Frequency-dependent ecological interactions increase the prevalence and shape the distribution of pre-existing drug resistance

Jeff Maltas[1,2,†], Dagim Shiferaw Tadele[1,3], Arda Durmaz[1], Christopher D. McFarland[2,5], Michael Hinczewski[4], and Jacob G. Scott[1,2,4,5,†]

[1]Cleveland Clinic, Translational Hematology Oncology Research, Cleveland, OH
[2]Case Western Reserve University, School of Medicine, Cleveland, OH
[3]Oslo University Hospital, Ullevål, Department of Medical Genetics, Oslo, Norway
[4]Case Western Reserve University, Department of Physics, Cleveland, OH
[5]Case Comprehensive Cancer Center, Cleveland, OH
[†]jeff.maltas@gmail.com, scottj10@ccf.org

ABSTRACT

The evolution of resistance remains one of the primary challenges for modern medicine from infectious diseases to cancers. Many of these resistance-conferring mutations often carry a substantial fitness cost in the absence of treatment. As a result, we would expect these mutants to undergo purifying selection and be rapidly driven to extinction. Nevertheless, pre-existing resistance is frequently observed from drug-resistant malaria to targeted cancer therapies in non-small cell lung cancer (NSCLC) and melanoma. Solutions to this apparent paradox have taken several forms from spatial rescue to simple mutation supply arguments. Recently, in an evolved resistant NSCLC cell line, we found that frequency-dependent ecological interactions between ancestor and mutant ameliorate the cost of resistance in the absence of treatment. Here, we hypothesize that frequency-dependent ecological interactions in general may play a major role in the prevalence of pre-existing resistance. We combine numerical simulations with robust analytical approximations to provide a rigorous mathematical framework for studying the effects of frequency-dependent ecological interactions on the evolutionary dynamics of pre-existing resistance. First, we find that ecological interactions significantly expand the parameter regime under which we expect to observe pre-existing resistance. Even when positive ecological interactions between mutants and ancestors are rare, these clones provide the primary mode of evolved resistance because their positive interaction leads to significantly longer extinction times. Next, we find that even in the case where mutation supply is sufficient to predict pre-existing resistance, frequency-dependent ecological forces still contribute a strong evolutionary pressure that selects for increasingly positive ecological effects. Finally, we genetically engineer several of the most common clinically observed resistance mechanisms to targeted therapies in NSCLC, a treatment notorious for pre-existing resistance, and where our theory predicts positive ecological interactions to be common. We find that all three engineered mutants display a positive ecological interaction with their ancestor, as predicted. Strikingly, as with our originally evolved resistant mutant, two of the three engineered mutants have ecological interactions that fully compensate for their substantial fitness costs. As a whole, these results suggest that frequency-dependent ecological effects may provide the primary mode by which pre-existing resistance emerges.

Introduction

The rapid, and often inevitable, evolution of therapy resistance is the primary threat to modern medicine's successful treatment of cancer, and infectious disease (e.g. bacterial, viral, fungal, and parasitic infections)[1–5]. The story of resistance and treatment failure is strikingly similar across biological kingdoms. A patient is diagnosed and undergoes an initially successful treatment, only for a small resistant subclone of the original disease to relapse, resulting in treatment failure. For decades, the response to this paradigm has been the development of novel, more efficient drugs, targeting orthogonal pathways in hopes of winning the evolutionary arms race. While this response has undeniably resulted in major success stories when considering individual cancers or infections, the overall outlook for drug-resistant disease remains grim[6–9].

As a result, growing efforts have been made to study these diseases in an *evolutionary* context, whereby scientists seek to understand the ecological and evolutionary forces that inevitably result in the untreatable disease state. Understanding these evolutionary forces that lead to resistance, should allow scientists and physicians to not only design more effective drugs, but perhaps more crucially, design more effect *treatments*. For example, recent work has focused on improving and prolonging the efficacy of our already established drugs via optimal dose scheduling[10–12], drug combinations[13–17], understanding spatial dynamics[18–20], understanding ecological interactions between competing subclones[21–24], and exploiting collateral sensitivity[25–29].

In a similar spirit, this work seeks to understand the evolutionary fates of potential resistance-conferring mutations that emerge *before* treatment has occurred. The fraction of these mutants that survive to see treatment are often the primary cause of treatment failure, referred to as "pre-existing resistance"[30–33]. While these resistant populations provide a large fitness advantage once treatment begins, they often carry a significant fitness disadvantage, or fitness cost (f_c), in the absence of treatment[34–38]. Nevertheless, resistance-conferring mutants often persist until treatment, at which time their treatment-sensitive ancestors are preferentially killed, resulting in the competitive release and relapse of the resistant population and inevitable treatment failure. Understanding how these resistant clones – with a fitness disadvantage – persist in the disease population prior to treatment may allow us to prevent resistance from emerging.

This interest is derived from recent work where we measured the frequency-dependent ecological interaction between an evolved EGFR tyrosine kinase inhibitor (TKI) resistant non-small-cell lung cancer (NSCLC) population and its TKI-sensitive ancestor[39]. The focus of that work was on the ecological interaction under TKI treatment, and the inevitable competitive release. Strikingly, we observed an interaction between the resistant mutant and its ancestor in the absence of any treatment. The resistant population was observed to grow about twenty percent slower than the ancestor when cultured separately, however when the resistant population was co-cultured with a majority ancestor population, that difference in fitness nearly vanished. This observation, referred to as negative frequency-dependent selection (negative because the selection for the mutant increases as the mutant frequency *decreases*), is a long-studied phenomenon[40–42], and has been described as the most "intuitively obvious explanation for polymorphisms in nature"[43]. Despite its long history and potential for potent evolutionary effects, frequency-dependent selection remains understudied in the context of drug resistance. This is especially surprising, because a resistant population typically first emerges as a single individual in a predominantly ancestor population, and as a result frequency-dependent ecological interactions have a profound potential to effect the dynamics of a resistant clone (**Fig. 1**).

In this work, we seek to develop a rigorous theory of pre-treatment evolution that incorporates frequency-dependent ecological interactions between the emerging resistant subclones and the ancestor from which they evolve. Using both a generalized Moran process and Wright-Fisher simulations, we show that mutants with the same intrinsic fitness (monoculture fitness) can have mean extinction times that vary by several orders of magnitude as a function of their ecological fitness (fitness when co-cultured in a predominantly ancestor environment). Next, we calculate the expected number of resistance-conferring mutants in the population as a function of the cost of resistance, as well as the population size, and rate at which resistance-conferring mutations occur. When comparing the result of this calculation both when we assume ecological interactions exist, and when they are forbidden, we identify a wide parameter space where pre-existence is only likely to occur if ecological-interactions are assumed. We then investigate the "many mutant regime" where pre-existence is likely even without ecological interactions, and demonstrate that these ecological interactions would play a prominent role in shaping the distribution of mutants, dramatically increasing the prevalence of mutants with high ecological fitness. Importantly, we show that these ecological effects drive the evolutionary outcomes even when mutants with high ecological fitness are rare. Surprisingly, despite the complexity of the model, we obtain analytical approximations for extinction rates, expected number of resistance-conferring mutants, and the distribution of observed mutants over the full range of ecological fitness. These analytical approximates both support our numerical simulations and allow us to extend our results to population sizes too large to simulate.

Finally, we test our theory experimentally by engineering several of the most common clinically-observed mutations to TKI-therapy in EGFR-driven NSCLC and compete these mutants against the TKI-sensitive ancestor. In all cases we observed an ecological interaction that resulted in mutant ecological fitnesses larger than their intrinsic fitness. In the case of BRAFV600E, despite harboring a fitness cost of approximately 20 percent, the ecological interaction was sufficiently large that its growth rate eclipsed the ancestor's. This is particularly striking because it suggests BRAFV600E will survive indefinitely as a small

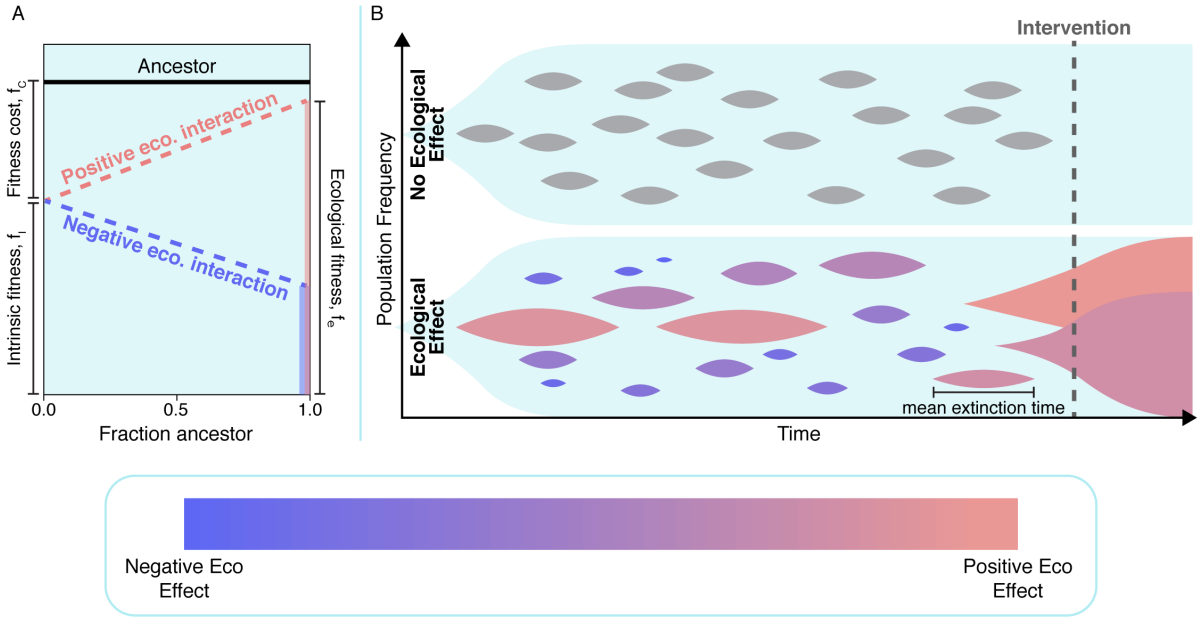

Figure 1. Cartoon abstraction demonstrating how frequency-dependent ecological interactions could increase the likelihood of pre-existing resistance. (A) Cartoon visualization of a typical frequency-dependent growth experiment. The ancestor (black line) is assumed to grow at a constant rate. Two hypothetical resistant mutants are depicted. Both mutants shared the same intrinsic fitness and fitness cost, however the positive ecological mutant (red, growth increases as the fraction of ancestor cells increases) has a significantly higher ecological fitness $f_e \approx 1$) than the negative ecological mutant (blue, growth decreases as the fraction of ancestor cells increases). As a result no mutants are present when a drug intervention is administered (vertical dashed line). (B) Top: Cartoon visualization of an evolving population with no ecological interactions. All mutants are assumed to have so non-insignificant fitness cost, f_c, and as a result go extinct. Bottom: The same evolving population, assuming ecological interactions are present. Note that an identical number of mutants emerge, however semi-rare mutants with positive ecological interactions demonstrate an increased time to extinction. As a result, when a drug intervention is administered, pre-existence is present.

fraction of the population, lurking until treatment results in competitive release of the untreatable resistant subclone. Taken together, these theoretical and experimental results argue that frequency-dependent ecological interactions between resistance mutants and their ancestor confer the primary mode by which resistance emerges in modern cancer therapeutics, and potentially all evolutionary diseases.

