PLASMID CARRIAGE AND THE UNORTHODOX USE OF FITNESS IN MICROBIOLOGY

CARLOS REDING

Experimental Medicine Division, Nuffield Department of Medicine, University of Oxford. UK.

Corresponding author: `reding@disroot.org`

The link between fitness and reproduction rate is a central tenet in microbiology, and indeed evolutionary biology: Mutants reproducing faster than the dominant wild-type are favoured by selection, but otherwise the mutation is lost. This link was given by Ronald Fisher in 1930 under the assumption that fitness can only change through mutations that boost or hinder growth rate, whence the use of logarithms on growth data by experimentalists. Here I show that logarithms are highly sensitive to sampling times, resulting in fitness estimates that are not constant over the growth of bacterial cultures. This variability invalidates typical selection measurements, and 'unfit' mutants can be co-maintained if they reach their equilibrium. And this is what I observed in competition assays between two *Escherichia coli* constructs, one of which harbours a non-transmissible plasmid that protects against tetracycline (pGW155B), without using antibiotic. Despite growing 40% slower than its drug-sensitive counterpart, the construct harbouring the plasmid persisted throughout the competition. And, perhaps more importantly, maintained the plasmid. My study suggests that reliance on growth rate masks that selection on plasmid carriage may be stronger than previously thought—explaining the seemingly-paradoxical abundance of plasmids in nature.

I. INTRODUCTION

The fitness of microbial genotypes is typically measured in terms of relative growth between two competing genotypes [1]. Population genetics [2–5] and microbiology [6, 7] textbooks define the fitness of microbial populations as their intrinsic growth rate. This notion stems from the use of the exponential growth model to study natural selection, where fitness (m, also known as the Malthusian parameter [8]) is given by the exponential model $N(t) = N_0 e^{mt}$. Here N_0 is the initial number of individuals of a given genotype, and $N(t)$ the number of individuals at time t. Thus, the ratio $W_{ij} = m_i/m_j$ gives the fitness of a genotype with respect to another (W_{ij}) [9] which is then used to calculate the selection coefficient (S_{ij}) as $S_{ij} = (m_i - m_j)/m_j = W_{ij} - 1$ [2, 9]. Prior theory shows this is indeed an over-simplification [10], but fitness concepts postulated by theory are difficult to implement in the laboratory [11, 12]. The result is an increasing gap between data and theory [13, 14]. So, given the simplicity of the rationale above, W and S are widely used in the laboratory to estimate relative fitness and the strength of natural selection in evolutionary and microbiological studies [9, 12, 15–18]. Here, researchers use of logarithms on growth data typically from pairwise competition assays [9, 15, 19–24] to estimate the Malthusian parameter m and then relative fitness. An approach that, it is argued [15], aggregates the contribution of each growth phase to fitness.

Other studies aimed to capture the specific contribution of each phase [25, 26], alas, intrinsic growth rate is a central parameter of such models.

If fitness, as intrinsic growth rate, is an inherited trait, it is reasonable to expect that mutants with greater fitness will be represented with greater proportion in the next generation—thus replacing the wild-type genotype [2, 8, 27]. This rationale underpins the study of plasmid maintenance in bacteria, and yet, it cannot explain the abundance of plasmids in nature [28–31]. Plasmids are extra-chromosomal elements that bacteria can acquire through horizontal gene transfer to thrive in unfavourable environments [28]. In favourable environments however, where plasmids do not add an advantage, they impose a fitness cost derived from the synthesis of additional genetic material during the replication and expression of plasmid-borne genes [32, 33]. Thus, conventional wisdom holds that plasmids should not exist without positive selection [28]—indeed a typical assumption also used in epidemiological models [34, 35]. But they do, and are ubiquitous in nature [36], creating a so-called *plasmid paradox* [28, 30]. Recent evidence showed that plasmids undergoing horizontal-gene transfer can propagate fast enough to persist despite the costs [37]. The problem is that most of the known plasmids are *not* transmissible [36]. So, how can we explain their abundance? The question here is whether a myriad of conditions exist to select for plasmid-carriage, or the logic above behind the notion of fitness is sub-optimal. After all, the availability of data testing the predictions of population genetics theory is extremely scarce [13, 14, 38], to a point that even whether plasmids do indeed incur fitness costs at all has been questioned in the past [19].

A corollary from the fitness metric above is that the ratio m_i/m_j is a constant, following the view that fitness is inheritable and therefore it can only change through mutations that affect m [9]. This ratio is rarely checked experimentally. Below I show it is not constant and, rather, m_i/m_j changes predictably in smaller timescales during the growth of populations. This variability revealed that plasmid-harbouring cells can be co-maintained when competing against genotypes reproduce faster contrary to what is currently postulated [28]. In this study, I co-maintained a construct of *Escherichia coli* harbouring a non-transmissible plasmid, with a tetracycline resistance gene, in competition assays without using antibiotics. Curiously, the selection coefficients for this constructs were negative, meaning it was being lost to selection. But the constructs were not only co-maintained, despite the 40% reduction in growth rate imposed by the plasmid with respect to its sensitive counterpart. The bacterium harbouring pGW155B kept the plasmid, even without exposure to tetracycline. Similarly, the sensitive construct was co-maintained in the presence of tetracycline concentrations that gave the drug-resistant construct an edge in terms of growth rate. In other words, my results suggest that there would be no such thing as a plasmid paradox and, rather, selection for plasmids—commonly studied through the exclusive use of growth rates to estimate fitness—may be stronger than previously thought.

II. RESULTS

Let me assume two competing genotypes, mutants A and B, that grow consistently with population genetics theory [2, 3]. The following system of ordinary differential equations de-

scribes the change in the number of individuals of each mutant over time:

$$\frac{dA}{dt} = m_a A, \quad \frac{dB}{dt} = m_b B, \tag{1}$$

where m is the population growth rate increase as introduced above, and $A(0) = B(0) = x$ where x is strictly positive. Note that m is derived from random birth (b) and death (d) rates, so that $m = b - d$ [2], and that it is independent of the number of individuals of each genotype. Each individual is equally likely to die or reproduce at any given time, and the age distribution is near to the equilibrium so that b and d are nearly constant. The growth dynamics of both mutants is illustrated in Figure 1A. If b and d are constant, m is a constant and therefore the relative fitness difference between genotypes $W_{ba} = m_b/m_a = k$ is also constant. In other words, if $m_b = \frac{ln[B(t)/B(0)]}{t}$ as derived from the exponential model introduced above [3, 9, 39] (see Methods), and $B(t)$ is the number of individuals of mutant B at an arbitrary time t, the relative fitness W_{ba} is the same regardless of t. The same applies to the fitness of A with respect to B. Indeed, if $m_b > m_a$, the number of mutant B individuals is higher than those of mutant A at all times as Figure 1A shows.

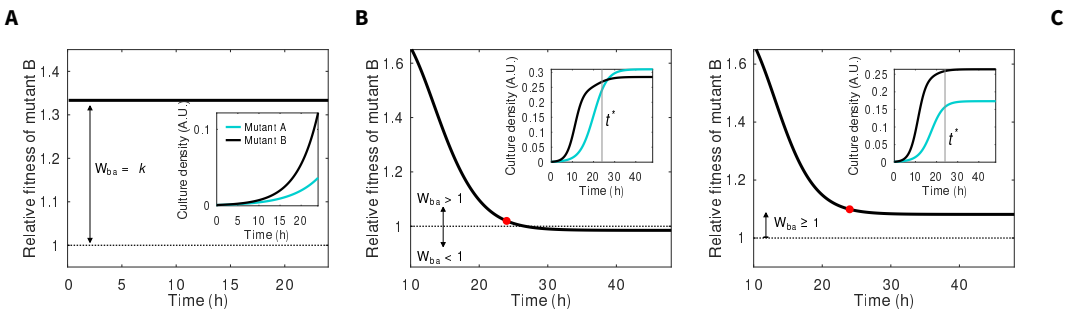

Figure 1. Change in relative fitness when populations have exponential and logistic growth. A) Change in relative fitness of mutant B (W_{ba}, black) over time when competing mutants have density-independent, exponential growth—assumed by population genetics theory for populations with continuous growth and overlapping generations [2–4]. I calculated the relative fitness as [9] $W_{ba} = m_b/m_a$, where m corresponds to the population growth rate for each mutant. The fitness of mutant A, reference, is shown as a black, dotted line. The inset shows the change in cell density, in arbitrary units (A. U.), during the competition. **B-C)** The same information is shown for mutants with density-dependent growth, with (B) and without (C) trade-off between reproduction rate and survival. The sampling time $t^* = 24$ hours, commonly used in microbiological assays, is noted by a red marker in the main plot, and a vertical, grey line in the insets. Model parameters for A) are $m_a = 0.15\ \mathrm{h}^{-1}$, $m_b = 0.2\ \mathrm{h}^{-1}$; for B) $m_a = 1.0275\ \mathrm{h}^{-1}$, $m_b = 1.05\ \mathrm{h}^{-1}$, $K_a = 12.5$ OD, $K_b = 5$ OD, $\alpha_{ab} = \alpha_{ba} = 0.15$; and for C) $m_a = 1.0275\ \mathrm{h}^{-1}$, $m_b = 1.05\ \mathrm{h}^{-1}$, $K_a = 10$ OD, $K_b = 5$ OD.

