environmental microbiology

Environmental Microbiology (2017) **19**(1), 218–236 doi:10.1111/1462-2920.13570

A putative 3-hydroxyisobutyryl-CoA hydrolase is required for efficient symbiotic nitrogen fixation in *Sinorhizobium meliloti* and *Sinorhizobium fredii* NGR234

Maryam Zamani, George C. diCenzo, Branislava Milunovic and Turlough M. Finan*
Department of Biology, McMaster University, 1280 Main St. W., Hamilton, Ontario, Canada L8S 4K1.

Summary

We report that the *smb20752* gene of the alfalfa symbiont *Sinorhizobium meliloti* is a novel symbiotic gene required for full N_2-fixation. Deletion of *smb20752* resulted in lower nitrogenase activity and smaller nodules without impacting overall nodule morphology. Orthologs of *smb20752* were present in all alpha and beta rhizobia, including the *ngr_b20860* gene of *Sinorhizobium fredii* NGR234. A *ngr_b20860* mutant formed Fix$^-$ determinate nodules that developed normally to a late stage of the symbiosis on the host plants *Macroptilium atropurpureum* and *Vigna unguiculata*. However an early symbiotic defect was evident during symbiosis with *Leucaena leucocephala*, producing Fix$^-$ indeterminate nodules. The *smb20752* and *ngr_b20860* genes encode putative 3-hydroxyisobutyryl-CoA (HIB-CoA) hydrolases. HIB-CoA hydrolases are required for L-valine catabolism and appear to prevent the accumulation of toxic metabolic intermediates, particularly methacrylyl-CoA. Evidence presented here and elsewhere (Curson *et al.*, 2014, PLoS ONE 9:e97660) demonstrated that Smb20752 and NGR_b20860 can also prevent metabolic toxicity, are required for L-valine metabolism, and play an undefined role in 3-hydroxybutyrate catabolism. We present evidence that the symbiotic defect of the HIB-CoA hydrolase mutants is independent of the inability to catabolize L-valine and suggest it relates to the toxicity resulting from metabolism of other compounds possibly related to 3-hydroxybutyric acid.

Received 10 August, 2016; accepted 6 October, 2016. *For correspondence. E-mail finan@mcmaster.ca; Tel. (+1) 905 525 9140 ext. 22932; Fax (+1) 905 522 6066.

© 2016 Society for Applied Microbiology and John Wiley & Sons Ltd

Introduction

The rhizobia are an agriculturally and ecologically important group of bacteria that survive as free-living cells in the soil and as endosymbiotic partners with legume species. During the endosymbiosis, the bacteria differentiate into bacteroids within the legume cells of a specialized organ known as the nodule (Oldroyd *et al.*, 2011). The differentiated bacteroids convert N_2 gas into ammonia which, unlike the N_2 gas, is available to the plant for use as a nitrogen source. Harvesting the rhizobium-legume symbiosis for agriculture has had, and will continue to provide, enormous benefits, including reducing costs and pollution associated with the use of nitrogen based fertilizers that rely on the Haber-Bosch process (Erisman *et al.*, 2008), and improving crop yields for small-holder farmers whom are unable to afford expensive fertilizers (Mutuma *et al.*, 2014). However, the potential of rhizobia is limited in that the symbiosis is restricted to legume plants (Werner *et al.*, 2014), meaning many of the major crop species such as cereals are unable to benefit from the rhizobia.

As the field of synthetic biology continues to rapidly advance, the long term goal of the rhizobial community to engineer biological nitrogen fixation with non-legume species is increasingly tenable. Several approaches are currently being investigated (Oldroyd and Dixon, 2014; Rogers and Oldroyd, 2014; Geddes *et al.*, 2015). One approach involves engineering non-legumes to be capable of functioning as the plant host during symbiotic nitrogen fixation (SNF) (Rogers and Oldroyd, 2014). For the bacterial partner, these synthetic symbioses could make use of existing rhizobia that have been modified for competitiveness in the engineered plant's soil microbiome. Alternatively, major bacterial isolates from the plant's soil microbiome can be engineered to perform SNF (Geddes *et al.*, 2015). The latter option will involve first elucidating the rhizobial genes necessary and sufficient to confer SNF (diCenzo, Zamani, *et al.*, 2016), and then identifying additional genes that promote an improved symbiotic partnership.

Sinorhizobium meliloti is a narrow host range legume symbiont, entering into symbiosis with plants of the

Medicago, *Melilotus*, and *Trigonella* genera. The genome of the model strain Rm1021 is divided into three primary replicons, a 3.7 Mb chromosome, a 1.7 Mb chromid (pSymB), and a 1.4 Mb megaplasmid (pSymA), and both the pSymA and pSymB replicons are essential for symbiosis (Rosenberg *et al.*, 1981; Finan *et al.*, 1986; Galibert *et al.*, 2001). The genetics of SNF is well-studied in this species (Barnett and Kahn, 2005), and the large number of genetic, genomic, and systems biology resources available for *S. meliloti* facilitates rapid and in-depth analysis of the role of individual genes (Djordjevic, 2004; Schroeder *et al.*, 2005; Cowie *et al.*, 2006; Pobigaylo *et al.*, 2006; Schlüter *et al.*, 2010; Sobrero *et al.*, 2012; Zhao *et al.*, 2012; Galardini *et al.*, 2013; Sugawara *et al.*, 2013; Yurgel *et al.*, 2013; Milunovic *et al.*, 2014; Roux *et al.*, 2014; Galardini *et al.*, 2015; Gemperline *et al.*, 2015; diCenzo, Zamani, *et al.*, 2016; Fei *et al.*, 2016).

Sinorhizobium fredii NGR234 is closely related to *S. meliloti*, containing a similar genome architecture (Schmeisser *et al.*, 2009). *S. fredii* NGR234 is a broad host range symbiont capable of forming effective nodules on legumes from over 120 genera (Pueppke and Broughton, 1999). The nodule morphology of legume species is classified into two major types, determinate and indeterminate (Ferguson *et al.*, 2010), and unlike *S. meliloti* that only forms indeterminate nodules, *S. fredii* NGR234 will form either determinate or indeterminate nodules depending on the plant host. These characteristics make *S. fredii* NGR234 an invaluable model for studying the generalizability of particular phenotypes across a broad range of symbioses.

With the goal of identifying the minimal necessary and sufficient symbiotic gene set, we recently described the construction of a *S. meliloti* mutant library carrying large, defined deletions of the pSymA and pSymB replicons (Milunovic *et al.*, 2014). Screening this library on *Medicago sativa* (alfalfa) allowed us to determine symbiotic phenotypes for the majority of the ~ 2,900 genes encoded by these two replicons (diCenzo, Zamani, *et al.*, 2016). With the exception of the B122 strain, none of the deletion mutant library strains displayed a severe symbiotic phenotype with *M. sativa* that was not expected based on previously described symbiotic genes. Deletion of the ~ 33 kb B122 region on pSymB resulted in a dramatic decrease in symbiotic efficiency. Here, we demonstrate that the locus within the B122 region responsible for the reduced SNF phenotype is a single gene that is conserved throughout the α- and β-rhizobia. Moreover, we show that deletion of the orthologous gene from *S. fredii* NGR234 resulted in a complete loss of symbiotic N_2-fixation (Fix^-) regardless of the plant host or nodule type. We fully characterize the symbiotic phenotypes associated with the deletion of this gene in both *S. meliloti* Rm1021 and *S. fredii* NGR234, and begin to characterize the biological function of the gene product.

Results

The smb20752 gene is required for a fully effective S. meliloti *– legume symbiosis*

In our recent screen of a *S. meliloti* deletion mutant library for symbiotic phenotypes with *M. sativa*, a severe symbiotic defect was noted for the strain RmP798 that has the ΔB122 deletion from nt. 1,529,711 to 1,572,422 of pSymB (Fig. 1A) (diCenzo, Zamani, *et al.*, 2016). *M. sativa* plants inoculated with RmP798 were stunted and light green (Fig. 1B) with shoot dry weights ~ 70% less than those inoculated with wild type (Fig. 1C). Antibiotic resistance profiles and PCR analysis confirmed that nodule isolates were RmP798, and all five of the nodule isolates re-inoculated on *M. sativa* maintained the reduced Fix (Fix^r) phenotype (data not shown). Thus, these data confirmed that deletion of the B122 region had a negative effect on symbiotic N_2 fixation.

To localize the specific locus responsible for the symbiotic phenotype, nine additional deletions were made within the B122 region (Fig. 1A). All seven of the deletions encompassing the 7.5 kb region from nt. 1,536,063 to 1,542,407 exhibited the Fix^r phenotype, whereas the two deletions that did not span this region (ΔB170 and ΔB171) resembled the wild type (Fig. 1A and C). This 6.3 kb region includes five complete coding regions (Fig. 1A), and each of these genes were disrupted by single cross-over plasmid integration. The only mutant to display the Fix^r phenotype had an insertion within *smb20752* (Fig. 1C), and the involvement of *smb20752* in the Fix^r phenotype was further shown through testing an independent, non-polar *smb20752* deletion (Fig. 1C). Expression of *smb20752 in trans* under control of the *dme* promoter from a multicopy plasmid fully complemented the Fix^r phenotype of the Δ*smb20752* and ΔB122 mutants (Fig. 1C), confirming that *smb20752* is the sole gene within the ΔB122 region responsible for the Fix^r phenotype.

Characterization of the Δ*smb20752 symbiotic phenotype with* M. sativa

A series of experiments were performed to gain insight into which stage(s) of the *M. sativa* symbiosis was affected by the Δ*smb20752* mutation. The Δ*smb20752* mutant did not appear to be impaired prior to nodule initiation as no difference was observed in the ability of the mutant to colonize the *M. sativa* root (Fig. 2A) nor did deletion of *smb20752* influence the rate that visible nodules appeared on inoculated *M. sativa* (Fig. 2B). Consistent with the dry weight phenotype, nitrogenase activity as measured via the rate of acetylene reduction was ~ 70% less per plant for *M.*

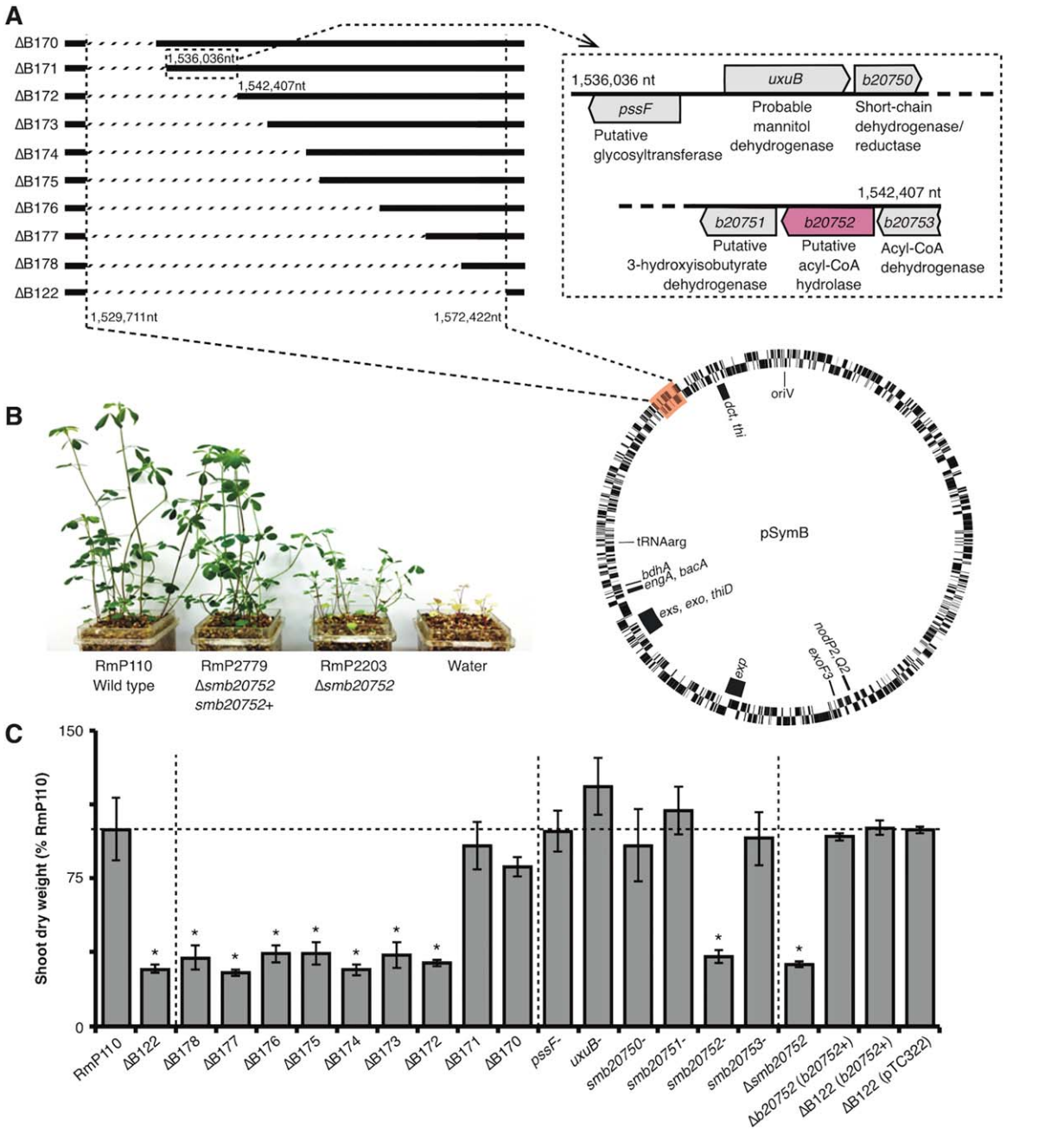

Fig. 1. Identification of the reduced Fix locus.
A. Schematic illustration of the workflow leading to the identification of *smb20752* as the Fix^r locus. The region on pSymB removed in the ΔB122 deletion is highlighted in orange. A series of sub-deletions were made within the B122 region, and these are represented by the lines in the upper left corner; the region deleted is indicated by the dashed lines. The localized Fix^r locus is boxed with dashed lines, and to the right, the genes within this region and the predicted functions of the gene products are indicated.
B. A photograph of *M. sativa* plants 28 days post-inoculation with various *S. meliloti* strains illustrating the Fix^r phenotype. The genotype of the *S. meliloti* strains are indicated below the plants.
C. The shoot dry weights of *M. sativa* inoculated with various *S. meliloti* strains involved in the identification of the Fix^r locus. Dry weights are represented as a percentage of the average weight of the *M. sativa* plants inoculated with wild type *S. meliloti* RmP110 (34.8 mg plant^{-1}). Data points represent the average of triplicate samples, and error bars indicate the standard error. Asterisks highlight samples that were statistically different from RmP110 as determined by an one-way ANOVA followed by a Tukey's-b post-hot test (Supporting Information Table S2). The dashed horizontal line indicates the position of 100% of RmP110, while the vertical dashed lines separate strains into relevant groupings. [Colour figure can be viewed at wileyonlinelibrary.com]

