Title: DNA methylation patterns in bacteria of the genus *Ensifer* during free-living growth and during nitrogen-fixing symbiosis with *Medicago* spp.

Running title: DNA methylation in the genus *Ensifer*

Authors: George C. diCenzo[1,2,#], Lisa Cangioli[2], Quentin Nicoud[3], Janis H.T. Cheng[1], Matthew J. Blow[4], Nicole Shapiro[4], Tanja Woyke[4], Benoît Alunni[3], Alessio Mengoni[2,*], and Peter Mergaert[3,*]

Affiliations: [1] Department of Biology, Queen's University, Kingston ON, Canada
[2] Department of Biology, University of Florence, Florence, Italy
[3] Université Paris-Saclay, CEA, CNRS, Institute for Integrative Biology of the Cell, Gif-sur-Yvette, France
[4] U.S. Department of Energy Joint Genome Institute, Berkeley CA, USA

 # Corresponding author: George C. diCenzo, george.dicenzo@queensu.ca

* These authors contributed equally to this work.

Keywords: Rhizobia, terminal differentiation, bacteroid, symbiotic nitrogen fixation, epigenetics, DNA methylation, Proteobacteria, cell cycle regulation, CcrM

ABSTRACT

Methylation of specific genomic DNA sequences is ubiquitous in bacteria and has known roles in immunity and regulation of cellular processes, such as the cell cycle. Here, we explored DNA methylation in bacteria of the genus *Ensifer*, including its potential role in regulating the process of terminal differentiation occurring during nitrogen-fixing symbiosis with legumes. Using single-molecule real-time sequencing, six unique genome-wide methylated motifs were identified across four *Ensifer* strains, five of which were strain-specific. These five motifs were nearly fully methylated across the genomes in all tested conditions, and they were not enriched in the promoter regions of symbiosis, carbon source, or cell cycle-regulated genes, suggesting that most DNA methylation is not a major regulatory mechanism in the genus *Ensifer*. Only the GANTC motif, recognized by the cell cycle-regulated CcrM methyltransferase, was methylated in all strains. In actively dividing cells, methylation of GANTC motifs increased progressively from the *ori* to *ter* region in each replicon, in agreement with a cell cycle-dependent regulation of CcrM. The GANTC methylation profile transited into a genome-wide pattern of near full methylation in the early stage of symbiotic differentiation, followed by a progressive decrease in methylation from the *ori* to *ter* regions of fully differentiated symbiotic bacteria. This is evidence of a dysregulated and constitutive CcrM activity during terminal differentiation, which we suggest is a driving factor for endoreduplication of terminally differentiated bacteroids.

IMPORTANCE

Nitrogen fixation by bacteria (rhizobia) in symbiosis with legumes is economically and ecologically important. In some cases, the symbiosis involves a complex bacterial transformation, known as terminal differentiation, that includes major shifts in the transcriptome and cell cycle. Epigenetic regulation via DNA methylation is an important regulatory mechanism contributing to the biology of diverse bacteria; however, the roles of DNA methylation in rhizobia and symbiotic nitrogen fixation have been poorly investigated. We show that aside from cell cycle regulation, DNA methylation is unlikely to be a major mechanism of transcriptional regulation in rhizobia and non-rhizobia of the genus *Ensifer*. However, we found strong evidence that the cell cycle methyltransferase CcrM is dysregulated during symbiosis, which may be a key factor driving the cell cycle switch in terminal differentiation and the establishment of effective rhizobium – legume symbioses. These novel results advance our understanding of this highly important, yet incompletely understood process.

55 **INTRODUCTION**

56 Methylation of genomic DNA is a pervasive phenomenon found in eukaryotes (1–3), archaea (4),
57 and bacteria (4, 5). The biological roles of DNA methylation are most extensively studied in
58 mammals, where it contributes to normal development and disease via its impact on gene
59 expression (6). In bacteria, DNA methylation is best known for its role in restriction-modification
60 (R-M) systems that are thought to provide defence against phage infection and limit horizontal
61 gene transfer through the degradation of invading non-methylated DNA (7). Several
62 methyltransferases (MTases) of R-M systems have also been implicated in phase variation in
63 pathogens through modulating gene expression (8). A recent study of over 200 bacterial and
64 archaeal species identified orphan MTases not belonging to R-M systems in nearly half of the
65 genomes (4). To date, biological functions have been attributed to very few orphan MTases,
66 namely, the Dam MTase of the γ-Proteobacteria and the CcrM MTase of the α-Proteobacteria (9).
67 The Dam MTase of *Escherichia coli* is notable for its role in regulation of DNA replication (10,
68 11) and DNA repair (12) by modulating the activity of other DNA-binding proteins. The CcrM
69 MTase was first identified in *Caulobacter crescentus* (13), with homologs since identified in
70 diverse α-Proteobacteria (14, 15). CcrM activity was shown to be cell cycle regulated in *C.*
71 *crescentus* and *Agrobacterium tumefaciens* (13, 16), leading to methylation of its cognate DNA
72 motif (the pentanucleotide GANTC) specifically during a short period at the end of DNA
73 replication. This leads to a switching of GANTC sites between fully methylated (methylated on
74 both strands) and hemi-methylated (methylated only on the template strand) as a result of DNA
75 replication (17), which serves to modulate gene expression in a cell cycle-dependent fashion (18–
76 20). Over- and under-expression of *ccrM* results in defects in DNA replication and cell division
77 (14, 16, 19), while its complete loss is generally lethal.

78 The rhizobia are a polyphyletic group of α-Proteobacteria and β-Proteobacteria that can
79 both live free in the soil and enter into an endosymbiotic interaction with legumes (21). This
80 interaction begins following an exchange of signals between the free-living partners (22), and it
81 culminates in the formation of a new organ known as a root nodule within which the cytoplasm of
82 plant cells contain thousands of N_2-fixing bacteria called bacteroids. Bacteroid formation results
83 in the differential expression of more than a thousand genes (23, 24) and global changes in cellular
84 metabolism (25). In legumes of the Inverted-Repeat Lacking Clade (IRLC) and the Dalbergioid
85 clade of the family *Papilionoideae*, bacteroid development involves an additional process of
86 terminal differentiation (26, 27); in other legume clades, bacteroid differentiation is less
87 pronounced and is reversible. Terminal bacteroid development, in contrast to reversible bacteroid
88 formation, involves cell enlargement (bacteroids are 5- to 10-fold longer than their free-living
89 counterparts) and genome endoreduplication (resulting in up to 24 copies of the genome per cell),
90 indicative of a cell cycle transition occurring during differentiation (26). Indeed, the correct
91 expression of cell cycle regulators in *Ensifer* (syn. *Sinorhizobium*) *meliloti*, a symbiont of
92 *Medicago* species of the IRLC, is essential for the formation of functional bacteroids (28, 29),
93 while over-expression of CcrM or disruption of the master cell cycle regulator CtrA can give rise
94 to bacteroid-like morphology in free-living cells (14, 30). Additionally, mutants in the *E. meliloti*

95 cell cycle regulators *divJ*, *cbrA*, and *cpdR1*, encoding three negative regulators of CtrA, form non-
96 functional nodules in which bacteroids do not differentiate properly (28, 29, 31), and genes
97 encoding several cell cycle regulators (including CcrM) are strongly downregulated in bacteroids
98 (24). The differentiation and cell cycle switch of bacteroids is controlled by the legume host
99 through the production of a large family of peptides, known as Nodule-specific Cysteine-Rich
100 (NCR) peptides (32–34).
101 Multiple studies have provided evidence that changes in the methylation status of the DNA
102 of legume nodule cells contributes to symbiotic development (35–37). Conversely, it remains
103 unknown if methylation of rhizobium DNA contributes to the regulation of N_2-fixation or
104 bacteroid development. We are aware of only one study (38) comparing DNA methylation of a
105 rhizobium (*Bradyrhizobium diazoefficiens* USDA110) between free-living and symbiotic states
106 (soybean nodules), and comparing these changes with differential expression data. Intriguingly,
107 the authors identified a DNA motif that was methylated specifically in bacteroids (38). However,
108 no clear evidence was presented that methylation of this (or any other) motif is involved in
109 transcriptional regulation, and the number of genes both differentially expressed and differentially
110 methylated in bacteroids did not appear to be different than expected by chance. While these data
111 may suggest that DNA methylation does not play a major role in regulating N_2-fixation by rhizobia,
112 they do not address the role of DNA methylation in terminal bacteroid differentiation, as *B.*
113 *diazoefficiens* undergoes reversible differentiation in soybean nodules (39).
114 Here, we use Pacific Biosciences Single-Molecule Real-Time (SMRT) sequencing to
115 detect genome-wide patterns of DNA methylation in four strains belonging to the genus *Ensifer*.
116 Our results indicate that DNA methylation is poorly conserved across the genus and suggests that
117 DNA methylation is not a major mechanism of regulating gene expression in these organisms aside
118 from cell cycle control. However, analysis of bacteroid samples suggests that constitutive
119 activation of the CcrM MTase may be a contributing factor driving terminal differentiation.
120

121 **RESULTS**

122 **The methylomes of the genus *Ensifer***

123 Our experimental design, as summarized in the Materials and Methods and **Figure S1**, was
124 developed to support an investigation into multiple potential roles of DNA methylation in plant-
125 associated bacteria from the genus *Ensifer* through the use of SMRT sequencing. This was
126 accomplished by: i) including DNA samples isolated from phylogenetically diverse wild-type
127 strains, ii) examining DNA methylation in a single strain across multiple conditions (exponential
128 phase growth versus stationary phase; growth with sucrose versus growth with succinate), iii)
129 investigating the impact of a large-scale genome reduction on DNA methylation patterns, and iv)
130 isolating DNA from bacteroids purified from legume nodules.
131 We began with a base modification analyses of four wild-type strains from three species,
132 including three nodule-forming strains (*E. meliloti* Rm2011, *E. meliloti* FSM-MA, *E. fredii*
133 NGR234) and one plant-associated, non-symbiotic strain (*E. adhaerens* OV14) (**Figure 1**). To
134 ensure consistency, all strains were grown to mid-exponential phase in a common minimal

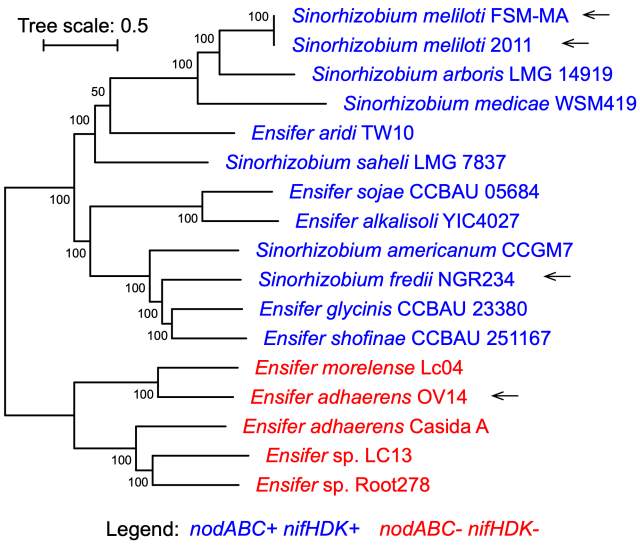


Figure 1. Phylogeny of the *genus Ensifer*. An unrooted maximum likelihood phylogeny of 17 representative *Ensifer* strains. The phylogeny represents the bootstrap best tree following 100 bootstrap replicates, prepared on the basis of the concatenated nucleotide alignments of 1566 core genes. Values represent the bootstrap support. N_2-fixing legume symbionts were identified by the presence of the symbiotic genes *nodABC* and *nifHDK*. They are indicated in blue, while red denotes non-symbiotic strains. The four wild-type strains used in this study are indicated with arrows.

Table 1. Methylated motifs identified in this study.

Motif*	Type†	Count¥	Frequency (motifs/kb)
E. meliloti 2011			
G**A**NTC CTN**A**G	m6A	11,169	1.67
RCG**C**CTC YGCGGAG	m4C	3,943	0.59
CGC**A**(N5)GTG GCGT(N5)C**A**C	m6A	1,085	0.16
E. meliloti FSM-MA			
G**A**NTC CTN**A**G	m6A	11,215	1.67
TCG**A**(N8)TCGA AGCT(N8)**A**GCT	m6A	2,612	0.39
E. fredii NGR234			
G**A**NTC CTN**A**G	m6A	11,111	1.61
CAG**A**(N7)GTTG GTCT(N7)C**A**AC	m6A	188	0.03
E. adhaerens OV14			
G**A**NTC CTN**A**G	m6A	8,475	1.10
WNCCG**A**TG WNGGCTAC	m6A	4,596	0.60

* The methylated nucleotides are indicated in boldface font.

† Indicates whether the modification is a N^6-methyladenoside (m6A) or N^4-methylcytosine (m4C).

¥ The total times the motif appears in the genome, regardless of methylation status.

137 medium with succinate as the carbon source. A total of six methylated motifs were identified, of
138 which five were m6A modifications and one was a m4C modification (**Table 1**). Five of the six
139 motifs were methylated specifically in one strain. Only the GANTC motif, recognized by the
140 highly conserved cell cycle-regulated CcrM methyltransferase (13, 14), was methylated in all four
141 strains. To further examine the conservation, or lack thereof, of DNA modification across the
142 genus *Ensifer*, we examined the distribution of methyltransferases in the model species *E. meliloti*.
143 Based on gene annotations, we identified 24 genes encoding putative MTases in a previous
144 pangenome analysis of 20 *E. meliloti* strains (40). Of these 24 genes, only one (*ccrM*) was found
145 in all 20 strains, while four were found in two strains and 19 were found in a single strain. These
146 results suggest that DNA methylation is unlikely to play a biologically significant role in the genus
147 *Ensifer* aside from cell cycle control via CcrM-mediated methylation, and phage defence.
148 In support of most DNA methylation not having a regulatory function in the genus *Ensifer*,
149 no global effect of carbon source (sucrose [glycolytic] vs. succinate [gluconeogenic]) was
150 observed on the DNA methylation pattern of *E. meliloti* Rm2011 (**Figure S2**), nor were any motifs
151 enriched in the promoter regions of genes previously shown to be differentially expressed when
152 grown with glucose vs. succinate (41). Similarly, no global differences in DNA methylation were
153 detected between *E. meliloti* Rm2011 and RmP3496, a Rm2011 derivative lacking the pSymA and
154 pSymB replicons that together account for 45% of the genome content of *E. meliloti* (42) (**Figure**
155 **S3**).
156
157 **Cell cycle regulation by the CcrM methyltransferase**
158 A progressive increase in the extent of methylation (herein defined as the estimated fraction of
159 reads mapping to a motif that were methylated) of GANTC sites was observed from the *ori* to *ter*
160 regions of the chromosomes of all four strains during mid-exponential growth (**Figures 2, S4-S6**).
161 There was a local drop in GANTC methylation around the 1.5 Mb mark in the *E. meliloti* FSM-
162 MA chromosome (**Figure S4**); however, this was seen in only two of three replicates and
163 corresponded to a region of high sequencing depth (**Figure S7**), suggesting the result is a
164 sequencing artefact. In contrast to exponential phase cultures, GANTC sites displayed near full
165 methylation (averaging ~ 95%) across the genome during early stationary phase, while all other
166 motifs displayed near full methylation (averaging 95-99%) across the genome regardless of growth
167 state. The observed pattern of GANTC methylation indicates a progressive switch from full to
168 hemi-methylated states as DNA replication proceeds (model provided as **Figure S8**), confirming
169 that the CcrM methyltransferase of the family *Rhizobiaceae* is cell cycle regulated as demonstrated
170 in *C. crescentus* (13, 17, 43). Strikingly, the GANTC methylation pattern differed across replicons
171 within each genome. For example, in *E. meliloti* Rm2011 the extent of GANTC methylation
172 ranged from 0.80 to 0.98 for pSymB and 0.80 to 0.96 for pSymA, while for the chromosome the
173 range was from 0.71 to 0.99 (**Figure 2**). This result suggests that replication of each replicon is
174 asynchronous, with replication of the secondary replicons being initiated later in the cell cycle than
175 that of the chromosome (model provided as **Figure S8**).
176

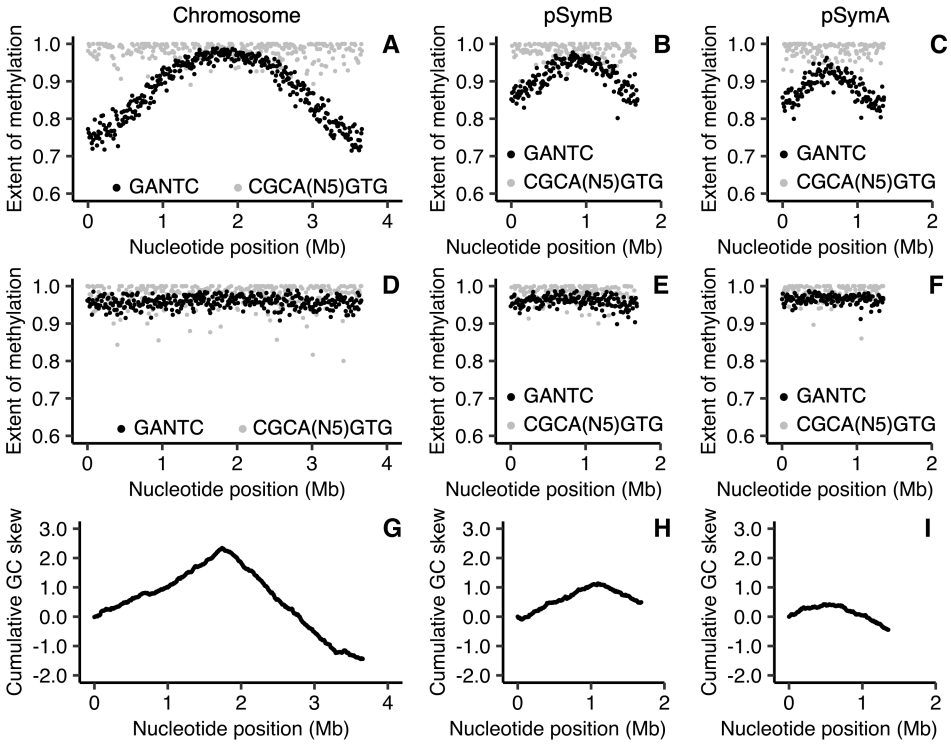


Figure 2. Genome-wide DNA methylation of *E. meliloti* Rm2011. (**A-F**) The extent of methylation is shown, using a 10 kb sliding window, of GANTC sites (black) and CGCA(N_5)GTG sites (grey) across the chromosome (**A**,**D**), pSymB (**B**,**E**), and pSymA (**C**,**F**) replicons of exponential phase (**A-C**) or early stationary phase (**D-F**) *E. meliloti* Rm2011. Averages from three biological replicates are shown. (**G-I**) Cumulative GC skews, shown using a 10 kb sliding window, across the *E. meliloti* Rm2011 chromosome (**G**), pSymB (**H**), and pSymA (**I**) replicons.

A previous study identified 462 cell cycle-regulated genes in *E. meliloti* through RNA-sequencing of synchronized cell populations (44), which were classified into six groups based on the timing of their expression. We identified 111 cell cycle-regulated genes, belonging to 78 transcripts, that contained at least one GANTC in the predicted promoter regions (defined as the 125 bp upstream of the transcript; **Dataset S1**), and distribution of these 111 genes across the six cell cycle gene expression groups was unbiased (44). As these 111 genes are both cell cycle regulated and contain a GANTC site, they represent an initial candidate CcrM regulon in *E. meliloti* although further work is required to validate the CcrM regulon.

