SUPPLEMENTARY INFORMATION

Compensatory mutations improve general permissiveness to antibiotic resistance plasmids

Loftie-Eaton, W., Bashford, K., Quinn, H., Dong, K., Millstein, J., Hunter, S.,, Maureen Thomason, Houra Merrikh, Ponciano, J.M., and Top, E. M.

I. Supplemental Methods: From sigma to s: contextualizing our model parameters within the traditional plasmid cost literature.

Traditional competition experiments are based on the exponential growth model

$$n_t = n_0 \mathrm{e}^{rt}, \qquad \text{(eq. 1)}$$

where r denotes the intrinsic growth rate, n_0 is the initial population size of the experiment and time t is assumed to be continuous. From this equation, it follows that

$$\frac{n_t}{n_0} = \mathrm{e}^{rt} \text{ or } \ln\frac{n_t}{n_0} = r(t-0), \qquad \text{(eq.2)}$$

where the 0 emphasizes that the t is really measuring the time elapsed since the beginning of the growth experiment. Equation (2) is the basis to easily estimate the intrinsic growth rate if at the end of a growth experiment one has initial population size and final population size estimates. To avoid confusion, let's denote the observed initial population size at time t_1 as n_1 and the final population size recorded at time t_2 as n_2. Then, the estimate of r can be computed by plugging these quantities into equation 2 and solving for its value:

$$\ln\left(\frac{n_2}{n_1}\right) = r(t_2 - t_1) \text{ or } \hat{r} = \frac{\ln n_2 - \ln n_1}{t_2 - t_1}.$$

This estimate is the slope of the population growth curve when it is plotted on a logarithmic (base e) scale. The ratio of two different estimates of the intrinsic growth rate is usually denoted as W. Now, in our model, time is measured in generations of plasmid-carrying cells. That is, one "average" generation time is set to be one unit of time in the model. Then, from one time step to the other, the number of plasmid-carrying cells simply doubles. The number of plasmid-free cells, however, will grow at a slightly different rate. If the plasmid imposes a "cost", that we will denote by σ, then the plasmid-free cells will grow at a rate slightly bigger than 2 between one time step and the other. The growth model of the plasmid-free cells (obviating conjugation and segregation) will then be given by:

$$n_t = n_{t-1} 2^{1+\sigma}. \qquad \text{(eq. 3)}$$

Note that this formulation is very general in that it does not preclude the difference in growth rate to be negative. That is, if σ were to be negative, then it would be indicating that the plasmid-carrying cells have an advantage of growth relative to the plasmid-free cells. Starting at time 0 to find population size at time 1, then iterating this operation by starting at time 1 to find the population size at time 2and so on, it is easy to find that

$$n_t = n_0\left(2^{1+\sigma}\right)^t, \text{ or } n_t = n_0\lambda^t, \qquad \text{(eq. 4)}$$

where $\lambda = 2^{1+\sigma}$ is the finite rate of increase. This growth model is a geometric growth model. Now, the exponential model can be linked with the geometric model. When t matches the same generation times than the discrete-time model,

$$n_t = n_0 e^{rt} = n_0\left(e^r\right)^t = n_0\lambda^t, \qquad \text{(eq. 5)}$$

where the finite rate of increase is now written as a function of the intrinsic growth rate. Equating these two ways of writing the finite rate of increase then gives

$$\ln\left(2^{1+\sigma}\right) = ln\left(e^r\right) \Leftrightarrow (1+\sigma)\ln 2 = r,$$

which implies that

$$\sigma = \frac{r - \ln 2}{\ln 2}. \qquad \text{(eq. 6)}$$

These two last equations then allow a direct transformation between our mathematical model cost estimates and the traditional competition experiments 'cost' estimates.

Traditional `cost' estimates take as the reference strain the plasmid-free strain, and aim at measuring how much slower than the plasmid-free strain does the plasmid-carrying strain grows. The ratio of the estimates of the intrinsic growth rates for both strains is used to estimate the `relative fitness', usually denoted as W in the following way: let r_{p^+} and r_{p^-} be the intrinsic growth rates of the plasmid carrying strain and plasmid-free strain respectively. Using equation 5, the finite rate of increase of these strains is usually defined as $\ln\lambda_{p^+}$ and $\ln\lambda_{p^-}^{\text{ref}}$, where the `ref' notation emphasizes that doubling time of the plasmid-free cells is taken as the reference time of the geometric growth model. In that case, denoting as S the plasmid cost, it follows that

$$W = \frac{r_{p^+}}{r_{p^-}} = \frac{\ln\lambda_{p^+}}{\ln\lambda_{p^-}^{\text{ref}}} = \frac{\ln 2^{1-s}}{\ln 2} = 1 - s.$$

In our models, on the other hand, the definition of the time interval is anchored in the growth rate of the plasmid-carrying cell, because it was the strain for which we had extensive records and had good information regarding the doubling time: we knew it was about 10 generations per day. In that case, the plasmid-carrying cells are assumed to double within one time step while the plasmid-free cells grow at a rate a little faster than 2. Writing, as we do in the main text, the cost as σ, it follows that the relative fitness can be written as

$$W = \frac{r_1}{r_2} = \frac{\ln\lambda_{p^+}^{\text{ref}}}{\ln\lambda_{p^-}} = \frac{\ln 2}{\ln 2^{1+\sigma}} = \frac{1}{1+\sigma}.$$

Equating both definitions of the relative fitness results in a simple equation to convert our `cost' to the cost usually reported in competition experiments:

$$1 - s = W = \frac{1}{1+\sigma}.$$

Then, solving for s as a function of σ gives:

$$s = \frac{\sigma}{1+\sigma} \quad \text{and}$$

whereas solving for σ as a function of s gives:

$$\sigma = \frac{s}{1-s}.$$

Because both the data from our competition assays and plasmid persistence assays have the same structure (counts of resistant and total colonies at multiple time points), we were able to write a joint likelihood for the data obtained from both assays that leverages the total information to obtain a single estimate of σ. This likelihood function is just the product of the two likelihood functions for each particular data set (See De Gelder *et al.* 2004, equation 9, or Ponciano *et al.* 2007, equation 12), written under the constraint that the cost parameter needs to be kept identical under both assays. This joint estimation considerably improves the precision of the estimates of cost (see Table S3). Combining the information from both assays is particularly helpful when the plasmid persistence assays do not change much over time. This is because the persistence assays, by beginning with 100% plasmid-bearing cells, impose a waiting time until enough plasmid-free cells appear to exert an observable fitness differential. In contrast, the competition assays begin with a

50/50 proportion of plasmid-free and plasmid-bearing strains, making it more likely that a significant difference in fitness will be measured when the plasmid has a non-zero fitness cost or benefit.