Results

Ecologically-dependent extinction time distributions with a generalized Moran process

We begin by considering a one-step birth-death process[44–46] with states $s \in \{0, 1, .., N\}$, where N is the total population size, s is the mutant population, and $N - s$ is the ancestor population. We do not consider mutation and as a result the states $s = 0$ (extinction) and $s = N$ (fixation) are absorbing. To account for ecological interactions, the mutant's growth rate is defined to be a function of $\frac{N-s}{N}$, or the fraction of the population that is of the ancestral type, and assumed to be linear, see **Fig. 1, right**. In addition, for simplicity, we define the ancestor's growth rate to be constant and, without loss of generality, normalized to 1 (see Materials and Methods for full model details).

First, we are interested in how the distribution of extinction times differs between recently emerged mutants with identical fitness costs, but distinct ecological interactions (**Fig. 2A, left**). In particular we assume one (neutral) mutant has no ecological interaction with the ancestor, and thus $f_c = f_e$ (**Fig. 2A, blue**), while the comparative (positive) mutant has an interaction that ameliorates the fitness cost of the mutant at extremely large ancestor fractions, $f_e = 1$ (**Fig. 2A, red**). In the case of a mutant with a positive ecological interaction, we see that the extinction time distribution is heavily right-skewed in comparison to a neutral ecological effect. As a result, if these two mutants were equally likely to emerge in a population, we would expect to observe a mutant with a positive ecological interaction significantly more often than an equivalent mutant with a neutral ecological interaction. However, ecological interactions are not always positive. Repeating this process in comparing a neutral mutant with a mutant that has a negative ecological interaction with the ancestor reveals distinct shift to shorter extinction times as one might intuitively expect (**Fig. S1**).

Extinction times depend on ecological interactions in a Wright-Fisher model

While formulating our system as a generalized Moran process allows for convenient closed-form solutions to quantities of interest such as extinction time distributions, this representation becomes computationally expensive as the population size approaches increasingly realistic values. In addition, we have completely ignored mutation, as well as more realistic conditions where many mutants are competing within an evolving population. As such, we switch to a Wright-Fisher formulation of our system[47–49]. In the Wright-Fisher model, populations are still constant in population size N, however each individual of the population is replaced every generation with offspring inheriting the parent's genotype with probability proportional to the parent's fitness. In addition, individuals acquire mutations with some probability μ and we assume mutant populations are sufficiently small that we can ignore mutant-mutant interactions. Still, several of the generalizations and abstractions from the generalized Moran process remain. Namely, the ancestor's growth is defined to be constant and normalized to 1, and the mutant growth rate is assumed to vary linearly between f_i and f_e (as a result, a mutant's growth is fully characterized by these two fitness values along with the fraction of the population that is ancestor).

Each simulation begins with an exclusively ancestor population and with each generation cells mutate with probability μ. Each mutant that arises has an intrinsic fitness drawn with uniform probability in $[0, 1 - f_c]$ and a corresponding ecological fitness drawn with uniform probability in $[0, 1]$. Each Wright-Fisher 'generation' consists of a mutation step, followed by an offspring/selection step. For each mutant that emerges we record its intrinsic and ecological fitness values and track its evolutionary trajectory, and thus extinction time (τ). A mutant that emerges but does not survive the subsequent selection step is defined to survive 0 generations. Employing this model we find that the mean extinction time varies nearly five orders of magnitude between the most positive (≈ 10 generations) and deleterious (≈ 0.001 generations) ecological interactions (**Fig. 2B**). In order to develop a more rigorous understanding of the evolutionary dynamics, we sought an analytical approximation for the extinction time of a mutant under the same Wright-Fisher conditions. Strikingly, we find a robust approximation across the whole range of f_e:

$$\tau(f_e) \approx \frac{3\ln(1 - f_e)}{f_e^2 - 3}. \tag{1}$$

Despite its simple form, this approximation agrees with simulation results with a typical error of 5% (**Fig. 2B**, full derivation and details found in the SI). Interestingly, the approximation is only a function of ecological fitness, and not mutation rate (assuming $\mu \ll 1$), population size, or fitness cost. This finding is supported by our simulation results (**Fig. 2C**).

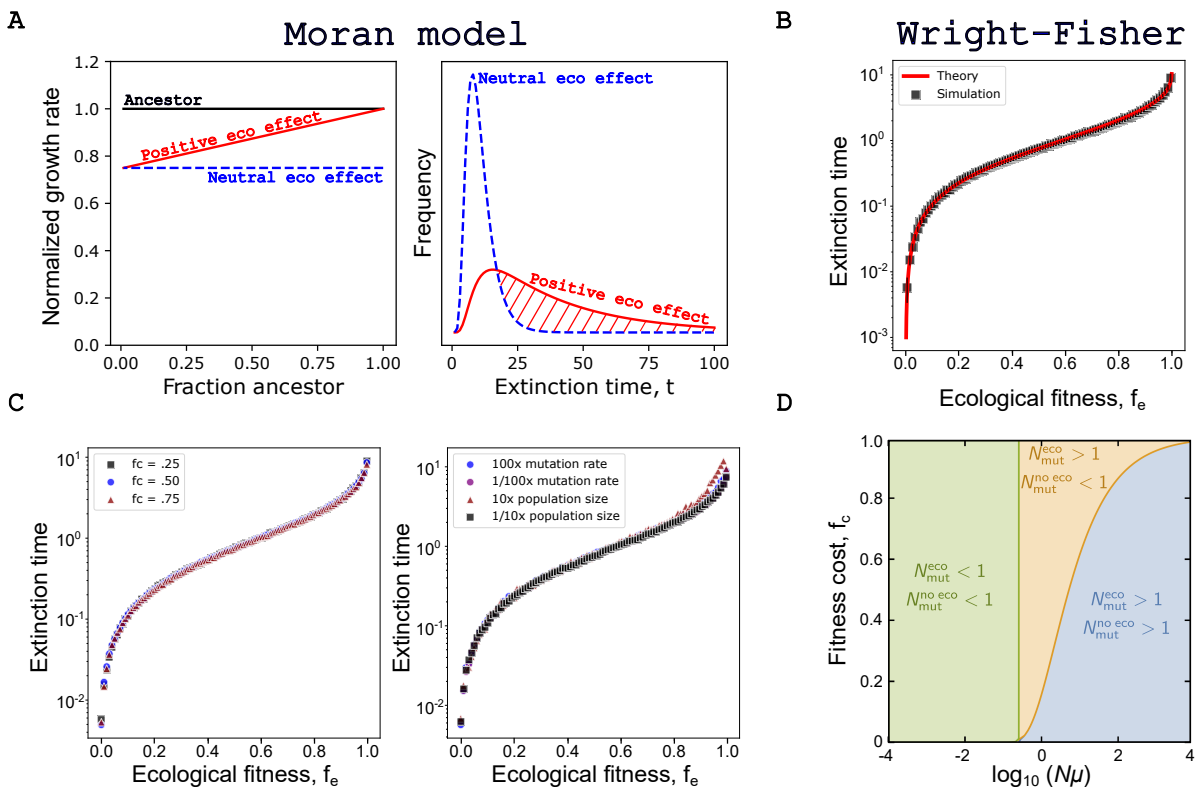

Figure 2. Analytical approximations and simulations predict that extinction times depend on ecological interactions. (A) Closed form extinction time distributions are calculated and visualized for a generalized Moran process (N=100, $f_c = 0.25$). The red distribution results from a mutant with a positive ecological interaction with the ancestor ($f_e = 1.0$), while the blue population has no ecological interaction with the ancestor ($f_e = 1 - fc = 0.75$). (B) Wright-Fisher simulations are used to numerically calculate the mean extinction time as a function of f_e (N=10000, $\mu = 10^{-6}$, 500 generations, f_i is drawn uniformly in $[0, 1 - f_c]$, f_i is drawn uniformly in $[0, 1]$). (C) Wright-Fisher simulations are repeated for varying values of f_c, μ, and N to confirm theoretical prediction that the extinction time distribution depends only on f_e. (D) Phase diagram depicting the three regimes of pre-existing resistance.

