Cell PRESS
Open ACCESS

EXTENDED EXPERIMENTAL PROCEDURES

TrfA Is Essential to the Plasmid Replication of pCM410 in Methylobacterium

The high frequency of *trfA* mutations, particularly ISMex25 insertions into its 3′ region, was rather surprising since this gene should be critical to plasmid replication in *Methylobacterium* (Lin and Helinski, 1992). To test its essentiality, we constructed 4 plasmids based on pCM410 that carried either a designed *trfA* deletion or three different ISMex25-inserted *trfA* (ISMex25::*trfA*) alleles from mini-plasmids (Figure 1D; see also Extended Experimental Procedures). These four variants were then introduced into the ancestral EM background, but none were able to be maintained stably, verifying the functional essentiality of TrfA to plasmid replication. The above results and the uneven distribution of ISMex25 insertions on pCM410 suggested that, instead of relying on TrfA, replication of $pCM410^{C}$ depends on the replication machinery present on p2META; moreover, disrupting the TrfA function might be required for stable maintenance of $pCM410^{C}$ in *Methylobacterium*. Interestingly, although plasmid constructs bearing ISMex25::*trfA* alleles could not replicate, introducing these plasmids into the ancestral EM background occasionally led to spontaneous integration into p2META through homologous recombination between ISMex25 copies present on these plasmids and p2META. One such synthetic cointegrate pHC24, conferred a fitness increase of 16.9% ± 1.5% relative to EM (Table S1). The ease of generating synthetic cointegrates resembling $pCM410^{C}$ through this manner allowed us to assess the fitness effect of cointegrate formation alone. Testing synthetic cointegrates avoided the complications of potential unidentified mutations present in other regions of the 49 kb $pCM410^{C}$.

Effects of the Flow Cytometry Forward Scatter Threshold on Fitness Assays

The small size of bacterial cells poses a technical issue for quantification using flow cytometry. The forward scatter (FSC, a good proxy for particle size) detection threshold of a cytometer needs to be set low enough for capturing cells of smaller sizes while the corresponding increase in instrument sensitivity inevitably results in the elevation of background noise (i.e., false positive particle counts). In order to select an appropriate level of sensitivity that balanced this tradeoff, we investigated the influence of the FSC detection threshold on cytometric quantification of strains with different cell sizes because our prior study indicated reduction in cell sizes during laboratory evolution of EM (Chou et al., 2011). Three strains of *Methylobacterium*, WT (cell length as 3.10 ± 0.05 μm, (Chou et al., 2011)), EM (4.50 ± 0.10 μm, (Chou et al., 2011)), and CM2291 (5.61 ± 0.13 μm; Chou et al., unpublished) were chosen to span the size range of strains characterized in this study. The fluorescent EM strain CM1232 that served as the reference in fitness assays was used as a standard for comparison here. Four cell suspensions with different compositions were prepared prior to the measurement: (1) CM1232, (2) a mixture of CM1232 with the WT strain, (3) a mixture of CM1232 with the EM strain, and (4) a mixture of CM1232 with CM2291. Each sample was quantified by LSR II flow cytometer (BD Biosciences) three times, each time for at least 50000 particle counts under five different FSC detection thresholds (200, 300, 500, 800, 1000). The following instrument settings remained unchanged throughout the experiment: FSC detector (250 V), side scatter detector (300 V), mCherry fluorescence detector (600 V), excitation laser (561 nm), long-pass filter (600 nm), and band-pass filter (610−620 nm). The actual flow rate of LSRII was also fixed although the reported flow rates estimated as counted particles per second ranged from 735 to 5556 due to changes in the instrument sensitivity. As the first cell suspension consisted of only florescent cells, counts of non-fluorescent particles in this sample reflected the magnitude of instrument noise with respect to FSC detection thresholds. Moreover, changes in the ratio of counts between fluorescent and non-fluorescent particles in the rest three mixed samples indicated the influence of FSC detection thresholds on quantification of strains with different cell sizes.

The magnitude of background noise elevated drastically as the FSC threshold dropped below 300 (Figure S2A). This increase in instrument sensitivity, and hence the noise level, caused higher count errors under low FSC thresholds (Figures S2A–S2D). On the contrary, increasing FSC thresholds lowered the instrument sensitivity, which led to strong detection bias toward larger cells (Figures S2B–S2D). This tendency was also reflected in the increasing FSC measures of all four strains with higher FSC thresholds (Figures S2E–S2G). Changes in the FSC threshold exerted little influence on counts of fluorescent versus non-fluorescent cells when two strains were of equal cell sizes (Figure S2C); however, this count ratio changed significantly with respect to FSC thresholds when two strains had different cell sizes (Figures S2B and S2D). Based on above results, all cytometric measurements reported in the study was based on FSC threshold 300 as this setting minimized both the background noise and the count bias. When comparing fitness quantified using FSC threshold 300 with an independent fitness assays of identical strains quantified by FSC threshold 800, results from the latter clearly underestimated the actual fitness values. This was due to an underestimation of cell numbers of strains harboring various pCM410 haplotypes in fitness assays since their cell sizes were smaller than the fluorescent reference EM strain. This is the basis for the nearly two-fold higher fitness value for $pCM410^{A1}$ stated here than in our earlier report (Chou et al., 2011).