References:

1. De Gelder, L., Ponciano, J. , Abdo, Z., Joyce, P., Forney, L.J., & Top E.M.. 2004. Combining mathematical models and statistical methods to understand and predict the dynamics of antibiotic sensitive mutants in a population of resistant bacteria during experimental evolution. Genetics 168: 1131-1144.
2. Ponciano J.M., De Gelder, L., Top, E. M. & Joyce P. The population biology of bacterial plasmids: a hidden Markov model approach. *Genetics* **176,** 957-968 (2007)

II. Supplemental Figures

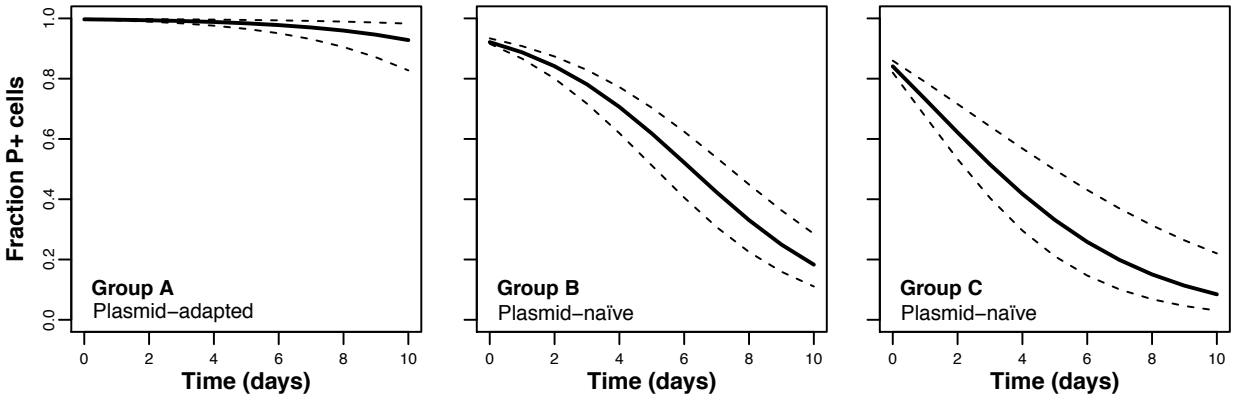

Fig. S1. Modeled plasmid persistence profiles for each group (See Fig. 2) show that plasmid persistence was significantly greater for all RP4-adapted strains (Group A) than in the plasmid-naïve strains in Groups B and C. Plasmid persistence profiles in Group B (evolved plasmids in ancestral host (H_AP_E)) and Group C (all ancestral or control hosts with the ancestral plasmids) were more similar yet distinct (also see Supplemental Tables 2, 3).

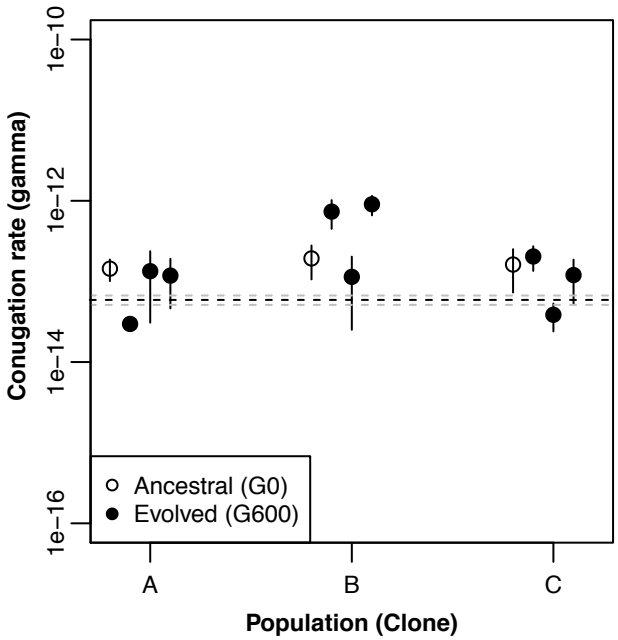

Fig. S2. Conjugation rate for the ancestor (empty circle) and three evolved (filled circle) clones from the three replicate populations. The black (mean) and grey (standard deviation) dotted lines represent the background conjugation frequency that occurs on the agar plates rather than in the liquid media and was measured by spreading separate donor and recipient cultures, at densities similar to the experimental mixed cultures, onto the same transconjugant-selective agar plate. No significant differences were observed between any evolved clones and their corresponding ancestor based on Welch t-tests. Evolved clones in each population are ordered sequentially from 1 to 3.