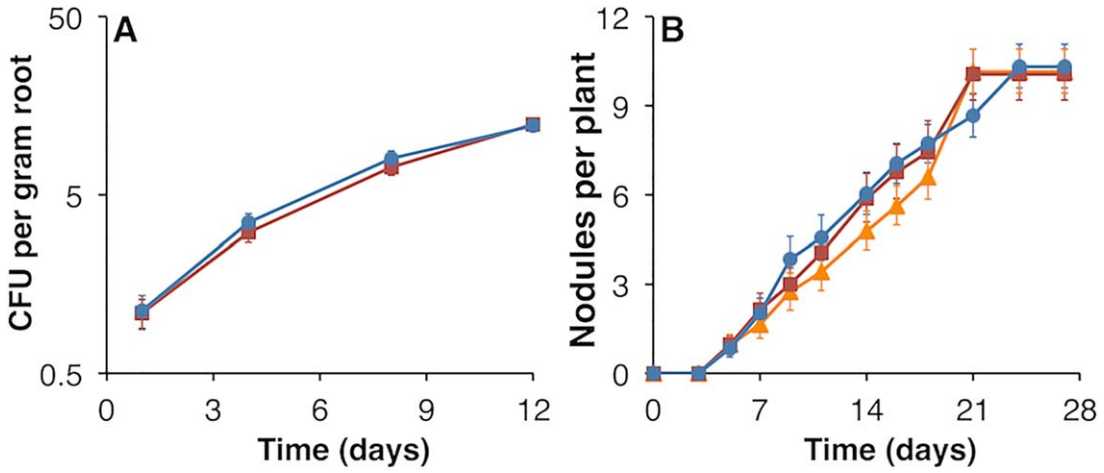

Fig. 2. The effect of deleting *smb20752* on the early symbiotic process.
A. The rate or *M. sativa* root colonization by *S. meliloti* RmP110 (blue) and Δ*smb20752* (red) is shown. Data points represent the mean of 12 plants, and the error bars indicate the standard error.
B. The rate of visible nodule appearance on *M. sativa* inoculated with *S. meliloti* RmP110 (blue), Δ*smb20752* (red), and Δ*smb20752* with *smb20752 in trans* (orange) is shown. Data points represent the mean of 18 plants, and the error bars indicate the standard error. [Colour figure can be viewed at wileyonlinelibrary.com]

sativa inoculated with Δ*smb20752* compared with those inoculated with the wild type RmP110 (Table 1). Additionally, there was a decrease in the total Δ*smb20752* nodule mass per plant due to a decrease in the average size of each nodule (Table 1). Nevertheless, even when standardized by nodule wet weight, the rate of acetylene reduction by nodules infected with Δ*smb20752* was still a third lower than the rate of acetylene reduction by wild type nodules (Table 1). However, the level of *nifH* expression, encoding the Fe-subunit of the nitrogenase enzyme, per mg nodule protein was not statistically different for nodules infected with the wild type RmP110 (16.2 ± 0.7 U per mg protein) compared with the Δ*smb20752* mutant (16.9 ± 0.4 U per mg of protein), as determined with a *nifH*$^+$::*gusA* transcriptional fusion. Overall, these data indicate that the Fix^r phenotype of Δ*smb20752* was largely due to the reduction of nodule tissue per plant, but also due to a decrease in nitrogenase activity, but not expression, in the mature *S. meliloti* bacteroids.

Results from a histological examination of *M. sativa* nodules were consistent with the above conclusion. With the exception of being smaller, confocal microscopy (CM)

Table 1. Phenotypes of temperate legumes inoculated with *S. meliloti* strains.

Host plant	Bacterial strain	SDW (mg plant^{-1})	Nodule wet mass (mg plant^{-1})	ARA (nmol h^{-1} plant^{-1})	ARA (nmol h^{-1} [mg nodule mass] $^{-1}$)
M. sativa	RmP110	72 ± 1 ^{a}	10.8 ±0.4 ^{a}	758 ± 37 ^{a}	71 ± 5 ^{a}
	RmP2203	22 ± 1 ^{b}	5.9 ±0.4 ^{b}	272 ± 12 ^{b}	47 ± 5 ^{b}
	RmP2779	69 ± 1 ^{a}	10.5 ± 0.3 ^{a}	693 ± 18 ^{a}	66 ± 3 ^{a}
	RmG994	9 ± 1 ^{c}	0 ^{c}	0 ^{c}	0 ^{c}
	None	9 ± 1 ^{c}	0 ^{c}	0 ^{c}	0 ^{c}
M. alba	RmP110	169 ± 14 ^{a}	10.4 ± 1.1 ^{a}	1125 ± 99 ^{a}	112 ± 19 ^{a}
	RmP2203	90 ± 1 ^{b}	7.3 ± 0.5 ^{b}	641 ± 14 ^{b}	88 ± 4 ^{a}
	RmP2779	150 ± 20 ^{a}	10.7 ± 0.3 ^{a}	1032 ± 108 ^{a}	97 ± 13 ^{a}
	RmG994	12 ± 1 ^{c}	0 ^{c}	0 ^{c}	0 ^{c}
	None	12 ± 1 ^{c}	0 ^{c}	0 ^{c}	0 ^{c}
M. truncatula	RmP110	91 ± 2 ^{a}	6.6 ± 0.1 ^{a}	946 ± 30 ^{a}	144 ± 4 ^{a}
	RmP2203	17 ± 1 ^{b}	2.8 ± 0.1 ^{b}	311 ± 5 ^{b}	112 ± 2 ^{b}
	RmP2779	93 ± 4 ^{a}	6.4 ± 0.1 ^{a}	890 ± 33 ^{a}	140 ± 4 ^{a}
	RmG994	11 ± 1 ^{b}	0 ^{c}	0 ^{c}	0 ^{c}
	None	11 ± 1 ^{b}	0 ^{c}	0 ^{c}	0 ^{c}

Values represent the means ± standard error of triplicate samples, with each sample consisting of 4–6 plants. Statistically unique groups ($\alpha < 0.05$) were identified with one-way ANOVAs followed by Tukey's-b post-hoc tests, and are represented by the superscript letters. Statistical analyses were performed independently for each plant and for each variable. Strains: RmP110 (wild type); RmP2203 (RmP110 Δ*smb20752*); RmP2779 (PmP110 Δ*smb20752* with *smb20752 in trans*); RmG994 (Rm1021 *tme dme* – a Fix$^-$ control); None (un-inoculated control). ARA, acetylene reduction activity; SDW, shoot dry weight.

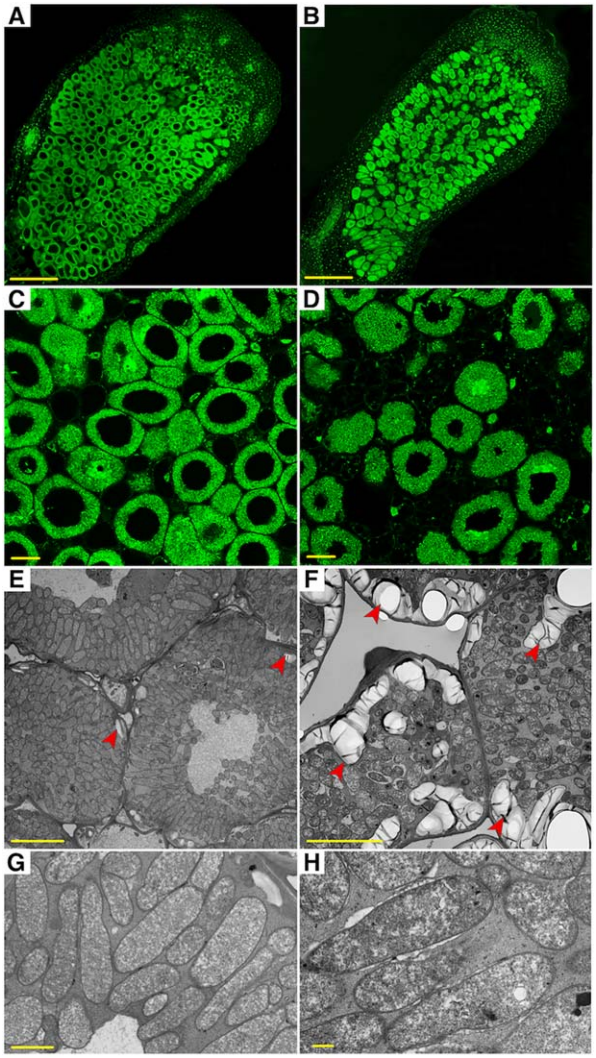

Fig. 3. *M. sativa* nodule structure and *S. meliloti* bacteroid morphology. (A–D) Confocal and (E–H) transmission electron (TEM) micrographs of *M. sativa* nodules induced by either (A,C,E,G) wild type or (B,D,F,H) *smb20752* mutant *S. meliloti* cells.
A,B. Confocal images showing the overall nodule structure.
C,D. Higher magnification confocal images from the nitrogen fixation zone of the nodule.
E,F. TEM images showing the morphology of plant cells from the nitrogen fixation zone of the nodule.
G,H. TEM images illustrating the morphology of the *S. meliloti* bacteroids in the nitrogen fixation zone of the nodule. Scale bars: (A,B) 250 μm (C,D) 25 μm (E,F) 10 μm (G) 2 μm (H) 0.5 μm.
[Colour figure can be viewed at wileyonlinelibrary.com]

revealed that the overall structure of nodules induced by Δ*smb20752* was similar to those induced by the wild type, and nodule plant cells were densely packed with bacteroids of either strain (Fig. 3A–D). Using transmission electron microscopy (TEM), it was observed that the bacteroids of the Δ*smb20752* mutant were elongated and differentiated similar to that of the wild type bacteroids (Fig. 3E–H). The only major difference noted in the TEM micrographs was an abundance of starch granules in the nodule plant cells infected with Δ*smb20752* that was not observed in cells infected with the wild type (Fig. 3E and F). Starch accumulation in amyloplasts in nodule plant cells is a typical feature of inefficient N_2-fixing nodules (Hirsch *et al.*, 1983), consistent with the observed decrease in nitrogenase activity per mg nodule mass (Table 1).

smb20752 is predominately expressed in the infecting, non-differentiated bacteria

To examine the expression pattern of *smb20752* throughout each stage of the symbiosis, *M. sativa* were inoculated with a *smb20752*$^{+}$::*gusA* transcriptional fusion (strain RmP1231). Compared with the expression of *nifH*::*gusA*, the expression of *smb20752* appeared low at all stages of the symbiosis as visualization of the β-glucuronidase activity from *smb20752*::*gusA* required overnight incubation with the chromogenic substrate X-gluc (5-bromo-4-chloro-3-indolyl-β-D-glucuronic acid). Nevertheless, expression of *smb20752* could be detected in the bacterial microcolonies within the curled root hair tips and throughout the infection thread (Fig. 4D and E), as well as within the infection zone of the nodules but not in the differentiated bacteroids of the nitrogen fixing zone (Fig. 4A). In contrast, expression of *nifH* was expectedly only observed in the nitrogen fixing zone (Fig. 4B, C, F, and G). There also appeared to be some overlap in the spatial expression patterns of *nifH* and *smb20752* (Fig. 4A–C), suggesting *smb20752* was still expressed in the early differentiating bacteria. Overall, these observations indicate that *smb20752* is expressed, albeit at a low level, throughout the early infection process and likely in the young differentiating bacteria, but is largely off in the mature bacteroids. This conclusion is supported by a nodule RNA-seq dataset that detected the majority of the *smb20752* expression prior to the nitrogen fixation zone of the nodule (Roux *et al.*, 2014).