Influences of Mutations of $pCM410^{A}$ on Predicted mRNA Folding and Translation of *fghA*

Translational repressors or mutations in the RBS that attenuate ribosome binding to mRNA have been shown to elevate mRNA degradation by increasing its exposure to RNase (Iost and Dreyfus, 1995; Cole and Nomura, 1986; Vytvytska et al., 2000; Kaberdin and Bläsi, 2006; Cho and Yanofsky, 1988). The magnitude of ribosome binding and the protection it conveys against mRNA degradation correlate positively with the translation initiation rate but negatively with the mRNA folding energy surrounding the RBS (Kudla et al.,

2009; Iost and Dreyfus, 1995; Yarchuk et al., 1992). Since mutations of pCM410^{A} all occurred in or adjacent to the RBS$_{fghA}$ and caused a significant reduction of the *flhA-fghA* transcript level, we suspected this decrease in the mRNA level resulted from a very similar reason. To computationally assess this possibility, we utilized RBS calculator (Salis et al., 2009) and Mfold (Zuker, 2003) programs to predict the translation initiation rate and the folding energy of the *flhA-fghA* transcript surrounding the RBS$_{fghA}$, respectively. In addition to correctly predicting the translational start site of both *flhA* and *fghA*, RBS calculator identified another potential in-frame translation start site of *fghA* 90 bp downstream of the annotated start codon, even though the strength of this second site was predicted to be 33-fold weaker. Except for the insertion of ISMex4 upstream of *fghA* (pCM410^{2A}), the remaining four pCM410^{A} mutations either nearly or completely abolished translation from the annotated start codon (Table S2). None of the pCM410^{A} mutations were predicted to generate effective in-frame translation start sites for *fghA*. On the other hand, the Mfold program predicted a substantial increase in the mRNA folding energy caused by the ISMex4 insertion upstream of *fghA*, while the other four pCM410^{A} mutations exerted negligible effects. Results from above analysis support our hypothesis that mutations of pCM410^{A} reduce the *flhA-fghA* transcript level by promoting mRNA degradation.

Plasmid Construction

The ancestral plasmid pCM410 that expressed the *flhA* and *fghA* genes from *P. denitrificans* by the P_{mxaF} promoter was generated previously (Chou et al., 2011). pHC96 that bears the P_{mxaF}-*flhA* cassette was constructed by ligating the 1.2 kb *Xba*I–*Xba*I fragment of pCM410 into pCM160 (Marx and Lidstrom, 2001) cut with *Xba*I. pHC97 that bears the P_{mxaF}–*fghA* cassette resulted from self-ligation of a 8.9 kb *Xba*I–*Xba*I fragment of pCM410. To tune the expression of the P_{mxaF}–*flhA-fghA* cassette, the cumate (*p*-isopropylbenzoate)-inducible system (Choi et al., 2006) was implanted into pCM410. The operator (CuO_{cmt}) of the cumate repressor CymR was synthesized by annealing two oligonucleotides, CuO_{cmt}-f (5′-GATCAACAAACAGACAATCTGGTCTGTTTGTAGGTAC-3′) and CuO_{cmt}-r (5′-CTACAAACAGACCAGATTGTCTGTTTGTT-3′), and ligated into the *Bam*HI–*Kpn*I site downstream of P_{mxaF} to generate pHC111. A 654-bp PCR fragment containing the RBS of the *fae* gene of *Methylobacterium* and the *cymR* gene was amplified using the genomic DNA of the *Methylobacterium* CymR1 strain (Choi et al., 2006) and primers cymR-f (5′-TGCTAG CAGGGAGAGACCCCGAATGGTGATCATGAGTCCAAAG-3′) and cymR-r (5′-ATTTATGGGCCCTAGCGCTTGAATTTCGC-3′). This fragment was subsequently cut and ligated into the *Nhe*I–*Psp*OMI site downstream of the P_{lac} promoter of pHC111 to generate pHC112. To examine whether this plasmid-borne cumate-inducible system resulted in homogenous, titratable expression across the population, the *flhA-fghA* cassette in pHC112 was replaced with the fluorescent reporter gene *mCherry* from pHC05 (Lee et al., 2009) to generate pHC113m. This was made by ligating the 746-bp EcoRI-KpnI fragment of pHC05 into pHC112 cut with EcoRI and KpnI. Using a LSR II flow cytometer (BD Biosciences) to quantify mCherry expression of pHC113m, we showed that this regulatory expression system resulted in homogenous expression for ~95% of the population while ~5% of the population exhibited no expression across different induction levels (data not shown). To convert pHC112 into a versatile inducible expression system for *Methylobacterium*, the *flhA-fghA* cassette in pHC112 was replaced with a multiple cloning site. This multiple cloning site was synthesized by annealing two complementary strands of oligonucleotides (5′-GCACGTGAATTCCATGGACCGGTACTAGTA GATCTTAAGAGCT-3′ and 5′-CTTAAGATCTACTAGTACCGGTCCATGGAATTCACGTGC-3′). It was then ligated into the pHC112 backbone cut with FspI and SacI to generate pHC115. Because mini-plasmids resulting from homologous recombination of two copies of ISMex25 on pCM410^{C} (plasmid cointegrates) were difficult to obtain directly from evolved isolates, equivalent versions were constructed in *E. coli* DH5α or 10-beta strains (New England Biolabs) by replacing the wild-type *trfA* gene of pCM410 with the *trfA*::ISMex25 alleles from these mini-plasmids. The presence of the ColE1 origin of replication in pCM410 and its derivatives is sufficient for their replication in *E. coli* without the *trfA* gene and the IncP origin of replication (*oriV*) (Marx and Lidstrom, 2002). PCR fragments containing *trfA*::ISMex25 alleles were amplified with primers S11 and S12 using genomic DNA of evolved isolates CM935, CM937, and CM956, all of which carried pCM410^{C} haplotypes. These fragments were then cut and ligated into pCM410 digested with *Sex*AI and *Sfi*I to generate pHC17, pHC18, and pHC19 (equivalent to the mini-plasmids derived from pCM410^{C} haplotypes in evolved isolates CM935, CM937, and CM956, respectively). To test whether the *trfA* gene is essential to replication of pCM410 in *Methylobacterium*, this gene was truncated by cutting pCM410 with *Sex*AI and *Sfi*I, using the Klenow fragment of DNA polymerase I (New England Biolabs) to fill in and chew back the respective overhanging ends of the resulting 9.4 kb *Sex*AI–*Sfi*I fragment, and finally self-ligating this blunted fragment to generate pHC27. In order to characterize the 1.1 kb deletion and the single nucleotide substitution present in pCM410^{B1}, these two mutations were separately introduced into pCM410 using PCR-driven overlap extension (Heckman and Pease, 2007) as they are merely 5-bp apart. Two overlap fragments required for synthesis of the fragment encompassing the 1.1 kb deletion were separately amplified from pCM410^{B1} using primers HC410p17 and HC410p14, and primers HC410p11 and HC410p15, respectively. These two fragments were then pooled together to perform PCR-driven overlap extension using primers HC410p11 and HC410p17. The resulting 0.8 kb fragment was cut and ligated into pCM410 cut with *Psi*I and *Sfi*I to generate pHC83 (as pCM410^{B2}). The fragment containing just the SNP was synthesized by a similar manner. The SNP (underlined) was incorporated in primers HC1002p1 (5′-GGCTTGCGATGGTTGCG-3′) and HC1002p2 (5′-CGCAACCATCGCAAGCC-3′) that span the overlapping region of two overlap fragments. The two overlap fragments were separately amplified with the pCM410 plasmid DNA using primers HC410p17 and HC1002p1, and primers HC410p11 and HC1002p2, respectively. These two fragments were then pooled together to perform PCR-driven overlap extension using primers HC410p11 and HC410p17. The resulting 1.9 kb fragment was cut and ligated into pCM410 cut with *Psi*I and *Sfi*I to generate pHC84 (as pCM410^{B3}).