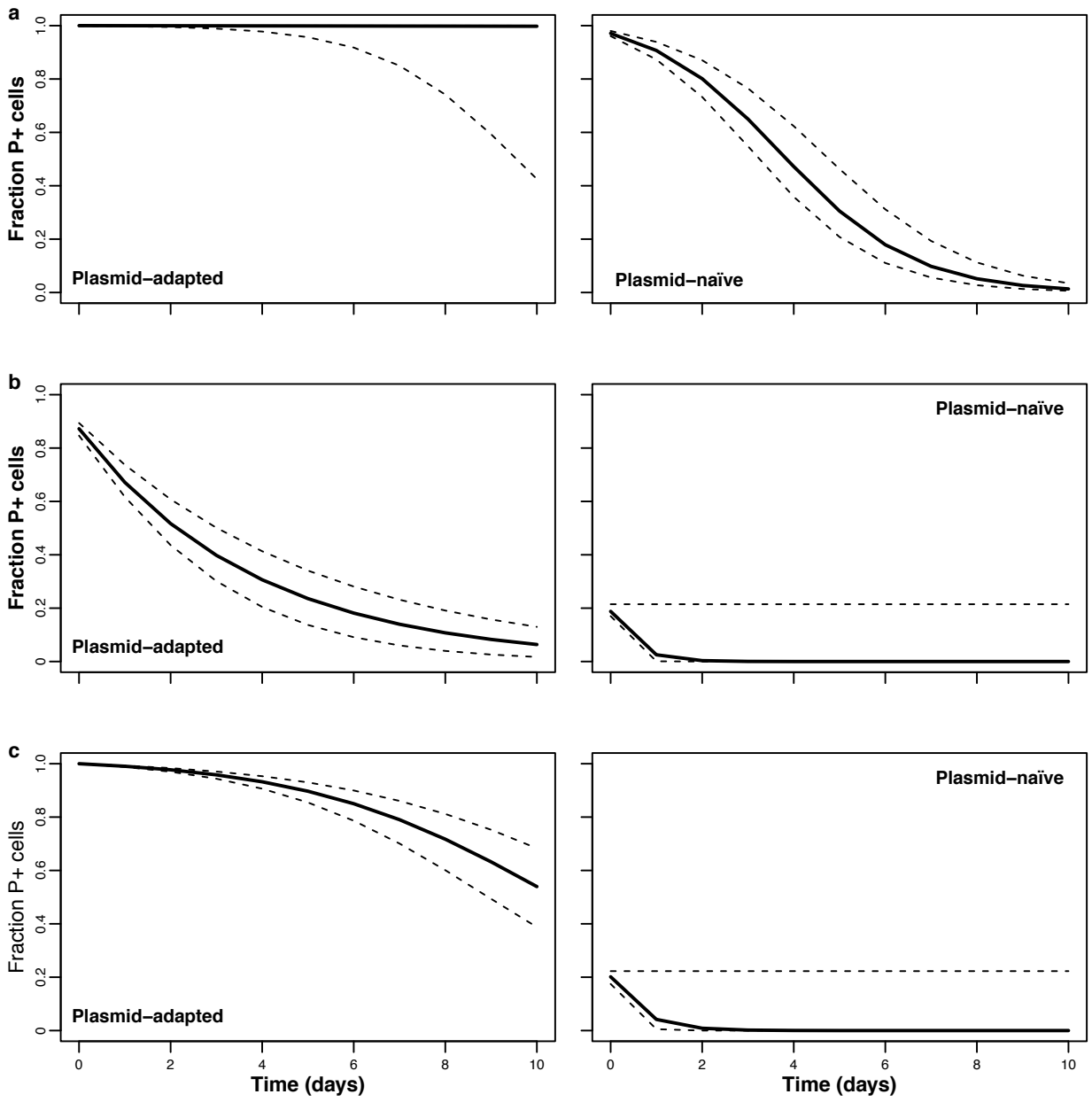

Fig. S3. Modeling of the group profiles for ancestral, evolved and control clones containing either of the three alternate plasmids, pB10, Rsa and RSF1010 (A to C, respectively) shows that plasmid persistence in RP4-adapted hosts was significantly greater than in the plasmid-naïve hosts (see also Supplemental Table 4).

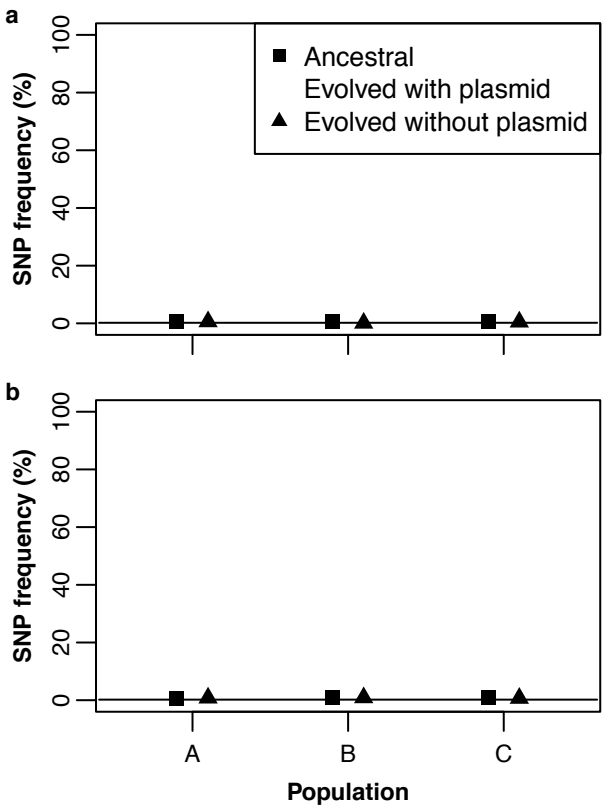

Fig. S4. The frequency of the (A) Xpd/Rad3$_{D672A}$ and (B) P$_{UvrDA\text{-}32C}$ alleles in each of the ancestral, evolved and control populations. The solid line in each panel represents the sequencing error frequency for that specific nucleotide substitution measured over the whole length of the amplicon.

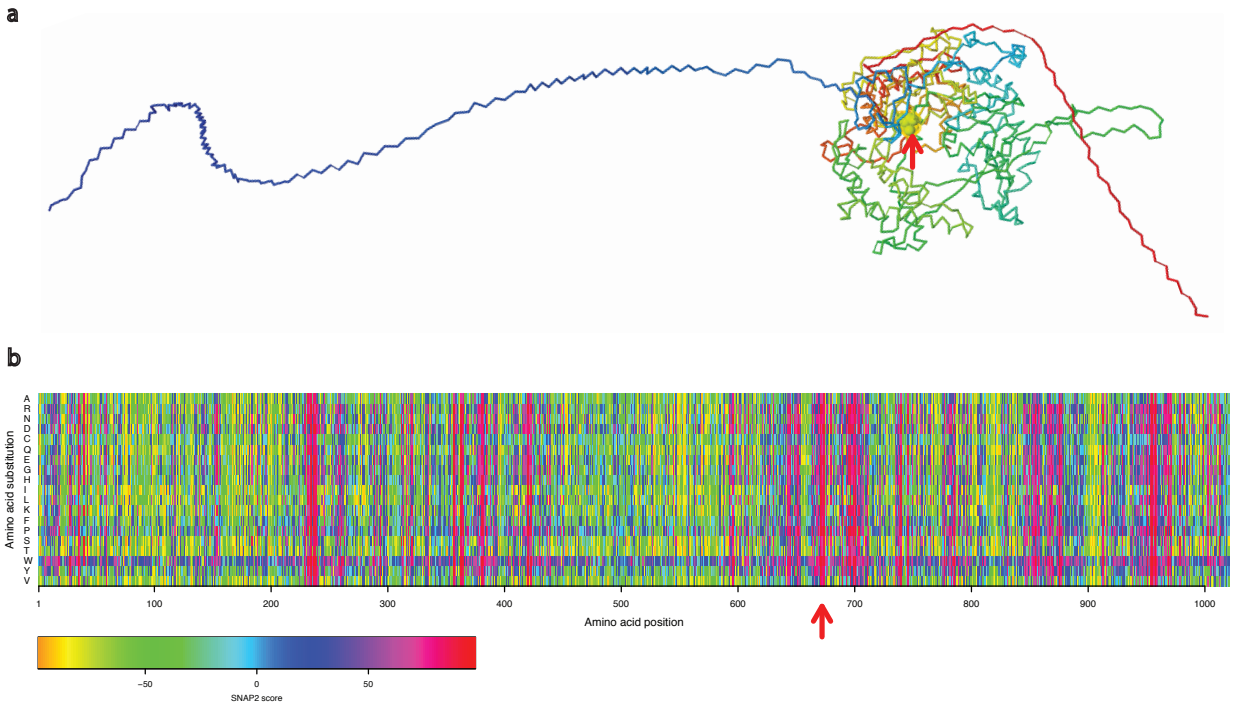

Fig. S5. Analysis of the CDS encoding the Xpd/Rad3-like helicase. (A) The Xpd/Rad3-like helicase domain (multicolored) is located at the C-terminus of the polypeptide. The structure of the N-terminus (dark blue) could not be predicted. (B). Effect of amino acid substitution on protein structure. A high SNAP2 score means that the specific substitution in that location is likely to have a strong effect on protein structure. The Xpd/Rad3$_{D672A}$ residue is indicated by a red arrow.