Ecological interactions can increase the probability of pre-existing resistance

Next we consider the model's implications for pre-existing resistance. Specifically, we are interested in quantifying the expected number of mutants in an evolving population. While it might be tempting to quickly conclude that including ecological interactions will necessarily increase the probability of pre-existing resistance because positive interactions will lead to longer extinction times, it is important to note that mutants with a high intrinsic fitness are more likely to acquire a relatively deleterious ecological fitness, than one that is beneficial. As such, a careful mathematical treatment is required. When the expected number of mutants in a population is low ($N_{\text{mut}} \ll 1$), potential resistance-conferring mutations are unlikely to be present at time of treatment. Contrarily, when the expected number of mutants is greater than 1, we expect treatment threatening resistance to be present when a drug is administered. We begin adapting our analytical model to calculate the mean number of mutants (see SI for full derivation and details). To begin we consider the case where no ecological interactions are present (f_i completely describes the growth rate of the mutant). In this case it can be shown that $N_{\text{mut}}^{\text{no eco}}$, the mean number of mutants ignoring ecological interactions, is:

$$N_{\text{mut}}^{\text{no eco}} = N\mu \left(-\frac{\ln f_c}{1 - f_c} - 1 \right). \tag{2}$$

Next, we seek to find an analytical approximation for $N_{\text{mut}}^{\text{eco}}$, the mean number of mutants assuming ecological interactions exist. In the case of a sufficiently small mutation rate, we can approximate the total mutant fraction as,

$$N_{\text{mut}}^{\text{eco}} \approx N\mu \left(-\frac{\ln(1 - f_{\max})}{f_{\max}} - 1 \right) \qquad \text{for } \mu \ll 1. \tag{3}$$

Here, $f_{\max}$ is the maximum value that f_e can take. While we can set $f_{\max}$ arbitrarily close to 1, it can never be exactly 1 for a well-defined normalization. Interestingly, for sufficiently small μ, the ratio $\frac{N_{\text{mut}}^{\text{eco}}}{N_{\text{mut}}^{\text{no eco}}}$ is constant with $N\mu$. While the simplicity of the approximation is appealing, unfortunately it breaks down as μ gets large. As a result, a more robust, though significantly more complex, approximation was derived (see SI for full derivation):

$$N_{\text{mut}}^{\text{eco}} \approx \frac{N\mu}{f_{\max}} W \left(\left(1 - f_{\max} + \mu f_c^{f_c/(f_c-1)} \right)^{-1} \right). \tag{4}$$

Here $W(x)$ is the Lambert W function, which is the solution y of the equation $ye^y = x$. This approximation allows for efficient calculation across several decades of μ within 10% of our numerical simulations. Employing these analytical approximations we identify three regimes of interest. The least interesting regime is the small $N\mu$ regime (**Fig. 2D, green**). Here the effective population size is insufficient to maintain a mutant subpopulation regardless of the strength or frequency of ecological interactions. While this is the least mathematically interesting regime, this corresponds to extremely rare pre-existing resistance and high likelihood of treatment success.

As $N\mu$ gets larger (**Fig. 2D, yellow**) we enter a regime where ecological interactions would suggest pre-existing resistance is likely ($N_{\text{mut}}^{\text{eco}} > 1$), while ignoring ecological interactions would suggest pre-existence is still rare ($N_{\text{mut}}^{\text{no eco}} < 1$). In this regime mutants have yet to become abundant, however, mutants with strong ecological interactions persist sufficiently long to threaten treatment efficacy. Representative simulation trajectories of this "rare mutant regime" are shown in **Fig. 3A**. Without ecological interactions (**Fig. 3A, top panel**) the mutation rate alone is insufficient to maintain a mutant subpopulation capable of threatening future treatment efficacy. However, with the introduction of ecological interactions (**Fig. 3A, bottom panel**), rare positive ecological mutants climb to significant fractions of the population, and have measurably longer extinction times that may threaten future treatments. As one might intuitively expect, the size of this regime where ecological effects drive pre-existence is heavily dependent on the imposed fitness cost of resistant mutants. We find that the larger the fitness cost imposed by resistance, the larger the comparative increase provided by allowing ecological effects.

Ecological interactions significantly influence the distribution of mutants

Next, we consider the final regime when $N\mu$ is large (**Fig. 2D, blue**). In this regime the mutational supply is sufficiently large to self-sustain a small, resistant subpopulation, regardless of ecological interactions (that is, both ($N_{\text{mut}}^{\text{no eco}} > 1$ and $N_{\text{mut}}^{\text{eco}} > 1$). Representative simulation trajectories of this "many mutant regime" are shown in **Fig. 3B**. At first glance one might assume this regime is uninteresting. In both cases mutants are sufficiently common to threaten future treatments, albeit ecological interactions significantly increase the steady-state fraction of resistant mutants. However, the results become more interesting when we consider the shape of the resistant subpopulation distribution. In each trajectory plot, the color is proportional to the

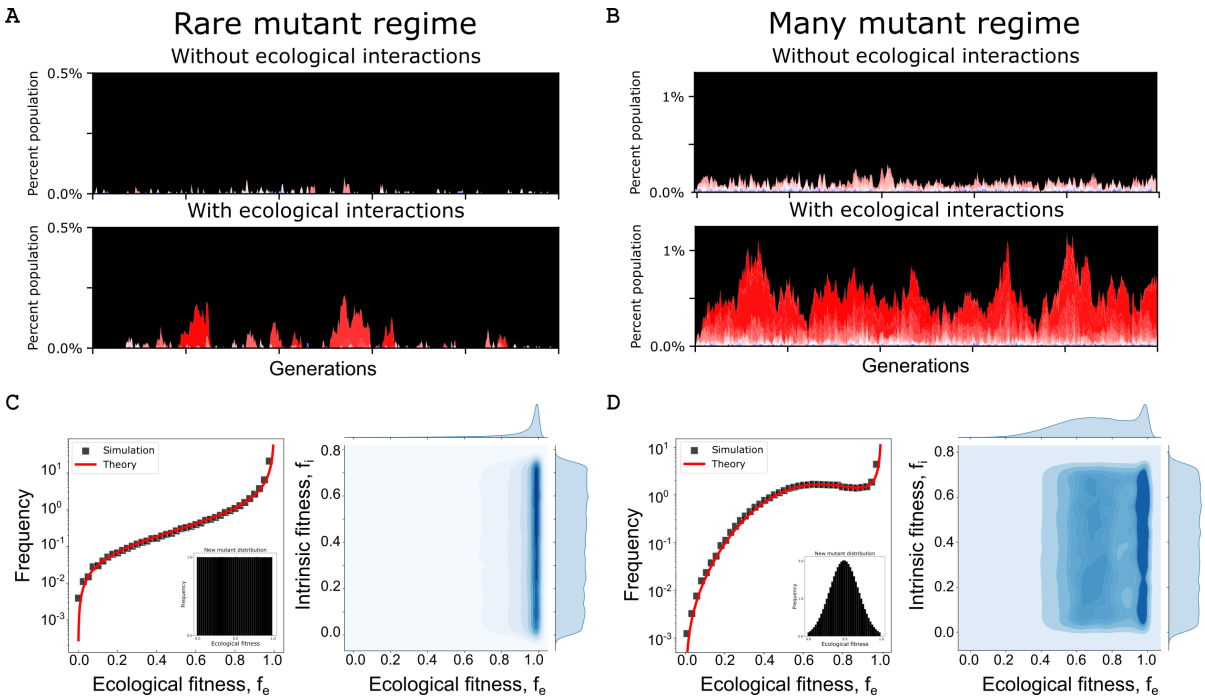

Figure 3. Positive ecological interactions make pre-existence more likely and dominate the stationary distribution of mutants. (A) Representative Wright-Fisher trajectory in the "rare mutant regime". Mutants exist in higher fractions and for longer periods with ecological interactions. Each mutant is colored by its ecological fitness where red represents an f_e value near 1 and blue represents an f_e value near 0. (B) Representative trajectory in the "many mutant regime". Strong positive ecological interactions dominate the stationary distribution of mutants. (C) Left: Stationary distribution of mutant ecological fitnesses when the mutant generating function is uniform in ecological fitness. Right: joint distribution density plot between intrinsic and ecological fitness. (D) Same as C, however the mutant generating function is now Gaussian centered about $f_e = 0.5$.

mutants ecological fitness with red representing an ecological fitness near 1 and blue representing an ecological fitness near 0. By inspection it is immediately clear that the most positive ecological mutants are over represented in the mutant population, considering they emerge with equal probability. However, we can do better and extract this relationship explicitly from our simulations (**Fig. 2C, left**). We find, similar to the impact of ecological effects on extinction times, that the frequency of a mutant spans multiple orders of magnitude as a mutant's ecological fitness varies from 0 to 1.

Extending our previous analytical work, it is straightforward to show that the stationary distribution of mutant ecological fitnesses goes as,

$$P(f_e) \approx \frac{f_e \mu}{f_{\max}(1-f_e)} \qquad \text{for } \mu \ll 1. \tag{5}$$

The above approximation works remarkably well despite the simplicity of its form. From this equation we find that the frequency of a mutant is invariant with respect to fitness cost and population size. This is shown explicitly via numerical simulations and visualization of the joint distribution of fitness cost and ecological fitness (**Fig. 2C, right**).

Non-uniform ecological distributions show similar qualitative results

An important context to keep in mind with the work is that up to this point we have assumed emerging mutants are assigned an ecological fitness with uniform probability in $[0,1]$. This assumption was not made for simplicity, but instead out of necessity. While evolutionary biologists have spent significant time both theorizing about, and measuring the distribution of fitness effects (DFE), very little time has been spent quantifying either the frequency or magnitude of ecological effects (distribution of ecological effects, DEE). As a result, it is difficult to even speculate on what the null model ought to be.

