The adaptive state determines the impact of mutations on evolving populations

Malgorzata Tyczynska Weh[1,2], Pragya Kumar[2,3], Viktoriya Marusyk[3], Andriy Marusyk[3,†,*], David Basanta[1,‡,*]

1: Department of Integrated Mathematical Oncology, H. Lee Moffitt Cancer Center & Research Institute, Tampa, FL, 33612, USA
2: Cancer Biology Ph.D. Program, University of South Florida, Tampa, FL, 33612, USA
3: Department of Tumor Microenvironment and Metastasis, H. Lee Moffitt Cancer Center and Research Institute, Tampa, FL 33612, USA

† andriy.marusyk@moffitt.org
‡ david@cancerevo.org

* Contributed equally

Abstract

Darwinian evolution results from an interplay between stochastic diversification of heritable phenotypes, impacting the chance of survival and reproduction, and fitness-based selection. The ability of populations to evolve and adapt to environmental changes depends on rates of mutational diversification and the distribution of fitness effects of random mutations. In turn, the distribution of fitness effects of stochastic mutations can be expected to depend on the adaptive state of a population. To systematically study the impact of the interplay between the adaptive state of a population on the ability of asexual populations to adapt, we used a spatial agent-based model of a neoplastic population adapting to a selection pressure of continuous exposure to targeted therapy. We found favorable mutations were overrepresented at the extinction bottleneck but depleted at the adaptive peak. The model-based predictions were tested using an experimental cancer model of an evolution of resistance to a targeted therapy. Consistent with the model's prediction, we found that enhancement of the mutation rate was highly beneficial under therapy but moderately detrimental under the baseline conditions. Our results highlight the importance of considering population fitness in evaluating the fitness distribution of random mutations and support the potential therapeutic utility of restricting mutational variability.

Introduction

Populations adapt to environmental selection pressures through natural selection for phenotypic variants with higher heritable fitness. Therefore, mutational diversification, the main process responsible for the generation of heritable phenotypic diversity, is essential for creating the substrate on which selection can operate. For asexual unicellular organisms or neoplastic populations of cancer cells, a random mutation can lead to three potential outcomes at the individual cell level: fitness increase, reduction, or no change[1,2]. The distribution of these fitness effects (DFE), i.e., the pool of mutations in a population before selection, can thus impact

populations' adaptation potential. This impact can be further influenced by the adaptive state of the population, originally conceptualized by Sewall Wright with the adaptive landscape metaphor[3]. For example, when a population is near extinction owing to a rapid environmental change or an introduction of new therapy-imposed selection pressure, the downside of negative mutations could be limited. Thus, higher mutation rates could enhance the ability of the population to adapt (Figure 1). In contrast, when a population is well-adapted, i.e., positioned at the peak of the adaptive landscape, the upside of positive mutation is negated, enhancing the contribution of negative mutations. Thus, increased mutation rates might reduce population fitness[4]. Critically high mutational frequencies can lead to population extinction (mutational meltdown)[5–7] for poorly adapted and well-adapted populations. The dependencies of the mutational impact on the adaptive state of the population have been previously hypothesized[8,9] but never proven.

Empirical inferences of the DFEs suggest mostly negative[10–14] or neutral[15–18] fitness effects of new mutations. However, these inferences have been largely limited to well-adapted populations, thus leading to a potential bias. Indeed, experimental studies demonstrate that the fitness effects of specific mutations vary with the environment[19,4] and have fitness costs[20,21]. To assess the interplay between mutation rates, DFE, and population adaptability, we used Agent-Based Modeling (ABM), a stochastic model of individual cells obeying specific actions and resulting in emergent collective behavior. Since space availability represents one of the key ecological resources that can substantially impact evolution, we chose to study the interplay with spatial ABM, where agents are restricted to a 2D grid domain.

Rather than studying a highly abstract population, we decided to focus on a specific scenario of a neoplastic population that faces a new therapeutic pressure that puts the population near an extinction threshold, such as an initiation of an initially effective targeted therapy. Cancer's adaptation to therapeutic pressures might reflect a competitive release of rare subpopulations fully resistant to therapies[22]. However, a growing body of evidence, including our work, indicates that, in clinical contexts with strong and durable initial remissions, therapy resistance evolves within populations of therapy-persisters[23–27], whose fitness, initially very close to the extinction threshold, can be improved with mutational changes that reduce therapeutic sensitivity.

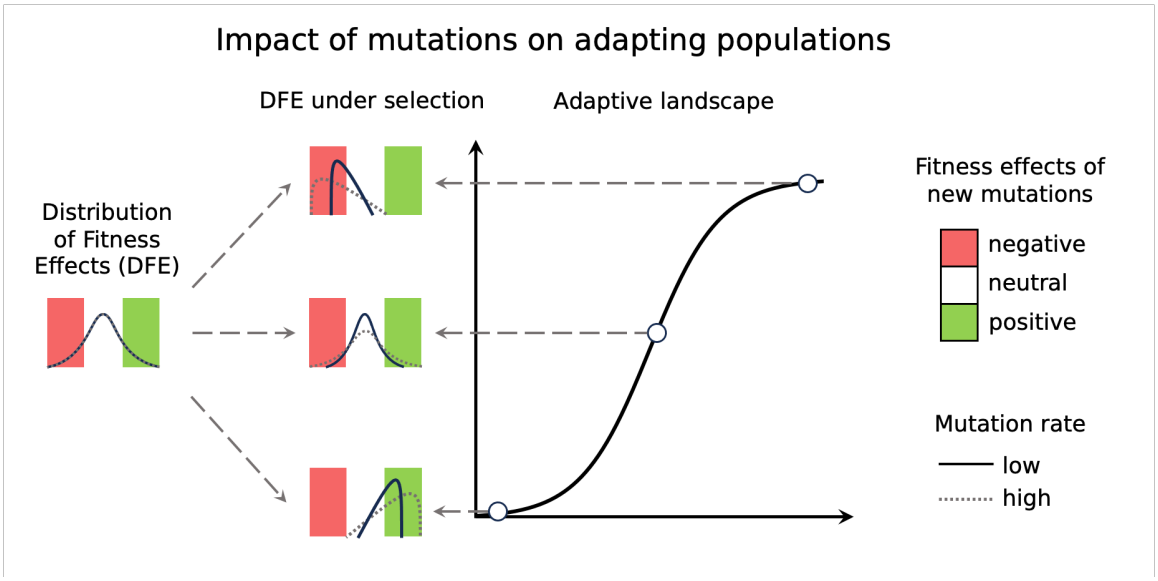

Figure 1: the hypothetical effect of mutations on adapting populations. The pool of mutations generated by the DFE is subject to selection, and the population's adaptive state can determine whether this selection results in positive, neutral, or negative changes. At the top of the adaptive landscape, mutations can stabilize or decrease population fitness. At the bottom, mutations can only benefit the populations. Mutation rates emphasize these effects. Consequently, a low mutation rate can be more beneficial for the population at the top of the adaptive landscape, whereas a high mutation rate — at the bottom.

For the mathematical analysis, we first constructed the ABM capable of simulating emerging adaptations due to the collective behavior of individuals acquiring mutations with various fitness effects. Next, we applied the model to study the effects of varied mutation rates on the scenario of persister cells adapting to treatment. This showed that high mutation rates can accelerate adaptations but result in substantial fitness costs long-term. Then, we extended our analytical efforts to systematically investigate the effects of mutations on adapting populations under various mutation rates, adaptive states, and DFEs. These efforts demonstrated that the adaptive state of the population determines whether the mutations will result in adaptive evolution. The mutation rates and DFEs modulate these adaptive effects.