The reduced Fix phenotype of Δsmb20752 is observed on multiple plant hosts

To examine whether the Fix^{r} phenotype of Δ*smb20752* was specific to *M. sativa*, the symbiotic properties of the mutant were also examined with *Melilotus alba* (white sweet clover) and *Medicago truncatula* (barrel medic). For both host plants, shoot dry weight analysis showed that plants inoculated with the Δ*smb20752* mutant were statistically smaller than plants inoculated with the wild type RmP110 (Table 1). However, the severity of the Fix^{r} phenotype was dependent on the plant host; the extent of the shoot dry weight phenotype of the Δ*smb20752* mutant was most severe with *M. truncatula* and least severe with *M. alba* (Table 1). Like *M. sativa*, the symbiotic phenotype of Δ*smb20752* with *M. truncatula* appeared to be related to

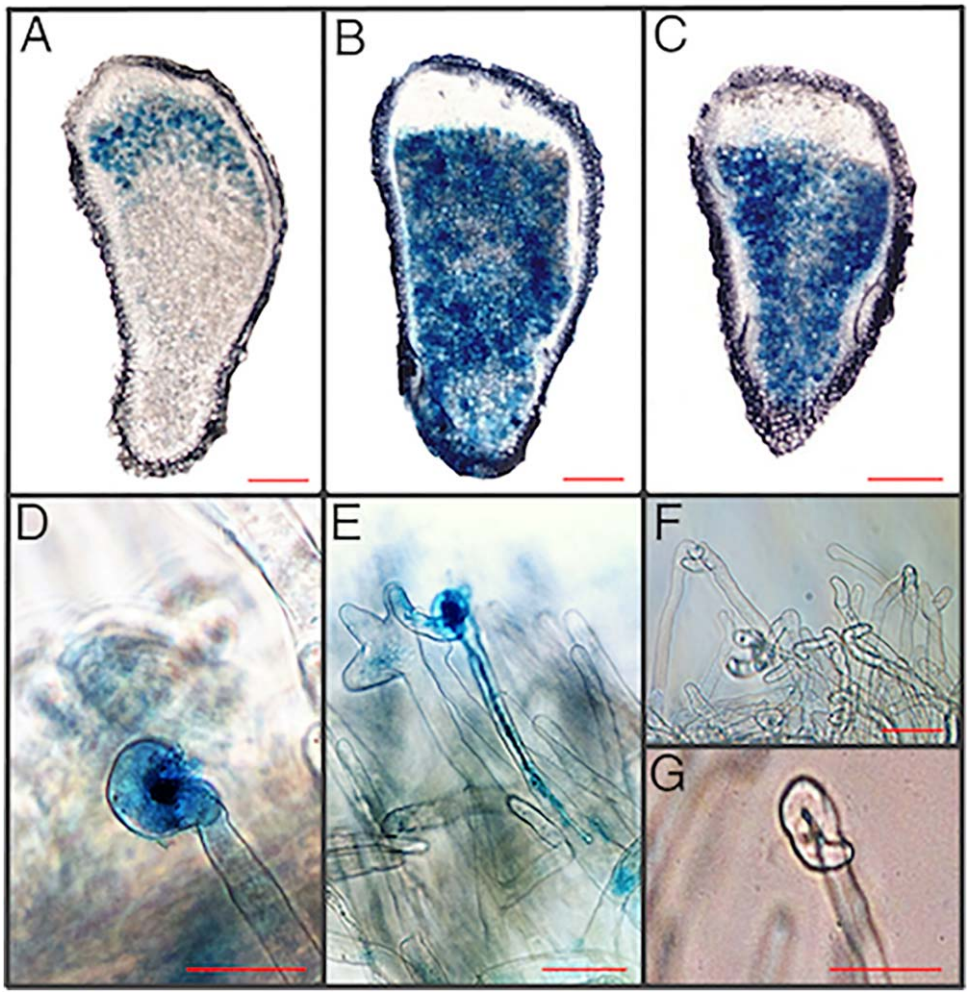

Fig. 4. Expression pattern of *smb20752* and *nifH* throughout symbiosis with *M. sativa*. The expression of the desired gene was monitored using *smb20752*$^{+}$ strains carrying a *gusA* transcriptional fusion to the desired gene and visualized with X-gluc staining. A–C. Images of *M. sativa* nodules infected with either the *S. meliloti* (A) RmP1231 (*smb20752*::*gusA*), (B) RmP319 (*nifH*::*gusA*), or (C) RmP3526 (Δ*smb20752-nifH*::*gusA*) fusion strains. D–G. Images of root hairs on *M. sativa* inoculated with either the *S. meliloti* (D,E) RmP1231 (*smb20752*::*gusA*) or (F,G) RmP319 (*nifH*::*gusA*) fusion strains. Scale bars: (A–C) 300 μm (D–G) 5 μm. [Colour figure can be viewed at wileyonlinelibrary.com]

both a decrease in the rate of nodule mass formation and the amount of nitrogenase activity per mg nodule mass (Table 1). This is in contrast to *M. alba*, where the rate of acetylene reduction per mg nodule mass was not statistically different between plants inoculated with the wild type and Δ*smb20752* (Table 1).

Deletion of the smb20752 ortholog of S. fredii NGR234 results in a loss of symbiotic N_2-fixation

To examine whether the involvement of the *smb20752* gene product in symbiosis is conserved, a non-polar deletion of the orthologous gene (*ngr_b20860*) in *S. fredii* NGR234 was produced. The symbiotic phenotypes of the mutant were examined with three tropical legume hosts, two of which form determinate nodules [*Macroptilium atropurpureum* (purple bush bean) and *Vigna unguiculata* (cowpea)] and one that forms indeterminate nodules [*Leucaena leucocephala* (white leadtree)]. This is in contrast to the host plants used with *S. meliloti*, all of which were temperate legumes with indeterminate nodules.

Transcriptional analysis indicated that *ngr_b20860* was not induced in the indeterminate nodules of *L. leucocephala* relative to the free-living cells (Table 3), consistent with the lack of *smb20752* induction in the nitrogen-fixing zone of alfalfa nodules infected with *S. meliloti* (Fig. 4). However, the use of whole nodule tissue, which is dominated by the nitrogen-fixing zone, for the transcriptional analysis would mask any low, localized induction, and it is therefore possible that *ngr_b20860* was induced in the infection threads similar to *smb20752* (Fig. 4). In contrast, a 30-60 fold induction of *ngr_b20860*, relative to free-living cells grown in TY complex medium, was observed in the determinate nodules of both *M. atropurpureum* and *V. unguiculata* (Table 3). Rather surprisingly, the deletion of *ngr_b20860* in *S. fredii* NGR234 resulted in a much more severe phenotype compared with the effect of deleting *smb20752* in *S. meliloti*. The *S. fredii* NGR234 Δ*ngr_b20860* mutant was unable to fix nitrogen in symbiosis with any of the three host species (Table 2, Fig. 5A and Supporting Information Fig. S1). This symbiotic phenotype was complemented upon expression of *smb20752 in trans* from the plasmid pTH2787 (Table 2).

Table 2. Phenotypes of tropical legumes inoculated with *S. fredii* NGR234 strains.

Host plant	Bacterial strain	SDW (mg plant^{-1})	Nodules per plant	Nodule wet mass (mg plant^{-1})	ARA (nmol h^{-1} plant^{-1})
M. atropurpureum (determinate nodules)	NGR234	240 ± 31 [a]	21 ± 2 [a]	298 ± 10 [a]	2618 ± 185 [a]
	P3821	227 ± 18 [a]	19 ± 3 [a]	322 ± 24 [a]	2063 ± 137 [b]
	P3820	34 ± 2 [b]	19 ± 1 [a]	309 ± 10 [a]	0 [c]
	None	30 ± 3 [b]	0 [b]	0 [b]	0 [c]
V. unguiculata (determinate nodules)	NGR234	743 ± 138 [a]	77 ± 4 [a]	901 ± 30 [a]	5350 ± 190 [a]
	P3821	674 ± 81 [a]	73 ± 3 [a]	845 ± 34 [a]	4700 ± 150 [b]
	P3820	251 ± 27 [b]	79 ± 3 [a]	803 ± 19 [a]	0 [c]
	None	268 ± 13 [b]	0 [b]	0 [b]	0 [c]
L. leucocephala (indeterminate nodules)	NGR234	391 ± 1 [a]	16 ± 2 [b]	62 ± 3 [a]	1903 ± 70 [a]
	P3821	323 ± 14 [b]	14 ± 5 [b]	62 ± 5 [a]	1508 ± 130 [b]
	P3820	34 ± 2 [c]	66 ± 11 [a]	3 ± 0.3 [b]	0 [c]
	None	36 ± 1 [c]	0 [c]	0 [b]	0 [c]

Values represent the means ± standard error of triplicate samples, with each sample consisting of 4–6 plants. Statistically unique groups ($\alpha < 0.05$) were identified with one-way ANOVAs followed by Tukey's-b post-hoc tests, and are represented by the superscript letters. Statistical analyses were performed independently for each plant and for each variable. Strains: NGR234 (wild type); P3820 (NGR234 Δ*ngr_b20860*); P3821 (NGR234 Δ*ngr_b20860* with *smb20752 in trans*); None (un-inoculated control). ARA, acetylene reduction activity; SDW, shoot dry weight.

The three legumes inoculated with the NGR234 Δ*ngr_b20860* mutant showed differences in their symbiotic phenotypes. *L. leucocephala* plants inoculated with the Δ*ngr_b20860* mutant had reduced nodule mass (~ 1/3rd) but four times the number of nodules compared with plants inoculated with the wild type (Table 2). The excessive nodulation is consistent with a lack of the autoregulation of nodulation (Kosslak and Bohlool, 1984; Reid *et al.*, 2011; Mortier *et al.*, 2012), suggesting that this symbiosis failed in a very early stage (van Brussel *et al.*, 2002). This is supported by TEM images that showed *L. leucocephala* nodules inoculated with Δ*ngr_b20860* were essentially devoid of bacterial cells (data not shown), consistent with a failure of the bacterial cells to be released into the plant cell cytoplasm.

In contrast, the total nodule number and nodule mass formed on *V. unguiculata* plants inoculated with wild type *S. fredii* NGR234 and the Δ*ngr_b20860* mutant were statistically the same (Table 2). However, nodules induced by the Δ*ngr_b20860* mutant were white in color (Fig. 5A) and no nitrogenase activity (acetylene reduction) was detected in these nodules (Table 2). Additionally, *nifH* was expressed at a very low level (Table 3), although higher than in free-living cells, consistent with the lack of detectable nitrogenase activity. CM and TEM images indicated a rapid nodule senescence as the overall organization of the nodule was lost (Fig. 5B), many plant cells appeared to be only partially full of bacteroids (Fig. 5C and D), and no peribacteroid membrane could be observed around the bacteroids (Fig. 5E). Additionally, examination of the root systems of *V. unguiculata* plants inoculated with Δ*ngr_b20860* four weeks post-inoculation, instead of six weeks, revealed the presence of a few light pink nodules, further supporting senescence of these nodules. As *S. fredii* NGR234 cells/bacteroids do not undergo major morphological changes during differentiation within *V. unguiculata* nodules (Oono *et al.*, 2010; Li *et al.*, 2013), the similarity between bacterial cell morphology of the mutant and wild type does not help address whether or not the Δ*ngr_b20860* cells fully differentiated prior to nodule senescence.

Similarly to *V. unguiculata*, the nodule number and nodule mass of *M. atropurpureum* plants inoculated with the Δ*ngr_b20860* mutant was not statistically different from plants inoculated with the wild type, but nodules induced by Δ*ngr_b20860* lacked detectable nitrogenase activity (Table 2, Fig. 5A). However, unlike *V. unguiculata* which were harvested at the same time post rhizobial inoculation, the histology of *M. atropurpureum* nodules was largely indistinguishable between plants inoculated with wild type

Table 3. Induction of the *ngr_b20860* and *nifH* genes during *S. fredii* symbiosis.

		Relative transcript expression	
Environment	Strain	*ngr_b20860*	*nifH*
TY liquid culture	NGR234	1.0 ± 0.8	1.0 ± 0.5
	P3820	0	0.9 ± 0.3
M. atropurpureum nodules	NGR234	60.1 ± 3.8	$2.7 \times 10^7 \pm 1.9 \times 10^6$
	P3820	ND	$2.7 \times 10^6 \pm 1.4 \times 10^5$
V. unguiculata nodules	NGR234	38.1 ± 3.8	$4.1 \times 10^7 \pm 4.6 \times 10^6$
	P3820	ND	$3.3 \times 10^4 \pm 3.7 \times 10^3$
L. leucocephala nodules	NGR234	1.7 ± 0.4	$1.2 \times 10^6 \pm 1.9 \times 10^5$
	P3820	ND	ND

The data are presented relative to the expression of the gene of interest in wild type *S. fredii* NGR234 grown in TY liquid culture. Values represent the means ± standard error of biological triplicates, with each biological replicate the average of technical triplicates. ND, not determined. Strains, NGR234 (wild type); P3820 (NGR234 Δ*ngr_b20860*).