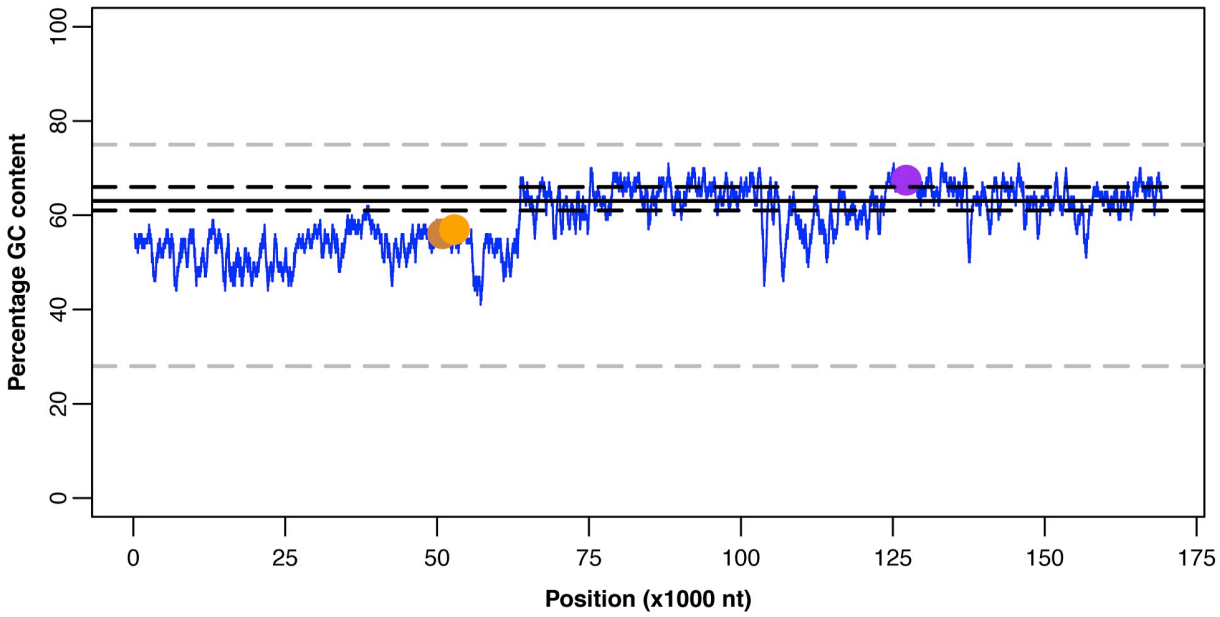

Fig. S6. Percentage GC content across the length of contig00024. The blue line represents the percentage GC content for contig00024. The solid black line represents the mean percentage GC content across the whole genome while the black and grey broken lines represent the first and third quartiles of the GC content across the whole genome, respectively. GC content was calculated using a 250 bp sliding window. All the SNPs found on this contig are indicated; the brown, orange and purple dots represent the Xpd/Rad3$_{D672A}$, P$_{UvrDA\text{-}32C}$ and BcsC$_{Q659Q}$ loci, respectively.

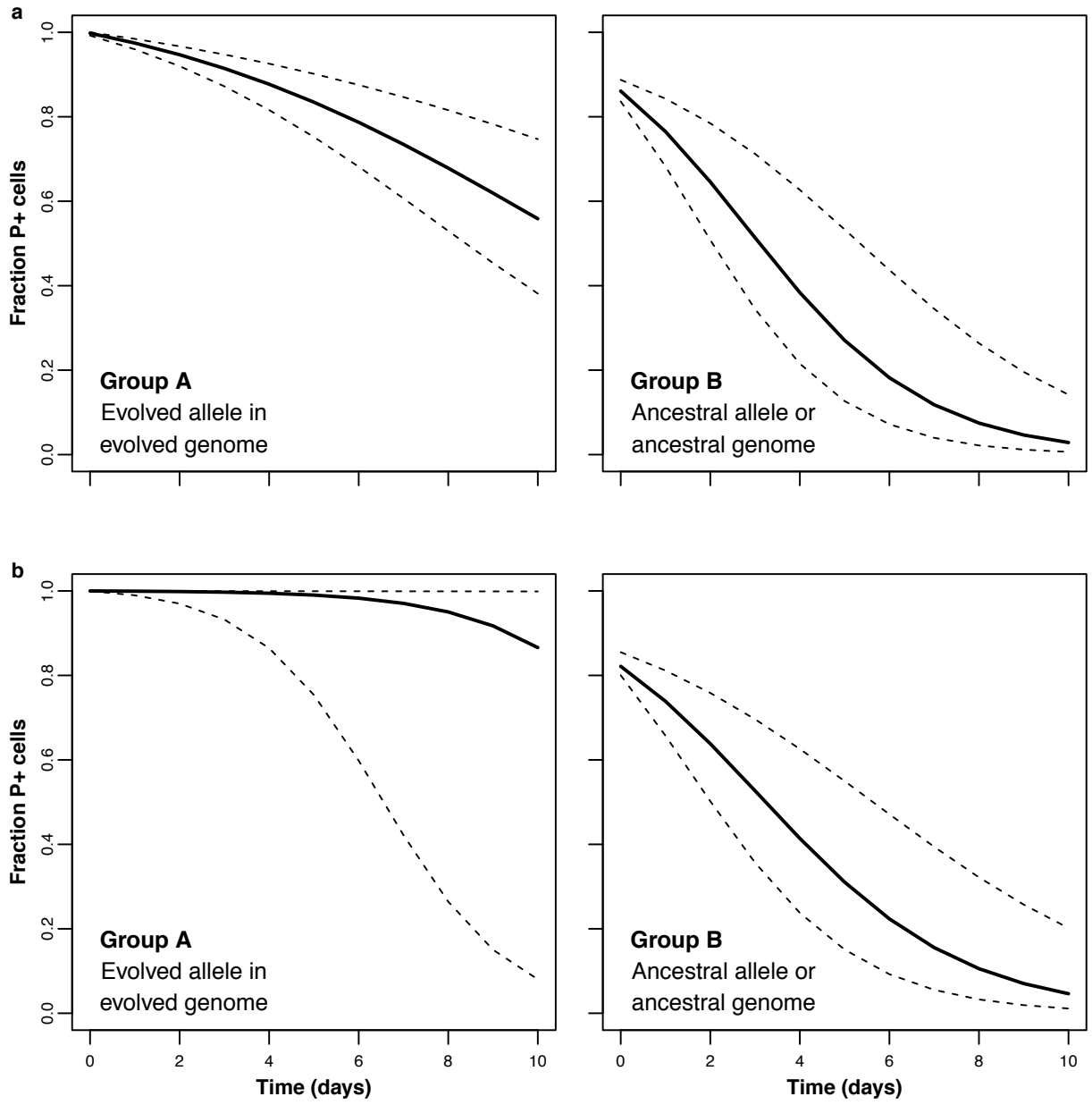

Fig. S7. Following the allelic exchange experiments, modeling of the grouped plasmid persistence profiles shows that Group A, which represents the evolved alleles in an evolved genome (A, Xpd/Rad3$_{D672A}$; B, P$_{uvrD_A\text{-}32C}$), exhibits greater plasmid persistence than Group B, which represents the ancestral helicase alleles (A, Xpd/Rad3; B, P$_{uvrD}$), in a control or ancestral genome, or the respective evolved alleles in the ancestral genome.