Now suppose that m depends on the number of individuals of each genotype. This is a reasonable assumption given that resources are depleted during microbial growth, and that resources are finite, limiting the abundance of each mutant over time:

$$\frac{dA}{dt} = m_a\left(1 - \frac{A^*}{K_a}\right)A, \quad \frac{dB}{dt} = m_b\left(1 - \frac{B^*}{K_b}\right)B. \tag{2}$$

Here K_a and K_b are the maximal population size attainable (*carrying capacity*) for mutants A and B, respectively, resulting from the limited availability of resources—note different efficiencies using the *same* resource can lead to different carrying capacities [40]. Now, the population growth rate m is corrected by the term $1 - (N_i/K_i)$. The Lotka-Volterra model of competition [41] includes the term $N_i^* = N_i + \alpha_{ij}N_j$, where $\alpha_{ij}N_j$ is the linear reduction in

growth rate—in terms of K—of species i by competing species j. This formalisation describes the limitation in growth imposed by the environment, due to finite resources, reducing m and the growth of both genotypes over time (Figure 1B and C). In this scenario, m_b is now corrected by changes in population size and, therefore, the above expression for W_{ba} must depend on time:

$$W_{ba} = \frac{m_b}{m_a} \rightarrow W_{ba}(t) = \overbrace{\frac{m_b}{m_a} \cdot \frac{1}{1+pt}}^{\substack{\text{contribution of} \\ m \text{ to } W}} + \overbrace{\frac{K_b^*}{K_a^*} \cdot \frac{pt}{1+pt}}^{\substack{\text{contribution of} \\ K^* \text{ to } W}} . \tag{3}$$

Here W_{ba} declines exponentially with time as Figures 1B and C illustrate, and eventually stabilises when the competing genotypes attain their carrying capacities K^*—new since their full potential cannot be realised in competition. Now, initially $W_{ba}(0) \approx m_b/m_a$ because $\frac{pt}{1+pt} \approx 0$ and $\frac{1}{1+pt} \approx 1$, with p being a parameter that controls the steepness of the change. As both genotypes grow, W_{ba} declines exponentially with $W_{ba}(t \rightarrow \infty) \approx K_b^*/K_a^*$ since now $\frac{pt}{1+pt} \approx 1$ and $\frac{1}{1+pt} \approx 0$. In other words, while fitness would depend on intrinsic growth rate (m) originally, once both genotypes are in equilibrium fitness will depend on the population size (K^*). Estimating fitness differences between two genotypes using equation (3), thus, requires prior knowledge of the carrying capacity.

This contrasts with the use of logarithms in the laboratory to estimate intrinsic growth rate as $m_b = \frac{ln[B(t)/B(0)]}{t}$ and $W_{ba} = m_b/m_a$ [9, 19, 21–24, 42–44]. Which means, there is an assumption of exponential growth as in Figure 1A. The result, however, is unreliable fitness estimates that can hinder, for example, comparative studies [45]. Figures 1B and C illustrate the importance of the sampling time for these estimates and how the use of logarithms can yield different—even conflicting—estimates. Thus, while it can be argued that estimations of m aim to *only* capture the exponential growth, the rest of the growth dynamics cannot be overlooked.

If, say, $m_b > m_a$ and $W_{ba} > 1$, which mutant, A or B, would be selected in pairwise competition assays? Despite the simplicity of equations (1) and (2), the answer is not straightforward. Figure 2 illustrates the change in relative abundance of two genotypes, A, and B, growing in competition for a common resource. These genotypes are haploid, with continuous growth and overlapping generations where m_a and m_b represent their population growth rates as introduced above. The competition is propagated at regular intervals c where the new initial conditions are given by $N_a(c_{+1}, 0) = N_a(c, t^*)d$, $N_b(c_{+1}, 0) = N_b(c, t^*)d$, with t^* being the time of growth allowed prior to the propagation and $0 \leq d \leq 1$ the dilution factor during the propagation step analog to experimental assays [9, 16, 23]. Following population genetics theory [2, 5, 6, 9], the outcome of this competition is straightforward: if, say, mutant B reproduces faster than mutant A ($m_b > m_a$), then $W_{ab} = m_a/m_b < 1$ and mutant A goes extinct (Figure 2A). Regardless of how long the competition is allowed to progress until the next propagation step, mutant A is always lost in competition with B. Only through the emergence of a new mutant A* with $m_a^* > m_b > m_a$ should this prediction change.

But when fitness changes through time the outcome of the competition becomes unclear, now depending on whether the competing genotypes reach their carrying capacities before the propagation: if the competition is propagated frequently, so that *both* genotypes are in active growth, the genotype reproducing faster will replace that reproducing slower consistently

with population genetics's notion of fitness—for this scenario. However, if one or both genotypes are allowed to reach their carrying capacities, both will be be co-maintained regardless of the in rates of population increase (Figures 2B and C). Here the relative frequency will be given by their carrying capacities K, so if reproduction and carrying capacity engage in a trade-off [40] for one genotype, this will be most frequent (Figure 2B). Importantly, these predictions are maintained regardless of the initial frequency of A (Figure S1), consistently with prior data [46].

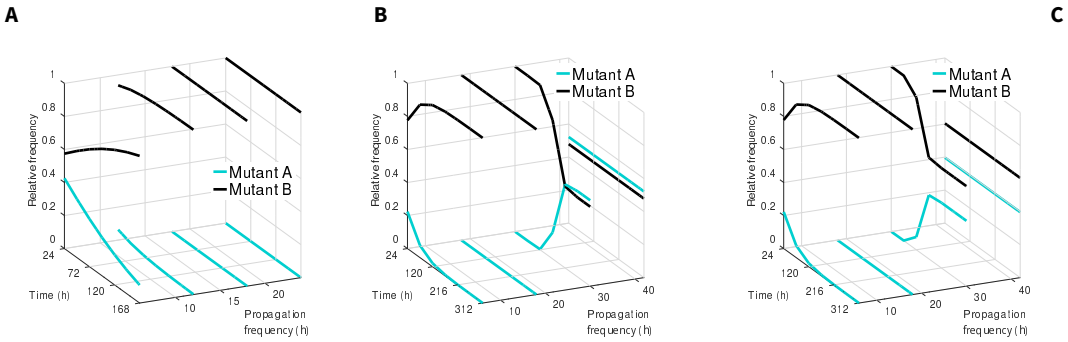

Figure 2. Change in relative frequency during competition depending on the length of the growth cycle. Changes in frequency over time for mutants A (cyan) and B (black) following propagations at arbitrary times t=6h, t=12h, t=18h, and t=24h. These times correspond to different stages of growth as Figure 1 shows. Each plot illustrate the case for genotypes with density-independent growth (A), or density-dependent growth where a genotype does engage in a trade-off between reproduction and survival (B) and when it does not (C). Parameter used in the model are those from Figure 1.