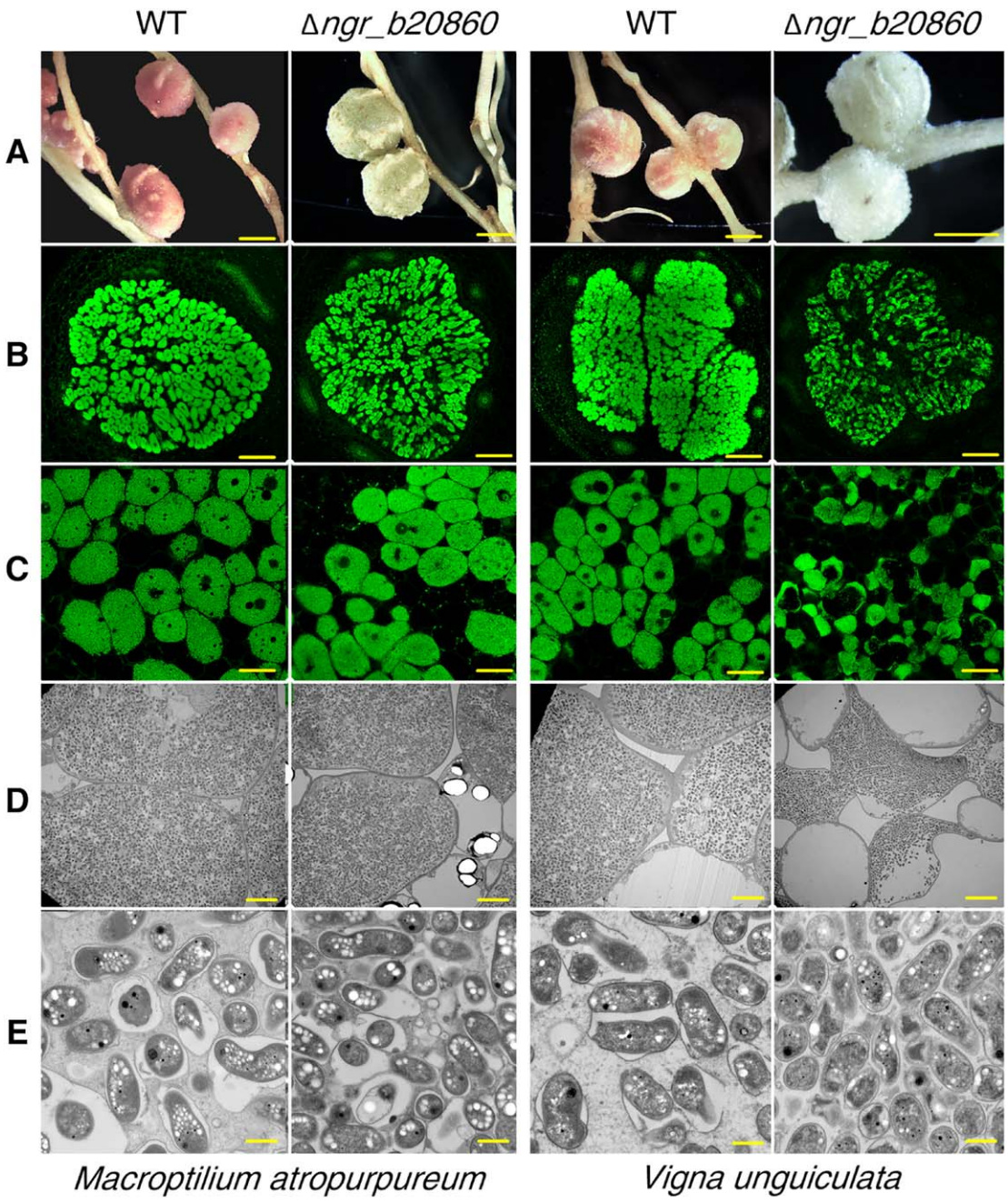

Fig. 5. *Macroptilium atropurpureum* and *V. unguiculata* nodule structure and *S. fredii* NGR234 bacteroid morphology. Images on the left are of *M. atropurpureum* nodules, while those on the right are of *V. unuiculata*. Within each plant, wild type nodules are shown on the left and Δ*ngr_b20860* are shown on the right.
A. Photographs of nodules induced by *S. fredii* NGR234 strains.
B. Confocal images showing the overall nodule structure.
C,D. Higher magnification (c) confocal images and (D) TEM images displaying the plant cell morphology in the nodules.
E. TEM images illustrating the morphology of the *S. fredii* NGR234 bacteroids. Scale bars: (A) 2.5 mm (B) 250 μm (C) 25 μm (D) 10 μm (E) 0.5 μm.
[Colour figure can be viewed at wileyonlinelibrary.com]

S. fredii NGR234 and the Δ*ngr_b20860* mutant (Fig. 5B–E). The only observable difference was the presence of some starch granules in empty plant cells of nodules induced by the mutant (Fig. 5D). While the TEM images provide little information on the differentiation state of the Δ*ngr_b20860* bacteroids (Oono *et al.*, 2010), RT-qPCR

indicated that *nifH* was partially induced in the Δ*ngr_b20860* bacteroids (Table 3). Thus, the complete lack of nitrogenase activity is not only due to lower nitrogenase expression, but also a result of the nitrogenase being unable to function, assuming nitrogenase is made.

Smb20752 is involved in L-valine and β-hydroxybutyrate catabolism

To begin identifying a possible biological role for the *smb20752* and *ngr_b20860* gene products, we examined whether deletion of these genes result in any carbon catabolic phenotypes. Previous MicroArray™ analysis has implicated the B122 region in the catabolism of branched chain amino acids (leucine, isoleucine, and valine) (diCenzo, Checcucci, *et al.*, 2016). Indeed, the gene products of *smb20752* and the upstream (*smb20753*) and downstream (*smb20751*) genes show homology to three of the enzymes involved in valine catabolism, and a *ngr_b20860* mutant is unable to grow with valine (Curson *et al.*, 2014). Here we observed that both the *S. meliloti* ΔB122 and Δ*smb20752* mutants showed poor growth on L-valine as the sole carbon source relative to the wild type, while both mutants grew well with succinate as the carbon source (Fig. 6). Similarly, a *ngr_b20860* mutant failed to grow with L-valine as the sole carbon source (Curson *et al.*, 2014). Both the ΔB122 and Δ*smb20752* mutants also showed poor growth with L-isoleucine (data not shown), while ΔB122 but not Δ*smb20752* had a growth phenotype with L-leucine (data not shown).

Previous work has also shown that four loci within the B122 region are involved in β-hydroxybutyrate (BHB) catabolism (Charles *et al.*, 1997), with only one of the loci, *bhbA*, having been identified (Charles and Aneja, 1999). The inability of Δ*smb20752* to grow in liquid medium with 15 mM BHB as the sole carbon source (Fig. 6) identified the *smb20752* gene product as one of the three unidentified loci within the B122 region involved in BHB metabolism. Peculiarly, when the concentration of BHB was lowered to 5 mM from 15 mM, the mutant was capable of growing, albeit somewhat slower than the wild type (Fig. 6). Similarly, the *S. fredii* Δ*ngr_b20860* mutant displays slow growth when grown with BHB as a sole carbon source (Supporting Information Fig. S2). Thus, *smb20752* and *ngr_b20860* are not essential for BHB catabolism but are perhaps involved in a detoxifying capacity.

The smb20752 gene is conserved in the rhizobia, but the genomic context is not

The phylogenetic history and conservation of *smb20752* was examined in a phylogenetically diverse and representative subset of the α- and β-rhizobial species (see

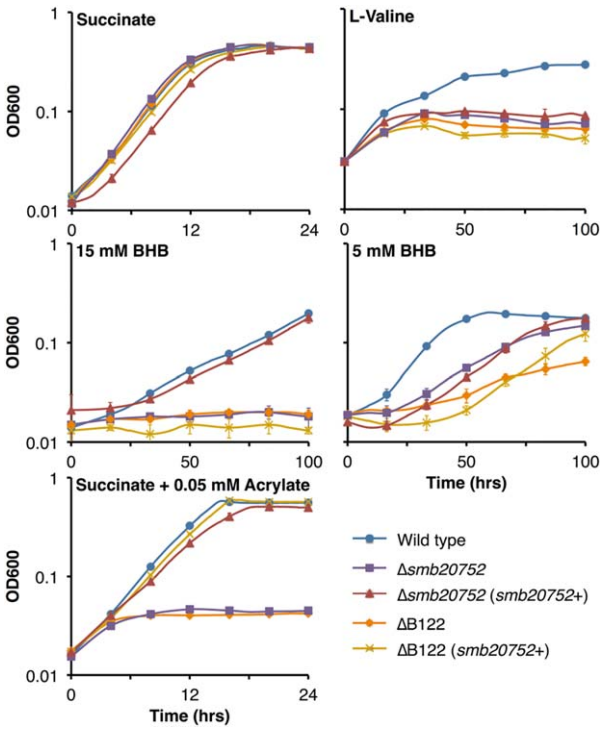

Fig. 6. The involvement of *smb20752* in L-valine and β-hydroxybutyrate catabolism.
The growth profiles of various *S. meliloti* strains are shown during growth in minimal medium containing various carbon sources. Data points represent the average of triplicate samples, with the error bars indicating the standard error. [Colour figure can be viewed at wileyonlinelibrary.com]

Experimental procedures). Proteins orthologous to Smb20752 were first identified with a Blast bidirectional best hit (Blast-BBH) approach. This protein appears to be highly conserved as orthologs of this protein were detected in the proteomes of all analyzed rhizobial genomes (Fig. 7 and Supporting Information Fig. S4). The topology of a RAxML maximum likelihood phylogeny based on the amino acid sequence of the identified Smb20752 orthologs is similar but not perfectly congruent to a multilocus sequence analysis (MSLA) of the same strains (Supporting Information Fig. S5). However, the phylogeny of Smc01153 orthologs, catalyzing the reaction in the L-valine catabolic pathway preceding the Smb20752 catalyzed reaction (Curson *et al.*, 2014), is also incongruent with the MLSA and the Smb20752 phylogeny (Supporting Information Fig. S5). Thus, the small difference between the Smb20752 and MLSA trees is likely not due to horizontal gene transfer, and instead *smb20752* appears to have been vertically transmitted to the modern day rhizobia from their common ancestor.

Despite *smb20752* orthologs being conserved across all examined rhizobia, the genomic context of the gene varied (Fig. 7). The *smb20752* orthologs were only co-localized with *smb20751* and *smb20753* homologs, which are

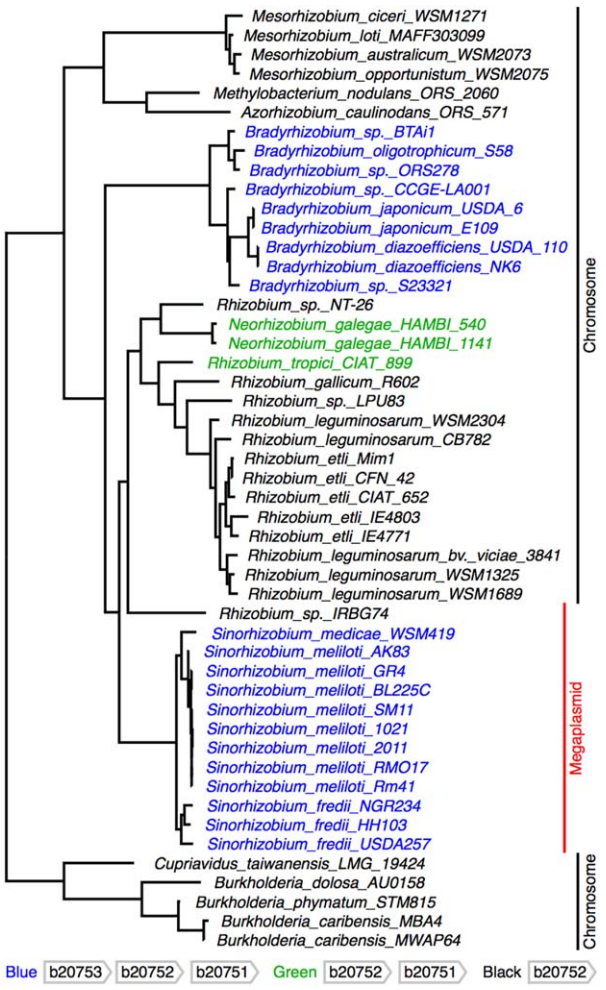

Fig. 7. Phylogenetic and genomic context analysis of *smb20752* orthologs.
A maximum likelihood bootstrap best tree based on the amino acid sequence of the identified Smb20752 orthologs, which was produced as described in the Experimental procedures section, is shown. The location of the Smb20752 orthologs are indicated along the right, and none of the orthologs on a chromosome are within a symbiotic island. Taxa are colour coded based on the genomic context of the *smb20752* as indicated at the bottom of the figure. A larger version of this figure with bootstrap values is provided as Figure S4. [Colour figure can be viewed at wileyonlinelibrary.com]

potentially also involved in branched chain amino acid metabolism (data not shown), in the *Sinorhizobium* and *Bradyrhizobium* species (Fig. 7). However, in nearly all other rhizobial strains the *smb20752* orthologs were not co-localized with *smb20751* or *smb20753* homologs, except in a few cases where just a *smb20751* homolog was present (Fig. 7). The observed variation in genomic context is consistent with the gene product of the *smb20752* orthologs being involved in biological functions that are independent from that of branched chain amino acid metabolism.

Discussion

The work described in this report arose through the analysis of symbiotic N_2-fixation (SNF) phenotypes of *S. meliloti* strains in which large regions of the pSymB replicon were deleted (diCenzo, Zamani, *et al.*, 2016). Relative to the wild type, nodules formed by strains lacking the 33 kb B122 region had a reduced symbiotic N_2-fixation (Fix) phenotype and this was surprising as no genes involved in symbiosis were known to lie in B122 region. Deletion of the *smb20752* gene alone conferred the reduced SNF phenotype and wild type N_2-fixation was restored by introduction of *smb20752* alone into the ΔB122 mutant (Fig. 1). While the *S. meliloti smb20752* mutation resulted in a reduced SNF phenotype on alfalfa, sweet clover, and barrel medic (Table 1), the symbiotic analysis of *S. fredii* NGR234 mutants lacking the *smb20752* ortholog, *ngr_b20860* (*vutE*) was more striking as these were completely defective in SNF (Table 2). Below we discuss the function of the *smb20752* gene and its ortholog *ngr_b20860* (*vutE*) and how this relates to the symbiotic phenotypes of the mutations in these genes.