For experimental validation of the key inferences, we assessed the dependence of the enhancement of mutation rates on the population fitness using an experimental model of ALK+ non-small cell lung cancers, the H3122 cell line. Specifically, we contrasted the impact of chemically enhanced mutagenesis between cells grown under optimal cell culture conditions versus cells exposed to near-extinction by the exposure to clinically relevant concentrations of different ALK inhibitors (ALKi). Consistent with the model's predictions, we found that the chemical mutagen (EMS), commonly used in experimental studies, enhances population fitness under exposure to the ALK, while moderately reducing it under optimal growth conditions. Whereas our simulations were focused on a specific context of neoplastic adaptation, our inferences should apply to other contexts of adaptation of asexual unicellular organisms, including the adaptation of bacteria to antibiotics. Our results reveal a complex, context-dependent relationship between

mutation rates and adaptation, highlighting the importance of considering a population's adaptive state when studying its evolution.

The mathematical model captures the impact of mutations on adapting populations

To understand the impact of elevated mutation rates during adaptation to treatment, we developed an Agent-Based Model (ABM): a stochastic model of agents governing emergent population behavior over time. In our model, individual agents—cells—are represented as a grid point in a two-dimensional space (Fig. 2a) and follow a set of possible actions representing a cell cycle: the cell can die, divide, and mutate with specific probabilities (rates). These actions are ordered as follows (Fig. 2b): if a cell doesn't die, it can mutate independently of cell replication; then, if there is available space, divide and acquire a replication-dependent mutation (Methods). Because it is unknown how much mutation rates vary during the cell cycle, we set one rate for both the replication-dependent and independent mutation (p_{mut}). p_{mut} is equal and constant for all cells, but cell division (p_{div}) and death (p_{die}) rates are cell-specific and can change if the cell is affected by a mutation. Because our ABM is spatial, the availability of unoccupied space on the grid can affect the population's growth, especially when population size reaches carrying capacity, i.e., the maximal number of individuals on the grid. Spatial effects and stochasticity are important for population dynamics in the absence of mutations; since these cannot impact individual cells' division or death rates, mutations are the only source of diversification in the model.

Next, we derived a metric to capture the initial population fitness, approximating the population's adaptive state: $W_0 = p_{div} - p_{die}/1-p_{die}$ (Supp. Material 1). Since W_0 is accurate if all cells are identical, we initialize all cells to have equal p_{div} and p_{die} values. $W_0 = 0$ results in approximately constant population over time (Fig. 2c), $W_0 < 0$ leads to extinction, and $W_0 > 0$ results in growth. W_0 should not be applied to capture the adaptive state of evolved populations due to stochasticity and emergent population heterogeneity. Without population turnover ($p_{die} = 0$), $W_0 = p_{div}$.

In our ABM, mutations change the individual cell fitness ($w := p_{div} - p_{die}$) by affecting either division or death probability with a quantity determined by a Mutation-Fitness Algorithm (MFA). This quantity is stochastic and generated from a specific distribution of fitness effects. The MFA results in an immediate and heritable fitness change in individual cells; the algorithm is sketched in Fig. 2d and proceeds as follows (see Methods for details and Supp. Mat. 1 for mathematical analysis): First, we compute the mutation's impact on cell fitness (Step 1 on Fig. 2d). To reflect the mutation's stochastic nature, we sample a random number $x_t \sim X$; depending on the value of x_t, the impact is negative, neutral, or positive. We compute the magnitude of this impact by linearly transforming x_t to y_t. Next (Fig. 2d step 2), depending on whether the mutation is positive or negative, we add or subtract the y_t to/from p_{div} or p_{die}; the probability of mutation changing either p_{div} or p_{die} is 50%. The new values of p_{div} or p_{die} and the number of mutations are inherited during cell division. There is no fitness penalty associated with the number of accumulated mutations.

We run a subset of initial simulations visually tracking evolution to ensure that our ABM can capture emerging adaptation. For example, the snapshots from a simulation with p_{mut} = 1e-3 (Fig. 2e) demonstrate population-level adaptations emerging from the gradual fitness increase. In addition, the set of Supplementary Movies 1 shows that the ABM model can capture either population extinction or adaptation when subject to elevated mutation rates. Hence, our ABM provides a platform for studying many evolutionary trajectories.

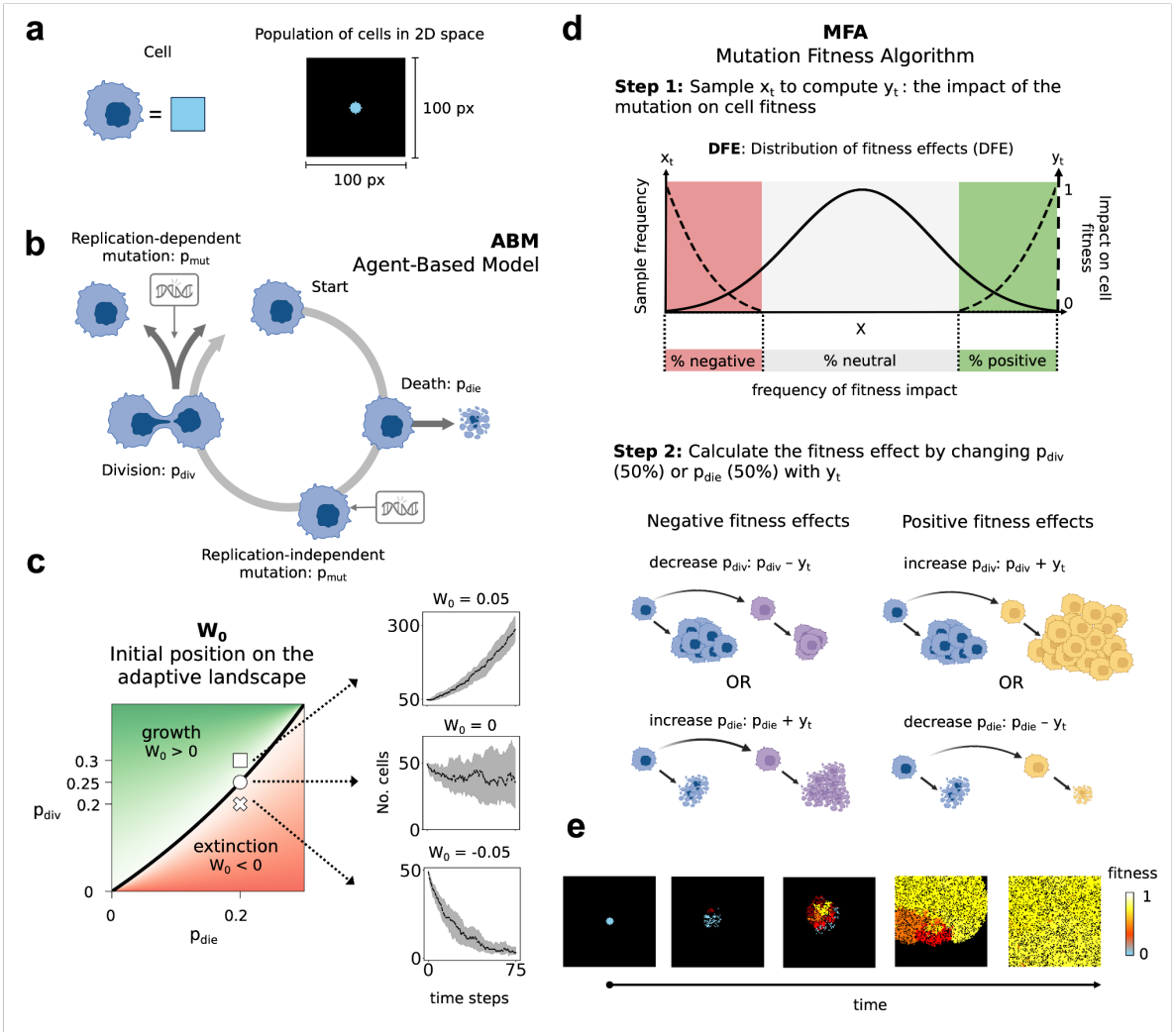

Figure 2: The *in silico* model to capture the impact of elevated mutation rates on adapting populations. a: representation of a cell as a pixel on the 100 by 100 px 2D grid. All simulations begin with cells clustered in the middle of the domain. **b**: the Agent-Based Model (ABM) flow chart. The actions are repeated for all cells. **c**: The initial adaptive state of the population (W_0) in the p_{div}, p_{die} space. The thick black line represents W_0 = 0 and separates the (p_{die}, p_{div}) combinations leading to population extinction (red) or growth (green). The big, white square, circle, and cross represent W_0 = 0.05, W_0 = 0, and W_0 = -0.05 values, respectively, and, when stimulated, result in growing, approximately constant, and decreasing populations (right); the gray area symbolizes 25-75% quantiles and the black line in the middle – median of 50 simulations. **d**: Visual representation of the Mutation Fitness Algorithm (MFA). Neutral mutations cause no change in fitness. **e**: snapshots of the ABM over time simulated with W_0 = 0 and p_{mut} = 1e-1.

