Persistence of large conjugative plasmids relies on the combination of active partitioning and addiction mechanisms

Johannes Effe[1], Mario Santer[1], Yiqing Wang[1,2], Theresa E. Feenstra[1], Nils F. Hülter[1], Tal Dagan[1*]

[1] Institute of General Microbiology, Kiel University, Kiel, Germany
[2] Present address: Institut Pasteur, Université de Paris Cité, CNRS UMR3525, Microbial Evolutionary Genomics, Paris, France

*Corresponding author: Tal Dagan, email: tdagan@ifam.uni-kiel.de

Abstract

Extrachromosomal and mobile genetic elements, including transposable elements, plasmids, and accessory chromosomes, are found across all domains of life. The replication and segregation of elements integrated into the host chromosome are ensured by the host life cycle. In contrast, the persistence of autonomously replicating elements relies on their ability to remain within the host population. Here, we assess the evolutionary success of different persistence strategies of prokaryotic plasmids. Through intracellular competitions between plasmid genotypes, we find that combining active partitioning during cell division with a toxin-antitoxin (TA) system for post-segregational killing is more effective than either strategy alone. Mathematical modeling of long-term plasmid evolution, calibrated with empirical plasmid persistence probabilities, further supports these findings. A survey of enterobacterial genomes indicates that partitioning and TA systems are core features of large conjugative plasmids. Indeed, we confirm the presence of a previously unrecognized type I TA system in IncX3 plasmids, which serve as important vectors of antibiotic resistance in human pathogens. These findings suggest that large conjugative plasmids likely encode TA systems, including those yet to be identified. The combination of TA and partitioning systems emerges as the most effective strategy for the persistence of low-copy extrachromosomal elements with unstable segregation.

Introduction

Plasmids are extrachromosomal genetic elements that reside in prokaryote cells. The persistence of plasmids in bacterial populations following invasion depends on the plasmid ability to replicate and segregate during cell division[1,2]. Failure in plasmid replication into monomeric entities lowers the number of plasmid-copies available for segregation, which may lead to imbalanced segregation and plasmid loss[3,4]. Segregational plasmid loss leads to the emergence of plasmid-free cells, hence plasmid persistence furthermore depends on the plasmid effect on their host fitness[5–7]. Plasmids supplying the host with a beneficial trait, e.g., resistance to antibiotics, may confer their host a fitness advantage that promotes the plasmid persistence. However, in the absence of selection pressure for the plasmid-encoded trait, plasmids that are neutral, or slightly deleterious, to their host fitness are at risk of extinction due to competition for resources with plasmid-free cells[8,9].

Plasmid persistence can be promoted by active partition systems that facilitate plasmid segregation via their effect on plasmid localization within the host cell. The systems typically comprise three main components: partition binding site (centromere), a centromere binding protein (CBP) and a nucleotide triphosphatase (NTPase). The partition complex is formed when the CBP binds the plasmid centromere(s) and interacts with the NTPase, which drives plasmid transport and positioning[10]. The presence of partition systems is associated with the plasmid size. For example, in *Enterobacteria*, plasmids larger than 25Kb typically harbor a partition system, which facilitates their transition via the nucleoid area and thus promotes their segregation[11]. Plasmids of that size typically have a low copy number, e.g., <5 copies in *Escherichia* plasmids[12]. Plasmid partition systems likely share a common origin with chromosomal partition systems that have a similar function, but are evolutionary diverged[13].

Another route for plasmid persistence is plasmid-encoded toxin-antitoxin (TA) systems that inhibit the growth of plasmid-free cells following segregational loss. Toxin-antitoxin systems typically comprise a toxin that inhibit an essential cellular process and an antitoxin that neutralizes its cognate toxin. TA systems are found also in bacterial chromosomes (or phages), where they assume diverse functions including the prokaryotic immune system and survival in growth limiting conditions[14]. The toxin is commonly a protein that induces a toxic effect on the cell, e.g., via competitive binding or degradation of essential proteins. The antitoxin can be either an RNA or a protein. In type I systems, an RNA antitoxin neutralizes a proteinaceous toxin by binding the toxin mRNA, thus inhibiting the toxin translation. Example is the plasmid-encoded *hok*/*sok* TA post-segregational killing[15]. Plasmids may harbor multiple TA systems and recent studies indicate a role of transposable elements in the evolution of plasmid TA repertoire[16,17].

Toxin-antitoxin and active partition systems increase plasmid stability via two different strategies[18]. Here we ask which (if any) of these plasmid traits is superior over the other. For that purpose, we conducted intracellular plasmid competitions between initially unstable non-mobile plasmids harboring either system or both. Both systems in our experimental setup were adopted from plasmid R1 that was originally isolated from a clinical *Salmonella* isolate[19]. The *parA* locus in R1 was adopted as the partition system; it includes the CBP-encoding *parR*, the actin-like ATPase *parM* and the ParR binding site *parC*[20]. The parB locus of R1, which includes the *hok*/*sok* toxin-antitoxin system was adopted as the TA system[21,22]. The Hok (Host killing) protein is a toxin that binds the cell membrane, thereby inhibiting the cell respiration. The antitoxin sok (suppression of killing) is an sRNA that inhibits *hok* translation by binding to its mRNA[23]. Using the empirical results from plasmid competition we developed a mathematical model for the effect of partition and TA systems on plasmid stability in evolutionary timescales.

Finally, we validate our prediction by observing the distribution of partition and TA systems in contemporary plasmids.

Results

Intracellular head-to-head competitions enable the comparison of plasmid fitness

To compare the effect of TA and Par systems on plasmid stability we established an experimental system for intracellular plasmid competitions. The recipient population is first transformed with the two competing plasmids simultaneously. The presence of the competing plasmids is confirmed by imposing selection for antibiotic resistance markers in the competing plasmids. The recipient population is then incubated for the competition duration with five replicates in the absence of antibiotics. At the end of the competition, the relative frequency of cell types in the replicate populations is determined (Fig. 1A). The plasmid that is found in a higher frequency of homogenic hosts is considered to have a higher fitness compared to the competing plasmid. This system enables us to rank plasmid traits pair-wise according to their relative fitness.

As the low and high reference for plasmid stability we utilized two pCON plasmid genotypes that are characterized by either low stability (unstable, U) or high stability (stable, S) (Fig. 1B). The backbone of both model plasmids originated from the pBBR1 broad-host range plasmid[24]. The pCON instability arises from conflicts between gene transcription and plasmid replication, which promotes the formation of plasmid multimers and result in segregational loss. The stable pCON genotype has a slightly greater distance between the origin of replication and the antibiotic resistance gene promoter, leading to a higher frequency of monomeric plasmids and high plasmid stability[25]. Competing plasmids harboring a toxin-antitoxin system (T), partition system (P), or both systems (PT) were constructed using the unstable plasmid backbone. To validate the effect of the plasmid stability system, all model plasmids were constructed with alternative selection markers: either *nptII* that confers resistance to kanamycin, or *cat* that confers resistance to chloramphenicol (Fig. 1B), which we denote by pCONn and pCONc, respectively, followed by the stability trait, e.g., pCONn-U (Fig. 1B)[26]. The model plasmids are isogenic with the exception of the antibiotic resistance markers and in/stability module. Most model plasmids have a negligible effect on their host fitness with the exception of pCONn-T and pCONc-PT that have a slightly deleterious effect on their host fitness (fitness effect ranging between 0 and -5.5%; Supplementary Fig. S1). The model plasmid copy number ranged between 3 and 8 plasmids (relative to the chromosome; Supplementary Table S1).

The stability of all model plasmids was evaluated from plasmid loss frequencies in a serial transfer experiment over five days. The unstable plasmid loss frequency was ca. 7-8% per transfer, on average, with little variation between pCONn-U and pCONc-U. The stable plasmids had an average loss frequency <1% for both plasmids (pCONn-S and pCONc-S) (Fig. 1C). The stability of all model plasmids equipped with a stability system was increased compared to the unstable model plasmids and similar to that of the stable model plasmid (Fig. 1C). The model plasmid stability was similar between plasmids harboring the same stability system (or none) that differ in their antibiotic resistance marker.

We first demonstrate our approach for pairwise plasmid fitness ranking with the results of plasmid competitions between the unstable and stable plasmid variants. The competition between pCONn-U and pCONc-S show that most replicate populations were dominated by homogenic hosts of the stable plasmid, with a smaller fraction of homogenic pCONn-U hosts ($P=7.75\times10^{-7}$, using paired sign test, n=25). The proportion of plasmid-free cells (non-hosts) ranged between 0 and 29%. In several replicates we observed heterogenic plasmid hosts,

where both competing plasmids co-exist. Plasmid co-existence in our experimental system can be the results of either incomplete segregation of the competing plasmids, or their fusion into heteromultimers[26]. To validate the effect of plasmid stability on the competition result, we also performed a reciprocal competition, pCONc-U and pCONn-S, using model plasmids where the antibiotic resistance markers are exchanged between the two plasmid backbones (Fig. 1B). The reciprocal competition yielded a majority of pCONn-S in most replicates, albeit the overall difference in the frequency of the competing plasmids was smaller (P=0.009, using paired sign test, n=19; Fig. 1D). Taken together, the competitions among the stable and unstable plasmids demonstrate a higher fitness of the stable plasmid, regardless of the antibiotic resistance selection marker in the plasmid backbone.