The analysis of the *smb20752* and *ngr_b20860* (*vutE*) genes and their free-living phenotypes as reported here and elsewhere (Curson *et al.*, 2014) strongly suggest that these orthologs encode the enzyme 3-hydroxyisobutyryl-CoA hydrolase (EC 3.1.2.4) that hydrolyzes 3-hydroxyisobutyryl-CoA (HIB-CoA) to give the free acid and CoA-SH. As shown in Fig. 8, this reaction occurs during the catabolism of the branched-chain amino acid L-valine to the tricarboxylic acid cycle intermediate succinyl-CoA (Massey *et al.*, 1976). Figure 8 includes the enzyme-gene designations recently assigned for *S. fredii* NGR234 by Curson *et al.* (2014). Most of the genes encoding for enzymes of this pathway are located together on the pSymB replicons of *S. meliloti* and *S. fredii*, and these include *smb20753* (*bauC*), *smb20752* (*vutE*), *smb20751* (*vutF*), *pccB*, *pccA*, and *bhbA* (Fig. 8). The role of the *S. meliloti smb20752* gene in L-valine catabolism is consistent with the poor growth of the *smb20752* mutant relative to the wild type with L-valine as the sole carbon source (Fig. 6), and Curson *et al.* (2014) recently showed that *ngr_b20860* (*vutE*) mutants of *S. fredii* NGR234 similarly do not grow with L-valine as the sole carbon source.

HIB-CoA hydrolase is a particularly important enzyme in mammalian, plant, and microbial cells metabolizing L-valine as HIB-CoA hydrolase activity appears to be required to protect cells against the toxic effects of the catabolic intermediate methacrylyl-CoA that reacts readily with free thiol groups of proteins (see Shimomura *et al.*, 1994; Zolman *et al.*, 2001; Curson *et al.*, 2014)]. High levels of methacrylyl-CoA hydratase (crotonase) and HIB-CoA hydrolase relative to other enzymes in the L-valine catabolic pathway are thought to rapidly remove

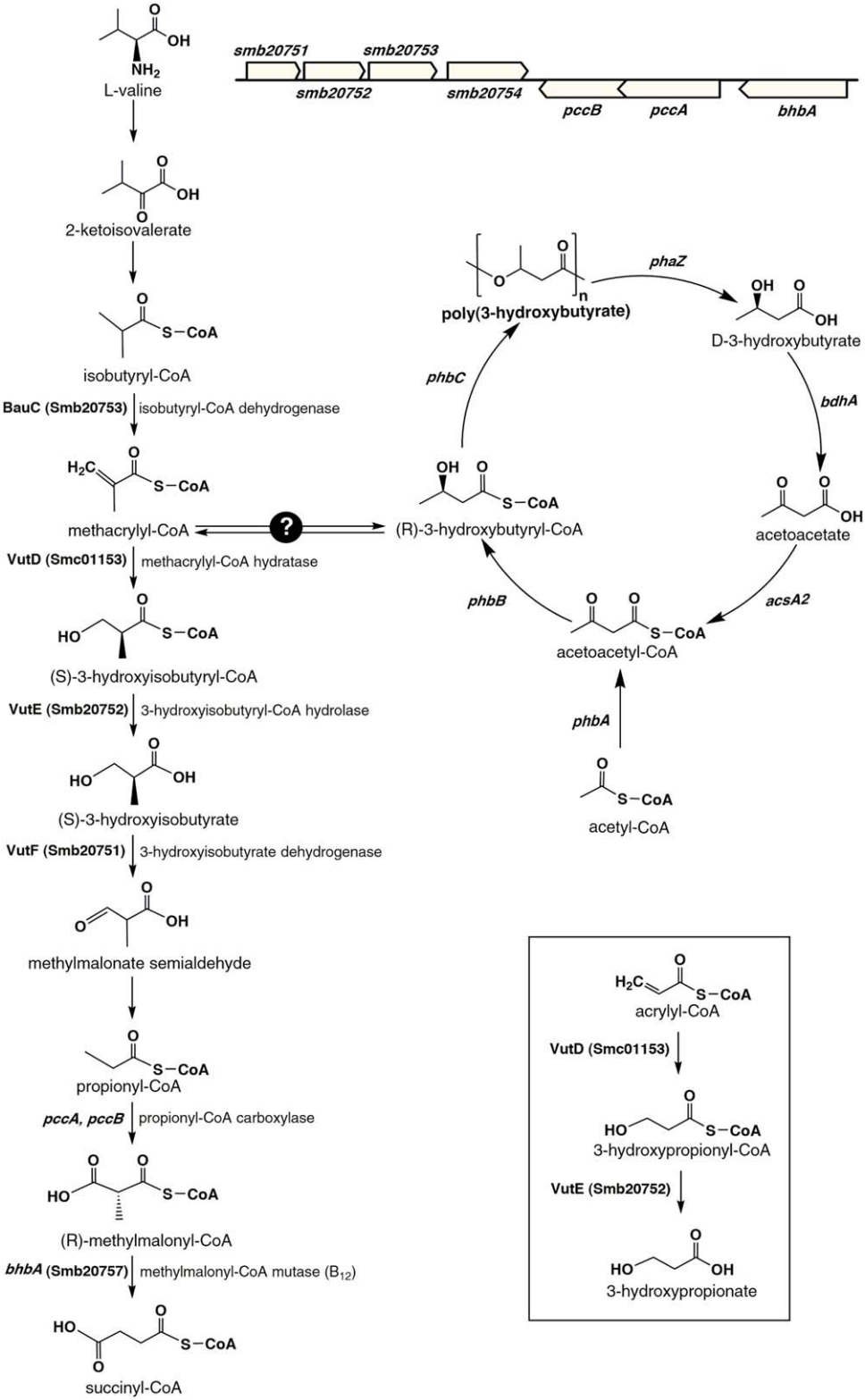

Fig. 8. Schematic of the metabolic pathways associated with Smb20752.
The L-*valine* (left) and acrylic acid (bottom right) catabolic pathways and the PHB metabolism cycle (right) are shown. A possible metabolic link, (indicated by the question mark) between these pathways is shown. An outline of the *smb20751-bhbA* gene region from the pSymB chromid of *S. meliloti* is shown in the upper right panel. Details of the pathways were taken from the literature (Massey *et al*., 1976; Trainer and Charles, 2006; Curson *et al*., 2014). [Colour figure can be viewed at wileyonlinelibrary.com]

methacrylyl-CoA and prevent its accumulation (Shimomura *et al*., 1994). Thus during L-valine metabolism in *S. meliloti* and *S. fredii*, HIB-CoA hydrolase mutants could accumulate HIB-CoA and methacrylyl-CoA and this would result in cellular toxicity. Curson *et al*. (2014) explored the protective role of HIB-CoA hydrolase in free-living bacterial

cells in studies involving acrylic acid. Acrylic acid is metabolized via the highly toxic molecules acrylyl-CoA and 3-hydroxypropionyl-CoA that are structurally similar to methacrylyl-CoA and HIB-CoA respectively (Fig. 8). They found that expression of bacterial *vutD* and *vutE* genes encoding methacrylyl-CoA hydratase and HIB-CoA hydrolase was sufficient to confer resistance to acrylate likely through the removal of the toxic product acrylyl-CoA (Curson *et al.*, 2014). Moreover, they reported that growth of *S. fredii* NGR234 *vutD* and *vutE* mutants displayed increased sensitivity to the presence of acrylate relative to the wild type strain (Curson *et al.*, 2014).

Similarly, we found that the *S. meliloti smb20752* mutant and the ΔB122 deletion mutant were severely sensitive to acrylate relative to the wild type (Fig. 6 and Supporting Information Fig. S3). Introduction of *smb20752 in trans* fully complemented growth of both the *smb20752* and the ΔB122 mutants at low acrylate concentrations (Fig. 6), and we also note that introduction of *smb20752* alone was sufficient to restore a full Fix^+ phenotype to the ΔB122 strain (Fig. 1). However, introduction of *smb20752 in trans* only partially complemented the *smb20752* and ΔB122 mutants at higher concentrations of acrylate, and similarly, it did not complement the L-valine growth phenotype of the *smb20752* mutant (Fig. 6 and Supporting Information Fig. S3). It therefore appears that the introduced *smb20752* gene was not expressed at the same level as the wild type gene; this maybe due to insufficient transcription as *smb20752* was not expressed from its native promoter. We also note that introduction of *smb20752* into the ΔB122 mutant expectedly did not complement the L-valine growth as other genes deleted as part of the ΔB122 deletion (*smb20753*, *smb20751*, *pccB*, *pccA*, and *bhbA*) are involved in L-valine catabolism (Fig. 8). As the same construct could fully complement the Fix phenotype of the ΔB122 mutant despite the inability to restore L-valine catabolism, the symbiotic phenotype must be independent of the ability to catabolize L-valine. And indeed, *S. meliloti* appears to receive little if any L-valine from the plant (las Nieves Peltzer *et al.*, 2008; Prell *et al.*, 2010).

In addition to the Fix and L-valine utilization phenotypes of *S. meliloti smb20752* and *S. fredii ngr_b20860* mutants, these genes are also required for optimal growth with 15 mM 3-hydroxybutyrate (BHB) as the sole carbon source (Fig. 6 and Supporting Information Fig. S2). These data are consistent with the previous work of Charles *et al.* (1997) where a genetic analysis revealed four linked loci (complementation groups) in *S. meliloti* that are required for growth with BHB. The only one of these four loci previously identified was *bhbA* (Charles and Aneja, 1999; Miyamoto *et al.*, 2003), which encodes for methylmalonyl-CoA mutase involved in the last step in L-valine catabolism (Fig. 8). The location of *smb20752* relative to *bhbA* indicates that this gene corresponds to the *bhbC* locus of Charles *et al.* (1997) and the lack of growth of the *smb20752* mutant with 15 mM 3-hydroxybutyrate (Fig. 6) is consistent with these previous data. However, we observed that partial growth of these strains occurred when the BHB concentration was reduced to 5 mM and this suggests that the BHB growth phenotypes of the *S. meliloti smb20752* and *S. fredii ngr_b20860* mutants is due to toxicity associated with BHB metabolism and not due to an inability to catabolize BHB.

BHB is a breakdown intermediate of the carbon storage compound poly-3-hydroxybutyrate (PHB) that plays a role in various stages of the rhizobial-legume symbiosis (Trainer and Charles, 2006). The ability to synthesize PHB is required for the catabolism of BHB as a carbon source by *S. meliloti* (Fig. 8; Trainer and Charles, 2006; Trainer *et al.*, 2010). The BHB and L-valine utilization phenotypes of the *smb20752* and *ngr_b20860* mutants suggest a link in the metabolism of these compounds. We are not aware of such a link but as noted previously, it is also unclear why *bhbA* mutants of *S. meliloti* do not grow on BHB (Charles and Aneja, 1999; Miyamoto *et al.*, 2003). We do, however, note that a link is present in the related α-proteobacterium *Methylobacterium extorquens* AM1, where the PHB synthesis intermediate 3-hydroxybutyryl-CoA can be converted to the L-valine toxic catabolic intermediate methacrylyl-CoA and then HIB-CoA (Korotkova *et al.*, 2002).

Whereas the Fix phenotype of the *smb20752* mutant cannot be linked to the inability to catabolize L-valine, it is tempting to associate the Fix phenotype with the PHB/BHB phenotype. *S. meliloti* cells contain visible PHB granules during growth in the infection thread and upon initial release into the plant cell, but these PHB granules are mobilized and disappear during the differentiation process (Paau *et al.*, 1978; Hirsch *et al.*, 1983; Vasse *et al.*, 1990; Wang *et al.*, 2007). This is consistent with the expression pattern of *smb20752*, which appears to be expressed in the infecting but not the differentiated bacteria (Fig. 4). A *S. meliloti* PHB synthesis mutant has a more severe symbiotic phenotype with *M. truncatula* than with *M. sativa* (Wang *et al.*, 2007), and similarly, a Δ*smb20752* mutant had a more severe symbiotic phenotype with *M. truncatula* than with *M. sativa* (Table 1). Unlike *S. meliloti*, *S. fredii* NGR234 accumulates PHB throughout the symbiosis with all three of the tested plants (Fig. 5 and Li *et al.*, 2013), and deletion of *ngr_b20860* in *S. fredii* NGR234 resulted in a Fix^- phenotype unlike the Fix^r phenotype of a *S. meliloti smb20752* mutant. Additionally, RNA-seq data is consistent with active turnover of PHB in *S. fredii* NGR234 bacteroids of *V. unguiculata* and *L. leucocephala* nodules (Li *et al.*, 2013). Expression of the *ngr_b20860* gene was induced in nodules of *V. unguiculata* and *M. atropurpureum*, although it was not induced in *L. leucocephala* nodules relative to free-living cells. Finally, deletion of

ngr_b20860 in *S. fredii* NGR234 resulted in a more severe symbiotic phenotype with *L. leucocephala* than with *M. atropurpureum* or *V. unguiculata* (Table 2). Similarly, disrupting *bdhA* involved in PHB degradation resulted in a symbiotic phenotype with *L. leucocephala* but not the other tested plants (Aneja and Charles, 2005), and the size of PHB granules are larger in *S. fredii* NGR234 bacteroids of *L. leucocephala* nodules compared with *V. unguiculata* nodules (Li *et al.*, 2013).

The above observations present a correlation between PHB utilization and the severity of the *smb20752* or *ngr_b20860* Fix phenotypes, suggestive of the Fix phenotypes of these mutants being associated with the build up of a toxic compound resulting from the metabolism of PHB. However, disrupting the *phaZ* gene that encodes the PHB depolymerase (Fig. 8) had no effect on the *smb20752* Fix phenotype (Supporting Information Table S3), and as *phaZ* mutants can synthesize but not catabolize PHB (Trainer *et al.*, 2010), any PHB-related toxicity in the *smb20752* mutant must arise from the PHB synthesis half of the PHB cycle (Fig. 8).