To test this prediction, I competed two constructs of *Escherichia coli* MC4100, Wyl and GB(c) (see Methods), and measured their variation in fitness through time. Figure 3A shows that both constructs reach their carrying capacities within 24 hours. The construct GB(c) carries the non-transmissible plasmid pGW155B [47] harbouring *tet(36)*, a ribosome-protection type resistance gene against tetracycline [47]. This plasmid lacks a *rep* gene to control tightly the partition of plasmids after cell division [48–50] (addgene vector database accession #2853). While pGW155B is a synthetic plasmid, many natural plasmids also lack partition systems [50, 51]. Both constructs have identical chromosome with exception of this plasmid, and the fluorescence gene they carry: cyan (*cfp*, GB(c)) or yellow (*yfp*, Wyl), to allow their identification in competition assays. Using a non-transmissible plasmid prevents cross-contamination by horizontal gene transfer between the resistant construct GB(c) and Wyl. Importantly, the use of constructs derived from MC4100 avoids interactions between competitors that may affect the outcome of the competition for reasons beyond pGW155B, like the production of bacteriocins [52].

pGW155B penalised the population growth rate of construct GB(c) by approximately 40% compared to Wyl (Figure 3B) as I found out using pure cultures. It might be argued that the 4-parameter logistic model I used does not fully capture the growth data. However, if I measure the maximum change between the reads (Euler's method, see Methods), the strong difference in growth rate is maintained. The duration of lag, and carrying capacity were also sensitive to plasmid carriage, regardless of whether I measured the growth using fluorescence or light scattering (Figure S2). Changes in cell size could confound optical density and fluorescence readings, but this phenomenon leaves a signature [53] in growth data that is absent in my

dataset (Figures S2, S3). The change in carrying capacity K can be linked to an increase in biomass yield (y) as Monod's expression [54] $y = K/S$ suggests, where S the supply of glucose. This metric is consistent with the data, given that construct GB(c) has to express and translate plasmid-borne genes with the same supply of glucose (Figures 3B and C). This means that pGW155B triggers a metabolic trade-off between growth rate and biomass yield redolent of rK-selection theory [40, 55].

For the competition assay, I mixed equal proportions (cell/cell) of two pure cultures, from construct GB(c)—harbouring pGW155B—and Wyl respectively, grown overnight, in media containing no antibiotic or 0.04 µg/mL of tetracycline (see Methods). I incubated the mixed culture at 30°C until both constructs reached their carrying capacity (K, ~ 24h as per data in Figure 3A), and then I propagated the competition into a new plate containing fresh media. I repeated the propagation step seven times totalling between 77 (GB(c)) and 119 (Wyl) generations given the reproduction rates in Figure 3B, where the generation time per hour is given by $m_i/24$, multiplied times the number of propagation steps. Considering the mutation rate of enterobacteria [56] and the emergence of mutants after 3 growth cycles [57], I constrained the fitness measurements in both constructs over the first 24h of competition and compared the outcome of the competition with predictions in Figure 2. Indeed, the relative fitness of both constructs changed through time consistently with Figure 1B and C. In the presence of tetracycline, the relative fitness of drug-sensitive construct Wyl was below GB(c), which harbours pGW155B, at all times as Figure 4A illustrates. This meant that GB(c) increased its relative frequency in following propagation steps and became the most abundant construct throughout the competition (Figure 4B) consistently with the theoretical predictions. Now,

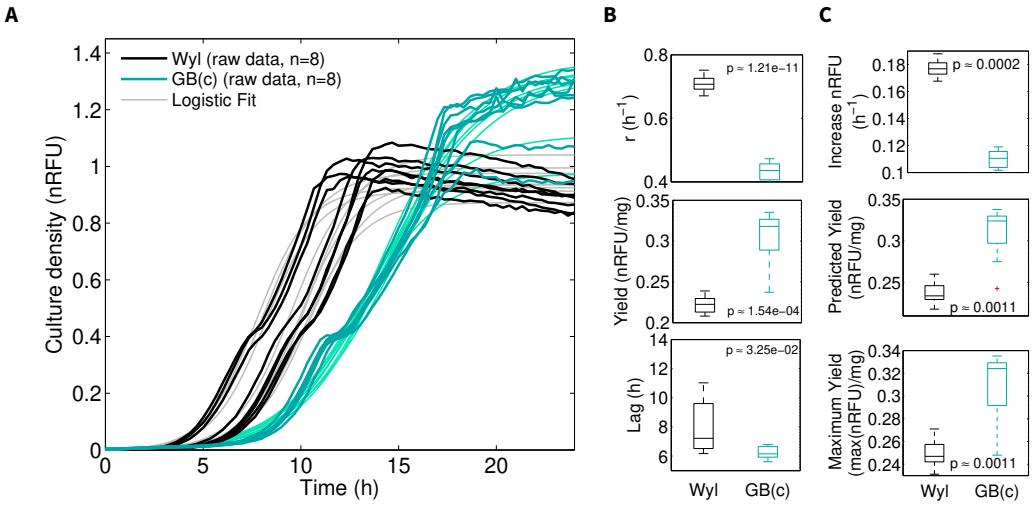

Figure 3. Asymmetric carriage costs of pGW155B. A) Overlapped growth curves of constructs Wyl (black) and GB(c) (cyan) as pure cultures, in the absence of tetracycline, over 24h. I estimated the *per capita* growth rate (m), population size in the equilibrium (K), biomass yield, and lag duration from a 4-parameter logistic model fitted to growth data (see Methods). Data fits for constructs Wyl and GB(c) are shown in grey and light cyan, respectively. **B)** Box plots showing the median (centre of the box), 25th and 75th percentile for each of the aforementioned parameters. The whiskers extend to the most extreme data points that are not outliers, and these are individually represented. The p value shown refers to a Two-sample t-test with unequal variance (8 replicates) that I used to test differences in the parameters between both constructs. **C)** Alternative metrics for growth rate and biomass yield: forward Euler approximation (top), data fit predicted yield (middle), and maximal yield across all time points in growth data (bottom). The p values correspond to Two-sample t-tests with unequal variance (8 replicates).

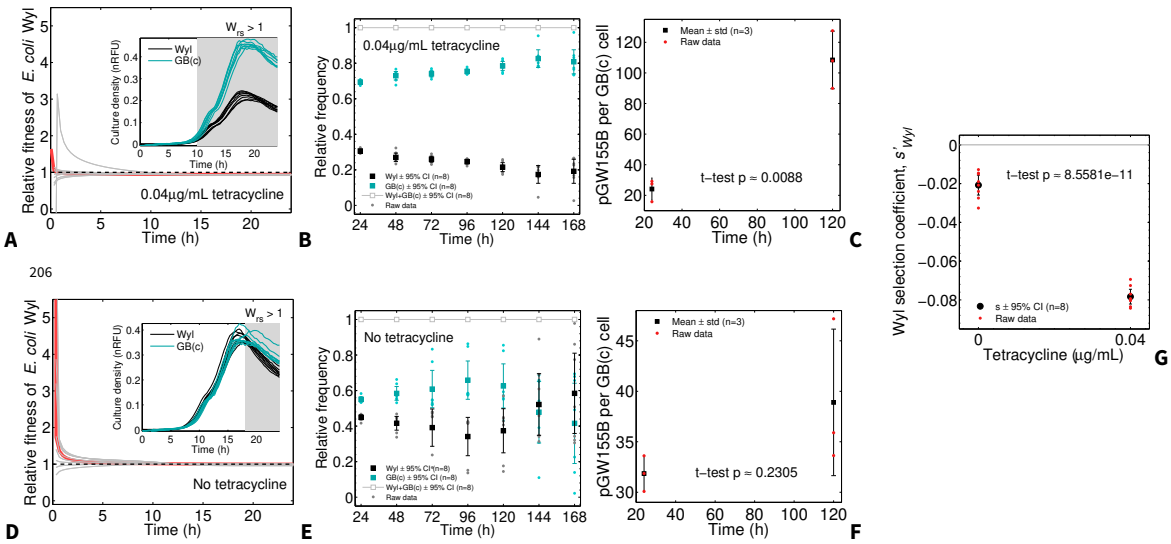

Figure 4. Time-dependent fitness data allows co-maintenance of both constructs without tetracycline. Relative fitness during first 24 hour pairwise-competition of construct Wyl, where both constructs grew in media supplemented with 0.04 µg/mL of tetracycline (**A**), and without antibiotic (**D**). Fitness for each replicate, derived from the logistic fit described in the Methods section as used in Figure 3A, is presented in grey with respect to the reference GB(c) (black dashed line), and the prediction from robustly fitting equation 3 to data in red with an adjusted R^2 of 0.655 (A, F-statistic against zero model = $1.14 \cdot 10^6, p \ll 10^{-100}$) and 0.902 (D, F-statistic against zero model = $1.42 \cdot 10^6, p \ll 10^{-100}$). The inset illustrate the growth curves for both constructs, highlighting in grey the period where the relative fitness of Wyl (W_{sr}) is lower than its tetracycline-resistant competitor's. Note that the frequency of construct Wyl remains above zero at all times. **B** and **E** show the change in relative frequency of both constructs through time when both constructs grew in media supplemented with (B) and without tetracycline (E). Error bars represent the mean ± 95% confidence intervals, with raw data points shown as red dots. **C** and **F** show the change in relative copy numbers of pGW155B borne by construct GB(c) during the pairwise-competition, calculated using quantitative PCR (see Methods). Error bars represent the mean ± standard deviation, with raw data shown in red. **G** illustrates the selection coefficient for construct GB(c) during the experiment (mean ± 95% confidence intervals, with raw data as red dots), a common selection metrics calculated as described in [9]. The p-value shown in **C**, **F**, and **G** corresponds to a Welch's t-test—two sample t-test assuming that samples come from a normal distribution with unknown means, and unknown and unequal variances.

the model does not implement any mutation processes but these occur during this assay. For example, the number of copies of pGW155B harboured by GB(c) increased 5- to 6-fold during the competition (2-sample t-test with unequal variances, p=0.0088, df=2.5826, t-statistic=-7.2441; with 3 replicates) driven by exposure to tetracycline (Figure 4C) as it would be reasonable to expect given the exposure to tetracycline. Despite the occurrence of these mutations, Wyl never goes extinct consistently with Figure 2 and its proportion is determine by its abundance at the propagation time.