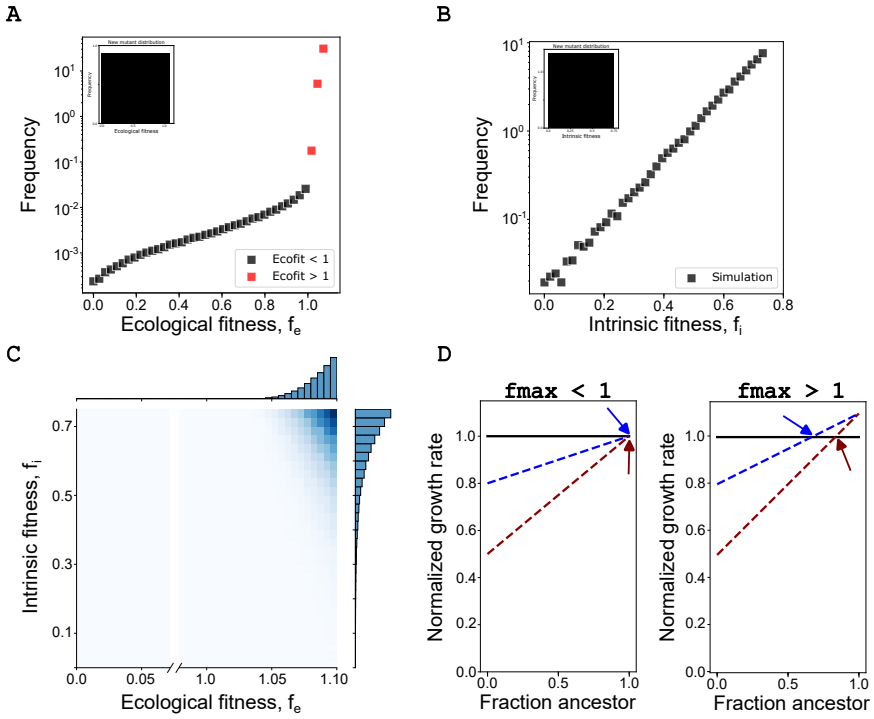

Figure 4. Positive ecological fitnesses above 1 result in a stable fixed point between mutant and ancestor. (A) Stationary distribution of mutant ecological fitnesses when the mutant generating function has uniform probability in $[0, 1.10]$. (B) Stationary distribution of mutant intrinsic fitnesses. (C) Joint distribution density plot between intrinsic and ecological fitness reveals the size of the fitness cost now has a significant impact on mutant survival. (D) Cartoon illustration of why two mutants with identical values of f_e can result in different extinction times. Colored arrows point to stabled fixed points between mutant and ancestor. Importantly, only the present mutant with the stable fixed point at the smallest value of f_e will achieve stability.

Crucially, the analytical approximations derived herein can be simply generalized to fit any assumed, or future measured, DEE. While we assumed a uniform distribution, a Gaussian model where the most positive and negative ecological interactions are rare relative to more modest, or non-interacting mutants, may be more accurate. As an example, the general stationary distribution of mutant ecological fitnesses would become,

$$P(f_e) \approx \frac{f_e \mu}{(1 - f_e)} \rho_0(f_e) \qquad \text{for } \mu \ll 1. \tag{6}$$

Here, $\rho_0(f_e)$ can be any theorized or measured distribution of ecological effects. As proof of principle, we numerically simulate the distribution of mutant ecological fitnesses under an assumed Gaussian DEE, and show the above analytical approximation still holds. The results are qualitatively similar to the uniform DEE and, strikingly, despite the rarity of mutants with positive ecological interactions, they still manage to dominate the predicted stationary distribution of mutants (**Fig. 2D**).

Sufficiently large positive ecological interactions result in a stable fixed point between mutant and ancestor

We now briefly consider the regime wherein the ecological fitness of a mutant can sample values greater than 1. Put another way, when the mutant population emerges, it may emerge into an environment where it out-competes its ancestor. Importantly, emerging mutants still have a nonzero fitness cost, so even the most positive ecological interactions cannot lead to a hard selective sweep. Though our earlier analytical approximations do not apply for $f_{\max} > 1$, the numerical simulations are robust in this regime. We find that the majority of stationary distribution mutants are mutants with ecological fitnesses larger than the ancestor, or $f_e > 1$ (**Fig. 4A**). This qualitative change in behavior above $f_e = 1$ can be explained in evolutionary game theory terms by a switch in the evolutionary game being played. When $f_e < 1$, the ancestor out-competes the mutant population at all population frequencies. As a result, it is a question of when, not if, the mutant population will be driven to extinction. When $f_e > 1$, however, the mutant population out-competes the ancestor at high ancestor frequencies, while the ancestor out-competes

the mutant at high mutant frequencies (as a result of the mutant fitness cost). This leads to a stable fixed point at some ancestor frequency where the two populations have an equal growth rate. This is particularly worrying in the case of therapy-resistant mutants, because it suggests if such a mutant emerges and survives the initial stochasticity of drift, it coexist at a sizeable frequency in the population until treatment.

Next, we found that, contrary to our previous results, the fitness cost of the mutant plays an important role in determining the stationary distribution of mutants (**Fig. 4B,C**). Here we see that only the mutants with the largest positive ecological interactions and smallest fitness costs (intrinsic fitness $= 1-$ fitness cost) are represented at meaningful frequencies. This result can be explained by the qualitative shift in evolutionary game for mutants where $f_e > 1$. Previously, regardless of the fitness cost, any mutant with $f_e \approx 1$ would grow at that ecological fitness, as the mutant population never became a meaningful fraction of the whole population (**Fig. 4D, left**). However, as hinted at in the numerical simulations, f_c and f_e *combine* to determine the stable fixed point between the mutant and the ancestor (**Fig. 4D, right**). As a result, even mutants with $f_e < 1$ are no longer characterized by their ecological fitness, instead they are characterized by their fitness at the frequency determined by the stable fixed point.

Clinically observed lung cancer mutations confer positive ecological interactions

Epidermal growth factor (EGFR) tyrosine kinase inhibitors (TKIs) are the first-line treatment for patients diagnosed with advanced non-small cell lung cancer (NSCLC). While the development of targeted TKIs has importantly extended overall survival times, these drugs are rarely curative[50] and patients often recur with TKI-resistant tumors. As a result, EGFR-mutant NSCLC is an ideal system for studying pre-existing resistance and a location we would expect to find the strongest evidence of our predictions — namely the presence of strongly positive ecological interactions between mutants and their ancestor.

To test our theory we genetically engineered (see Materials and Methods) three of the most commonly clinically observed resistance mutations found in response to TKIs:[51,52] BRAF-V600E, KRAS-G12V, and PIK3CA-E545K. Then, using our previously described evolutionary game assay[23,39], we measured the ecological interaction between each of these mutants and the ancestor PC9 cell line from which they emerged.

We found that each of the three engineered mutants had varying levels of positive ecological interactions with their ancestor (**Fig. 5, top**). The most extreme result of the three was undoubtedly the BRAFV600E mutant, in which we observed a sufficiently strong ecological interaction to out-compete the ancestor in predominantly ancestor populations ($f_e > 1$). This observation suggests, at least in a laboratory environment, that if BRAFV600E emerges and survives stochastic evolutionary forces, it will stably coexist at approximately 20% of the population.

A similarly strong positive ecological interaction was observed between KRASG12V and the PC9 ancestor. While the ecological fitness did not eclipse the ancestor, the ecological interaction was strong enough to completely ameliorate its fitness cost of approximately 20%. Finally, the PIK3CA-E545K mutant showed the weakest ecological interaction, though still slightly positive in nature. While we already reported on the ecological interaction between the evolved mutant and its ancestor, as it was the motivator of this study[39], we performed additionally sequencing analysis (WXS and RNA-seq) and identified several common clinical mutations present, distinct from the engineered mutations: MET overexpression, CCND1 amplification, and KRASG12D mutation (**Fig. 5, bottom**). Taken together, these experimental results match our theoretical predictions and strongly support the hypothesis that frequency-dependent ecological interactions are a primary mode by which resistance is conferred in evolutionary diseases.

Discussion

While much work has gone into quantifying clinically problematic resistant bacteria, cancers and viruses, we nearly always characterize these clones in monoculture - entirely outside the eco-evolutionary forces that selected for (or against) them in the first place. In this work we set out to provide the foundation for a rigorous and generalizable mathematical framework that incorporates frequency-dependent ecological interactions and can be used to study their role in pre-existing resistance. This work both compliments and builds off of recent studies from a wide range of disciplines ranging from theoretical population genetics and ecology to clinical trials across several biological kingdoms. We demonstrate that the presence of ecological interactions can significantly increase the probability of pre-existing resistance, in addition to shaping the distribution of mutants likely to be present before treatment. We derive analytical approximations of several quantities of interest including extinction time, mean mutant population numbers, and the underlying distribution of mutants each as a function of ecological fitness. Importantly, these results can easily be generalized to any theorized distribution of ecological effects, or future experimentally measured distribution. As an important example, we show that even when we assume positive ecological interactions are rare, they still end up as a plurality of the stationary mutant frequency distribution. Finally, in an model system for pre-existing resistance, we show common clinically observed mutants harbor positive, frequency-dependent ecological interactions when co-cultured with their ancestor, providing strong evidence for our theory in cancer. In addition, recent exciting work in bacteria