It is also interesting to note the difference between determinate and indeterminate nodule types. Expression of *ngr_b20860* was induced in both determinate nodules tested, and the symbiotic phenotypes of the Δ*ngr_b20860* mutant suggests a role for this gene late in the symbiosis. In contrast, the expression of *ngr_b20860* or *smb20752* was not induced in the nitrogen-fixing zones of the tested indeterminant nodules, and the severe symbiotic phenotype (absence of bacteroids) of the *ngr_b20860* mutant in *L. leucocephala* nodules and the reduced nodule size of *M*. sativa nodules induced by the *smb20752* mutant suggests a requirement for these genes early during the infection process in these symbioses.

Overall, it therefore appears likely that the build up of a toxic compound in the *S. meliloti smb20752*/*ngr_b20860* mutants is the cause of the symbiotic phenotypes, however the identify and precise nature of the affected metabolic pathway requires further study. In doing so, it is expected that novel insights into the rhizobium – legume symbiosis will be obtained that will help progress ongoing attempts at engineering synthetic symbioses.

Experimental procedures

Bacterial growth conditions, media, and genetic manipulations

All media (LB, LBmc, TY, M9), antibiotic concentrations, and growth conditions for *S. meliloti* and *E. coli* were as previously described (diCenzo *et al.*, 2014). Unless stated otherwise, the concentration of carbon sources in M9 media were: 15 mM glucose, 15 mM succinate, 15 mM DL-β-3-hydroxybutyrate, 10 mM L-valine, 10 mM L-leucine, 10 mM L-isoleucine. *S. fredii* NGR234 and derivatives were grown at 30°C in TY medium or *Rhizobium* minimal medium (RMM) (Trainer, 2009). Antibiotic concentrations for selection of *S. fredii* NGR234 strains were: 50 μg/ml gentamicin, 25 μg/ml kanamycin, 50 μg/ml rifampicin, 1000 μg/ml streptomycin, and 10 μg/ml tetracycline (Stanley *et al.*, 1988). Growth curves were performed and analyzed as described elsewhere (diCenzo *et al.*, 2014). DNA manipulations and recombinant techniques, bacterial matings, and ΦM12 transductions were performed as described before (Finan *et al.*, 1984; Sambrook *et al.*, 1989; Cowie *et al.*, 2006; Milunovic *et al.*, 2014; diCenzo, Zamani, *et al.*, 2016).

Bacterial strains and plasmids

Bacterial strains and plasmids are described below and in Supporting Information Table S1. *S. meliloti* RmP798 (ΔB122; pSymB nt. 1,529,711 to 1,572,422) was described previously (Milunovic *et al.*, 2014). The nine deletions within the B122 region were constructed via the Flp/FRT recombination system (Milunovic *et al.*, 2014). For this, plasmid pTH1942 was constructed by ligating an ∼ 1,500 nt PCR product (pSymB nt. 1,528,150 to 1,529,711) into pTH1937 (Milunovic *et al.*, 2014), which contains a FRT site, via SpeI and EcoRI. Plasmid pTH1942 was then conjugated into nine *S. meliloti* fusion library strains each containing a single FRT site (Cowie *et al.*, 2006). The intervening region between the two FRT sites in the resulting strains was deleted with Flp via pTH1944 (Milunovic *et al.*, 2014), producing deletions ΔB170 through ΔB178.

Sinorhizobium meliloti knockout mutants of RmP110 (Yuan *et al.*, 2006) were constructed through single cross-over plasmid integration within a central portion of the appropriate open reading frame. The appropriate PCR product was ligated into NsiI/BglII digested plasmid pTH1703 (Cowie *et al.*, 2006), and the resulting vectors introduced into the genome of wild type *S. meliloti* RmP110 via conjugation. *S. meliloti* transcriptional fusion reporter strains were constructed in the same way as the knockout mutants, except that integration of the plasmid into the genome was designed such that expression of the complete open reading frame was not disrupted. *S. meliloti* RmP3536 (Δ*smb20752 nifH*::*gusA*) was constructed by recombining the *nifH*::*gusA* fusion from RmP319 (Cowie *et al.*, 2006) into *S. meliloti* RmP2203 (Δ*smb20752*) via transduction.

The non-polar Δ*smb20752* deletion mutant was made through double cross-over recombination using *sacB* negative selection (Quandt and Hynes, 1993). Regions upstream and downstream of *smb20752* were sequentially ligated into pJQ200mp18 (Quandt and Hynes, 1993), and single and double cross-over recombinants of the resulting vector in the *S. meliloti* RmP110 background were obtained as described before (Quandt and Hynes, 1993; diCenzo and Finan, 2015).

A non-polar deletion of the *ngr_b20860* gene was constructed by cloning the *ngr_b20860* open reading frame ± approximately 500 nt (PCR amplified) into pUCP30T (Gm^R) via HindIII and EcoRI. The *ngr_b20860* open reading replaced with a FRT-Nm^R-FRT cassette using using λ *red* recombinase (Datsenko and Wanner, 2000). This construct was then transferred to *S. fredii* NGR234, and Gm^S Nm^R double recombinants isolated. The Nm^R gene was then deleted via Flp expressed from pTH2505 (Zhang *et al.*, 2012), following which pTH2505 was cured from the cell.

For expression of *smb20752 in trans*, the *dme* promoter region and the *smb20752* open reading frame were individually PCR amplified and cloned into the broad host range vector pTH1582 (Yuan *et al.*, 2005) via HindIII/BglII. The *dme* promoter region was used as it was unclear where the native *smb20752* promoter was located. The resulting plasmid (pTH2787) was then conjugated into the desired *S. meliloti* or *S. fredii* recipient.

Plant materials and growth conditions

Temperate plant materials were prepared and grown largely as described elsewhere (Yarosh *et al.*, 1989; diCenzo *et al.*, 2015). All seeds (*Medicago sativa* cv. Iroquois, *M. truncatula* cv. Jemalong, *M. alba* cv. Polara) were sterilized with 95% ethanol for 5 min, then 2.5% hypochlorite for 20 min. Sterilized seeds were washed, spread on 1.5% agar, incubated overnight at 4°C in the dark, then germinated at room temperature in the dark for either 48 h (*M. sativa*) or 72 h (*M. truncatula* and *M. alba*). Leonard assemblies were prepared as described before (diCenzo *et al.*, 2015), and 6-8 seedlings were potted in each. Two to three nights later, $\sim 1 \times 10^9$ CFU of *S. meliloti* were added to each Leonard assembly. Plants were grown for either 28 (*M. sativa*), 35 (*M. alba*), or 42 (*M. truncatula*) days in a Conviron growth chamber with a day (18 h, 21°C) and night (6 h, 17°C) cycle.

Tropical plant materials were treated largely as described before (Zhang *et al.*, 2012). *L. leucocephala* (Lain.) de Wit, *M. atropurpureum* cv. Aztec Atro, and *Vigna unguiculata* cv. Red Caloona seeds were scarified by treatment with concentrated H_2SO_4 for 5–10 min (*M. atropurpureum*), 10 min (*V. unguiculata*), or 15–20 min (*L. leucocephala*), thoroughly rinsed with sterile ddH_2O, soaked in sterile ddH_2O until swollen, surface sterilized with 3% sodium hypochlorite for 10 min, and repeatedly rinsed with sterile ddH_2O for 1 h. Sterilized seeds were spread on 1.5% agar, incubated at 4°C in the dark for up to 2 days, and then germinated in the dark for 2 days at room temperature. 4-6 seedlings were transferred to each Leonard assembly (diCenzo *et al.*, 2015), and inoculated with $\sim$ 1 x 10^9 *S. fredii* per pot 2 days later. Plants were grown in a Conviron growth chamber for either 42 (*M. atropurpureum* and *V. unguiculata*) or 70 (*L. leucocephala*) days with a day (16h, 28°C) and night (8h, 20°C) cycle.

Symbiotic assays

All symbiotic assays were performed as described previously (diCenzo *et al.*, 2015) and are explained briefly here. Shoot dry weights were determined by cutting off the shoots and drying at 50°C for 10 days prior to weighing. For acetylene reduction measurements, all roots from each Leonard assembly were cut off, immediately placed in a capped tube with 10% acetylene, and the rate of acetylene reduction over the subsequent 15 min was determined. All nodules on the root systems within each Leonard assembly were counted, picked off the roots, and weighed to determine nodule wet weight. *S. meliloti* were isolated from nodules by crushing surface sterilized nodules in TY + 300 mM sucrose, and plating dilutions on TY agar plates. To examine nodulation kinetics, individual sterile seedlings were placed on 1% agar slants of Jensen's medium (Jensen, 1942) in 18 $\times$ 150 mm test tubes. Each *S. meliloti* strain was inoculated onto thirteen seedlings, incubated in a Conviron growth chamber as described above, and visible nodules counted every few days.

Root colonization

To examine the ability of the *S. meliloti* strains to colonize *M. sativa* roots, *M. sativa* seedlings were planted in Leonard assemblies and inoculated with *S. meliloti* as described above. Plants were carefully removed from the Leonard assemblies 1, 3, 8, and 12 days post inoculation, and the roots excised at the crown. The roots were then weighed, transferred to 1 ml of phosphate buffered saline (PBS; 137 mM NaCl, 2.7 mM KCl, 10 mM Na_2HPO_4, 2 mM KH_2PO_4), and incubated in a sonicating water bath (Bransonic Ultrasonic cleaner M2800) for 30 min at room temperature. The PBS suspension was then serial diluted, and *S. meliloti* cells enumerated by plating on selective LB Sm agar plates.

β-Glucuronidase assays

β-glucuronidase activity in alfalfa nodules was determined following a previously published protocol (Cowie *et al.*, 2006). Ten nodules per sample were collected in a 1.5 ml microfuge tube and crushed in 750 μl MMS buffer (40 mM MOPS, 20 mM KOH, 2 mM MgSO4, 0.3 M sucrose; pH 7.0). Plant tissue was pelleted via centrifugation (400 *g* for 2 min), and 500 μl of the supernatant was transferred to a new tube. Sodium dodecyl sulphate was added to the supernatant to a concentration of 0.01%, and incubated on ice for 5 min. β-glucuronidase assays were performed in 96-well microtitre plates, and consisted of 10 μl of the previously prepared supernatant and 40 μl GUS buffer (Cowie *et al.*, 2006). Assays were terminated by the addition of 50 μL of 1 M Na_2CO_3, and absorbance at 420 nm was measured with a BioTek Cytation 3. The protein concentration of the supernatant was determined through a Bradford assay (Bradford, 1976), and specific activities were calculated as described before (Cowie *et al.*, 2006).

Histochemical analyses

Histochemical staining to detect β-glucuronidase activity in *M. sativa* nodules and infection threads was based off a previously described method (Boivin *et al.*, 1990). For examination of infection threads, sterile seedlings growing on 1% agar slants of Jensen's medium (Jensen, 1942) in 18 x 150 mm test tubes were inoculated with the appropriate *S. meliloti* strain and incubated in a Conviron growth chamber for five days as described above. Whole roots were then stained with X-gluc (5-bromo-4-chloro-3-indolyl-β-D-glucuronic acid) overnight ($\sim$ 12 to 16 h) as described before (Boivin *et al.*, 1990), with the roots either not fixed, or first gently fixed on ice with 1.25% ice cold glutaraldehyde in a 0.2 M sodium cacodylate buffer. Root sections were cleared in a solution of 1% sodium hypochlorite for 5 min, and then imaged using a Nikon TE2000 inverted microscope.

For histochemical staining of nodules, fresh nodules were removed from the plant roots and immediately mounted upon

a specimen plate using Instant Krazy Glue (Elmer's Products Canada). Nodules were sliced into 90 μm longitudinal sections using a vibrating blade microtome (Leica VT1000), and stained for 3–6 h at 30°C in staining buffer (200 mM sodium phosphate, 10 mM EDTA, 0.5 mM $K_3[Fe(CN)_6]$, 1.5 mM $K_4[Fe(CN)_6]$, pH 7.0). Stained sections were then washed twice in 200 mM sodium phosphate buffer (pH 7.0) for several hours. Nodule sections were cleared in a solution of 1% sodium hypochlorite for 5 min, and then imaged using a Nikon TE2000 inverted microscope.

Confocal microscopy and transmission electron microscopy

Confocal microscopy (CM) and transmission electron microscopy (TEM) were performed as described before (diCenzo *et al.*, 2015). A single change was made to the TEM procedure; to improve penetration of the glutaraldehyde fixative, large nodules were cut in half and all samples were placed in a vacuum oven overnight or longer as required. CM images were taken with a Lecia TCS SP2 confocal laser microscope at the McMaster Biology Confocal Facility, and TEM images taken using a JEOL JEM 1200 EX TEMSCAN transmission electron microscope operating at an accelerating voltage of 80 kV.

