Supplementary Information

Evolved plasmid-host interactions reduce plasmid interference cost

Hirokazu Yano^{1,2,4}, Katarznya Wegrzyn³, Wesley Loftie-Eaton^{1,2}, Jenny Johnson¹, **Gail E. Deckert1,2, Linda M. Rogers1,2, Igor Konieczny3 , and Eva M. Top1,2**

¹ Department of Biological Sciences, ² Institute for Bioinformatics and Evolutionary Studies, University of Idaho, Moscow, Idaho 83844, USA.

³ Department of Molecular and Cellular Biology, Intercollegiate Faculty of Biotechnology, University of Gdansk, 24 Kladki, 80-822 Gdansk, Poland.

⁴ Faculty of Life and Environmental Sciences, University of Tsukuba, Tsukuba, Ibaraki, 305-8571 Japan

List of contents

Plasmid and strain constructions

Table S1. Strains and plasmids used in this study

Table S2. Oligonucleotides used in this study.

Table S3. Statistical evaluation of effect of DnaB overexpression on plasmid persistence dynamics

Table S4. Maximum likelihood estimates (MLEs) of plasmid persistence parameters*^a*

Fig. S1. TrfA expression levels in *S. oneidensis* strain MR-1 carrying pMS0506 or its evolved variants.

Fig. S2. Distribution of cell length of *Shewanella oneidensis* MR-1 harboring ancestral plasmid pMS0506, its evolved variant pEvo-Sh1, pBBR1-MCS2, and its derivatives carrying *trfA1*(M124L), *trfA2*, or *trfA1(*Δ43/M124L).

Fig. S3. Loss of TrfA1 production improves persistence of pBP136Km in *P. putida* populations.

Fig. S4. ELISA showing reduced binding affinity to DnaB of TrfA1 variants.

Fig. S5. Effect of DnaB overproduction on the persistence of ancestral plasmid pMS0506.

Fig. S6. Effect of DnaB overproduction on the persistence of pEvo-Sh1.

Fig. S7. Alignment of DnaB proteins of representative gammaproteobacterial species.

Plasmid and strain constructions

1. Constructs used for fitness cost and plasmid copy number analysis.

DNA cloning was generally performed using *E. coli* strain EC100. EC100D *pir*+ cells were used for DNA cloning into pHY835LAC (Table S1). Phusion DNA polymerase (New England Biolab) was used for PCR. The *trfA* genes were amplified from pMS0506 and its evolved variants (pEvo-Sh1, pEvo-Sh11, pEvo-Sh5, pEvo-Sh13, pEvo-Sh3) by PCR, and the PCR products were cloned into pHSG398. All strains and plasmids are listed in Table S1. The M124L mutation was introduced into *trfA1* and its variants on the pHSG398 derivative, by amplifying the pHSG398 derivative with primers trfAM124LF, trfAM124LF (Table S2), phosphorylation and self-ligation of the PCR products. After the sequences were confirmed to be correct, the insert region in the pHSG398 derivatives was amplified using primers SacI-SD-pBP136trfA1F, and BamHindTrfAR. The PCR products were digested with *Sac*I and *Hin*dIII, and cloned into *Sac*I-*Hin*dIII site region of pBBR1-MCS2. The *trfA2* genes were amplified using primers SacI-SD-pBP136trfA2, BamHindTrfAR using pMS0506 as template, and cloned into pBBR1-MCS2. This gave rise to pHY1010, pHY1011, pHY1012, pHY1014, pHY1015, pHY987, and pHY988. Using the same method, *trfA1* and its variants were cloned into the *Sac*I-*Hin*dIII sites region within the mini-Tn*7* region of pHY835LAC. This gave rise to pHY884, pHY885, pHY886, pHY888, pHY890, pHY892, pHY897, and pHY895.

For construction of copy number reference plasmid pHY873, the 2-kb fragment containing *atp* genes of strain MR-1 was amplified from genomic DNA of MR-1 using primers EcoMR1oriCF, EcoMR1oriCR (Table S2). The PCR product was digested with *Eco*RI and the 1.8-kb fragment was cloned into the *Eco*RI site of pHY872 (Yano *et al*, 2012). This gave rise to pHY873.

The *dnaB* gene of *S. oneidensis* was PCR amplified using primers SacSDdnaBMR1F, HinddnaBMR1R, and the PCR product was cloned into pHSG398. After the sequence of the cloned *dnaB* was confirmed to be correct, the *Sac*I-*Hin*dIII region of the pHGS398 derivative was moved into the equivalent region of pHY835LAC. This gave rise to pHY1018.

