BIOLOGICAL SCIENCES: Evolution

A tug-of-war between driver and passenger mutations in cancer and other adaptive processes

Christopher D McFarland[1], Leonid Mirny[1,2,3] and Kirill S Korolev[3,4]

[1] Graduate Program in Biophysics, Harvard University, Boston, MA 02115
[2] Department of Physics, Massachusetts Institute of Technology, Cambridge, MA 02139
[3] Institute for Medical Engineering and Science, Massachusetts Institute of Technology, Cambridge, MA 02139
[4] Currently at Department of Physics, Boston University, Boston, MA 02215

Corresponding Author: Leonid Mirny leonid@mit.edu, E25-526, MIT, 77 Massachusetts ave, Cambridge, MA 02139, (617) 452-4862

A tug-of-war between driver and passenger mutations in cancer and other adaptive processes

Christopher D McFarland [*], **Leonid Mirny** [* † ‡ §], **Kirill S Korolev** [† ¶]

[*]Graduate Program in Biophysics, Harvard University, Boston, MA 02115,[†]Department of Physics &,[‡]Institute for Medical Engineering and Science, Massachusetts Institute of Technology (MIT), Cambridge, MA 02139, USA,[§]Corresponding Author, and [¶]Currently at Department of Physics, Boston University, Boston, MA 02215

Submitted to Proceedings of the National Academy of Sciences of the United States of America

Cancer progression is an example of a rapid adaptive process where evolving new traits is essential for survival and requires a high mutation rate. Precancerous cells acquire a few key mutations that drive rapid population growth and carcinogenesis. Cancer genomics demonstrates that these few 'driver' mutations occur alongside thousands of random 'passenger' mutations—a natural consequence of cancer's elevated mutation rate. Some passengers can be deleterious to cancer cells, yet have been largely ignored in cancer research. In population genetics, however, the accumulation of mildly deleterious mutations has been shown to cause population meltdown. Here we develop a stochastic population model where beneficial drivers engage in a tug-of-war with frequent mildly deleterious passengers. These passengers present a barrier to cancer progression that is described by a critical population size, below which most lesions fail to progress, and a critical mutation rate, above which cancers meltdown. We find support for the model in cancer age-incidence and cancer genomics data that also allow us to estimate the fitness advantage of drivers and fitness costs of passengers. We identify two regimes of adaptive evolutionary dynamics and use these regimes to rationalize successes and failures of different treatment strategies. We find that a tumor's load of deleterious passengers can explain previously paradoxical treatment outcomes and suggest that it could potentially serve as a biomarker of response to mutagenic therapies. Collective deleterious effect of passengers is currently an unexploited therapeutic target. We discuss how their effects might be exacerbated by both current and future therapies.

cancer | evolution | genomics | mathematical modeling

Significance Statement

During adaptation, populations start in hostile conditions and needs to evolve new traits to survive. A population of cancer cells within a body is an example of such a population: cancer needs to evolve new traits to survive and progress. These new traits are acquired by mutations in a few specific "driver" genes. However, many other genes are mutated and randomly altered during this process, potentially damaging cancer cells. The role that these damaging "passenger" mutations play in cancer, and other adaptive processes, is unknown. Here we show that driver mutations engage in a tug-of-war with damaging passengers. This tug-of-war explains many phenomena in oncology, suggesting how to target existing therapies and develop new therapies to exploit damaging passengers.

Introduction

While many populations evolve new traits via a gradual accumulation of changes, some adapt very rapidly. Examples include viral adaptation during infection (1); the emergence of antibiotic resistance (2); artificial selection in biotechnology (3); and cancer (4). Rapid adaptation is characterize by three key features: (i) the availability of strongly advantageous traits accessible by a few mutations, (ii) an elevated mutation rate (5, 6), and (iii) a dynamic population size (7). Traditional theories of gradual adaptation are not applicable under these conditions, and new approaches are needed to address pressing problems in medicine and biotechnology.

Cancer progression is an example of a rapidly adapting population: cancers develop as many as ten new traits (8), often have a high mutation rate (8-10), and a population size that is rapidly changing over time. This process is driven by a handful of mutations and chromosomal abnormalities in cancer-related genes (oncogenes and tumor suppressors), collectively called *drivers*. From an evolutionary point of view, drivers are mutations that are beneficial to cancer cells because their phenotypes increase the cell proliferation rate or eliminate brakes on proliferation (8). Drivers, however, arise alongside thousands of other mutations and alterations dispersed through the genome that have no immediate beneficial effect, collectively called *passengers*.