RNA extraction

For isolation of RNA from free-living *S. fredii* cells, the desired strains were subcultured to an $OD_{600} \sim 0.05$ in RMM-succinate, and grown to an $OD_{600} \sim 0.4$. For each culture, the equivalent of 1 ml at an OD_{600} of 1 was collected, centrifuged at 10 000 *g* for 2 min, and the cell pellet flash frozen in liquid nitrogen and stored at −80°C until use. The cell pellets were resuspended in 50 μl "hot SDS lysis" buffer [20 mM Tris-HCl (pH 8), 400 mM NaCl, 40 mM EDTA (Ethylenediaminetetraacetic acid), 1% SDS (Sodium dodecyl sulfate), and 1% BME (β-mercaptoethanol)], placed in boiling water for 75 seconds to lyse the cells, and diluted with 50 μl nuclease free ddH_2O. This solution was then mixed with 350 μl of "lysis buffer" from the PureLink® RNA Mini Kit (Life Technologies), 250 μl of 100% ethanol, and purified using the PureLink® RNA Mini Kit according to the manufacturer's instructions. RNA was collected in 45 μl of nuclease free ddH_2O, and DNA contamination was removed by treatment with the TURBO DNA-*free*™ Kit (Life Technologies) using the routine DNase treatment as described by the manufacturer. The quality of the purified RNA was examined on a MOPS-formaldehyde agarose gel and by UV spectrometry, and the absence of contaminating DNA was confirmed via PCR with published primers (diCenzo, Zamani, *et al.*, 2016) amplifying the *engA* and *rmlC* genes.

Root nodules of *V. unguiculata* and *M. atropurpureum* were collected 6 weeks post inoculation with *S. fredii*, whereas *L. leucocephala* were collected 10 weeks post inoculation. Nodules were immediately flash frozen in liquid nitrogen and stored at −80°C until use. For RNA purification, nodules from four plants per sample were combined and a total of 300 mg of nodules per sample were taken. Nodule samples were individually ground in liquid nitrogen with a mortar and pestle to a very fine powder. Both plant and bacterial RNA was extracted from the ground nodule material using a Plant/Fungi Total RNA Purification Kit (Norgen, Bioteck Corp. Canada) according to the manufacturer's instructions. Two DNase treatments were performed on the purified RNA, the first using an on column RNase-Free DNase Set (Qiagen, Hilden, Germany) and the second with the TURBO DNA-*free*™ Kit (Life Technologies). The quality of the purified RNA was examined on a MOPS-formaldehyde agarose gel and by UV spectrometry, and the absence of contaminating bacterial DNA was confirmed via PCR as described above

RT-qPCR

The synthesis of cDNA was performed with using the SuperScript® IV First-Strand Synthesis System (Invitrogen) according to the manufacturer's instructions. Each cDNA synthesis reaction included 600 ng total RNA and 2.5 ng/μl random hexamers.

The qPCR reactions were performed using PerfeCTa® SYBR® Green SuperMix (Quanta Biosciences) according to the manufacturer's instructions. The final volume of each reaction was 20 μl, and each reaction included 5 μl of a 5-fold dilution of the cDNA synthesis reaction and 300 nM of each primer. Target genes included *nifH* (forward primer: 5′-GTG GAT GAG CCT GGA ATT G; reverse primer: 5′-TAT GCC GCC AAC AAC ATC) and *ngr_b20860* (forward primer: 5′-TTT CGG ATT GCG GTC CTT GT; reverse primer: 5′-GTT GAA CGA ATG CCT GGA GC), and the reference gene was *rpoD* (forward primer: 5′-ACA TCA CCA ATG TCG GCG GTG AAG; reverse primer: 5′-TGC AGC TTG CGG AGC TTC TTG TAG). The qPCR reactions were performed using a BioRad Real Time PCR machine, and cycling conditions were: 95°C for 3 min, followed by 40 cycles consisting of 15 seconds at 95°C and 45 seconds at 60°C

Bioinformatics analyses

To examine the conservation of *smb20752* and of the genomic context of *smb20752* in other rhizobial species, proteins homologous to Smb20752 (Accession No. NP_437984.1) were identified in target species using a Blast-BBH approach. As hundreds of draft rhizobial genomes are now publicly available, this analysis was limited by primarily selecting those whose genome is annotated as finished on the National Center for Biotechnology Information (NCBI) Genome database. This strain set was supplemented with a few additional genomes if they were at least in scaffolds to improve coverage of the rhizobial species. Genomes were downloaded through the NCBI ftp site and the Blast-BBH approach performed using the command line BLAST+ software (Camacho *et al.*, 2009) and custom shell and perl scripts. Hits were mapped back to the corresponding replicon (chromosome or megaplasmid) using custom perl scripts, and orthologs were manually examined to see if they were situated within a symbiotic island.

The genomic localization of the *smb20752* orthologs was examined using custom perl and shell scripts to pull out the genes upstream and downstream of the *smb20752* ortholog in the corresponding .gff file, and the proteins encoded by

these genes were searched against the *S. meliloti* proteome with BLASTp to determine their best hit.

To construct a phylogeny based on the amino acid sequence of the Smb20752 orthologs, the orthologs were first aligned using MAFFT-linsi (Katoh and Standley, 2013) and trimmed with trimAl and the automated1 algorithm (Capella-Gutiérrez *et al*., 2009). Phylogenetic analysis of the resulting trimmed alignment was performed with the RAxML BlackBox webserver (Stamatakis *et al*., 2008) and the maximum likelihood bootstrap best tree is presented following 100 bootstrap replicates. The same process was employed to construct the Smc01153 phylogenetic tree.

To construct a MLSA tree, the MarkerScanner.pl script of the AMPHORA2 pipeline was first used to detect 31 highly conserved bacterial proteins in the analyzed strains (Wu and Scott, 2012). Those proteins not present in single copy in each of the strains were discarded, leaving 27 proteins. Each set of proteins were aligned with MAFFT-linsi (Katoh and Standley, 2013) and trimmed with trimAl and the automated1 algorithm (Capella-Gutiérrez *et al*., 2009), following which the trimmed alignments were concatenated with the help of the Geneious R8 software. The concatenated alignment was submitted to the RAxML BlackBox webserver (Stamatakis *et al*., 2008) and the maximum likelihood bootstrap best tree is presented following 100 bootstrap replicates.

Acknowledgements

We are thankful to Marcia Reid for assistance with the transmission electron microscopy, and to Trevor Charles for providing the pTC322 cosmid clone and *phaZ* mutant used during our studies. This work was funded by the Natural Sciences and Engineering Research Council of Canada through grants to TMF, and a NSERC CGS-D scholarship to GCD.

References

Aneja, P., and Charles, T.C. (2005) Characterization of *bdhA*, encoding the enzyme d-3-hydroxybutyrate dehydrogenase, from *Sinorhizobium* sp. strain NGR234. *FEMS Microbiol Lett* **242:** 87–94.

Barnett, M.J., and Kahn, M.L. (2005) pSymA of *Sinorhizobium meliloti*: nitrogen fixation and more. In *Genomes and Genomics of Nitrogen-Fixing Organisms*. Nitrogen Fixation: Origins, Applications, and Research Progress. Berlin/Heidelberg: Springer Netherlands, pp. 113–132.

Boivin, C., Camut, S., Malpica, C.A., Truchet, G., and Rosenberg, C. (1990) *Rhizobium meliloti* genes encoding catabolism of trigonelline are induced under symbiotic conditions. *Plant Cell* **2:** 1157–1170.

Bradford, M.M. (1976) A rapid and sensitive method for the quantitation of microgram quantities of protein utilizing the principle of protein-dye binding. *Anal Biochem* **72:** 248–254.

van Brussel, A.A.N., Tak, T., Boot, K.J.M., and Kijne, J.W. (2002) Autoregulation of root nodule formation: signals of both symbiotic partners studied in a split-root system of *Vicia sativa* subsp. nigra. *Mol Plant Microbe Interact* **15:** 341–349.

Camacho, C., Coulouris, G., Avagyan, V., Ma, N., Papadopoulos, J., Bealer, K., and Madden, T.L. (2009) BLAST+: architecture and applications. *BMC Bioinform* **10:** 421.

Capella-Gutiérrez, S., Silla-Martínez, J.M., and Gabaldón, T. (2009) trimAl: a tool for automated alignment trimming in large-scale phylogenetic analyses. *Bioinformatics* **25:** 1972–1973.

Charles, T.C., and Aneja, P. (1999) Methylmalonyl-CoA mutase encoding gene of *Sinorhizobium meliloti*. *Gene* **226:** 121–127.

Charles, T.C., Cai, G.Q., and Aneja, P. (1997) Megaplasmid and chromosomal loci for the PHB degradation pathway in *Rhizobium (Sinorhizobium) meliloti*. *Genetics* **146:** 1211–1220.

Cowie, A., Cheng, J., Sibley, C.D., Fong, Y., Zaheer, R., Patten, C.L., *et al*. (2006) An integrated approach to functional genomics: construction of a novel reporter gene fusion library for *Sinorhizobium meliloti*. *Appl Environ Microbiol* **72:** 7156–7167.

Curson, A.R.J., Burns, O.J., Voget, S., Daniel, R., Todd, J.D., McInnis, K., *et al*. (2014) Screening of metagenomic and genomic libraries reveals three classes of bacterial enzymes that overcome the toxicity of acrylate. *PLoS One* **9:** e97660.

Datsenko, K.A., and Wanner, B.L. (2000) One-step inactivation of chromosomal genes in *Escherichia coli* K-12 using PCR products. *Proc Natl Acad Sci USA* **97:** 6640–6645.

diCenzo, G.C., Checcucci, A., Bazzicalupo, M., Mengoni, A., Viti, C., Dziewit, L., *et al*. (2016) Metabolic modelling reveals the specialization of secondary replicons for niche adaptation in *Sinorhizobium meliloti*. *Nat Commun* **7:** 12219.

diCenzo, G.C., and Finan, T.M. (2015) Genetic redundancy is prevalent within the 6.7 Mb *Sinorhizobium meliloti* genome. *Mol Genet Genomics* **290:** 1345–1356.

diCenzo, G.C., MacLean, A.M., Milunovic, B., Golding, G.B., and Finan, T.M. (2014) Examination of prokaryotic multipartite genome evolution through experimental genome reduction. *PLoS Genet* **10:** e1004742.

diCenzo, G.C., Zamani, M., Cowie, A., and Finan, T.M. (2015) Proline auxotrophy in *Sinorhizobium meliloti* results in a plant-specific symbiotic phenotype. *Microbiology* **161:** 2341–2351.

diCenzo, G.C., Zamani, M., Milunovic, B., and Finan, T.M. (2016) Genomic resources for identification of the minimal N2-fixing symbiotic genome. *Environ Microbiol* **18:** 2534–2547.

Djordjevic, M.A. (2004) *Sinorhizobium meliloti* metabolism in the root nodule: a proteomic perspective. *Proteomics* **4:** 1859–1872.

Erisman, J.W., Sutton, M.A., Galloway, J., Klimont, Z., and Winiwarter, W. (2008) How a century of ammonia synthesis changed the world. *Nat Geosci* **1:** 636–639.

Fei, F., diCenzo, G.C., Bowdish, D.M.E., McCarry, B.E., and Finan, T.M. (2016) Effects of synthetic large-scale genome reduction on metabolism and metabolic preferences in a nutritionally complex environment. *Metabolomics* **12:** 23.

Ferguson, B.J., Indrasumunar, A., Hayashi, S., Lin, M.H., Lin, Y.H., Reid, D.E., and Gresshoff, P.M. (2010) Molecular analysis of legume nodule development and autoregulation. *J Integr Plant Biol* **52:** 61–76.

Finan, T.M., Hartweig, E., LeMieux, K., Bergman, K., Walker, G.C., and Signer, E.R. (1984) General transduction in *Rhizobium meliloti*. *J Bacteriol* **159:** 120–124.

Finan, T.M., Kunkel, B., De Vos, G.F., and Signer, E.R. (1986) Second symbiotic megaplasmid in *Rhizobium meliloti* carrying exopolysaccharide and thiamine synthesis genes. *J Bacteriol* **167:** 66–72.

Galardini, M., Brilli, M., Spini, G., Rossi, M., Roncaglia, B., Bani, A., *et al*. (2015) Evolution of intra-specific regulatory networks in a multipartite bacterial genome. *PLoS Comput Biol* **11:** e1004478.

Galardini, M., Pini, F., Bazzicalupo, M., Biondi, E.G., and Mengoni, A. (2013) Replicon-dependent bacterial genome evolution: the case of *Sinorhizobium meliloti*. *Genome Biol Evol* **5:** 542–558.

Galibert, F., Finan, T.M., Long, S.R., Pühler, A., Abola, A.P., Ampe, F., *et al*. (2001) The composite genome of the legume symbiont *Sinorhizobium meliloti*. *Science* **293:** 668–672.

Geddes, B.A., Ryu, M.H., Mus, F., Garcia Costas, A., Peters, J.W., Voigt, C.A., and Poole, P. (2015) Use of plant colonizing bacteria as chassis for transfer of N_2-fixation to cereals. *Curr Opin Biotechnol* **32:** 216–222.

Gemperline, E., Jayaraman, D., Maeda, J., Ané, J.M., and Li, L. (2015) Multifaceted investigation of metabolites during nitrogen fixation in *Medicago* via high resolution MALDI-MS imaging and ESI-MS. *J Am Soc Mass Spectrom* **26:** 149–158.

Hirsch, A.M., Bang, M., and Ausubel, F.M. (1983) Ultrastructural analysis of ineffective alfalfa nodules formed by *nif*:: Tn*5* mutants of *Rhizobium meliloti*. *J Bacteriol* **155:** 367–380.

Jensen, H.L. (1942) Nitrogen fixation in leguminous plants. I. General characters of root-nodule bacteria isolated from species of *Medicago* and *Trifolium* in Australia. *Proc Linn Soc NSW* **67:** 98–108.

Katoh, K., and Standley, D.M. (2013) MAFFT multiple sequence alignment software version 7: improvements in performance and usability. *Mol Biol Evol* **30:** 772–780.

Korotkova, N., Chistoserdova, L., Kuksa, V., and Lidstrom, M.E. (2002) Glyoxylate regeneration pathway in the methylotroph *Methylobacterium extorquens* AM1. *J Bacteriol* **184:** 1750–1758.