We found it striking that the *E. adhaerens* OV14 genomes had 2,636 to 2,740 fewer GANTC sites than the other three strains, despite having the largest genome size. Normalized by genome length, there are 1.10 GANTC sites per kb in the *E. adhaerens* OV14 genome (**Table 1**), which is similar to the 1.12 GANTC sites per kb in *C. crescentus* NA1000. In contrast, the three legume symbionts contained more than 1.60 GANTC sites per kb across their genomes (**Table 1**). This result prompted us to examine the frequency and distribution of GANTC sites across 157 *Ensifer* genomes. As defined previously (45), the genus *Ensifer* can be broadly sub-divided into

192 two monophyletic clades; the "symbiotic" clade (113 strains) in which nearly all strains are legume
193 symbionts, and the "non-symbiotic" clade (44 strains) in which nearly all strains are non-
194 symbionts. Consistent with previous results (18), GANTC sites occurred less frequently in all
195 genomes (0.90 to 1.83 GANTC sites per kb) than expected in a random sequence of nucleotides
196 (~ 3.5 GANTC sites per kb). Moreover, GANTC sites were ~ 2-fold more common in intergenic
197 regions than in coding regions (**Figure 3**). Strikingly, there was a strong and statistically significant
198 difference (p-value < 1 x 10^{-10}; two-sample t-test) in the frequency of GANTC sites across the
199 genomes of strains belonging to the symbiotic clade compared to strains of the non-symbiotic
200 clades (**Figure 3**), with an overall average of 1.70 and 1.06 GANTC sites per kb in the symbiotic
201 and non-symbiotic clades, respectively. The difference in the frequency of GANTC sites between
202 the two clades could not be explained by differences in the GC content of these organisms, as both
203 clades had an average GC content of 61.9% (**Figure S9**), suggesting that the difference reflects
204 underlying differences in the evolution, and possibly the biology, of these two clades.
205

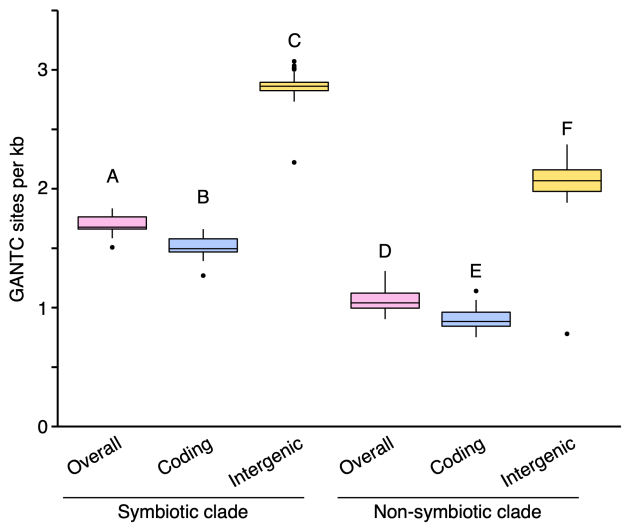


Figure 3. GANTC frequency in the genus *Ensifer*. Box plots summarizing the frequency of GANTC sites (as GANTC sites per kb) in 157 *Ensifer* strains is shown. The monophyletic "symbiotic" and "non-symbiotic" clades as defined previously (45), are represented by 111 and 44 genomes respectively. The densities of GANTC sites across the entire genome (pink), within coding regions (blue), and within intergenic regions (yellow) are shown. Statistically different values (p < 0.05) are denoted by uppercase letters as determined by a one-way ANOVA followed by a Tukey's HSD post hoc test.

206
207 **Contributions of DNA methylation to bacteroid differentiation**
208 The only previously published study to examine the role of rhizobium DNA methylation during
209 symbiosis using SMRT sequencing did so in a symbiosis where the bacteria do not undergo
210 terminal differentiation (38, 39). To evaluate whether DNA methylation potentially contributes to
211 regulation of terminal differentiation, we determined the DNA methylation patterns of *E. meliloti*
212 Rm2011 and FSM-MA bacteroids purified from whole *Medicago sativa* nodules. *E. meliloti* FSM-
213 MA bacteroids were additionally purified from *Medicago truncatula* nodules to determine whether
214 the host plant influences bacteroid DNA methylation. *E. meliloti* Rm2011 bacteroids were not
215 isolated from *M. truncatula* nodules as, unlike FSM-MA, Rm2011 forms a poor symbiosis with
216 *M. truncatula* (46, 47).
217 Moreover, we exploited the spatially distinct developmental zones that are present in

218 indeterminate nodules (48), like those formed by *M. sativa* and *M. truncatula*. At the tip of these
219 nodules, a bacteria-free meristem is present, responsible for the continuous growth of the nodule.
220 Immediately below is the infection and differentiation zone II where nodule cells become infected
221 and bacteria differentiate into the large, polyploid bacteroids. The tip and zone II of nodules
222 appears white. Adjacent to the white zone II is the easily recognizable pinkish zone III (due to the
223 presence of the oxygen-carrying leghemoglobin) where mature bacteroids fix nitrogen. This
224 nodule tissue organization provided an opportunity to examine how DNA methylation patterns
225 may differ between differentiating and differentiated bacteroids. To this end, *E. meliloti* Rm2011
226 and FSM-MA bacteroids were isolated from nodules hand-sectioned at the white-pink border;
227 bacteroids isolated from the white sections represent the infecting and differentiating bacteroids
228 (zone II) while those isolated from the pink sections represent the mature, hence terminally
229 differentiated, N_2-fixing bacteroids (zone III).
230 Fluorescence microscopy and flow cytometry confirmed that nodule sectioning resulted in
231 the isolation of distinct bacteroid populations (**Figures 4, S10-S12**). Nearly all of the bacteroids
232 isolated from whole-nodule samples and zone III samples were enlarged and polyploid, and most
233 were positive for propidium iodide (PI) staining as expected for terminally differentiated
234 bacteroids (26). In contrast, bacteroids of the zone II samples contained a mix of cell types differing
235 in their size, ploidy level, and PI staining. These data confirmed that the whole-nodule samples
236 and zone III samples consisted predominately of mature N_2-fixing bacteroids, whereas the zone II
237 samples contained a mix of cells at various stages of bacteroid differentiation.
238 With the exception of the GANTC sites (described below), no global difference was
239 observed in the methylation patterns of bacteroids versus free-living cells (**Figures 5, 6, S13-S16**).
240 Although a lower percentage of each motif was detected as methylated in the bacteroid samples
241 compared to free-living samples, this was correlated with lower sequencing depth (**Table S1**) and
242 thus unlikely to be biologically meaningful. Unlike *B. diazoefficiens* USDA110 (38), we did not
243 identify motifs that were methylated specifically in the *E. meliloti* bacteroids. In addition, we found
244 little evidence for any of the methylated motifs being enriched in the promoter regions of *E.*
245 *meliloti* Rm2011 genes up-regulated or down-regulated in bacteroids relative to free-living cells,
246 as identified in published transcriptomic data for *E. meliloti* Rm1021 (23), a near-isogenic relative
247 of strain Rm2011 also derived from the nodule isolate SU47. These data suggest that most DNA
248 methylation is unlikely to be a significant factor in directly regulating gene expression in *E. meliloti*
249 bacteroids.
250
251 **CcrM methyltransferase activity is dysregulated during terminal differentiation**
252 Bacteroid development involves cell enlargement and genome endoreduplication, indicative of a
253 cell cycle transition occurring during differentiation (26). Indeed, *ccrM* and other genes encoding
254 cell cycle regulators are strongly downregulated in terminally differentiated *E. meliloti* bacteroids
255 (24). We were therefore interested to examine whether GANTC methylation by the CcrM MTase
256 was disrupted in bacteroids. Our data revealed a surprising genome-wide pattern of GANTC
257 methylation in *E. meliloti* Rm2011 and FSM-MA bacteroids, which differed from free-living cells

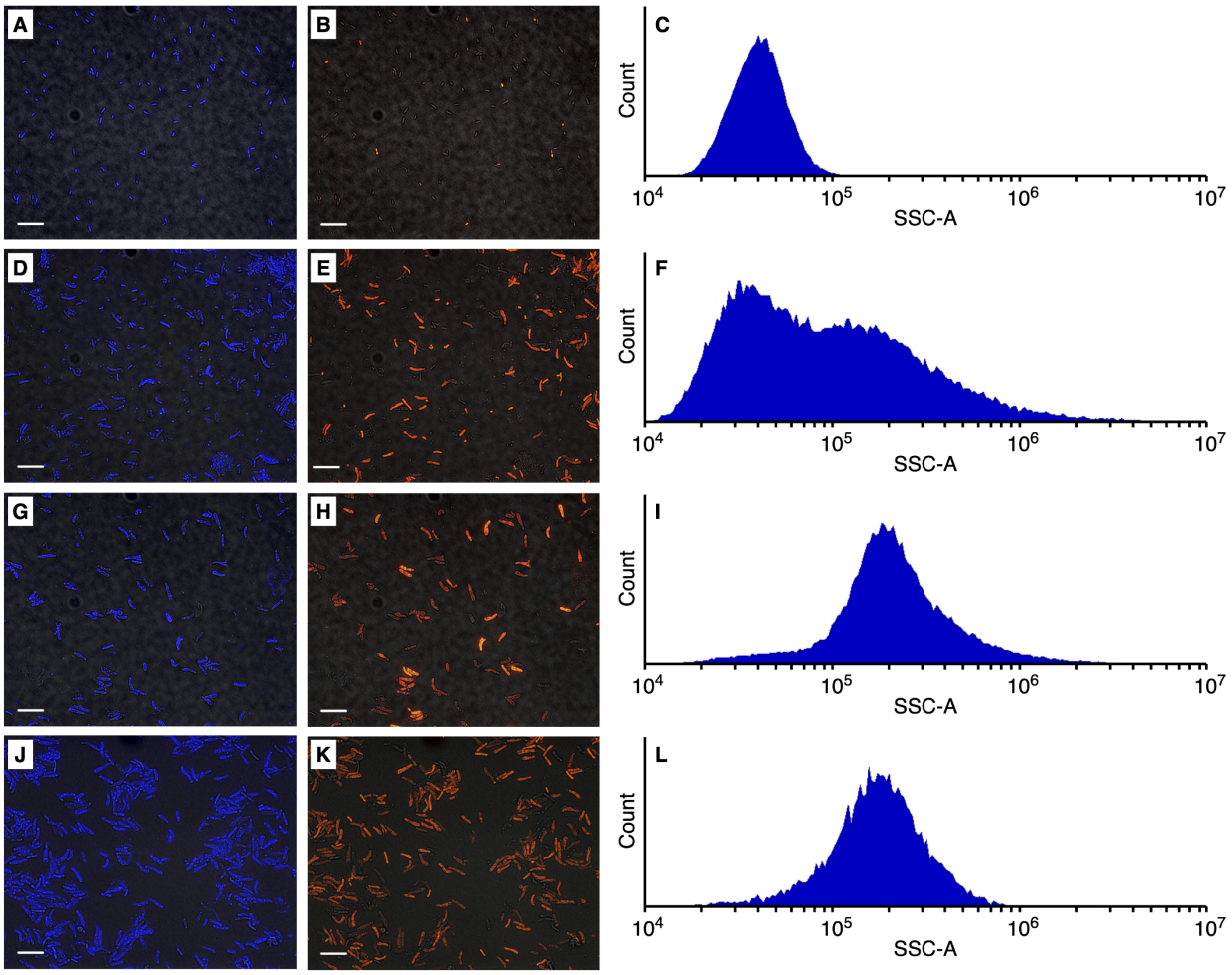


Figure 4. Morphology of *E. meliloti* Rm2011 cell populations. Free-living *E. meliloti* Rm2011 cells (**A-C**), and *E. meliloti* Rm2011 cells purified from *M. sativa* zone II nodule sections (**D-F**), *M. sativa* zone III nodule sections (**G-I**), or *M. sativa* whole nodules (**J-L**). Ploidy data is provided in Figure S9. Micrographs show *E. meliloti* Rm2011 cell populations stained with the DNA binding dyes DAPI (blue) and PI (red). The scale bar represents 10 nm. (**A**,**D**,**G**,**J**) DAPI fluorescence overlaid with DIC (differential interference contrast) images. (**B**,**E**,**H**,**K**) PI fluorescence overlaid with DIC images. (**C**,**F**,**I**,**L**) Histograms summarizing the distribution of flow cytometry side scattering values, providing an estimation of cell morphology, of heat-killed *E. meliloti* Rm2011 populations. Graphs are based on 50,000 cells. Data for *E. meliloti* FSM-MA is shown in Figure S10.

in either the exponential or stationary phases of growth (**Figures 5, 6, S13-S16**). The majority of GANTC sites had moderate to high levels of methylation in zone II, zone III, or whole nodule samples, averaging 0.71 to 0.95 across each replicon (**Figures 5, 6, Tables S2, S3**). Moreover, a progressive decrease in the extent of chromosomal methylation of the GANTC sites was observed from the *ori* to *ter* in nearly all bacteroid samples, which differs from the patterns seen in exponential (progressive increase from *ori* to *ter*) and stationary (consistent methylation) phase cells. The exception was the *E. meliloti* FSM-MA zone II bacteroid sample, which displays a

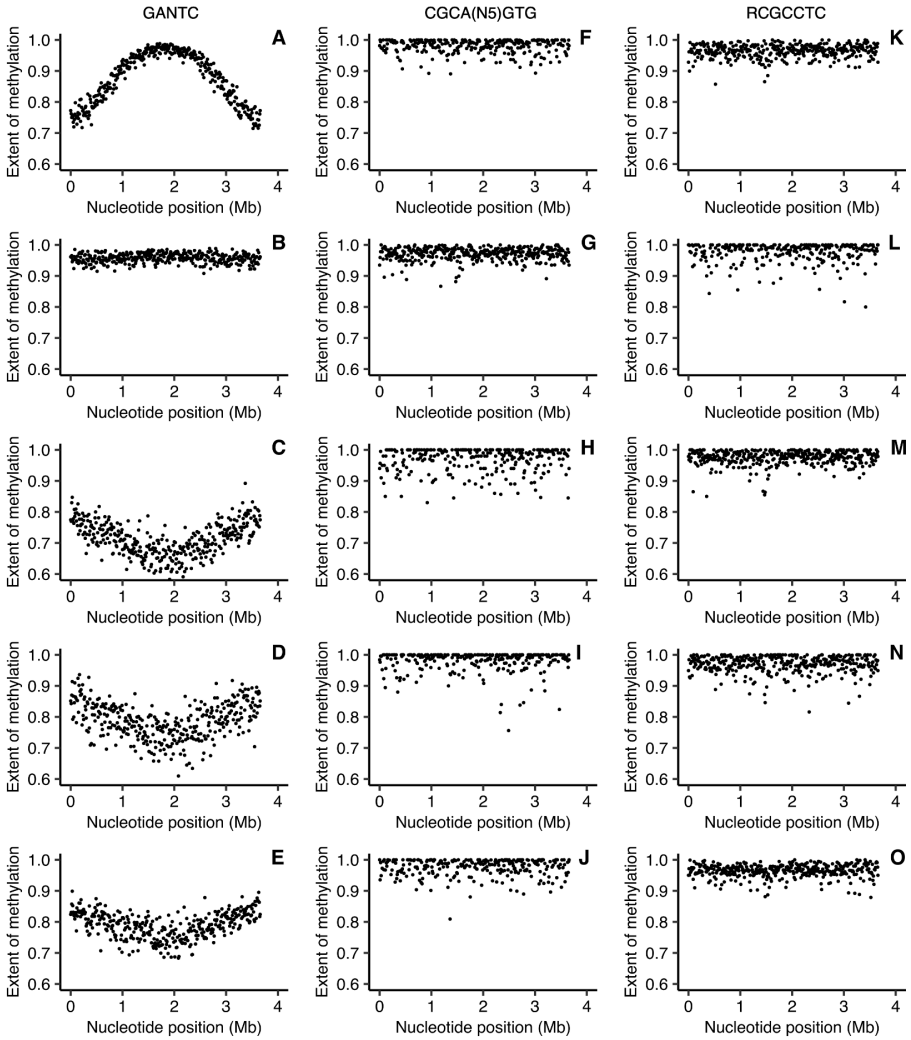


Figure 5. Chromosome-wide DNA methylation of *E. meliloti* Rm2011 bacteroids. The extent of methylation of (**A**-**E**) GANTC, (**F**-**J**) CGCA(N_5)GTG, and (**K**-**O**) RCGCCTC motifs across the *E. meliloti* Rm2011 chromosome is shown using a 10 kb sliding window. Averages from three biological replicates are shown for free-living and whole nodule samples; data represents one replicate for the zone II and zone III nodule sections. (**A**,**F**,**K**) Free-living cells harvested in mid-exponential phase. (**B**,**G**,**L**) Free-living cells harvested in early stationary phase. (**C**,**H**,**M**) Bacteroids isolated from *M. sativa* zone II nodule sections. (**D**,**I**,**N**) Bacteroids isolated from *M. sativa* zone III nodule sections. (**E**,**J**,**O**) Bacteroids isolated from *M. sativa* whole nodule samples. Data for pSymB and pSymA are shown in Figures S12 and S13.

consistently high level of GANTC methylation across the genome (**Figure 6C**). This pattern, which is different from those in exponential phase cells as well as mature bacteroids, could correspond to the methylation status of an early stage of bacteroid differentiation. Supporting this hypothesis, the flow cytometry data was consistent with the *E. meliloti* FSM-MA zone II bacteroids having a lower average DNA content than the *E. meliloti* Rm2011 zone II bacteroids (**Figure S17**), suggesting that the former sample represents an earlier stage of differentiation than

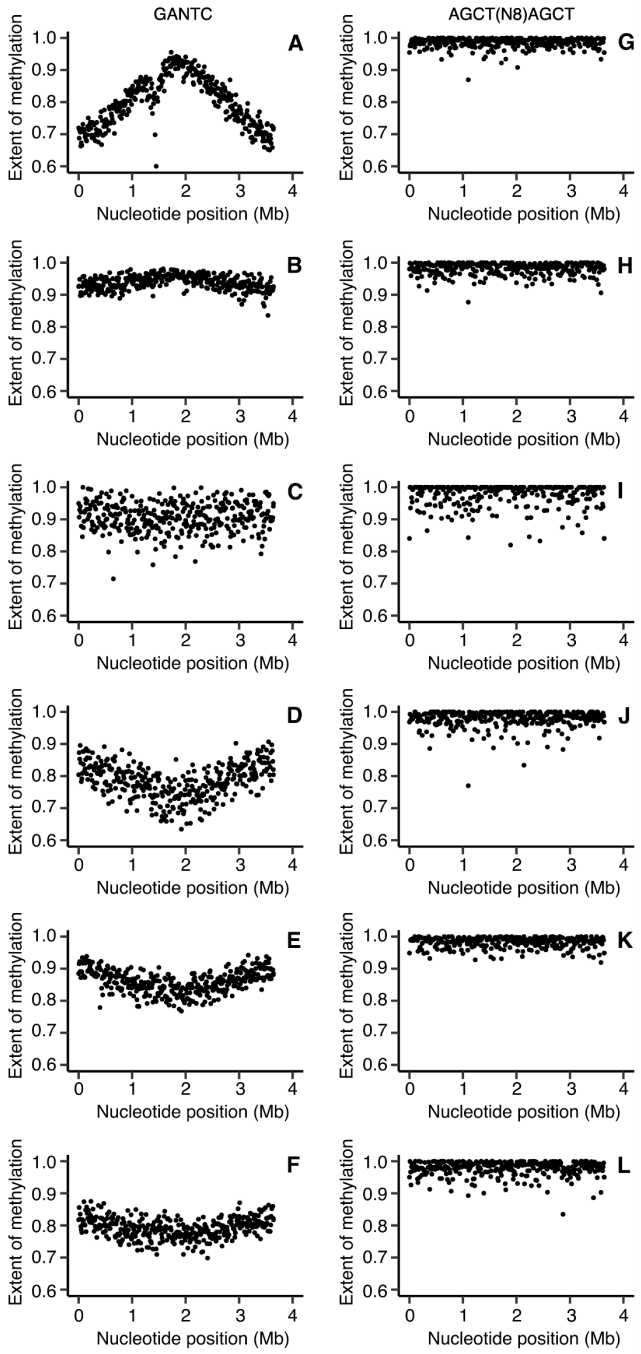


Figure 6. Chromosome-wide DNA methylation of *E. meliloti* FSM-MA bacteroids. The extent of methylation of (**A-F**) GANTC and (**G-L**) AGCT(N_8)AGCT motifs across the *E. meliloti* FSM-MA chromosome is shown using a 10 kb sliding window. Averages from three biological replicates are shown for free-living and whole nodule samples; data represents one replicate for the zone II and zone III nodule sections. (**A**,**G**) Free-living cells harvested in mid-exponential phase. (**B**,**H**) Free-living cells harvested in early stationary phase. (**C**,**I**) Bacteroids isolated from *M. sativa* zone II nodule sections. (**D,J**) Bacteroids isolated from *M. sativa* zone III nodule sections. (**E**,**K**) Bacteroids isolated from *M. sativa* whole nodule samples. (**F**,**L**) Bacteroids isolated from *M. truncatula* whole nodule samples. Data for pSymB and pSymA are shown in Figures S14 and S15.

the latter sample. In contrast to GANTC, the extent of methylation of the second m6A modified motif in each genome was consistently high, irrespective of condition or replicon (**Figures 5, 6, Tables S2, S3**). Similarly, sequencing depth was consistent across the length of each replicon (**Figure S7**). These observations indicate that the changes in GANTC methylation patterns are biologically meaningful and not simply a sequencing artefact, and they are suggestive of a dysregulation of CcrM during terminal bacteroid differentiation.