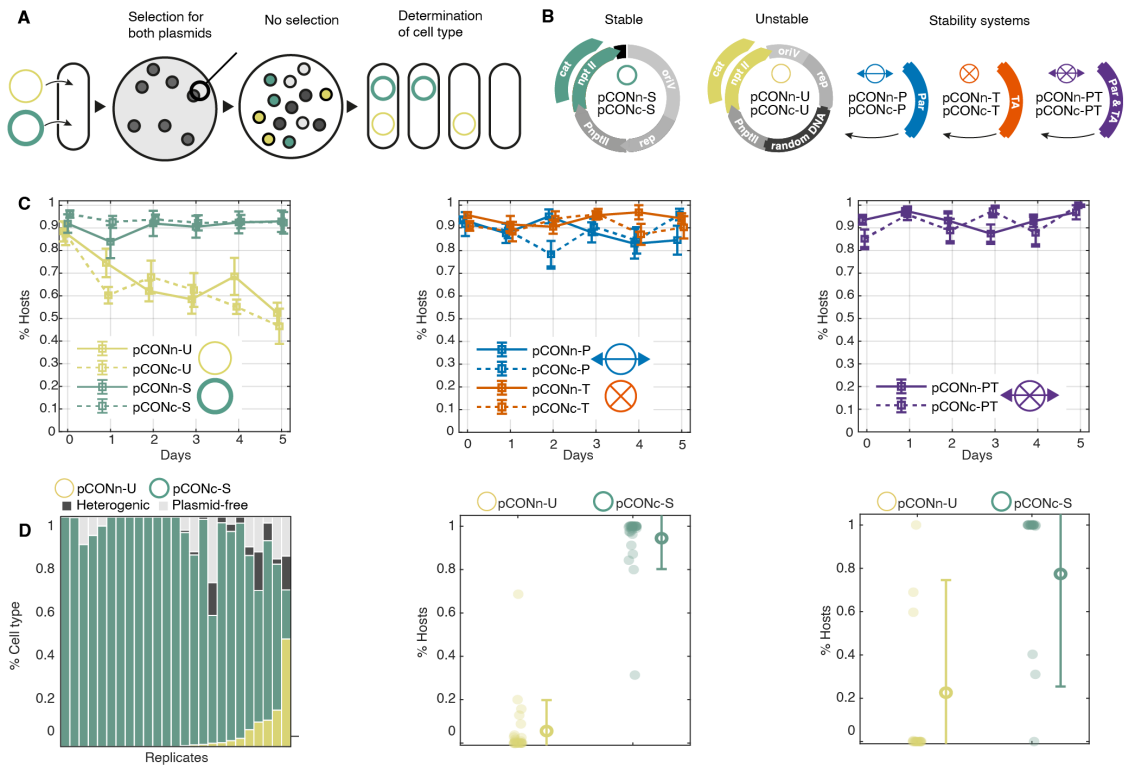

Figure 1. An experimental system for pairwise *in cellulo* competitions between plasmid genotypes. A) The plasmids are introduced simultaneously into a naïve host population that is cultured under selection for both plasmid genotypes. Replicate colonies are picked and cultivated in non-selective conditions. The result of the plasmid competition is evaluated according to the frequency of host types at the end of the experiment. These include plasmid-specific hosts, heterogenic hosts of both plasmids and plasmid-free cells. B) The set of competitor model plasmids. The plasmid backbone consists of the pBBR1 oriV and *rep* along with the P_{nptII} promoter followed by an *nptII* or *cat* antibiotic resistance gene. In the pCON-S variants, non-coding DNA between the oriV and the antibiotic resistance gene. The remaining plasmids vary in the locus between the *rep* gene and the pnptII promoter. In the pCON-U variants a segment of random DNA is inserted; in the pCON-P plasmid the parA locus of plasmid R1 is inserted, which contains the genes *parM*, *parR* and the parC DNA binding site; the plasmids pCONn-T and pCONc-T both contain the parB locus of plasmid R1, which consists of the *hok*/*sok* toxin antitoxin system; plasmids of the PT type contain both the parA and the parB locus. Plasmid symbols are shown next to the plasmid types. The model plasmid size ranges between 2,733 bp (pCONc-S) and 5,273 bp

(pCONn-PT) (see Supplementary Table S1 for details). C) Model plasmid stability as evaluated in a serial transfer experiment under non-selective conditions (n=8 replicates except for pCONc-R with n=5 and pCONc-PT with n=7. D) (left) A bar graph showing the results of pCONc-U vs pCONn-S competition. Each bar corresponds to a replicate population with shaded area in the bar proportional to the proportion of cell types in the population. (middle) The results of pCONc-U vs pCONn-S presented as a scatter plot. Host frequencies are shown with dots. The mean frequency is shown with a circle, with error bars corresponding to standard deviation. (right) Results of the reciprocal competition between pCONn-U vs pCONc-S are presented in the form of a scatter plot.

Pairwise competitions reveal superiority of plasmids harboring both partition and TA systems

To evaluate the effect of an active partition system on plasmid fitness, the plasmid pCONc-P was competed against the unstable pCONn-U and the stable pCONn-S. The competition against the unstable plasmid revealed that the proportion of homogenic pCONc-P hosts was nearly 100%, with only rare homogenic pCONn-U hosts across all replicates. In contrary, the competition against the stable plasmid yielded mostly homogenic pCONn-S hosts and only a minor proportion of homogenic pCONc-P hosts (Fig. 2A.1-2). The reciprocal competition between plasmids, where the stability systems are exchanged between the plasmid backbones, reveal no significant difference in the proportion of pCONn-P homogenic hosts in the competition against pCONc-U (Fig. 2B; Supplementary Fig. S3). The reciprocal competition of pCONn-P against the stable plasmid (pCONc-S) yielded a majority of pCONc-S homogenic hosts (Supplementary Fig. S3), thus confirming the advantage of the stable plasmid over the plasmid comprising a partition system.

The competition of plasmids harboring a TA system against the unstable plasmids yielded a significantly higher frequency of pCONc-T homogenic hosts compared to pCONn-U hosts (Fig. 2A.3). The fitness advantage of the TA harboring plasmid compared to the unstable plasmid was further confirmed in the reciprocal competition (Supplementary Fig. S3). These results are in agreement with previous results on the ability of plasmids harboring TA to outcompete a TA-lacking, otherwise isogenic plasmid[27]. In contrast, the competition between plasmid pCONc-T and the stable plasmid pCONn-S yielded similar proportions of homogenic hosts of the two plasmids, with pCONn-S having a slightly higher mean proportion of homogenic hosts (Fig. 2A.4). The reciprocal competition between pCONn-T and pCONc-S also reveals a higher proportion of the stable plasmid hosts compared to the TA plasmid hosts (Supplementary Fig. S3). The competitions against pCONn/c-T often resulted in heteromultimers of the competing plasmids that were observed with high frequency and were stable over time (Supplementary Fig. S4). Further comparison of plasmid fitness between the two stability systems showed that the plasmid carrying the toxin-antitoxin (pCONn-T) had a significant advantage compared to the plasmid carrying the partition system (pCONc-P) (Fig. 2A.5), which was confirmed in the reciprocal competition (Fig. S3). Taken together, the TA system increased the fitness of the unstable plasmid similarly to the active partition system but is superior to the partition system in a direct competition.

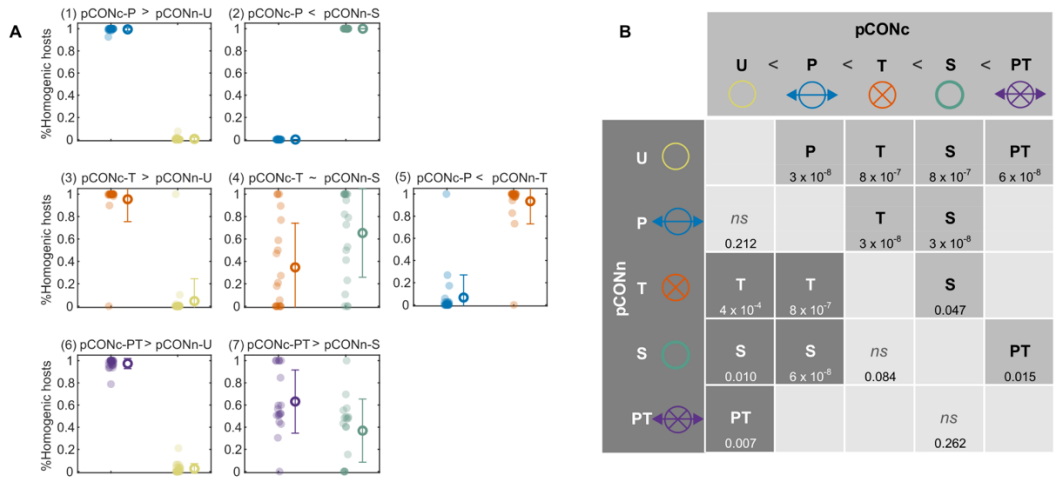

Figure 2. Plasmid fitness ranking from pairwise competitions. A) Results of pairwise competitions between different pCONc and pCONn plasmid variants. Each dot corresponds to the proportion of homogenic hosts carrying a given plasmid variant in an individual head-to-head competition, while the mean frequency is indicated by the circle. Error bars show the standard deviation. Proportion of cell types, including heterogenic and plasmid-free cells, are presented in Supplementary Fig. S2. B) Ranking of plasmid stability determinants, denoted by inequality symbols, derived from pairwise competitions of pCONn variants (rows) against pCONc variants (columns). The table cells indicate the superior competing variant having significantly higher proportion of homogenic hosts at the end of the experiment (P-values below using two-sided paired sign test, n=25 in all pairs of variants, except pCONn-S vs. pCONc-PT where n=20).

The competition between plasmids that harbor a combination of both systems–the active partition system and the TA system–, and an unstable competitor plasmid reveals a clear advantage of plasmids pCONc-PT and pCONn-PT over the unstable plasmid variants (Fig. 2.6; S3). Competing the stabilized PT plasmids against the stable plasmid variant revealed a higher proportion of the pCONc-PT hosts compared to pCONn-S (Fig. 2.7). In the reciprocal competition, the mean proportion of homogenic hosts of both plasmids was not significantly different (Fig. S3). Hence, compared to the other stabilized plasmids, the plasmid harboring both stability systems had the highest success compared to the stable plasmid. The competitions allow for an unambiguous ranking of the plasmid stability determinants in the short-term evolution experiments. Mathematically speaking, the homogeneous relation, >, that indicates which stability determinant is more advantageous than another is transitive (e.g., S > T and T > U, then S > U holds). The set of pairwise competition experiments show that all stability traits PT, S, P, T were advantageous over the unstable model plasmid in one or both of the competitions. The P trait was inferior to both S and T traits, while the S trait had an advantage over the T trait. Finally, plasmids encoding the PT stability trait had a significant fitness advantage when competed against a stable plasmid (S) (Fig. 2b).