Similarly, when both constructs grew in mixed cultures, now without tetracycline exposure, the relative fitness of Wyl was higher, equal, or lower than plasmid-harbouring construct GB(c) through time, as Figure 4D shows. Consistently with the predictions both constructs where co-maintained, and with similar relative frequencies, despite the difference in growth rate (Figure 4E). Note here that while the initial deviation is small, without strong selection by antibiotic exposure each replicate evolves differently, with some replicates of Wyl being lost, while others are co-maintained. In other words, GB(c) preserved pGW155B despite the lack of tetracycline and the difference in growth rates (2-sample t-test with unequal variances, p=0.2305, df=2.2351, t-statistic=-1.6353; with 3 replicates. Figure 4F). This con-

trasts with the negative selection coefficient [2, 9] for construct Wyl (Figure 4G), for example, calculated as its relative fitness with respect to GB(c) measured after 24h of competition minus 1, as introduced above. This common metric suggests that Wyl is being selected against, but data show this construct is indeed being co-maintained—albeit at a lower frequency with respect to GB(c) consistently with predictions in Figure 2.

III. DISCUSSION

This notion of fitness applies within a very specific set of assumptions [8]: Haploid genotypes with continuous growth and overlapping generations. Diploid genotypes or those with more complex reproductive cycles require different assumptions, and yet, due to the simplicity of the method, a similar metric—if not the same—is also used to measure fitness from yeast to corals [i.e. 20, 21, 58]. Furthermore, it is noteworthy to note that fitness, taken as the aforementioned Malthusian parameter, indeed varies through evolutionary timescales due to mutations [9, 59]. The scope of this manuscript is the variations in fitness that occur at smaller timescales, where fitness is assumed constant and often used to make evolutionary predictions [19, 22–24, 60, 61]. Note that both genotypes have glucose as sole carbon source and the growth conditions are homogeneous to avoid the formation of different niches that can be exploited for coexistence [46].

Fitness, as measured in the laboratory, is known to be sensitive to t^* [62]. But is not used to test models that are widely used in population genetics. Rather, it is presented as an experimental quirk: If t^* is too long one competitor may drive the other to extinction, whereas if t^* is too short the measurement may not be precise [62]. This means fitness is not a scientific metric. Now, one could argue that variations in fitness measurements are caused by frequency-dependent selection [46], albeit it also depends on t^* without explanation. It has been suggested [25, 26] that each growth phase contributes to fitness, and yet, population growth rate is the main parameter in the models presented [25, 26] where a contribution of zero from growth rate translates into zero fitness. Equation 3 provides a more robust metric, given the correction of m by changes in population size, whereas the ratio of growth rates m_i/m_j would be more appropriate with chemostats [63] where the assumptions of exponential growth are closely met. The result is that fitness estimates, and selection coefficients that depend of such estimates, can be highly misleading. See how, according to Figure 4G, construct Wyl is selected against according to its selection coefficient. However, data show it is being co-maintained. Is this phenomenon behind the plasmid paradox?

If the fitness measured experimentally varies within a season or cycle, it means cells harbouring plasmids—paying a cost for their maintenance—are mistakenly assumed to go extinct without positive selection. On the contrary, what I suggest here is that if all genotypes reach their equilibrium, before the selection event, then plasmids can be maintained. Even in the absence of mutations or differences in carrying capacities, explaining maintenance of pGW155B without antibiotic on the one hand, and the persistence of the drug-sensitive construct Wyl when exposed tetracycline shown in Figure 4B. Indeed, the use of more tetracycline would have cleared Wyl. The point is that, despite the substantial difference in fitness, the population of Wyl levelled off and was co-maintained consistently with the predictions in Figure 2. Only emerging mutants that are unable to reach their equilibrium before the se-

lection step, would be lost to selection—which data from Figure 4E suggest this might have occurred in one of the replicates. The lack of mutations assumed here is, naturally, unrealistic. But it sets the expectation that a mutant needs not go extinct because of lower growth rate—as is currently thought—and, rather, can co-maintained if it can reach an equilibrium.

Another interesting observation was the dynamism in numbers of plasmids borne by construct GB(c) when exposed to tetracycline, as Figure 4C and F illustrate. While the number of plasmids borne did not change significantly in the absence of tetracycline, exposing GB(c) to the drug forced a 5-fold increase in the number of plasmids per cell. Not only this phenomenon could help create and maintain reservoirs of plasmids within populations. Perhaps more importantly, it illustrates that cells testing positive for a plasmid-borne gene, as it is the case in the clinic with antibiotic-resistant genes with molecular or *in silico* techniques [64], may differ in resistance levels merely by differences in plasmid copy number—suggesting that plasmid copy number should also be monitored.

In conclusion, fitness estimates decline exponentially during the growth of a population, until the competing genotypes attain their carrying capacities. This means that selection events that occur before fitness is steady likely fall within the classical notion of fitness from population genetics theory, and it would therefore be reasonable to expect the fixation or extinction of one of the genotypes. Similarly, if selection events occur *after* fitness is steady then it is likely that both competing genotypes are co-maintained. With knowledge of their intrinsic growth rate (m) and carrying capacities (K^*), equation 3 can help produce more accurate fitness estimates. My study suggests that, given the misleading predictions of conventional fitness estimates and the pervasive use of growth rate in the laboratory as a proxy for fitness, selection for plasmids could be higher than previously thought.

IV. METHODS

Media and Strains. I used the strains of *Escherichia coli* GB(c) and Wyl [65] (a gift from Remy Chait and Roy Kishony), and M9 minimal media supplemented with 0.4% glucose and 0.1% casamino acids (w/v). Note the construct Wyl harbours pCS-λ [66]. I made tetracycline stock solutions from powder stock (Duchefa, Ref. # 0150.0025) at 5mg/mL in 50% ethanol, filter sterilised, and stored at -20^oC. Subsequent dilutions were made from this stock in supplemented M9 minimal media and kept at 4^oC.

Batch transfer protocol. I inoculated a 96-well microtitre plate containing 150 µg/mL of supplemented M9 media with a mixture of two overnight cultures, one of *E. coli* GB(c) and another of *E. coli* Wyl (1µL containing approx. 2×10^6 cells, Figure S4). To avoid unforeseen variation in the initial number of copies of pGW155B between assays, the overnight culture for GB(c) was supplemented with 100ng/mL of tetracycline as described elsewhere [16]. Then, I centrifuged the cultures, and resuspended the pelleted cells in M9 minimal media prior adding to the microtitre plate to avoid the introduction of residual antibiotic and metabolic byproducts. I incubated the plate at 30^oC in a commercial spectrophotometer and measured the optical density of each well at 600nm (OD_{600}), yellow florescence for the Wyl strain (YFP excitation at 505nm, emission at 540nm), and cyan fluorescence for the GB(c) strain (CFP at 430nm/480nm) every 20min for 24h. After each day I transferred 1.5µL of

each well, using a 96-well pin replicator, into a new microtitre plate containing fresh growth medium and tetracycline.

The use of by-hand laboratory techniques, such as colony counting, were impractical given the continuous measurements required for the experiments in this study and are also error-prone [67]. Thus, given the existing correlation between cell number, optical density, and relative fluorescence measurements (Supplementary Figures S4 and S5); I used a spectrophotometer to automatically produce proxies of population density. While these devices have inherent limitations, Supplementary Figure S4 shows that OD_{600} and fluorescence readings are reasonable indicators of bacterial population density: Readings correlate positively, and indeed linearly, with live cell counts measured in colony-forming units (CFU) per ml.

