

Supplementary Material for

DNA methylation in *Ensifer* species during free-living growth and during nitrogen-fixing symbiosis with *Medicago* spp.

George C. diCenzo, Lisa Cangioli, Quentin Nicoud, Janis H.T. Cheng, Matthew J. Blow, Nicole Shapiro, Tanja Woyke, Emanuele G Biondi, Benoît Alunni, Alessio Mengoni, and Peter Mergaert

George diCenzo

Email: george.dicenzo@queensu.ca

This PDF file includes:

Figures S1 to S22

Tables S1 to S8

Legends for Datasets S1 to S3

Supplementary References

Other supplementary materials for this manuscript include the following:

Datasets S1, S2, and S2

Table S1. Putative methyltransferases in the *Ensifer meliloti* pangenome and their distribution.

Ortholog group	Annotation	Number of strains	Strains (locus tag)
1	Cell cycle-regulated methyltransferase CcrM	20	AK83 (Sinme_0634), B399 (BWO90_18765), B401 (BWO76_18415), BL225C (SinmeB_0539), FSM-MA (SMB554_04735), GR4 (C770_GR4Chr0926), HM006 (CDO22_06835), KH35c (CDO23_12830), KH46 (CDO24_07725), M162 (CDO25_05105), M270 (CDO26_11555), RMO17 (DU99_04775), RU11/001 (SMRU11_19915), Rm1021 (SMc00021), Rm41 (CDO27_16945), SM11 (SM11_chr0582), T073 (CDO28_05930), USDA1021 (CDO29_08105), USDA1106 (CDO30_03265), USDA1157 (CDO31_02955)
2	DNA (cytosine-5-)-methyltransferase	2	RU11/001 (SMRU11_31240), SM11 (SM11_chr0923)
3	Modification methylase	1	M270 (CDO26_10750)
4	DNA (cytosine-5-)-methyltransferase	2	USDA1106 (CDO30_09145), Rm1021 (SMc03763)
5	DNA (cytosine-5-)-methyltransferase	1	M270 (CDO26_06525)
6	Modification methylase	1	RU11/001 (SMRU11_01595)
7	DNA methyltransferase	1	T073 (CDO28_02045)
8	DNA modification methylase	2	RU11/001 (SMRU11_31210), SM11 (SM11_chr0917)
9	DNA methylase N-4/N-6	1	AK83 (Sinme_2322)
10	Modification methylase	1	FSM-MA (SMB554_07065)
11	Site-specific DNA-methyltransferase	1	HM006 (CDO22_08570)
12	Site-specific DNA-methyltransferase	1	M270 (CDO26_04310)
13	DNA (cytosine-5-)-methyltransferase	1	RU11/001 (SMRU11_29360)
14	Site-specific DNA methylase	1	GR4 (C770_GR4pA023)
15	DNA cytosine methyltransferase	1	KH46 (CDO24_06020)
16	Modification methylase	1	KH46 (CDO24_13835)
17	DNA cytosine methyltransferase	1	HK46 (CDO24_34910)
18	Modification methylase	1	M270 (CDO26_04310)
19	DNA methylase	1	RU11/001 (SMRU11_01860)
20	SAM-dependent DNA methyltransferase	1	USDA1106 (CDO30_04960), Rm1021 (SMc02296)
21	N-6 DNA methylase	1	FSM-MA (SMB554_16155)
22	Type I restriction-modification system methyltransferase subunit	1	GR4 (C770_GR4Chr0590)
23	N-6 DNA methylase	1	KH35c (CDO23_13780)
24	SAM-dependent DNA methyltransferase	1	M162 (CDO25_03260)

Table S2. 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
<i>E. meliloti</i> FSM-MA	MM9-succinate 1	GCA_002215195.1	2,414,122,786	SAMN12793050
<i>E. meliloti</i> FSM-MA	MM9-succinate 2	GCA_002215195.1	1,973,904,434	SAMN12793077
<i>E. meliloti</i> FSM-MA	MM9-succinate 3	GCA_002215195.1	2,276,105,908	SAMN12793051
<i>E. meliloti</i> FSM-MA	MM9-succinate stationary 1	GCA_002215195.1	983,326,511	SAMN18104108
<i>E. meliloti</i> FSM-MA	MM9-succinate stationary 2	GCA_002215195.1	1,149,234,588	SAMN18104109
<i>E. meliloti</i> FSM-MA	MM9-succinate stationary 3	GCA_002215195.1	1,112,880,576	SAMN18104110
<i>E. meliloti</i> FSM-MA	<i>M. sativa</i> whole nodules 1	GCA_002215195.1	669,928,620	SAMN12793096
<i>E. meliloti</i> FSM-MA	<i>M. sativa</i> whole nodules 2	GCA_002215195.1	619,883,638	SAMN12793097
<i>E. meliloti</i> FSM-MA	<i>M. sativa</i> whole nodules 3	GCA_002215195.1	2,014,796,418	SAMN12792983
<i>E. meliloti</i> FSM-MA	<i>M. truncatula</i> whole nodules 1	GCA_002215195.1	1,223,643,922	SAMN13167594
<i>E. meliloti</i> FSM-MA	<i>M. truncatula</i> whole nodules 2	GCA_002215195.1	993,061,044	SAMN13168102
<i>E. meliloti</i> FSM-MA	<i>M. truncatula</i> whole nodules 3	GCA_002215195.1	1,038,438,145	SAMN13167746
<i>E. meliloti</i> FSM-MA	<i>M. sativa</i> distal nodule sections 1	GCA_002215195.1	454,620,103	SAMN15739021
<i>E. meliloti</i> FSM-MA	<i>M. sativa</i> proximal nodule sections 1	GCA_002215195.1	2,644,422,028	SAMN14511029
<i>E. meliloti</i> FSM-MA	<i>M. truncatula</i> A17 whole nodules 1	GCA_002215195.1	900,330,949	SAMN19249705
<i>E. meliloti</i> FSM-MA	<i>M. truncatula dnf1</i> mutant whole nodules 1	GCA_002215195.1	1,422,524,809	SAMN19249700
<i>E. meliloti</i> FSM-MA	<i>M. truncatula dnf2</i> mutant whole nodules 1	GCA_002215195.1	668,179,310	SAMN19249701
<i>E. meliloti</i> FSM-MA	<i>M. truncatula dnf4</i> mutant whole nodules 1	GCA_002215195.1	1,390,180,451	SAMN19249702
<i>E. meliloti</i> FSM-MA	<i>M. truncatula dnf5</i> mutant whole nodules 1	GCA_002215195.1	3,414,132,613	SAMN19249703
<i>E. meliloti</i> FSM-MA	<i>M. truncatula dnf7</i> mutant whole nodules 1	GCA_002215195.1	1,369,820,112	SAMN19249704
<i>E. meliloti</i> Rm2011	MM9-succinate 1	GCA_000346065.1	1,859,862,543	SAMN12793063
<i>E. meliloti</i> Rm2011	MM9-succinate 2	GCA_000346065.1	2,307,929,125	SAMN12793100
<i>E. meliloti</i> Rm2011	MM9-succinate 3	GCA_000346065.1	2,174,907,026	SAMN12793065
<i>E. meliloti</i> Rm2011	MM9-succinate stationary 1	GCA_000346065.1	1,303,384,501	SAMN18104105
<i>E. meliloti</i> Rm2011	MM9-succinate stationary 2	GCA_000346065.1	1,223,856,206	SAMN18104106
<i>E. meliloti</i> Rm2011	MM9-succinate stationary 3	GCA_000346065.1	1,217,560,804	SAMN18104107
<i>E. meliloti</i> Rm2011	MM9-sucrose 1	GCA_000346065.1	1,302,160,101	SAMN12793017
<i>E. meliloti</i> Rm2011	MM9-sucrose 2	GCA_000346065.1	1,673,395,654	SAMN12793016
<i>E. meliloti</i> Rm2011	MM9-sucrose 3	GCA_000346065.1	2,965,562,894	SAMN12793099
<i>E. meliloti</i> Rm2011	<i>M. sativa</i> whole nodules 1	GCA_000346065.1	814,031,492	SAMN12793035
<i>E. meliloti</i> Rm2011	<i>M. sativa</i> whole nodules 2	GCA_000346065.1	798,075,475	SAMN12793034
<i>E. meliloti</i> Rm2011	<i>M. sativa</i> whole nodules 3	GCA_000346065.1	1,138,170,578	SAMN12793018
<i>E. meliloti</i> Rm2011	<i>M. sativa</i> distal nodule sections 1	GCA_000346065.1	6,610,799,406	SAMN16773451
<i>E. meliloti</i> Rm2011	<i>M. sativa</i> proximal nodule sections 2	GCA_000346065.1	1,544,246,716	SAMN13167882
<i>E. meliloti</i> RmP3496	MM9-sucrose 1	GCA_000346065.1	1,138,116,503	SAMN12792987
<i>E. meliloti</i> RmP3496	MM9-sucrose 2	GCA_000346065.1	1,328,727,153	SAMN12793078
<i>E. meliloti</i> RmP3496	MM9-sucrose 3	GCA_000346065.1	2,234,156,208	SAMN12793033

<i>E. fredii</i> NGR234	MM9-succinate 1	GCA_000018545.1	1,741,191,577	SAMN12792974
<i>E. fredii</i> NGR234	MM9-succinate 2	GCA_000018545.1	2,484,318,454	SAMN12793102
<i>E. fredii</i> NGR234	MM9-succinate 3	GCA_000018545.1	2,279,238,884	SAMN12793103
<i>E. adhaerens</i> OV14	MM9-succinate 1	GCA_000583045.1	1,883,494,303	SAMN12793104
<i>E. adhaerens</i> OV14	MM9-succinate 2	GCA_000583045.1	2,914,499,268	SAMN12792975
<i>E. adhaerens</i> OV14	MM9-succinate 3	GCA_000583045.1	2,272,958,147	SAMN12793036

Table S3. 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
<i>M. sativa</i> distal nodule sections	0.712	0.883	0.882
<i>M. sativa</i> proximal nodule sections	0.787	0.925	0.918
<i>M. sativa</i> whole nodules	0.791	0.921	0.920
CGCA(N₅)GTG			
Free-living (mid-exponential)	0.978	0.980	0.981
Free-living (stationary)	0.974	0.981	0.982
<i>M. sativa</i> distal nodule sections	0.966	0.959	0.961
<i>M. sativa</i> proximal nodule sections	0.978	0.974	0.983
<i>M. sativa</i> whole nodules	0.973	0.972	0.976

Table S4. 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
<i>M. sativa</i> distal nodule sections	0.906	0.948	0.942
<i>M. sativa</i> proximal nodule sections	0.784	0.915	0.895
<i>M. sativa</i> whole nodules	0.860	0.946	0.936
<i>M. truncatula</i> whole nodules	0.793	0.917	0.911
TCGA(N₈)TCGA			
Free-living (mid-exponential)	0.984	0.984	0.981
Free-living (stationary)	0.980	0.984	0.984
<i>M. sativa</i> distal nodule sections	0.973	0.973	0.980
<i>M. sativa</i> proximal nodule sections	0.978	0.915	0.977
<i>M. sativa</i> whole nodules	0.980	0.964	0.963
<i>M. truncatula</i> whole nodules	0.976	0.974	0.981

Table S5. Average extent of methylation of GANTC motifs in *E. meliloti* FSM-MA bacteroids purified from *M. truncatula* *dnf* mutant nodules and wild-type *M. truncatula* A17 nodules.