Chromosome, pSymB, and pSymA sequencing depth are unequal in *E. meliloti* bacteroids

We noticed that in each bacteroid sample, the average extent of GANTC methylation for the chromosomes of the two strains were lower (by 0.04 to 0.13) than that of pSymA or pSymB, and unlike the chromosome, the extent of GANTC methylation was relatively constant across pSymA and pSymB (**Figures S13-S16** compared to **Figures 5 and 6**). These results suggest that, unlike in free-living cells, replication of the three replicons is not well coordinated during terminal differentiation. In agreement with this hypothesis, the mean sequencing depth across pSymA and pSymB was on average ~ 33% lower than that of the chromosome in all replicates of the *E. meliloti* whole-nodule bacteroid samples (**Table 2**). Assuming sequencing depth is correlated with copy number, this observation suggests that *E. meliloti* bacteroids carry only two copies each of pSymA and pSymB per three copies of the chromosome.

Table 2. Relative sequencing depth of each *E. meliloti* replicon.

		Relative mean sequencing depth *		
Strain	Sample	Chromosome	pSymA	pSymB
Rm2011	Mid-exponential	1.00 ± 0.00	1.00 ± 0.03	0.98 ± 0.03
Rm2011	Stationary	1.00 ± 0.00	1.06 ± 0.02	1.02 ± 0.03
Rm2011	*M. sativa* bacteroids	1.00 ± 0.00	0.64 ± 0.02	0.62 ± 0.02
FSM-MA	Mid-exponential	1.00 ± 0.00	1.02 ± 0.03	0.87 ± 0.05
FSM-MA	Stationary	1.00 ± 0.00	1.10 ± 0.00	1.01 ± 0.03
FSM-MA	*M. sativa* bacteroids	1.00 ± 0.00	0.79 ± 0.04	0.71 ± 0.02
FSM-MA	*M. truncatula* bacteroids	1.00 ± 0.00	0.66 ± 0.02	0.59 ± 0.01

* Sequencing depth is presented as a fraction, relative to the sequencing depth of the chromosome in the same sample. Values are the means of triplicate samples ± standard deviation.

DISCUSSION

In this study, we examined the genome-wide DNA methylation patterns in the free-living cells of four *Ensifer* strains, and in bacteroids of two *E. meliloti* strains, and detected a total of six methylated motifs. We were able to predict cognate MTases for most of the motifs based on genome annotations, the exception being the WNCCGATG motif of *E. adhaerens* OV14. The CGCA(N_5)GTG motif of *E. meliloti* Rm2011 is presumably methylated by Smc02296 (HsdM), a predicted m6A MTase belonging to the HsdRSM type I R-M system that is known to be functional and reduce transformation efficiency (49, 50). The RCGCCTC motif of *E. meliloti* Rm2011 is possibly methylated by Smc03763, a predicted cysteine-specific MTase located upstream of the gene *vsr* that putatively encodes a very short patch repair protein. Neither of these proteins are found in the other three strains examined here. The motifs TCGA(N_8)TCGA of *E. meliloti* FSM-MA and CAGA(N_7)GTTG of *E. fredii* NGR234 are likely methylated by SMB554_16155 and NGR_c01340, respectively, which are 88% identical at the amino acid level. Homologs of these two proteins are not found in the other two strains.

With the exception of GANTC, each methylated motif was detected as methylated only in

310 a single strain. Moreover, MTases are, with the exception of CcrM, not conserved among *E.*
311 *meliloti* strains. The lack of conservation suggests that most DNA methylation does not have a
312 major regulatory role in the genus *Ensifer*, aside from its role in cell cycle regulation. Supporting
313 this conclusion, no motif was enriched in the promoter regions of symbiosis, carbon source, or cell
314 cycle-regulated genes, and we did not detect any motifs that were methylated specifically in
315 bacteroids. However, we cannot rule out that one or more methylated motifs may influence specific
316 gene expression during free-living growth, differentiation, or N_2-fixation through extended motifs
317 or proximity to other promoter elements, similar to the interplay between CcrM and GcrA during
318 cell cycle regulation in *C. crescentus* (20, 51).
319 As previously published studies have provided evidence for a cell cycle transition
320 occurring during terminal bacteroid differentiation (26), we were particularly interested in CcrM,
321 a cell cycle-regulated MTase that is broadly conserved in the α-Proteobacteria, and its cognate
322 DNA motif, GANTC. By identifying GANTC sites in the promoter regions of a previously
323 identified set of 462 cell cycle-regulated genes (44), we identified a candidate CcrM regulon in *E.*
324 *meliloti* consisting of 111 genes. However, further studies are required to better delineate the CcrM
325 regulon in *E. meliloti* as the presence of a GANTC site is not diagnostic of CcrM regulation;
326 GANTC sites were found in the promoter regions of 904 transcripts that did not display cell cycle
327 regulation, and the promoter regions of cell cycle regulated genes were not enriched in GANTC
328 sites relative to the whole *E. meliloti* genome. Studies in *C. crescentus* suggest that the impact of
329 the fully or hemi-methylated status of GANTC sites on gene expression is mediated, at least in
330 part, through modulating the activity of the transcription regulator GcrA (20, 51). However, not
331 all promoter sites containing a GANTC motif are regulated by GcrA in *C. crescentus*, with the
332 relationship dependent on an extended YGAKTCK motif and the precise position of this motif
333 relative to other promoter elements (51). Likely, CcrM-mediated gene regulation in the genus
334 *Ensifer* is similarly dependent upon additional sequence elements beyond the GANTC motif.
335 Consistent with past observations (18), GANTC sites were under-represented in the
336 genomes of 157 *Ensifer* strains, particularly within coding regions. More surprising, however, was
337 the strong difference in the frequency of GANTC sites between the previously defined (45)
338 symbiotic and non-symbiotic clades in the genus *Ensifer*, with the frequency of GANTC sites
339 being ~ 60% higher in the symbiotic clade. As methylation of the GANTC motif by CcrM is known
340 to influence gene expression in *C. crescentus* (18), our observations suggest that CcrM has a
341 greater impact on modulating gene expression in the symbiotic clade compared to the non-
342 symbiotic clade. Although further work is required to understand the biological significance of the
343 greater frequency of GANTC sites in the symbiotic clade, it is tempting to speculate it is associated
344 with legume symbiosis.
345 In support of a link between CcrM activity and terminal bacteroid differentiation, our data
346 is consistent with CcrM activity influencing terminal bacteroid differentiation. The overall
347 moderate to high rates of GANTC methylation in all *E. meliloti* bacteroid samples, coupled with
348 the lack of a chromosome-wide pattern in the *E. meliloti* FSM-MA zone II sample, suggests that
349 CcrM remains constitutively active throughout most of terminal differentiation. Considering that

350 over-expression of CcrM can give rise to bacteroid-like morphology in free-living cells (14), we
351 hypothesize that constitutive CcrM MTase activity is one (of potentially multiple) factor(s) driving
352 polyploidization of bacteroids (**Figure S8**). CcrM activity is confined to a short window in the cell
353 cycle since the *ccrM* gene is expressed in the late phase of genome replication (44) and the CcrM
354 protein is degraded by the Lon protease prior to cell division (52). Thus, constitutive CcrM MTase
355 activity in differentiating bacteroids could be obtained through an aberrant expression of the gene
356 or alternatively a lack of proteolytic degradation of the CcrM protein. In agreement with the latter
357 possibility, Lon protease was identified as a target of the NCR247 peptide (34). It is tempting to
358 speculate that NCR peptides like NCR247 inhibit Lon protease activity, thereby stabilizing CcrM
359 and trigger the bacteroid differentiation. However, the CcrM MTase does not appear to remain
360 active in fully differentiated bacteroids, with the lower GANTC methylation near the chromosomal
361 *ter* regions suggesting that loss of CcrM MTase activity occurs slightly prior to completion of
362 genome endoreduplication (model provided as **Figure S8**). These hypotheses are consistent with
363 *M. truncatula – E. meliloti* nodule zone-specific RNA-sequencing data (24), which showed that
364 *ccrM* expression in the root distal portion of zone II is ~ 2-fold higher than in the root proximal
365 portion of zone II, and ~ 10-fold higher than in zone III.
366 Our analyses also provide insight into the genome replication dynamics of *E. meliloti*
367 during free-living growth and terminal bacteroid differentiation. Interestingly, our data are
368 consistent with a loss of coordination of replication of the three replicons during terminal bacteroid
369 differentiation, leading to unequal copy numbers with two copies of pSymA and pSymB per three
370 copies of the chromosome in bacteroids. This contrasts with free-living cells, where copy number
371 of the three replicons was approximately equal based on average sequencing depth. In contrast to
372 our results, a previous comparison of the relative abundance of the three replicons in free-living
373 *E. meliloti* Rm1021 versus bacteroids, no changes were detected using comparative genome
374 hybridization with microarrays (26). We believe that the difference between our present data and
375 the previous analysis is due to the subtlety of the differences and the lower sensitivity of the
376 microarray hybridization method compared to high throughput sequencing.
377 We also observed that the extent of GANTC methylation at the *ori* of pSymA and pSymB
378 is higher than at the *ori* of the chromosome, while the *ter* of the pSymA and pSymB has a slightly
379 lower extent of GANTC methylation than the *ter* of the chromosome. As GANTC methylation
380 occurs at a fixed stage of the cell cycle corresponding to the end of chromosome replication, our
381 observations indicate that pSymA and pSymB replication is initiated later in the cell cycle than
382 initiation of chromosome replication, while their replication terminates slightly before completion
383 of chromosomal replication and the activation of CcrM (model provided as **Figure S8**). These
384 results provide additional support for the notion of spatiotemporal regulation of DNA replication
385 and partitioning in the multipartite *E. meliloti* genome as proposed previously (44, 53). Similarly,
386 replication of chromosome II of *Vibrio cholerae* is delayed relative to chromosome I, leading to
387 the replication of these two replicons terminating at approximately the same time (54). Thus, co-
388 ordinating the timing or replication termination may be a general feature of multipartite genomes.
389

390 **MATERIALS AND METHODS**

391 **Experimental design**

392 The overall experimental design is summarized in **Figure S1**. Genomic DNA was isolated from
393 four wild-type *Ensifer* strains to explore how DNA methylation varies across this genus; to allow
394 direct comparison, the four strains were grown to mid-exponential phase in minimal medium with
395 succinate as a carbon source. To investigate how DNA methylation patterns differ between
396 actively dividing and non-dividing cells, genomic DNA was isolated from *E. meliloti* Rm2011
397 growth to either mid-exponential phase or stationary phase. Genomic DNA was isolated from *E.*
398 *meliloti* Rm2011 grown to mid-exponential phase with either a glycolytic (sucrose) or
399 gluconeogenic (succinate) carbon source to investigate whether DNA methylation may play a role
400 in regulating carbon metabolism. Furthermore, a *E. meliloti* Rm2011 derivative lacking the pSymA
401 and pSymB replicons (named RmP3496), as well as a complemented strain in which pSymA and
402 pSymB were re-introduced (named RmP3499), were studied to gain insight into whether these
403 replicons contribute to DNA methylation patterns in *E. meliloti* Rm2011; these strains were grown
404 with sucrose (instead of succinate) as RmP3496 lacks the succinate transporter.

405 In addition to the free-living samples, *E. meliloti* bacteroid samples purified from legume
406 nodules were collected to investigate changes in DNA methylation during bacteroid differentiation
407 and nitrogen fixation. To do so, *E. meliloti* Rm2011 and *E. meliloti* FSM-MA bacteroids were
408 isolated from *M. sativa* whole nodules. *E. meliloti* FSM-MA bacteroids were additionally purified
409 from *M. truncatula* whole nodules to examine the impact of host plant on bacteroid DNA
410 methylation patterns. *E. meliloti* Rm2011 bacteroids were only isolated from *M. sativa* nodules as,
411 unlike FSM-MA, Rm2011 forms a poor symbiosis with *M. truncatula* (46, 47). Moreover, *E.*
412 *meliloti* Rm2011 and *E. meliloti* FSM-MA bacteroids were isolated from *M. sativa* nodule sections
413 (sectioned at the white – pink border to separate the root distal infection and differentiation zone
414 II [white] from the root proximal nitrogen-fixing zone III [pink]) to facilitate an analysis of how
415 DNA methylation patters differ between differentiating bacteroids and fully differentiated and
416 nitrogen-fixing bacteroids.

417

418 **Bacterial strains and growth conditions**

419 Bacterial strains used in this work are listed in **Table S4**. All strains were routinely grown on TY
420 with 2 μM $CoCl_2$ as it was required for *E. meliloti* RmP3496 (42). The MM9 minimal medium
421 (42) consisted of the following: 40 mM MOPS, 20 mM KOH, 19.2 mM NH_4Cl, 85.6 mM NaCl, 2
422 mM KH_2PO_4, 1 mM $MgSO_4$, 0.25 mM $CaCl_2$, 1 μg ml^{-1} biotin, 42 nM $CoCl_2$, 38 μM $FeCl_3$, 10
423 μM thiamine-HCl, and either 10 mM sucrose (MM9-sucrose) or 20 mM disodium succinate
424 (MM9-succinate). Prior to inoculation of plants with *E. meliloti*, the strains were grown in YEB
425 medium (55).

426

427 **DNA isolation from free-living cells**

428 Overnight cultures of all strains grown in MM9-succinate or MM9-sucrose media were diluted
429 into 10 mL of the same medium to a starting OD_{600nm} of 0.025 (0.05 for RmP3499) and incubated

430 overnight at 30˚C with shaking (130 rpm). The next day, cultures were diluted into 40 mL of the
431 same medium in 100 mL flasks to the OD_{600nm} values as listed in **Table S5**. To obtain mid-
432 exponential phase samples, cultures were harvested after 15.5-16 hours of growth at OD_{600nm}
433 values between 0.37 and 0.69 (**Table S5**). To obtain stationary phase samples, cultures were
434 harvested after 24 hours of growth at OD_{600nm} values of ~ 1.4. In all cases, cultures were streaked
435 on TY plates to check for contamination and then centrifuged (8,200 *g*, 10 minutes, 4°C); the full
436 40 mL was centrifuged for mid-exponential phase cultures, whereas only 15 mL was centrifuged
437 for stationary phase cultures. The majority of the supernatant was discarded, and the pellet
438 resuspended in the remaining volume, transferred to a 2 mL tube, centrifuged again (16,200 *g*,
439 room temperature, one minute), and the supernatant discarded. Three biological replicates, each
440 starting from a separate overnight culture, were performed. DNA was isolated using
441 phenol:chloroform extractions and ammonium acetate precipitations as described elsewhere (56),
442 and the DNA pellets (following RNase A treatment) were resuspended in 200 µL of 10 mM Tris-
443 HCl, pH 8.5.
444
445 **DNA isolation from bacteroids**
446 *M. sativa* cv. Gabès and *M. truncatula* cv. A17 were used for all experiments. Seeds were scarified,
447 surface sterilized, and germinated on Kalys agar as described previously (47). Fifty mL of
448 overnight cultures of *E. meliloti* Rm2011 or FSM-MA, grown in YEB, were centrifuged (4,000 *g*,
449 20 minutes, room temperature) and resuspended in ~ 1,200 mL of sterile, distilled water to obtain
450 a cell suspension at an OD_{600nm} of ~ 0.1. Germinated seedlings were immersed for one hour in the
451 appropriate rhizobial cell suspension, and then planted in a perlite:sand (2:1) mixture. Plants were
452 grown in a greenhouse for five to six weeks, with occasional watering with a 1 g L^{-1} nutrient
453 solution (PlantProd solution [N-P-K, 0-15-40; Fertil]).
454 For whole nodule samples, pink nodules were collected from 53-60 plants per replicate 34-
455 35 days post-inoculation. Nodules were collected from the roots and stored in tubes in liquid
456 nitrogen until collection was complete, at which point they were stored at -80°C until use. For
457 sectioned nodule samples, pink nodules were collected from 103 *M. sativa* plants for each of the
458 microsymbionts 35 to 40 days post-inoculation. Nodules were manually sectioned at the white-
459 pink border. Nodule sections were stored in tubes over dry ice or liquid nitrogen until collection
460 was complete, at which point they were stored at -80°C until use. Average plant shoot dry weights
461 for all samples are listed in **Table S6**. Bacteroids were isolated from the nodule samples using
462 Percoll gradient centrifugation as described elsewhere (26). The recovered bacteroids were
463 resuspended in 50-100 µL of Bacteroid Extraction Buffer (BEB; 125 mM KCl, 50 mM Na-
464 succinate, 50 mM TES, pH 7.0), and either used immediately for microscopy, flow cytometry, and
465 DNA isolation, or stored at -80°C until use.
466 Nucleic acids were initially purified from all bacteroid samples using Epicentre
467 MasterPure™ Complete DNA and RNA Purification Kits, following the protocol for DNA
468 isolation from cell samples. For sectioned nodule samples, pure DNA was isolated by using the
469 manufacturer's protocol for the complete removal of RNA. For whole nodule samples, the isolated

470 DNA was further purified by treating the nucleic acids samples with RNase A, after which pure
471 DNA was isolated by using phenol:chloroform extractions followed by ammonium acetate DNA
472 precipitations as described elsewhere (56). In all cases, the final DNA pellets were resuspended in
473 200 µL of 10 mM Tris-HCl, pH 8.5. Three biological replicates were performed for bacteroids
474 isolated from whole nodules, whereas only one replicate was performed for bacteroids isolated
475 from sectioned nodules due to difficulties in obtaining sufficient DNA yields.
476
477 **DNA sequencing, modification detection, and motif analysis**
478 DNA sequencing was performed at the U.S. Department of Energy Joint Genome Institute (JGI)
479 or in-house at the University of Florence (the stationary phase samples) using Pacific Biosciences
480 (PacBio) sequencing technology (57). Genomic DNA was sheared to 3 kb using a Covaris LS220
481 (Covaris Inc., Woburn, MA, USA) or 15 kbp (for stationary phase samples) using g-TUBEs
482 (Covaris Inc., Woburn, MA, USA). Sheared DNA was treated with exonuclease to remove single-
483 stranded ends and DNA damage repair mix followed by end repair and ligation of barcoded blunt
484 adapters using SMRTbell Template Prep Kit 2.0 (PacBio, Menlo Park, CA, USA). Libraries were
485 purified with AMPure PB beads (Beckman Coulter, Brea, CA, USA) and eight libraries with
486 different barcodes were pooled at equimolar ratios and purified with AMPure PB beads. For most
487 samples, SMRTbell template libraries were prepared using a Sequel Binding Kit 3.0 (PacBio,
488 Menlo Park, CA, USA), and sequenced on a Sequel instrument using a v3 or v4 sequencing primer,
489 1M v3 SMRT cells, and Version 3.0 sequencing chemistry with 1x360 or 1x600 sequencing movie
490 run times. The exceptions were the *E. meliloti* Rm2011 zone II and *E. meliloti* FSM-MA zone III
491 bacteroid samples. For these samples, SMRTbell template libraries were prepared using a Sequel
492 II Binding Kit 2.0 (PacBio, Menlo Park, CA, USA), and then sequenced on a Sequel II instrument
493 using the tbd-sample dependent sequencing primer, 8M v1 SMRT cells, and Version 2.0
494 sequencing chemistry with 1x900 sequencing movie run times.
495 DNA modification detection and motif analysis were performed using the PacBio SMRT
496 Link software (PacBio, Menlo Park, CA, USA). Briefly, raw reads were filtered using SFilter to
497 remove short reads and reads derived from sequencing adapters. Filtered reads were aligned
498 against the appropriate reference genome (**Table S1**) using BLASR (58) and modified sites were
499 then identified through kinetic analysis of the aligned DNA sequence data (59); the number of
500 mapped bases per sample is provided in **Table S1**. Modified sites were then grouped into motifs
501 using MotifFinder. These motifs represent the recognition sequences of MTase genes active in the
502 genome (60). Downstream analyses were performed using custom Perl and R scripts.
503
504 **Flow cytometry**
505 Flow cytometry was performed as described previously (26). Freshly prepared bacteroid samples
506 were diluted in 200 µL of BEB, heat-treated for 10 minutes in a 70°C water bath, and then stained
507 with the DNA-binding dye diamidino-2-phenylindole (DAPI). Cell size and ploidy level of the
508 bacteroid samples were determined using flow cytometry with a Beckman Coulter CytoFLEX S
509 instrument. Measurements consisted of 50,000 cells. Data analysis was performed using the