Growth Rate and Fitness Assays

Prior to growth rate assays and fitness assays, all strains were acclimated in minimal media supplemented with carbon sources used in the ensuing assays. Three replicate cultures of each strain were sampled periodically and the change in OD_{600} was measured using a Bio-Rad microplate reader model 680. Competition experiments were performed by following a previously described procedure (Lee et al., 2009). Briefly, after one round of acclimation, test strains and a reference strain expressing a fitness-neutral fluorescent protein were mixed by a 1:1 volume ratio, diluted 1:64 into 9.6 ml of minimal media supplemented with 15 mM methanol. The ratios of non-fluorescent cells in mixed populations were measured by passing population samples before (R_0) and after (R_1) competition growth through a LSR II flow cytometer (BD Biosciences) for at least 5×10^4 cell counts per sample. Due to the small size of bacterial cells, the forward scatter threshold of LSRII was adjusted to 300 since this setting ensured the detection of cells of smaller sizes while minimizing the background noise of the instrument (Figure S2). Fitness values (W) relative to the reference strain were calculated by a previously described equation assuming an average of 64-fold size expansion of mixed populations during competitive growth (Lee et al., 2009):

$$W = \frac{\log\left(\frac{R_1 \cdot 64}{R_0}\right)}{\log\left(\frac{(1 - R_1) \cdot 64}{1 - R_0}\right)}$$

The fluorescent EM strain CM1232 (Chou et al., 2011) was used regularly as the reference to measure fitness values of the Δ*mptG* strain (CM624) bearing pCM410 derivatives. CM2055, a fluorescent WT *Methylobacterium* (CM1180 (Lee et al., 2009)) bearing an empty control plasmid pCM160 (Marx and Lidstrom, 2001) was the reference for quantifying the fitness cost of FlhA and FghA overexpression in WT *Methylobacterium* bearing pHC96 and pHC97, respectively. Because *Methylobacterium* can lose IncP plasmids spontaneously in the absence of selection, when performing fitness assays with the above three strains growth media were supplemented with kanamycin (25 mg/ml) in addition to 15 mM methanol during the acclimation phase. Competition experiments then proceeded in growth media supplemented with just 15 mM methanol.

Characterization of Class C Haplotypes

The junctions between the pCM410 and the p2META backbones of $pCM410^C$ were confirmed by PCR amplification and sequencing using primers HC11p1 (5′-ATCGTCGGCTCGGTCGGTTTC-3′) and HC11p2, primers HC410p7 and HC11p3 (5′-GCCTGAGGTTCC CAGGAAGC-3′), primers HC11p3 and HC11p2, or primers HC11p1 and HC410p7 depending on the orientation of ISMex25 transposition. In order to test whether plasmids with disrupted *trfA* alleles can replicate in *Methylobacterium*, three pCM410-derived plasmids with *trfA*::ISMex25 alleles (pHC17, pHC18, and pHC19) and with a truncated *trfA* allele (pHC27) were conjugated into the Δ*mptG* strain CM624 via tri-parental mating (Chistoserdov et al., 1994). The mating mixtures were then selected on 2% (w/v) agar plates containing 14.8 mM succinate and 50 μg/ml kanamycin. As all such plasmids lost the ability to self-replicate in *Methylobacterium*, selection for kanamycin resistance led to maintenance of the *kan* gene of pHC17, pHC18, pHC19 through integration of whole plasmids into p2META via homologous recombination of ISMex25 copies, generating cointegrates pHC24, pHC26, and pHC25, respectively. On the contrary, conjugation of pHC27 into CM624 did not result in any viable clones and thus no successful integration under kanamycin selection.