<i>M. truncatula</i> genotype	Chromosome	pSymA	pSymB
GANTC			
<i>dnf1</i>	0.859	0.868	0.874
<i>dnf5</i>	0.956	0.961	0.959
<i>dnf2</i>	0.870	0.939	0.941
<i>dnf7</i>	0.817	0.919	0.923
<i>dnf4</i>	0.814	0.915	0.924
A17	0.804	0.917	0.920
TCGA(N8)TCGA			
<i>dnf1</i>	0.974	0.977	0.970
<i>dnf5</i>	0.984	0.985	0.983
<i>dnf2</i>	0.975	0.978	0.976
<i>dnf7</i>	0.974	0.972	0.974
<i>dnf4</i>	0.971	0.978	0.965
A17	0.973	0.967	0.969

Table S6. Bacterial strains.

Strain	Genotype	Source
<i>Ensifer adhaerens</i> OV14	Wild type OV14; not a nitrogen-fixing legume symbiont	(1)
<i>Ensifer fredii</i> NGR234	Wild type NGR234 <i>rif-I</i> ; Rif ^R	(2)
<i>Ensifer meliloti</i> FSM-MA	Wild type FSM-MA; Cm ^R	(3)
<i>Ensifer meliloti</i> Rm2011	Wild type SU47 <i>str-3</i> ; Sm ^R	Lab collection
<i>Ensifer meliloti</i> RmP3496	Rm2011 lacking pSymA and pSymB; Sm ^R Sp ^R	(4, 5)

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

Table S7. Densities and doubling times of bacterial cultures grown for the isolation of DNA.

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

* Each Replicate 1 sample consisted of two 40 mL cultures (inoculated from the same starter culture) that were combined prior to DNA isolation. The density and doubling time values for this replicate are based on just one of the cultures for each sample.

† All density values are OD_{600nm} measurements.

Table S8. Plant shoot dry weights.

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

* ND: Not determined.

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

Variation in DNA methylation across the genus *Ensifer*



Grown to mid-exponential phase in minimal medium with succinate as the carbon source

Impact of growth stage (actively divided versus non-dividing) on DNA methylation



Grown to **mid-exponential phase** in minimal medium with succinate as the carbon source

Grown to **stationary phase** in minimal medium with succinate as the carbon source

Role of DNA methylation in regulating carbon metabolism



Grown to mid-exponential phase in minimal medium with **succinate** as the carbon source

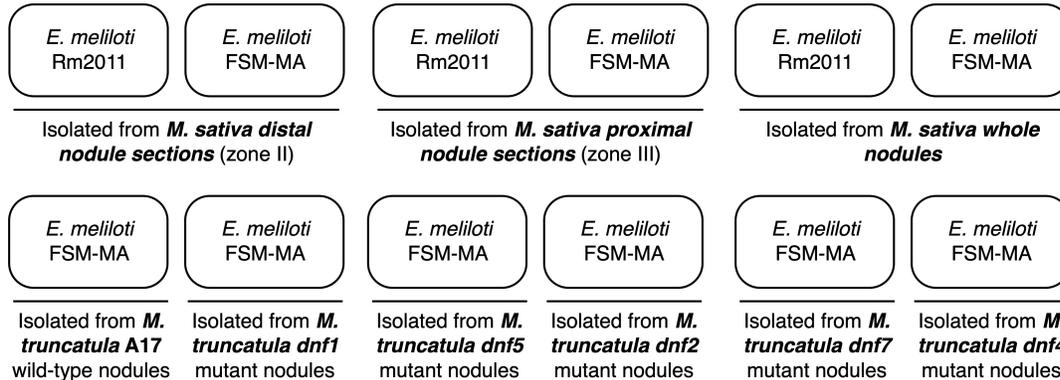
Grown to mid-exponential phase in minimal medium with **sucrose** as the carbon source

Influence of secondary replicons on DNA methylation patterns

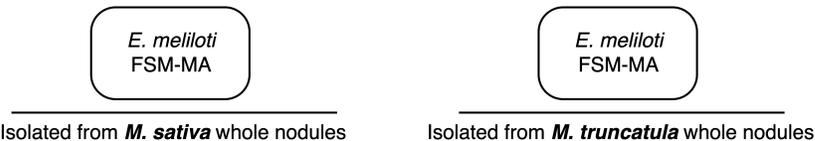


Grown to mid-exponential phase in minimal medium with sucrose as the carbon source

Contributions of DNA methylation changes to regulation of bacteroid differentiation and nitrogen-fixation



Impact of plant host on bacteroid DNA methylation



Isolated from *M. sativa* whole nodules

Isolated from *M. truncatula* whole nodules

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

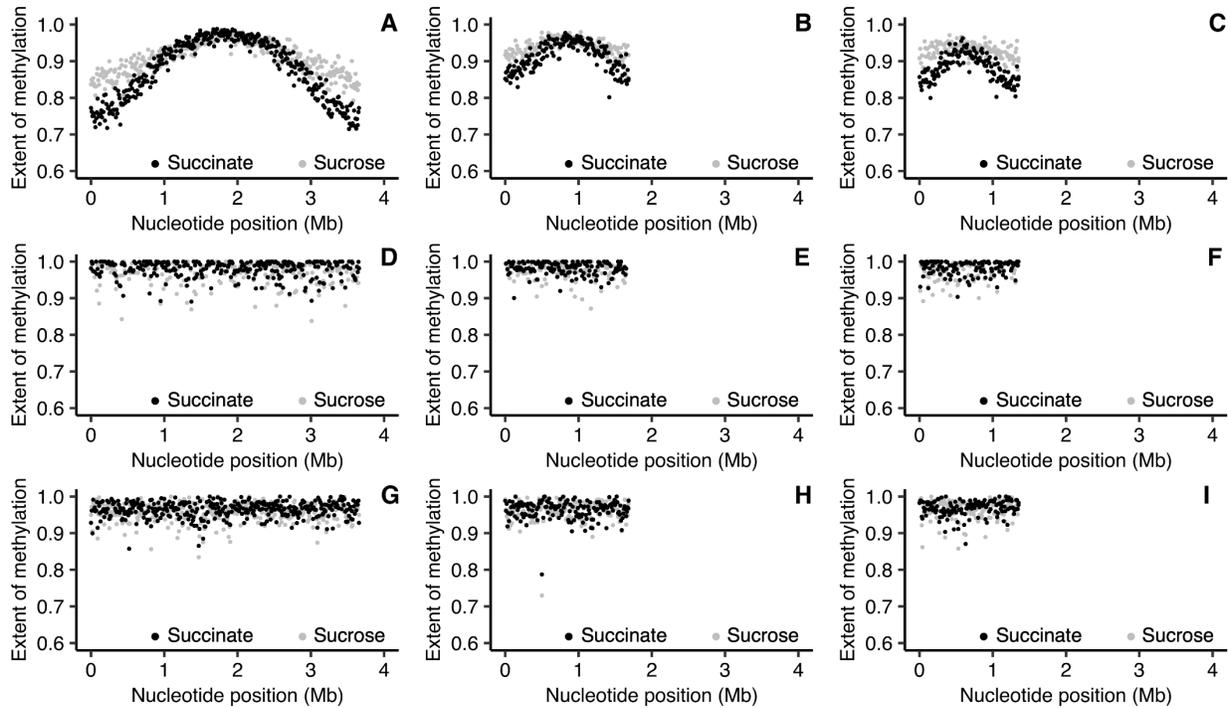


Figure S2. Impact of carbon source on genome-wide DNA methylation of *E. meliloti* Rm2011. The extent of methylation is shown, using a 10 kb sliding window, of cells grown to mid-exponential phase and provided succinate (black) or sucrose (grey) as the sole source of carbon. (A-C) Data for the GANTC motif for the chromosome (A), pSymB (B), and pSymA (C). (D-F) Data for the CGCA(N₅)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).

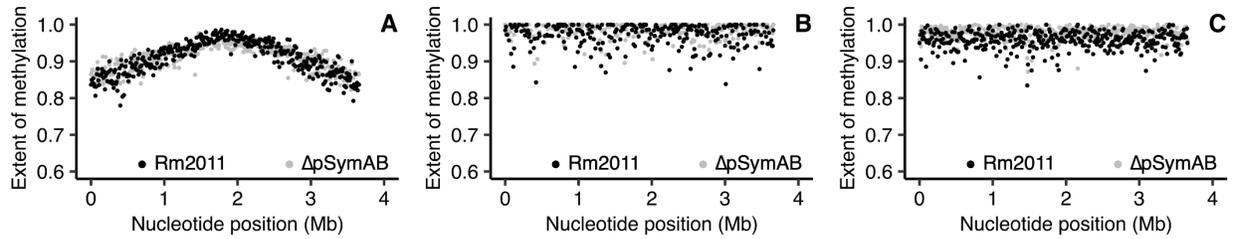


Figure S3. Impact of pSymA and pSymB removal on chromosome-wide DNA methylation of *E. meliloti* Rm2011. The extent of methylation for the chromosome, shown using a 10 kb sliding window, is provided for wild type *E. meliloti* Rm2011 (black) or *E. meliloti* ΔpSymAB (grey) grown to mid-exponential phase and provided sucrose as the sole source of carbon. **(A)** Data for the GANTC motif. **(B)** Data for the CGCA(N₅)GTG motif. **(C)** Data for the RCGCCTC motif.