Passengers have been previously assumed to be neutral and largely ignored in cancer research, yet growing evidence suggests that they may sometimes be deleterious to cancer cells and, thus, play an important role in both neoplastic progression and clinical outcomes. In an earlier study, we showed that deleterious passengers can readily accumulate during tumor progression and found that many passengers present in cancer genomes exhibit signatures of damaging mutations (11). Additionally, chromosomal gains and losses that are pervasive in cancer can be passengers, and have been shown to be highly damaging to cancer cells (12). Lastly, cancers with high levels of chromosomal alterations exhibit better clinical outcomes in breast, ovarian, gastric, and non-small cell lung cancer (13). Passenger mutations and chromosomal abnormalities can be deleterious via a variety of mechanisms: direct loss-of-function (14), proteotoxic cytotoxicity from protein disbalance and aggregation (15), or by inciting an immune response (16).

While the role of deleterious mutations in cancer is largely unknown, their effects on natural populations has been extensively studied in genetics (5, 17-19). The accumulations of deleterious mutations can cause the extinction of a population via processes known as Muller's ratchet and mutational meltdown (17, 20, 21) It was recently proposed that inevitable accumulation of deleterious mutations in natural populations should be offset by new beneficial mutations, leading to long-term population stability (19). Here we consider a rapid adaptation of a population with a variable size and subject of a high mutation rate. A rapidly adapting population faces a double bind: it must quickly acquire, often exceeding rare, adaptive mutations and yet avoid mutational meltdown. As a result,

Reserved for Publication Footnotes

adaptive processes frequently fail. Indeed, less than 0.1% of species on earth have adapted fast enough to avoid extinction (22) and, similarly, only about 0.1% of precancerous lesions ever advance to cancer (23). To control cancer or pathogens, we should understand the constraints that evolution imposes on their rapid adaptation.

Here we investigate how asexual populations such as tumors rapidly evolve new traits while avoiding mutational meltdown. Unlike classical theories of gradual adaptation, the evolutionary model we develop has three key features: (i) rare, strongly advantageous driver mutations, (ii) a high mutation rate that makes moderately deleterious passengers relevant, and (iii) a population size that varies with the fitness of individual cells. We found that a tug-of-war between beneficial drivers and deleterious passengers creates two major regimes of population dynamics: an adaptive regime, where the probability of adaptation (cancer) is high; and a non-adaptive regime, where adaptation (cancer) is exceedingly rare.

Adaptive and non-adaptive regimes are separated by a critical population size or barrier to cancer progression that most lesions fail to overcome, and a critical mutation rate that leads to mutational meltdown. We found strong evidence of these phenomena in age-incidence curves and recent cancer genomics data. Agreement of the model with these data allows us to estimate the selective advantages of drivers as 10-50%, a range consistent with recent direct experimental measurements (24). Genomic data also show that deleterious passengers are approximately 100 times weaker. Our model offers a new interpretation of cancer treatment strategies and explains a previously paradoxical relationship between cancer mutation rates and therapeutic outcomes. Most importantly, it suggests that deleterious passengers offer a new, unexploited avenue of cancer therapy.

Results

Model. We consider a dynamic population of cells that can divide, mutate (in a general sense, i.e. including alterations, epigenetic changes, etc), and die stochastically. Mutations occur during cell divisions with a per-locus rate μ. The number of **d**river loci in the genome, i.e. a driver target size, is T_d, while the target size for deleterious **p**assengers is T_p. Hence, the genome-wide driver and passenger mutation rates are $\mu_d = \mu T_d$ and $\mu_p = \mu T_p$ respectively. A driver increases an individual's growth rate by $s_d \sim 0.1$, while a new passenger decreases the growth rate by $s_p \sim 10^{-4}-10^{-1}$. Here we only consider fixed values of s_d and s_p because previous work showed that drivers and passengers sampled from various fitness distributions (exponential, normal, and Gamma) exhibit essentially the same dynamics (11). The net effect of multiple mutations on cell fitness w is given by $w = (1+s_d)^{n_d}(1+s_p)^{n_p}$, where n_d and n_p are the total number of drivers and passengers in a cell.