2. Constructs used for biochemical assays

For biochemical assays, monomeric TrfA was used. For this purpose we introduced a mutation into each *trfA* gene on the pHSG398 derivatives constructed above by PCR amplification using primers Q279D/S292Lr and Q279D/S292Lf (Table S2), and subsequent phosphorylation and self-ligation of the PCR products. After confirmation of mutations in the pHSG398 derivatives, their respective *trfA* region was amplified using primers EcoNdeHisTrfA1F and BamHindTrfAR. The PCR products were digested with *Nde*I and *BamH*I, and then cloned into *Nde*I-*Bam*HI sites region of pET11a. This gave rise to pHY915, pHY921, pHY916, pHY917 and pHY919. The *dnaB* gene of MR-1 was

amplified with primers EcoNdeDnaBmr1F and HindDnaBmr1hisR (Table S2). The PCR products were digested, and then cloned into *Nde*I-*Hind*III sites of pET22b. This gave rise to pHY1032.

3. Insertion of mini-Tn7-*trfA* in the chromosome

To insert mini-Tn7 carrying *trfA* or *dnaB* into chromosomes, pHY835LAC derivatives were first introduced into *E. coli* strain BW29427 in the presence of diaminopimelic acid. Tn7 transposase expression plasmid was also introduced into BW29427. Two donor strains BW29427 (pHY835 derivatives), BW29427 (pUX-BF13) and one recipient strain, either MR-1 or BW25113 were grown overnight and one ml of each culture was pelleted and pooled in 1.5 ml tube. Cell suspensions were placed on LBA containing 100 mM diaminopimelic acids. The folowing day, mating mixtures were serially diluted in saline and plated on LBA with 30 μg/ml Gm. The colonies were re-streaked once on the same selective agar to purify. Presence of mini-Tn7 in the specific site in the chromosome (*att*Tn*7*) was confirmed by PCR using primer pairs, E coli glmS and Tn7R109 (Table S2) for *E. coli* BW25113, and MR-1glmS and Tn7R109 for *S. oneidensis* MR-1.

Table S1. Strains and plasmids used in this study

pHY924 pHY872 carrying the 2.5 kb *oriC* region of the *E.coli* chromosome; Tc^r (Yano *et al*, 2012) ^a Antibiotic resistance: Ap, ampicillin; Km, kanamycin; Gm, gentamycin; Cm, chloramphenicol, Sm, streptomycin. Rif, rifampicin. Tc, tetracycline. *dnaB⁺* and *dnaB* indicate presence or absence of additional *dnaB* copy in the chromosome.

^b *trfA* genotypes: *trfA1*wt/*trfA2*, wild-type *trfA* allele encoding both TrfA1 and TrfA2; M124L, a point mutation from ATG to CTC in the 124th codon, substituting methionine to leucine of the product and eliminating translation of TrfA2; R31P, a point mutation from G to C at position 92 of the *trfA1* frame, substituting the 31th arginine with proline of the product; A25T, a point mutation from G to A at position 73 of the *trfA1* frame, substituting the 25th alanine with threonine of the product; Δ5, an deletion from 93 to 107, deleting five aa of the product; Δ43, an deletion from position 97 to 225 in the *trfA1* frame, deleting 43 aa of the product; Δ77, an deletion from position 24 to 254 in the *trfA1* frame, deleting 77 aa of the product; Q279D/S292L, a mutation from C to G at position 835, substituting $279th$ glutamine to asparagine, and mutations from TC to CT at positions 874-875, substituting the 292th serine to leucine of TrfA1; these mutations make TrfA1 monomeric. The same nomenclature was applied to the equivalent mutations in TrfA2.

Table S2. Oligonucleotides used in this study.

Pair	Reference [Host (plasmid) /IPTG	Subject [Host (plasmid) / IPTG conc.	BIC.sep	BIC.joint	ΔBIC (BIC.sep -	Welch t.test p.value
	conc./ $Expt^2$.]	/Expt.]			BIC.joint)	
(i)	HY0759(pMS0506)/0 uM IPTG/B	HY1014(pMS0506)/0 uM IPTG/B	136.9	377.7	$-240.8*$	$0.011*$
(ii)	HY0759(pMS0506)/10 uM IPTG/A	HY1014(pMS0506)/0 uM IPTG/A	205.0	212.4	$-7.3*$	0.123
(iii)	HY0759(pMS0506)/10 uM IPTG/A	HY1014(pMS0506)/10 uM IPTG/A	227.4	252.3	$-24.8*$	$0.034*$
(iv)	HY0759(pMS0506)/100 uM IPTG/A	HY1014(pMS0506)/100 uM IPTG/A	236.0	355.9	$-119.9*$	$0.018*$
(v)	HY1014(pMS0506)/0 uM IPTG/A	HY1014(pMS0506)/10 uM IPTG/A	215.7	212.8	2.8	0.171
(vi)	HY1014(pMS0506)/10 uM IPTG/A	HY1014(pMS0506)/100 uM IPTG/A	245.6	274.6	$-28.9*$	$0.028*$
(vii)	HY1014(pMS0506)/100 uM IPTG/A	HY1014(pEvo-Sh1)/100 uM IPTG/A	237.2	466.3	$-229.0*$	$0.013*$
(viii)	HY0759(pMS0506)/10 uM IPTG/A	HY0759(pMS0506)/100 uM IPTG/A	217.7	208.7	8.9	0.218