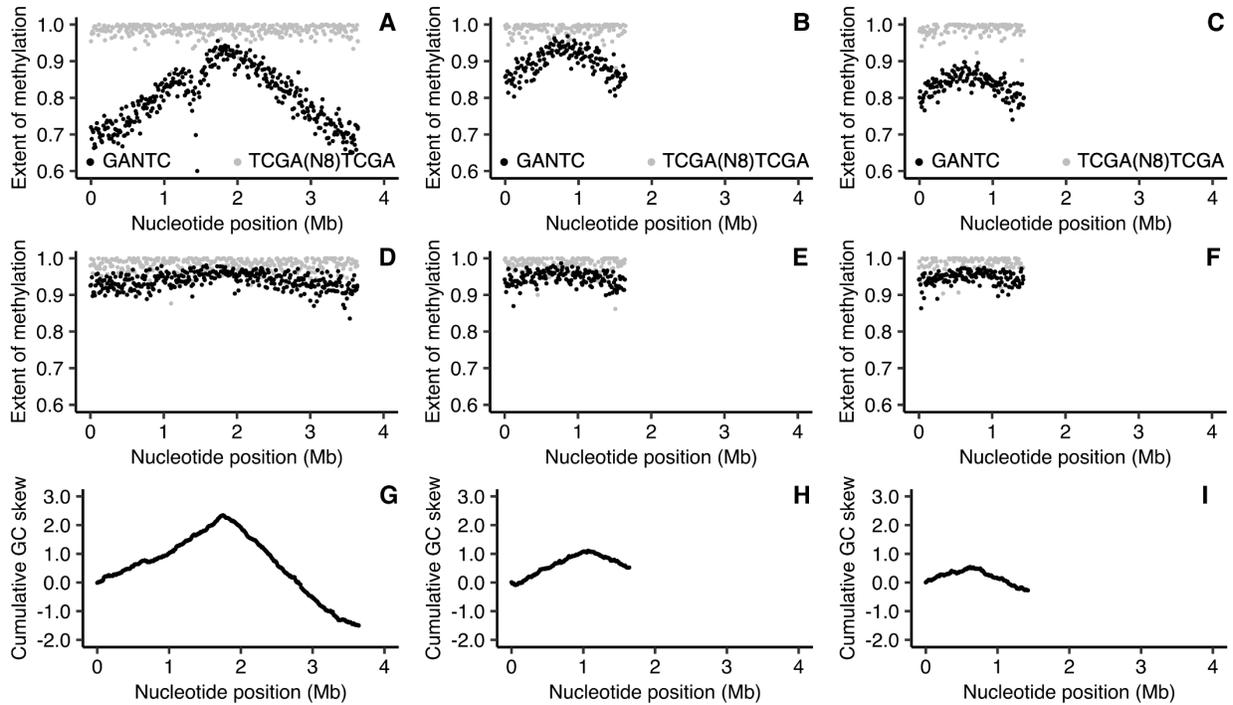


Figure S4. Genome-wide DNA methylation of *E. meliloti* FSM-MA. (A-F) The extent of methylation is shown, using a 10 kb sliding window, of GANTC sites (black) and TCGA(N₈)TCGA sites (grey) across the chromosome (A,D), pSymB (B,E), and pSymA (C,F) replicons of exponential phase (A-C) or early stationary phase (D-F) *E. meliloti* FSM-MA. Averages from three biological replicates are shown. (G-I) Cumulative GC skews are shown, using a 10 kb sliding window, across the *E. meliloti* FSM-MA chromosome (G), pSymB (H), and pSymA (I) replicons.

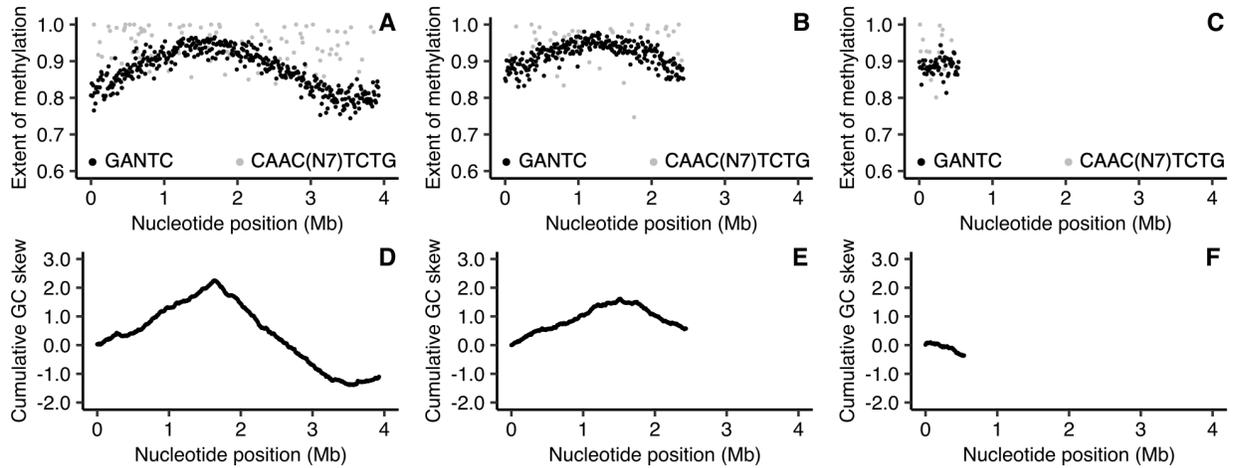


Figure S5. Genome-wide DNA methylation of *E. fredii* NGR234. (A-C) The extent of methylation is shown, using a 10 kb sliding window, of GANTC sites (black) and CAAC(N₇)TCTG sites (grey) across the chromosome (A), pNGR234b (B), and pNGR234a (C) replicons of exponential phase *E. fredii* NGR234. Averages 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. Nucleotide positions refer to the nucleotide positions in the genome assembly files available through NCBI, which does not necessarily correlate with the location of the origin of replication. Instead, the origin of replication and replication terminus regions of each replicon are expected to be represented by the low and high points, respectively, on the cumulative GC skews.

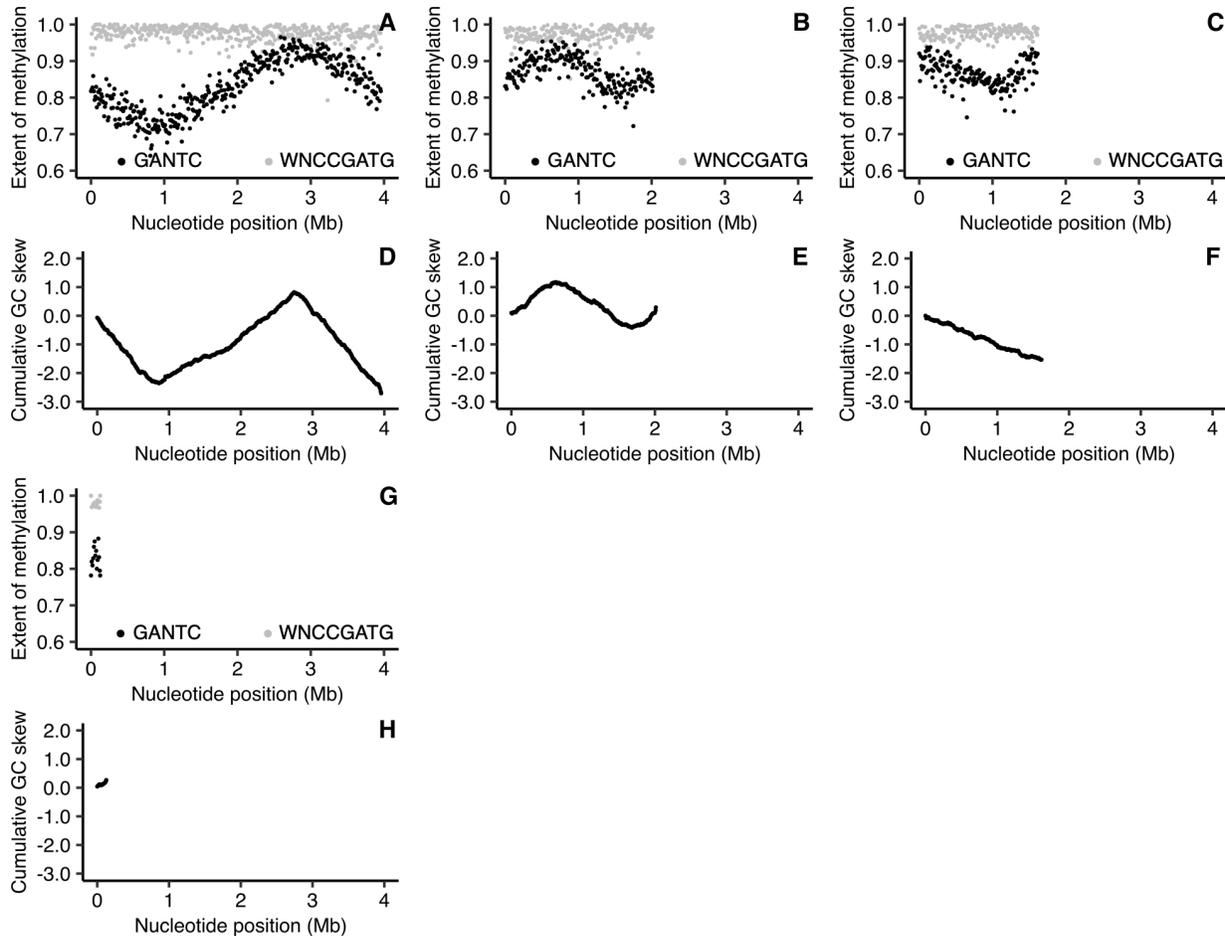


Figure S6. Genome-wide DNA methylation of *E. adhaerens* OV14. (A-C,G) The extent of methylation is shown using a 10 kb sliding window, of GANTC sites (black) and WNCCGATG sites (grey) across chromosome 1 (A), chromosome 2 (B), pOV14b (C), and pOV14c (G) replicons of exponential phase *E. adhaerens* OV14. Averages from three biological replicates are shown. (D-F,H) Cumulative GC skews are shown, using a 10 kb sliding window, across the *E. adhaerens* OV14 chromosome 1 (D), chromosome 2 (E), pOV14b (F), and pOV14c (H) replicons. Nucleotide positions refer to the nucleotide positions in the genome assembly files available through NCBI, which does not necessarily correlate with the location of the origin of replication. Instead, the origin of replication and replication terminus regions of each replicon are expected to be represented by the low and high points, respectively, on the cumulative GC skews.