Quantification of Plasmid Copy Numbers and Transcript Levels

The primers used to detect transcripts of the *fhlA*, *fghA*, and *rpsB* genes were HC410p20 and HC410p21, HC410p22 and HC410p23, and HCAM111 (5′-TGACCAACTGGAAGACCATCTCC-3′) and HCAM113 (5′-TTGGTGTCGATCACGAACAGCAG-3′), respectively. Detection of plasmid DNA was targeted at the *kan* gene using primers HC410p18 and HC410p19. The *rpsB* gene (encodes 30S ribosomal protein S2) in the *Methylobacterium* chromosome META1 was chosen as the reference gene for normalizing transcript levels as well as plasmid copy numbers. Two-step real-time PCR experiments were performed in three replicates with the PerfeCTa SYBR Green SuperMix (Quanta Biosciences) on a DNA Engine Opticon2 (MJ Research). Data analysis was performed with the Opticon Monitor v. 2.02 (MJ Research). The average threshold cycle (Ct) value for each gene was calculated from triplicate reactions by following a previously described method (Chou et al., 2009). The ΔCt value described the difference between Ct of the target gene and Ct of the reference *rpsB* gene. The ΔΔCt value described the difference between the ΔCt of the engineered strain CM702 and that of CM624 carrying pCM410-derived plasmids. The relative difference in transcript levels was calculated as $2^{\Delta\Delta Ct}$. Plasmid copy numbers were measured by the following procedures. Genomic DNA was quantified by a Nanodrop ND-1000 (Thermo Scientific), and 25 ng of genomic DNA was added to each real-time PCR reaction. To establish a standard curve, 1, 0.1, 0.01, and 0.001 ng of pCM410 (equivalent to 9.09×10^7, 9.09×10^6, 9.09×10^5, and 9.09×10^4 plasmid molecules, respectively) were mixed with 25 ng of the genomic DNA (equivalent to 3.03×10^6 genome copies) of the WT strain CM502. The standard curve is a plot of ΔCt (i.e., Ct_{kan}–Ct_{rpsB}) versus plasmid molecules in the Log_2 scale (Slope = −1.122, $R^2 = 0.997$, data not shown). For each sample, by interpolating its ΔCt value against the SC, the absolute quantity of plasmid DNA can be estimated using the following equation:

$$\text{Plasmid copy number per genome} = 2^{\frac{\Delta Ct - SC\ intercept}{SC\ slope}} \div (3.03 \times 10^6)$$

Cluster Analysis and Principle Component Analysis

The plasmid copy numbers, *flhA* and *fghA* transcript levels, and FlhA and FghA enzyme activities of the ancestor pCM410 and its derivatives in the ancestral EM background were normalized such that the group sum of each phenotypic measurement was zero and the sum of the squares equaled 1. Cluster analysis was performed using Cluster 3.0 (Eisen et al., 1998). The similarity matrices of this normalized dataset were computed based on the Euclidean distance or the Pearson correlation. Hierarchical clustering was then performed using centroid linkage, single linkage, complete linage, or average linkage methods where phenotypes were weighted equally. The resulting dendrograms were visualized using Java TreeView (Saldanha, 2004). Principle component analysis of the normalized dataset was performed using the 'princomp' function of the R package (R Development Core Team, 2009).

SUPPLEMENTAL REFERENCES

Cho, K.O., and Yanofsky, C. (1988). Sequence changes preceding a Shine-Dalgarno region influence *trpE* mRNA translation and decay. J. Mol. Biol. *204*, 51–60.

Choi, Y.J., Morel, L., Bourque, D., Mullick, A., Massie, B., and Míguez, C.B. (2006). Bestowing inducibility on the cloned methanol dehydrogenase promoter (P_{mxaF}) of *Methylobacterium extorquens* by applying regulatory elements of *Pseudomonas putida* F1. Appl. Environ. Microbiol. *72*, 7723–7729.

Cole, J.R., and Nomura, M. (1986). Changes in the half-life of ribosomal protein messenger RNA caused by translational repression. J. Mol. Biol. *188*, 383–392.

Heckman, K.L., and Pease, L.R. (2007). Gene splicing and mutagenesis by PCR-driven overlap extension. Nat. Protoc. *2*, 924–932.

Kudla, G., Murray, A.W., Tollervey, D., and Plotkin, J.B. (2009). Coding-sequence determinants of gene expression in *Escherichia coli.* Science *324*, 255–258.

Lee, M.C., Chou, H.H., and Marx, C.J. (2009). Asymmetric, bimodal trade-offs during adaptation of *Methylobacterium* to distinct growth substrates. Evolution *63*, 2816–2830.

Marx, C.J. (2008). Development of a broad-host-range *sacB*-based vector for unmarked allelic exchange. BMC. Res. Notes *1*, 1.

Marx, C.J., and Lidstrom, M.E. (2001). Development of improved versatile broad-host-range vectors for use in methylotrophs and other Gram-negative bacteria. Microbiology *147*, 2065–2075.

Marx, C.J., and Lidstrom, M.E. (2002). Broad-host-range *cre-lox* system for antibiotic marker recycling in gram-negative bacteria. Biotechniques *33*, 1062–1067.

R Development Core Team (2009). R: A Language and Environment for Statistical Computing (Vienna: R Foundation for Statistical Computing).

Saldanha, A.J. (2004). Java Treeview—extensible visualization of microarray data. Bioinformatics *20*, 3246–3248.

Vytvytska, O., Moll, I., Kaberdin, V.R., von Gabain, A., and Bläsi, U. (2000). Hfq (HF1) stimulates *ompA* mRNA decay by interfering with ribosome binding. Genes Dev. *14*, 1109–1118.

Yarchuk, O., Jacques, N., Guillerez, J., and Dreyfus, M. (1992). Interdependence of translation, transcription and mRNA degradation in the *lacZ* gene. J. Mol. Biol. *226*, 581–596.