A mathematical model of plasmid segregation enables a comparison among pCON stability determinants in evolutionary timescales

To predict the plasmid competitions on long time scales compared to the experimental system, we performed simulations with a model of random plasmid segregation[28], which we extended to describe the influence of toxin-antitoxin and partition systems (Fig. 3a). In models of random plasmid assortment, the probability of plasmid segregation at cell division is 2^{1-n_r}, which

arises from the distribution of plasmid copies into daughter cells and relies on the number of plasmid copies, n_r. Plasmid fusion or clustering of sister plasmids may decrease the number of plasmid units available for segregation, such that the probability of plasmid-free daughter cells increases. The model presented here accommodates the model plasmid properties by including the following parameters: the plasmid ability to replicate into monomeric units is parametrized by $1 - p_{\mathrm{def}}$, where p_{def} is the probability for plasmid copies to be removed from the plasmid pool available for segregation after plasmid replication. For example, if $p_{\mathrm{def}} = 0$, all $2n_c$ copies would be available for segregation. Conversely, if $p_{\mathrm{def}} = 1$, newly replicated plasmids would be entirely removed from the pool of independent plasmid units. The probability at which pairs of plasmid copies are actively assorted to different daughter cells in a non-random fashion (balanced assortment) is parametrized by p_{par}. Finally, the probability at which plasmid-free cells are killed post segregation if they divide from hosts of a plasmid harboring a TA system is parametrized by p_{eff}. Our model describes the population dynamics of cells with (up to) two model plasmids. While the model describes the full stochasticity of plasmid assortment to daughter cells, the host population growth is assumed to be deterministic with cell divisions in each generation in the absence of population bottlenecks or cell death (see details in the Methods and in the supplemental text).

The virtual plasmids are defined by parametrizing p_{def}, p_{par}, and p_{eff}, which were estimated using the model plasmid persistence data (Fig. 3C; Fig. 1C). The model does not include the effect of resistance markers (i.e., pCON$^{n}/_{c}$ plasmid backbones have the same characteristics). Hence, we denote the (virtual) plasmid variants only by their stability determinant. The plasmid copy number was set to $n_{\mathrm{c}} = 4 \approx \frac{2}{3}\bar{n}_{\mathrm{PCN}}$, which is based on the mean model plasmid copy number, $\bar{n}_{\mathrm{PCN}} = 5.6 \pm 2.1 \triangleq \frac{3}{2}n_{\mathrm{c}}$ (Supplementary Table S1; the factor $\frac{3}{2}$ arises because the PCN is measured across all stages of the cell cycle). The virtual model plasmid S is solely defined by the copy number parameter, $n_c = 4$ (Fig. 3B). For the other variants, U, P, T, and PT, we estimated the monomerization deficiency to be $p_{\mathrm{def}} = 0.332 \pm 0.053$ ($\mathrm{CI}_{95\%}$) from the pCONn-U/pCONc-U persistence data (Fig. 3C, see Methods). For the P and PT plasmid variants we estimated the partitioning efficiency to be $p_{\mathrm{par}} = 0.435 \pm 0.105$ from the pCONn-P/pCONc-P persistence data. Finally, we estimated the efficiency of post-segregational killing by the T and PT variants to be $p_{\mathrm{eff}} = 0.884 \pm 0.063$ from the pCONn-T/pCONc-T persistence data (see details in the supplemental text). The estimated plasmid parameters were used for the simulation of long-term plasmid persistence (Fig. 3B) and pair-wise plasmid competitions (Fig. 3D).

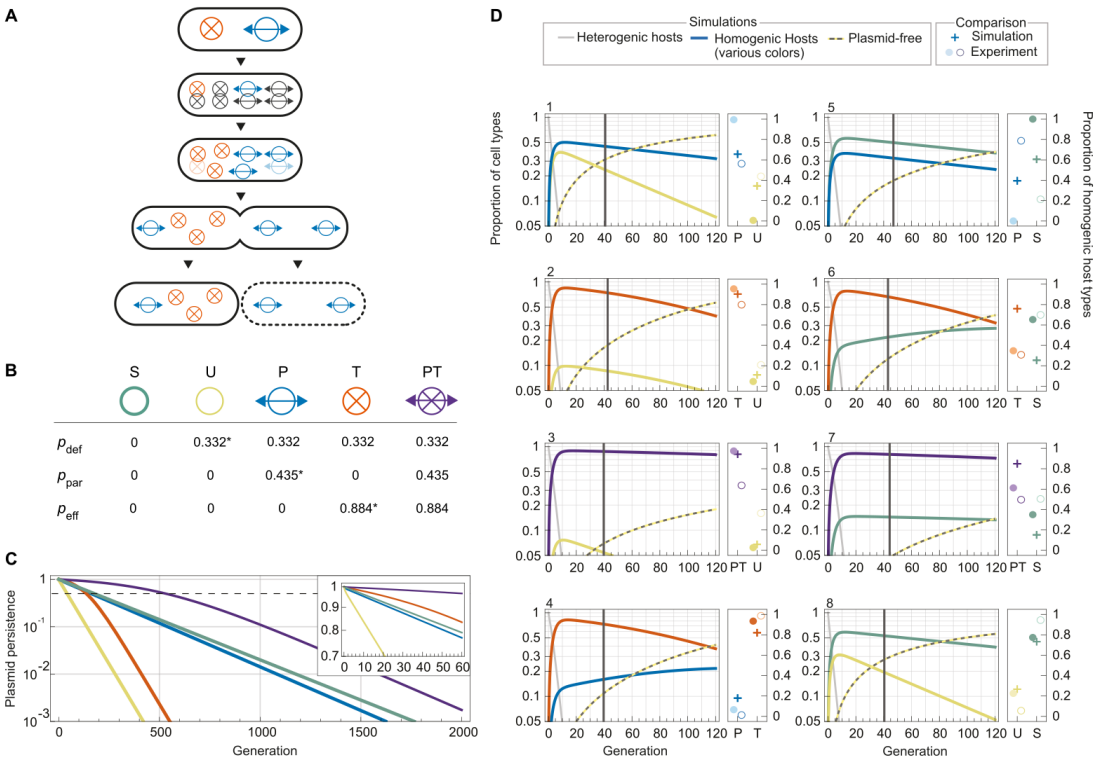

Figure 3. Model of plasmid segregation and simulations of plasmid persistence and head-to-head competitions. A) Cell life cycle in the plasmid segregation model for the example of a plasmid-heterogenic cell carrying two copies of different plasmid variants encoding a partition and toxin-antitoxin system at cell birth. The life cycle describes 1. random plasmid replication, 2. monomerization (depletion of plasmid replicates due to incomplete multimer resolution) and 3. random or balanced assortment followed by cell division into two daughter cells and 4. post segregational killing. B) Each virtual plasmid variant, S, U, P, T, and PT, is characterized by a parameter triplet $(p_{\mathrm{def}}, p_{\mathrm{par}}, p_{\mathrm{eff}})$. The parameters p_{def}, p_{par}, p_{eff} were estimated based on experimental persistence data of the respective pCON$^{n}/_{c}$ variants (indicated by an asterisk, see Methods). C) Plasmid persistence in plasmid-homogenic cells for various plasmid variants in the short term (inset) and long term. The dashed line indicates plasmid loss of 50% (plasmid half-life). D) Prediction of the cell genotype dynamics (left) in head-to-head competitions (various plots) and comparison between simulated and experimental genotype proportions when experiments ended. The black vertical lines show generations passed until the end of the plasmid competition experiments.

TA systems temporarily compensate for plasmid instability while partition systems promote long term plasmid stability

To compare the effect of plasmid stability determinants on plasmid loss dynamics in the short term and long term we first simulated the persistence of all plasmid variants. For the stable plasmid variant (S), the segregation rate, $r^{(\mathrm{S})} = 2^{1-2n_c} = \frac{1}{128}$, calculated from classical random segregation models, results in an exponential decay of the plasmid host proportion, $h^{(S)}(g) = \left(1-\frac{r}{2}\right)^{g} = \left(1-\frac{1}{256}\right)^{g}$, over generations, g (Fig. 3C). The unstable plasmid variant (U) with

an estimated monomerization deficiency of $p_{\mathrm{def}} = 0.332$ has a ca. 4-fold higher segregation rate $r^{(\mathrm{S})} = \frac{1}{30.7}$. The stable plasmid variant (S) has a higher half-life (177 generations until the host cell proportion is 50%) compared to the unstable variant (U) (42 generations), as expected. Our estimation of the partitioning efficiency of partition system encoding plasmids (P) results in an increased half-life of 118 generations. The persistence dynamics of the TA-encoding plasmid variant (T) are similar to the P-variant for the first 100 generations. When plasmid-free cells start dominating the population the T plasmid persistence changes, however, and the host cell frequency decays exponentially at the same rate as the unstable plasmid (U). The change in the plasmid persistence trend in the model stems from two consecutive processes. Initially, plasmid-free cells arise from host cells due to random segregation, however, those are mostly killed post segregation by the toxin. The emergence of plasmid-free cells in the later phase is primarily from surviving plasmid-free cells (lacking the toxin). Hence, the model predicts that the TA system prolongs the plasmid persistence for approximately 100 generations compared to the unstable plasmid. Note the TA system does not alter the rate of segregational plasmid loss; its effect on plasmid persistence is best understood by the resulting dynamics of host and plasmid-free cells in the population. The PT plasmid has the highest persistence with a half-life of 531 generations that is more than 3-fold compared to the P plasmid. Although the host cell proportion of the PT-variant is still substantially higher after hundreds of generations, the stable plasmid variant (S) has a higher persistence in the very long term compared to the PT-plasmid. Taken together, the stabilization systems provided to the unstable plasmid – whether acting individually or synergistically (as in the PT variant) – cannot fully compensate for the monomerization deficiency but significantly enhance the plasmid's persistence on small- and intermediate-time scales.

Predicted long-term evolutionary reflect major trends of head-to-head competitions

To predict the outcome of intracellular pairwise competitions between plasmids with different stability determinants, we simulated the plasmid segregation dynamics in initially heterogenic hosts, mirroring the experimental head-to-head competitions. The simulations captured the simultaneous processes of plasmid segregation and loss (Fig. 3D, left panels). In all simulated competitions, heterogenic plasmid hosts rapidly declined to proportions below 10% (after 5–10 generations), while homogenic hosts temporarily dominate the population. Additionally, the simulations predicted a rise in plasmid-free cells to frequencies between 5% and 50 % by the end of the experiments. The simulated proportions of homogenic hosts were similar to the empirical frequencies at the corresponding generation times with only few exceptions. In head-to-head competitions where the reciprocal experiments yielded the same outcome (excluding PT vs. S), the simulations generally reproduced the empirical result with the exception of the competition T vs. S (Fig. 3D, plot 1-4, 6, and 8). In the T vs. S competition, the simulated proportions of the homogenic hosts converge during the experimental generation times and intersect later at generation 132; thus, the S-plasmid hosts, which were found to be dominant in the experiment, are predicted to also dominate the population but only at later generation times (plot 6). Similar dynamics were observed for the competitions P vs. T (as well as for PT vs. S) with the P-variant (or S-variant) eventually dominating the population. These patterns are similar to the persistence dynamics of the respective variants in homogenic hosts (Fig. 3D plots 4,7, cf. Fig. 1C). In the competition PT vs. S, plasmid-host cells homogenic for the PT variant dominated for a very long period but were eventually outcompeted by the S-variant at generation 5,210. Taken together, the combination of partition and TA system is predicted to confer the plasmid long-term stability.