At the edge of extinction, higher mutation rates facilitate adaptation at the expense of decreasing maximal population fitness

We first apply our model to investigate how mutations impact adaptations of persister cells to treatment. The stress caused by treatment reduces the population to a small number of persister cells that grow with approximately zero net growth rate. Their rate then increases as they become adapted to treatment. To model this scenario, we initialize the simulations with 49 cells, representing a total of 0.49% of the carrying capacity of 10000 cells. We then set their initial adaptive state to $W_0 = 0$ (i.e., zero net growth) with $p_{div} = 0.25$ and $p_{die} = 0.2$. We also set $p_{die\ min} = 0.196$; this ensures a high cell turnover and is inspired by experimentally derived estimates of tumor death rates[28,29]. Because $W_0 = 0$ corresponds to the fitness of persister cell populations under treatment, we do not model treatment explicitly.

While many potential DFEs can govern adaptations in persister cell populations, we decided to initialize the MFA with the DFE of a simple case of a distribution having equal, small chances of negative and positive fitness effects[15]. This choice results in the frequency of neutral mutations as 99% and 0.5% of negative and 0.5% of positive mutations (see Supp. Mat. 1 and Methods). We initialized each ABM with mutation rates $p_{mut} = 0$, or in a range 1e-4–1 (see Methods for the statistical details).

Consistent with the expectation of mutations enhancing the ability of populations to adapt at the edge of extinction, an increase in the mutation rate prevents population extinction (Fig. 3a) and accelerates adaptation (Fib 3b). On the other hand, the steady-state values of population fitness (Fig. 3c-d) and population numbers (Supp. Fig. 5a) decrease with increasing mutation rates, suggesting potential fitness costs of high mutation rates. To measure how adaptive the evolution is, we derived a metric Max Δw: maximum average change in fitness (Fig. 3e and Methods). Max $\Delta w > 0$ indicates fitness increase leading to adaptation, and Max $\Delta w < 0$—population fitness decline. As expected, Max Δw increases with the increase in mutation rate (Fig. 3f, Supp. Fig. 3b).

To investigate the causes of the fitness costs, we analyzed the number of accumulated positive and negative mutations (Fig. 3g). While we observe a substantial accumulation of positive mutations for all mutation rates, the average number of negative mutations is several orders of magnitude smaller. Despite this, populations with high mutation rates accumulate negative mutations, suggesting that the fitness costs emerge from the total load of negative mutations (Supplementary Movies 2). The number of positive mutations acquired at the time of Max Δw approximates how many mutations are required for adaptive evolution (Supp. Fig. 3b). Our analysis shows that regardless of the mutation rate, 1-3 mutations are necessary for adaptations, which matches the estimated number of mutational events contributing to tumor expansion[30] or the emergence of antibiotic resistance[31,32,33] from the literature.

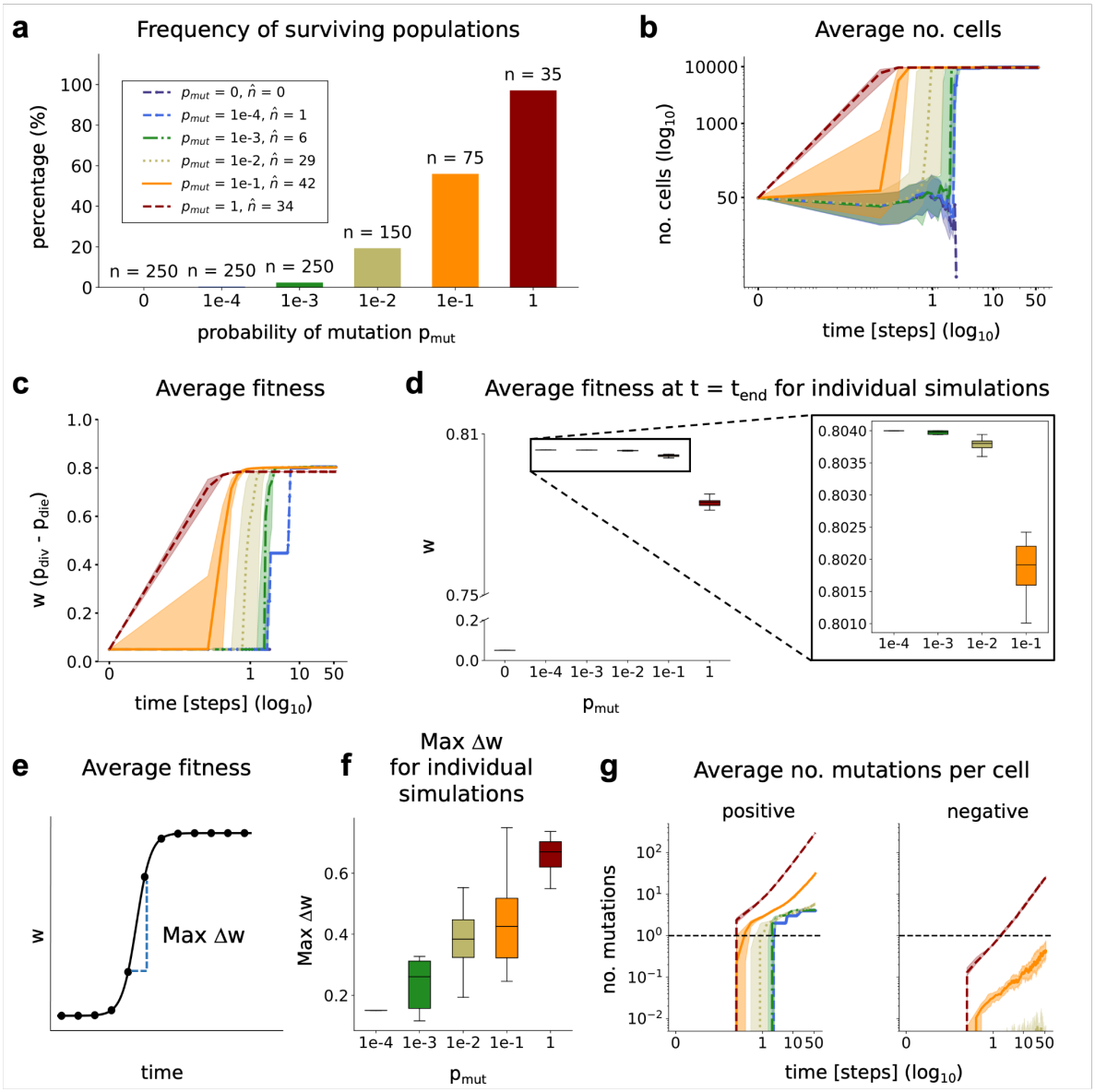

Figure 3: The qualitative predictions of the impact of the elevated mutation rates on the persister cell adaptations to treatment. The persister cell state under treatment is represented by a small population number (49 cells) and zero net growth rate (W_0 = 0). **a**: the percentage of replicates surviving until the end of the simulation ($\hat{n}$; t_{end} = 40000 time steps) out of (n) simulations. **b**: the average population number over time ($\log_{10}$-$\log_{10}$). **c**: the average fitness over log10 time, and **d**: the corresponding distribution of the median average fitness at the end of the simulation (t_{end}) across $\hat{n}$ replicates (left), with a focus on the p_{mut} = 1e-4 — 1e-1 (right). The differences between distributions with $\hat{n}$ > 5 are significant on the 1% level (two-way KS statistics; see supplementary table 2). **e**: the average number of positive (left) and negative (right) on the $\log_{10}$-$\log_{10}$ scale. The color scheme for all graphs follows **a**. **f**: Illustrative example of the Max Δw. **g**: Max Δw for the individual simulations over p_{mut}. For **b**, **c**, and **e**, the thick lines represent the median average across $\hat{n}$ simulations, and the colored areas are the corresponding 25-75% quantiles. For all graphs, we scaled the time step unit to hypothetical years, assuming 1 time step = ½ day.