Growth parameter estimation. Yellow and cyan fluorescence protein genes were constitutively expressed given the approximately constant ratio between fluorescence and optical density (Figure S5) in the conditions tested in the article. This allowed me to use fluorescence data as a proxy for cell density in mixed culture conditions. I normalised fluorescence readings using a conversion factor, n_f, calculated by growing each construct in a range of glucose concentrations, and regressing the linear model $\mathrm{RFU} = n_\mathrm{f} \cdot \mathrm{OD} + c$, where RFU is relative fluorescence units data, OD optical density data, n_f the conversion factor between fluorescence and optical density, and c the crossing point with the y–axis when $\mathrm{OD} = 0$. I imported the resulting time series data set (Figures S3) into `MATLAB R2014b` to subtract background and calculate fitness as described in the main text as follows.

Once corrected, I used the expression $m_i = \frac{ln[N_i(t)/N_i(0)]}{t}$ to estimate the *per capita* growth rate m_i as per in [9, 19, 22–24] although I used different values for t to test the theoretical predictions. This expression derives from the exponential model in equation 1 as it follows. We can isolate the term with m for, say, mutant A, as $A(t) = A_0 e^{m_a t} \rightarrow \frac{A(t)}{A_0} = e^{m_a t}$. To remove the number e we then apply logarithms to this expression leading to $ln\frac{A(t)}{A_0} = m_a t$. Re-arranging the terms we are left with $\frac{ln[A(t)/A_0]}{t} = m_a$, which is the expression used by experimentalists.

To measure the carriage costs of pGW155B, I fitted to data the logistic model

$$g(t) = \frac{b + K}{(1 + C \cdot e^{-m \cdot [t-l]})}.$$

where the term $g(t)$ denotes the change in culture growth as a function of time (in OD, YFP, or CFP units), b the inoculum size used to subtract the background, C is a parameter, m the *per capita* growth rate, K the maximal population size attained and l the duration of the lag phase. Alternatively, I measured the rate of increase using the forward Euler approximation $m_i = max_{0 \leq t \leq 24h}\left(\frac{N_i(t)/N_i(t-1)}{\Delta t}\right)$, where $\Delta t = 20\mathrm{min} = 1/3$ is the time between reads. To estimate the biomass yield I divided OD data in stationary phase by the glucose supplied [68]. Alternatively, I used the highest density across all time points and the density reported by the data fit, both divided by glucose supply, as alternative metrics for biomass yield.

Finally, I divided the length of the growth cycle (24h) by the *per capita* growth rate, multiplied by the number of propagations (seven) to estimate the total number of duplication events—where one duplication means one generation—for each construct.

DNA material extraction. For each concentration, I sampled three representative 150 µg/mL replicates that I divided in two groups: for chromosome and plasmid DNA extraction. I used 'GeneJet DNA' (ThermoScientific, Ref. #K0729) and 'GeneJet Plasmid' (ThermoScientific, Ref. #K0502) extraction kits to extract chromosomal and plasmid DNA (pDNA) from the samples, respectively, and used Qubit to quantify DNA and pDNA yields. Both extracts were diluted accordingly in extraction buffer, as per in manufacturer instructions, to normalise DNA across samples.

Quantitative PCR and plasmid copy number estimation. I used `primer3` to design two pairs of primers with melting temperature (T_m) of 60°C and non-overlapping probes with T_m of 70°C. The amplicon ranges between 100 to 141bp depending on the locus (Table 1). Two reaction mixes were prepared using the kit 'Luminaris Color Probe Low ROX' (ThermoScientific, Ref. #K0342), adding 0.3µM of each primer and 0.2µM of the probe as per manufacturer specifications. Following a calibration curve for each reaction (Figure S6) I added 0.01ng of chromosomal or plasmid DNA material to each of the reaction mixes.

Table 1. Primers and probes designed using `Primer3`. Amplicon ranging from 100 to 141bp. T_m indicates the estimated melting temperature.

Target gen	Sequence (5' → 3')	T_m (^{o}C)	Feature
tatB	CGATGAAGCGTTCCTACGTT	60.27	Forward
	TCATGCGCAGCTTCATTATC	59.94	Reverse
	AAGGCGAGCGATGAAGCGCA	70.70	Probe
tet(36)	ATTGGGCATCTATTGGCTTG	59.22	Forward
	CCGATTCACAGGCTTTCTTG	60.76	Reverse
	AGCCTTTGCCAATTGGGGCG	70.37	Probe

To estimate the relative copies of pGW155B per GB(c) cell, I calculated the corresponding proportion of chromosomal DNA corresponding to the GB(c)-type from data in Figure 4 and used the expression [23]

$$cn = \frac{(1+E_c)^{Ctc}}{(1+E_p)^{Ctp}} \times \frac{S_c}{S_p},$$

where cn is the number of plasmid copies per chromosome, S_c and S_p are the size of the chromosome and pGW155B amplicon in bp, E_c and E_p the efficiency of the qPCR taken from data in Figure S6, and Ctc and Ctp are the cycles at which I first detected product amplification (C_t).

Funding. The author received no financial support for the research, authorship, and/or publication of this article.

Availability of data and materials. Data used to generate Figures 3 and 4 is available as a supplementary table. An Octave/MATLAB implementation of equations 1 and 2 is available at
`https://gitlab.com/rc-reding/papers/-/tree/master/FisherTheorem`.

Competing interests. The author declares no competing interests.

REFERENCES

1. Chevin, L.-M. On measuring selection in experimental evolution. *Biol. Lett.* **7,** 210–213 (2011).
2. Crow, J. F., Kimura, M., *et al. An introduction to population genetics theory.* (New York, Evanston and London: Harper & Row, Publishers, 1970).
3. Nagylaki, T. in *Introduction to Theoretical Population Genetics* 5–27 (Springer, 1992).
4. Rice, S. H. *Evolutionary theory: mathematical and conceptual foundations* ISBN: 9780878937028 (Oxford University Press, 2004).
5. Templeton, A. *Population Genetics and Microevolutionary Theory* ISBN: 9780471409519 (Wiley, 2006).
6. Dawes, I. & Sutherland, I. *Microbial Physiology* ISBN: 9780632024636 (Wiley, 1992).
7. Borriello, S., Murray, P. & Funke, G. *Topley and Wilson's Microbiology and Microbial Infections, 2 Volume Set: Bacteriology* ISBN: 9780470660263 (Wiley, 2007).
8. Fisher, R. A. *The Genetical Theory of Natural Selection* 308–315 (At The Clarendon Press, 1930).
9. Lenski, R. E., Rose, M. R., Simpson, S. C. & Tadler, S. C. Long-Term Experimental Evolution in Escherichia coli. I. Adaptation and Divergence During 2,000 Generations. *Am. Nat.* **138,** pp. 1315–1341 (1991).
10. Orr, H. A. Fitness and its role in evolutionary genetics. *Nat. Rev. Genet.* **10,** 531–539 (2009).
11. Wiser, M. J. & Lenski, R. E. A comparison of methods to measure fitness in Escherichia coli. *PloS One* **10,** e0126210 (2015).
12. Concepción-Acevedo, J., Weiss, H. N., Chaudhry, W. N. & Levin, B. R. Malthusian parameters as estimators of the fitness of microbes: a cautionary tale about the low side of high throughput. *PLoS One* **10,** e0126915 (2015).
13. Kokko, H. & López-Sepulcre, A. The ecogenetic link between demography and evolution: can we bridge the gap between theory and data? *Ecol. Lett.* **10,** 773–782 (2007).
14. Orr, H. A. The population genetics of beneficial mutations. *Philos. Trans. R. Soc. Lond., B, Biol. Sci.* **365,** 1195–1201 (2010).
15. Vasi, F., Travisano, M. & Lenski, R. E. Long-term experimental evolution in Escherichia coli. II. Changes in life-history traits during adaptation to a seasonal environment. *Am. Nat.* **144,** 432–456 (1994).
16. Chait, R., Craney, A & Kishony, R. Antibiotic interactions that select against resistance. *Nature* **446,** 668–671 (2007).
17. Wang, H., Avican, K., Fahlgren, A., *et al.* Increased plasmid copy number is essential for Yersinia T3SS function and virulence. *Science* **353,** 492–495 (2016).
18. Adkar, B. V., Manhart, M., Bhattacharyya, S., Tian, J., Musharbash, M. & Shakhnovich, E. I. Optimization of lag phase shapes the evolution of a bacterial enzyme. *Nat. Ecol. Evol.* **1,** 1–6 (2017).

19. Andersson, D. I. & Levin, B. R. The biological cost of antibiotic resistance. *Curr. Opin. Microbiol.* **2,** 489–493 (1999).