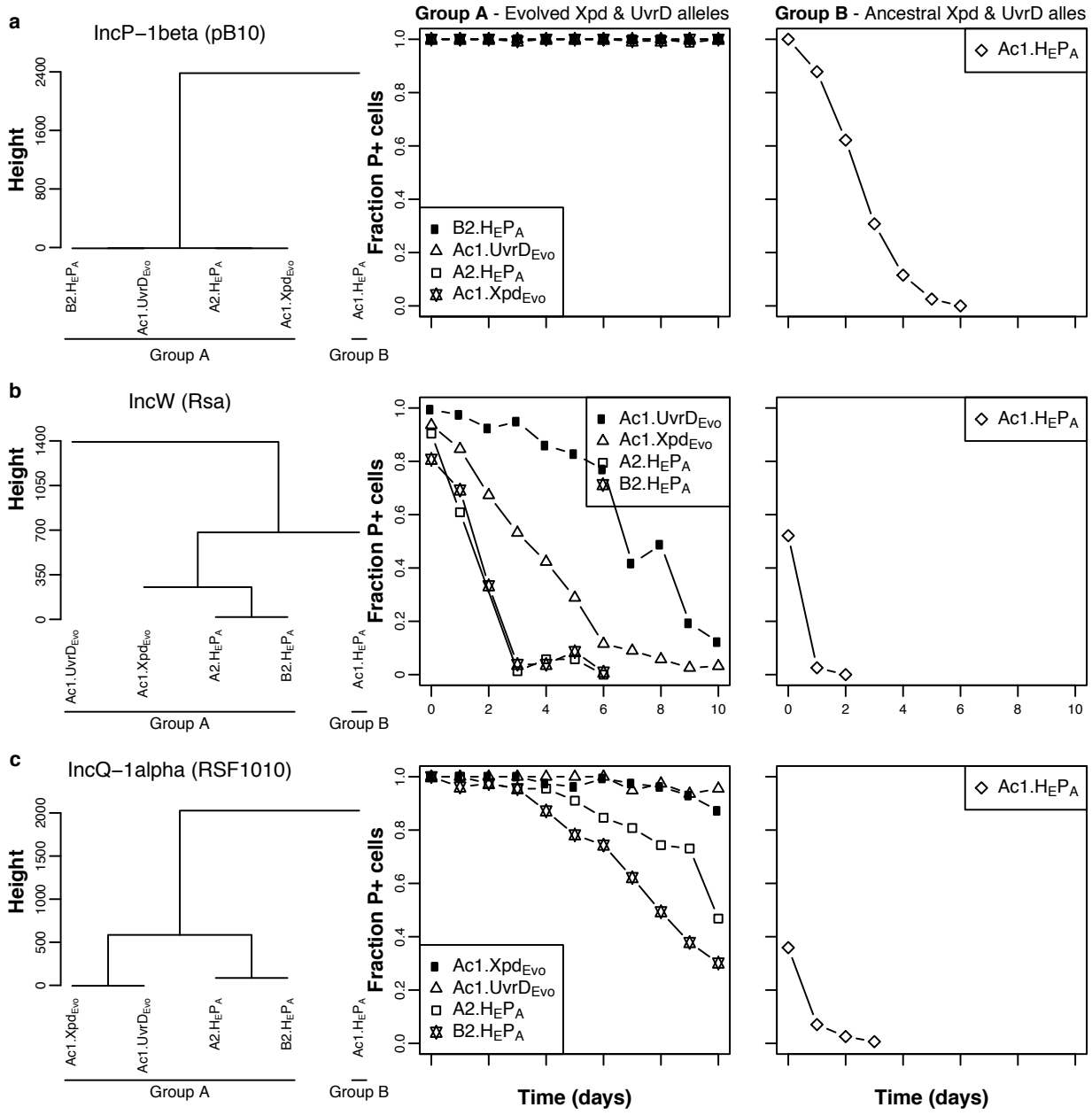

Fig. S8. The persistence dynamics of plasmids (a) pB10, (b) Rsa and (c) RSF1010 in control strain Ac1 containing either the evolved alleles Xpd/Rad3$_{D672A}$ or P$_{uvrD_A-32C}$ are more similar to those of the plasmid-adapted strains containing equivalent alleles, than to the Ac1 control strain containing the ancestral Xpd and UvrD alleles. Each data point represents the mean fraction of plasmid-containing (P+) cells (n = 3).

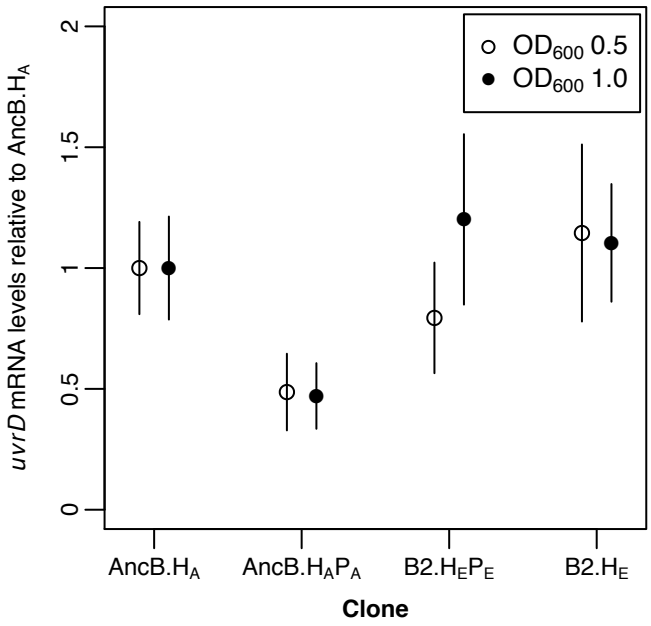

Fig. S9. *uvrD* mRNA levels for a plasmid-containing ancestral clone (AncB.H_AP_A)) and plasmid-containing and plasmid-free evolved clones (B2.H_EP_E and B2.H_E, respectively) relative to a plasmid-free ancestral clone (AncB.H_A), measured for mid-exponential (OD_{600} 0.5) and stationary phase (OD_{600} 1.0) cultures. Gene expression levels were normalized between samples using the 16S rRNA levels. Each data point consists of three biological and two technical repeats. Error bars indicate the standard error. The only strain that showed a significantly different *uvrD* gene expression level relative to the ancestor was the plasmid-containing ancestor AncB.HAPA (Adjusted P value < 0.05).