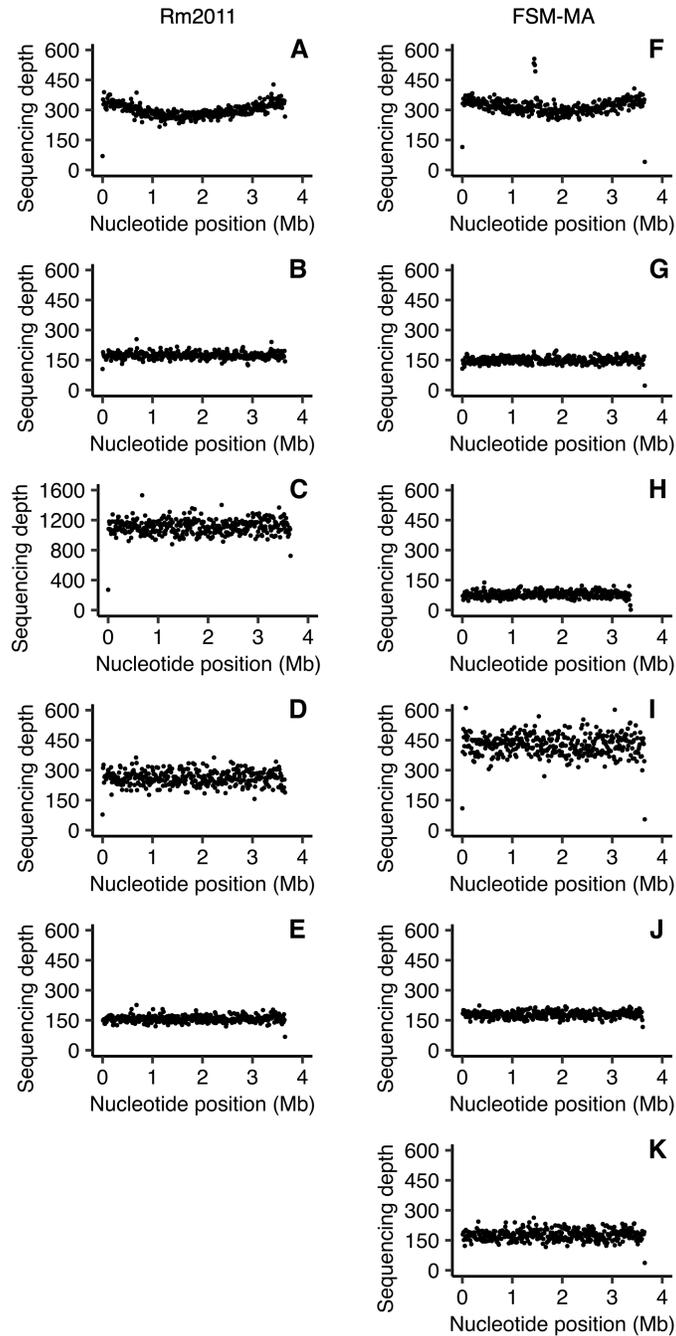


Figure S7. Sequencing depth across the *E. meliloti* RM2011 and FSM-MA chromosome. The sequencing depths (i.e., number of reads mapping to a given nucleotide) of the (A-E) *E. meliloti* Rm2011 chromosome and (F-K) *E. meliloti* FSM-MA chromosome are shown, using a 10 kb sliding window. Averages from 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.

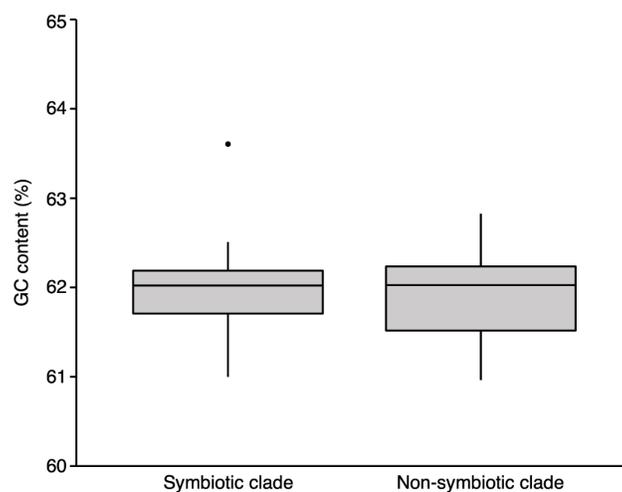


Figure S8. GC content in the genus *Ensifer*. Box plots summarizing the GC content of the genomes of 157 *Ensifer* strains are shown. The monophyletic “symbiotic” and “non-symbiotic” clades as defined previously (12), are represented by 111 and 44 genomes respectively.

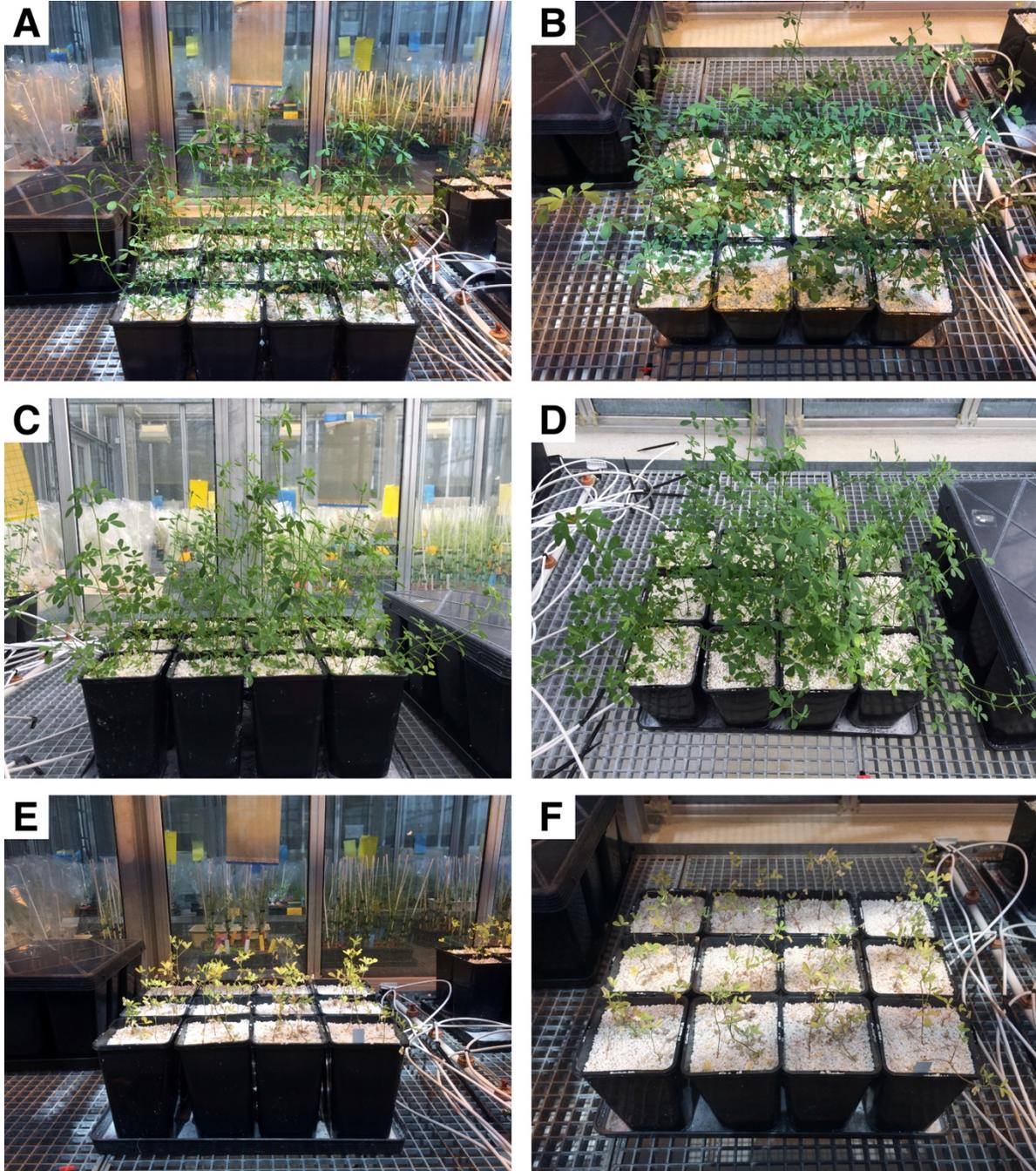


Figure S9. Photos of *M. sativa* plants from which nodules were collected for sectioning. Photographs of representative *M. sativa* plants either inoculated with (A,B) *S. meliloti* Rm2011 or (C,D) *S. meliloti* FSM-MA, or (E,F) uninoculated controls. Nodules from these plants were collected and sectioned to isolate zone II and zone III bacteroid samples.