The impact of mutations on adaptive evolution depends on the adaptive state of the population

Having validated the intuitive prediction that mutations benefit populations at the edge of extinction but reduce their fitness in the long term, we sought to investigate the impact of mutations on evolving populations systematically. Initially, we assumed a symmetric DFE distribution with rare deleterious fitness effects. This is likely wrong because prior studies suggest mostly neutral or deleterious fitness effects of new mutations[10-12, 15-17]. Thus, we considered a variety of distributions which we further divided into two groups (methods) (Fig. 4a): the first group is made of distributions with equal frequencies of positive and negative effects. Thus, it is unbiased about either negative or positive fitness effects. The second group comprises distributions biased towards deleterious fitness effects; we refer to these groups as “unbiased” and “biased.”; we have not investigated DFEs with a bias towards positive fitness effects due to the lack of such DFEs in the literature. Methods and Supp. Mat. 1 provides a mathematically precise definition of all the distributions.

In the unbiased group, lower frequencies of neutral mutations lead to an increased pool of positive and negative mutations. Because of selection, we expect an increased speed of evolution with a decreasing frequency of neutral mutations. Thus, we set the distributions from the unbiased group to have neutral fitness effects of 99%, 80%, 33%, and 0% (corresponding to 0.5%, 10%, 33%, and 50% of positive/negative effects, respectively).

The group of DFEs with a bias towards deleterious fitness effects reduces the pool of mutations leading to adaptive changes. Since the chances of successful adaptations can depend on 1) the variety of negative fitness effects and 2) the frequency of the positive and neutral mutations, we further divided the biased DFEs into two groups: one having a major bias towards very deleterious mutations (labeled as “Major”) and one less biased, allowing for a large variety of negative fitness effects (labeled as “Minor”). The distributions from the biased group always have 40% negative effects but can have either 40% neutral effects (i.e., 20% positive effects) or 0% neutral effects (i.e., 60% positive effects).

Since we speculate that the mutation's impact on evolving populations varies with the adaptive state of the population, we also vary W_0 (recall Fig. 2c). We keep the high population turnover by setting p_{die} = 0.2 and vary the initial p_{div} to model different degrees of adaptation (Fig. 4b): population extinction (W_0 = -0.15), zero net growth (W_0 = 0), intermediate population growth (W_0 = 0.1, 0.45) and maximum possible growth, representing the top position on the adaptive landscape (W_0 = 0.75).

We again apply the Max Δw to test whether the evolution is adaptive in all 240 sets of simulations (six different p_{mut}, eight different DEFs, and five W_0 values), each initialized with 49 cells. In addition to each Max Δw measurement, we record the trajectories of the average population sizes, fitness, number of positive/negative mutations, and more; we provide an example of illustrative

cases in the Supp. Fig. 5. Finally, Methods and Supp. Fig. 4 assesses the required number of simulation replicates for statistical significance.

Fig. 4b demonstrates that the impact of mutations on adaptive evolutions varies with the adaptive state of the population. At the bottom of the adaptive landscape (W_0 = -0.15, 0), higher mutation rates increase the odds of adaptation. For better-adapted populations (W_0 = 0.1, 0.45), mutations still lead to adaptive evolution, but the lower magnitude of Max Δw implies a less dramatic effect. At the top of the adaptive landscape (W_0 = 0.75), frequent mutations result in maladaptive changes (Max Δw < 0); only populations with infrequent or absent mutations can maintain high initial fitness (Max Δw ~ 0). These results indicate that mutations determine the effects of mutation based on the adaptive state of the population.

Regardless of the DFE, the graphs of Max Δw over the adaptive landscape are more aligned to the W_0 = 0 line as the mutation rate decreases; p_{mut} = 0 all Max Δw are 0 if W_0 > 0 for all DFEs (Supp. Fig. 8). This behavior is not unexpected: a lower rate of diversification reduces the pool of mutations that the selection can act upon and, consequently — the reduced propagation of adaptive/maladaptive effects of mutations.

Changes in the DFE can cause variation in Max Δw for a given mutation rate and adaptive state: for the unbiased DFE, p_{mut} = 1e-1, the Max Δw is maximized at 80% of neutral mutations for all $W_0 \neq 0.75$, and for p_{mut} = 1e-2 under the same conditions - at 0% of neutral mutations. For p_{mut} = 1e-1 and p_{mut} = 1 (Supp. Fig. 8), we observe mutational meltdowns for some, but not all forms of DFEs: for the biased DFEs at p_{mut} = 1, the meltdown happens for all DFEs except the DFE with a minor bias towards negative fitness effects and the 0% of neutral mutations. Collectively, this shows that the relationship between the DFE and the potential for adaptive evolution is nonlinear: how the specific distribution will dictate evolution is difficult to predict, even if the mutation rate and the adaptive state of the populations are known.