289 **Partition and TA systems are core functions of large plasmid types**

290 To examine the outcome of long-term plasmid evolution on the repertoire of partition and toxin-
291 antitoxin (TA) systems, we surveyed for their presence in 6,784 plasmids from 2,441 isolates
292 of *Klebsiella*, *Escherichia* and *Salmonella* (*KES*)[29]. Partition systems were identified in 67%
293 (4,579) of the plasmids, all of which were classified into plasmid taxonomic units (PTUs) of
294 large plasmids (>19Kb). Most of the PTUs had a median of one partition system per plasmid
295 (Fig. 4a). Multiple partition systems were observed, e.g., in a PTU-HI2 plasmid that harbors
296 two functional replicons RepHI2 and RepHI2A encoding a major partition system (ParABS)
297 and a minor system (ParMRC)[30], The presence of both partition systems has been shown to
298 confer highly stable segregation in a study of the related plasmid R27[31]. Plasmids in PTU-E15,
299 -FK and -FE that can have up to five partition systems (Fig. 4A).

300 The survey for TA systems revealed 44 TA families in 66% (4,456) of the plasmids,
301 most of which are in PTUs of large plasmids. TA systems were rarely found in PTUs of small
302 plasmids (except for PTU-E4 and E1). Plasmids without TA systems are exceptional among
303 PTUs of large plasmids, examples are PTU-X3, E15, E16, and N1 (Fig. 4A). The maximum
304 number of TA systems observed in a single plasmid was 11, e.g., in *Klebsiella pneumoniae*
305 strains sampled in epidemiological surveys for antibiotic resistant strains[32,33].

306 To investigate whether the apparent absence of TA systems in some large plasmids is
307 due to limitations of our survey approach, we examined plasmids in PTU-X3 where no TA
308 system could be identified based on currently known TA systems[34]. Plasmids in PTU-X3
309 (IncX3) are typically large and conjugative; many harbor antibiotic resistance *gene* $bla_{\text{NDM-5}}$
310 and were reported in human infections and domesticated animals[35,36]. The presence of type II
311 TA was reported in <10% of the IncX3 plasmids[37]. A detailed examination of distant homologs
312 to known TA systems in X3 plasmids revealed a gene coding for DinQ-like toxin (termed *dqlB*),
313 bearing high structural similarity to a chromosomal type I TA system *dinQ-agrB* in *E. coli*[38]. An
314 *agrB* antitoxin RNA was further identified next to plasmid *dinQ*-like open reading frame. The
315 genetic neighborhood of the *dqlB-agrB* includes transposable elements and is conserved
316 across 38 (out of 116) X3 plasmids (Fig. 4B; Supplementary Table. S5).
317

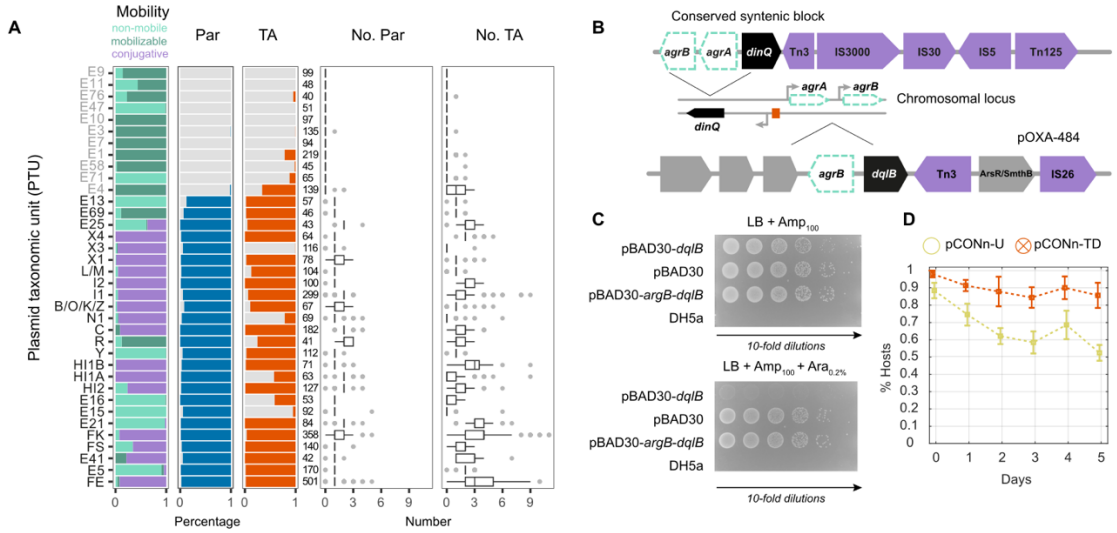

Figure 4. Large plasmids likely harbor both partition and TA systems. (A) Presence of partition (Par) and TA systems in *KES* plasmid taxonomic units. PTUs of small plasmids (<19Kb) are written in gray; PTUs of large plasmids (≥19Kb) are written in black. Horizontal stacked bar graphs show the fraction of plasmid mobility in the PTU and the systems presence in plasmids and the distribution of system instances per plasmid. (B) Structure of a conserved syntenic block including *dinQ*-like gene in X3 plasmids. The chromosomal *dinQ* in *E. coli* is presented together with the *agrB-dqlB* locus in plasmid pOXA-484 in *E. coli* strain EC-JS316. (C) Toxicity assay of the *agrB-dqlB* locus from pOXA-484. Growth of *E. coli* DH5α cells carrying either no plasmid, the empty vector control plasmid (pBAD30), the *argB-dqlB* locus from plasmid pOXA-484 in *E. coli* EC-JS316 (pBAD30- *agrB-dqlB*), or the *dqlB* carrying plasmid (pBAD30-*dqlB*). The samples were plated on medium selective for ampicillin resistance (LB+Amp_{100}) and selective plates supplemented with Arabinose (LB+Amp_{100}+$Ara_{0,2\%}$). The expression of *dqlB* locus is induced in the presence of Arabinose via P_{BAD}. The presented results are representative of three independent replicates. (D) Stability of plasmid pCONn-TD comprising the *agrB-dqlB* locus from plasmid pOXA-484 as evaluated in a serial transfer experiment under non-selective conditions (n=6).

To validate a TA function of the *argB-dqlB*, we identified a relevant strain, *E. coli* EC-JS316, that harbors an IncX3 plasmid with the locus. This strain belongs to *E. coli* clone ST10 and carries plasmid pOXA-484 (IncX3) that confers the host resistance to multiple antibiotics[39]. The *dqlB* locus from that plasmid was cloned into an expression vector under the regulation of the P_{BAD} promoter with or without the putative antitoxin *agrB*. Induction of the *dinQ*-like toxin expression led to growth inhibition of the host, while induction of the complete *agrB-dqlB* expression showed no signs of growth inhibition (Fig. 4C). Our results thus confirm a TA function of the *agrB-dqlB* locus in pOXA-484. To further validate the plasmid stability function of the *agrB-dqlB* TA system, we cloned the native pOXA-484 *agrB-dqlB* locus into the pCONn backbone (termed pCONn-TD). Quantifying the persistence of pCONn-TD showed an increase in plasmid stability compared to pCONn-U (Fig. 4D). We conclude that the DinQ-like TA system has a role in the persistence of IncX3 pOXA-484-like antibiotic resistance plasmids, in agreement with recent studies reporting similar results[40,41]. We conclude that large conjugative plasmids likely harbor TA systems, even if those cannot be identified by sequence similarity to the so far recognized systems.

Discussion

Extrachromosomal and mobile genetic elements are frequent across all domains of life. The replication and segregation of genetic elements that integrate into the host chromosome is secured by the basic cellular process; examples are transposons in bacteria (e.g., Tn7[42]) and insertion elements in Archaea (e.g., ISC1217[43]), starships in fungi[44], P-elements in plants[45] and SINEs in human[46]. In contrast, the persistence of autonomously replicating elements largely depends on their ability to persist within the host lineage. Studies of extrachromosomal genetic elements that evolved into an integral component of the host genome suggest that such elements harbor mechanisms for the synchronization of their replication with the host chromosome and their segregation with cell division. Examples are the secondary chromosome in the bacterium *Vibrio cholera*[47]), accessory chromosomes in fungi[48], B-chromosomes in plants and animals[49] as well as the eukaryotic organelles[50]. Nonetheless, the specific molecular mechanisms at the origins of such elements play often remain poorly understood. Here we provide empirical insights and theoretical predictions into the evolutionary success of extrachromosomal elements depending on their persistence strategy with the example of prokaryotic plasmids.

Our study shows that mechanisms that increase plasmid stability are likely to increase also the evolutionary success (or fitness) of the plasmid. When the deviation between plasmid stability determinants is large, e.g., as the difference between of unstable and the stabilized pCON variants, the results of the competition where in favor of the plasmid having a higher stability. Predicting the evolutionary trajectory in competitions between plasmids having a similar stability is, however, more challenging. Our approach of intracellular competitions supplies a methodological solution for the comparison of plasmid fitness also between plasmids having a similar stability. The performed competitions reveal that the stabilized plasmids outcompeted the unstable plasmid, hence both partition and TA systems have the potential to increase the fitness of plasmids that are impaired in their replication and multimer resolution (such as the pCON backbone[25]). At the same time, the stabilized plasmids were inferior in the competition against the stable pCON$^{n}/_{c}$ -S, hence, active partition and TA systems do not increase the fitness of a preliminarily unstable plasmid compared to a plasmid that replicates and forms multimers effectively. The combination of both partition and TA systems was revealed to be superior in the short-term experiments, even above the highly stable plasmid. This result may be explained by the contribution of a partition systems to reduce the frequency of segregants (plasmids-free hosts)[18], which is beneficial if the toxin effectivity is low (e.g., due to the toxin instability[51]).