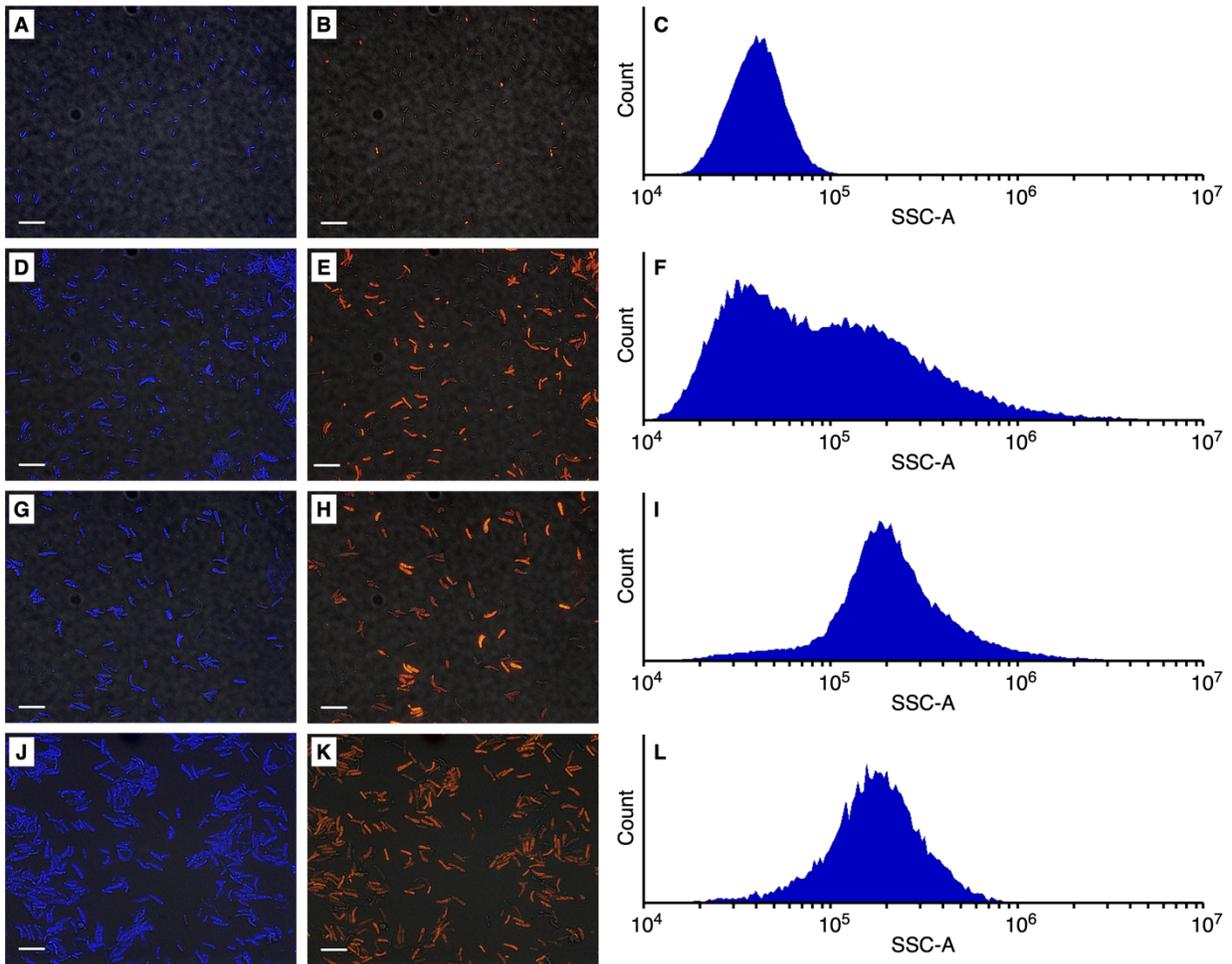


Figure S10. Morphology of *E. meliloti* Rm2011 cell populations. Free-living *E. meliloti* Rm2011 cells (A-C), and *E. meliloti* Rm2011 cells purified from *M. sativa* zone II nodule sections (D-F), *M. sativa* zone III nodule sections (G-I), or *M. sativa* whole nodules (J-L). Ploidy data is provided in Figure S11. Micrographs show *E. meliloti* Rm2011 cell populations stained with the DNA binding dyes DAPI (blue) and PI (red). The scale bar represents 10 μ m. (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.

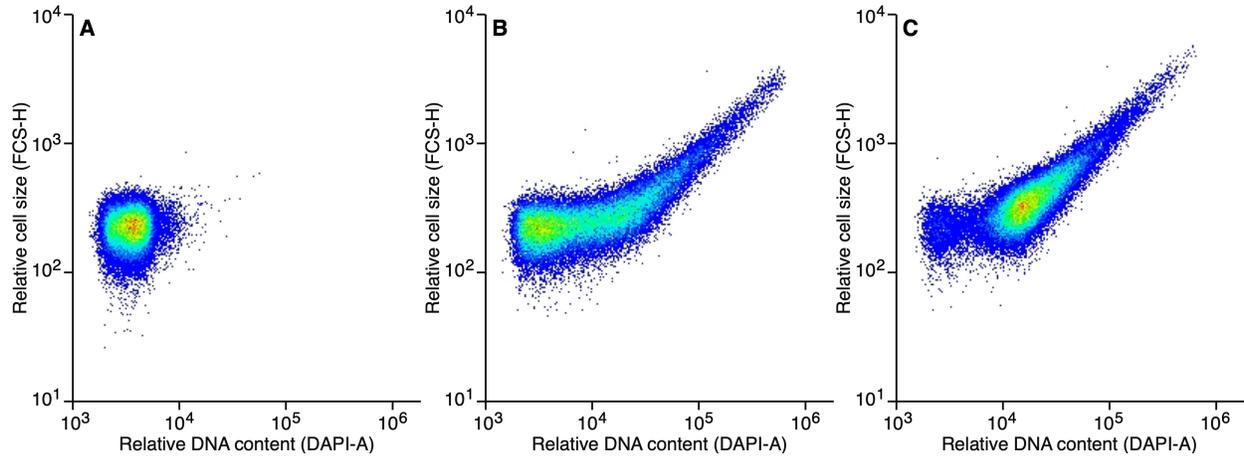


Figure S11. Cell size and DNA content of *E. meliloti* Rm2011 cell populations. Pseudo-coloured scatterplots displaying the DNA content (X-axis) and cell size (Y-axis) of various *E. meliloti* Rm2011 populations, based on flow cytometry readings of 50,000 cells. The colour scheme indicates the number 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).

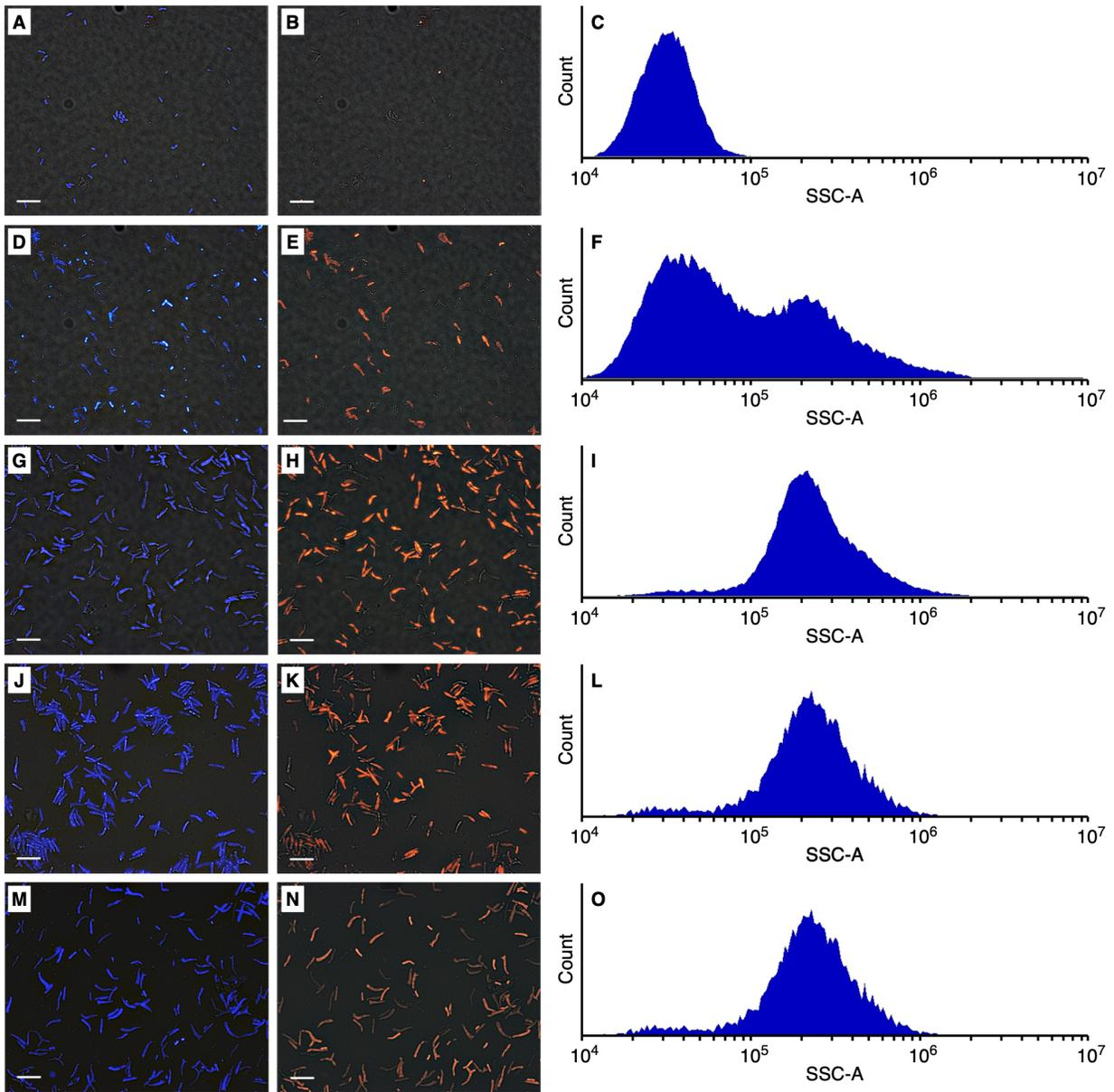


Figure S12. Morphology of *E. meliloti* FSM-MA cell populations. Free-living *E. meliloti* FSM-MA cells (A-C), and *E. meliloti* FSM-MA cells purified from *M. sativa* zone II nodule sections (D-F), *M. sativa* zone III nodule sections (G-I), *M. sativa* whole nodules (J-L), or *M. truncatula* whole nodules (M-O). Ploidy data is provided in Figure S13. Micrographs show *E. meliloti* FSM-MA cell populations stained with the DNA binding dyes DAPI (blue) and PI (red). The scale bar represents 10 μ m. (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.

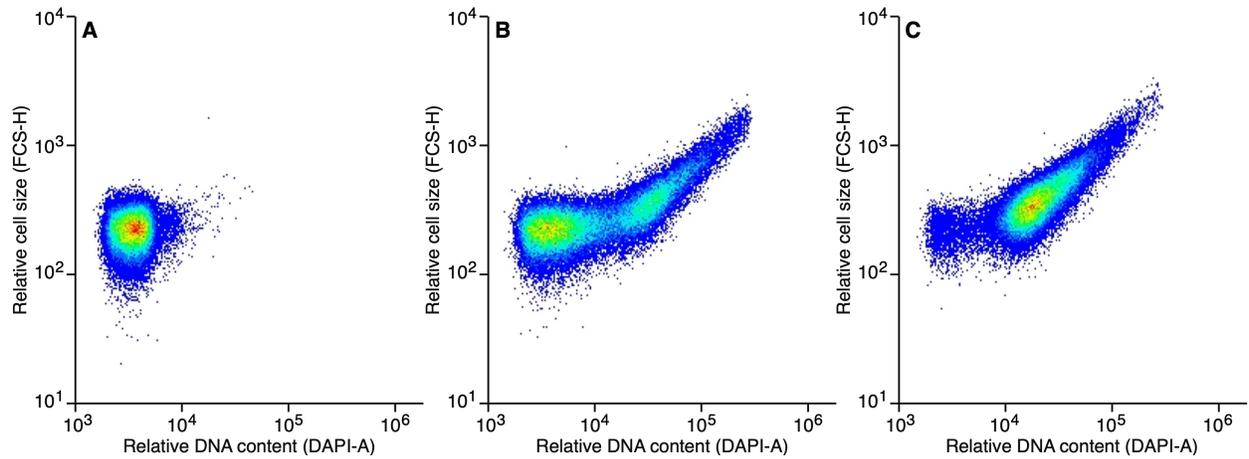


Figure S13. Cell size and DNA content of *E. meliloti* FSM-MA cell populations. Pseudo-coloured scatterplots displaying the DNA content (X-axis) and size (Y-axis) of various *E. meliloti* FSM-MA populations, based on flow cytometry readings of 50,000 cells. The colour scheme indicates the number of values 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).