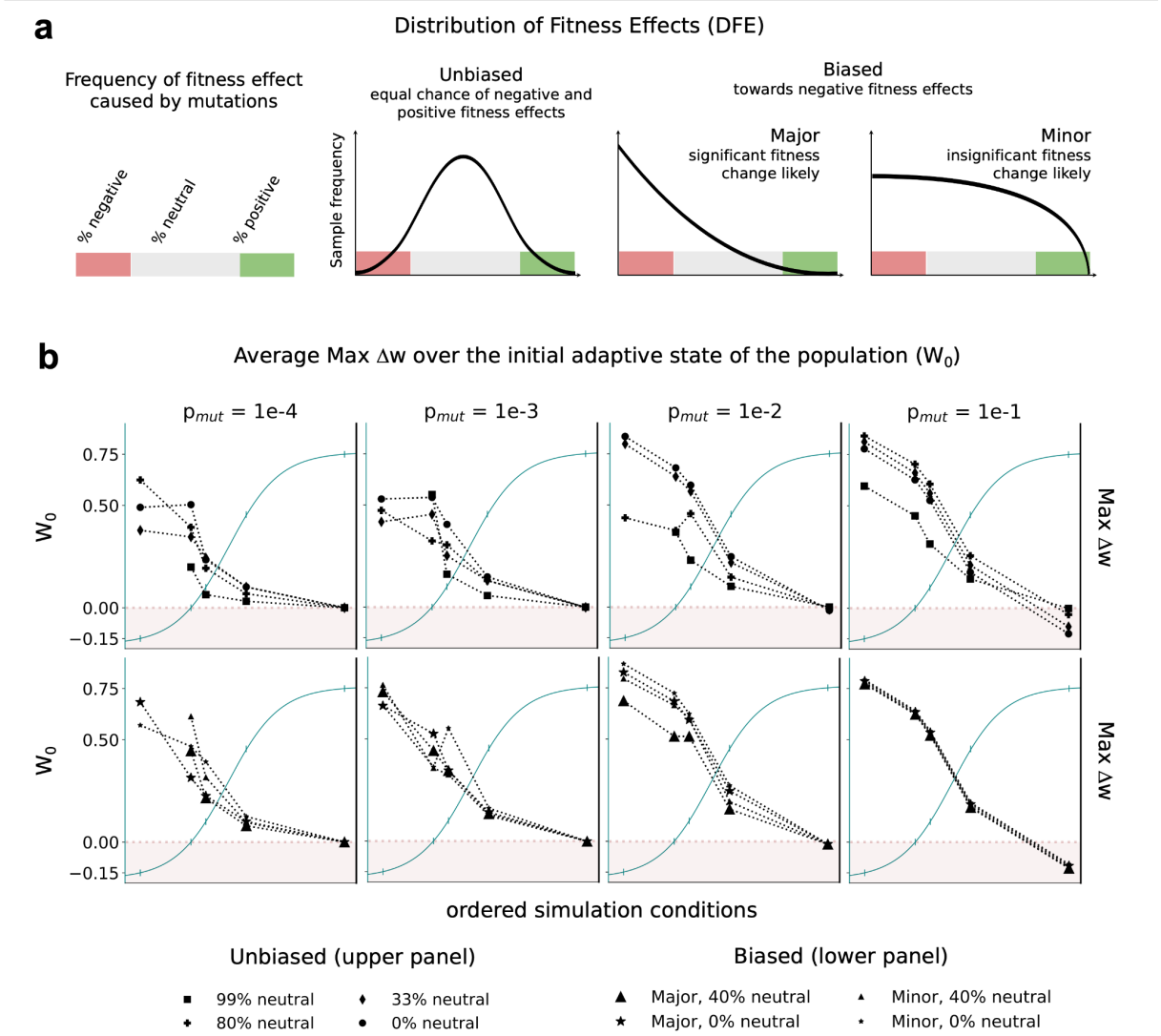

Figure 4: The impact of elevated mutation rates changes with the position on the adaptive landscape and the choice of the distribution of fitness effects.
a: a visual representation of the groups of DFEs. **b**: Max Δw (right y-axis, thick, black) of *n* simulations initialized and ordered relative to the adaptive landscape. The adaptive landscape is represented by W_0 (left y-axis, thin, teal). The Max Δw and W_0 axes share the same numeric values. The black symbols represent a specific choice of the DFE. The W_0 (teal, thin marks) values for simulations are obtained with p_{die} = 0.2, and p_{div} = (0.1, 0.25, 0.35, 0.7, 1). For each combination of W_0 and DFE, the *n* is 35 (p_{mut} = 1), 75 (p_{mut} = 1e-1), 150 (p_{mut} = 1e-2), or 250 (p_{mut} = 1e-3, 1e-4, 0). Some graphs are missing because of population extinction before the end of the simulation. The graphs for the extreme conditions (p_{mut} = 0 and 1) are displayed in Supp. Fig. 7.

Biological experiments qualitatively support the mathematical model predictions about the impact of mutations during adaptations to treatment

To empirically test the theoretical findings of the dependence of the impact of enhanced mutation rates on population fitness, we decided to contrast the impact of enhanced mutations between optimal growth conditions and therapy-induced selection bottleneck under an experimental context that closely mirrors the scenario explored in our in silico simulations. To this end, we used an experimental cancer model, the H3122 cell line, a commonly used model to study acquired therapy resistance in targetable ALK+ lung cancers. This established cell line is characterized by a stable, high growth rate under standard in vitro cell culture conditions. Exposure to clinically relevant ALKi concentrations eliminates the majority of therapy-naive cells; the surviving minor fraction initially displays a persistence phenotype, i.e., strong suppression of cell proliferation and near zero net growth rates[25]. Therapy-naive cells do not contain subpopulations with pre-existent resistance. Instead, the populations of persisters gradually adapt to the drugs by accumulating multiple resistance-promoting mutational and expression-level changes, eventually approaching the baseline growth rates of the therapy-naive H3122 cells[25]. Based on the in silico results, we expected that elevation of stochastic mutation rates, induced by continuous exposure to non-toxic concentrations of a chemical mutagen EMS, should enhance the speed of adaptation of therapy-exposed cells while having either no effect or cause a modest reduction of population fitness of cells under the baseline growth condition (Fig. 5A, left). To detect the impact of the enhanced mutation rates, we adopted a clonogenic assay used in our earlier study[25], where fitness differences could be captured by comparing the clonogenic fraction and the size of colonies (Fig. 5a, right)[25].

Consistent with the theoretical predictions, EMS exposure led to a slight reduction of clonogenic survival of H3122 cells and the overall population size (combined area of all colonies); while the distribution of colony sizes trended smaller under EMS treatment, the difference was not statistically significant (Fig. 5b-d, Supp. Fig. 8). As expected, treatment with three different ALKi (alectinib, lorlatinib, and crizotinib) reduced the clonogenic fraction and inhibited the growth of clonogenic cells (reflected in colony sizes), compared to the DMSO control (Figure 5b-d, Supp. Fig. 8). At the same time, clonogenic fraction, total population size, and size of individual colonies were significantly larger under the EMS exposure, indicating enhanced adaptability of the populations under therapeutic bottlenecks (Fig. 5b-d, Supp. Fig. 8). In summary, experimental studies supported the in silico inferences, indicating that the impact of enhanced mutation rates is dependent on the state of the population fitness.

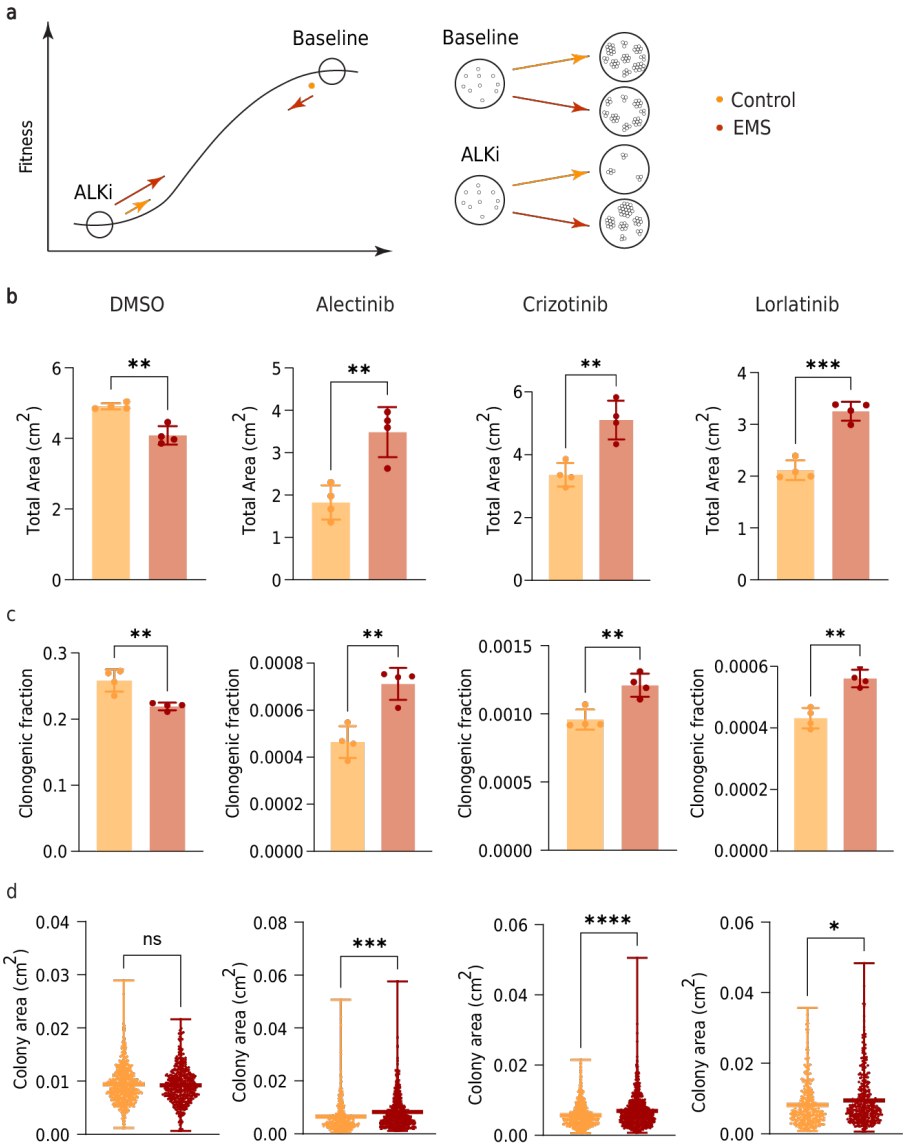

Fig. 5: *In vitro* model of the mutagen effect on the persister cells' adaptation to treatment. a: A schemata of the experimental design. H3122 cells were treated with DMSO (baseline) or ALKi, Alectinib (0.5μM), Crizotinib (0.5μM) or Lorlatinib (2μM), alone or in combination with EMS (1μM). For the EMS groups, cells were pre-treated with EMS for 48 hours. Cell colonies were analyzed. **b:** Total area covered by the cells. **c:** Clonogenic fraction (number of colonies as fraction of cells plated). **d:** Area of each colony. Two-tailed t-test was performed. p-value – 0.1234 (ns), 0.0332 (*), 0.0021 (**), 0.0002 (***), <0.0001 (****).