The mathematical model reveals systematic differences in the influence of partition and toxin-antitoxin systems on plasmids' long-term persistence. Partition systems slow down the plasmid decay in the population by reducing the segregation rate, while toxin-antitoxin systems only delay plasmid loss via their negative effect of post-segregational killing on newly generated plasmid-free cells. Since the TA system *hok*/*sok* used in our experiments is a highly efficient toxin-antitoxin system, it allowed only a few plasmid-free lineages to arise in the initial population of plasmid hosts and, thus, effectively suppressed plasmid loss. Toxin-antitoxin systems can be most efficient in stabilizing the plasmid compared to the active partition system when plasmid hosts make up (almost) the entire population. Our findings, thus, suggest that TA systems play a role in plasmid stability after periods of strong positive selection for plasmid hosts. When the proportion of plasmid host cells becomes small, plasmid-free cells are mainly the offspring of plasmid-free cells. In such a scenario, the frequency of a plasmid harboring a TA system decays at nearly the same rate as the unstable plasmid because post-segregational

382 killing only affects a small proportion of the plasmid-free subpopulation. For plasmids encoding
383 both TA and partition system (PT), the partition system additionally stabilizes the plasmid with
384 the result that low plasmid-host proportions are reached later, thus the TA system remain
385 efficient for additional generations, which results in the very successful strategy of the PT-
386 variant and is in line with previous experimental result[18]. Taken together, our mathematical
387 model provides a new framework to evaluate the interplay of partition systems, toxin-antitoxin
388 systems, and plasmid copy number on the plasmid persistence.
389 Since the model plasmids share an identical backbone (Fig. 1B), our results offer
390 insights into the allele dynamics of stability determinants and their outcome in the long range
391 for the evolution of plasmid gene content. The stability determinants can be considered as
392 plasmid alleles that segregate under non-selective conditions after a balancing selection event
393 (i.e., starting from similar frequencies of both alleles). Heteroplasmy of plasmid alleles can be
394 maintained under rapidly alternating selection for both alleles, as has been demonstrated for
395 antibiotic resistance alleles[52]; the maintenance of heteroplasmy requires, however, that the
396 selection for both alleles is strong[53]. The experimental and simulated plasmid competitions
397 presented here show that random segregation of plasmid variants with (on average) six copies
398 under non-selective conditions is rapid, where the competition outcome is determined by the
399 dynamics of homogenic plasmid-hosts (Fig. 3D). A higher plasmid copy number leads to
400 extended states of plasmid heteroplasmy (termed also *heterozygocity window*[54]), as
401 demonstrated for allele dynamics of a plasmid having ca. 15 copies per cell[55]. Random
402 segregation of plasmid alleles during cell division (termed *segregational drift*[56]) may lead to a
403 rapid loss of plasmid alleles even if they confer a fitness advantage to the host[28]. The mode
404 of plasmid replication and segregation is furthermore expected to affect the heterozygocity
405 window length[54]. However, empirical data on the mode of plasmid replication and segregation
406 remains lacking, with only few exceptions[28], hence the presence of mechanisms for
407 maintenance of plasmid heteroplasmy (if any) remains unknown. A recent study that
408 investigated allele dynamics of the eukaryotic organelles in clonal plants by calibrating a
409 mathematical model similar to the one presented here with deeply sequenced samples
410 concluded that heteroplasmy of the mitochondrion and plastid in the plant tissue is rather
411 rare[57]. Taken together, we conclude that plasmid heteroplasmy is rather rare and the fate of
412 plasmid variants is most likely determined by the dynamics of plasmid persistence in
413 homogenic hosts.
414 Our experimental and theoretical results show that the combination of post-
415 segregational killing mechanism (TA) and active partitioning systems is superior over all other
416 plasmid persistence strategies (under conditions that are non-selective for the plasmid
417 presence). The prevalence of such combinations in large, mostly conjugative, plasmids,
418 further supports our observations (Fig. 4A). Indeed, our model plasmids are rather small (2.7-
419 5.3 Kb; Fig. 1B), however, they share a similar copy number of ca. 5 copies per cell with large
420 *Escherichia* plasmids[12]. Hence, we consider the plasmid copy number a key determinant of
421 our results and predictions. At the same time, our simulation of long-term plasmid persistence
422 predicts that the most successful stability determinants do not ensure infinite persistence and
423 highlight important systematic differences in the effect of TA and partition systems on plasmid
424 evolutionary success. Indeed, plasmids in many of the large plasmid PTUs may harbor and
425 disseminate antibiotic resistance genes[29,58], hence their long-term persistence is likely
426 promoted in addition by periodic selection events (and to dispersal across host populations)
427 (e.g., as shown for pOXA-48[59]). We conclude that the combination of TA and partition systems
428 being the most effective strategy for the persistence of low-copy extrachromosomal elements
429 whose segregation with the host is unstable.

Materials and Methods

Bacterial strains and culture conditions

Bacterial cultures were grown routinely at 37 °C while shaking or on Lysogeny Broth (LB) agar plates unless stated otherwise. For selective conditions, the following concentrations were used: kanamycin 10 µg/ml; chloramphenicol 10 µg/ml; trimethoprim 125 µg/ml; tetracycline 10 µg/ml. The *E. coli* strain MG1655 *recA*::*tetA* derived from the *E. coli* strain MG1655 (DSM No. 18039, German Collection of Microorganisms and Cell Cultures, DSMZ) was used as plasmid host in all competition experiments. The construction of the model plasmids was performed with *E. coli* DH5α[60] as cloning host. Additionally, the toxicity assay was also conducted in DH5α. All primers used in this study are listed in supplementary table S1.

Construction of strains and plasmids

To generate the *E. coli* MG1655 *recA*::*tetA* strain the parental strain MG1655 was first transformed with the helper plasmid pTKRED[61], harboring genes encoding the λ-Red enzymes for recombineering. The *tetA* gene along with its promoter P_{lacIQ1} was amplified using pTKLP-*tetA* as template. For the amplification of *tetA* with its promoter the primers NH300 and NH301 were used along with Phusion polymerase. After acquiring the MG1655 strain containing pTKRED a liquid culture with IPTG (2 mM) was inoculated. Electrocompetent cells were generated after an overnight incubation at 30 °C. The electrocompetent cells were then transformed with the purified PCR product. The transformed cells were incubated at 30 °C and plated onto LB agar plates with tetracycline. The transformants were cured of pTKRED by incubation at elevated temperature (37 °C).

A trimethoprim resistant strain used for the assessment of plasmid fitness effects was generated through triparental mating of the host strain, *E. coli* MG1655 *recA*::*tetA*, with the two *E. coli* strains MFDpir pGP-Tn7-Tp[62] and MFDpir pTNS2[63]. The three strains were incubated overnight in liquid culture at their respective optimal temperature. Following this, the three strains were washed twice by centrifuging for 5 min at 3000 rpm and resuspension in Phosphate Buffered Saline (PBS). Afterwards, the cells were resuspended in 500 µl of PBS and mixed in a 1:1:1 ratio based on their OD_{600}. The strain mix was spotted onto a non-selective agar plate with diaminopimelic acid (DAP) and incubated at 30 °C overnight. After the mating, the conjugation spot was resuspended, the cells plated onto selective agar plates without DAP and incubated at 40 °C overnight. This step hinders the MFDpir strain growth due to their requirement of DAP, as well as the replication of plasmids pGP-Tn7-TP and pTNS2 due to their temperature sensitivity, leaving only plasmid-free recipient strain alive on the plate.

For the construction of the model plasmids the plasmids pCON (termed here pCONn) (GenBank accession number MK697350)[25] and pCON2 (termed here pCONc)[26] were used as the plasmid backbone. The plasmids pCONn-U and pCONc-U were generated via Gibson assembly. Here the plasmids pCONn and pCONc were amplified using the primer pair #104/#105. As insert a sequence of the plasmid pLC[56] was amplified using the primer pair #101/art2_F(JI). This insert contains non-coding DNA sequence and an ampicillin resistance gene (total 2,258 bp).

The plasmids pCONn-P and pCONc-P were generated by inserting the parA locus of plasmid R1 into the respective model plasmids. The *parA* locus contains *parM*, *parR* and *parC*, which constitute an active partitioning system, was amplified using the primer pair #98/#106. The PCR product and the plasmids pCONn and pCONc were digested using Hind III and the insert was cloned into the plasmid backbone. The construction of pCONn-T and pCONc-T was done via Gibson assembly (NEBuilder R © protocol; New England Biolabs) using the primer

476 pair #99/#100 to amplify the *parB* locus of plasmid R1, containing the *hok*/*sok* toxin-antitoxin
477 system. The primer pair #117/#118 was used to amplify plasmid pCONn and plasmid pCONc.
478 To construct the plasmids pCONn-PT and pCONc-PT, the plasmids pCONn-P and pCONc-P
479 were amplified using the primer pair #114/#115. The parB locus of plasmid R1 was amplified
480 using the primer pair #112/#113. The PCR products were then used for a Gibson assembly.
481 The donor of the *agrB-dqlB* locus was the self-transmissible pOXA-484 plasmid, a
482 previously described IncX3 plasmid[39]. The locus was cloned into plasmid pBAD30 under the
483 control of a P_{BAD} promoter[64]. The construct pBAD30-*dqlB* carries the *dqlB* gene. Primers
484 TF21/22 were used for the *dqlB* insert, while the backbone was digested using the restriction
485 enzymes XbaI and EcoRI-HF (Supplementary Table S2). The construct pBAD30-*argB-dqlB*
486 incorporates the complete *agrB-dqlB* locus downstream of the P_{BAD} promoter. Primers
487 TF57/58 were used for the complete locus insert and pBAD30 was digested with Eco53kI
488 (Supplementary Table S2). The plasmid pCONn-TD was constructed from pCONn with the
489 complete *agrB-dqlB* locus incorporates in the same locus as the other plasmid stability
490 modules (Fig. 1B). The plasmid pCONn-TD was constructed using TF53/54 for the complete
491 *agrB-dqlB* locus insert and TF49/50 for the backbone (Supplementary Table S2). All pBAD30
492 derived plasmids, as well as pCONn-TD, were constructed via Gibson Assembly®.