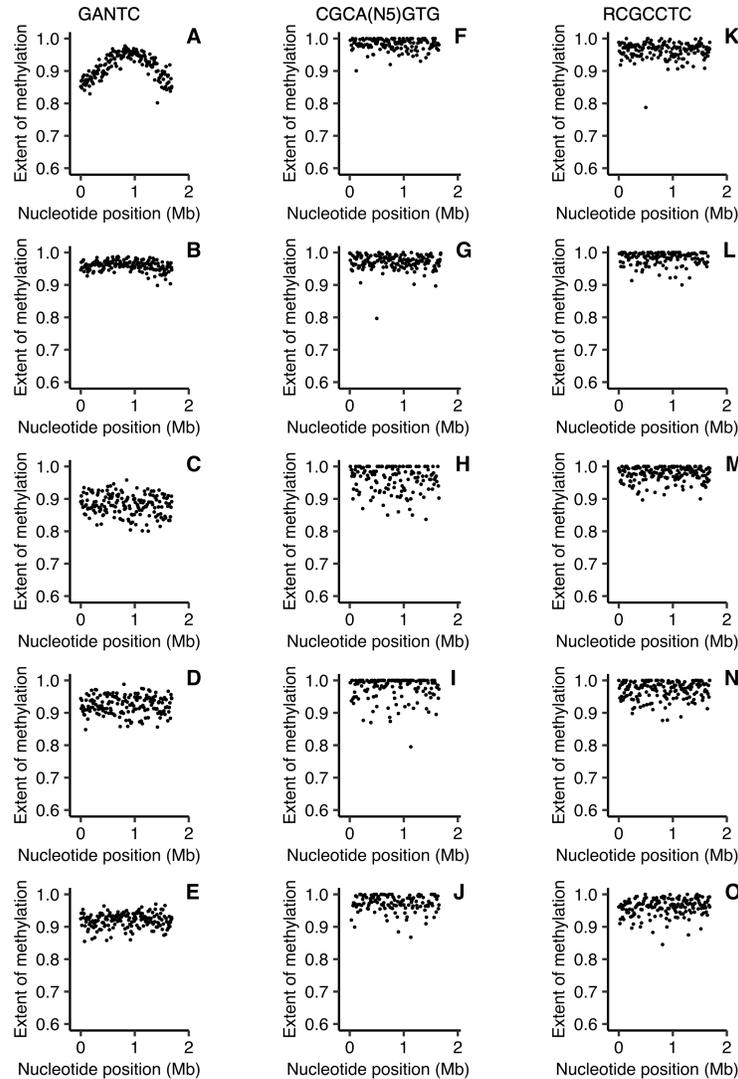


Figure S14. pSymB-wide DNA methylation of *E. meliloti* Rm2011 bacteroids. The extent of methylation of (A-E) GANTC, (F-J) CGCA(N₅)GTG, and (K-O) RCGCCTC motifs across the *E. meliloti* Rm2011 pSymB replicon is shown using a 10 kb sliding window. Averages from three biological replicates are shown for free-living and whole nodule samples; data represents one replicate for the zone II and zone III nodule sections. (A,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.

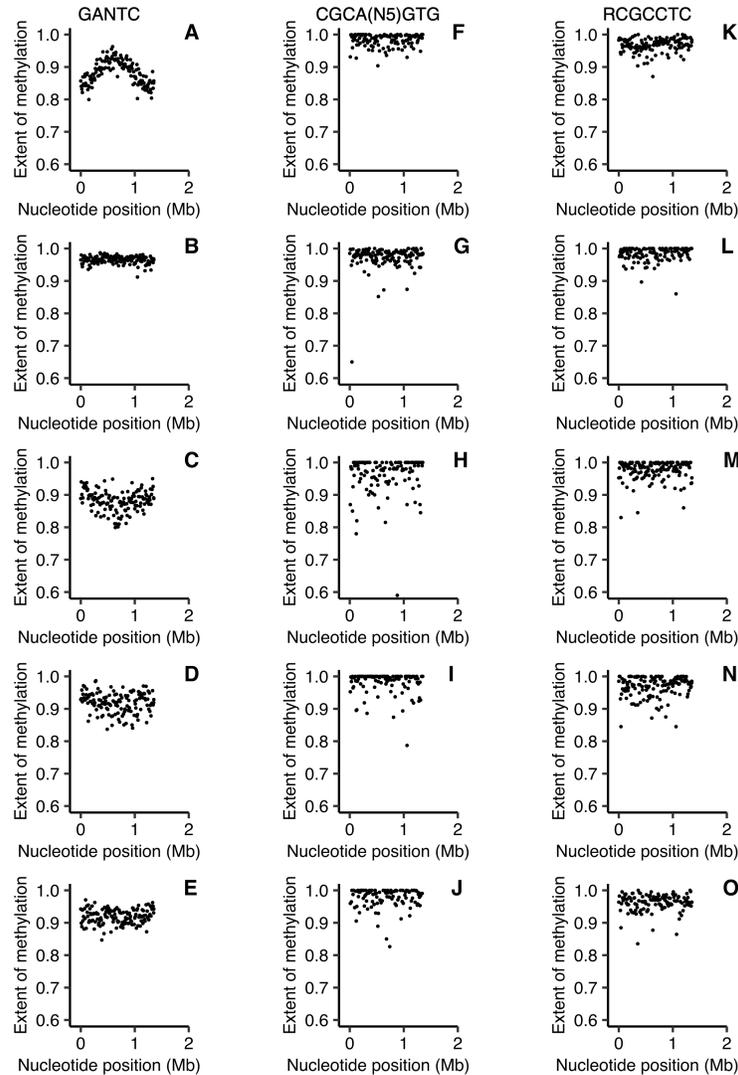


Figure S15. pSymA-wide DNA methylation of *E. meliloti* Rm2011 bacteroids. The extent of methylation of (A-E) GANTC, (F-J) CGCA(N₅)GTG, and (K-O) RCGCCTC motifs across the *E. meliloti* Rm2011 pSymA replicon is shown using a 10 kb sliding window. Averages from three biological replicates are shown for 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.

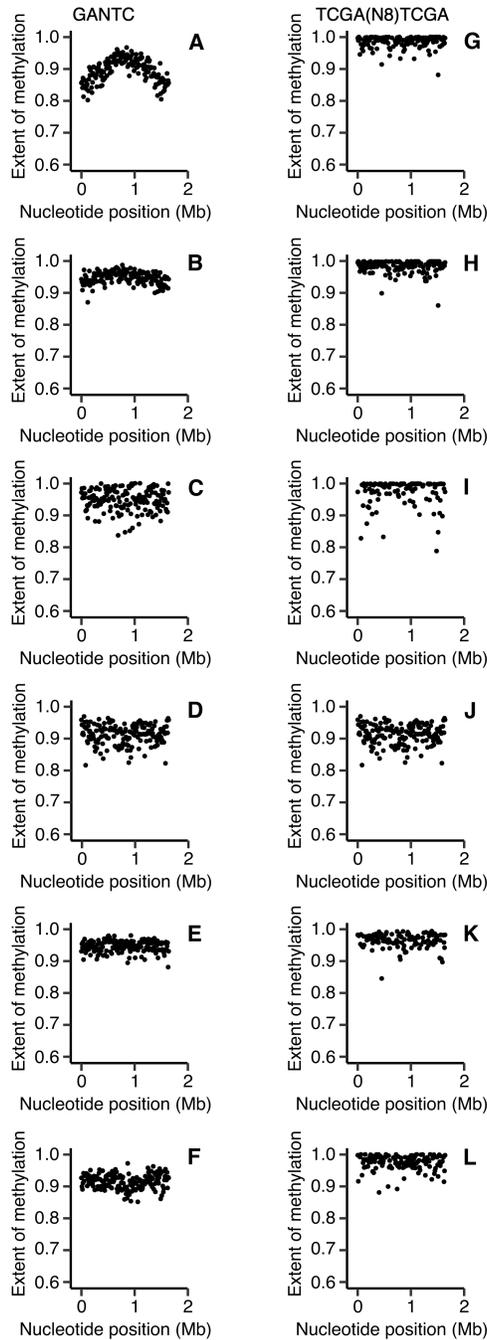


Figure S16. pSymB-wide DNA methylation of *E. meliloti* FSM-MA bacteroids. The extent of methylation of (A-F) GANTC and (G-L) TCGA(N₈)TCGA motifs across the *E. meliloti* FSM-MA pSymB replicon is shown using a 10 kb sliding window. Averages from three biological replicates are shown for free-living and 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.

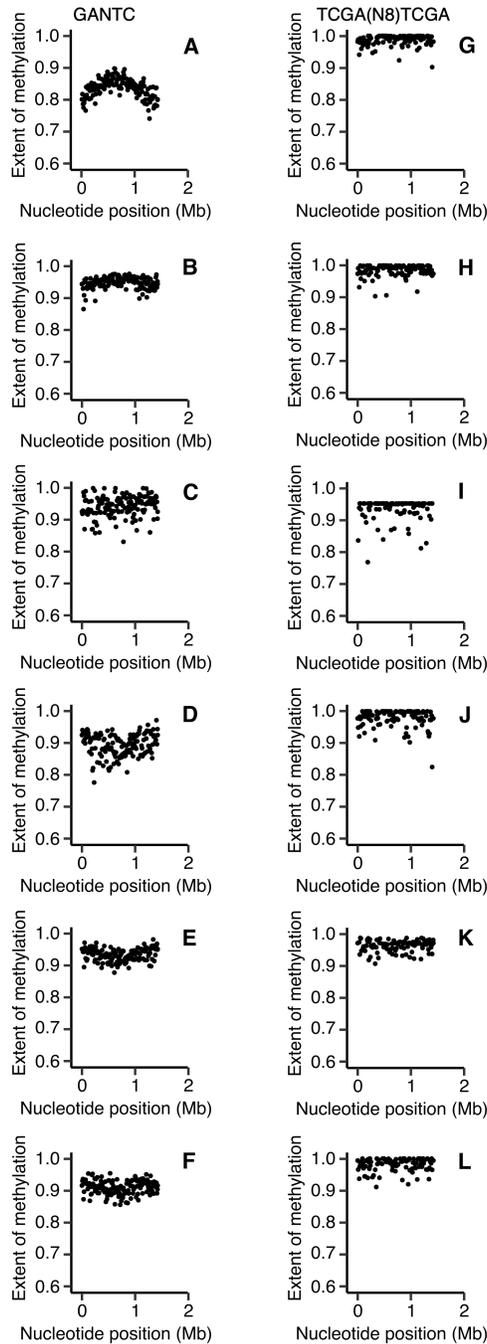


Figure S17. pSymA-wide DNA methylation of *E. meliloti* FSM-MA bacteroids. The extent of methylation of (A-F) GANTC and (G-L) TCGA(N₈)TCGA motifs across the *E. meliloti* FSM-MA pSymA replicon is shown using a 10 kb sliding window. Averages from three biological replicates are shown for free-living and whole nodule samples; data represents one replicate for the zone II and zone III nodule sections. (A,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.