510 CytExpert 2.2.0.97 software.
511
512 **Fluorescence microscopy**
513 One µL of each freshly prepared bacteroid sample was mixed with 1 µL of 50 µg mL^{-1} DAPI and
514 1 µL of 100 µg mL^{-1} propidium iodide (PI), which are both DNA binding dyes. Samples were
515 visualized at 100x magnification under oil immersion using a Nikon Eclipse 80*i* fluorescence
516 microscope with the NIS-Elements BR 4.00.01 software and a Digital Sight DS-U3 camera.
517
518 **Phylogenetic analysis**
519 The nucleotide fasta files of representative *Ensifer* species were downloaded from the National
520 Centre for Biotechnology Information (NCBI) Genome database. A core gene phylogeny was
521 constructed using a previously prepared pipeline (61) reliant on the use of Roary 3.11.3 (62),
522 Prokka 1.12-beta (63), PRANK 140110 (64), trimAl (65), and RAxML 8.2.9 (66). The phylogeny
523 was visualized with the iTol webserver (67). Identification of nodulation (*nodABC*) and
524 nitrogenase genes (*nifHDK*) was performed with a published pipeline (61) reliant on the use of
525 HMMER 3.1b2 (68), and the Pfam-A 31.0 (69) and TIGERFAM 15.0 (70) databases.
526
527 **Data availability**
528 Most sequencing data is available through the JGI Genome Portal (genome.jgi.doe.gov/portal/)
529 under Proposal 503835, as well as through NCBI (see **Table S1** for Biosample accessions). The
530 stationary phase data is available only through the NCBI (BioProject accession PRJNA706182;
531 see **Table S1** for Biosample accessions). All custom scripts to perform the analyses described in
532 this study are available through GitHub (github.com/diCenzo-
533 Lab/003_2021_Ensifer_DNA_methylation).
534
535 **ACKNOWLEDGMENTS**
536 We are grateful to E. Mullins (Teagasc, MTA2018233) for *E. adhaerens* OV14 and M. Bourge
537 (I2BC) for help with the flow cytometry experiments. Most sequencing work was performed at
538 the U.S. Department of Energy (DOE) Joint Genome Institute (JGI), through the CSP New
539 Investigator program (proposal: 503835). The work conducted by the U.S. Department of Energy
540 Joint Genome Institute, a DOE Office of Science User Facility, is supported under Contract No.
541 DE-AC02-05CH11231. This work was partially supported by the "MICRO4Legumes" grant
542 (Italian Ministry of Agriculture), by the grant "Dipartimento di Eccellenza 2018–2022" from the
543 Italian Ministry of Education, University and Research (MIUR), and by a Discovery Grant from
544 Natural Science and Engineering Research Council of Canada. The present work benefited from
545 the Imagerie-Gif core facility supported by the Agence Nationale de la Recherche (ANR-11-
546 EQPX-0029/Morphoscope, ANR-10-INBS-04/FranceBioImaging, ANR-11-IDEX-0003-
547 02/Saclay Plant Sciences). G.C.D was supported by a NSERC postdoctoral fellowship and a
548 European Molecular Biology Organization (EMBO) short term fellowship for part of this work.
549 L.C. was supported by the MICRO4Legumes grant (Italian Ministry of Agriculture). Q.N. was

supported by a PhD fellowship from the Paris-Saclay University. J.H.T.C. was supported by a NSERC Undergraduate Summer Research Award. P.M. and B.A. were supported by Saclay Plant Sciences (SPS) and grant ANR-17-CE20-0011 from the Agence Nationale de la Recherche.

REFERENCES

1. Greenberg MVC, Bourc'his D. 2019. The diverse roles of DNA methylation in mammalian development and disease. Nat Rev Mol Cell Biol 20:590–607.
2. Zhang H, Lang Z, Zhu J-K. 2018. Dynamics and function of DNA methylation in plants. Nat Rev Mol Cell Biol 19:489–506.
3. Tang Y, Gao X-D, Wang Y, Yuan B-F, Feng Y-Q. 2012. Widespread existence of cytosine methylation in yeast DNA measured by gas chromatography/mass spectrometry. Anal Chem 84:7249–7255.
4. Blow MJ, Clark TA, Daum CG, Deutschbauer AM, Fomenkov A, Fries R, Froula J, Kang DD, Malmstrom RR, Morgan RD, Pósfai J, Singh K, Visel A, Wetmore K, Zhao Z, Rubin EM, Korlach J, Pennacchio LA, Roberts RJ. 2016. The epigenomic landscape of prokaryotes. PLOS Genet 12:e1005854.
5. Sánchez-Romero MA, Cota I, Casadesús J. 2015. DNA methylation in bacteria: from the methyl group to the methylome. Curr Opin Microbiol 25:9–16.
6. Gopalakrishnan S, Van Emburgh BO, Robertson KD. 2008. DNA methylation in development and human disease. Mutat Res 647:30–38.
7. Vasu K, Nagaraja V. 2013. Diverse functions of restriction-modification systems in addition to cellular defense. Microbiol Mol Biol Rev 77:53–72.
8. Atack JM, Tan A, Bakaletz LO, Jennings MP, Seib KL. 2018. Phasevarions of bacterial pathogens: methylomics sheds new light on old enemies. Trends Microbiol 26:715–726.
9. Adhikari S, Curtis PD. 2016. DNA methyltransferases and epigenetic regulation in bacteria. FEMS Microbiol Rev 40:575–591.
10. Campbell JL, Kleckner N. 1990. *E. coli oriC* and the *dnaA* gene promoter are sequestered from *dam* methyltransferase following the passage of the chromosomal replication fork. Cell 62:967–979.
11. Kang S, Lee H, Han JS, Hwang DS. 1999. Interaction of SeqA and Dam methylase on the hemimethylated origin of *Escherichia coli* chromosomal DNA replication. J Biol Chem 274:11463–11468.
12. Lahue R, Au K, Modrich P. 1989. DNA mismatch correction in a defined system. Science 245:160–164.
13. Zweiger G, Marczynski G, Shapiro L. 1994. A *Caulobacter* DNA methyltransferase that functions only in the predivisional cell. J Mol Biol 235:472–485.
14. Wright R, Stephens C, Shapiro L. 1997. The CcrM DNA methyltransferase is widespread in the alpha subdivision of proteobacteria, and its essential functions are conserved in *Rhizobium meliloti* and *Caulobacter crescentus*. J Bacteriol 179:5869–5877.
15. Brilli M, Fondi M, Fani R, Mengoni A, Ferri L, Bazzicalupo M, Biondi EG. 2010. The

591 diversity and evolution of cell cycle regulation in alpha-proteobacteria: a comparative
592 genomic analysis. BMC Syst Biol 4:52.
593 16. Kahng LS, Shapiro L. 2001. The CcrM DNA methyltransferase of *Agrobacterium tumefaciens*
594 Is essential, and Its activity is cell cycle regulated. J Bacteriol 183:3065–3075.
595 17. Kozdon JB, Melfi MD, Luong K, Clark TA, Boitano M, Wang S, Zhou B, Gonzalez D, Collier
596 J, Turner SW, Korlach J, Shapiro L, McAdams HH. 2013. Global methylation state at base-
597 pair resolution of the *Caulobacter* genome throughout the cell cycle. Proc Natl Acad Sci USA
598 110:E4658–E4667.
599 18. Gonzalez D, Kozdon JB, McAdams HH, Shapiro L, Collier J. 2014. The functions of DNA
600 methylation by CcrM in *Caulobacter crescentus*: a global approach. Nucleic Acids Res
601 42:3720–3735.
602 19. Gonzalez D, Collier J. 2013. DNA methylation by CcrM activates the transcription of two
603 genes required for the division of *Caulobacter crescentus*: regulation of cell division by DNA
604 methylation. Mol Microbiol 88:203–218.
605 20. Fioravanti A, Fumeaux C, Mohapatra SS, Bompard C, Brilli M, Frandi A, Castric V, Villeret
606 V, Viollier PH, Biondi EG. 2013. DNA binding of the cell cycle transcriptional regulator
607 GcrA depends on N6-adenosine methylation in *Caulobacter crescentus* and other
608 Alphaproteobacteria. PLOS Genet 9:e1003541.
609 21. Wang ET, Young JPW. 2019. History of rhizobial taxonomy, p. 23–39. *In* Wang, ET, Tian,
610 CF, Chen, WF, Young, JPW, Chen, WX (eds.), Ecology and Evolution of Rhizobia:
611 Principles and Applications. Springer, Singapore.
612 22. Oldroyd GED. 2013. Speak, friend, and enter: signalling systems that promote beneficial
613 symbiotic associations in plants. Nat Rev Microbiol 11:252–263.
614 23. Barnett MJ, Toman CJ, Fisher RF, Long SR. 2004. A dual-genome Symbiosis Chip for
615 coordinate study of signal exchange and development in a prokaryote–host interaction. Proc
616 Natl Acad Sci USA 101:16636–16641.
617 24. Roux B, Rodde N, Jardinaud M-F, Timmers T, Sauviac L, Cottret L, Carrère S, Sallet E,
618 Courcelle E, Moreau S, Debellé F, Capela D, de Carvalho-Niebel F, Gouzy J, Bruand C,
619 Gamas P. 2014. An integrated analysis of plant and bacterial gene expression in symbiotic
620 root nodules using laser-capture microdissection coupled to RNA sequencing. Plant J
621 77:817–837.
622 25. diCenzo GC, Tesi M, Pfau T, Mengoni A, Fondi M. 2020. Genome-scale metabolic
623 reconstruction of the symbiosis between a leguminous plant and a nitrogen-fixing bacterium.
624 Nat Commun 11:2574.
625 26. Mergaert P, Uchiumi T, Alunni B, Evanno G, Cheron A, Catrice O, Mausset A-E, Barloy-
626 Hubler F, Galibert F, Kondorosi A, Kondorosi E. 2006. Eukaryotic control on bacterial cell
627 cycle and differentiation in the *Rhizobium*-legume symbiosis. Proc Natl Acad Sci USA
628 103:5230–5235.
629 27. Czernic P, Gully D, Cartieaux F, Moulin L, Guefrachi I, Patrel D, Pierre O, Fardoux J,
630 Chaintreuil C, Nguyen P, Gressent F, Da Silva C, Poulain J, Wincker P, Rofidal V, Hem S,

631 Barrière Q, Arrighi J-F, Mergaert P, Giraud E. 2015. Convergent evolution of endosymbiont
632 differentiation in Dalbergioid and inverted repeat-lacking clade legumes mediated by nodule-
633 specific cysteine-rich peptides. Plant Physiol 169:1254–1265.
634 28. Pini F, Frage B, Ferri L, De Nisco NJ, Mohapatra SS, Taddei L, Fioravanti A, Dewitte F,
635 Galardini M, Brilli M, Villeret V, Bazzicalupo M, Mengoni A, Walker GC, Becker A, Biondi
636 EG. 2013. The DivJ, CbrA and PleC system controls DivK phosphorylation and symbiosis
637 in *Sinorhizobium meliloti*. Mol Microbiol 90:54–71.
638 29. Kobayashi H, De Nisco NJ, Chien P, Simmons LA, Walker GC. 2009. *Sinorhizobium meliloti*
639 CpdR1 is critical for co-ordinating cell cycle progression and the symbiotic chronic infection.
640 Mol Microbiol 73:586–600.
641 30. Pini F, De Nisco NJ, Ferri L, Penterman J, Fioravanti A, Brilli M, Mengoni A, Bazzicalupo
642 M, Viollier PH, Walker GC, Biondi EG. 2015. Cell cycle control by the master regulator
643 CtrA in *Sinorhizobium meliloti*. PLOS Genet 11:e1005232.
644 31. Gibson KE, Campbell GR, Lloret J, Walker GC. 2006. CbrA is a stationary-phase regulator of
645 cell surface physiology and legume symbiosis in *Sinorhizobium meliloti*. J Bacteriol
646 188:4508–4521.
647 32. Van de Velde W, Zehirov G, Szatmari A, Debreczeny M, Ishihara H, Kevei Z, Farkas A,
648 Mikulass K, Nagy A, Tiricz H, Satiat-Jeunemaître B, Alunni B, Bourge M, Kucho K, Abe
649 M, Kereszt A, Maróti G, Uchiumi T, Kondorosi E, Mergaert P. 2010. Plant peptides govern
650 terminal differentiation of bacteria in symbiosis. Science 327:1122–1126.
651 33. Penterman J, Abo RP, De Nisco NJ, Arnold MFF, Longhi R, Zanda M, Walker GC. 2014.
652 Host plant peptides elicit a transcriptional response to control the *Sinorhizobium meliloti* cell
653 cycle during symbiosis. Proc Natl Acad Sci USA 111:3561–3566.
654 34. Farkas A, Maróti G, Dürgő H, Györgypál Z, Lima RM, Medzihradszky KF, Kereszt A,
655 Mergaert P, Kondorosi E. 2014. *Medicago truncatula* symbiotic peptide NCR247 contributes
656 to bacteroid differentiation through multiple mechanisms. Proc Natl Acad Sci USA
657 111:5183–5188.
658 35. Satgé C, Moreau S, Sallet E, Lefort G, Auriac M-C, Remblière C, Cottret L, Gallardo K, Noirot
659 C, Jardinaud M-F, Gamas P. 2016. Reprogramming of DNA methylation is critical for nodule
660 development in *Medicago truncatula*. Nat Plants 2:16166.
661 36. Nagymihaly M, Veluchamy A, Györgypál Z, Ariel F, Jégu T, Benhamed M, Sz\Hucs A,
662 Kereszt A, Mergaert P, Kondorosi E. 2017. Ploidy-dependent changes in the epigenome of
663 symbiotic cells correlate with specific patterns of gene expression. Proc Natl Acad Sci USA
664 114:4543–4548.
665 37. Pecrix Y, Staton SE, Sallet E, Lelandais-Brière C, Moreau S, Carrère S, Blein T, Jardinaud M-
666 F, Latrasse D, Zouine M, Zahm M, Kreplak J, Mayjonade B, Satgé C, Perez M, Cauet S,
667 Marande W, Chantry-Darmon C, Lopez-Roques C, Bouchez O, Bérard A, Debellé F, Muños
668 S, Bendahmane A, Bergès H, Niebel A, Buitink J, Frugier F, Benhamed M, Crespi M, Gouzy
669 J, Gamas P. 2018. Whole-genome landscape of *Medicago truncatula* symbiotic genes. Nat
670 Plants 4:1017–1025.

671 38. Davis-Richardson AG, Russell JT, Dias R, McKinlay AJ, Canepa R, Fagen JR, Rusoff KT,
672 Drew JC, Kolaczkowski B, Emerich DW, Triplett EW. 2016. Integrating DNA methylation
673 and gene expression data in the development of the soybean-*Bradyrhizobium* N_2-fixing
674 symbiosis. Front Microbiol 7:518.
675 39. Barrière Q, Guefrachi I, Gully D, Lamouche F, Pierre O, Fardoux J, Chaintreuil C, Alunni B,
676 Timchenko T, Giraud E, Mergaert P. 2017. Integrated roles of BclA and DD-
677 carboxypeptidase 1 in *Bradyrhizobium* differentiation within NCR-producing and NCR-
678 lacking root nodules. Sci Rep 7:9063.
679 40. diCenzo GC, Zamani M, Checcucci A, Fondi M, Griffitts J, Finan TM, Mengoni A. 2019.
680 Multi-disciplinary approaches for studying rhizobium - legume symbioses. Can J Microbiol
681 65:1–33.
682 41. diCenzo GC, Muhammed Z, Østerås M, O'Brien SAP, Finan TM. 2017. A key regulator of
683 the glycolytic and gluconeogenic central metabolic pathways in *Sinorhizobium meliloti*.
684 Genetics 207:961–974.
685 42. diCenzo GC, MacLean AM, Milunovic B, Golding GB, Finan TM. 2014. Examination of
686 prokaryotic multipartite genome evolution through experimental genome reduction. PLOS
687 Genet 10:e1004742.
688 43. Mohapatra SS, Fioravanti A, Biondi EG. 2014. DNA methylation in *Caulobacter* and other
689 Alphaproteobacteria during cell cycle progression. Trends Microbiol 22:528–535.
690 44. De Nisco NJ, Abo RP, Wu CM, Penterman J, Walker GC. 2014. Global analysis of cell cycle
691 gene expression of the legume symbiont *Sinorhizobium meliloti*. Proc Natl Acad Sci USA
692 111:3217–3224.
693 45. Fagorzi C, Ilie A, Decorosi F, Cangioli L, Viti C, Mengoni A, diCenzo GC. 2020. Symbiotic
694 and nonsymbiotic members of the genus *Ensifer* (syn. *Sinorhizobium*) are separated into two
695 clades based on comparative genomics and high-throughput phenotyping. Genome Biol Evol
696 12:2521–2534.
697 46. Moreau D, Voisin A-S, Salon C, Munier-Jolain N. 2008. The model symbiotic association
698 between *Medicago truncatula* cv. Jemalong and *Rhizobium meliloti* strain 2011 leads to N-
699 stressed plants when symbiotic N2 fixation is the main N source for plant growth. J Exp Bot
700 59:3509–3522.
701 47. Kazmierczak T, Nagymihaly M, Lamouche F, Barrière Q, Guefrachi I, Alunni B, Ouadghiri
702 M, Ibijbijen J, Kondorosi E, Mergaert P, Gruber V. 2017. Specific host-responsive
703 associations between *Medicago truncatula* accessions and *Sinorhizobium strains*. Mol Plant
704 Microbe Interact 30:399–409.
705 48. Vasse J, De Billy F, Camut S, Truchet G. 1990. Correlation between ultrastructural
706 differentiation of bacteroids and nitrogen fixation in alfalfa nodules. J Bacteriol 172:4295–
707 4306.
708 49. Ferri L, Gori A, Biondi EG, Mengoni A, Bazzicalupo M. 2010. Plasmid electroporation of
709 *Sinorhizobium* strains: The role of the restriction gene *hsdR* in type strain Rm1021. Plasmid
710 63:128–135.