493 **Preparation of electrocompetent cells**

494 For the preparation of electrocompetent cells, liquid cultures (2 mL) of the respective bacterial
495 strain were inoculated and grown overnight with constant shaking (200 rpm in Innova 44). If
496 the strain contained a plasmid, selective pressure for plasmid maintenance was applied. After
497 the overnight incubation, 1 ml of the grown liquid cultures was used to inoculate fresh LB
498 medium (1 l) in a 2 l flask. After 2 h the OD_{600} of the new culture was measured and once an
499 OD_{600} of 0.6 had been reached, the flask was transferred onto ice and incubated with constant
500 shaking for 20 min. The culture was afterwards split into four Falcon Tubes and centrifuged in
501 a pre-cooled (4 °C) centrifuge (Heraeus Multifuge X3R with the Fiberlite F15-8x50cy rotor;
502 Thermo Scientific) at 4000 g for 10 min. The supernatant was discarded, and the cells were
503 washed twice with cold 10% glycerol. The cells were afterwards pooled again by resuspending
504 them in 1 ml of 10% glycerol, i.e., the cells in the first tube were resuspended in 1 ml of 10%
505 glycerol and this suspension was then used to resuspend the next tube and so on. The final
506 suspension of cells was then transferred into a 2 ml Eppendorf reaction tube. This tube was
507 again centrifuged in a pre-cooled (4 °C) centrifuge (Heraeus Fresco 21; Thermo Scientific) at
508 4000 g for 10 min and the supernatant discarded. The cells were afterwards resuspended in
509 0.4 ml of cold 10% glycerol, aliquoted into 40 µl each and stored at -80 °C [65].

510 **Transformation**

511 The transformation during plasmid construction was performed with a heat-shock protocol
512 using chemically competent cells[66]. Competent cells aliquots were thawed, mixed with the
513 DNA and incubated on ice for 30 minutes. The transformation tubes were shifted to a water
514 bath at 42 °C and incubated between 45 – 90 s, followed by 2 min rest on ice. The
515 transformation mix was then transferred into fresh LB medium and incubated for 1 h at 37 °C
516 and 200 rpm. After the outgrowth period, appropriate dilutions and volumes were plated on
517 selective LB agar plates and incubated overnight.
518 The introduction of model plasmids into the naïve host population was performed with
519 electroporation. The electrocompetent cells were mixed with the respective amounts of
520 plasmid DNA and transferred into cuvettes with a gap of 1 mm. The cells were electroporated
521 at 1350 V, 25 uF and 600 Ω using the BioRad Genepulser Xcell. The transformation mix was

522 transferred into 1 ml of fresh LB medium and incubated for 1 – 3 h at 37 °C with 200 rpm. The
523 cultures were then either exposed to positive selection in liquid media or plated on selective
524 LB agar plates.

525 **DNA sequencing and plasmid copy number**

526 In order to estimate the plasmid copy number (PCN) host populations were grown in liquid
527 media with either Km or Cm for 16 h. The bacterial culture had therefore reached the stationary
528 phase. From the cultures, chromosomal (cDNA) and plasmid DNA (pDNA) was extracted
529 using the Promega WIZARD DNA extraction kit as instructed by the manufacturer. The
530 acquired DNA sequenced by Illumina (Eurofins) to an average sequencing depth of 469.31
531 cDNA and 2698.64 pDNA (Supplementary Table 1). This allowed us to calculate the average
532 PCN for the plasmid backbone. The analysis of the sequencing data was done using galaxy
533 version 24.1[67]. Plasmid sequencing for validation was done by extracting plasmids using the
534 Monarch Miniprep Kit (New England Biolabs) followed by making use of either the Whole
535 Plasmid Sequencing or overnight sequencing service of Eurofins.

536 **Quantification of plasmid loss**

537 For the plasmid stability assays, a population of plasmid host were incubated overnight
538 incubation at 37 °C and 200 rpm in liquid media under selective pressure with the respective
539 antibiotic. If the cells were taken from a cryostock, an additional day of growth on a selective
540 plate was added prior to the preadaptation in liquid medium. Following the preadaptation, the
541 OD_{600} of the cultures was used to calculate the volume required for inoculation of fresh medium
542 to an OD of 0.01. The ratio of hosts within the population at t_0 was determined by differential
543 plating after inoculation of the liquid culture. Every 24 h for five days, i.e., t_0 to t_{120} the
544 proportion of hosts was determined by plating on LB plates and selective plates containing
545 antibiotics according to the tested plasmid: 10 ng/µl Km for pCONn variants; 10 ng/µl Cm for
546 pCONc variants. In each day, fresh LB liquid media was inoculated with the previous culture,
547 using a dilution of 1:100. The results are supplied in supplementary dataset S1.

548 **Plasmid fitness effect on the host**

549 The plasmid carrying strains and the plasmid-free strain were competed against a plasmid-
550 free strain, that is isogenic except for a Tp^R gene supplying resistance to trimethoprim.
551 Bacterial cultures were first preadapted in liquid media for three days. The preadaptation was
552 done with and without antibiotics depending on whether the respective strain carries a plasmid.
553 The cultures were transferred to fresh liquid LB media every 24 h with dilution of 1:100. After
554 the preadaptation the strains were diluted 1:10 and the OD_{600} of the strains was measured
555 and adjusted to 0.1. The competing strains were used to inoculate 1960 µl of fresh LB media,
556 with 20 µl from strain 1 and another 20 µl from strain 2. The competition mix was then divided
557 into two 1000 µl volumes, where one was incubated overnight and the other was used to
558 determine the frequency of competing strains at the starting point t_0. The strains' frequency at
559 t_0 and t_{24} was determined by plating the competition mix in appropriate dilutions onto non-
560 selective LB agar plates, LB agar plates with either Km (10 ng/µl) or Cm (10 ng/µl) and LB
561 agar plates with Trim (125 ng/µl). The relative fitness (*w*) was calculated using the Malthusian
562 parameter formula[68] and compared between plasmid genotypes with two-sided Wilcoxon test.
563 The results are supplied in supplementary dataset S2.

564 **Intracellular plasmid competitions**

565 In the intracellular head-to-head plasmid competition, an *E. coli* MG1655 *recA*::*tetA* host
566 population was electroporated and transformed with the two competing plasmids, with 1 ng of

plasmid DNA (or more) each, at the same time. This was followed by 3 h of incubation in non-selective liquid LB medium to allow for plasmid establishment in the host. After the outgrowth period, antibiotics were added to a concentration of 10 ng/µl for Km and 3 ng/µl for Cm. The liquid culture was then incubated overnight for ~16 h. The culture was then diluted and plated on double selective media (Km 10 ng/µl, Cm 10 ng/µl) and incubated at 20 °C for 3 days to ensure that only hosts carrying both plasmids survive, while keeping the growth rate minimal. Five out of the colonies that grew on the double selective plate (termed initial colonies) were excised and resuspended in PBS (see also Fig. 1A). The suspended initial colonies were plated on non-selective agar LB plates that were incubated overnight at 37 °C. After the incubation period, five colonies per initial colony were excised and resuspended in PBS. The frequency of hosts and plasmid-free cells were quantified by plating on non-selective LB agar plates and selective LB agar plates with either Km (10 ng/µl), Cm (10 ng/µl) or both antibiotics. Additionally, the presence of competing plasmids was validated in the initial colonies in a similar manner. The results are supplied in supplementary datasets S3, S4.

Model simulation

We simulate the population dynamics of plasmid hosts and plasmid-free cells using a model of random segregation extended for incomplete plasmid monomerization post replication, active partitioning and post segregational killing (induced by partition systems and toxin antitoxin systems respectively). The model describes the population dynamics of bacterial cells characterized by the abundances i_{p_1} and i_{p_2} of plasmid copies of two different plasmid variants, referred to by p_1 and p_2, at cell birth. Plasmids copies replicate following a Polya urn scheme the cell harbors $2n_c$ plasmid copies, i.e. plasmids are randomly selected for replication from the cell's plasmid pool and replicates are retrieved into the pool before the next replication event, which results in $J_{p_1}^{\mathrm{rep}}$ plasmid copies of variant p_1 and $J_{p_2}^{\mathrm{rep}} = 2\mathrm{n_c} - J_{p_1}^{\mathrm{rep}}$ of variant p_2 in the cell; note that $J_{p_1}^{\bullet}, J_{p_2}^{\bullet}$ here and below are random variables. For the stable plasmid variant, $p_{1,2} = S$, plasmid copies monomerize (optimally) leaving $J_S^{\mathrm{mon}} = J_S^{\mathrm{rep}}$ plasmid units that segregate randomly at cell division. For the other plasmid variants, $p_{1,2} \in U, P, T, PT$, we model suboptimal monomerization by removal of from the pool of the (new $J_{p_1}^{\mathrm{rep}} - i_{p_1}$) plasmid replicates by binomial sampling with parameter p_{def} (monomerization deficiency) leaving $J_{\mathrm{p}_{1,2}}^{\mathrm{mon}}$ copies (of the respective plasmid variants) before cell division. For the plasmid variant harboring a partition system, $p_1 \in P, PT$, plasmids form $\lfloor J_{p_1}^{\mathrm{mon}}/2 \rfloor$ pairs and plasmid copies of pairs are independently successfully partitioned (divided) at cell division at probability p_{par}, i.e., they are distributed to different daughter cells, while the remaining copies are randomly distributed, resulting in $J_{p_1}^{\mathrm{div}}$ copies in one (of the two) daughter cells (see supplemental text for all probability distributions). The mathematical derivation of the probabilities $p_{i \to j}$ that a cell harboring i_{p_1} and i_{p_2} plasmid copies of variant 1 and 2 at cell birth, respectively, denoted by $\mathrm{i} = (i_{p_1}, i_{p_2})$ gives rise to a daughter cell of type $j = (j_{p_1}, j_{p_2})$ with j_{p_1} and j_{p_2} plasmid copies of variant 1 and 2 allows us to compute the expected number of j-type cells produced at cell division of an i-type cell. For plasmid variants encoding at toxin-antitoxin cells, $p_{1,2} \in T, PT$, cells may die from post-segregational killing. If the daughter cells lacks the plasmid ($j_{p_{1,2}} = 0$) while the parental cell carried a plasmid encoding the TA system ($i_{p_{1,2}} > 0$, see supplemental text for the case that both plasmid variants, p_1 and p_2, carry the TA system, which we do not examine in the results shown), cells only survive with a small chance $\mathrm{p}_{\mathrm{i}\to\mathrm{j}}^{\mathrm{surv}} = 1 - \mathrm{p}_{\mathrm{eff}}$ due to post-segregational killing escape. We defined the expected number of j-type cells that is produced at cell division of an i-type cell,

$$2p_{i\to j}p_{i\to j}^{\text{surv}} =: A_{ji},$$

which used to compute the cell-type proportions after g generations,

$$\mathbf{x}(g) = \frac{\boldsymbol{A}^g \boldsymbol{x}_0}{|\boldsymbol{A}^g \boldsymbol{x}_0|},$$

where $\mathbf{x}$ is a vector containing the proportions of all cell-types, A is a matrix with elements A_{ji} and the vector $\mathbf{x_0}$ describes the initial population state ($x_{0,i} = 1$, for $i = (i_{p_1}, i_{p_2}) = (1,1)$ for initial populations of heteroplasmic hosts and for $i = (i_{p_1}, i_{p_2}) = (1,0)$ for populations of homoplasmic hosts). Mathematical derivations and details on the estimation of the parameters can be found in the supplemental text. Simulations were performed using Mathematica® (V. 14).