III. Supplemental Tables, except for Tables S5 and S9, which are in Supplementary file S2 (Excel).

Table S1. Estimated plasmid persistence dynamics and Bayesian Information Criteria (BIC) for the ancestral and evolved plasmid-host permutations, and the three groups into which they clustered. Parameter estimates are based on the persistence data (Figs. 2, S1).

		λ[A]			σ[B]			$T_{1\%}$[C]			
Strain permutation[D]	Group[E]	CI (2.5%)[F]	MLE[G]	CI (97.5%)	CI (2.5%)	MLE	CI (97.5%)	CI (2.5%)	MLE	CI (97.5%)	BIC[H]
AncA.H_AP_A	C	3.03E-09	5.60E-09	7.21E-03	3.83E-02	6.19E-02	6.67E-02	11	15	26	184
AncB.H_AP_A	C	1.62E-09	1.47E-03	7.20E-03	4.87E-02	6.61E-02	7.41E-02	10	13	20	206
AncC.H_AP_A	C	1.05E-09	6.73E-08	4.83E-03	6.03E-02	7.28E-02	7.86E-02	10	13	16	198
A1.H_EP_E	A	2.66E-05	2.33E-04	3.75E-04	5.07E-16	3.00E-08	8.91E-03	>100[I]	>100	>100	91
A2.H_EP_E	A	3.53E-09	2.21E-05	3.31E-04	2.67E-11	3.16E-02	5.50E-02	23	46	>100	83
A3.H_EP_E	A	1.24E-12	2.16E-11	1.35E-04	5.03E-11	3.67E-02	9.75E-02	15	45	>100	46
B1.H_EP_E	A	1.01E-09	4.67E-05	9.97E-05	3.80E-21	5.37E-09	8.67E-02	17	>100	>100	33
B2.H_EP_E	A	1.04E-12	4.81E-06	3.72E-05	2.65E-02	8.09E-02	2.78E-01	7	25	37	38
B3.H_EP_E	A	6.32E-35	2.33E-05	5.30E-05	8.32E-49	4.06E-15	1.14E-01	13	>100	>100	25
C1.H_EP_E	A	6.18E-12	9.18E-10	4.09E-04	5.11E-10	2.63E-02	4.34E-02	27	53	>100	91
C2.H_EP_E	A	1.57E-06	1.35E-05	4.76E-04	2.52E-11	9.39E-02	1.28E-01	11	20	>100	71
C3.H_EP_E	A	2.28E-12	3.32E-11	2.21E-05	8.01E-02	1.06E-01	1.52E-01	12	19	12	76
A1.H_EP_A	A	4.56E-05	1.33E-04	1.00E-03	3.86E-03	6.30E-02	8.51E-02	14	24	>100	130
A2.H_EP_A	A	2.38E-05	6.89E-04	1.49E-03	4.48E-15	1.83E-02	4.52E-02	23	56	>100	152
A3.H_EP_A	A	3.67E-05	1.05E-04	4.16E-04	9.88E-08	5.04E-02	7.06E-02	19	30	>100	82
B1.H_EP_A	A	3.46E-10	4.72E-06	1.05E-04	3.76E-17	7.13E-02	2.05E-01	8	28	>100	33
B2.H_EP_A	A	4.68E-05	1.31E-04	7.84E-04	2.15E-07	6.04E-02	8.23E-02	15	25	>100	96
B3.H_EP_A	A	6.48E-05	1.78E-04	3.65E-04	1.49E-13	2.16E-02	4.64E-02	28	60	>100	94
C1.H_EP_A	A	5.63E-11	9.54E-05	1.19E-03	4.03E-09	2.32E-02	3.34E-02	29	50	>100	136
C2.H_EP_A	A	1.29E-04	3.55E-04	4.50E-04	3.12E-11	2.15E-10	2.61E-02	44	>100	>100	132
C3.H_EP_A	A	1.59E-08	5.08E-04	1.11E-03	6.19E-08	1.56E-02	3.73E-02	28	67	>100	132
A1.H_AP_E	B	1.15E-10	2.78E-10	2.67E-03	4.96E-02	6.14E-02	6.61E-02	14	17	20	216
A2.H_AP_E	B	9.77E-11	1.55E-10	3.07E-03	5.88E-02	7.17E-02	7.90E-02	11	14	17	223
A3.H_AP_E	B	7.00E-11	4.80E-11	2.83E-03	3.88E-02	5.39E-02	5.76E-02	15	19	25	214
B1.H_AP_E	B	1.41E-10	1.19E-09	3.53E-03	3.34E-02	5.01E-02	5.57E-02	15	20	29	215
B2.H_AP_E	B	3.57E-11	5.51E-10	3.48E-03	4.52E-02	5.72E-02	6.39E-02	13	17	22	238
B3.H_AP_E	B	1.09E-10	9.95E-10	2.66E-03	3.58E-02	5.26E-02	5.68E-02	15	19	27	213
C1.H_AP_E	B	9.93E-11	3.84E-10	2.87E-03	5.12E-02	6.39E-02	6.98E-02	13	17	19	280
C2.H_AP_E	B	7.44E-10	9.43E-10	3.42E-03	3.73E-02	5.30E-02	5.61E-02	15	19	26	200
C3.H_AP_E	B	1.19E-10	1.38E-09	2.65E-03	4.20E-02	5.71E-02	6.14E-02	15	19	23	225
Ac.H_EP_A	C	1.34E-08	7.54E-03	1.65E-02	3.10E-02	5.36E-02	7.57E-02	7	12	32	189
Bc.H_EP_A	C	9.15E-03	2.09E-02	2.22E-02	1.24E-09	3.42E-09	2.63E-02	12	21	49	186
Cc.H_EPA	C	5.53E-03	1.32E-02	1.49E-02	1.22E-08	1.25E-08	2.12E-02	16	33	81	208
Group dynamics	A	1.28E-07	2.64E-05	1.29E-04	3.20E-02	4.33E-02	4.83E-02	28	34	49	2017
Group dynamics	B	3.16E-10	4.98E-11	1.14E-03	5.25E-02	5.71E-02	5.84E-02	16	18	20	2109
Group dynamics	C	5.68E-03	9.17E-03	1.29E-02	2.13E-02	3.21E-02	4.21E-02	12	17	26	1439

[A] λ, segregational loss frequency; [B] σ, plasmid cost; [C] $T_{1\%}$, time (in days) to 1% plasmid retention, or 99% plasmid loss; based on initial and final fraction of P- cells of 0 and 0.99, respectively; [D], H_AP_A ancestral host and plasmid, H_AP_E ancestral host and evolved plasmid, H_EP_E evolved host and plasmid, H_EP_A evolved host and ancestral plasmid; [E] Grouping is based on cluster analysis, see Fig. 2A; [F]CI, confidence interval; [G] Maximum Likelihood Estimate; [H] Bayesian information criteria; [I] exceeds 100 days.