711 50. Brumwell SL, MacLeod MR, Huang T, Cochrane RR, Meaney RS, Zamani M,
712 Matysiakiewicz O, Dan KN, Janakirama P, Edgell DR, Charles TC, Finan TM, Karas BJ.
713 2019. Designer *Sinorhizobium meliloti* strains and multi-functional vectors enable direct
714 inter-kingdom DNA transfer. PLOS One 14:e0206781.
715 51. Haakonsen DL, Yuan AH, Laub MT. 2015. The bacterial cell cycle regulator GcrA is a σ^{70}
716 cofactor that drives gene expression from a subset of methylated promoters. Genes Dev
717 29:2272–2286.
718 52. Wright R, Stephens C, Zweiger G, Shapiro L, Alley MR. 1996. *Caulobacter* Lon protease has
719 a critical role in cell-cycle control of DNA methylation. Genes Dev 10:1532–1542.
720 53. Frage B, Döhlemann J, Robledo M, Lucena D, Sobetzko P, Graumann PL, Becker A. 2016.
721 Spatiotemporal choreography of chromosome and megaplasmids in the *Sinorhizobium*
722 *meliloti* cell cycle. Mol Microbiol 100:808–823.
723 54. Rasmussen T, Jensen RB, Skovgaard O. 2007. The two chromosomes of *Vibrio cholerae* are
724 initiated at different time points in the cell cycle. EMBO J 26:3124–3131.
725 55. Krall L, Wiedemann U, Unsin G, Weiss S, Domke N, Baron C. 2002. Detergent extraction
726 identifies different VirB protein subassemblies of the type IV secretion machinery in the
727 membranes of *Agrobacterium tumefaciens*. Proc Natl Acad Sci USA 99:11405–11410.
728 56. Cowie A, Cheng J, Sibley CD, Fong Y, Zaheer R, Patten CL, Morton RM, Golding GB, Finan
729 TM. 2006. An integrated approach to functional genomics: construction of a novel reporter
730 gene fusion library for *Sinorhizobium meliloti*. Appl Environ Microbiol 72:7156–7167.
731 57. Eid J, Fehr A, Gray J, Luong K, Lyle J, Otto G, Peluso P, Rank D, Baybayan P, Bettman B,
732 Bibillo A, Bjornson K, Chaudhuri B, Christians F, Cicero R, Clark S, Dalal R, deWinter A,
733 Dixon J, Foquet M, Gaertner A, Hardenbol P, Heiner C, Hester K, Holden D, Kearns G, Kong
734 X, Kuse R, Lacroix Y, Lin S, Lundquist P, Ma C, Marks P, Maxham M, Murphy D, Park I,
735 Pham T, Phillips M, Roy J, Sebra R, Shen G, Sorenson J, Tomaney A, Travers K, Trulson
736 M, Vieceli J, Wegener J, Wu D, Yang A, Zaccarin D, Zhao P, Zhong F, Korlach J, Turner S.
737 2009. Real-time DNA sequencing from single polymerase molecules. Science 323:133–138.
738 58. Chaisson MJ, Tesler G. 2012. Mapping single molecule sequencing reads using basic local
739 alignment with successive refinement (BLASR): application and theory. BMC Bioinform
740 13:238.
741 59. Flusberg BA, Webster DR, Lee JH, Travers KJ, Olivares EC, Clark TA, Korlach J, Turner
742 SW. 2010. Direct detection of DNA methylation during single-molecule, real-time
743 sequencing. Nat Methods 7:461–465.
744 60. Clark TA, Murray IA, Morgan RD, Kislyuk AO, Spittle KE, Boitano M, Fomenkov A, Roberts
745 RJ, Korlach J. 2012. Characterization of DNA methyltransferase specificities using single-
746 molecule, real-time DNA sequencing. Nucleic Acids Res 40:e29–e29.
747 61. diCenzo GC, Debiec K, Krzysztoforski J, Uhrynowski W, Mengoni A, Fagorzi C, Gorecki A,
748 Dziewit L, Bajda T, Rzepa G, Drewniak L. 2018. Genomic and biotechnological
749 characterization of the heavy-metal resistant, arsenic-oxidizing bacterium *Ensifer* sp. M14.
750 Genes 9:379.

751 62. Page AJ, Cummins CA, Hunt M, Wong VK, Reuter S, Holden MTG, Fookes M, Falush D,
752 Keane JA, Parkhill J. 2015. Roary: rapid large-scale prokaryote pan genome analysis.
753 Bioinformatics 31:3691–3693.
754 63. Seemann T. 2014. Prokka: rapid prokaryotic genome annotation. Bioinformatics 30:2068–
755 2069.
756 64. Löytynoja A. 2014. Phylogeny-aware alignment with PRANK., p. 155–170. *In* Russell, D
757 (ed.), Methods in Molecular Biology (Methods and Protocols). Humana Press, Totowa, NJ.
758 65. Capella-Gutiérrez S, Silla-Martínez JM, Gabaldón T. 2009. trimAl: a tool for automated
759 alignment trimming in large-scale phylogenetic analyses. Bioinformatics 25:1972–1973.
760 66. Stamatakis A. 2014. RAxML version 8: a tool for phylogenetic analysis and post-analysis of
761 large phylogenies. Bioinformatics 30:1312–1313.
762 67. Letunic I, Bork P. 2016. Interactive tree of life (iTOL) v3: an online tool for the display and
763 annotation of phylogenetic and other trees. Nucleic Acids Res 44:W242–W245.
764 68. Eddy SR. 2009. A new generation of homology search tools based on probabilistic inference.
765 Genome Inform 23:205–211.
766 69. Finn RD, Coggill P, Eberhardt RY, Eddy SR, Mistry J, Mitchell AL, Potter SC, Punta M,
767 Qureshi M, Sangrador-Vegas A, Salazar GA, Tate J, Bateman A. 2016. The Pfam protein
768 families database: towards a more sustainable future. Nucleic Acids Res 44:D279–D285.
769 70. Haft DH, Selengut JD, Richter RA, Harkins D, Basu MK, Beck E. 2013. TIGRFAMs and
770 genome properties in 2013. Nucleic Acids Res 41:D387–D395.
771

772 **SUPPLEMENTARY MATERIALS**

773 **Table S1.** Number of mapped subread bases per sample.

Strain	Condition & Replicate	Genome Accession (NCBI Assembly)	Mapped Subreads Bases	NCBI Biosample Accessions for data generated in this study
E. meliloti FSM-MA	MM9-succinate 1	GCA_002215195.1	2,414,122,786	SAMN12793050
E. meliloti FSM-MA	MM9-succinate 2	GCA_002215195.1	1,973,904,434	SAMN12793077
E. meliloti FSM-MA	MM9-succinate 3	GCA_002215195.1	2,276,105,908	SAMN12793051
E. meliloti FSM-MA	MM9-succinate stationary 1	GCA_002215195.1	983,326,511	SAMN18104108
E. meliloti FSM-MA	MM9-succinate stationary 2	GCA_002215195.1	1,149,234,588	SAMN18104109
E. meliloti FSM-MA	MM9-succinate stationary 3	GCA_002215195.1	1,112,880,576	SAMN18104110
E. meliloti FSM-MA	*M. sativa* whole nodules 1	GCA_002215195.1	669,928,620	SAMN12793096
E. meliloti FSM-MA	*M. sativa* whole nodules 2	GCA_002215195.1	619,883,638	SAMN12793097
E. meliloti FSM-MA	*M. sativa* whole nodules 3	GCA_002215195.1	2,014,796,418	SAMN12792983
E. meliloti FSM-MA	*M. truncatula* whole nodules 1	GCA_002215195.1	1,223,643,922	SAMN13167594
E. meliloti FSM-MA	*M. truncatula* whole nodules 2	GCA_002215195.1	993,061,044	SAMN13168102
E. meliloti FSM-MA	*M. truncatula* whole nodules 3	GCA_002215195.1	1,038,438,145	SAMN13167746
E. meliloti FSM-MA	*M. sativa* distal nodule sections 1	GCA_002215195.1	454,620,103	SAMN15739021
E. meliloti FSM-MA	*M. sativa* proximal nodule sections 1	GCA_002215195.1	2,644,422,028	SAMN14511029
E. meliloti Rm2011	MM9-succinate 1	GCA_000346065.1	1,859,862,543	SAMN12793063
E. meliloti Rm2011	MM9-succinate 2	GCA_000346065.1	2,307,929,125	SAMN12793100
E. meliloti Rm2011	MM9-succinate 3	GCA_000346065.1	2,174,907,026	SAMN12793065
E. meliloti Rm2011	MM9-succinate stationary 1	GCA_000346065.1	1,303,384,501	SAMN18104105
E. meliloti Rm2011	MM9-succinate stationary 2	GCA_000346065.1	1,223,856,206	SAMN18104106
E. meliloti Rm2011	MM9-succinate stationary 3	GCA_000346065.1	1,217,560,804	SAMN18104107
E. meliloti Rm2011	MM9-sucrose 1	GCA_000346065.1	1,302,160,101	SAMN12793017
E. meliloti Rm2011	MM9-sucrose 2	GCA_000346065.1	1,673,395,654	SAMN12793016
E. meliloti Rm2011	MM9-sucrose 3	GCA_000346065.1	2,965,562,894	SAMN12793099
E. meliloti Rm2011	*M. sativa* whole nodules 1	GCA_000346065.1	814,031,492	SAMN12793035
E. meliloti Rm2011	*M. sativa* whole nodules 2	GCA_000346065.1	798,075,475	SAMN12793034
E. meliloti Rm2011	*M. sativa* whole nodules 3	GCA_000346065.1	1,138,170,578	SAMN12793018
E. meliloti Rm2011	*M. sativa* distal nodule sections 1	GCA_000346065.1	6,610,799,406	SAMN16773451
E. meliloti Rm2011	*M. sativa* proximal nodule sections 2	GCA_000346065.1	1,544,246,716	SAMN13167882
E. meliloti RmP3496	MM9-sucrose 1	GCA_000346065.1	1,138,116,503	SAMN12792987
E. meliloti RmP3496	MM9-sucrose 2	GCA_000346065.1	1,328,727,153	SAMN12793078
E. meliloti RmP3496	MM9-sucrose 3	GCA_000346065.1	2,234,156,208	SAMN12793033
E. meliloti RmP3499	MM9-sucrose 1	GCA_000346065.1	1,302,160,101	SAMN12792986
E. meliloti RmP3499	MM9-sucrose 2	GCA_000346065.1	1,673,395,654	SAMN12793101
E. meliloti RmP3499	MM9-sucrose 3	GCA_000346065.1	2,965,562,894	SAMN12793098
E. fredii NGR234	MM9-succinate 1	GCA_000018545.1	1,741,191,577	SAMN12792974
E. fredii NGR234	MM9-succinate 2	GCA_000018545.1	2,484,318,454	SAMN12793102
E. fredii NGR234	MM9-succinate 3	GCA_000018545.1	2,279,238,884	SAMN12793103
E. adhaerens OV14	MM9-succinate 1	GCA_000583045.1	1,883,494,303	SAMN12793104
E. adhaerens OV14	MM9-succinate 2	GCA_000583045.1	2,914,499,268	SAMN12792975
E. adhaerens OV14	MM9-succinate 3	GCA_000583045.1	2,272,958,147	SAMN12793036

774

775 **Table S2.** Average extent of methylation of m6A modified motifs in *E. meliloti* Rm2011.

Condition	Chromosome	pSymB	pSymA
GANTC			
Free-living (mid-exponential)	0.878	0.916	0.883
Free-living (stationary)	0.957	0.960	0.964
M. sativa distal nodule sections	0.712	0.883	0.882
M. sativa proximal nodule sections	0.787	0.925	0.918
M. sativa whole nodules	0.791	0.921	0.920
CGCA(N_5)GTG			
Free-living (mid-exponential)	0.978	0.980	0.981
Free-living (stationary)	0.974	0.981	0.982
M. sativa distal nodule sections	0.966	0.959	0.961
M. sativa proximal nodule sections	0.978	0.974	0.983
M. sativa whole nodules	0.973	0.972	0.976

776

777 **Table S3.** Average extent of methylation of m6A modified motifs in *E. meliloti* FSM-MA.

Condition	Chromosome	pSymB	pSymA
GANTC			
Free-living (mid-exponential)	0.803	0.898	0.836
Free-living (stationary)	0.938	0.947	0.947
M. sativa distal nodule sections	0.906	0.948	0.942
M. sativa proximal nodule sections	0.784	0.915	0.895
M. sativa whole nodules	0.860	0.946	0.936
M. truncatula whole nodules	0.793	0.917	0.911
TCGA(N_8)TCGA			
Free-living (mid-exponential)	0.984	0.984	0.981
Free-living (stationary)	0.980	0.984	0.984
M. sativa distal nodule sections	0.973	0.973	0.980
M. sativa proximal nodule sections	0.978	0.915	0.977
M. sativa whole nodules	0.980	0.964	0.963
M. truncatula whole nodules	0.976	0.974	0.981

778

779 **Table S4.** Bacterial strains.

Strain	Genotype	Source
Ensifer adhearens OV14	Wild type OV14; not a nitrogen-fixing legume symbiont	(1)
Ensifer fredii NGR234	Wild type NGR234 *rif-1*; Rif^R	(2)
Ensifer meliloti FSM-MA	Wild type FSM-MA; Cm^R	(3)
Ensifer meliloti Rm2011	Wild type SU47 *str-3*; Sm^R	Lab collection
Ensifer meliloti RmP3496	Rm2011 lacking pSymA and pSymB; Sm^R Sp^R	(4, 5)
Ensifer meliloti RmP3499	RmP3496 with pSymA and pSymB re-introduced; Sm^R Sp^R	(5)

780 Cm – Chloramphenicol; Sm – Streptomycin; Sp – Spectinomycin; Rif – Rifampicin

781 **Table S5.** Densities and doubling times of bacterial cultures grown for the isolation of DNA.

		Replicate 1 *			Replicate 2			Replicate 3		
Strain	Carbon source	Starting density †	Final density	Doubling time (h)	Starting density	Final density	Doubling time (h)	Starting density	Final density	Doubling time (h)
E. meliloti Rm2011	Succinate	0.015	0.554	2.98	0.014	0.525	2.94	0.014	0.499	2.98
E. meliloti FSM-MA	Succinate	0.004	0.571	2.13	0.003	0.420	2.20	0.003	0.369	2.26
E. fredii NGR234	Succinate	0.019	0.631	3.09	0.016	0.650	2.92	0.016	0.634	2.94
E. adhaerens OV14	Succinate	0.002	0.686	1.86	0.002	0.564	1.84	0.002	0.552	1.85
E. meliloti Rm2011	Sucrose	0.014	0.581	2.91	0.013	0.512	2.90	0.013	0.516	2.90
E. meliloti RmP3499	Sucrose	0.010	0.648	2.56	0.008	0.494	2.61	0.008	0.46	2.66
E. meliloti RmP3496	Sucrose	0.039	0.574	3.98	0.035	0.389	4.44	0.035	0.401	4.38

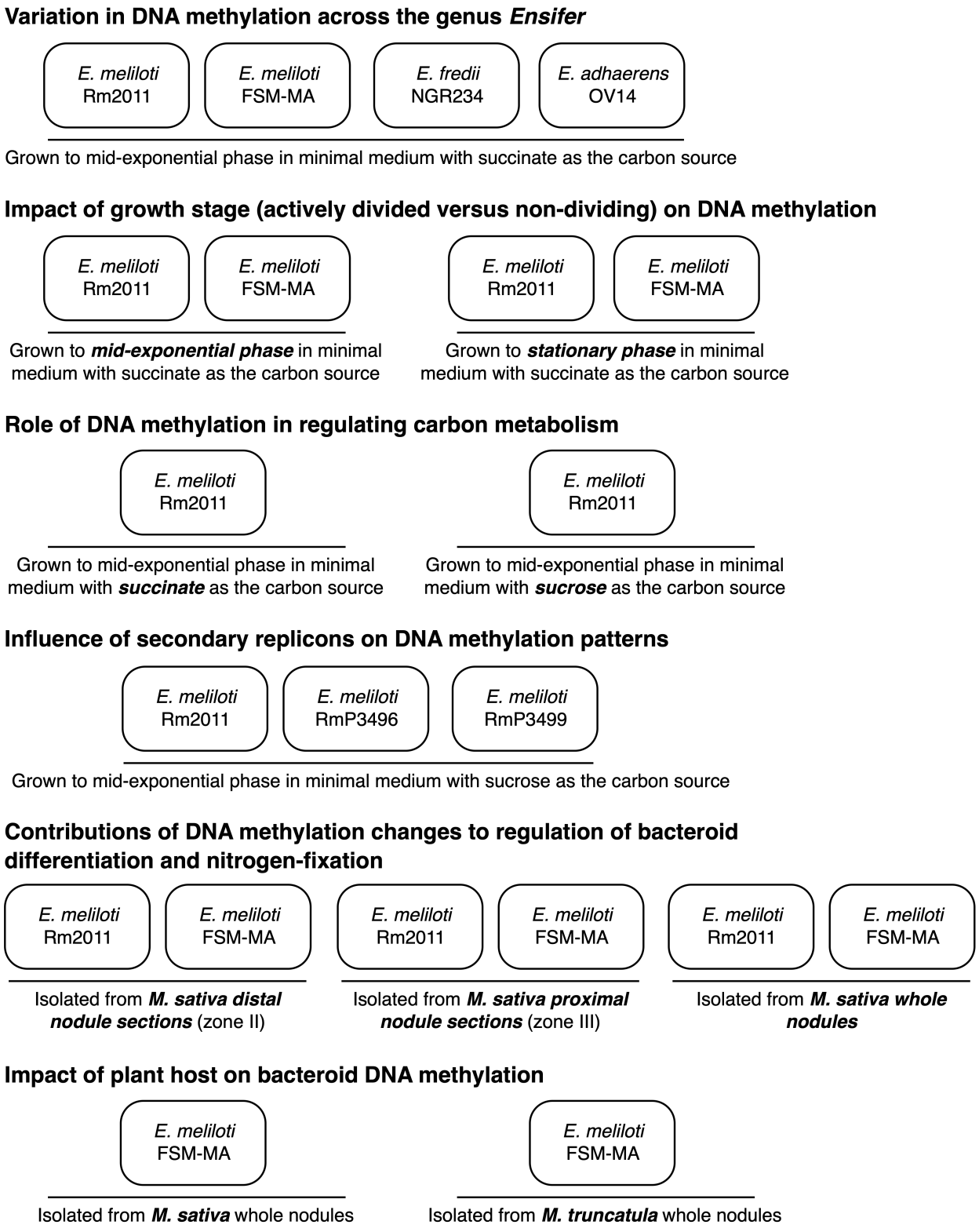
782 * Each Replicate 1 sample consisted of two 40 mL cultures (inoculated from the same starter culture) that were combined prior to DNA
783 isolation. The density and doubling time values for this replicate are based on just one of the cultures for each sample.
784 † All density values are OD_{600nm} measurements.

785 **Table S6.** Plant shoot dry weights.

Plant	Bacterium	Replicate	Number of plants	Average shoot dry weight (mg)
Plants for whole nodules †				
M. sativa	*E. meliloti* Rm2011	1	55	57
M. sativa	*E. meliloti* Rm2011	2	52	60
M. sativa	*E. meliloti* Rm2011	3	52	54
M. sativa	*E. meliloti* FSM-MA	1	60	84
M. sativa	*E. meliloti* FSM-MA	2	60	80
M. sativa	*E. meliloti* FSM-MA	3	55	90
M. truncatula	*E. meliloti* FSM-MA	1	60	58
M. truncatula	*E. meliloti* FSM-MA	2	60	60
M. truncatula	*E. meliloti* FSM-MA	3	55	68
M. sativa	Uninoculated	1	10	ND *
M. truncatula	Uninoculated	1	5	ND *
Plants for sectioned nodules†				
M. sativa	*E. meliloti* Rm2011	1	103	180
M. sativa	*E. meliloti* FSM-MA	1	103	162
M. sativa	Uninoculated	1	55	31

786 * ND: Not determined.

787 † Plants for isolation of whole nodules and plants for isolation of sectioned nodules were grown
788 independently at separate periods of the year.



789
790
791 **Figure S1. Experimental design overview.** A schematic overview of the experimental design of
792 this study, summarizing the comparisons that were performed, and which strains/conditions
793 correspond to each comparison.