Discussion

Our computational and experimental studies demonstrate a fundamental principle: the fitness effects of mutations are not fixed but rather shaped by a population's adaptive state. This finding validates and extends beyond the intuitive expectation that mutation outcomes should vary based on a population's position in the adaptive landscape. We found favorable mutations were overrepresented when populations were near extinction due to therapeutic pressure but depleted when populations were well-adapted. At this extinction threshold, where only a small fraction of cells survived therapy, positive mutations had outsized effects while negative mutations were largely masked, accelerating the pace of adaptation to treatment. The effect is reversed for populations approaching a fitness peak, leading to a moderate reduction of population fitness at high mutation rates by suppression of purifying selection. Our findings highlight the importance of considering the adaptive state of populations when interpreting experimental DEF inferences. While the impact of the adaptive state of populations can profoundly impact the evolutionary consequences of random mutations[8,9,34], this dependence is often overlooked. While our study focused on neoplastic populations exposed to therapy, our conclusions should be generalizable to other contexts involving asexual populations, such as yeasts and microbes.

An accurate understanding of the effects of mutation rates on evolutionary dynamics requires adequate quantitative knowledge of the DFE. However, empirical inferences of DFEs are complicated by the strong dependence of observed DFE on the adaptive state of populations. Most random mutations can be expected to reduce fitness or be fitness-neutral, as disrupting complex molecular processes is easier than improving them[35]. This intuition is supported by multiple empirical studies, revealing a strong bias towards neutral or mildly deleterious effects of random mutagenesis[10-18]. However, intuition is unsuitable for quantitative inferences, and most experimental studies have operated with well-adapted populations near or at the fitness peak[11,13,14,17]. This scenario facilitates well-controlled inferences but limits the ability to detect the effect of positive mutations. Similarly, computational statistical inferences such as mutational analyses of cancer genomes[10,1518] are likely confounded by a reflection of relatively long periods of relative stasis with tumor expansion occurring at a local fitness plateau.

The use of spatial ABM enabled us to sidestep some of the significant limitations of both experimental and statistical inferences of DFE. ABM-based studies enable simultaneous tracking of the fitness of individuals, mutational occurrences, and their effects. This approach allows for resolution unachievable by the population-genetics-based mathematical models of fitness effects of new mutations, such as the Gillespie mutational landscape model[36] and Fisher's geometric model[37]. Importantly, our studies highlight the importance of considering space explicitly: despite having a division probability that significantly exceeds the probability of death (Fig. 2c), populations exhibited near-zero growth due to the lack of free adjacent space available for cell division. While the importance of space is often overlooked, space represents a critical ecological resource[7] that profoundly affects eco-evolutionary dynamics in contexts of cellular populations with relatively low motility. Consequently, a lack of spatial considerations can lead to biased inferences and misunderstanding of the drivers of the evolutionary dynamics.

Our findings have important therapeutic implications. The evolutionary tradeoff of enhanced mutational variability is reflected in the phenomenon of stress-induced mutagenesis, i.e., the elevation of mutation rate under therapeutic and environmental stress[38-40], with a return to baseline once populations have adapted to the stressor. Consequently, therapeutic targeting of mutational processes can be beneficial when cancers are most responsive to cytotoxic and targeted therapies while becoming less effective or even counter-effective after the evolutionary rescue.

Like any modeling study, our work is a simplification of reality. Our analyses of the adaptive state of the populations are limited by the lack of precise knowledge about the underlying Distribution of Fitness Effects (DEFs). Different types of mutational processes can have vastly different DEFs, which we have not considered in our simulations. For example, compared to nucleotide level substitutions and small insert/deletions, the effects of large-scale chromosomal aberrations are likely skewed to detrimental effects. Our study did not consider the epistatic interactions between individual mutations, spatial habitats with different phenotypic optima, or interactions between distinct populations. In neoplastic populations, plasticity-mediated non-mutational changes in gene expression patterns might overshadow the effects of mutational variability. Despite the limitations, our work provides empirical support for an intuitive notion that the adaptive state of a population shapes how mutations affect its evolution. Explicit consideration of the adaptive state of a population can improve the efficiency of therapeutic approaches directed against mutational processes while reducing the potential side effects.

Methods

The Agent-Based Model

Grid definition & Agent Flowchart

An agent represents a single cell as a pixel in a 2D grid. A grid is an assembly of pixels of a specific size representing a tissue. The model is on-lattice, i.e., cell positions are discretized and correspond to specific pixel coordinates. One agent can occupy at most one pixel and neighborhoods are calculated using Moore approximation. We use a 100 x 100 px domain for all simulations, and each simulation is initialized with 49 cells assembled in a circle at the center of the domain. This ABM is implemented using the HAL library in Java [41], and the actions of the ABM are iterated over time and for all alive cells.

The flow chart of our ABM, illustrated in Fig. 1A., proceeds as follows: First, a cell dies if a $p_1 < p_{die}$, where $p_1 \sim U(0,1)$; dead cells are removed from the domain immediately. If a cell survives, then it can mutate independently of cell division with a probability p_{mut}^i, then divide if $p_2 < p_{div}$, $p_2 \sim U(0,1)$, and if there is available space. If a cell divides, it can mutate again with a probability p_{mut}^d. In this study, we assume $p_{mut}^i = p_{mut}^d = p_{mut}$. We restrict p_{div} and p_{die} to $p_{div} \in [0,1]$ and $p_{die} \in$

$[p_{die\ min},1]$, where $p_{die\ min}$ is set to 0.196. This relatively high $p_{die\ min}$ value ensures a high cell turnover, characteristic of multiple cancer cell lines [29,42].

Mutation Fitness Algorithm (MFA) and the Distribution of Fitness Effects (DFE)

The MFA algorithm models the mutation-caused quantifiable fitness change in an individual cell: every time a cell acquires a mutation, it impacts cell fitness by changing p_{div} or p_{die}. How much fitness changes depends on the underlying DFE, i.e., the frequency of the positive (beneficial), negative (deleterious), or neutral fitness effects. Our DFE is a β(a,b) distribution for all simulations. For results associated with Fig. 2, we set a = b = 3 and define the fitness effects of a mutation $x_t = X_t \in [0,1]$ as negative if $x_t \in [0,x_l]$, neutral if $x_t \in]x_l, x_u[$, and positive if $x_t \in [x_u,1]$; here, x_l = 0.082829 and x_u = 0.917171 following[15]. We vary a, b, x_l, and x_u when investigating the interplay between the mutation rates, DFEs, and the position on the adaptive landscape (Fig. 3). For all unbiased DFEs (equal frequency of negative and positive mutations), a = b = 3. For the "Major bias" (prevalence of highly impactful negative mutations), a = 1, b = 3, and for the "Minor bias" (prevalence of negative mutations with a variety of impact): a = 1, b = 1.3. We match the specific positive/negative/neutral effect frequencies by solving the cumulative probability distributions for x_l, and x_u values (Supplementary Material 1). After sampling x_t a linear transformation of x_t to y_t determines how p_{div} or p_{die} will change. For negative fitness effects, $L^- := [0, x_l] \rightarrow [0, 1]$, $L^-(x_t) := (x_l - x_t) / x_l$, and for the positive effects L^+: $[x_u, 1] \rightarrow [0, 1]$, $L^+(x_t) = (x_t - x_u) / (1 - x_u)$; neutral fitness effects are having $y_t = 0$. Because L^+ and L^- are linear, they are frequency preserving, implying that mutations with smaller impacts happen more frequently than those with large ones. Finally, y_t can affect p_{div} or p_{die} with a 50% frequency; this corresponds to the $p_{div} = p_{div} - y_t$ (50%) or $p_{die} = p_{die} + y_t$ (50%) with $y_t = L^-(x_t)$ (negative effects), and $p_{die} = p_{die} - y_t$ (50%) or $p_{div} = p_{div} + y_t$ (50%) with $y_t = L^+(x_t)$ (positive effects). If p_{div} +/- y_t exceeds 0 or 1, and p_{die} +/- y_t exceeds p_{die_min} (0.196) or 1, the respective p_{div} and p_{die} values are set to 0, 1, or p_{die_min}, respectively.