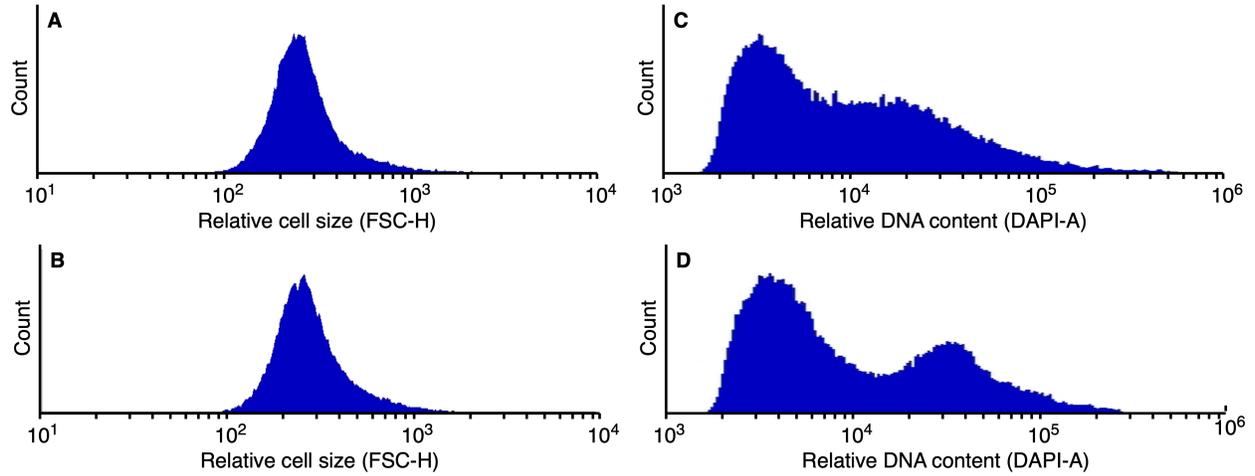


Figure S18. Cell size and DNA content of *E. meliloti* zone II bacteroids. (A-B) Histograms summarizing the distribution of flow cytometry side scattering values of heat-killed *E. meliloti* Rm2011 (A) and *E. meliloti* FSM-MA (B) zone II bacteroid populations. (C-D) Histograms summarizing the distribution 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.

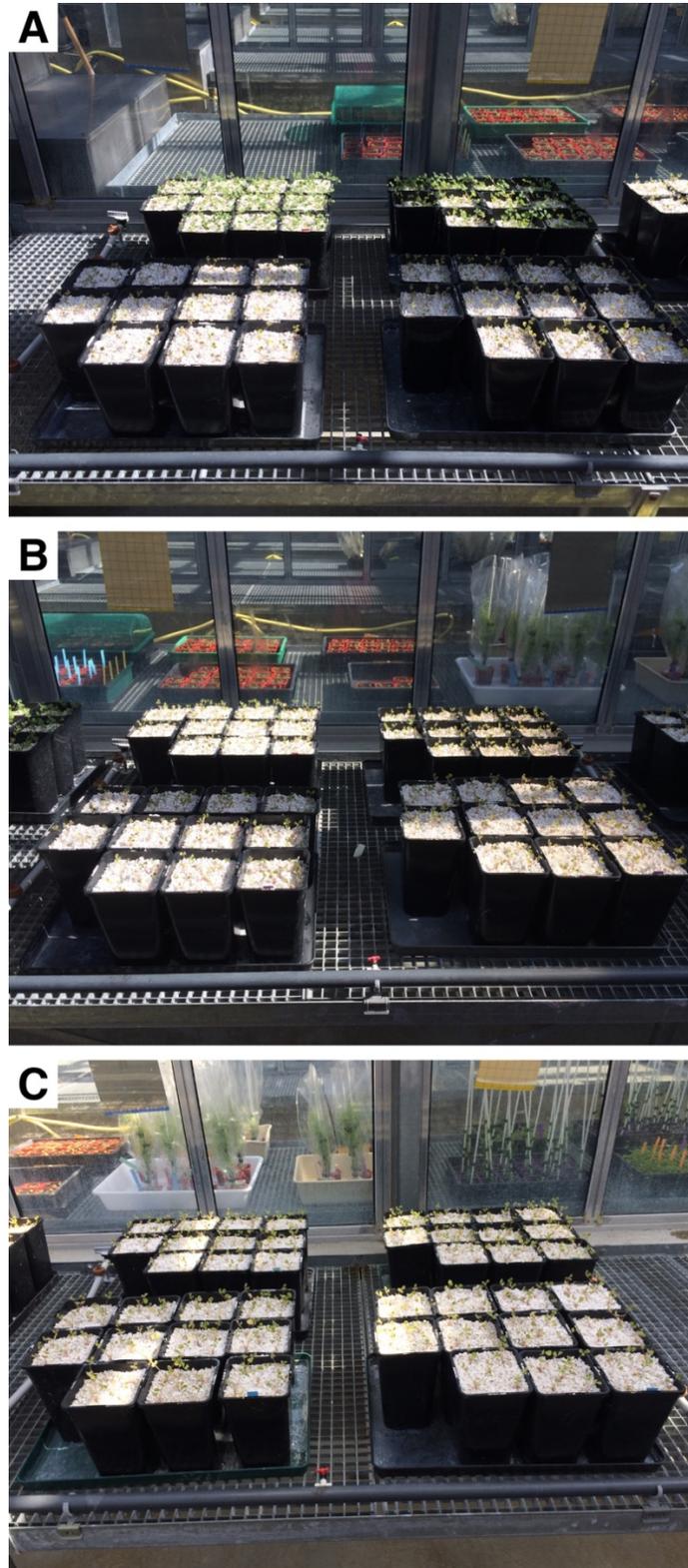


Figure S19. Photos of *M. truncatula* plants grown for nodule collection. Photographs of *M. truncatula* (A) wild type A17 (top) or *dnf1* (bottom), (B) *dnf2* (top) or *dnf4* (bottom), and (C) *dnf5* (top) or *dnf7* (bottom) are shown. Nodules from these plants were collected for bacteroid isolation.

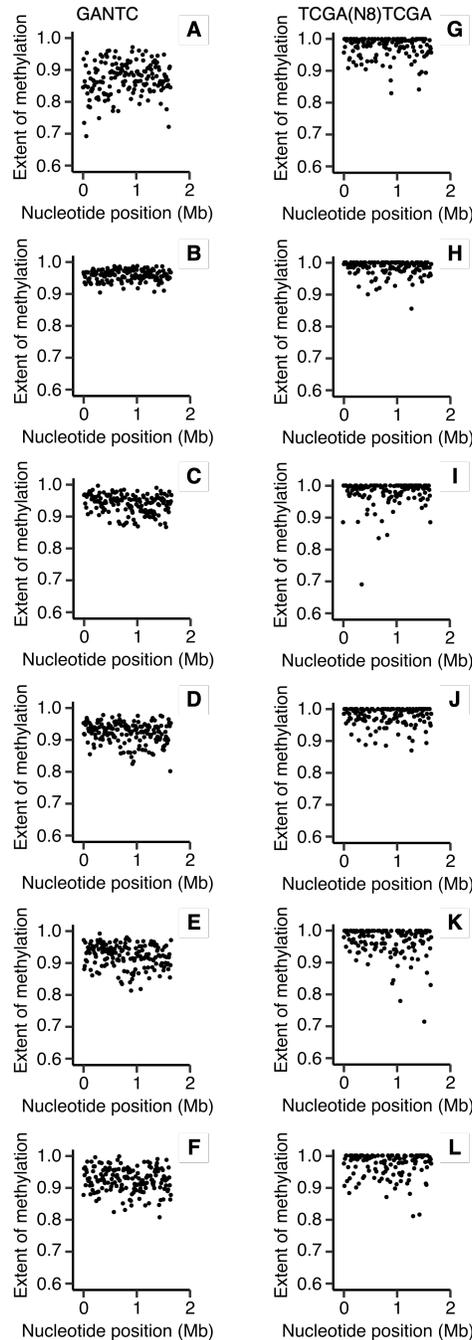


Figure S20. pSymB-wide DNA methylation of *E. meliloti* FSM-MA bacteroids purified from *M. truncatula* *dnf* mutant nodules. The extent of methylation of (A-F) GANTC and (G-L) TCGA(N₈)TCGA motifs across the *E. meliloti* FSM-MA pSymB replicon is shown using a 10 kb sliding window. (A,G) Bacteroids isolated from *M. truncatula* *dnf1* mutant nodules. (B,H) Bacteroids isolated from *M. truncatula* *dnf5* mutant nodules. (C,I) Bacteroids isolated from *M. truncatula* *dnf2* mutant nodules. (D,J) Bacteroids isolated from *M. truncatula* *dnf7* mutant nodules. (E,K) Bacteroids isolated from *M. truncatula* *dnf4* mutant nodules. (F,L) Bacteroids isolated from wild-type *M. truncatula* A17 nodules.