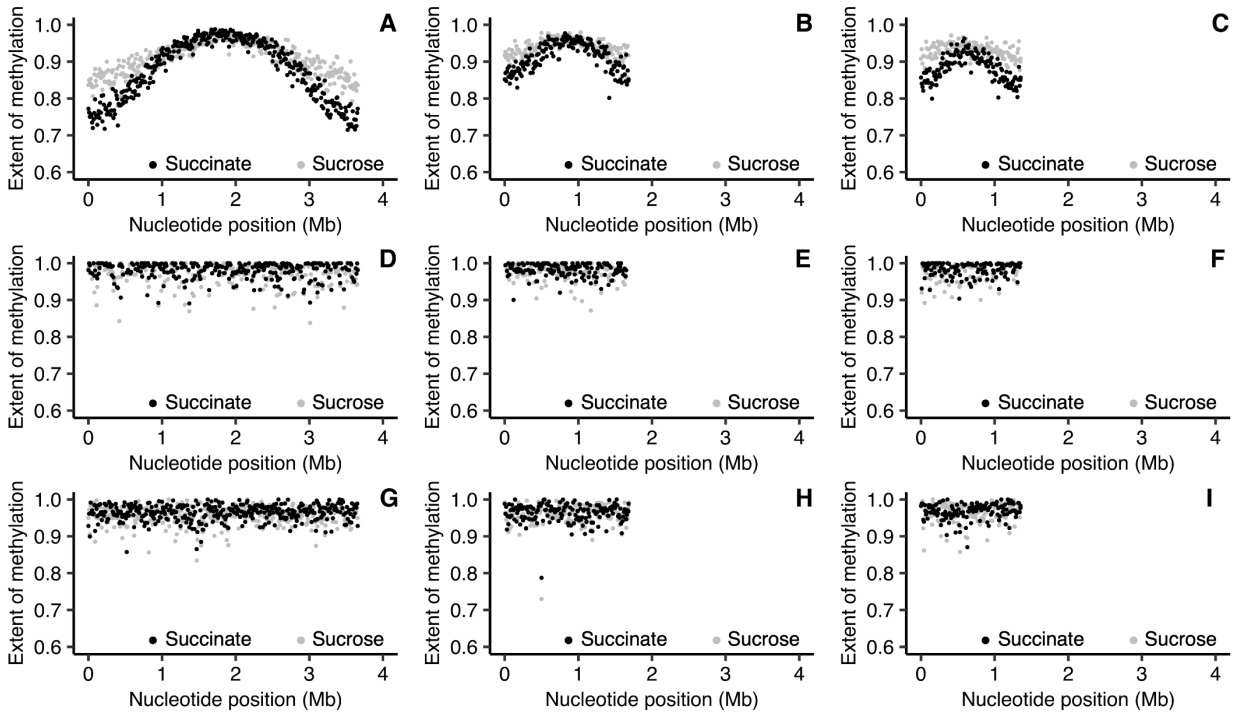


Figure S2. Impact of carbon source on genome-wide DNA methylation of *E. meliloti* Rm2011. The extent of methylation is shown, using a 10 kb sliding window, of cells grown to mid-exponential phase and provided succinate (black) or sucrose (grey) as the sole source of carbon. (**A**-**C**) Data for the GANTC motif for the chromosome (A), pSymB (B), and pSymA (C). (**D**-**F**) Data for the CGCA(N_5)GTG motif for the chromosome (D), pSymB (E), and pSymA (F). (**G**-**I**) Data for the RCGCCTC motif for the chromosome (G), pSymB (H), and pSymA (I).

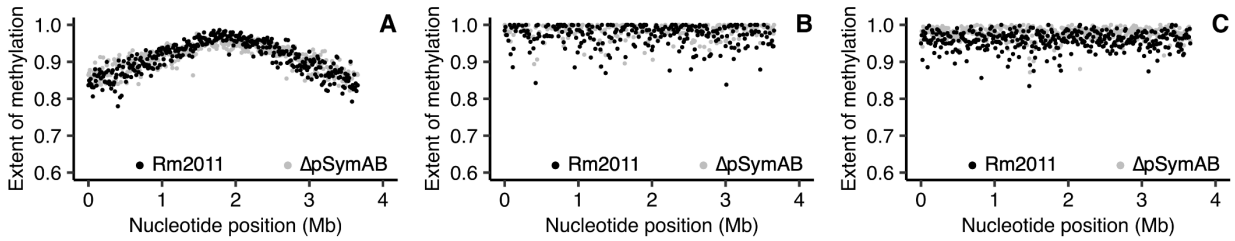
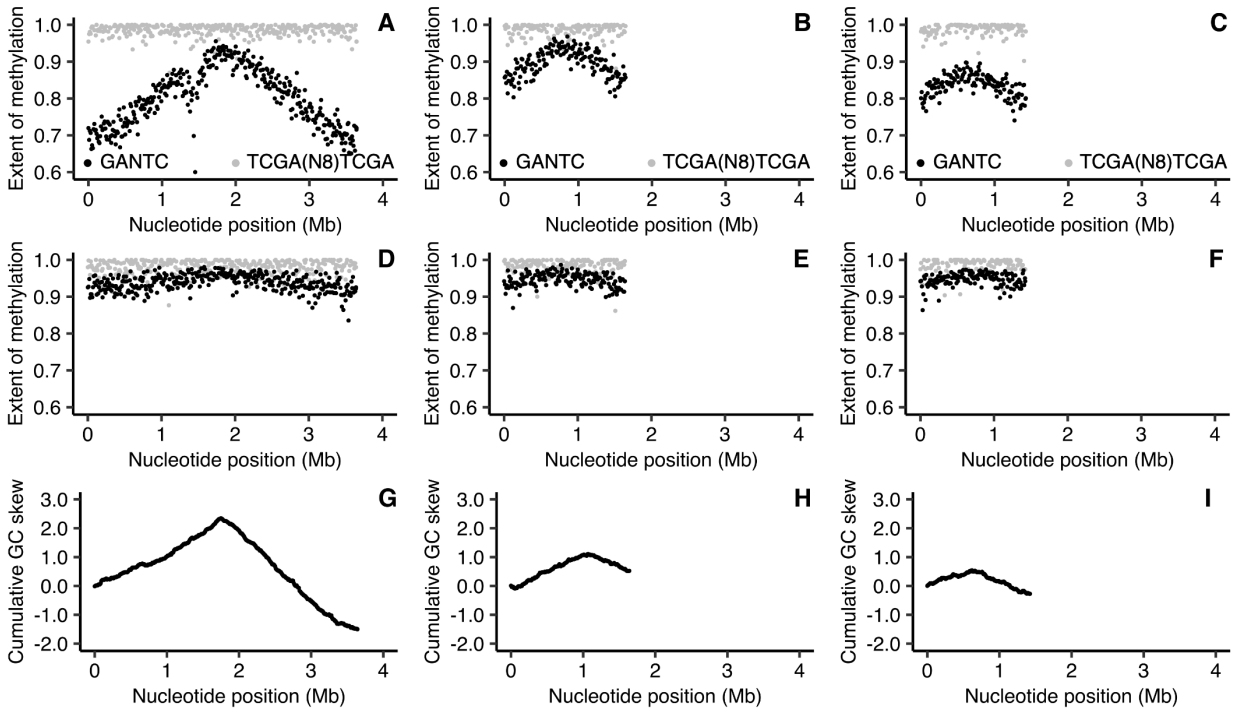


Figure S3. Impact of pSymA and pSymB removal on chromosome-wide DNA methylation of *E. meliloti* Rm2011. The extent of methylation for the chromosome, shown using a 10 kb sliding window, is provided for wild type *E. meliloti* Rm2011 (black) or *E. meliloti* ΔpSymAB (grey) grown to mid-exponential phase and provided sucrose as the sole source of carbon. (**A**) Data for the GANTC motif. (**B**) Data for the CGCA(N_5)GTG motif. (**C**) Data for the RCGCCTC motif.



809
810
811 **Figure S4. Genome-wide DNA methylation of *E. meliloti* FSM-MA.** (**A**-**F**) The extent of
812 methylation is shown, using a 10 kb sliding window, of GANTC sites (black) and
813 TCGA(N_5)TCGA sites (grey) across the chromosome (**A**,**D**), pSymB (**B**,**E**), and pSymA (**C**,**F**)
814 replicons of exponential phase (**A**-**C**) or early stationary phase (**D**-**F**) *E. meliloti* FSM-MA.
815 Averages from three biological replicates are shown. (**G**-**I**) Cumulative GC skews are shown, using
816 a 10 kb sliding window, across the *E. meliloti* FSM-MA chromosome (**G**), pSymB (**H**), and
817 pSymA (**I**) replicons.

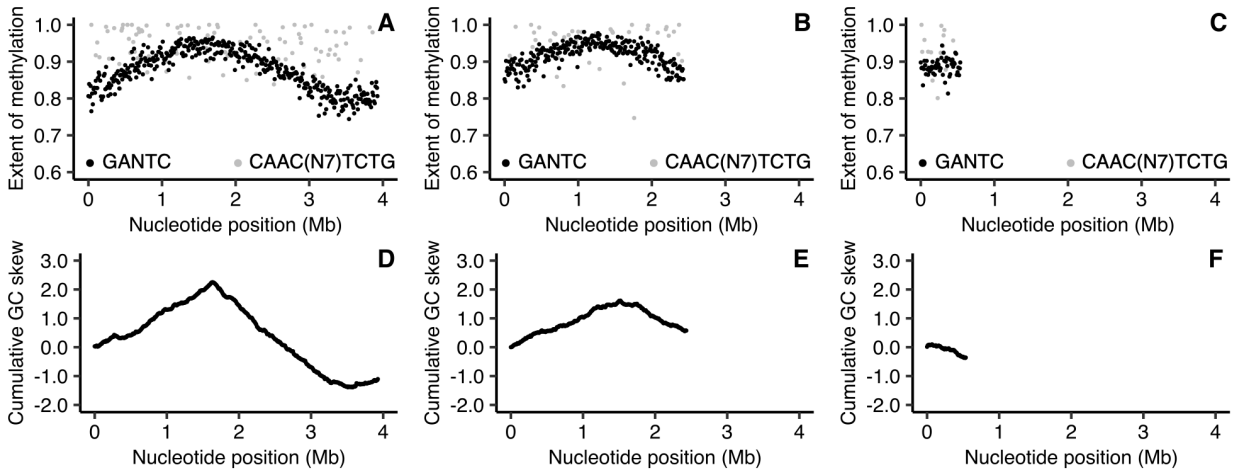
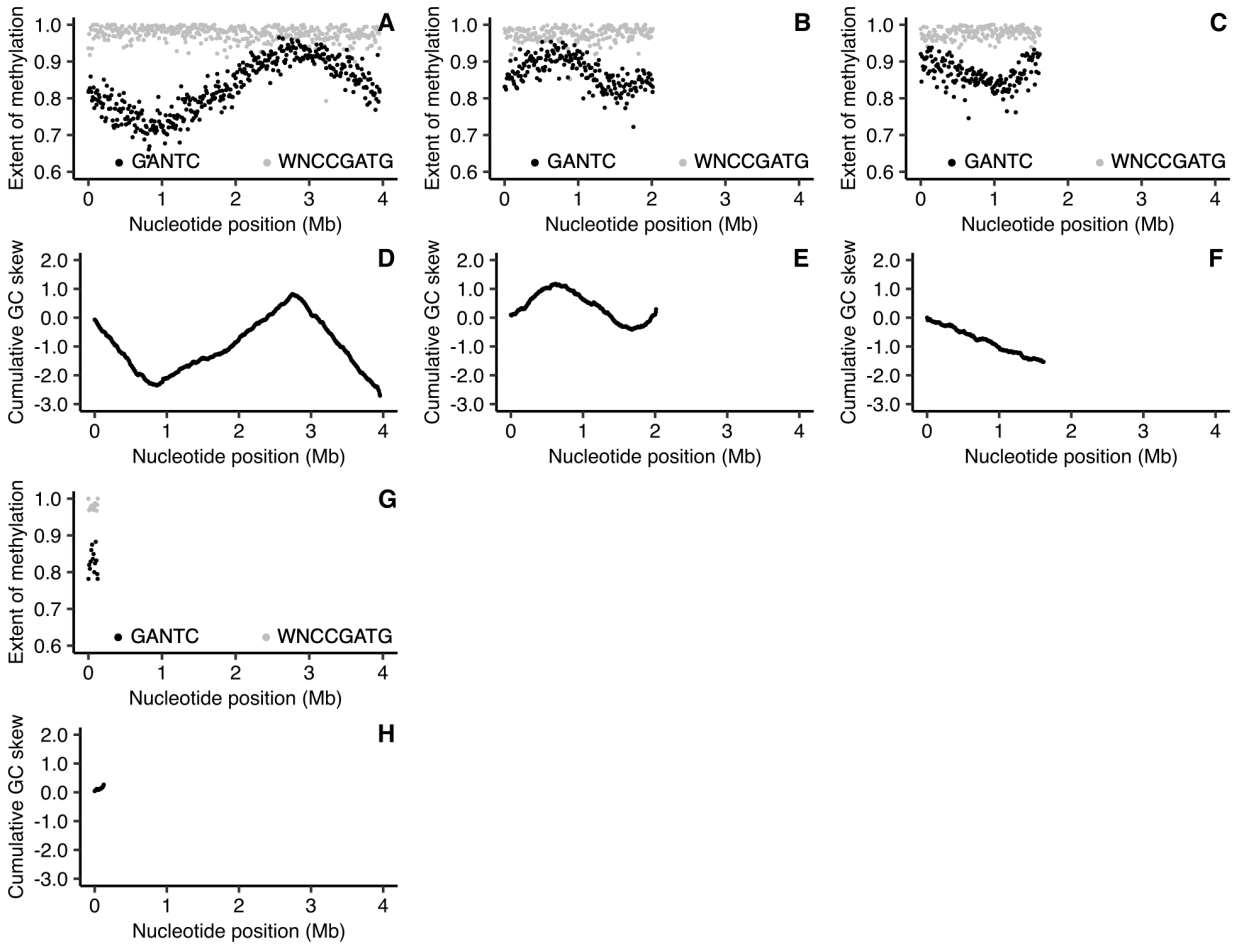


Figure S5. Genome-wide DNA methylation of *E. fredii* NGR234. (**A**-**C**) The extent of methylation is shown, using a 10 kb sliding window, of GANTC sites (black) and CAAC(N_7)TCTG sites (grey) across the chromosome (**A**), pNGR234b (**B**), and pNGR234a (**C**) replicons of exponential phase *E. fredii* NGR234. Averages from three biological replicates are shown. (**D**-**F**) Cumulative GC skews are shown, using a 10 kb sliding window, across the *E. fredii* NGR234 chromosome (**D**), pNGR234b (**E**), and pNGR234a (**F**) replicons.



826
827
828 **Figure S6. Genome-wide DNA methylation of *E. adhaerens* OV14.** (**A**-**C**,**G**) The extent of
829 methylation is shown using a 10 kb sliding window, of GANTC sites (black) and WNCCGATG
830 sites (grey) across chromosome 1 (**A**), chromosome 2 (**B**), pOV14b (**C**), and pOV14c (**G**) replicons
831 of exponential phase *E. adhaerens* OV14. Averages from three biological replicates are shown.
832 (**D**-**F**,**H**) Cumulative GC skews are shown, using a 10 kb sliding window, across the *E. adhaerens*
833 OV14 chromosome 1 (**D**), chromosome 2 (**E**), pOV14b (**F**), and pOV14c (**H**) replicons.

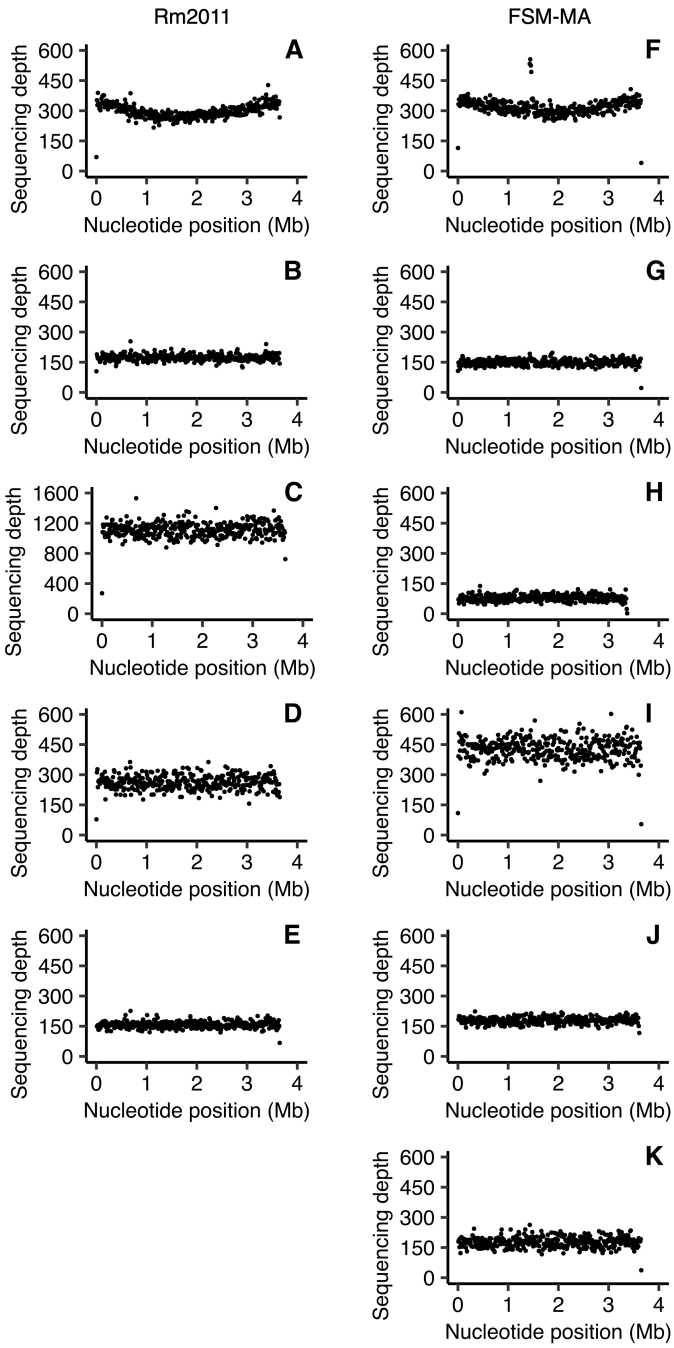
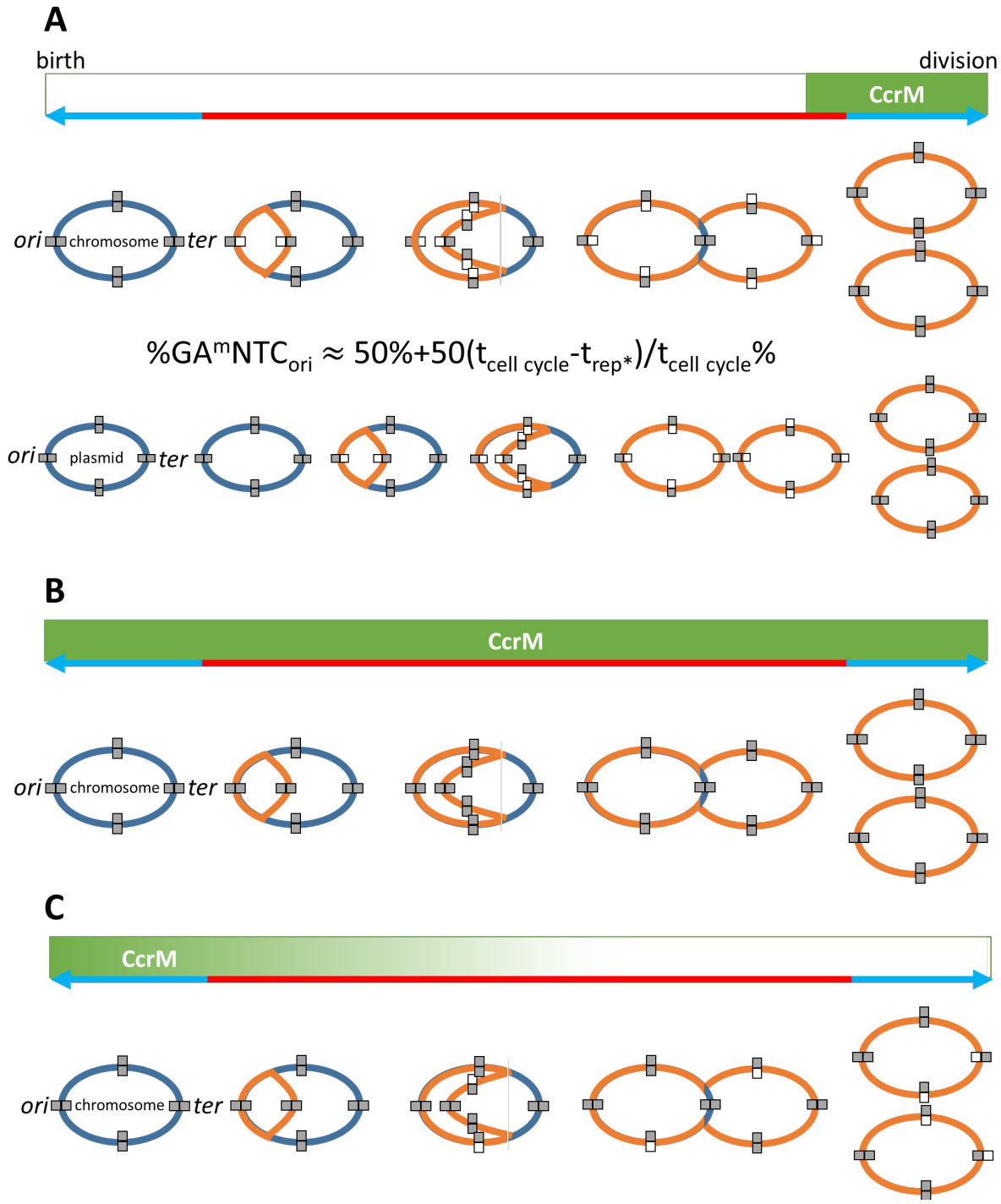
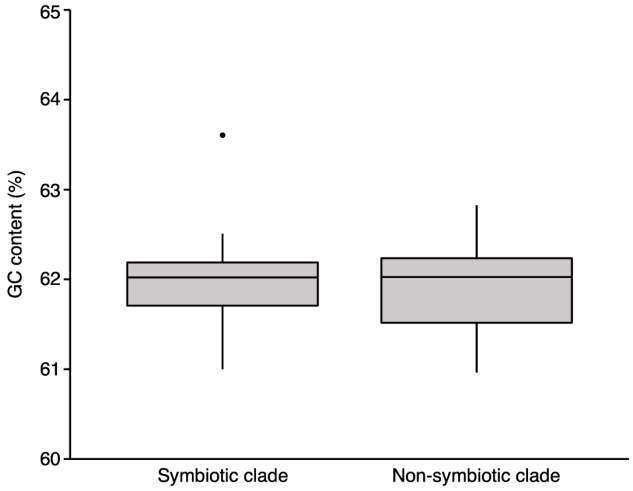