Zuker, M. (2003). Mfold web server for nucleic acid folding and hybridization prediction. Nucleic Acids Res. *31*, 3406–3415.

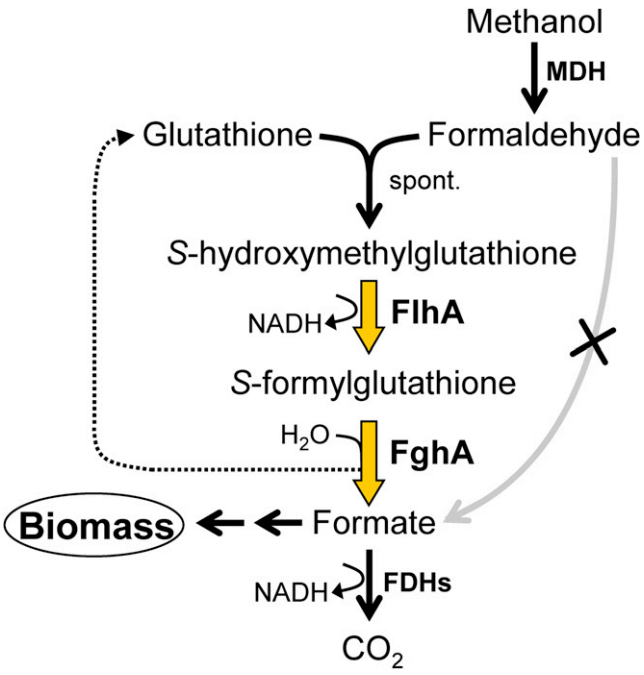

Figure S1. Methanol Metabolism in the EM Strain, Related to Figure 1

The GSH-linked formaldehyde oxidation pathway from is indicated with yellow arrows. The disabled native pathway is indicated by the gray arrow with the "X". MDH, methanol dehydrogenase; spont., spontaneous reaction; FlhA, *S*-hydroxymethyl glutathione dehydrogenase; FghA, *S*-formyl-GSH hydrolase; FDHs, formate dehydrogenases.

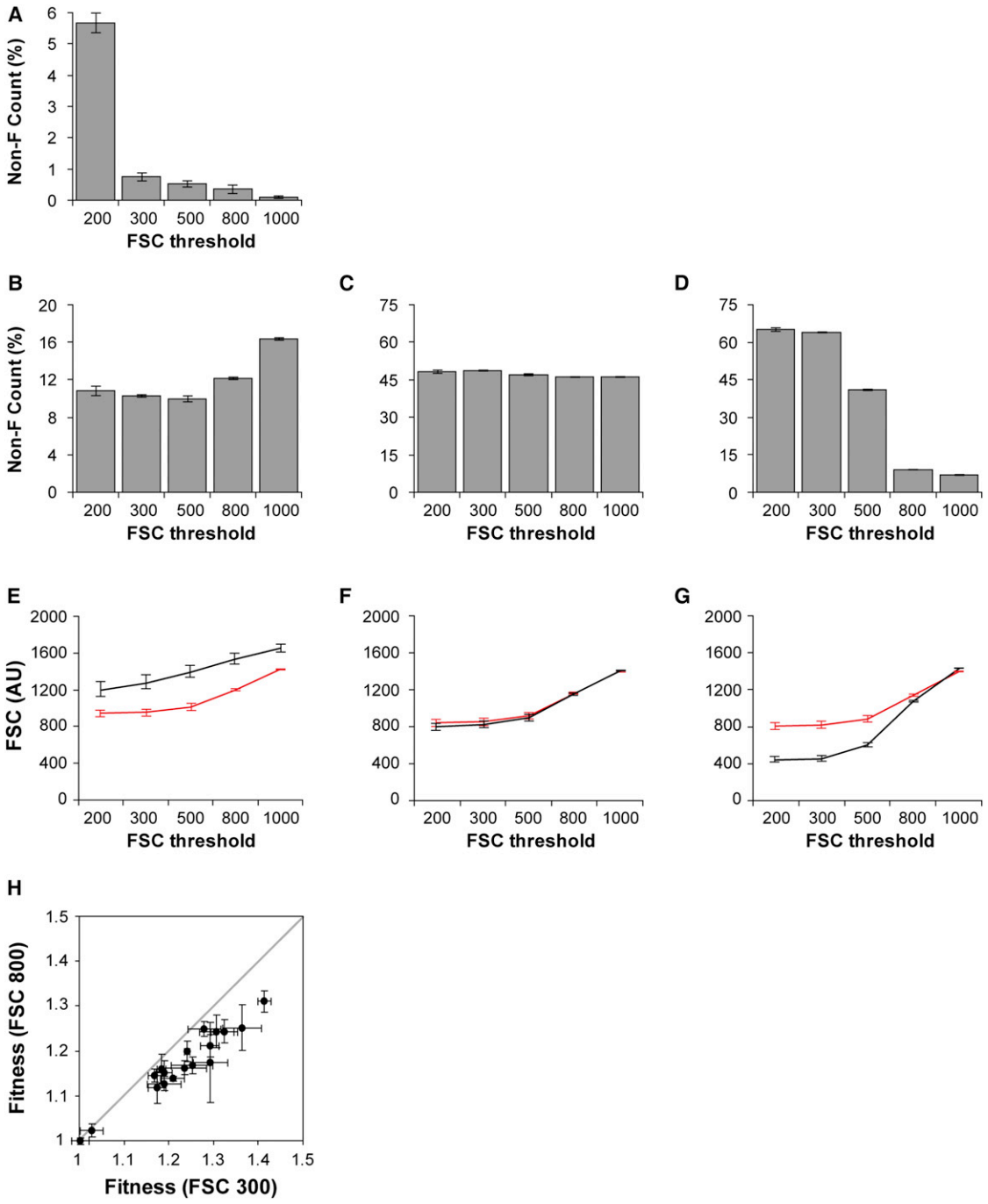

Figure S2. Influence of Flow Cytometry Settings on Fitness Assays, Related to Figure 2

(A–D) Counts of non-fluorescent (Non-F) particles in samples consisting of (A) just the fluorescent EM strain, or the fluorescent EM strain mixed with strains of (B) larger, (C) equal, and (D) smaller cell sizes (i.e., CM2291, EM, and WT strains, respectively) under various forward scatter (FSC) thresholds. Error bars are 95% confidence interval.