Annotation of active partition systems and toxin-antitoxin systems in plasmid genomes

For the survey of active partition systems and toxin-antitoxin systems, we used a previously established *KES* dataset of enterobacterial genomes. The *KES* dataset includes 1114 chromosomes and 3098 plasmids from *Escherichia*; 755 chromosomes and 2693 plasmids from *Klebsiella*; 572 chromosomes and 993 plasmids from *Salmonella* (RefSeq database; version 01/2021). Protein coding genes in the *KES* genomes were clustered into 32,623 gene families as previously described[29]. Briefly, reciprocal best hits (RBHs) of protein sequences between all replicon pairs were identified using MMseqs2[69] (v.13.45111, with module easy-rbh applying a threshold of E-value $\leqslant 1 \times 10^{-10}$). RBHs were further compared by global alignment using parasail-python[70] (v. 1.2.4, with the Needleman-Wunsch algorithm). Sequence pairs with $\geqslant$ 30% identical amino acids were clustered into gene families using a high-performance parallel implementation of the Markov clustering algorithm[71] (HipMCL with parameter –abc -I 2.0).

We first collected NTPase and CBP protein sequences, associated with type I, II, III, and a novel type (StbA) partition systems from the literatures[10,72,73]. These sequences were aligned against the 32,623 clusters of homologous gene families using BLAST [74], followed by manual inspection of gene location and orientation, we kept only the neighboring pairs of NTPase and CBP encoded on the same strand (except StbA system). Our survey resulted into 34 partition systems in the plasmids (Supplementary table S3). Note that our search did not include the *par* sites (centromere like DNA sequence). Active partition systems were exclusively found in plasmids with size larger than 19 kbp, with exceptions of seven plasmids smaller than 19 kbp, which were identified as putative truncated assemblies after BLAST searches against NCBI nr database.

Utilizing the Toxin-Antitoxin Database (TADB) and the tool TAfinder (version 2.0)[34] we identified TA systems belong to 44 families in 4,456 plasmids (a full list of TAs is in supplementary table S4). After organizing TA system-carrying plasmids into PTUs, we manually searched for toxin gene annotation in those PTUs where no TA system was detected. A *dqlB* gene encoding DinQ-like toxin was annotated after a recent RefSeq update, *dqlB* shares 41.2% global identity with the amino acids sequence of DinQ toxin in *E. coli* K-12 strain MG1655 chromosome. ESMfold[75] revealed the same helix structure of the two proteins, TM-align[76] further showed their highly structural similarity with the TM-scores 0.96 and 0.85. We used BLAST against the Rfam database to predict the two small RNA sequence, *agrA* and *agrB*, in the potential *dqlB* locus on PTU-X3 plasmids. This resulted in a prediction of *argB* antisense RNA locus in the neighborhood of *dqlB* but failed in predicting an *agrA* locus.

References

1. Bergstrom, C. T., Lipsitch, M. & Levin, B. R. Natural selection, infectious transfer and the existence conditions for bacterial plasmids. *Genetics* **155**, 1505–1519 (2000).
2. Paulsson, J. Multileveled selection on plasmid replication. *Genetics* **161**, 1373–1384 (2002).
3. Summers, D. K. & Sherratt, D. J. Multimerization of high copy number plasmids causes instability: CoIE1 encodes a determinant essential for plasmid monomerization and stability. *Cell* **36**, 1097–1103 (1984).
4. Summers, D. K., Beton, C. W. H. & Withers, H. L. Multicopy plasmid instability: the dimer catastrophe hypothesis. *Molecular Microbiology* **8**, 1031–1038 (1993).
5. Stewart, F. M. & Levin, B. R. The population biology of bacterial plasmids: *a priori* conditions for the existence of conjugationally transmitted factors. *Genetics* **87**, 209–228 (1977).
6. Loftie-Eaton, W. *et al.* Compensatory mutations improve general permissiveness to antibiotic resistance plasmids. *Nat Ecol Evol* **1**, 1354–1363 (2017).
7. Alonso-del Valle, A. *et al.* Variability of plasmid fitness effects contributes to plasmid persistence in bacterial communities. *Nat Commun* **12**, 2653 (2021).
8. Rodríguez-Beltrán, J., DelaFuente, J., León-Sampedro, R., MacLean, R. C. & San Millan, A. Beyond horizontal gene transfer: the role of plasmids in bacterial evolution. *Nat Rev Microbiol* (2021) doi:10.1038/s41579-020-00497-1.
9. Castañeda-Barba, S., Top, E. M. & Stalder, T. Plasmids, a molecular cornerstone of antimicrobial resistance in the One Health era. *Nat Rev Microbiol* **22**, 18–32 (2024).
10. Bouet, J.-Y. & Funnell, B. E. Plasmid Localization and Partition in *Enterobacteriaceae*. *EcoSal Plus* **8**, 10.1128/ecosalplus.ESP-0003–2019 (2019).
11. Planchenault, C. *et al.* Intracellular positioning systems limit the entropic eviction of secondary replicons toward the nucleoid edges in bacterial cells. *J Mol Biol* **432**, 745–761 (2020).
12. Ramiro-Martínez, P., Quinto, I. de, Gama, J. A. & Rodríguez-Beltrán, J. Universal rules govern plasmid copy number. 2024.10.04.616648 Preprint at https://doi.org/10.1101/2024.10.04.616648 (2024).
13. Gerdes, K., Møller-Jensen, J. & Jensen, R. B. Plasmid and chromosome partitioning: surprises from phylogeny. *Mol Microbiol* **37**, 455–466 (2000).
14. Jurėnas, D., Fraikin, N., Goormaghtigh, F. & Van Melderen, L. Biology and evolution of bacterial toxin–antitoxin systems. *Nat Rev Microbiol* **20**, 335–350 (2022).
15. Gerdes, K., Helin, K., Christensen, O. W. & Løbner-Olesen, A. Translational control and differential RNA decay are key elements regulating postsegregational expression of the killer protein encoded by the *parB* locus of plasmid R1. *J Mol Biol* **203**, 119–129 (1988).
16. Lima-Mendez, G. *et al.* Toxin-Antitoxin gene pairs found in Tn3 family transposons appear to be an integral part of the transposition module. *mBio* **11**, 10.1128/mbio.00452-20 (2020).
17. Kurata, T. *et al.* A hyperpromiscuous antitoxin protein domain for the neutralization of diverse toxin domains. *Proc Natl Acad Sci USA* **119**, (2022).
18. Brendler, T., Reaves, L. & Austin, S. Interplay between plasmid partition and postsegregational killing systems. *J Bacteriol* **186**, 2504–2507 (2004).
19. Datta, N. & Kontomichalou, P. Penicillinase synthesis controlled by infectious R factors in *Enterobacteriaceae*. *Nature* **208**, 239–241 (1965).

20. Dam, M. & Gerdes, K. Partitioning of plasmid R1. Ten direct repeats flanking the *parA* promoter constitute a centromere-like partition site *parC*, that expresses incompatibility. *J Mol Biol* **236**, 1289–1298 (1994).
21. Gerdes, K. *et al.* Mechanism of postsegregational killing by the *hok* gene product of the *parB* system of plasmid R1 and its homology with the *relF* gene product of the *E. coli relB* operon. *EMBO J* **5**, 2023–2029 (1986).
22. Thisted, T. & Gerdes, K. Mechanism of post-segregational killing by the *hok/sok* system of plasmid R1. *J Mol Biol* **223**, 41–54 (1992).
23. Gerdes, K., Rasmussen, P. B. & Molin, S. Unique type of plasmid maintenance function: postsegregational killing of plasmid-free cells. *Proc Natl Acad Sci USA* **83**, 3116–3120 (1986).
24. Antoine, R. & Locht, C. Isolation and molecular characterization of a novel broad-host-range plasmid from *Bordetella bronchiseptica* with sequence similarities to plasmids from Gram-positive organisms. *Mol Microbiol* **6**, 1785–1799 (1992).
25. Wein, T., Hülter, N. F., Mizrahi, I. & Dagan, T. Emergence of plasmid stability under non-selective conditions maintains antibiotic resistance. *Nat Commun* **10**, 2595–13 (2019).
26. Hülter, N. F., Wein, T., Effe, J., Garoña, A. & Dagan, T. Intracellular competitions reveal determinants of plasmid evolutionary success. *Front Microbiol* **11**, 2062 (2020).
27. Cooper, T. F. & Heinemann, J. A. Postsegregational killing does not increase plasmid stability but acts to mediate the exclusion of competing plasmids. *Proc Natl Acad Sci USA* **97**, 12643–12648 (2000).
28. Garoña, A., Santer, M., Hülter, N. F., Uecker, H. & Dagan, T. Segregational drift hinders the evolution of antibiotic resistance on polyploid replicons. *PLOS Genetics* **19**, e1010829 (2023).
29. Wang, Y. & Dagan, T. The evolution of antibiotic resistance islands occurs within the framework of plasmid lineages. *Nat Commun* **15**, 4555 (2024).
30. Gilmour, M. W., Thomson, N. R., Sanders, M., Parkhill, J. & Taylor, D. E. The complete nucleotide sequence of the resistance plasmid R478: defining the backbone components of incompatibility group H conjugative plasmids through comparative genomics. *Plasmid* **52**, 182–202 (2004).
31. Lawley, T. D. & Taylor, D. E. Characterization of the double-partitioning modules of R27: Correlating plasmid stability with plasmid localization. *J Bacteriol* **185**, 3060–3067 (2003).
32. Sheppard, A. E. *et al.* Complete Genome Sequence of KPC-Producing *Klebsiella pneumoniae* Strain CAV1193. *Genome Announcements* **4**, 10.1128/genomea.01649-15 (2016).
33. AbuOun, M. *et al.* A genomic epidemiological study shows that prevalence of antimicrobial resistance in *Enterobacterales* is associated with the livestock host, as well as antimicrobial usage. *Microb Genom* **7**, 000630 (2021).
34. Guan, J. *et al.* TADB 3.0: an updated database of bacterial toxin–antitoxin loci and associated mobile genetic elements. *Nucleic Acids Research* **52**, D784–D790 (2024).
35. Li, X. *et al.* Dissemination of blaNDM-5 gene via an IncX3-type plasmid among non-clonal Escherichia coli in China. *Antimicrobial Resistance & Infection Control* **7**, 59 (2018).
36. Wang, Y. *et al.* Comprehensive resistome analysis reveals the prevalence of NDM and MCR-1 in Chinese poultry production. *Nat Microbiol* **2**, 1–7 (2017).
37. Bustamante, P. & Iredell, J. R. Carriage of type II toxin-antitoxin systems by the growing group of IncX plasmids. *Plasmid* **91**, 19–27 (2017).
38. Weel-Sneve, R. *et al.* Single transmembrane peptide DinQ modulates membrane-dependent activities. *PLOS Genet* **9**, e1003260 (2013).