While the p_{div} and p_{die} changes happen immediately, they might be effective first during the next iteration; see supplementary material 1 for the relevant mathematical details. In addition to fitness changes, cells accumulate all mutations, allowing us to track the number of positive, negative, and neutral mutations over time.

Data collection

Each run of the ABM allows the collection of the average cell data (see Supp. Fig. 1a). The average cell data (e.g., mean p_{div}) allows us to infer the general population behavior over long time scales quickly and is collected every 100th time step until time step 40000, where 1 time step = ½ of a day = 1/730 of a year. Data from each replicate is saved in csv format. Simulation replicates inferred with this approach are denoted by the hat symbol ($\hat{}$). Some replicates can go extinct because of the stochasticity, space limitations, and zero net growth assumption. Thus, the total number of analyzed replicates ($\hat{n}$) often differs from the total simulation replicates executed (n).

Determining the no. simulations (statistical significance)

We estimated the number of required simulations by capturing the variance in the final number of cells (i.e., at time 54 years = 40000-time steps) as a function of the number of simulation replicates (Supp. Fig. 4) and observed for $p_{mut} > 0$, the variance converges; however, the asymptotic value of variance and the speed of convergence depend on the p_{mut}: higher p_{mut} leads to a higher variance value yet faster variance stabilization. In addition to the p_{mut}, the value of stabilized variance depends on the DFE (Supp. Fig. 4). Hence, the apriori determination of the sufficient number of simulations for in-depth statistical analysis is limited. Because populations are more prone to extinction with decreasing W_0, determining a sufficient number of simulations with low W_0 can induce a bias; thus, we use $W_0 = 0.75$ (initial $p_{div} = 1$, $p_{die} = 0.2$) to determine sufficient simulations as n = 250 for p_{mut} = 0, 1e-4 and 1e-3, n = 150 for p_{mut} = 1e-2, n = 75 for p_{mut} = 1e-1, and n = 35 for pmut = 1.

Max Δw

We compute the Max Δw as follows: 1) we first calculate the average fitness over time w(t) of a population with equal p_{mut} and W_0, then 2) we compute the forward differences, i.e., Δw = w(t+1) - w(t), as well as their absolute values, Δ|w| = |w(t+1) - w(t)|; while Δ|w| is strictly positive, Δw is not, and can thus indicate population fitness increase or decline. In 3), we select the maximum of the Δ|w|, (Max Δ|w|) and report the equivalent Δw value, corresponding to Max Δw. Fitness averages of populations that faced extinctions for all simulations were excluded from this analysis. Because Δw is a numerical approximation calculated from the simulation data, Max Δw depends on the number of data points and sampling size used for this calculation. To ensure that Max Δw is a representative measure of adaptiveness (i.e., that a significantly larger than zero value of Max Δw exists), we analyzed the individual Δw curves; examples of these curves (Supp. Fig. 3a) show that our data collection time step is sufficient to use Max Δw for further analysis.

In vitro assay

ALK+ non-small cell lung cancer cell line H3122 was used for the *in vitro* assays. The cell line was maintained in RPMI media (Fisher), supplemented with 10% fetal bovine serum (FBS) (Fisher) and insulin (ThermoScientific). Mutations were triggered using Ethyl methanesulfonate (EMS) (Fisher). The sensitivity of H3122 cells with addition of EMS (1μM) was assessed with DMSO (baseline) and ALKi, Crizotinib (0.5μM), Alectinib (0.5μM) and Lorlatinib (2μM). Briefly, cells were seeded for the colony formation assay at 0.6 million cells per 6 cm plate and treated with ALKi or ALKi + EMS and, at 2,000 cells per 6 cm plate, treated with DMSO and EMS. For the EMS groups, cells were pretreated with EMS for 48 hours to facilitate mutagenesis before seeding for the assay. Treatment was repeated with fresh media twice a week. Crystal violet staining was performed after 2-3 weeks for baseline groups and 3-5 weeks for the ALKi groups. Images were taken and analyzed using ImageJ (Fiji). Colony numbers, total area covered by the cells and areas of each colony were quantified. Clonogenic fraction was calculated as the fraction of the number of colonies by cells plated.

Tables

Table 1: List of all symbols and abbreviations

Name	Symbol	Value	Remarks
Anaplastic Lymphoma Kinase	ALK	N/A	
Non-Small Cell Lung Cancer	NSCLC	N/A	
Tyrosine Kinase Inhibitors	TKI	N/A	
Individual cell probability of division at time t	$p_{div} := p_{div}(t)$	Varies	Time index omitted for readability,
Individual cell probability of death at time t	$p_{die} := p_{die}(t)$	Varies	Time index omitted for readability
Individual cell fitness at time t	$w := w(t)$	$w = p_{div} - p_{die}$	Time index omitted for readability
Initial adaptive state of the population/initial population fitness	W_0	$W_0 = p_{div} - p_{die}/(1-p_{die})$	See Supp. Material for derivation
Minimal probability of division	p_{div_min}	0	
Maximal probability of division	p_{div_max}	1	
Minimal probability of death	p_{die_min}	0.196	
Maximal probability of death	p_{die_max}	1	
Division-dependent mutation rate	p_{mut}^{d}	{0, 1e-4, 1e-3, 1e-2, 1e-1, 1}	1. Constant for all times t, 2. Equal to the division-independent mutation rate
Division-independent mutation rate	p_{mut}^{i}	{0, 1e-4, 1e-3, 1e-2, 1e-1, 1}	1. Constant for all times t, 2. Equal to the division-dependent mutation rate
Mutation Fitness algorithm	MF	N/A	
Distribution of Fitness Effects (of a single mutation)	DFE	N/A	
Domain width	x_w	100 px	
Domain height	y_h	100 px	

Carrying capacity	K	10000	K = X * Y
The initial number of cells	$N_{cells,0}$	49	Approx 0.5% of carrying capacity
Average data recording frequency (time step)	T_{avg}	100	
Average data end time of the simulation	$T_{avg,end}$	40000	
beta distribution parameters	a,b	(3,3), (1,3) or (1,1.3)	See supplementary material 1
The potential fitness effect of a mutation	x_t	Varies	1. It depends on the x_u, x_l, and DFE. 2. Either positive, negative, neutral
Frequency interval for the positive mutations	$[x_u,1]$	Varies	See methods and supplementary material 1
Frequency interval for the negative mutations	$[0,x_l]$	Varies	See methods and supplementary material 1
Specific fitness effect of a mutation	y_t	Varies	Obtained from a linear transformation of x_t

Table 2: Two way KS statistics comparing fitness distributions at the end of time (related to Fig. 3d)

p_{mut} pairs tested	ks2 stats	ks2 p-value	1% significant?
1:1e-1	1.0	0.0	TRUE
1:1e-2	1.0	1.88737914186277E-15	TRUE
1:1e-3	1.0	2.716354218113E-06	TRUE
1:1e-4	Excluded (number of surviving simulation replicates too small)		
1:0	Excluded (number of surviving simulation replicates too small)		
1e-1:1e-2	1.0	0.0	TRUE
1e-1:1e-3	1.0	1.62979101658856E-07	TRUE
1e-1:1e-4	Excluded (number of surviving simulation replicates too small)		
1e-1:0	Excluded (number of surviving simulation replicates too small)		
1e-2:1e-3	0.7988505747126440	0.0007466916385323600	TRUE
1e-2:1e-4	Excluded (number of surviving simulation replicates too small)		
1e-2:0	Excluded (number of surviving simulation replicates too small)		

Code and Supplementary Movies

The code to generate the ABM, the code used for the statistical analysis, and the supplementary movies 1 (simulation examples) and supplementary movies 2 (fitness evolution) are publicly accessible at github.com/gortzaah/AdaptationMutation. The remaining supplementary material is displayed as a supplementary pdf.