(E–G) Geometric means of FSC values (expressed in arbitrary units, AU) of fluorescent cells (red) and non-fluorescent cells (black) in samples consisting of the fluorescent EM strain mixed with strains of (E) larger, (F) equal, and (G) smaller cell sizes (i.e. CM2291, EM, and WT strains, respectively) under various FSC thresholds.

(H) Fitness measurements of pCM410 derivatives quantified at FSC thresholds 300 and 800.

Error bars indicate 95% confidence intervals.

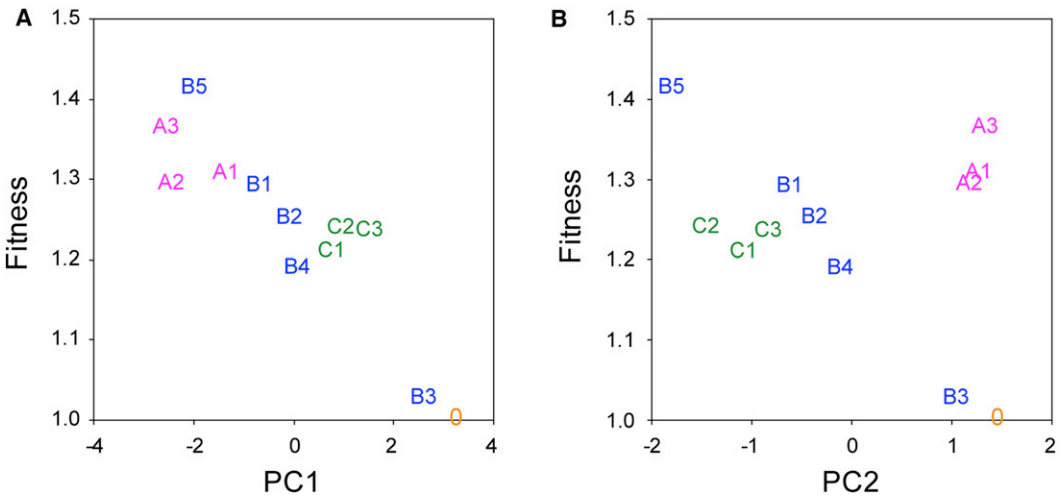

Figure S3. Fitness Correlates Strongly with Gene Expression, Related to Figure 4B

(A) The correlation between principle component 1 (PC1) and fitness.

(B) The correlation between principle component 2 (PC2) and fitness.

Table S1. Genotypes and Phenotypes of pCM410 and its derivatives.

Haplotypes			Source			Plasmid copy	Transcript		Enzyme activity		Fitness
Class	**Index**	**[a]Mutations**	**Pop**	**Gen**	**Iso**		***flhA***	***fghA***	**FlhA**	**FghA**	
Ancestral pCM410	0					7.477 ± 0.680	1.000 ± 0.091	1.000 ± 0.160	0.816 ± 0.056	2.074 ± 0.189	1.002 ± 0.020
A	A1	11 bp deletion (3791–3801bp)	4	600	**CM1145**, CM1735, CM1736, CM1737	8.098 ± 1.628	0.373 ± 0.155	0.418 ± 0.128	0.449 ± 0.045	0.266 ± 0.023	1.307 ± 0.039
		11 bp deletion (3794–3804 bp)	1	600	CM1129, **CM1726**						1.279 ± 0.037
		37 bp deletion (3766–3802 bp)	6	600	CM1157, CM1741, **CM1742**, CM1743						1.324 ± 0.030
	A2	ISMex4 insertion between 3771 and 3772 bp	4	300	**CM1007**	7.170 ± 1.660	0.092 ± 0.010	0.098 ± 0.024	0.414 ± 0.040	0.174 ± 0.021	1.293 ± 0.038
	A3	T→C at 3804 bp	5	600	CM1151, **CM1738**, CM1739, CM1740	8.012 ± 0.907	0.096 ± 0.025	0.181 ± 0.036	0.359 ± 0.009	0.094 ±0.002	1.364 ± 0.043
B	B1	1109 bp deletion (9141–113 bp); T→G at 9136 bp	2	300	**CM1002**	3.962 ± 0.289	0.457 ± 0.037	0.485 ± 0.073	0.432 ± 0.050	0.817 ± 0.129	1.292 ± 0.021