20. Steinmetz, L. M., Scharfe, C., Deutschbauer, A. M., *et al.* Systematic screen for human disease genes in yeast. *Nat. Genet.* **31,** 400–404 (2002).

21. Hall, D. W. & Joseph, S. B. A high frequency of beneficial mutations across multiple fitness components in Saccharomyces cerevisiae. *Genetics* **185,** 1397–1409 (2010).

22. Gullberg, E., Cao, S., Berg, O. G., *et al.* Selection of resistant bacteria at very low antibiotic concentrations. *PLoS Pathog.* **7,** e1002158 (2011).

23. Millán, A. S., Peña-Miller, R., Toll-Riera, M, *et al.* Positive selection and compensatory adaptation interact to stabilize non-transmissible plasmids. *Nat. Commun.* **5** (2014).

24. Ram, Y., Dellus-Gur, E., Bibi, M., *et al.* Predicting microbial growth in a mixed culture from growth curve data. *Proc. Natl. Acad. Sci. U.S.A.* **116,** 14698–14707 (2019).

25. Li, Y., Venkataram, S., Agarwala, A., *et al.* Hidden complexity of yeast adaptation under simple evolutionary conditions. *Current Biology* **28,** 515–525 (2018).

26. Lin, J., Manhart, M. & Amir, A. Evolution of microbial growth traits under serial dilution. *Genetics* **215,** 767–777 (2020).

27. Kimura, M. On the change of population fitness by natural selection. *Heredity* **12,** 145–167 (1958).

28. Harrison, E. & Brockhurst, M. A. Plasmid-mediated horizontal gene transfer is a coevolutionary process. *Trends Microbiol.* **20,** 262–267 (2012).

29. MacLean, R. C. & San Millan, A. Microbial evolution: towards resolving the plasmid paradox. *Curr. Biol.* **25,** R764–R767 (2015).

30. Probing the plasmid paradox. *Nat. Ecol. Evol.* **5,** 1559–1559 (2021).

31. Brockhurst, M. A. & Harrison, E. Ecological and evolutionary solutions to the plasmid paradox. *Trends Microbiol.* **30,** 534–543 (2022).

32. Lili, L. N., Britton, N. F. & Feil, E. J. The persistence of parasitic plasmids. *Genetics* **177,** 399–405 (2007).

33. Vogwill, T. & MacLean, R. C. The genetic basis of the fitness costs of antimicrobial resistance: a meta-analysis approach. *Evol. Appl.* **8,** 284–295 (2015).

34. Levin, B. R., Lipsitch, M, Perrot, V, *et al.* The population genetics of antibiotic resistance. *Clin. Infect. Dis.* **24 Suppl 1,** S9–16 (1997).

35. Tepekule, B., Abel zur Wiesch, P., Kouyos, R. D. & Bonhoeffer, S. Quantifying the impact of treatment history on plasmid-mediated resistance evolution in human gut microbiota. *Proc. Natl. Acad. Sci. U.S.A.* **116,** 23106–23116 (2019).

36. Smillie, C., Garcillán-Barcia, M. P., Francia, M. V., Rocha, E. P. & de la Cruz, F. Mobility of plasmids. *Microbiol. Mol. Biol. Rev.* **74,** 434–452 (2010).

37. Lopatkin, A. J., Meredith, H. R., Srimani, J. K., Pfeiffer, C., Durrett, R. & You, L. Persistence and reversal of plasmid-mediated antibiotic resistance. *Nat. Commun.* **8,** 1689 (2017).

38. Wakeley, J. The limits of theoretical population genetics. *Genetics* **169,** 1–7 (2005).

39. Hansen, S. K., Rainey, P. B., Haagensen, J. A. & Molin, S. Evolution of species interactions in a biofilm community. *Nature* **445,** 533–536 (2007).

40. Reding-Roman, C., Hewlett, M., Duxbury, S., Gori, F., Gudelj, I. & Beardmore, R. The unconstrained evolution of fast and efficient antibiotic-resistant bacterial genomes. *Nat. Ecol. Evol.* **1,** 0050 (2017).

41. Gilpin, M. E. & Ayala, F. J. Global models of growth and competition. *Proc. Natl. Acad. Sci. U.S.A.* **70,** 3590–3593 (1973).

42. Lenski, R. E. Experimental studies of pleiotropy and epistasis in Escherichia coli. I. Variation in competitive fitness among mutants resistant to virus T4. *Evolution* **42,** 425–432 (1988).

43. Wiser, M. J., Ribeck, N. & Lenski, R. E. Long-term dynamics of adaptation in asexual populations. *Science* **342,** 1364–1367 (2013).

44. Lenski, R. E., Wiser, M. J., Ribeck, N., *et al.* Sustained fitness gains and variability in fitness trajectories in the long-term evolution experiment with Escherichia coli. *Proc. R. Soc. B.* **282,** 20152292 (2015).

45. Hinz, A., Amado, A., Kassen, R., Bank, C. & Wong, A. Unpredictability of the fitness effects of antimicrobial resistance mutations across environments in Escherichia coli. *bioRxiv.* https://www.biorxiv.org/content/early/2023/10/27/2023.10.25.563906 (2023).

46. Ribeck, N. & Lenski, R. E. Modeling and quantifying frequency-dependent fitness in microbial populations with cross-feeding interactions. *Evolution* **69,** 1313–1320 (2015).

47. Whittle, G, Whitehead, T. R., Hamburger, N, Shoemaker, N. B., Cotta, M. A., *et al.* Identification of a new ribosomal protection type of tetracycline resistance gene, tet(36), from swine manure pits. *Appl. Environ. Microbiol.* **64,** 4151–4158 (2003).

48. Del Solar, G. & Espinosa, M. Plasmid copy number control: an ever-growing story. *Mol. Microbiol* **37,** 492–500 (2000).

49. Garcillán-Barcia, M. P., Alvarado, A. & de la Cruz, F. Identification of bacterial plasmids based on mobility and plasmid population biology. *FEMS Microbiol. Rev.* **35,** 936–956 (2011).

50. Million-Weaver, S. & Camps, M. Mechanisms of plasmid segregation: have multicopy plasmids been overlooked? *Plasmid* **75,** 27–36 (2014).

51. Guynet, C. & de la Cruz, F. Plasmid segregation without partition. *Mob. Genet. Elements* **1,** 236–241 (2011).

52. Dobson, A., Cotter, P. D., Ross, R. P. & Hill, C. Bacteriocin Production: a Probiotic Trait? *Appl. Environ. Microbiol.* **78,** 1–6 (2012).

53. Stevenson, K., McVey, A. F., Clark, I. B., Swain, P. S. & Pilizota, T. General calibration of microbial growth in microplate readers. *Sci. Rep.* **6,** 1–7 (2016).

54. Monod, J. *Recherches sur la croissance des cultures bactériennes* PhD thesis (Paris, 1941).

55. Reznick, D., Bryant, M. J. & Bashey, F. r-and K-selection revisited: the role of population regulation in life-history evolution. *Ecology* **83,** 1509–1520 (2002).

56. Perfeito, L., Fernandes, L., Mota, C. & Gordo, I. Adaptive mutations in bacteria: high rate and small effects. *Science* **317,** 813–815 (2007).

57. Reding, C., Catalán, P., Jansen, G., *et al.* The Antibiotic Dosage of Fastest Resistance Evolution: gene amplifications underpinning the inverted-U. *Mol. Biol. Evol.* **9,** 3847–3863. ISSN: 0737-4038 (Mar. 2021).

58. Edmunds, P. J. Intraspecific variation in growth rate is a poor predictor of fitness for reef corals. *Ecology* **98,** 2191–2200 (2017).

59. Cao, H. & Plague, G. R. The fitness effects of a point mutation ins *Escherichia coli* change with founding population density. *Genetica* **144,** 417–424 (2016).

60. Bergstrom, C. T., Lipsitch, M. & Levin, B. R. Natural selection, infectious transfer and the existence conditions for bacterial plasmids. *Genetics* **155,** 1505–1519 (2000).

61. Blanquart, F., Achaz, G., Bataillon, T. & Tenaillon, O. Properties of selected mutations and genotypic landscapes under Fisher's geometric model. *Evolution* **68,** 3537–3554 (2014).

62. McDonald, M. J. Microbial experimental evolution–a proving ground for evolutionary theory and a tool for discovery. *EMBO Rep.* **20,** e46992 (2019).

63. Moser, H. *The dynamics of bacterial populations maintained in the chemostat.* (Carnegie institution. Washington, D.C., 1958).