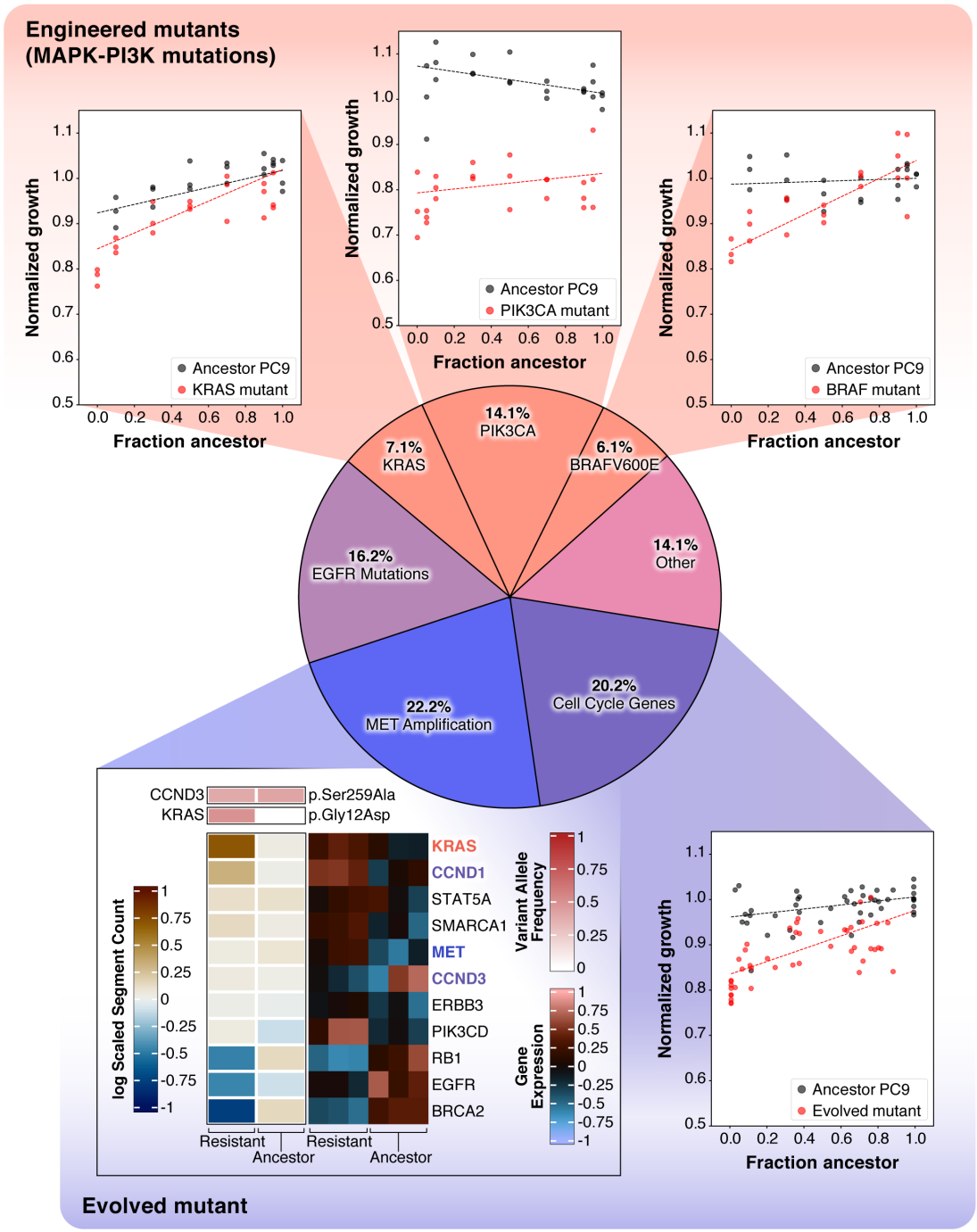

Figure 5. As predicted, common clinically observed resistance mutations in NSCLC harbor strong postitive ecological interactions with their ancestor in a model system of pre-existing resistance. Pie chart: Visual representation of the known resistance mechanisms to Osimertinib, a third generation TKI and the current standard of care for EGFR-positive NSCLC. Top: Measured positive ecological interactions between engineered resistant mutants and their ancestor. From left to right — KRAS-G12V, PIK3CA-E545K, BRAF-V600F. Bottom right: Measured positive ecological interaction between evolved gefitinib-resistant NSCLC PC9 population and its ancestor. Previously reported in[39]. Bottom left: Fresh sequencing analysis reveals the gefitinib-resistant evolved mutant has additional clinically observed resistance mutations including: KRASG12D, MET amplification, and CCND1 amplification (cell cycle genes).

provides additional evidence, as frequency-dependent interactions resulted in maintenance of otherwise costly antibiotic resistant populations in *Escherichia coli*[53] and *Pseudomonas aeruginosa*[54].

It is also important to address several limitations of our work. As we mentioned earlier, the distribution of ecological effects (DEE) has never been experimentally measured. As a result, assumptions regarding the distributional parameters have to made in order to calculate meaningful quantities of interest. While we did our best to combat this by developing analytical models that are agnostic to this distribution, the quantitative aspect of our results are subject to the specifics of a model. Our hope is that the analytical and numerical results herein, when combined with the promising experimental work in NSCLC, motivate future measurements of the DEE across diverse model systems. Similarly, our own experimental validation is constrained to one subsystem. Our predictions are broad and should apply to many evolving populations where pre-existence is evolved. Therefore it is important that future studies should should aim to test these theories not just in other cancers, but in other organisms from HIV to drug resistant bacteria. It is possible that these principles provide the most explanatory power in cancer and bacteria where it is common to find highly dense heterogeneous populations in contrast to viruses, for example. Finally, while the model aims to generally capture major evolutionary forces that may underlie pre-existing resistance, it is still an abstraction of a much more complex clinical scenario where the immune system, spatial dynamics, and treatment adherence, to name only a few, can play major roles.

Materials and Methods

Cell culture

PC-9 are human adenocarcinoma cells derived from undifferentiated lung tissue were obtained from Sigma (Sigma, USA). PC-9 cells were cultured in RPMI-1640 medium supplemented with 10% heat inactivated fetal bovine serum (FBS) and 1% penicillin streptomycin solution at 37°C with humidity containing 5% CO2. Cells were split every four days to maintain optimum confluency of ≈80-90%.

Engineering of mutant cell lines

To establish PC-9 cells stably expressing target genes, HEK-293T cells were co-transfected using TransIT-Lenti transfection reagent (Mirus, USA), with 500ng psPAX2 (addgene, USA), 100ng PMD2 (addgene, USA) and 500ng of target genes. Viral particles were collected after 48hrs and used to transduce PC-9 cells. Then, to establish ancestor PC-9 cells stably expressing nuclear localized GFP, cells were transduced with pLVX-eGFP-Hygro (Vectorbuild, USA). In addition, to establish query cells expressing fluorescently labeled PC-9 cells with a gene of interest cells were co-transduced with pLVX-mCherry-Hygro or pLVX-mCherry-Puro and each of pLVX-PIK3CA-E545K-Bsd (Vectorbuild, USA). Next, 72hrs after transduction, cells were selected with 200μg/ml hygromycin, 5μg/ml puromycin and 5μg/ml blasticidin.

Drug sensitivity assay

Cells were harvested at 70-80% confluence, stained with trypan blue (Corning, USA), and counted with a TC20 Automated Cell Counter (Bio-Rad, USA). Luminescent based cell viability assays using CellTiter-Glo (CTG) reagent (Promega, USA) were performed in 96 well plate (Corning, USA). A total number of $3,000$ cells were plated in 90µL of complete medium per well in three replicate per drug concentration with Multidrop reagent dispenser (Thermo Fishers, USA). After 3hrs of incubation, 10µL of gefitinib, osimertinib and erlotinib (Cayman, USA) diluted in complete RPMI-1640 medium were added to the cells. Compounds were tested in a threefold dilution in a range of $0-1.8\mu$M, $0-3\mu$M and $0-10\mu$M for gefitinib, osimertinib and erlotinib respectively. After 72hrs of incubation, 25µL CTG reagent was add to the cells; incubated for 10 minutes at room temperature and luminescence was measured.

Game assay

PC-9 mutants stably expressing nuclear localized fluorescent signal ancestor PC-9 stably expressing nuclear localized GFP were co-cultured at different initial proportion of ancestor cells at a density of $1,500$ cells in 90μL of fresh medium. After 3hrs of incubation, 10μL of DMSO diluted in complete RPMI-1640 medium (final DMSO concentration of 0.1% v/v) were added to the cells in three replicates per initial proportion. Then time-lapse microscopy images were obtained for GFP and mCherry using BioSpa automated incubator (BioTek, USA) every 4 hours over the course of 96 hrs. Then, images were processed with the open-source software CellProfiler[55]. Images were background subtracted, converted to 8-bit, contrast enhanced, and thresholded, then raw cell numbers were extracted.

Generalized Moran model details

In this work we consider a well known generalized Moran process, a model previously used to study frequency dependent evolutionary dynamics[44–46]. Briefly, we consider a one-step birth-death process with states $s \in \{0,1,..,N\}$ and characterized by birth and death rates b_i and d_i with $i \in \{1,1,..,N-1\}$. As a result this model describes a fixed population size N, with s resistant mutants and $N-s$ ancestor population. We forgo mutation rate and as a consequence the states $s=0$ and $s=N$ are absorbing. We consider an evolutionary game with a $2x2$ payoff matrix such that:

$$\begin{array}{cc} & \begin{array}{cc} \text{R} & \text{A} \end{array} \\ \begin{array}{c} \text{R} \\ \text{A} \end{array} & \begin{pmatrix} (1-f_c) & f_e \\ 1 & 1 \end{pmatrix} \end{array}$$

As a result, we can write the expected payoffs in a population of s mutants and N-s ancestor individuals as:

$$P(R) = \frac{s-1}{N-1}(1-f_c) + \frac{N-s}{N-1} f_e \qquad\qquad P(A) = 1.$$

RNA-Seq

Paired-end reads are preprocessed using *fastp* to trim and quality filter the reads. Following the filtering, reads are aligned to GRCh38 reference genome via *STAR* aligner. Read quantification is done using *Salmon* on the extracted transcriptome locations from spliced STAR alignment. Gene-level abundance are aggregated from bootstrapped transcript abundances using R package *tximport*. The *Arriba* tool is coupled with spliced alignments for fusion-transcript detection as well. Pathway level expression activities are quantified using R package *GSVA* and *msigdbR* for the Hallmark Pathways. The R package *ComplexHeatmap* was used to generate heatmaps.

Whole exome sequencing

Paired-end whole-exome reads of ancestor and parental lines were preprocessed using *fastp* similar to RNA-Seq. Alignment to GATK (GATK best practices bucket) version of GRCh38 reference is done using *bwa-mem* aligner. Following the alignment, variant calling pipeline according to the GATK workflow including duplicate marking and variant calling via HaplotypeCaller was conducted. Variants passing filtering based on hard-filtering are further annotated using Variant Effect Predictor (VEP) tool. Exome alignments are further input to *CNVkit* for copy-number alterations. Using a flat-reference for bias correction log2 scaled abundances are generated for ancestor and resistant strains. Copy number segments are captured using circual binary segmentation and assigned to genes mapping to the segment.

Code availability

Code used in this study will be made openly available on GitHub.

Acknowledgments

This work was made possible by the National Institute of Health T32CA094186 (JAM), The Research Council of Norway 325628/IAR (DST), 5R37CA244613-03 (JGS), 5T32GM007250-46 (JGS), and American Cancer Society RSG-20-096-01 (JGS).