Table S3. Statistical evaluation of effect of DnaB overexpression on plasmid persistence dynamics*¹*

1. BIC.sep: the BIC of the model that assumes that two given plasmid persistence data sets are governed by different dynamics; BIC.joint: the "null" model that the stability dynamics in both data sets are the same. More negative ΔBIC values are indicative of larger differences between two plasmid persistence dynamics. *: The difference is significant. We also conducted a Welch t.test on the plasmid-containing fractions on Day 10 or Day 4 (for pair (i)).

2. Strains HY0759 and HY1014 have a mini-Tn7 inserted in their chromosome, respectively without and with an extra copy of *dnaB*. Experiments performed in parallel are indicated by the same letter. Data from experiment A are shown in Fig. 3. We could not obtain data for pMS0506 in strain HY0759 with 0 uM IPTG condition due to a technical mistake in experiment A. An additional experiment B was conducted on different days and by a different person. Data set "HY0759(pMS0506)/10 uM IPTG/B" was used as control to evaluate data set "HY1014(pMS0506)/0 uM IPTG/B".

Table S4. Maximum likelihood estimates (MLEs) of plasmid persistence parameters*^a*

^{a.} The model used is the following segregation-selection (SS) model (Ponciano *et al*, 2007): $X_t = [2(1 - n_{t-1}) + 2^{1+s}n_{t-1}][n_{t-1}2^{1+s} + 2(1 - n_{t-1})]$, where Xt is a fraction of plasmid-free cell at generation time *t*, n_t is the plasmid-containing fraction at generation time *t*. The number of plasmid-free colonies among 52 colonies over time was used to estimate parameters. MLEs were obtained after 100 times bootstrapping. We used the SS model for pMS0506 and pEvo-Sh1 since these plasmids are not transmissible. We also used the SS model and not the HT (horizontal transfer) model to estimate the parameters for experiment C because it was a better fit for the data. c.i.: Credibilty interval.

^{*b*.} A, B, C: data sets with the same letter were generated at the same time. Rows with rep1, rep2, rep3 show parameter estimates for each of triplicate assays. Note that high initial plasmid free fraction in pMS0506-carrying population is reproducible. Other rows are the estimates for the pooled data of the triplicate assays.

c. The time (in generations) it takes for the plasmid to be lost from 99% of the population given the average estimated parameters shown in the previous columns.

Fig. S1. TrfA expression levels in *S. oneidensis* **strain MR-1 carrying pMS0506 or its evolved variants.** In each lane, 125 ug of total cell extract was loaded. TrfA quantity standards (STDs; 1.25, 0.63, 0.31, 0.16 pmol for both purified proteins TrfA1 and TrfA2) are loaded in the left four lanes. TrfA proteins were detected as previously described (Yano *et al*, 2012). A. (i) Example of blot for pMS0506 and four evolved plasmids pEvo-Sh11, pEvo-Sh5, pEvo-Sh13, pEVo-Sh1. Quantitation of the blot is shown in panel A (i). B (i) Example of blot for pEvo-Sh3, pEvo-Sh14, pEvo-Sh15. (ii) Quantitation for blot shown in panel B (i). The relative TrfA levels below the blots are means from four replicate quantitative Western blotting experiments.

```
Fig. S2
```


Fig. S2. Distribution of cell length of *Shewanella oneidensis* MR-1 harboring ancestral plasmid pMS0506 (**A**), its evolved variant pEvo-Sh1 (**B**), pBBR1-MCS2 (**F**), and its derivatives carrying *trfA1*(M124L) (**C**), *trfA2* (**D**), or *trfA1(*Δ43/M124L) (**E**). Examples of cell images are shown next to the corresponding histogram. Note that the coefficient of variance (CV) of TrfA1-producing cells was consistently higher than that of strains without TrfA1. The difference in distribution was compared using one-tailed two-sample Kolmogorov-Smirnov test: pMS0506 vs pEvo-Sh1, *P* = 0.034; TrfA1 vs TrfA2, *P* < 0.001; TrfA1 vs Δ43, *P* < 0.001; TrfA1 vs pBBR1-MSC2, *P* < 0.001. Scale bar indicates 3.3 um.