			2	600	CM1729, CM1730, CM1731						
	B2	[b] pHC83, 1109 bp deletion (9141–113 bp)				5.007 ± 0.276	0.581 ± 0.095	0.722 ± 0.071	0.442 ± 0.034	0.854 ± 0.090	1.252 ± 0.047
	B3	[b] pHC84, T→G at 9136 bp				6.775 ± 0.603	0.828 ± 0.188	1.103 ± 0.282	0.785 ± 0.114	1.585 ± 0.258	1.028 ± 0.026
	B4	T→G at 8314 bp; T→G at 8576 bp	3	600	**CM1734**	5.525 ± 0.915	0.627 ± 0.047	0.725 ± 0.067	0.463 ± 0.083	0.862 ± 0.070	1.190 ± 0.037
	B5	ISMex25 insertion between 8302 and 8303 bp	6	300	**CM1013**	1.726 ± 0.213	0.263 ± 0.084	0.426 ± 0.090	0.340 ± 0.011	0.129 ± 0.003	1.414 ± 0.016
C		Cointegrate, ISMex25 insertion between 8106 and 8107 bp	7	120	**CM937**						1.241 ± 0.006
		Cointegrate, ISMex25 insertion between 8154 and 8155 bp	1	120	CM935, CM936						
	C1		1	600	**CM1727**, CM1728	4.551 ± 0.526	0.997 ± 0.209	0.878 ± 0.178	0.403 ± 0.038	1.010 ± 0.118	1.209 ± 0.027
		Cointegrate, ISMex25 insertion between 8274 and 8275 bp	8	120	CM956						
			8	600	**CM1747**, CM1748, CM1749						1.174 ± 0.020
	C2	Cointegrate, ISMex25 insertion between 8302 and 8303 bp	6	180	**CM2128**	3.810 ± 1.021	1.054 ± 0.156	0.983 ± 0.210	0.441 ± 0.048	0.838 ± 0.073	1.239 ± 0.042
		Cointegrate, ISMex25 insertion between	7	600	**CM1744**,						1.184 ±

	8343 and 8344 bp			CM1745, CM1746						0.010
	Cointegrate, ISMex25 insertion between 8374 and 8375 bp	7	120	CM955						1.190 ± 0.018
C3	Cointegrate, ISMex25 insertion between 732 and 733 bp; T→G at 8314 bp; T→G at 8576 bp	3	600	**CM1139**, CM1732, CM1733	4.598 ± 0.493	0.888 ± 0.154	1.207 ± 0.140	0.538 ± 0.033	1.027 ± 0.079	1.235 ± 0.050
	[c]pHC24, ISMex25 insertion between 8154 and 8155 bp									1.195 ± 0.040

Evolved isolates whose plasmids were subject to phenotypic measurements are indicated by bold texts. The indexed haplotypes correspond to those described in Figure 2. Pop, population; Gen, generation; Iso, evolved isolate; ND, not determined.

[a]The locations of mutations on pCM410 are described in terms of the plasmid file deposited in Genbank (Accession no. FJ389188).

[b]plasmids constructed by introducing each of the two mutations of $pCM410^{B1}$ into the ancestral pCM410.

[c]synthetic plasmid cointegrate, resulting from homologous recombination between two ISMex25 copies present on the plasmid pHC17 and on the native pMETA2 plasmid.

Table S2. Influences of Mutations of pCM410[A] on Predicted mRNA Folding and Translation of *fghA*, Related to Figure 4A

Haplotype index	[a]Mutations	[b]mRNA folding energy (kcal/mol)		[c]Translation
		600 nt	2000 nt	
0	No mutation; ancestral pCM410	-255.66	-844.94	1
A1	11 bp deletion (3791–3801bp)	-260.28	-842.79	0.038
	11 bp deletion (3794–3804 bp)	-261.36	-845.69	0
	37 bp deletion (3766–3802 bp)	-268.29	-840.87	0.005
A2	ISMex4 insertion between 3771 and 3772 bp	-251.16	-970.18	1
A3	T→C at 3804 bp	-254.82	-844.78	0

[a]The locations of mutations on pCM410 are described in terms of the plasmid file deposited in Genbank (Accession no. FJ389188).

[b]The folding energy (ΔG, expressed as negative terms) of mRNA surrounding the RBS_{fghA} at 30 °C was calculated using the Mfold program (Zuker, 2003). Two fragment lengths, 600 nt and 2000 nt, were chosen for analysis. The 600 nt mRNA fragment encompassed the 300 nt upstream and 300 nt downstream regions of the annotated fghA translation start site. The 2000 nt mRNA fragment encompassed the 1700 nt upstream and the 300 nt downstream regions of the annotated *fghA* translation start site. Only the energy of the most stable mRNA folding state was reported.

[c]The activity of translation initiation predicted by RBS calculator (Salis et al., 2009) at the annotated *fghA* translation start site was expressed relative to that of the ancestral pCM410.