References

1. Fisher, M. C. *et al.* Tackling the emerging threat of antifungal resistance to human health. *Nat. Rev. Microbiol.* **20**, 557–571 (2022).
2. Goldberg, D. E., Siliciano, R. F. & Jacobs, W. R. Outwitting evolution: fighting drug-resistant tb, malaria, and hiv. *Cell* **148**, 1271–1283 (2012).
3. Darby, E. M. *et al.* Molecular mechanisms of antibiotic resistance revisited. *Nat. Rev. Microbiol.* 1–16 (2022).
4. Pennings, P. S. Hiv drug resistance: problems and perspectives. *Infect. disease reports* **5** (2013).
5. Menard, D. & Dondorp, A. Antimalarial drug resistance: a threat to malaria elimination. *Cold Spring Harb. perspectives medicine* **7**, a025619 (2017).
6. Sandru, A., Voinea, S., Panaitescu, E. & Blidaru, A. Survival rates of patients with metastatic malignant melanoma. *J. medicine life* **7**, 572 (2014).
7. Boucher, H. W. *et al.* Bad bugs, no drugs: no eskape! an update from the infectious diseases society of america. *Clin. infectious diseases* **48**, 1–12 (2009).
8. Organization, W. H. *Antimicrobial resistance: global report on surveillance* (World Health Organization, 2014).
9. Arnold, M. *et al.* Progress in cancer survival, mortality, and incidence in seven high-income countries 1995–2014 (icbp survmark-2): a population-based study. *The Lancet Oncol.* **20**, 1493–1505 (2019).
10. Maltas, J., Singleton, K. R., Wood, K. C. & Wood, K. B. Drug dependence in cancer is exploitable by optimally constructed treatment holidays. *bioRxiv* DOI: 10.1101/2022.07.01.498458 (2022). https://www.biorxiv.org/content/early/2022/07/03/2022.07.01.498458.full.pdf.
11. Iram, S. *et al.* Controlling the speed and trajectory of evolution with counterdiabatic driving. *Nat. Phys.* **17**, 135–142 (2021).
12. Gatenby, R. A., Silva, A. S., Gillies, R. J. & Frieden, B. R. Adaptive therapy. *Cancer research* **69**, 4894–4903 (2009).
13. Feder, A. F., Harper, K. N., Brumme, C. J. & Pennings, P. S. Understanding patterns of hiv multi-drug resistance through models of temporal and spatial drug heterogeneity. *eLife* **10**, e69032, DOI: 10.7554/eLife.69032 (2021).

14. Anderson, G. R. *et al.* A landscape of therapeutic cooperativity in kras mutant cancers reveals principles for controlling tumor evolution. *Cell reports* **20**, 999–1015 (2017).

15. Hegreness, M., Shoresh, N., Damian, D., Hartl, D. & Kishony, R. Accelerated evolution of resistance in multidrug environments. *Proc. Natl. Acad. Sci.* **105**, 13977–13981 (2008).

16. Torella, J. P., Chait, R. & Kishony, R. Optimal drug synergy in antimicrobial treatments. *PLoS computational biology* **6**, e1000796 (2010).

17. Dean, Z., Maltas, J. & Wood, K. B. Antibiotic interactions shape short-term evolution of resistance in e. faecalis. *PLoS pathogens* **16**, e1008278 (2020).

18. Fu, F., Nowak, M. A. & Bonhoeffer, S. Spatial heterogeneity in drug concentrations can facilitate the emergence of resistance to cancer therapy. *PLoS Comput. Biol* **11**, e1004142 (2015).

19. Zhang, Q. *et al.* Acceleration of emergence of bacterial antibiotic resistance in connected microenvironments. *Sci.* **333**, 1764–1767 (2011).

20. De Jong, M. G. & Wood, K. B. Tuning spatial profiles of selection pressure to modulate the evolution of drug resistance. *Phys. review letters* **120**, 238102 (2018).

21. Korolev, K. S., Xavier, J. B. & Gore, J. Turning ecology and evolution against cancer. *Nat. Rev. Cancer* **14**, 371–380 (2014).

22. Wargo, A. R., Huijben, S., De Roode, J. C., Shepherd, J. & Read, A. F. Competitive release and facilitation of drug-resistant parasites after therapeutic chemotherapy in a rodent malaria model. *Proc. Natl. Acad. Sci.* **104**, 19914–19919 (2007).

23. Kaznatcheev, A., Peacock, J., Basanta, D., Marusyk, A. & Scott, J. G. Fibroblasts and alectinib switch the evolutionary games played by non-small cell lung cancer. *Nat. ecology & evolution* **3**, 450–456 (2019).

24. Good, B. H., McDonald, M. J., Barrick, J. E., Lenski, R. E. & Desai, M. M. The dynamics of molecular evolution over 60,000 generations. *Nat.* **551**, 45–50 (2017).

25. Maltas, J. & Wood, K. B. Pervasive and diverse collateral sensitivity profiles inform optimal strategies to limit antibiotic resistance. *PLoS biology* **17**, e3000515 (2019).

26. Maltas, J. & Wood, K. B. Dynamic collateral sensitivity profiles highlight challenges and opportunities for optimizing antibiotic sequences. *bioRxiv* DOI: 10.1101/2021.12.19.473361 (2021). https://www.biorxiv.org/content/early/2021/12/21/2021.12.19.473361.full.pdf.

27. Scarborough, J. A. *et al.* Identifying states of collateral sensitivity during the evolution of therapeutic resistance in ewing's sarcoma. *Iscience* **23**, 101293 (2020).

28. Maltas, J., McNally, D. M. & Wood, K. B. Evolution in alternating environments with tunable interlandscape correlations. *Evol.* **75**, 10–24 (2021).

29. Zhao, B. *et al.* Exploiting temporal collateral sensitivity in tumor clonal evolution. *Cell* **165**, 234–246 (2016).

30. Lin, J. J. & Shaw, A. T. Resisting resistance: targeted therapies in lung cancer. *Trends cancer* **2**, 350–364 (2016).

31. Boshuizen, J. *et al.* Reversal of pre-existing ngfr-driven tumor and immune therapy resistance. *Nat. communications* **11**, 3946 (2020).

32. Jangir, P. K. *et al.* Pre-existing chromosomal polymorphisms in pathogenic *E. coli* potentiate the evolution of resistance to a last-resort antibiotic. *eLife* **11**, e78834, DOI: 10.7554/eLife.78834 (2022).

33. Robinson, M., Tian, Y., Delaney IV, W. E. & Greenstein, A. E. Preexisting drug-resistance mutations reveal unique barriers to resistance for distinct antivirals. *Proc. Natl. Acad. Sci.* **108**, 10290–10295 (2011).

34. Szakacs, G. *et al.* Targeting the achilles heel of multidrug-resistant cancer by exploiting the fitness cost of resistance. *Chem. reviews* **114**, 5753–5774 (2014).

35. Sampah, M. E. S., Shen, L., Jilek, B. L. & Siliciano, R. F. Dose–response curve slope is a missing dimension in the analysis of hiv-1 drug resistance. *Proc. Natl. Acad. Sci.* **108**, 7613–7618 (2011).

36. Melnyk, A. H., Wong, A. & Kassen, R. The fitness costs of antibiotic resistance mutations. *Evol. applications* **8**, 273–283 (2015).

37. Weinreich, D. M., Delaney, N. F., DePristo, M. A. & Hartl, D. L. Darwinian evolution can follow only very few mutational paths to fitter proteins. *science* **312**, 111–114 (2006).

38. Kühnert, D. *et al.* Quantifying the fitness cost of hiv-1 drug resistance mutations through phylodynamics. *PLoS pathogens* **14**, e1006895 (2018).

39. Farrokhian, N. *et al.* Measuring competitive exclusion in non–small cell lung cancer. *Sci. Adv.* **8**, eabm7212, DOI: 10.1126/sciadv.abm7212 (2022). https://www.science.org/doi/pdf/10.1126/sciadv.abm7212.

40. Ayala, F. J. & Campbell, C. A. Frequency-dependent selection. *Annu. review Ecol. systematics* **5**, 115–138 (1974).

41. Clarke, B. C., Shelton, P. & Mani, G. Frequency-dependent selection, metrical characters and molecular evolution. *Philos. Transactions Royal Soc. Lond. B, Biol. Sci.* **319**, 631–640 (1988).

42. Levin, B. Frequency-dependent selection in bacterial populations. *Philos. Transactions Royal Soc. Lond. B, Biol. Sci.* **319**, 459–472 (1988).

43. Trotter, M. V. & Spencer, H. G. Frequency-dependent selection and the maintenance of genetic variation: exploring the parameter space of the multiallelic pairwise interaction model. *Genet.* **176**, 1729–1740 (2007).

44. Traulsen, A., Claussen, J. C. & Hauert, C. Coevolutionary dynamics in large, but finite populations. *Phys. Rev. E* **74**, 011901 (2006).

45. Ashcroft, P., Traulsen, A. & Galla, T. When the mean is not enough: Calculating fixation time distributions in birth-death processes. *Phys. Rev. E* **92**, 042154, DOI: 10.1103/PhysRevE.92.042154 (2015).

46. Nowak, M. A., Sasaki, A., Taylor, C. & Fudenberg, D. Emergence of cooperation and evolutionary stability in finite populations. *Nat.* **428**, 646–650 (2004).

47. Imhof, L. A. & Nowak, M. A. Evolutionary game dynamics in a wright-fisher process. *J. mathematical biology* **52**, 667–681 (2006).

48. Wright, S. Evolution in mendelian populations. *Genet.* **16**, 97 (1931).

49. Fisher, R. A. Xxi.—on the dominance ratio. *Proc. royal society Edinb.* **42**, 321–341 (1923).

50. Hayashi, H. *et al.* Overall treatment strategy for patients with metastatic nsclc with activating egfr mutations. *Clin. Lung Cancer* **23**, e69–e82 (2022).

51. Chmielecki, J. *et al.* Candidate mechanisms of acquired resistance to first-line osimertinib in egfr-mutated advanced non-small cell lung cancer. *Nat. Commun.* **14**, 1070 (2023).

52. Camidge, D. R., Pao, W. & Sequist, L. V. Acquired resistance to tkis in solid tumours: learning from lung cancer. *Nat. reviews Clin. oncology* **11**, 473–481 (2014).

53. Leale, A. M. & Kassen, R. The emergence, maintenance, and demise of diversity in a spatially variable antibiotic regime. *Evol. Lett.* **2**, 134–143 (2018).

54. Dimitriu, T. *et al.* Negative frequency dependent selection on plasmid carriage and low fitness costs maintain extended spectrum β-lactamases in escherichia coli. *Sci. reports* **9**, 17211 (2019).

55. Stirling, D. R. *et al.* Cellprofiler 4: improvements in speed, utility and usability. *BMC bioinformatics* **22**, 1–11 (2021).