64. Carattoli, A., Zankari, E., García-Fernández, A., *et al.* In silico detection and typing of plasmids using PlasmidFinder and plasmid multilocus sequence typing. *Antimicrob. Agents Chemother.* **58,** 3895–3903 (2014).

65. Chait, R., Shrestha, S., Shah, A. K., Michel, J.-B. & Kishony, R. A Differential Drug Screen for Compounds That Select Against Antibiotic Resistance. *PLoS One* **5,** e15179 (2010).

66. Kishony, R. & Leibler, S. Environmental stresses can alleviate the average deleterious effect of mutations. *J. Biol.* **2,** 1–10 (2003).

67. Brugger, S. D., Baumberger, C., Jost, M., Jenni, W., Brugger, U. & Mühlemann, K. Automated counting of bacterial colony forming units on agar plates. *PLoS One* **7,** e33695 (2012).

68. Monod, J. The Growth of Bacterial Cultures. *Annu. Rev. Microbiol.* **3,** 371–394 (1949).

I. SUPPLEMENTARY FIGURES

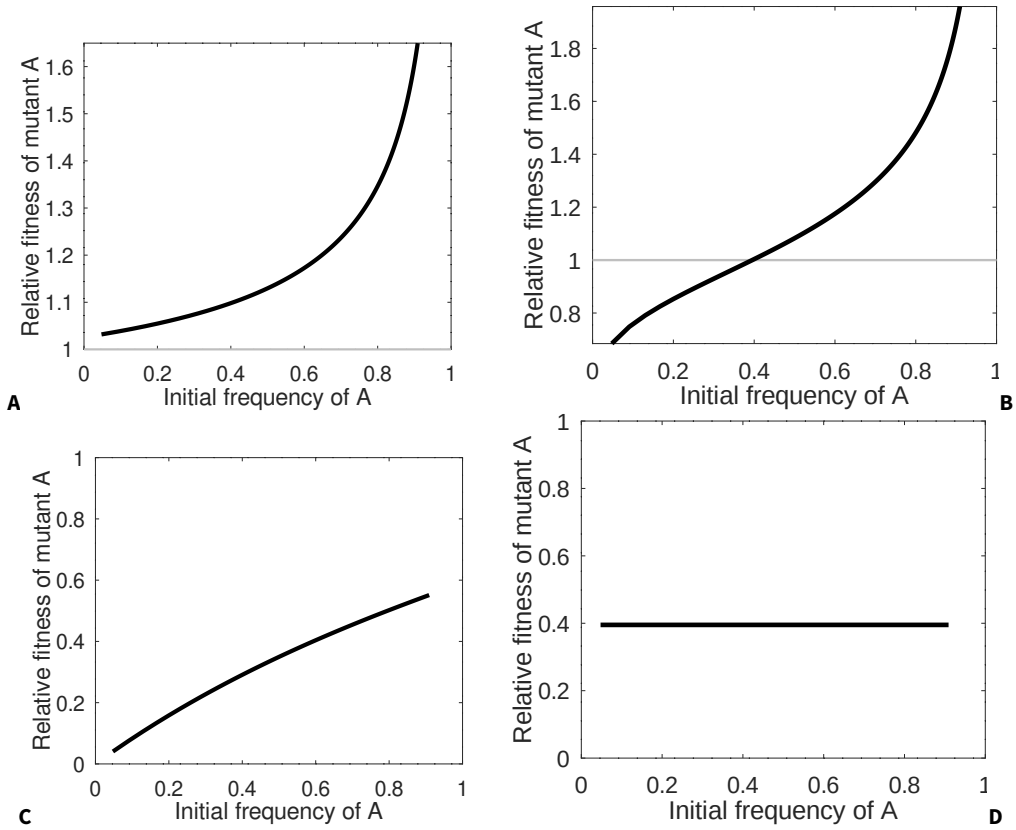

Figure S1. Relative fitness of mutant A at different initial frequencies. A–B) Relative fitness of mutant A measured after seven propagation steps, analog to those described in the main manuscript, when growth rate and population size engage in a trade-off (A) and when they do not (B). The frequency on the x-axis is that of mutant A at t=0. The fitness of each competitor (m_i) was calculated as explained in the main manuscript as $m_i = \frac{ln[N_i(t)/N_i(0)]}{\Delta t}$ where $N_i(t)$ is the density of each mutant at time $t = 7, N_i(0)$ that at time t=0, and $\Delta t = 7$. This was used to calculate the relative fitness of A as $W_{ab} = m_a/m_b$ as described in the text. The code is available in github in the link provided in the text, and the parameters used are those from Figure 1B and C. **C–D)** Relative frequency of mutant A at the end of the seventh propagation when growth rate and population size engage in a trade-off (C) and when they do not (D), shown in the y-axis, as a function of its frequency at t=0.

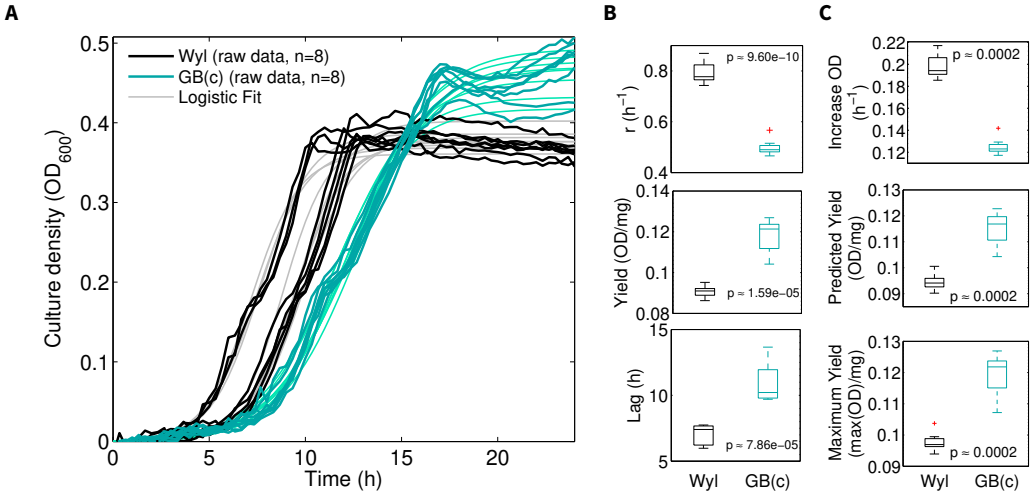

Figure S2. Asymmetric carriage costs of pGW155B (optical density data). A) Growth curves of constructs Wyl (black) and GB(c) (cyan) in pure culture, in the absence of tetracycline, over 24h. I estimated the maximum growth rate (r), population size in the equilibrium (K), biomass yield, and lag duration from a 4-parameter logistic model fitted to growth data (see Methods). Data fits for constructs Wyl and GB(c) are shown in grey and light cyan, respectively. **B)** Box plots showing the median (centre of the box), 25th and 75th percentile for each of the aforementioned parameters. The whiskers extend to the most extreme data points that are not outliers, and these are individually represented. The p value shown refers to Two-sample t-tests with unequal variance (8 replicates) that I used to test differences in the parameters between both constructs. **C)** Alternative metrics for growth rate and biomass yield: forward Euler approximation (top), data fit predicted yield (middle), and maximal yield across all time points (bottom). The p values correspond to Two-sample t-tests with unequal variance (8 replicates).