Figure S7. Sequencing depth across the *E. meliloti* RM2011 and FSM-MA chromosome. The sequencing depths (i.e., number of reads mapping to a given nucleotide) of the (**A**-**E**) *E. meliloti* Rm2011 chromosome and (**F**-**K**) *E. meliloti* FSM-MA chromosome are shown, using a 10 kb sliding window. Averages from three biological replicates are shown for free-living and whole nodule samples; data represents one replicate for the zone II and zone III nodule sections. (**A**,**F**) Free-living cells harvested in mid-exponential phase. (**B**,**G**) Free-living cells harvested in early stationary phase. (**C**,**H**) Bacteroids isolated from *M. sativa* zone II nodule sections. (**D**,**I**) Bacteroids isolated from *M. sativa* zone III nodule sections. (**E**,**J**) Bacteroids isolated from *M. sativa* whole nodule samples. (**K**) Bacteroids isolated from *M. truncatula* whole nodule samples.



845
846
847 **Figure S8. Model describing the methylation patterns observed in cultures and bacteroids.**
848 **(A)** Methylation pattern in free-living cells in exponential and stationary phases of growth. The
849 blue/red line indicates the relevant cell cycle phases: red is the S-phase or genome replication
850 phase; blue is the gap phases and division. The replication progression and methylation status of
851 the circular chromosome and a plasmid is depicted during the cell cycle progression from cell birth
852 until division. The replicons at cell birth are in dark blue and the newly replicated DNA is in

853 orange. Full grey rectangles indicate the fully methylated state and grey/white rectangles indicate
854 the hemimethylated state. Replicons start the cell cycle in a fully methylated state across the
855 genome and freshly replicated genome segments become hemimethylated. CcrM activity (green
856 box) is confined to a short window in the cell cycle because the *ccrM* gene is expressed in the late
857 phase of genome replication (6) and because the protein is degraded by the Lon protease prior to
858 cell division (7). The formula expresses the extent of methylation at the origin (*ori*) of replication
859 in an asynchronous bacterial population in culture. The methylation level shuttles between 50%
860 and 100% depending on the proportion of cells in the culture that have a hemimethylated origin
861 due to ongoing DNA replication prior to the activation of CcrM, and the complementary proportion
862 of cells that have a fully methylated origin because they either have not yet started DNA replication
863 or they have completed (or nearly completed) DNA replication and have activated CcrM. The
864 fraction of cells in the different cell phases is in turn proportional to the duration (t) of the
865 corresponding cell cycle stages (t_{rep*} is the time from initiation of replication till activation of
866 CcrM, just before replication termination). Based on this formula, our experimentally measured
867 extent of methylation at the *ori* of the chromosome in exponential phase cells, and the measured
868 doubling time of *E. meliloti*, we estimate that it takes about 1 hour to replicate the chromosome,
869 corresponding to a DNA polymerase rate of ~ 450 to 500 nucleotides per second. This compares
870 well with the estimated DNA polymerase rates of *Escherichia coli* and *Caulobacter crescentus*,
871 which are ~ 600 and ~ 350 nucleotides per second, respectively (8, 9). In the genus *Ensifer*, the
872 extent of methylation at the *ori* of the megaplasmids is higher than in the chromosome, while the
873 terminus (*ter*) of the megaplasmids has a slightly lower extent of methylation. Since DNA
874 methylation happens at the fixed stage of the cell cycle at the end of chromosome replication, it
875 follows that plasmid replication is initiated later in the cell cycle than initiation of chromosome
876 replication, and that their replication terminates slightly before termination of the chromosome and
877 the activation of CcrM. This is consistent with independent analyses revealing spatiotemporal
878 regulation of DNA replication and partitioning in *E. meliloti* (6, 10). The *ter* of the chromosome
879 remains near fully methylated during the complete cell cycle because its replication coincides with
880 the activity window of CcrM. In a stationary phase culture, methylation approaches 100% because
881 few cells are in the replication phase and the majority of cells are in a gap phase of the cell cycle.
882 **(B)** Methylation pattern in an early stage of bacteroid differentiation. Early stage bacteroids have
883 a high extent of methylation across the genome suggesting that the activity of CcrM is extended
884 to include also the replication phase of the cell cycle. This can be the result of an aberrant
885 expression of the *ccrM* gene or a lack of proteolytic degradation of the CcrM protein. **(C)**
886 Methylation pattern in a late stage of bacteroid differentiation. In mature bacteroids, the
887 methylation extent at the *ori* is high but characteristically, the *ter* of the chromosome in bacteroids
888 has a reduced methylation extent. This can be the result of a drop in CcrM activity during the last
889 chromosome replication cycle of the endoreduplication process in differentiating bacteroids.
890 Possibly, the drop of CcrM activity can also be the cause of the arrest of the endoreduplication
891 process. Reduced methylation at the *ter* region is generally not observed in the plasmids,
892 suggesting that plasmid endoreduplication finishes before chromosome endoreduplication and that
893 CcrM activity stops after the end of plasmid replication. The iconography of the illustrations is
894 based on Figure 1 of Mohapatra et al. 2014 (11).



895
896
897 **Figure S9. GC content in the genus *Ensifer*.** Box plots summarizing the GC content of the
898 genomes of 157 *Ensifer* strains are shown. The monophyletic "symbiotic" and "non-symbiotic"
899 clades as defined previously (12), are represented by 111 and 44 genomes respectively.

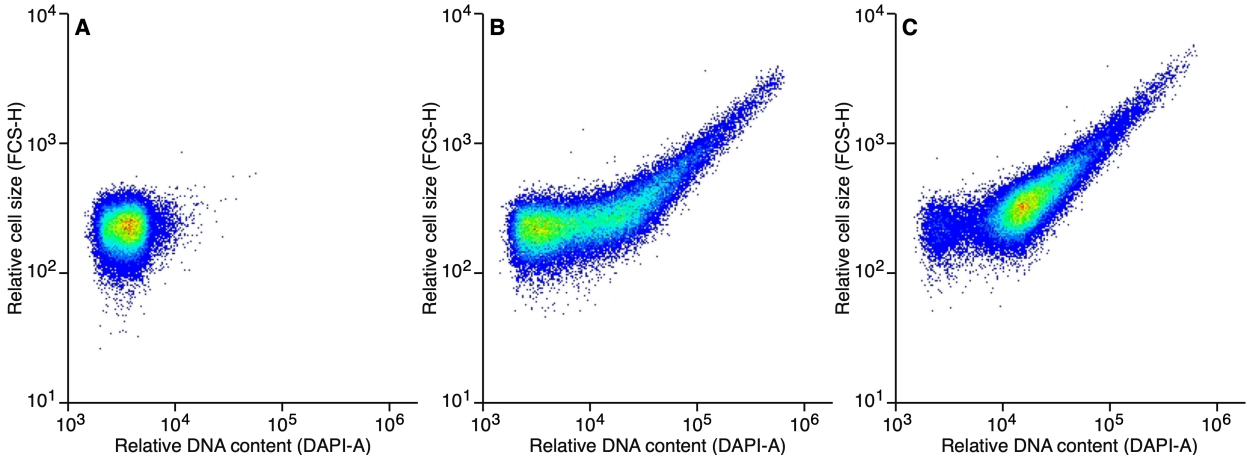


Figure S10. Cell size and DNA content of *E. meliloti* Rm2011 cell populations. Pseudo-coloured scatterplots displaying the DNA content (X-axis) and cell size (Y-axis) of various *E. meliloti* Rm2011 populations, based on flow cytometry readings of 50,000 cells. The colour scheme indicates the number of values plotted at a given location of the graph, with blue to red indicating lower to higher density. Free-living *E. meliloti* Rm2011 cells (**A**), *E. meliloti* Rm2011 cells purified from *M. sativa* zone II nodule sections (**B**), and *E. meliloti* Rm2011 cells purified from *M. sativa* zone III nodule sections (**C**).

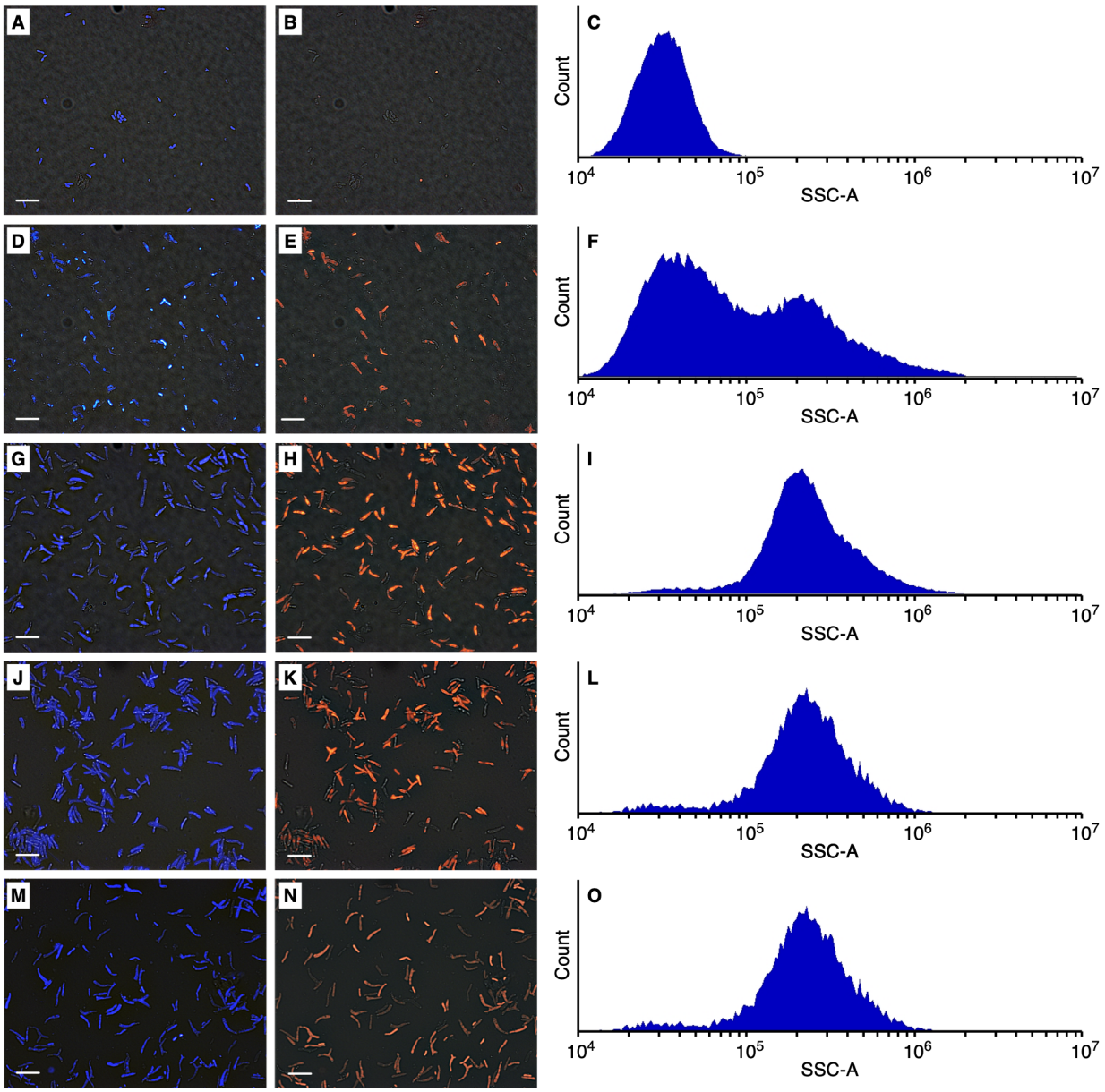


Figure S11. Morphology of *E. meliloti* FSM-MA cell populations. Free-living *E. meliloti* FSM-MA cells (**A-C**), and *E. meliloti* FSM-MA cells purified from *M. sativa* zone II nodule sections (**D-F**), *M. sativa* zone III nodule sections (**G-I**), *M. sativa* whole nodules (**J-L**), or *M. truncatula* whole nodules (**M-O**). Ploidy data is provided in Figure S12. Micrographs show *E. meliloti* FSM-MA cell populations stained with the DNA binding dyes DAPI (blue) and PI (red). The scale bar represents 10 nm. (**A**,**D**,**G**,**J**,**M**) DAPI fluorescence overlaid with DIC (differential interference contrast) images. (**B**,**E**,**H**,**K**,**N**) PI fluorescence overlaid with DIC images. (**C**,**F**,**I**,**L**,**O**) Histograms summarizing the distribution of flow cytometry side scattering values of heat-killed *E. meliloti* FSM-MA populations. Graphs are based on 50,000 cells.

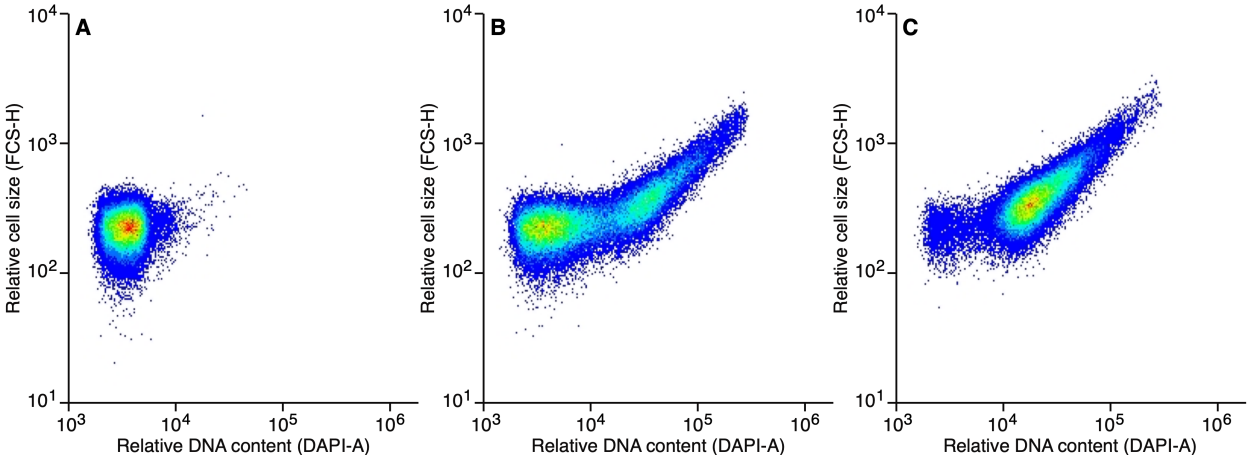
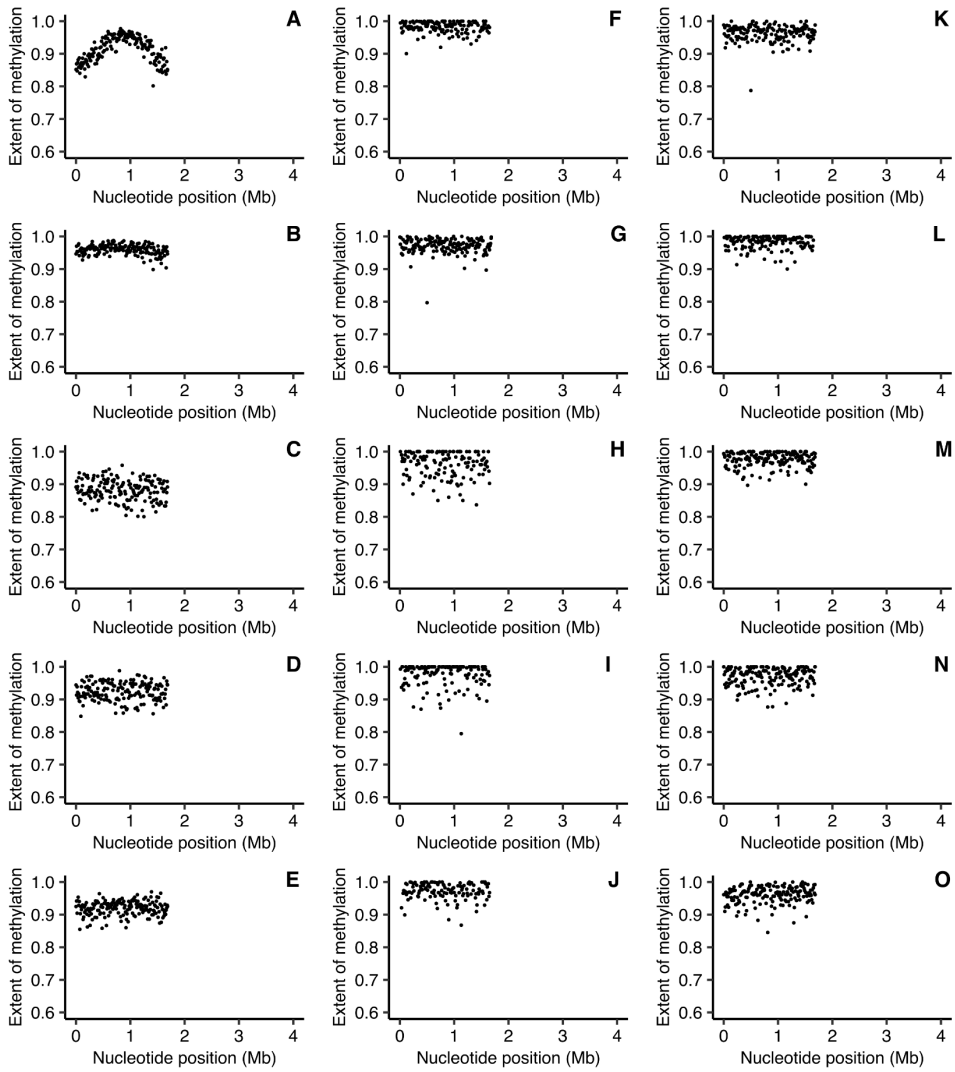


Figure S12. Cell size and DNA content of *E. meliloti* FSM-MA cell populations. Pseudo-coloured scatterplots displaying the DNA content (X-axis) and size (Y-axis) of various *E. meliloti* FSM-MA populations, based on flow cytometry readings of 50,000 cells. The colour scheme indicates the number of values plotted at a given location of the graph, with blue to red indicating lower to higher density. Free-living *E. meliloti* FSM-MA cells (**A**), *E. meliloti* FSM-MA cells purified from *M. sativa* zone II nodule sections (**B**), and *E. meliloti* FSM-MA cells purified from *M. sativa* zone III nodule sections (**C**).