References

1. Loewe, L. & Hill, W. G. The population genetics of mutations: good, bad and indifferent. *Philos Trans R Soc Lond B Biol Sci* **365**, 1153–1167 (2010).
2. Genetic Mutation. https://www.nature.com/scitable/topicpage/genetic-mutation-1127/.
3. Wright, S. Evolution in Mendelian Populations. *Genetics* **16**, 97–159 (1931).
4. Giraud, A. *et al.* Costs and benefits of high mutation rates: adaptive evolution of bacteria in the mouse gut. *Science* **291**, 2606–2608 (2001).
5. Swanstrom, R. & Schinazi, R. F. Lethal mutagenesis as an antiviral strategy. *Science* **375**, 497–498 (2022).
6. Lynch, M., Bürger, R., Butcher, D. & Gabriel, W. The mutational meltdown in asexual populations. *J. Hered.* **84**, 339–344 (1993).
7. Maley, C. C. *et al.* Classifying the evolutionary and ecological features of neoplasms. *Nat Rev Cancer* **17**, 605–619 (2017).
8. Rozhok, A. & DeGregori, J. Somatic maintenance impacts the evolution of mutation rate. *BMC Evol Biol* **19**, 172 (2019).
9. Bao, K., Melde, R. H. & Sharp, N. P. Are mutations usually deleterious? A perspective on the fitness effects of mutation accumulation. *Evol Ecol* **36**, 753–766 (2022).
10. Watson, C. J. *et al.* The evolutionary dynamics and fitness landscape of clonal hematopoiesis. *Science* **367**, 1449–1454 (2020).
11. Wloch, D. M., Szafraniec, K., Borts, R. H. & Korona, R. Direct estimate of the mutation rate

and the distribution of fitness effects in the yeast Saccharomyces cerevisiae. *Genetics* **159**, 441–452 (2001).

12. Eyre-Walker, A., Woolfit, M. & Phelps, T. The distribution of fitness effects of new deleterious amino acid mutations in humans. *Genetics* **173**, 891–900 (2006).
13. Loewe, L. & Charlesworth, B. Inferring the distribution of mutational effects on fitness in Drosophila. *Biol Lett* **2**, 426–430 (2006).
14. Trindade, S., Perfeito, L. & Gordo, I. Rate and effects of spontaneous mutations that affect fitness in mutator Escherichia coli. *Philos Trans R Soc Lond B Biol Sci* **365**, 1177–1186 (2010).
15. Martincorena, I. *et al.* Universal Patterns of Selection in Cancer and Somatic Tissues. *Cell* **171**, 1029–1041.e21 (2017).
16. Robert, L. *et al.* Mutation dynamics and fitness effects followed in single cells. *Science* **359**, 1283–1286 (2018).
17. Rojas, J. *et al.* Comparative modeling reveals the molecular determinants of aneuploidy fitness cost in a wild yeast model. *Cell Genom* 100656 (2024).
18. Williams, M. J. *et al.* Measuring the distribution of fitness effects in somatic evolution by combining clonal dynamics with dN/dS ratios. *Elife* **9**, (2020).
19. Hinz, A., Amado, A., Kassen, R., Bank, C. & Wong, A. Unpredictability of the Fitness Effects of Antimicrobial Resistance Mutations Across Environments in Escherichia coli. *Mol. Biol. Evol.* **41**, (2024).
20. Marcusson, L. L., Frimodt-Møller, N. & Hughes, D. Interplay in the selection of fluoroquinolone resistance and bacterial fitness. *PLoS Pathog.* **5**, e1000541 (2009).
21. Marusyk, A. *et al.* Irradiation alters selection for oncogenic mutations in hematopoietic progenitors. *Cancer Res* **69**, 7262–7269 (2009).
22. Greaves, M. & Maley, C. C. Clonal evolution in cancer. *Nature* **481**, 306–313 (2012).
23. Ramirez, M. *et al.* Diverse drug-resistance mechanisms can emerge from drug-tolerant

cancer persister cells. *Nat. Commun.* **7**, 10690 (2016).

24. Sharma, S. V. *et al.* A chromatin-mediated reversible drug-tolerant state in cancer cell subpopulations. *Cell* **141**, 69–80 (2010).
25. Vander Velde, R. *et al.* Resistance to targeted therapies as a multifactorial, gradual adaptation to inhibitor specific selective pressures. *Nat. Commun.* **11**, 2393 (2020).
26. França, G. S. *et al.* Cellular adaptation to cancer therapy along a resistance continuum. *Nature* **631**, 876–883 (2024).
27. Shaffer, S. M. *et al.* Rare cell variability and drug-induced reprogramming as a mode of cancer drug resistance. *Nature* **546**, 431–435 (2017).
28. Miti, T., Desai, B., Miroshnychenko, D., Basanta, D. & Marusyk, A. Dissecting the Spatially Restricted Effects of Microenvironment-Mediated Resistance on Targeted Therapy Responses. *Cancers* **16**, (2024).
29. Liu, S., Edgerton, S. M., Moore, D. H., 2nd & Thor, A. D. Measures of cell turnover (proliferation and apoptosis) and their association with survival in breast cancer. *Clin. Cancer Res.* **7**, 1716–1723 (2001).
30. Knudson, A. G., Jr. Mutation and cancer: statistical study of retinoblastoma. *Proc Natl Acad Sci U S A* **68**, 820–823 (1971).
31. National Research Council, Division on Earth and Life Studies, Division of Medical Sciences, Commission on Life Sciences & Committee to Study the Human Health Effects of Subtherapeutic Antibiotic Use in Animal Feeds. *The Effects on Human Health of Subtherapeutic Use of Antimicrobials in Animal Feeds*. (National Academies Press, 1980).
32. Springer, B. *et al.* Mechanisms of streptomycin resistance: selection of mutations in the 16S rRNA gene conferring resistance. *Antimicrob. Agents Chemother.* **45**, 2877–2884 (2001).
33. Chung, H. *et al.* Rapid expansion and extinction of antibiotic resistance mutations during treatment of acute bacterial respiratory infections. *Nat. Commun.* **13**, 1231 (2022).
34. Rozhok, A. I. & DeGregori, J. Toward an evolutionary model of cancer: Considering the

mechanisms that govern the fate of somatic mutations. *Proc Natl Acad Sci U S A* **112**, 8914–8921 (2015).

35. Wagner, A. Robustness and evolvability: a paradox resolved. *Proc Biol Sci* **275**, 91–100 (2008).
36. Gillespie, J. H. Molecular evolution over the mutational landscape. *Evolution* **38**, 1116 (1984).
37. Fisher, R. A. *The Genetical Theory of Natural Selection*. (Clarendon Press, Oxford, 1930).
38. Fitzgerald, D. M., Hastings, P. J. & Rosenberg, S. M. Stress-Induced Mutagenesis: Implications in Cancer and Drug Resistance. *Annu Rev Cancer Biol* **1**, 119–140 (2017).
39. McGranahan, N. & Swanton, C. Clonal Heterogeneity and Tumor Evolution: Past, Present, and the Future. *Cell* **168**, 613–628 (2017).
40. Pribis, J. P., Zhai, Y., Hastings, P. J. & Rosenberg, S. M. Stress-Induced Mutagenesis, Gambler Cells, and Stealth Targeting Antibiotic-Induced Evolution. *mBio* **13**, e0107422 (2022).
41. Bravo, R. R. *et al.* Hybrid Automata Library: A flexible platform for hybrid modeling with real-time visualization. *PLoS Comput. Biol.* **16**, e1007635 (2020).
42. Gallaher, J. A., Brown, J. S. & Anderson, A. R. A. The impact of proliferation-migration tradeoffs on phenotypic evolution in cancer. *Sci. Rep.* **9**, 2425 (2019).