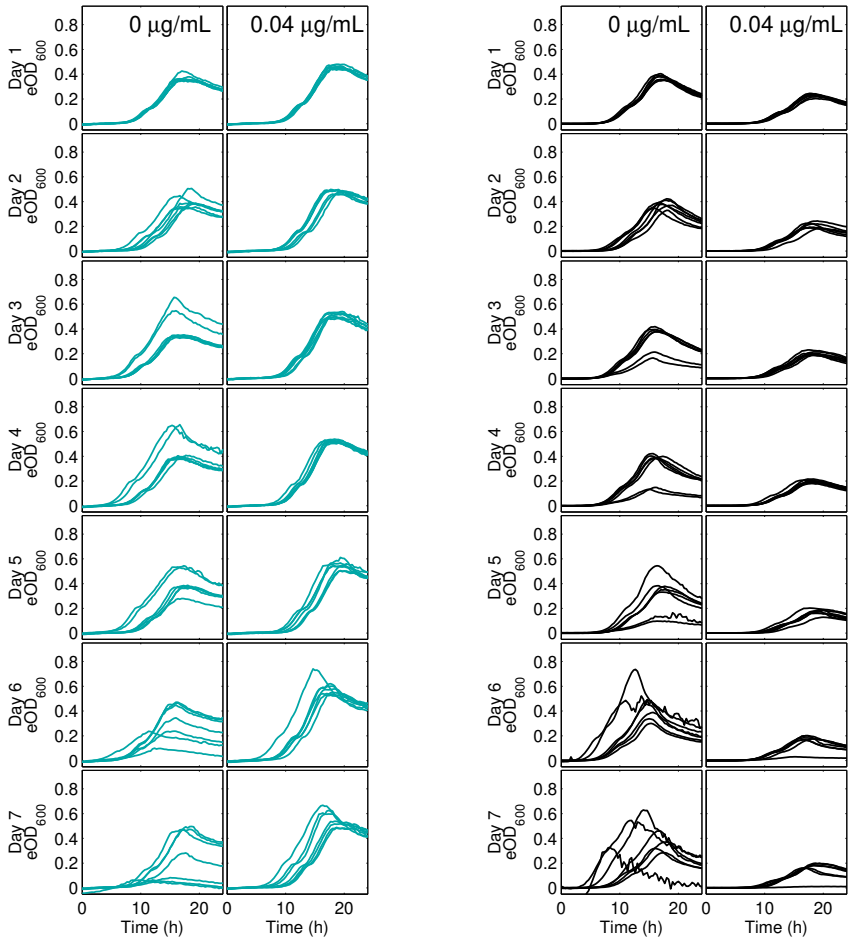

Figure S3. Raw data of each construct during the competition. Raw optical density data for GB(c) (cyan) and Wyl (black), measured every 20 min and converted from fluorescence readings. Each individual plot represents the change in abundance of each construct over 24 hours. Data corresponding to each condition, with and without tetracycline, are aggregated as columns.

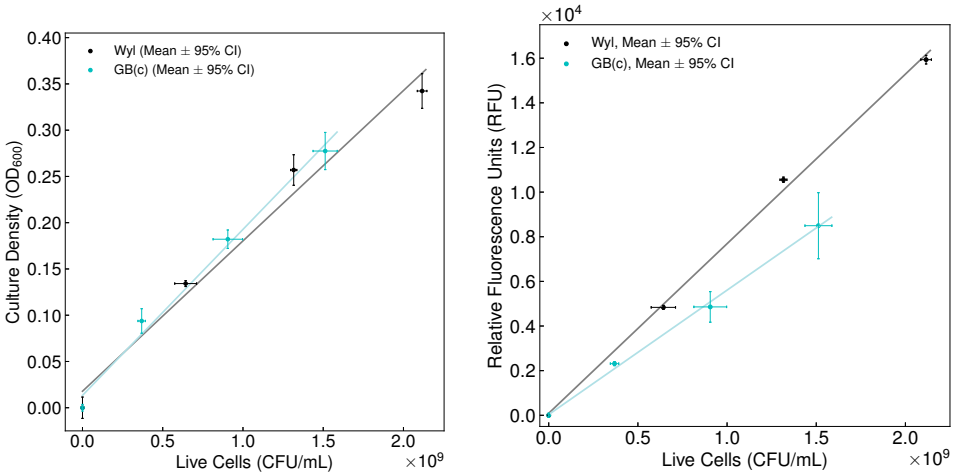

Figure S4. Calibration curve to translate optical density and fluorescence data to number of *Escherichia coli* cells. I fitted the linear model $a = bx + c$ to optical density and colony counting data (dots) to calculate the number of optical density units (OD_{600}, left) and relative fluorescence units (RFU, right) per cell. a denotes the optical density readings measured at 600nm, c the crossing point with the $y-$axis when $x = 0$, and b the conversion factor between optical density and number of cells (x). I interpolated optical density and fluorescence readings to calculate the number of cells within a culture as $x = (a - c)/b$. For the strain Wyl, $b = 1.62 \times 10^{-10}\ OD \cdot mL \cdot CFU^{-1}$ and $c = 1.78 \times 10^{-2}\ OD$, whereas for GB(c) $b = 1.79 \times 10^{-10}\ OD \cdot mL \cdot CFU^{-1}$ and $c = 1.33 \times 10^{-2}\ OD$. For the case of relative fluorescence units, $b = 7.597 \times 10^{-6}\ RFU \cdot mL \cdot CFU^{-1}$ and $c = 90.891\ RFU$ for strain Wyl; whereas $b = 5.573 \times 10^{-6}\ RFU \cdot mL \cdot CFU^{-1}$ and $c = 33.706\ RFU$ for GB(c).

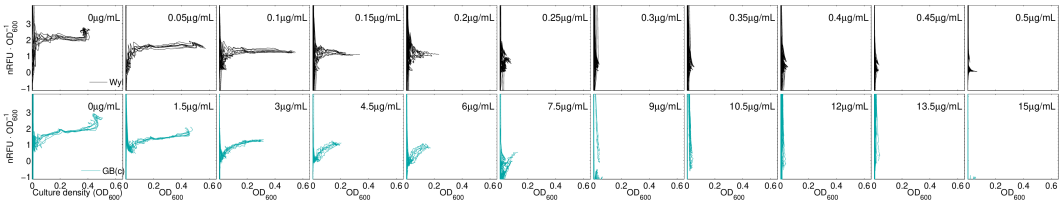

Figure S5. Changes in relative fluorescence over time in GB(c) and Wyl strains in pure culture conditions. Raw change in florescence, per optical density units, measured every 20min for 24h for the Wyl- (black) and GB(c)-type. Each column represents the data set for each tetracycline concentration used.

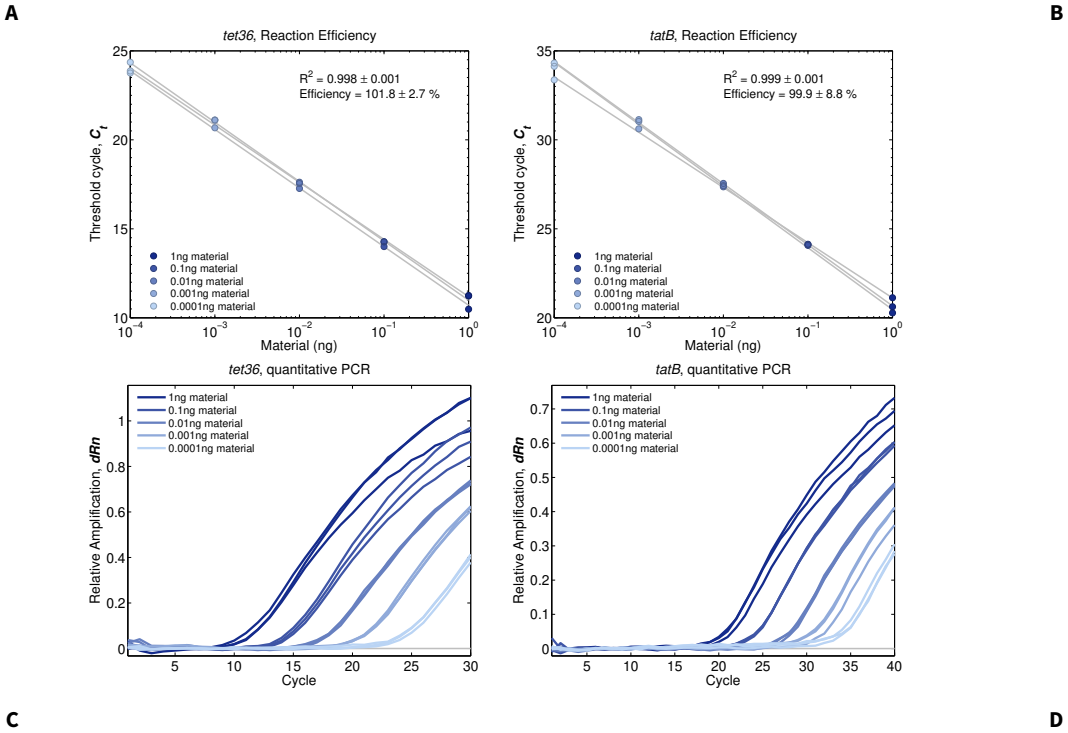

Figure S6. Quantitative PCR calibration curves for *tet(36)* and *tatB*. Reaction efficiency for the set of primers and probes, listed in Table 1 the main Methods section, for *tet(36)* (A) and *tatB*. The efficiency was calculated as $E_f = 10^{-1/Slope} - 1$, and the slope term calculated by fitting a linear model to qPCR threshold cycle (C_t) data. The mean ± standard deviation for the adjusted coefficient of determination R^2 and efficiency are shown in the figures. The amplification curves for each reaction, using 3 replicates, are shown in C) and D), respectively.