidase terminates the symbiosis-specific respiratory chain of *Bradyrhizobium japonicum*. *J. Bacteriol.* 178, 1532–1538 (1996).

82. Pierre, O., Zuffery, R., Thöny-Meyer, L., Appleby, C. A. & Hennecke, H. A high-affinity C4-type cytochrome oxidase terminates the symbiosis-specific respiratory chain of *Bradyrhizobium japonicum*. *J. Bacteriol.* 178, 1532–1538 (1996).

83. Fedorova, E. et al. Identi...
Acknowledgements
GCD was supported by a postdoctoral fellowship from the Natural Science and Engineering Research Council of Canada, and funding from Queen’s University. AM was supported by a grant from Fondazione Cassa di Risparmio di Firenze (project name: Metatrack). This work was partially supported by a grant from the Italian Ministry of Agriculture (MICRO4Legumes, act n. n.89267, 19-12-2019).

Author contributions
G.D., A.M., and M.F. conceived the study. G.D., T.P., M.F. prepared the metabolic models. G.D., M.T., M.F. ran simulations with the model. G.D., A.M., and M.F. interpreted the results. G.D., A.M., and M.F. wrote the manuscript. All authors critically revised the manuscript and approved the final version.

Competing interests
The authors declare no competing interests.

Additional information
Supplementary information is available for this paper at https://doi.org/10.1038/s41467-020-16484-2.

Correspondence and requests for materials should be addressed to A.M. or M.F.

Peer review information Nature Communications thanks Jörg Schwender and the other, anonymous, reviewer for their contribution to the peer review of this work. Peer reviewer reports are available

Reprints and permission information is available at http://www.nature.com/reprints

Publisher’s note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made. The images or other third party material in this article are included in the article’s Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the article’s Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this license, visit http://creativecommons.org/licenses/by/4.0/.

© The Author(s) 2020