56. Neuts, M. F. *Matrix-geometric solutions in stochastic models: an algorithmic approach* (Dover, 1994).

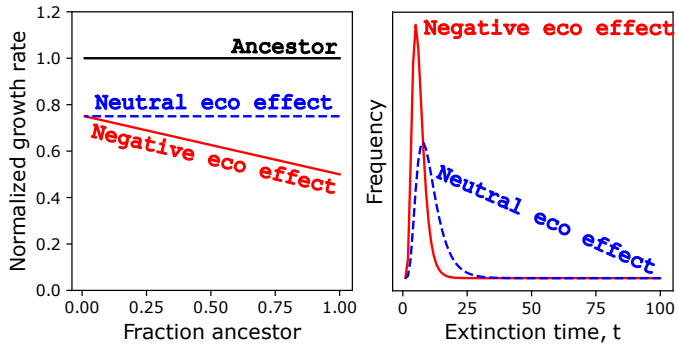

Figure S1. Similar qualitative trends exist when comparing neutral and negative mutants. Closed form extinction time distributions are calculated and visualized for a generalized Moran process (N=100, $f_c = 0.25$). The red distribution results from a mutant with a negative ecological interaction with the ancestor ($f_e = 0.5$), while the blue population has no ecological interaction with the ancestor ($f_e = 1 - fc = 0.75$).

Supplemental Material: Frequency-dependent ecological interactions increase the prevalence and shape the distribution of pre-existing drug resistance

Simulations

Here we provide additional complimentary results from our numerical simulations in the main paper.

Additional generalized Moran process results

Similar to the main text figure, we compare the extinction time distributions of two emerging mutants in an initially ancestor population. Here we compare a neutral mutant with a negative mutant.

While the effect is smaller that the comparison between positive and neutral mutants in the main text, we find similar qualitative trends in that extinction time distribution for neutral mutants is shifted to longer extinction times when compared to the negative mutant (**Fig. S1**).

Analytical theory

The following sections describe the analytical theory supporting the numerical simulations in the main text. Important formulas that are used to fit the simulation results are highlighted by boxes.

Stationary distribution and total number of mutants

We seek to derive an approximate analytical expression for the stationary probability density of mutants $P(f)$ in the Wright-Fisher simulations described in the main text. We assume ecological interactions are present, and that sufficient time has passed for a stationary state to be reached. $P(f)df$ is defined as the fraction of the total population that consists of mutants with instantaneous fitnesses between f and $f + df$. Integrating this distribution gives the mean total number of mutants $N_{\text{mut}}^{\text{eco}} = N \int_0^{f_{\max}} df\, P(f)$, where N is the fixed total population size. For the theoretical calculations we consider an upper bound on the mutant fitness $f_{\max} < 1$ to allow for a well-defined normalization (as explained in more detail below), though we can set $f_{\max}$ arbitrarily close to 1 in order to fit the numerical results.

Every generation of the model consists of a mutation step followed by selection. Let us consider first the mutation part. If $P(f)$ is the current distribution, mutation modifies it to a new distribution,

$$P^m(f) = P(f) + \mu\rho_\alpha(f) - \mu P(f). \tag{S1}$$

Here μ is the mutation probability for a single cell in one generation, and $\rho_\alpha(f)$ is the probability density of a new mutant having fitness f, given a total fraction of mutants $\alpha \equiv \int_0^{f_{\max}} df\, P(f) = N_{\text{mut}}^{\text{eco}}/N$. The second and third terms on the right-hand side in **Eq.** (S1) are respectively the gain and loss due to new mutations.

The dependence of $\rho_\alpha(f)$ on α reflects the role of ecological interactions. In the simplest linear model, a new mutant is assigned a fitness function $f(\alpha) = \alpha f_i + (1-\alpha) f_e$, where f_i (the intrinsic fitness) is randomly drawn from a uniform distribution between 0 and $1-f_c$, and f_e (the ecological fitness) is randomly drawn from a uniform distribution between 0 and $f_{\max}$. Here f_c, where $0 < f_c < 1$, is the cost associated with intrinsic fitnesses, with $1 - f_c < f_{\max}$. For this definition of $f(\alpha)$, the distribution $\rho_\alpha(f)$ is given by

$$\rho_\alpha(f) = \begin{cases} \frac{f}{\alpha(1-\alpha) f_{\max}(1-f_c)} & f < \alpha(1-f_c) \\ \frac{1}{(1-\alpha) f_{\max}} & \alpha(1-f_c) \le f < (1-\alpha) f_{\max} \\ \frac{\alpha(1-f_c)+(1-\alpha) f_{\max} - f}{\alpha(1-\alpha) f_{\max}(1-f_c)} & f \ge (1-\alpha) f_{\max} \end{cases}. \tag{S2}$$

For $0 < \alpha < 1$ the distribution $\rho_\alpha(f)$ has a trapezoidal shape, rising linearly from zero for small f, then plateauing in the middle region, before decreasing linearly to zero at $f_{\max}$. In the two limits $\alpha = 0$ and $\alpha = 1$ it reverts to a uniform distribution between 0 and $f_{\max}$ (for $\alpha = 0$) or between 0 and $1 - f_c$ (for $\alpha = 1$).

The second part of the dynamics is the selection step, which makes a further modification of the mutant distribution, yielding

$$P^s(f) = \frac{f P^m(f)}{1 - \int_0^{f_{\max}} df\, P^m(f) + \int_0^{f_{\max}} df\, f P^m(f)}. \tag{S3}$$

The above form reflects the fact that the ancestors after the mutation step, comprising a fraction $1 - \int_0^{f_{\max}} df\, P^m(f)$ of the total, have fitness 1, and mutants with fitness f have a chance of surviving into the next generation proportional to f. In order for the system to be in a stationary state, the distribution $P^s(f)$ in the next generation must end up being the same as the starting distribution $P(f)$ in the current generation. Using **Eqs.** (S1)-(S3), we can express the condition $P^s(f) = P(f)$ compactly as

$$P(f) = \frac{f\left[(1-\mu)P(f) + \mu \rho_\alpha(f)\right]}{1-\beta}, \tag{S4}$$

where $\beta \equiv \int_0^{f_{\max}} df\, (1-f)\left[(1-\mu)P(f) + \mu\rho_\alpha(f)\right])$. **Eq.** (S4) can be solved for $P(f)$,

$$P(f) = \frac{f \mu \rho_\alpha(f)}{1 - \beta - f(1-\mu)}. \tag{S5}$$

Note that both α and β on the right-hand side of **Eq.** (S5) depend implicitly on $P(f)$. In order for the solution to be self-consistent, we plug **Eq.** (S5) into the definitions of α and β, which leads to a closed system of equations for these two quantities:

$$\alpha = \int_0^{\text{fmax}} df\, \frac{f \mu \rho_\alpha(f)}{1-\beta - f(1-\mu)}, \qquad \beta = \int_0^{\text{fmax}} df\, \frac{(1-\beta)(1-f)\mu\rho_\alpha(f)}{1-\beta-f(1-\mu)}. \tag{S6}$$

The equation for β is satisfied exactly when $\beta = \mu$, using the fact that $\int_0^{f_{\max}} df\, \rho_\alpha(f) = 1$. Plugging $\beta = \mu$ into the α equation, we can carry out the integral analytically, leading to the following relation:

$$\begin{aligned} \alpha = \frac{\mu}{\alpha(\alpha-1)(1-\mu)(1-f_c) f_{\max}} \Bigg[& \alpha f_c \Big(\ln\left[\alpha(f_c - 1) + 1\right] \\ & - \ln\left[\alpha(f_c + f_{\max} - 1) - f_{\max} + 1\right] + (\alpha-1) f_{\max} \Big) - (\alpha - 1) \Big(\ln\left[\alpha(f_c-1)+1\right] \\ & + (f_{\max} - 1) \ln\left[\alpha(f_c + f_{\max} - 1) - f_{\max} + 1\right] + \alpha f_{\max} \Big) \\ & + ((\alpha-1) f_{\max} + 1) \ln\left[(\alpha-1) f_{\max} + 1\right] \Bigg]. \end{aligned} \tag{S7}$$

There is no explicit analytical solution for α from the above equation, but there are ways to derive approximate solutions that work in different limits. We consider two such limits in turn.

Small mutant fractions ($\mu \ll 1$): When $\mu \ll 1$, the mean number of mutants in the population becomes small, and their fraction α, given by the solution of **Eq.** (S7), scales like $\alpha \propto \mu$. In this limit **Eq.** (S7) gives us $\alpha = \mu(-f_{\max}^{-1}\ln(1-f_{\max})-1)+\mathcal{O}(\mu^2)$, so we get an approximate expression for the total mutant fraction (main text Eq. (3)):

$$\boxed{N_{\text{mut}}^{\text{eco}} \approx N\mu\left(-\frac{\ln(1-f_{\max})}{f_{\max}}-1\right) \qquad \text{for } \mu \ll 1.} \tag{S8}$$

To leading order in μ, we can approximate the stationary distribution of mutant fitnesses in **Eq.** (S5) as (main text Eq. (5)):

$$\boxed{P(f) \approx \frac{f\mu}{f_{\max}(1-f)} \qquad \text{for } \mu \ll 1.} \tag{S9}$$

Note that since $\alpha \ll 1$, $\rho_\alpha(f) \approx \rho_0(f)$, and the distribution of instantaneous fitnesses f in **Eq.** (S9) is approximately also the distribution of ecological fitnesses f_e. In order for **Eq.** (S9) to be normalizable, and hence its integral giving **Eq.** (S8) to be well-defined, we need $f_{\max}$ strictly smaller than 1. In the numerical simulations only a finite number of mutants are sampled overall during the course of the evolutionary trajectories, and an effective value of $f_{\max} \approx 0.995$ was found to provide good fits between the theory and simulation results.

Large mutant fractions ($\mu \lesssim 1$): We would like to extend the results above to larger values of μ and α, to cover simulation cases where the mutant fractions α are on the order of 10% of the total. We note that as α gets larger, the distribution of new mutant fitnesses $\rho_\alpha(f)$ in **Eq.** (S2) gets increasingly suppressed at fitnesses near zero and $f_{\max}$. This makes the precise value of $f_{\max}$ less important for determining α, and we can approximate **Eq.** (S7) by taking the $f_{\max} \to 1$ limit. If we then keep the expressions to leading order in μ, assuming that $\alpha \sim \mathcal{O}(\mu)$, **Eq.** (S7) becomes

$$\alpha \approx \frac{\mu}{1-f_c}\big((f_c-1)\ln(\alpha)+f_c\ln f_c\big). \tag{S10}$$

The solution to this equation has the form

$$\alpha \approx \mu W\left(\left(\mu f_c^{f_c/(f_c-1)}\right)^{-1}\right). \tag{S11}$$

Here $W(x)$ is the Lambert W function, which is the solution y of the equation $ye^y = x$. For $x > 0$ (which is the only case that arises in our problem) the function is single-valued (the so-called zero branch of the solution).