Kosslak, R.M., and Bohlool, B.B. (1984) Suppression of nodule development of one side of a split-root system of soybeans caused by prior inoculation of the other side. *Plant Physiol* **75:** 125–130.

Li, Y., Tian, C.F., Chen, W.F., Wang, L., Sui, X.H., and Chen, W.X. (2013) High-resolution transcriptomic analyses of *Sinorhizobium* sp. NGR234 bacteroids in determinate nodules of *Vigna unguiculata* and indeterminate nodules of *Leucaena leucocephala*. *PLoS One* **8:** e70531.

Massey, L.K., Sokatch, J.R., and Conrad, R.S. (1976) Branched-chain amino acid catabolism in bacteria. *Bacteriol Rev* **40:** 42–54.

Milunovic, B., diCenzo, G.C., Morton, R.A., and Finan, T.M. (2014) Cell growth inhibition upon deletion of four toxin-antitoxin loci from the megaplasmids of *Sinorhizobium meliloti*. *J Bacteriol* **196:** 811–824.

Miyamoto, E., Watanabe, F., Charles, T.C., Yamaji, R., Inui, H., and Nakano, Y. (2003) Purification and characterization of homodimeric methylmalonyl-CoA mutase from *Sinorhizobium meliloti*. *Arch Microbiol* **180:** 151–154.

Mortier, V., Holsters, M., and Goormachtig, S. (2012) Never too many? How legumes control nodule numbers. *Plant Cell Environ* **35:** 245–258.

Mutuma, S.P., Okello, J.J., Karanja, N.K., and Woomer, P.I. (2014) Smallholder farmers' use and profitability of legume inoculants in western Kenya. *Afr Crop Sci J* **22:** 205–214.

de las Nieves Peltzer, M., Roques, N., Poinsot, V., Aguilar, O.M., Batut, J., and Capela, D. (2008) 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.

Oldroyd, G.E., and Dixon, R. (2014) Biotechnological solutions to the nitrogen problem. *Curr Opin Biotechnol* **26:** 19–24.

Oldroyd, G.E.D., Murray, J.D., Poole, P.S., and Downie, J.A. (2011) The rules of engagement in the legume-rhizobial symbiosis. *Annu Rev Genet* **45:** 119–144.

Oono, R., Schmitt, I., Sprent, J.I., and Denison, R.F. (2010) Multiple evolutionary origins of legume traits leading to extreme rhizobial differentiation. *New Phytol* **187:** 508–520.

Paau, A.S., Cowles, J.R., and Raveed, D. (1978) Development of bacteroids in alfalfa (*Medicago sativa*) nodules. *Plant Physiol* **62:** 526–530.

Pobigaylo, N., Wetter, D., Szymczak, S., Schiller, U., Kurtz, S., Meyer, F., *et al*. (2006) Construction of a large signature-tagged mini-Tn*5* transposon library and its application to mutagenesis of *Sinorhizobium meliloti*. *Appl Environ Microbiol* **72:** 4329–4337.

Prell, J., Bourdès, A., Kumar, S., Lodwig, E., Hosie, A., Kinghorn, S., *et al*. (2010) Role of symbiotic auxotrophy in the *Rhizobium*-legume symbioses. *PLoS One* **5:** e13933.

Pueppke, S.G., and Broughton, W.J. (1999) *Rhizobium* sp. strain NGR234 and *R. fredii* USDA257 share exceptionally broad, nested host ranges. *Mol Plant Microbe Interact* **12:** 293–318.

Quandt, J., and Hynes, M.F. (1993) Versatile suicide vectors which allow direct selection for gene replacement in Gram-negative bacteria. *Gene* **127:** 15–21.

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

Rogers, C., and Oldroyd, G.E.D. (2014) Synthetic biology approaches to engineering the nitrogen symbiosis in cereals. *J Exp Bot* **65:** 1939–1946.

Rosenberg, C., Boistard, P., Dénarié, J., and Casse-Delbart, F. (1981) Genes controlling early and late functions in symbiosis are located on a megaplasmid in *Rhizobium meliloti*. *Mol Gen Genet* **184:** 326–333.

Roux, B., Rodde, N., Jardinaud, M.F., Timmers, T., Sauviac, L., Cottret, L., *et al*. (2014) 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.

Sambrook, J., Fritsch, E.F., and Maniatis, T. (1989) *Molecular Cloning: A Laboratory Manual*. New York: Cold Spring Harbor Laboratory.

Schlüter, J.P., Reinkensmeier, J., Daschkey, S., Evguenieva-Hackenberg, E., Janssen, S., Jänicke, S., *et al*. (2010) A genome-wide survey of sRNAs in the symbiotic nitrogen-

fixing alpha-proteobacterium *Sinorhizobium meliloti*. *BMC Genomics* **11:** 245.

Schmeisser, C., Liesegang, H., Krysciak, D., Bakkou, N., Le Quéré, A., Wollherr, A., *et al*. (2009) *Rhizobium* sp. strain NGR234 possesses a remarkable number of secretion systems. *Appl Environ Microbiol* **75:** 4035–4045.

Schroeder, B.K., House, B.L., Mortimer, M.W., Yurgel, S.N., Maloney, S.C., Ward, K.L., and Kahn, M.L. (2005) Development of a functional genomics platform for *Sinorhizobium meliloti*: construction of an ORFeome. *Appl Environ Microbiol* **71:** 5858–5864.

Shimomura, Y., Murakami, T., Fujitsuka, N., Nakai, N., Sato, Y., Sugiyama, S., *et al*. (1994) Purification and partial characterization of 3-hydroxyisobutyryl-coenzyme A hydrolase of rat liver. *J Biol Chem* **269:** 14248–14253.

Sobrero, P., Schlüter, J.P., Lanner, U., Schlosser, A., Becker, A., and Valverde, C. (2012) Quantitative proteomic analysis of the Hfq-regulon in *Sinorhizobium meliloti* 2011. *PLoS One* **7:** e48494.

Stamatakis, A., Hoover, P., and Rougemont, J. (2008) A rapid bootstrap algorithm for the RAxML Web servers. *Syst Biol* **57:** 758–771.

Stanley, J., Dowling, D.N., and Broughton, W.J. (1988) Cloning of *hemA* from *Rhizobium* sp. NGR234 and symbiotic phenotype of a gene-directed mutant in diverse legume genera. *Mol Gen Genet* **215:** 32–37.

Sugawara, M., Epstein, B., Badgley, B.D., Unno, T., Xu, L., Reese, J., *et al*. (2013) Comparative genomics of the core and accessory genomes of 48 *Sinorhizobium* strains comprising five genospecies. *Genome Biol* **14:** R17.

Trainer, M.A. (2009) Carbon metabolism and desiccation tolerance in the nitrogen-fixing rhizobia *Bradyrhizobium japonicum* and *Sinorhizobium meliloti*. PhD Thesis. Waterloo, ON, Canada: University of Waterloo.

Trainer, M.A., Capstick, D., Zachertowska, A., Lam, K.N., Clark, S.R., and Charles, T.C. (2010) Identification and characterization of the intracellular poly-3-hydroxybutyrate depolymerase enzyme PhaZ of *Sinorhizobium meliloti*. *BMC Microbiol* **10.**

Trainer, M.A., and Charles, T.C. (2006) The role of PHB metabolism in the symbiosis of rhizobia with legumes. *Appl Microbiol Biotechnol* **71:** 377–386.

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

Wang, C., Saldanha, M., Sheng, X., Shelswell, K.J., Walsh, K.T., Sobral, B.W.S., and Charles, T.C. (2007) Roles of poly-3-hydroxybutyrate (PHB) and glycogen in symbiosis of *Sinorhizobium meliloti* with *Medicago* sp. *Microbiology* **153:** 388–398.

Werner, G.D.A., Cornwell, W.K., Sprent, J.I., Kattge, J., and Kiers, E.T. (2014) A single evolutionary innovation drives the deep evolution of symbiotic N_2-fixation in angiosperms. *Nat Commun* **5:** 4087.

Wu, M., and Scott, A.J. (2012) Phylogenomic analysis of bacterial and archaeal sequences with AMPHORA2. *Bioinformatics* **28:** 1033–1034.

Yarosh, O.K., Charles, T.C., and Finan, T.M. (1989) Analysis of C_4-dicarboxylate transport genes in *Rhizobium meliloti*. *Mol Microbiol* **3:** 813–823.

Yuan, Z.C., Zaheer, R., and Finan, T.M. (2005) Phosphate limitation induces catalase expression in *Sinorhizobium meliloti*, *Pseudomonas aeruginosa* and *Agrobacterium tumefaciens*. *Mol Microbiol* **58:** 877–894.

Yuan, Z.C., Zaheer, R., and Finan, T.M. (2006) Regulation and properties of PstSCAB, a high-affinity, high-velocity phosphate transport system of *Sinorhizobium meliloti*. *J Bacteriol* **188:** 1089–1102.

Yurgel, S.N., Mortimer, M.W., Rice, J.T., Humann, J.L., and Kahn, M.L. (2013) Directed construction and analysis of a *Sinorhizobium meliloti* pSymA deletion mutant library. *Appl Environ Microbiol* **79:** 2081–2087.

Zhang, Y., Aono, T., Poole, P., and Finan, T.M. (2012) NAD(P)$^+$-malic enzyme mutants of *Sinorhizobium* sp. Strain NGR234, but not *Azorhizobium caulinodans* ORS571, maintain symbiotic N_2 fixation capabilities. *Appl Environ Microbiol* **78:** 2803–2812.

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

Zolman, B.K., Monroe-Augustus, M., Thompson, B., Hawes, J.W., Krukenberg, K.A., Matsuda, S.P., and Bartel, B. (2001) *chy1*, an Arabidopsis mutant with impaired beta-oxidation, is defective in a peroxisomal beta-hydroxyisobutyryl-CoA hydrolase. *J. Biol. Chem* **276:** 31037–31046.

Supporting information

Additional Supporting Information may be found in the online version of this article at the publisher's web-site:

Fig. S1. Tropical legumes inoculated with various *S. fredii* NGR234 strains. Pictures of the plants were taken 28 days post inoculation with the different *S. fredii* NGR234 strains. The genotype of the strain inoculated in each pot is indicated. (a) *V. unguiculata*, (b) *M. atropurpureum*, and (c) *L. leucocephala*.

Fig. S2. The involvement of *ngr_b20860* in β-hydroxybutyrate catabolism. Growth of *S. fredii* NGR234 wild type (blue circles), Δ*ngr_b20860* (orange squares), and Δ*ngr_b20860* with *smb20752 in trans* (green triangles) in M9 media with either (a) glucose or (b) β-hydroxybutyrate as the sole carbon source. Data points represent the mean of triplicate samples, with the error bars indicating the standard deviation.

Fig. S3. Sensitivity of *S. meliloti* strains to acrylate. Growth curves for (A) wild type *S. meliloti*, (B) the Δ*smb20752* mutant, and (C) the Δ*smb20752* mutant with *smb20752* expressed *in trans* from a plasmid are shown. All cultures were grown in M9 minimal medium with succinate as the sole carbon source and with varying concentrations of acrylate as indicated. Data points represent the mean of triplicate samples, with the error bars indicating the standard deviation.

Fig. S4. Phylogenetic analysis of rhizobial Smb20752 orthologs. Rhizobial proteins orthologous to Smb20752 were identified via a Blast-BBH approached, aligned with mafft (Katoh and Standley, 2013), trimmed with trimAl (Capella-Gutiérrez *et al*., 2009), and an unrooted maximum likelihood phylogeny produced with RAxML through the BlackBox server (Stamatakis *et al*., 2008) as described in the materials and methods. The location of the Smb20752 orthologs

are indicated along the right, and none of the orthologs on a chromosome are within a symbiotic island. Taxa are colour coded based on the genomic context of the *smb20752* as indicated at the bottom of the figure. Bootstrap values are indicated from 100 bootstrap replicates.

Fig. S5. Multilocus sequence analysis of the rhizobial strains examined in this study. The AMPHORA2 pipeline (Wu and Scott, 2012) was used to identify and collect orthologs of 27 highly conserved bacterial proteins (Frr, InfC, PyrG, RplA, RplB, RplC, RplD, RplE, RplF, RplK, RplL, RplM, RplN, RplP, RplS, RplT, RpmA, RpoB, RpsB, RpsC, RpsE, RpsI, RpsJ, RpsK, RpsM, RpsS, Tsf) that are present in all genomes used in this work. Each set of orthologs were individually aligned with mafft (Katoh and Standley, 2013) and trimmed with trimAl (Capella-Gutiérrez *et al.*, 2009). The 27 alignments were then concatenated and an unrooted maximum likelihood phylogeny was produced with RAxML through the BlackBox server (Stamatakis *et al.*, 2008). Bootstrap values are indicated from 100 bootstrap replicates.

Fig. S6. Phylogenetic analysis of rhizobial Smc01153 orthologs. Rhizobial proteins orthologous to Smc01153 were identified via a Blast-BBH approached, aligned with mafft (Katoh and Standley, 2013), trimmed with trimAl (Capella-Gutiérrez *et al.*, 2009), and an unrooted maximum likelihood phylogeny produced with RAxML through the BlackBox server (Stamatakis *et al.*, 2008) as described in the materials and methods. Bootstrap values are indicated from 100 bootstrap replicates.

Table S1. Bacterial strains and plasmids.

Table S2. Statistical analysis of the shoot dry weight phenotypes of *M. sativa* plants inoculated with various *S. meliloti* strains.

Table S3. Effect of *phaZ* disruption on the *smb20752* Fix phenotype.