Table S2. BIC for the segregation and selection (SS) and horizontal transfer (HT) models based on the plasmid persistence profile for each ancestral[a] and evolved[b] plasmid-containing clone.

Clone	BIC SS	BIC HT
AncA.H_AP_A	184	191
A1.H_EP_E	91	98
A2.H_EP_E	83	90
A3.H_EP_E	46	53
AncB.H_AP_A	206	213
B1.H_EP_E	33	40
B2.H_EP_E	38	45
B3.H_EP_E	25	32
AncC.H_AP_A	198	205
C1.H_EP_E	91	98
C2.H_EP_E	70	78
C3.H_EP_E	76	83

[a] Ancestral host and plasmid (H_AP_A)

[b] Evolved host and plasmid (H_EP_E)

Table S3. Segregational loss frequency (λ) and cost (σ) MLEs for the ancestral[a] and evolved[b] clones and the corresponding BIC. Parameters were jointly estimated from competition (Fig. 3) and persistence data (Figs. 2, S1)

Strain	λ			σ			BIC
	CI Low (2.5%)	MLE	CI High (97.5%)	CI Low (2.5%)	MLE	CI High (97.5%)	
AncA.H_AP_A	4.70E-08	6.16E-08	4.16E-03	4.16E-03	5.25E-02	6.55E-02	252
A1.H_EP_E	1.87E-04	4.79E-04	8.59E-04	8.59E-04	-4.44E-02	-2.47E-02	144
A2.H_EP_E	1.20E-04	4.36E-04	6.83E-04	6.83E-04	-4.69E-02	-2.81E-02	143
A3.H_EP_E	1.94E-10	1.28E-04	2.47E-04	2.47E-04	-3.54E-02	-2.58E-02	294
AncB.H_AP_A	1.67E-09	3.21E-07	4.65E-03	4.65E-03	6.52E-02	7.70E-02	301
B1.H_EP_E	6.57E-12	2.09E-04	4.51E-04	4.51E-04	-1.36E-01	-1.11E-01	84
B2.H_EP_E	9.78E-05	2.71E-04	4.79E-04	4.79E-04	-8.60E-02	-6.42E-02	98
B3.H_EP_E	1.07E-13	8.33E-05	1.77E-04	1.77E-04	-1.11E-01	-8.44E-02	78
AncC.H_AP_A	7.00E-10	5.20E-07	5.87E-03	5.87E-03	6.19E-02	7.89E-02	282
C1.H_EP_E	9.02E-04	1.52E-03	2.08E-03	2.08E-03	-1.23E-01	-1.02E-01	146
C2.H_EP_E	1.50E-03	2.29E-03	3.37E-03	3.37E-03	-1.53E-01	-1.25E-01	218
C3.H_EP_E	8.78E-04	1.46E-03	2.04E-03	2.04E-03	-1.09E-01	-8.48E-02	186

[a] Ancestral host and plasmid (H_AP_A)

[b] Evolved host and plasmid (H_EP_E)

Table S4. Estimated plasmid persistence dynamics and BIC for the three alternate plasmids in the ancestral-, evolved- and control hosts, and the groups into which they clustered.

			λ			σ			$T_{1\%}$			
Plasmid	**Host**	**Group[A]**	**CI (2.5%)**	**MLE**	**CI (97.5%)**	**CI (2.5%)**	**MLE**	**CI (97.5%)**	**CI (2.5%)**	**MLE**	**CI (97.5%)**	**BIC**
pB10	AncA	B	1.80E-10	1.84E-09	3.22E-04	1.46E-01	1.62E-01	1.81E-01	8	10	23	127
	AncB	B	1.32E-08	7.77E-04	1.41E-03	1.35E-01	1.47E-01	1.63E-01	7	9	22	165
	AncC	B	1.54E-03	2.34E-03	3.24E-03	1.15E-01	1.27E-01	1.44E-01	8	9	11	166
	A2*	A	6.32E-35	5.27E-07	5.84E-05	8.32E-49	9.71E-02	7.08E-01	3	24	>100	22
	B2*	A	6.32E-35	1.91E-13	2.30E-05	8.32E-49	3.28E-17	1.00E-01	16	>100	>100	19
	C2*	A	3.55E-11	3.50E-05	7.02E-05	2.91E-24	6.82E-14	1.00E-01	15	>100	>100	29
	Ac1	B	4.05E-03	6.54E-03	9.25E-03	1.53E-01	1.80E-01	2.13E-01	5	6	7	91
	Bc1	B	3.30E-10	2.24E-09	2.44E-03	1.28E-01	1.42E-01	1.54E-01	7	8	10	124
	Cc1	B	1.19E-02	2.10E-02	2.54E-02	5.45E-09	8.56E-03	3.21E-02	10	17	38	256
	Group dynamics	A	1.13E-14	2.33E-05	3.50E-05	1.31E-25	3.29E-05	9.77E-02	16	>100	>100	56
	Group dynamics	B	1.66E-03	3.08E-03	4.21E-03	8.91E-02	9.53E-02	1.02E-01	9	10	12	2056
Rsa[B]	A2*	A	1.23E-06	1.80E-02	3.92E-02	9.35E-02	1.65E-01	2.38E-01	3	5	12	90
	B2*	A	2.22E-02	4.06E-02	5.22E-02	7.97E-09	2.42E-02	7.84E-02	5	8	20	222
	C2*	A	6.35E-03	1.18E-02	1.57E-02	2.24E-04	1.71E-02	3.82E-02	13	22	71	184
	Ac1	B	4.36E-08	2.36E-07	2.83E-01	4.22E-08	5.52E-01	7.77E-01	0	1	>100	33
	Bc1	B	2.79E-06	1.08E-05	8.36E-01	2.09E-01	6.01E-01	6.35E-01	0	1	4	159
	Cc1	B	5.15E-08	1.33E-01	1.62E-01	1.27E-07	2.78E-07	3.84E-01	1	2	>100	36
	Group dynamics	A	1.95E-02	2.58E-02	2.76E-02	3.78E-09	1.22E-08	1.62E-02	12	17	23	739
	Group dynamics	B	3.03E-07	1.82E-01	2.20E-01	7.43E-07	4.04E-06	3.61E-01	1	1	>100	311
RSF1010	AncA	B	1.45E-07	1.60E-01	1.92E-01	6.55E-08	6.87E-07	3.48E-01	1	2	>100	48
	AncB	B	2.71E-07	9.49E-06	3.39E-01	5.27E-06	2.99E-01	2.87E+00	0	1	>100	23
	AncC	B	1.50E-06	8.87E-06	7.07E-01	2.32E-06	4.38E-01	7.18E+00	0	1	>100	27
	A2*	A	3.88E-04	5.92E-04	2.92E-03	9.23E-06	5.85E-02	6.87E-02	15	22	>100	132
	B2*	A	8.81E-04	1.28E-03	1.69E-03	5.47E-02	6.29E-02	7.30E-02	15	18	22	149
	C2*	A	1.11E-04	2.46E-04	1.25E-03	2.20E-06	5.10E-02	7.06E-02	17	27	>100	112
	Ac1	B	4.47E-08	1.33E-01	1.59E-01	2.95E-08	1.99E-07	3.20E-01	1	2	>100	50
	Bc1	B	1.72E-06	8.41E-06	8.50E-01	1.04E-01	5.28E-01	1.47E+00	0	1	5	35
	Cc1	B	1.65E-08	1.13E-01	1.49E-01	2.53E-08	1.59E-07	2.36E-01	1	3	>100	58
	Group dynamics	A	6.35E-04	8.03E-04	1.04E-03	4.55E-02	5.25E-02	5.85E-02	20	23	27	1124
	Group dynamics	B	8.56E-08	1.46E-01	1.63E-01	1.20E-07	1.78E-07	2.74E-01	1	2	>100	584