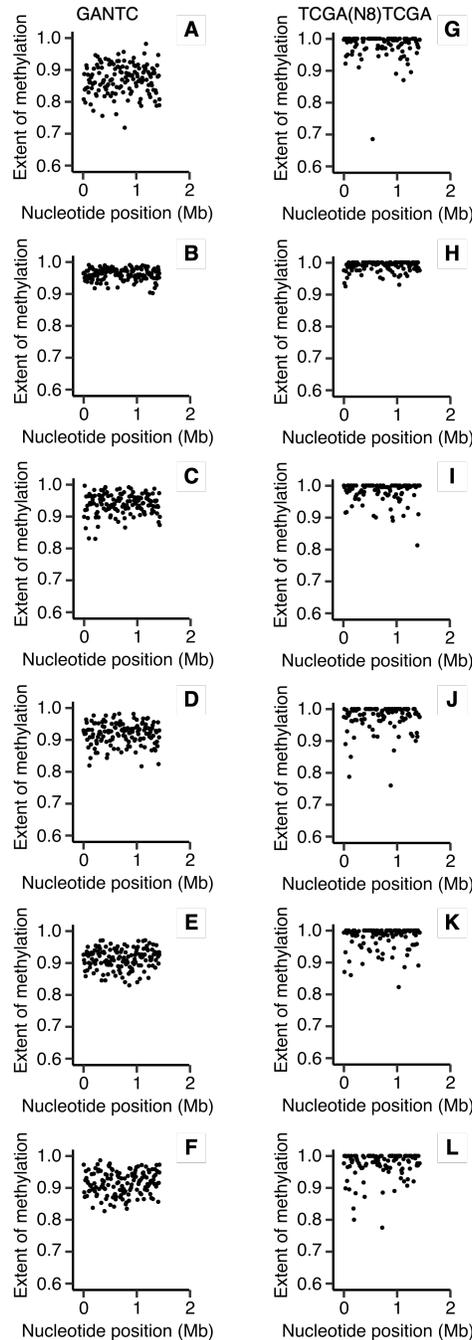


Figure S21. pSymA-wide DNA methylation of *E. meliloti* FSM-MA bacteroids purified from *M. truncatula* *dnf* mutant nodules. The extent of methylation of (A-F) GANTC and (G-L) TCGA(N₈)TCGA motifs across the *E. meliloti* FSM-MA pSymA replicon is shown using a 10 kb sliding window. (A,G) Bacteroids isolated from *M. truncatula* *dnf1* mutant nodules. (B,H) Bacteroids isolated from *M. truncatula* *dnf5* mutant nodules. (C,I) Bacteroids isolated from *M. truncatula* *dnf2* mutant nodules. (D,J) Bacteroids isolated from *M. truncatula* *dnf7* mutant nodules. (E,K) Bacteroids isolated from *M. truncatula* *dnf4* mutant nodules. (F,L) Bacteroids isolated from wild-type *M. truncatula* A17 nodules.

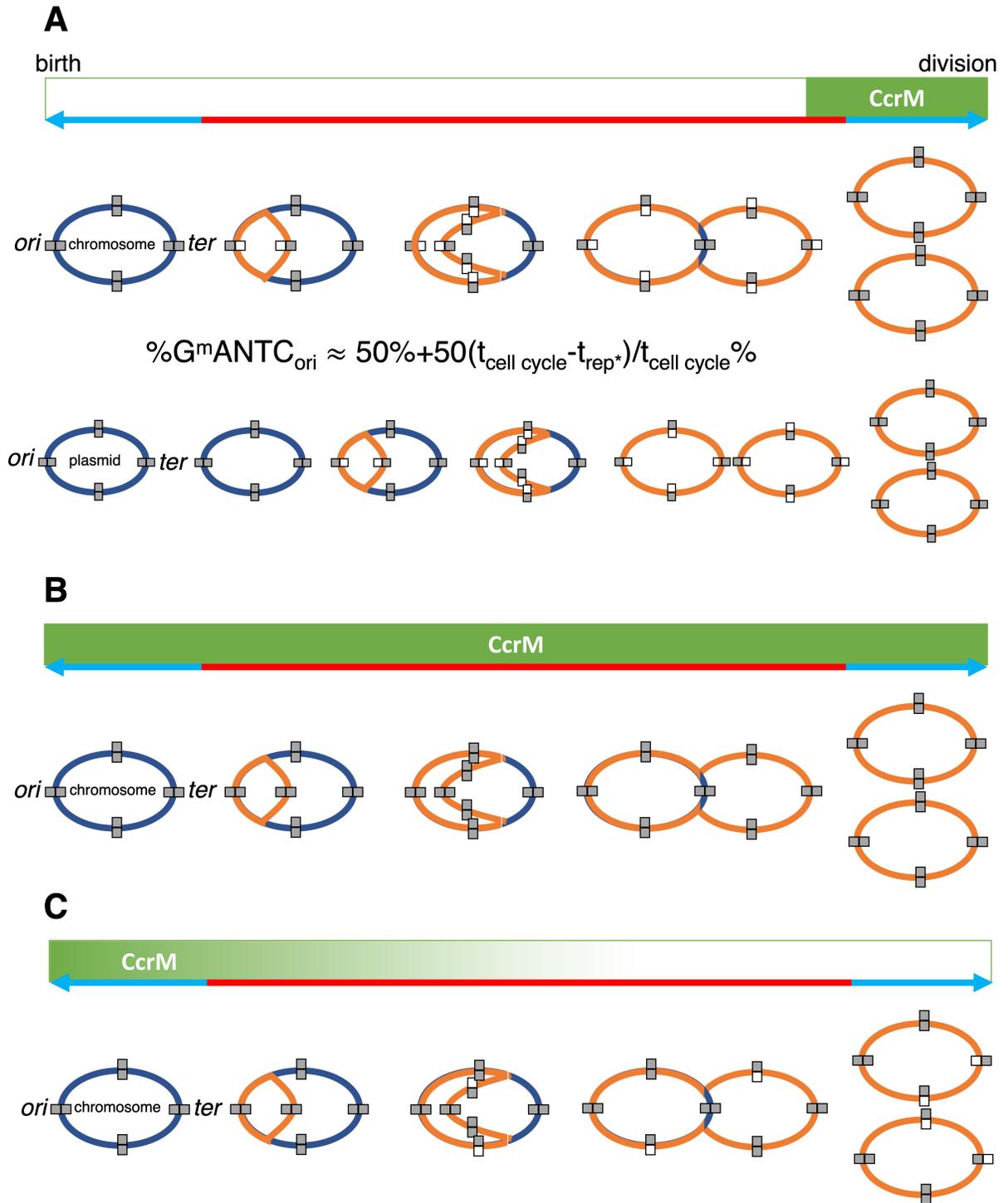


Figure S22. Model describing the methylation patterns observed in cultures and bacteroids. (A) Methylation pattern in free-living cells in exponential and stationary phases of growth. The blue/red line indicates the relevant cell cycle phases: red is the S-phase or genome replication phase; blue is the gap phases and division. The replication progression and methylation status of the circular chromosome and a plasmid is depicted during the cell cycle progression from cell birth

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

Dataset S1 (separate file). Cell cycle regulated genes belonging to transcripts with at least one GANTC site situated within the 125 bp upstream region. The first column indicates the gene locus tag, while the second column indicates the cell cycle expression group of the gene as defined by De Nisco et al (2014) (6).

Dataset S2 (separate file). The predicted extent of methylation (defined as the estimated fraction of reads mapping to a motif that were methylated) of each GANTC site across the entire genome in all replicates of all conditions is shown for all strains used in this study. A value of 0 is used when methylation of a site was not predicted; however, a site not being called as methylated does not necessarily indicate that the site is not methylated.

Dataset S3 (separate file). An archive of the raw flow cytometry data (as FCS files) used to generate the figures presented in this study.

SUPPLEMENTARY REFERENCES

1. Wendt T, Doohan F, Mullins E. 2012. Production of *Phytophthora infestans*-resistant potato (*Solanum tuberosum*) utilising *Ensifer adhaerens* OV14. *Transgenic Res* 21:567–578.
2. Trinick MJ. 1980. Relationships amongst the fast-growing rhizobia of *Lablab purpureus*, *Leucaena leucocephala*, *Mimosa* spp., *Acacia farnesiana* and *Sesbania grandiflora* and their affinities with other rhizobial groups. *J Appl Bacteriol* 49:39–53.
3. Nagymihaly M, Vásárhelyi BM, Barrière Q, Chong T-M, Bálint B, Bihari P, Hong K-W, Horváth B, Ibjibijen J, Amar M, Farkas A, Kondorosi E, Chan K-G, Gruber V, Ratet P, Mergaert P, Kereszt A. 2017. The complete genome sequence of *Ensifer meliloti* strain CCMM B554 (FSM-MA), a highly effective nitrogen-fixing microsymbiont of *Medicago truncatula* Gaertn. *Stand Genomic Sci* 12:75.
4. diCenzo GC, MacLean AM, Milunovic B, Golding GB, Finan TM. 2014. Examination of prokaryotic multipartite genome evolution through experimental genome reduction. *PLOS Genet* 10:e1004742.
5. diCenzo GC, Zamani M, Milunovic B, Finan TM. 2016. Genomic resources for identification of the minimal N₂-fixing symbiotic genome. *Environ Microbiol* 18:2534–2547.
6. De Nisco NJ, Abo RP, Wu CM, Penterman J, Walker GC. 2014. Global analysis of cell cycle gene expression of the legume symbiont *Sinorhizobium meliloti*. *Proc Natl Acad Sci USA* 111:3217–3224.
7. Wright R, Stephens C, Zweiger G, Shapiro L, Alley MR. 1996. *Caulobacter* Lon protease has a critical role in cell-cycle control of DNA methylation. *Genes Dev* 10:1532–1542.
8. Li S, Brazhnik P, Sobral B, Tyson JJ. 2008. A Quantitative Study of the Division Cycle of *Caulobacter crescentus* Stalked Cells. *PLoS Comput Biol* 4:e9.
9. Reyes-Lamothe R, Possoz C, Danilova O, Sherratt DJ. 2008. Independent Positioning and Action of *Escherichia coli* Replisomes in Live Cells. *Cell* 133:90–102.
10. Frage B, Döhlemann J, Robledo M, Lucena D, Sobetzko P, Graumann PL, Becker A. 2016. Spatiotemporal choreography of chromosome and megaplasmids in the *Sinorhizobium meliloti* cell cycle. *Mol Microbiol* 100:808–823.
11. Mohapatra SS, Fioravanti A, Biondi EG. 2014. DNA methylation in *Caulobacter* and other Alphaproteobacteria during cell cycle progression. *Trends Microbiol* 22:528–535.
12. Fagorzi C, Ilie A, Decorosi F, Cangioli L, Viti C, Mengoni A, diCenzo GC. 2020. Symbiotic and nonsymbiotic members of the genus *Ensifer* (syn. *Sinorhizobium*) are separated into two clades based on comparative genomics and high-throughput phenotyping. *Genome Biol Evol* 12:2521–2534.