While **Eq.** (S11) works in the $f_{\max} \to 1$ limit, ideally we would like an expression that works for $f_{\max}$ close to, but not exactly 1, and for the entire range of $\mu \lesssim 1$ including small μ. Since we know $N_{\text{mut}}^{\text{eco}} = N\alpha$ for $\mu \ll 1$ and $f_{\max} < 1$ from **Eq.** (S8), we posit the following approximate form for $N_{\text{mut}}^{\text{eco}}$ (main text Eq. (4)):

$$\boxed{N_{\text{mut}}^{\text{eco}} \approx \frac{N\mu}{f_{\max}} W\left(\left(1-f_{\max}+\mu f_c^{f_c/(f_c-1)}\right)^{-1}\right).} \tag{S12}$$

By construction this is consistent with **Eq.** (S11) when $f_{\max} \to 1$. When $1-f_{\max}$ is small but nonzero and $\mu \ll 1$, we can use the fact that $W(x)$ diverges like $\ln x$ for large positive x to see that **Eq.** (S12) gives $N_{\text{mut}}^{\text{eco}} \approx -N\mu f_{\max}^{-1}\ln(1-f_{\max})$ for $\mu \to 0$. This recovers the dominant contribution in Eq. (S8) when $1-f_{\max}$ is small.

Thus in principle **Eq.** (S12) should work for a wide range of μ and different values of f_c. **Fig. S2** depicts a comparison of **Eq.** (S12) to simulation results, and the analytical approximation is within 10% of the numerical value across five decades of μ and both small and large f_c (as shown by the errors shown in the inset). For small μ the $N_{\text{mut}}^{\text{eco}}$ curves start out independent of f_c and proportional to μ, as expected from **Eq.** (S8). With increasing μ, the curves bend downwards in a way that depends on f_c, as the population of mutants becomes a non-negligible fraction of the total.

Non-uniform ecological/intrinsic fitness distributions: The derivations above can be easily extended to non-uniform distributions of the ecological and/or intrinsic fitnesses. This modifies the shape of $\rho_\alpha(f)$, depending on the specific distributions from which f_i and f_e are drawn for new mutants. In general, imagine f_e is drawn from a distribution $\rho_0(f_e)$, like the Gaussian example considered in the main text. In the limit $\mu \ll 1$, when the total mutant fraction α is small, **Eq.** (S9) for the distribution becomes (main text Eq. (6)):

$$\boxed{P(f_e) \approx \frac{f_e\mu}{(1-f_e)}\rho_0(f_e) \qquad \text{for } \mu \ll 1.} \tag{S13}$$

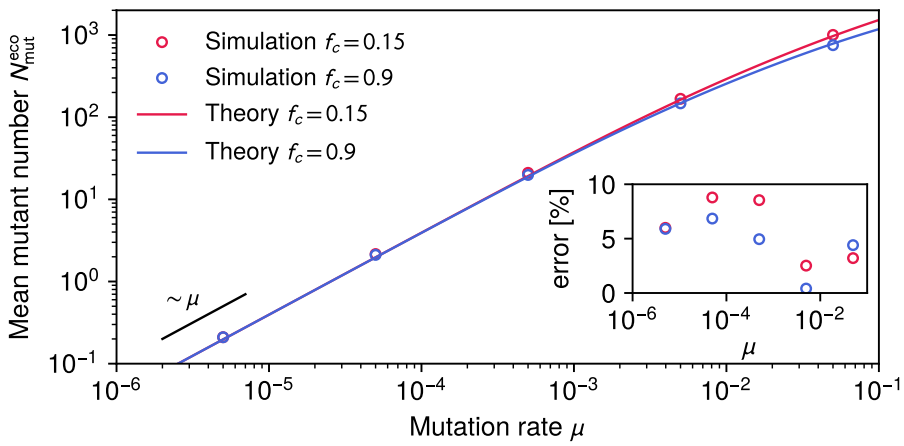

Figure S2. Comparison of the approximate analytical theory, **Eq.** (S12) (curves), and the simulation results (circles) for the mean number of mutants $N_{\rm mut}^{\rm eco}$ as a function of mutation rate μ. We set $f_{\max} = 0.995$, $N = 10^4$, and use two different values for the fitness cost, $f_c = 0.15$ and 0.9 (red and blue respectively). The inset shows the absolute percent error of the analytical approximation with respect to the numerical values. The black line on the left shows the scaling $N_{\rm mut}^{\rm eco} \propto \mu$ expected in the small μ regime from **Eq.** (S8).

Other results from the theory can be similarly generalized.

No ecological interactions: In the absence of ecological interactions, the fitness function $f(\alpha) = f_i$ for all α. If f_i is drawn from a uniform distribution between 0 and $1 - f_c$, the distribution $\rho_\alpha(f)$ becomes uniform: $\rho_\alpha(f) = 1/(1 - f_c)$. An analogous calculation to the one above yields the mean number of mutants in this scenario (main text Eq. (2)):

$$\boxed{N_{\rm mut}^{\rm no\ eco} = N\mu\left(-\frac{\ln f_c}{1-f_c} - 1\right).} \tag{S14}$$

This number is a useful baseline for gauging the relative effectiveness of ecological interactions in enhancing mutant populations.

Mean time to extinction of an individual mutant

The final quantity we would like to calculate theoretically is the mean number of generations that an individual mutant survives after first arising. Note that if a mutant is generated by the mutation step, but does not survive the selection step immediately afterwards, we say its lifetime is zero generations. To simplify the calculation, we consider the regime $\mu \ll 1$, where the chance that a mutant disappears via a second mutation is negligible. And by the definition of the model, the same type of mutant cannot be generated again from either the ancestor or other mutant populations. So in this regime the mutant persists with a randomly fluctuating population until in one of the selection steps none of its population is chosen to survive to the next generation.

Let us focus on a single mutant type with fitness f. If there are ℓ such mutants in the current generation, the probability $W_{k\ell}$ that there will be k mutants of this type in the next generation is given by the binomial distribution characteristic of Wright-Fisher dynamics,

$$W_{k\ell} = \binom{N}{k}\left(1 - \frac{f\ell}{N(1-\beta)}\right)^{N-k}\left(\frac{f\ell}{N(1-\beta)}\right)^{k}. \tag{S15}$$

For $\mu \ll 1$ we can assume that $\beta \ll 1$, $f \approx f_e$ and that $N \gg k$ for any k that has a non-negligible probability, since the number of mutants of a single type at any given time will be a tiny fraction of the total. We can then approximate **Eq.** (S15) as

$$W_{k\ell} \approx \frac{(f_e\ell)^k}{k!}e^{-f_e\ell}, \tag{S16}$$

which is just the limit in which the binomial distribution looks like a Poisson distribution. The probabilities $W_{k\ell}$ can be interpreted as components of an $N \times N$ transition matrix W. Since the vast majority of this matrix will consist of probabilities exponentially close to zero, we can focus on the states $k, \ell = 0, \ldots, M$ for some $M \ll N$. Thus we will consider W instead to be

an $(M+1)\times(M+1)$ matrix, choosing M large enough to get a satisfactory approximation to the non-truncated system. This reduces the problem to an $(M+1)$ state discrete time Markov process.

Note that $W_{k0}=0$, so $\ell=0$ (extinction) is an absorbing state. From the theory of phase-type distributions[56], the mean number of generations to extinction can be calculated using the $M\times M$ submatrix S of the tranpose W^T, defined via $S_{k\ell}=W_{\ell k}$ for $k,\ell=1,\ldots,M$. Starting from a population of 1 (after the mutation step) at generation zero, the mean time to extinction is given by

$$\tau(f_e)=\mathbf{z}^T(I-S)^{-1}\mathbf{e}-1. \tag{S17}$$

Here I is the $M\times M$ identity matrix, $\mathbf{z}$ is an M-dimensional vector with a 1 in the first element and zero elsewhere, and $\mathbf{e}$ is an M-dimensional vector with 1 for all its elements. The -1 at the end of **Eq.** (S17) is due to the counting convention where extinction during the first selection step is considered to be an extinction time of zero.

For a given choice of M, it turns out the matrix inverse in **Eq.** (S17) can be calculated analytically. Even though the resulting expression becomes unwieldy for large M, it can always be Taylor expanded around $f_e=0$ to give relatively simple results. The Taylor coefficient of order f_e^n in the expansion remains unchanged for any choice of $M\geq n$. Thus we can find the Taylor expansion of $\tau(f_e)$ in the non-truncated system up to any chosen order M, simply by Taylor expanding **Eq.** (S17).

The first few terms of this Taylor expansion are:

$$\tau(f_e)=f_e+\frac{f_e^2}{2}+\frac{2f_e^3}{3}+\frac{f_e^4}{8}+\frac{19f_e^5}{30}+\frac{f_e^6}{144}+\frac{107f_e^7}{280}+\cdots. \tag{S18}$$

While the lowest terms are sufficient to describe the mean extinction time of mutants with $f_e\ll 1$, progressively more terms are required to approximate $\tau(f_e)$ as f_e approaches 1 from below. In fact, technically in this approximation the series $\tau(f_e)$ diverges at $f_e=1$, since we have effectively taken $N\to\infty$ in **Eq.** (S16). In practice this is not a problem since we only consider mutants with fitnesses up to $f_{\max}<1$. Rather than working with the Taylor expansion directly, we constructed an analytical approximation designed to agree with the expansion through order f_e^3, and capture the divergence at large f_e (main text Eq. (1)):

$$\boxed{\tau(f_e)\approx\frac{3\ln(1-f_e)}{f_e^2-3}.} \tag{S19}$$

Despite its simple form, this approximation agrees with the simulation results for $\mu\ll 1$ across the whole range of f_e with a typical error of 5%.