[A] Based on cluster analysis, see Fig. 4.

[B] Due to the high instability of plasmid Rsa in the ancestors, its persistence dynamics in this host could not be determined.

Table S5 is in Supplementary file S2.

Table S6. Estimated plasmid persistence dynamics and BIC for the allelic exchange mutants.

		λ			σ			$T_{1\%}$			
Strain	**Group[A]**	**CI (2.5%)**	**MLE**	**CI (97.5%)**	**CI (2.5%)**	**MLE**	**CI (97.5%)**	**CI (2.5%)**	**MLE**	**CI (97.5%)**	**BIC**
$Anc.Xpd_{Anc}$ [B]	B	3.98E-09	1.50E-08	7.20E-03	3.51E-02	6.03E-02	6.62E-02	12	47	83	178
$Anc.Xpd_{Evo}$	B	1.72E-09	1.86E-08	6.20E-03	6.00E-02	8.20E-02	8.82E-02	10	34	52	176
$Ac1.Xpd_{Evo}$	A	2.22E-03	3.07E-03	4.22E-03	3.31E-02	4.43E-02	5.39E-02	17	21	28	158
$Ac1.Xpd_{Anc}$	B	3.04E-09	1.05E-02	2.22E-02	4.50E-02	7.91E-02	1.09E-01	6	10	67	134
$A2.H_EP_A$	A	3.02E-06	6.89E-04	1.38E-03	6.31E-11	1.83E-02	3.92E-02	25	57	>100	152
$Anc.UvrD_{Anc}$	B	2.20E-09	1.84E-08	7.85E-03	3.95E-02	6.48E-02	6.99E-02	12	43	77	157
$Anc.UvrD_{Evo}$	B	5.67E-09	1.97E-03	8.61E-03	5.23E-02	7.22E-02	8.34E-02	10	15	56	163
$Ac1.UvrD_{Evo}$	A	7.19E-14	3.68E-10	3.97E-04	3.38E-12	2.54E-01	3.71E-01	4	14	>100	36
$Ac1.UvrD_{Anc}$	B	2.28E-08	9.82E-03	2.17E-02	5.15E-03	3.20E-02	5.47E-02	9	18	>100	152
$B2.H_EP_A$	A	4.84E-05	1.31E-04	8.31E-04	1.04E-06	6.04E-02	7.90E-02	15	25	>100	96

[A] Grouping is based on cluster analysis, see Fig. 6; [B] The evolved or ancestral alleles of $Xpd/Rad3_{D672A}$ and $P_{uvrD\_A\text{-}32C}$ are denoted as Xpd_{Evo}, Xpd_{Anc}, $UvrD_{Evo}$ or $UvrD_{Anc}$.

Table S7. ΔBIC for comparisons of the Xpd/Rad3 allelic exchange mutant persistence profiles.

	$Anc.Xpd_{Anc}$	**$Ac1.Xpd_{Anc}$**	**$A2.H_EP_A$**
$Anc.Xpd_{Evo}$	-30	NA	NA
$Ac1.Xpd_{Evo}$	NA	-902	-455

Table S8. ΔBIC BIC for comparisons of the $P_{UvrD\_A\text{-}32C}$ allelic exchange mutant persistence profiles.

	$Anc.UvrD_{Anc}$	**$Ac1.UvrD_{Anc}$**	**$B2.H_EP_A$**
$Anc.UvrD_{Evo}$	0.53	NA	NA
$Ac1.UvrD_{Evo}$	NA	-2071	-39

Table S9 is in Supplementary file S2.

Table S10. Primers used in this study.

Primers [A]	Purpose/Sequence [B]
16S-F	Quantitation of 16S rRNA: GGAAGGGCAGTAAGCGAATA
16S-R	Quantitation of 16S rRNA: CGCTTGCACCCTCTGTATTA
CS1-H2helicase1-F	Xpd/Rad3 SNP quantification; 5'- ACACTGACGACATGGTTCTACAACAATCTTGAAGTCGTCTTACGGA-3'
CS2-H2helicase1-R	Xpd/Rad3 SNP quantification; 5'- TACGGTAGCAGAGACTTGGTCTTGAGCCCTACATTCTCAAATGCTA-3'
CS1-H2helicase2prom-F	UvrD operon promoter SNP quantification; 5'- ACACTGACGACATGGTTCTACACTGAGGCCATCAACAAGCTTG-3'
CS2-H2helicase2prom-R	UvrD operon promoter SNP quantification; 5'- TACGGTAGCAGAGACTTGGTCTCGATGACGGTGATCTCGGTAAC-3'
pPS04_aph_F	Verify pPS04 insertion; 5'-GTGTTATGAGCCATATTCAACGGG-3'
pPS04_aph_R	Verify pPS04 insertion; 5'-GCCAGTGTTACAACCAATTAACCA-3'
UvrD_helicase_F	Quantitation of *uvrD*; 5'-GACGACATCACGAAATACGAGAAG-3'
UvrD_helicase_R	Quantitation of *uvrD*; 5'-GACCTTGATGATCTGGGTCACTC-3'
UvrD_Prom_F	Cloning 573-bp UvrD operon promoter fragment into pPS04; 5'-TCTAGACAATTTGCGGACATTCTATCG -3'
UvrD_Prom_R	Cloning 573-bp UvrD operon promoter fragment into pPS04; 5'-CATATGATTCCCACAGTGTTTGAGTTTAA-3'
Xpd/Rad3_F	Cloning 800-bp Xpd/Rad3 fragment into pPS04; 5'- TCTAGACTGGACAATCTTGAAGTCGTCTTAC -3'
Xpd/Rad3_R	Cloning 800-bp Xpd/Rad3 fragment into pPS04; 5'- CATATGAGGCCATCTTTGTTAACGATGTTG -3'

[A] All primers were designed specifically for this study; [B] Underlined regions represent the CS1 and CS2 tags for attaching Illumina barcodes.