39. Sommer, J. *et al.* OXA-484, an OXA-48-type carbapenem-hydrolyzing class D β-lactamase from *Escherichia coli*. *Front. Microbiol.* **12**, (2021).
40. Baffert, Y. *et al.* Genetic determinants of pOXA-48 plasmid maintenance and propagation in *Escherichia coli*. 2025.01.23.634512 Preprint at https://doi.org/10.1101/2025.01.23.634512 (2025).
41. Calvo-Villamañán, A., Sastre-Dominguez, J., Barrera-Martín, Á., Costas, C. & Millan, Á. S. Dissecting pOXA-48 fitness effects in clinical enterobacteria using plasmid-wide CRISPRi screens. 2025.01.23.633114 Preprint at https://doi.org/10.1101/2025.01.23.633114 (2025).
42. Peters, J. E. Tn7. *Microbiol Spectr* **2**, 10.1128/microbiolspec.mdna3-0010–2014 (2014).
43. Schleper, C., Röder, R., Singer, T. & Zillig, W. An insertion element of the extremely thermophilic archaeon *Sulfolobus solfataricus* transposes into the endogenous β-galactosidase gene. *Molec. Gen. Genet.* **243**, 91–96 (1994).
44. Gluck-Thaler, E. *et al.* Giant starship elements mobilize accessory genes in fungal genomes. *Mol Biol Evol* **39**, msac109 (2022).
45. Liu, P., Cuerda-Gil, D., Shahid, S. & Slotkin, R. K. The epigenetic control of the transposable element life cycle in plant genomes and beyond. *Annu Rev Genet* **56**, 63–87 (2022).
46. Feusier, J. *et al.* Pedigree-based estimation of human mobile element retrotransposition rates. *Genome Res.* **29**, 1567–1577 (2019).
47. Val, M.-E. *et al.* A checkpoint control orchestrates the replication of the two chromosomes of Vibrio cholerae. *Sci Adv* **2**, e1501914 (2016).
48. Komluski, J., Stukenbrock, E. H. & Habig, M. Non-Mendelian transmission of accessory chromosomes in fungi. *Chromosome Res* **30**, 241–253 (2022).
49. Ferree, P. M. *et al.* What is a B chromosome? Early definitions revisited. *G3 Genes|Genomes|Genetics* **14**, jkae068 (2024).
50. Yoshida, Y. & Mogi, Y. How do plastids and mitochondria divide? *Microscopy* **68**, 45–56 (2019).
51. Kozlova, S., Morozova, N., Ispolatov, Y. & Severinov, K. Dependence of post-segregational killing mediated by Type II restriction–modification systems on the lifetime of restriction endonuclease effective activity. *mBio* **15**, e01408-24 (2024).
52. Rodríguez-Beltrán, J. *et al.* Multicopy plasmids allow bacteria to escape from fitness trade-offs during evolutionary innovation. *Nat Ecol Evol* **2**, 873–881 (2018).
53. Dewan, I. & Uecker, H. Evolutionary rescue of bacterial populations by heterozygosity on multicopy plasmids. 2024.05.29.596466 Preprint at https://doi.org/10.1101/2024.05.29.596466 (2024).
54. Santer, M., Kupczok, A., Dagan, T. & Uecker, H. Fixation dynamics of beneficial alleles in prokaryotic polyploid chromosomes and plasmids. *Genetics* **222**, iyac121 (2022).
55. Garoña, A., Hülter, N. F., Romero Picazo, D. & Dagan, T. Segregational Drift Constrains the Evolutionary Rate of Prokaryotic Plasmids. *Mol Biol Evol* **38**, 5610–5624 (2021).
56. Ilhan, J. *et al.* Segregational drift and the interplay between plasmid copy number and evolvability. *Mol Biol Evol* **36**, 472–486 (2019).
57. Khachaturyan, M., Santer, M., Reusch, T. B. H. & Dagan, T. Heteroplasmy is rare in plant mitochondria compared with plastids despite similar mutation rates. *Mol Biol Evol* **41**, msae135 (2024).
58. Lopatkin, A. J. *et al.* Persistence and reversal of plasmid-mediated antibiotic resistance. *Nat Commun* **8**, 1689 (2017).

59. León-Sampedro, R. *et al.* Pervasive transmission of a carbapenem resistance plasmid in the gut microbiota of hospitalized patients. *Nat Microbiol* **6**, 606–616 (2021).
60. Hanahan, D. Studies on transformation of *Escherichia coli* with plasmids. *J Mol Biol* **166**, 557–580 (1983).
61. Tas, H., Nguyen, C. T., Patel, R., Kim, N. H. & Kuhlman, T. E. An Integrated System for Precise Genome Modification in *Escherichia coli*. *PLOS ONE* **10**, e0136963 (2015).
62. Crépin, S., Harel, J. & Dozois, C. M. Chromosomal complementation using Tn *7* transposon vectors in enterobacteriaceae. *Appl Environ Microbiol* **78**, 6001–6008 (2012).
63. Choi, K.-H. *et al.* A Tn7-based broad-range bacterial cloning and expression system. *Nat Methods* **2**, 443–448 (2005).
64. Guzman, L. M., Belin, D., Carson, M. J. & Beckwith, J. Tight regulation, modulation, and high-level expression by vectors containing the arabinose PBAD promoter. *J Bacteriol* **177**, 4121–4130 (1995).
65. Dower, W. J., Miller, J. F. & Ragsdale, C. W. High efficiency transformation of *E. coli* by high voltage electroporation. *Nucleic Acids Res* **16**, 6127–6145 (1988).
66. Inoue, H., Nojima, H. & Okayama, H. High efficiency transformation of *Escherichia coli* with plasmids. *Gene* **96**, 23–28 (1990).
67. The Galaxy Community. The Galaxy platform for accessible, reproducible, and collaborative data analyses: 2024 update. *Nucleic Acids Res* **52**, W83–W94 (2024).
68. Bouma, J. E. & Lenski, R. E. Evolution of a bacteria/plasmid association. *Nature* **335**, 351–352 (1988).
69. Steinegger, M. & Söding, J. MMseqs2 enables sensitive protein sequence searching for the analysis of massive data sets. *Nat Biotechnol* **35**, 1026–1028 (2017).
70. Daily, J. Parasail: SIMD C library for global, semi-global, and local pairwise sequence alignments. *BMC Bioinformatics* **17**, 81 (2016).
71. Azad, A., Pavlopoulos, G. A., Ouzounis, C. A., Kyrpides, N. C. & Buluç, A. HipMCL: a high-performance parallel implementation of the Markov clustering algorithm for large-scale networks. *Nucleic Acids Research* **46**, e33 (2018).
72. Guynet, C., Cuevas, A., Moncalián, G. & Cruz, F. de la. The stb Operon Balances the Requirements for Vegetative Stability and Conjugative Transfer of Plasmid R388. *PLOS Genetics* **7**, e1002073 (2011).
73. Brooks, A. C. & Hwang, L. C. Reconstitutions of plasmid partition systems and their mechanisms. *Plasmid* **91**, 37–41 (2017).
74. Camacho, C. *et al.* BLAST+: architecture and applications. *BMC Bioinformatics* **10**, 421 (2009).
75. Rives, A. *et al.* Biological structure and function emerge from scaling unsupervised learning to 250 million protein sequences. *Proc Natl Acad Sci USA* **118**, e2016239118 (2021).
76. Zhang, Y. & Skolnick, J. TM-align: a protein structure alignment algorithm based on the TM-score. *Nucleic Acids Res* **33**, 2302–2309 (2005).

Data and Material Availability

The data underlying this article are available in the article and in its online supplementary material.

Acknowledgements

We thank Stephan Göttig from the University Hospital, Goethe University Frankfurt am Main, for contribution of *E. coli* EC-JS316 isolate. We thank Lisa Hartmann for critical comments on the manuscript. We thank Susanne Landis for graphical illustrations. This research was supported in part through high-performance computing resources available at the Kiel University computing center.

Funding

German Science Foundation; RTG 2501 TransEvo, grant number: 456882089.
European Research Council; pMolEvol, grant number: 101043835.

Author contributions

JE, NH and TD conceived the study and designed plasmid competitions. JE performed plasmid competitions and analyzed their results. JE, MS and TD conceived model simulations. MS designed and performed model simulations. YW and TD conceived the plasmid comparative genomes. YW designed and performed plasmid comparative genomics and detection of TA systems. TF and NH designed TA system validation experiments. TF performed TA validation experiments and analyzed their results with assistance from NH. All authors wrote the manuscript.

Competing interests

The authors declare no competing interests.