Fig. S3.

Fig. S3. Loss of TrfA1 production improves persistence of pBP136Km in *P. putida* **populations.** Triplicate assays were shown as distinct curves. Persistence parameter estimates were shown in Table S4.

Fig. S4.

Fig. S4. ELISA showing reduced binding affinity to DnaB of TrfA1 variants. Data point indicates mean and standard deviation. M124L/Q279D/S292L were introduced into TrfA1 and its variants.

Fig. S5. **Effect of DnaB overproduction on the persistence of ancestral plasmid pMS0506.** Three replicate data sets obtained at the same time are separately shown within the same plot. The hosts used were *S. oneidensis* HY0759 for "DnaB- ", and HY1041 for "DnaB⁺" condition, respectively. In the experiment A: 0 uM IPTG condition (Top left), we could not obtain data for "DnaB" due to a technical mistake; instead, the "DnaB" condition at 10 uM IPTG condition was shown in the plot as reference.

Fig. S6. Effect of DnaB overproduction on the persistence of pEvo-Sh1. Upper panels: Three replicate data obtained at the same time are separately shown within the same plot (only two for "DnaB" at 100 uM IPTG). The hosts used were the same as Fig.S5. Lower panels, left: Data of pMS0506 and pEvo-Sh1 obtained under the same conditions are shown within the same plot for comparison.

Fig. S7. Alignment of DnaB proteins of representative gammaproteobacterial species. E.col, *Escherichia coli* MG1655; S.one, *Shewanella oneidensis* MR-1; P.put, *Pseudomonas putida* KT2440; P.aer, *Pseudomonas aeruginosa* PAO1; V.col, *Vibrio cholerae* O1 biovar El Tor str. N16961; H.inf, *Heamophilus influenzae* Rd KW20; M.cat, *Moraxella catarrhalis*; A.hyd, *Aeromonas hydrophila* subsp. hydrophila ATCC 7966; P.hal, *Pseudoalteromonas haloplanktis* TAC125.

References for Supplementary Information

- Amy, P.S., Schulke, J.W., Frazier, L.M., and Seidler, R.J. (1985) Characterization of aquatic bacteria and cloning of genes specifying partial degradation of 2,4 dichlorophenoxyacetic acid. *Appl Environ Microbiol* **49**: 1237-1245.
- Baba, T., Ara, T., Hasegawa, M., Takai, Y., Okumura, Y., Baba, M. *et al* (2006) Construction of *Escherichia coli* K-12 in-frame, single-gene knockout mutants: the Keio collection. *Mol Syst Biol* **2**: 2006.0008.
- Bagdasarian, M., Lurz, R., Ruckert, B., Franklin, F.C., Bagdasarian, M.M., Frey, J., and Timmis, K.N. (1981) Specific-purpose plasmid cloning vectors. II. Broad host range, high copy number, RSF1010-derived vectors, and a host-vector system for gene cloning in *Pseudomonas*. *Gene* **16**: 237-247.
- Kovach, M.E., Elzer, P.H., Hill, D.S., Robertson, G.T., Farris, M.A., Roop, R.M., and Peterson, K.M. (1995) Four new derivatives of the broad-host-range cloning vector pBBR1MCS, carrying different antibiotic-resistance cassettes. *Gene* **166**: 175-176.
- Ponciano, J.M., De Gelder, L., Top, E.M., and Joyce, P. (2007) The population biology of bacterial plasmids: a hidden Markov model approach. *Genetics* **176**: 957-968.
- Sota, M., Tsuda, M., Yano, H., Suzuki, H., Forney, L.J., and Top, E.M. (2007) Region-specific insertion of transposons in combination with selection for high plasmid transferability and stability accounts for the structural similarity of IncP-1 plasmids. *J Bacteriol* **189**: 3091-3098.
- Sota, M., Yano, H., Hughes, J.M., Daughdrill, G.W., Abdo, Z., Forney, L.J., and Top, E.M. (2010) Shifts in the host range of a promiscuous plasmid through parallel evolution of its replication initiation protein. *ISME J* **4**: 1568-1580.
- Yano, H., Deckert, G.E., Rogers, L.M., and Top, E.M. (2012) Roles of long and short replication initiation proteins in the fate of IncP-1 plasmids. *J Bacteriol* **194**: 1533-1543.