Table S3. List of Primers for Sequencing pCM410, Related to Figure 1

Primers	Sequence	Corresponding location on pCM410
S14	5'-GGCTTCGACGGCGTTTCTG-3'	94 to 76 bp
HC410p26	5'-CACCGCTAACCTGTCTTTTAACCTG-3'	507 to 531 bp
HC410p17	5'-CGGGAGGGTTCGAGAAGG-3'	635 to 618 bp
HC410p12	5'-GCAGAAGTGGTCAGCTTG-3'	731 to 714 bp
HC410p24	5'-TCGCTGAGATAGGTGCCTC-3'	897 to 915 bp
HC410p25	5'-TGGAACGAAAACTCACGTTAAGGGA-3'	1065 to 1041 bp
S17	5'-TTTCCGAAGGTAACTGGCTTC-3'	1205 to 1225 bp
S16	5'-TGCTACAGAGTTCTTGAAGTG-3'	1300 to 1280 bp
HC410p10	5'-AGTCAGTGAGCGAGGAAG-3'	1851 to 1868 bp
HC160p1	5'-GCGGATCTCACACAGG-3'	2434 to 2449 bp
S1	5'-GAGAAGACAAGATGAGAACC-3'	2611 to 2630 bp
S4	5'-CGGGCCTTCCAGGTTGAC-3'	2696 to 2679 bp
S2	5'-ACCGCCAAGGTGGAAATCG-3'	3165 to 3183 bp
S18	5'-GATCGACGCCGATGATCATG-3'	3279 to 3260 bp
S3	5'-ACATCAACAAGGGCTTCGAC-3'	3685 to 3704 bp
HC410p16	5'-CCAGCGTCATGGGTTCTTC-3'	3812 to 3794 bp
HC410p22	5'-GAATGGGCGGCGGAATACG-3'	4013 to 4031 bp
HC410p20	5'-GCTTCTATGTCGATGCGACCGA-3'	4116 to 4137 bp
HC410p23	5'-GGTGACATAGTGCCACATCCTGA-3'	4180 to 4158 bp
HC410p21	5'-GAAGGTCATGGCGATGGTCAG-3'	4291 to 4271 bp
S5	5'-AGCAGGGCTACGATCACAG-3'	4560 to 4578 bp
M13Forward	5'-GTAAAACGACGGCCAGT-3'	4710 to 4694 bp
HC160p2	5'-GCTGCAAGGCGATTAAGTTG-3	4762 to 4743 bp
HC410p13	5'-GGCACCGAATGCGTATGATTC-3'	5591 to 5611 bp
S7	5'-AGTAAAGCGCCGGCTGCTG-3'	5670 to 5688 bp
S6	5'-GGACCTGTTGAACGAGGTC-3'	5755 to 5737 bp
HC410p18	5'-GAAAACTCACCGAGGCAGTTCCATAG-3'	6671 to 6696 bp
S9	5'-GACTGAATCCGGTGAGAATG-3'	6811 to 6830 bp
HC410p19	5'-TCAGTCGTCACTCATGGTGATTTCTCA-3'	6816 to 6790 bp
S8	5'-AATGAATAACGGTTTGGTTGATG-3'	6925 to 6903 bp
S11	5'-GACTTCCGGCAAGCTATACG-3'	7869 to 7888 bp
HC11p2	5'-CCTCTGTTTATCGGCAGTTCGTAG-3'	7913 to 7936 bp
S10	5'-CACTTGCTTCGCTCAGTATC-3'	7977 to 7658 bp
HC410p11	5'-GCTTCGTGTGTTTCAGCAAC-3'	8828 to 8847 bp
HC410P4	5'-CAGGTCAGCGAGGCCAAG-3'	8873 to 8856 bp

Primers	Sequence	Corresponding location on pCM410
HC410P7	5'-GCGAGGAACTATGACGACCAAG-3'	8928 to 8907 bp
HC410p14	5'-GGCTTGCGATGGTTTCG-3'	9122 to 8138 bp
S13	5'-CGATGGTTTCGGCATCCTC-3'	9128 to 9146 bp
HC410p15	5'-CGAAACCATCGCAAGCC-3'	9138 to 9122 bp
S12	5'-CCCGCAAGGAGGGTGAATG-3'	9225 to 9207 bp
HC410p8	5'-CACGAATACCAGCTCCGC-3'	9381 to 9398 bp
S15	5'-TCCGTCGACCCTTTCCGAC-3'	10131 to 13 bp
HC410p9	5'-GACGGATCTTTTCCGCTG-3'	10136 to 10119 bp

Table S4. List of Strains and Plasmids, Related to Figure 1

Strain or plasmid	Description	Source or reference
Strains		
CM502	Wild-type, *crtI*^{502}	(Marx, 2008)
CM624	*crtI*^{502}, Δ*mptG*	(Chou et al., 2011)
CM701	Δ*mptG*, pCM410	(Chou et al., 2011)
CM702	*crtI*^{502}, Δ*mptG*, pCM410	(Chou et al., 2011)
CM1180	*crtI*^{502}, *katA*::(*loxP*-t_{rrnB}-P_{tacA}-*Venus*-t_{T7})	(Lee et al., 2009)
CM1232	*crtI*^{502}, *katA*::(*loxP*-t_{rrnB}-P_{tacA}-*mCherry*-t_{T7}), Δ*mptG*, pCM410	(Chou et al., 2011)
CM2055	*crtI*^{502}, *katA*::(*loxP*-t_{rrnB}-P_{tacA}-*Venus*-t_{T7}), pCM160	This study
Plasmids		
pCM160	P_{mxaF} expression vector; [a]Km^r	(Marx and Lidstrom, 2001)
pCM410	P_{mxaF}-*flhA-fghA*; Km^r	(Chou et al., 2011)
pHC05	P_{tacA}-*mCherry*; [b]Tc^r	(Lee et al., 2009)
pHC17	pCM410 with the *trfA*::ISMex25^{935} allele	This study
pHC18	pCM410 with the *trfA*::ISMex25^{937} allele	This study
pHC19	pCM410 with the *trfA*::ISMex25^{956} allele	This study
pHC24	Cointegrate formed by pHC17 and p2META	This study
pHC25	Cointegrate formed by pHC19 and p2META	This study
pHC26	Cointegrate formed by pHC18 and p2META	This study
pHC27	pCM410 with the truncated *trfA* allele	This study
pHC83 (pCM410^{B2})	pCM410 with the 1109-bp deletion from pCM410^{B1}	This study
pHC84 (pCM410^{B3})	pCM410 with the single base substitution from pCM410^{B1}	This study
pHC96	pCM160 with *flhA* from pCM410	This study
pHC97	pCM410 with *flhA* removed	This study
pHC111	pCM410 with the CuO_{cmt} operator	This study
pHC112	pHC111 with the repressor gene *cymR*	This study
pHC113m	pHC112 with the *flhA-fghA* cassette replaced by *mCherry*	This study
pHC115	pHC112 with the *flhA-fghA* cassette replaced by a polylinker	This study
p2META	A native plasmid of *Methylobacterium*	(Vuilleumier et al., 2009)

[a]Km^r, kanamycin resistance

[b]Tc^r, tetracycline resistance.