929
930
931 **Figure S13. pSymB-wide DNA methylation of *E. meliloti* Rm2011 bacteroids.** The extent of
932 methylation of (**A-E**) GANTC, (**F-J**) CGCA(N_5)GTG, and (**K-O**) RCGCCTC motifs across the
933 *E. meliloti* Rm2011 pSymB replicon is shown using a 10 kb sliding window. Averages from three
934 biological replicates are shown for free-living and whole nodule samples; data represents one
935 replicate for the zone II and zone III nodule sections. (**A**,**F**,**K**) Free-living cells harvested in mid-
936 exponential phase. (**B**,**G**,**L**) Free-living cells harvested in early stationary phase. (**C**,**H**,**M**)
937 Bacteroids isolated from *M. sativa* zone II nodule sections. (**D**,**I**,**N**) Bacteroids isolated from *M.*
938 *sativa* zone III nodule sections. (**E**,**J**,**O**) Bacteroids isolated from *M. sativa* whole nodule samples.

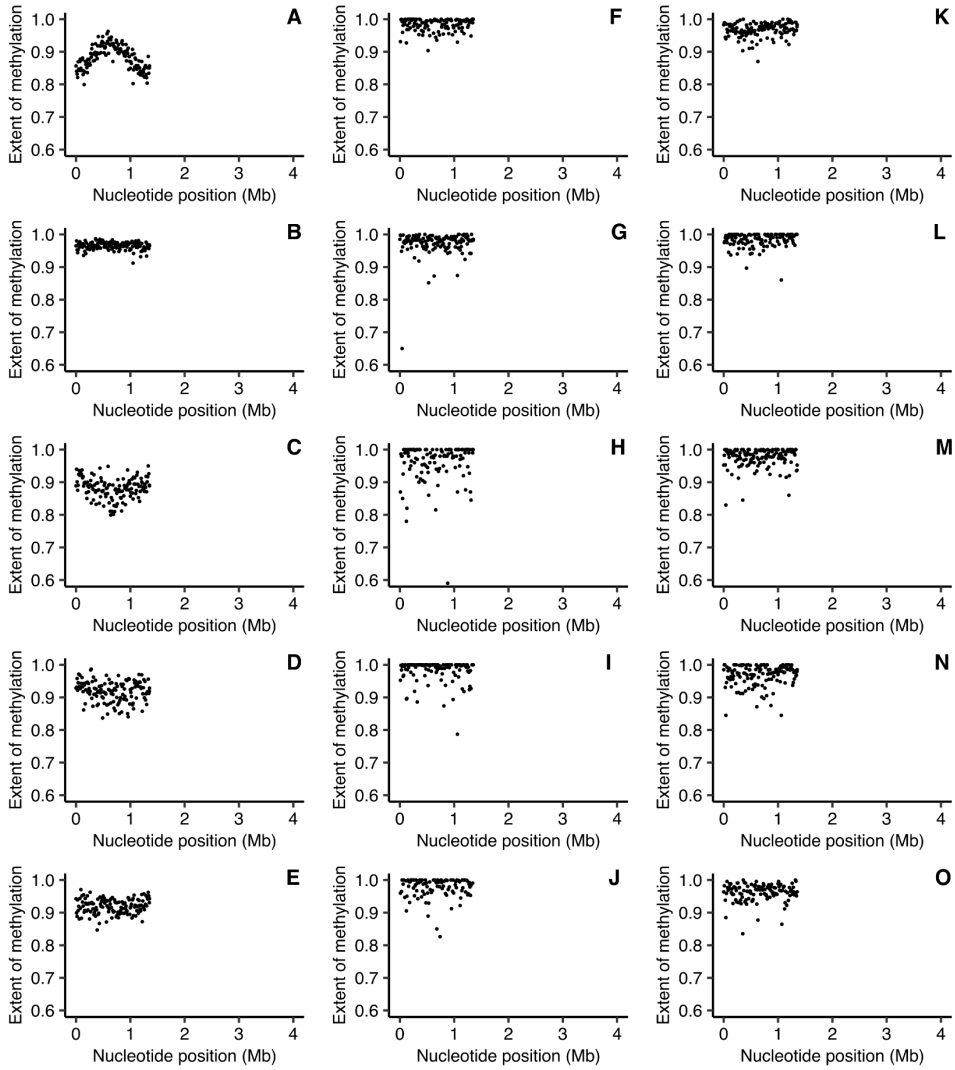


Figure S14. pSymA-wide DNA methylation of *E. meliloti* Rm2011 bacteroids. The extent of methylation of (**A-E**) GANTC, (**F-J**) CGCA(N_5)GTG, and (**K-O**) RCGCCTC motifs across the *E. meliloti* Rm2011 pSymA replicon is shown using a 10 kb sliding window. Averages from three biological replicates are shown for free-living and whole nodule samples; data represents one replicate for the zone II and zone III nodule sections. (**A**,**F**,**K**) Free-living cells harvested in mid-exponential phase. (**B**,**G**,**L**) Free-living cells harvested in early stationary phase. (**C**,**H**,**M**) Bacteroids isolated from *M. sativa* zone II nodule sections. (**D**,**I**,**N**) Bacteroids isolated from *M. sativa* zone III nodule sections. (**E**,**J**,**O**) Bacteroids isolated from *M. sativa* whole nodule samples.

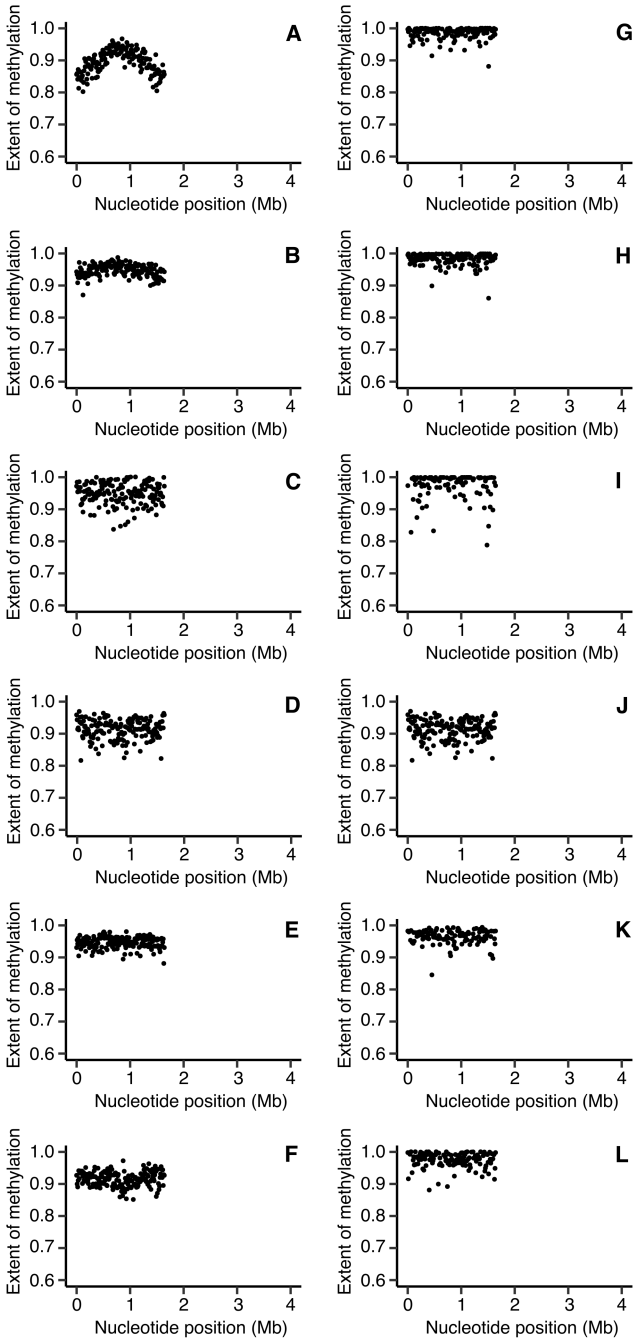
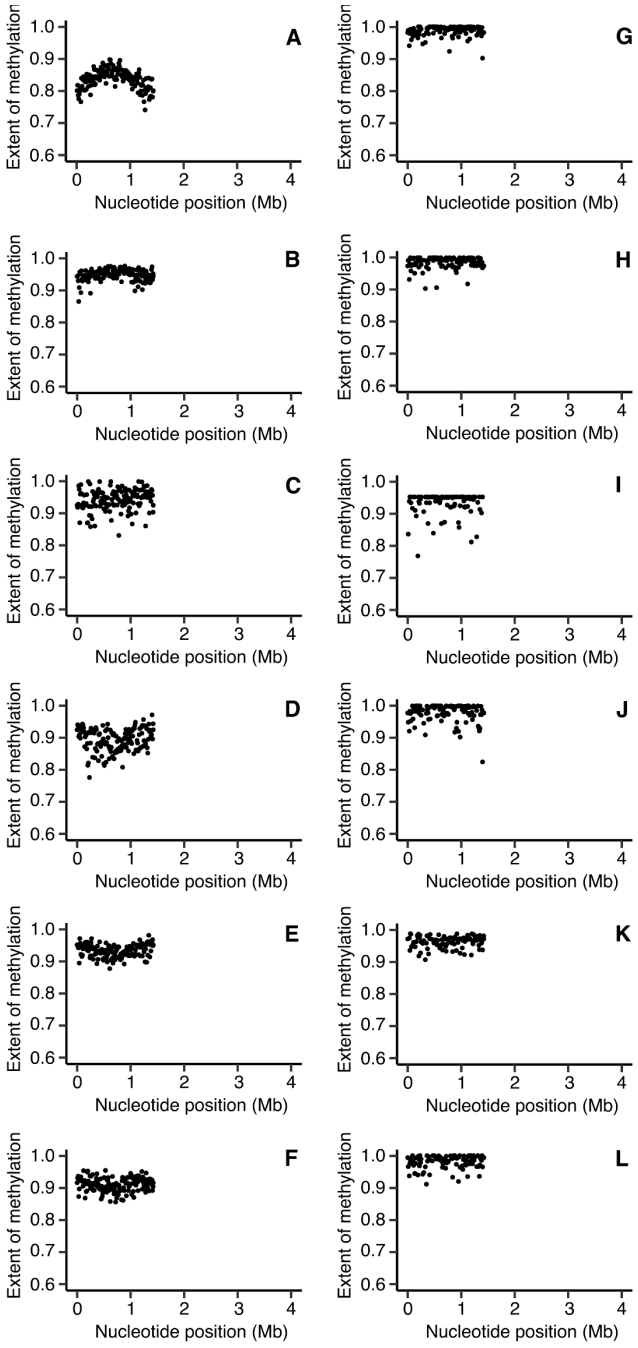


Figure S15. pSymB-wide DNA methylation of *E. meliloti* FSM-MA bacteroids. The extent of methylation of (**A**-**F**) GANTC and (**G**-**L**) AGCT(N_8)AGCT motifs across the *E. meliloti* FSM-MA pSymB replicon is shown using a 10 kb sliding window. Averages from three biological replicates are shown for free-living and whole nodule samples; data represents one replicate for the zone II and zone III nodule sections. (**A**,**G**) Free-living cells harvested in mid-exponential phase. (**B**,**H**) Free-living cells harvested in early stationary phase. (**C**,**I**) Bacteroids isolated from *M. sativa* zone II nodule sections. (**D**,**J**) Bacteroids isolated from *M. sativa* zone III nodule sections. (**E**,**K**) Bacteroids isolated from *M. sativa* whole nodule samples. (**F**,**L**) Bacteroids isolated from *M. truncatula* whole nodule samples.



960
961
962 **Figure S16. pSymA-wide DNA methylation of *E. meliloti* FSM-MA bacteroids.** The extent of
963 methylation of (**A**-**F**) GANTC and (**G**-**L**) AGCT(N_8)AGCT motifs across the *E. meliloti* FSM-
964 MA pSymA replicon is shown using a 10 kb sliding window. Averages from three biological
965 replicates are shown for free-living and whole nodule samples; data represents one replicate for
966 the zone II and zone III nodule sections. (**A**,**G**) Free-living cells harvested in mid-exponential
967 phase. (**B**,**H**) Free-living cells harvested in early stationary phase. (**C**,**I**) Bacteroids isolated from
968 *M. sativa* zone II nodule sections. (**D**,**J**) Bacteroids isolated from *M. sativa* zone III nodule
969 sections. (**E**,**K**) Bacteroids isolated from *M. sativa* whole nodule samples. (**F**,**L**) Bacteroids
970 isolated from *M. truncatula* whole nodule samples.

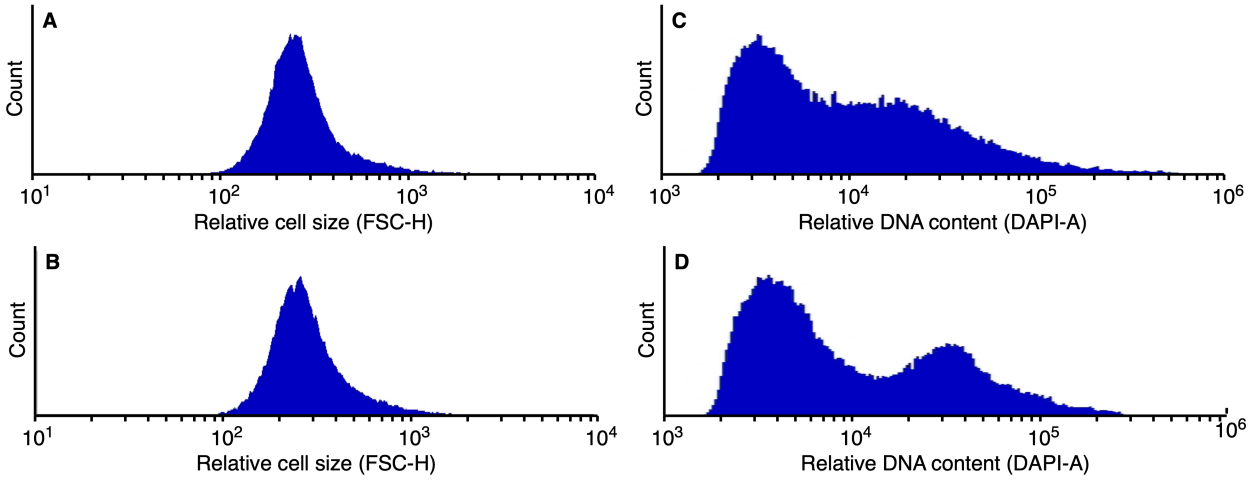


Figure S17. Cell size and DNA content of *E. meliloti* zone II bacteroids. (**A**-**B**) Histograms summarizing the distribution of flow cytometry side scattering values of heat-killed *E. meliloti* Rm2011 (**A**) and *E. meliloti* FSM-MA (**B**) zone II bacteroid populations. (**C**-**D**) Histograms summarizing the distribution of flow cytometry DAPI fluorescence values of heat-killed *E. meliloti* Rm2011 (**C**) and *E. meliloti* FSM-MA (**D**) zone II bacteroid populations. Graphs are based on 50,000 cells.

979 **SUPPLEMENTARY REFERENCES**
980

981 1. Wendt T, Doohan F, Mullins E. 2012. Production of *Phytophthora infestans*-resistant potato
982 (*Solanum tuberosum*) utilising *Ensifer adhaerens* OV14. Transgenic Res 21:567–578.

983 2. Trinick MJ. 1980. Relationships amongst the fast-growing rhizobia of *Lablab purpureus*,
984 *Leucaena leucocephala*, *Mimosa* spp., *Acacia farnesiana* and *Sesbania grandiflora* and their
985 affinities with other rhizobial groups. J Appl Bacteriol 49:39–53.

986 3. Nagymihaly M, Vásárhelyi BM, Barrière Q, Chong T-M, Bálint B, Bihari P, Hong K-W,
987 Horváth B, Ibijbijen J, Amar M, Farkas A, Kondorosi E, Chan K-G, Gruber V, Ratet P,
988 Mergaert P, Kereszt A. 2017. The complete genome sequence of *Ensifer meliloti* strain
989 CCMM B554 (FSM-MA), a highly effective nitrogen-fixing microsymbiont of *Medicago*
990 *truncatula* Gaertn. Stand Genomic Sci 12:75.

991 4. diCenzo GC, MacLean AM, Milunovic B, Golding GB, Finan TM. 2014. Examination of
992 prokaryotic multipartite genome evolution through experimental genome reduction. PLoS
993 Genet 10:e1004742.

994 5. diCenzo GC, Zamani M, Milunovic B, Finan TM. 2016. Genomic resources for identification
995 of the minimal N_2-fixing symbiotic genome. Environ Microbiol 18:2534–2547.

996 6. De Nisco NJ, Abo RP, Wu CM, Penterman J, Walker GC. 2014. Global analysis of cell cycle
997 gene expression of the legume symbiont *Sinorhizobium meliloti*. Proc Natl Acad Sci USA
998 111:3217–3224.

999 7. Wright R, Stephens C, Zweiger G, Shapiro L, Alley MR. 1996. *Caulobacter* Lon protease has
1000 a critical role in cell-cycle control of DNA methylation. Genes Dev 10:1532–1542.

1001 8. Li S, Brazhnik P, Sobral B, Tyson JJ. 2008. A quantitative study of the division cycle of
1002 *Caulobacter crescentus* stalked cells. PLoS Comput Biol 4:e9.

1003 9. Reyes-Lamothe R, Possoz C, Danilova O, Sherratt DJ. 2008. Independent positioning and
1004 action of *Escherichia coli* replisomes in live cells. Cell 133:90–102.

1005 10. Frage B, Döhlemann J, Robledo M, Lucena D, Sobetzko P, Graumann PL, Becker A. 2016.
1006 Spatiotemporal choreography of chromosome and megaplasmids in the *Sinorhizobium*
1007 *meliloti* cell cycle. Mol Microbiol 100:808–823.

1008 11. Mohapatra SS, Fioravanti A, Biondi EG. 2014. DNA methylation in *Caulobacter* and other
1009 Alphaproteobacteria during cell cycle progression. Trends Microbiol 22:528–535.

1010 12. Fagorzi C, Ilie A, Decorosi F, Cangioli L, Viti C, Mengoni A, diCenzo GC. 2020. Symbiotic
1011 and nonsymbiotic members of the genus *Ensifer* (syn. *Sinorhizobium*) are separated into two
1012 clades based on comparative genomics and high-throughput phenotyping. Genome Biol Evol
1013 12:2521–2534.