The birth and death rates of a cell in our model depend not only on fitness, but also on the population size N via a Gompertzian growth function often used to describe cancerous populations (25) (see **SI** for details). At large N, deaths exceed births and tumors must adapt (or innovate) via new drivers to progress to larger population sizes. Thus, populations in our model expand and shrink in two ways: on a short time-scale due to stochastic cell divisions and deaths, and on a long time-scale due to the accumulation of advantageous and deleterious mutations. Previous models of advantageous and deleterious mutations have not considered a varying population size (26, 27).

In cancer and other adapting populations the target size for advantageous mutations (drivers) is much smaller than the target size for deleterious mutations ($T_d \ll T_p$). If driver loci include a few specific sites (~ 10 per gene) in all cancer-associated genes (approximately 100, (28)), then collectively drivers will constitute less than one one-millionth of the genome. Conversely, as much as 10% of the human genome is well-conserved and likely deleterious when mutated (29, 30). In natural populations, T_p should still remain much greater than T_d simply because natural selection optimizes genomes to their environment, implying that most changes will be neutral or damaging. Indeed, most protein coding mutations and alterations were deleterious or neutral when investigated in fly (31), yeast (32), and bacterial genomes (33). We consider only moderately deleterious loci here ($s_p \approx 10^{-4} - 10^{-1}$)—which account for most nonsynonymous mutations (34, 35). Deleterious mutations outside of this range either do not fixate or negligibly alter progression (11). Hence, we used a conservative size of $T_p \approx 10^5 - 10^7$ loci to account for passengers with fitness effects outside of this range that we are neglecting (see **SI** and **Table S1** for details of parameters estimation). This quantity is still much greater than T_d. Finally, we explored a variety of driver fitness advantages, as estimates in the literature ranged from 0.0001 (36) to 1 (24).

A critical population size. **Figure 1A** shows the dynamics $N(t)$ of individual populations starting at different initial sizes N^0, which correspond to different potential hyperplasia sizes (we begin trajectories immediately after a stem cell acquires its first driver, see **SI** for a discussion of dynamics before this time point). Populations exhibit two ultimate outcomes: growth to macroscopic size (i.e. cancer progression), or extinction. The prevalence of either outcome is determined by a critical population size N^*, about which larger populations ($N > N^*$) generally commit to rapid growth and smaller populations ($N < N^*$) generally commit to extinction.

To understand the cause of this critical population size N^*, we looked at the short-term dynamics of populations. All trajectories show a reversed saw-toothed pattern (**Fig. 1B**), which result from a tug-of-war between drivers and passengers (11). When a new driver arises and takes over the population, the population size increases to a new stationary value. In between these rare driver events, the population size gradually decreases due to the accumulation of deleterious passengers. The relative rate of these competing processes determines whether a population commits to rapid growth or goes extinct.

We can identify the location of N^* by considering the average change in population size over time ($< dN/dt >$), which is simply the average population growth due to driver accumulation (v_d) minus the population decline due to passenger accumulation (v_p). Fixation of a new driver causes an immediate jump in population size $\Delta N = N s_d$. These jumps occur randomly at a nearly constant rate $f = \mu_d N s_d$, given by the driver occurrence rate $\mu_d N$, multiplied by a driver's fixation probability $s_d/(1+s_d) \approx s_d$. Hence, the velocity due to drivers is $v_d = f\Delta N = \mu_d N^2 s_d^2$. Similarly, passengers' velocity ($v_p = \mu_p N s_p$) is a product of their rate of occurrence ($\mu_p N$); their effect on population size ($N s_p$); and their probability of fixation ($\sim 1/N$, a more accurate measure of this probability is used below and provided in the **SI**). Thus, we obtain:

$$\left\langle \frac{dN}{dt} \right\rangle = \mu_p s_p N \left(\frac{N}{N^*} - 1 \right) \qquad [1]$$

$$N^* = \frac{T_p s_p}{T_d s_d^2} \qquad [2]$$

where N^* is the critical population size.

Because the population velocity is negative below N^* and positive above N^* there is an effective barrier for cancer (**Fig. 1C**): smaller populations tend to shrink, while larger populations tend to expand. Simulations support our conclusion that the probability of cancer increases with N^0 and sharply transitions at N^* (**Fig. 1D**). Indeed, drastically different probability curves collapse onto a single curve once N^0 is rescaled by N^* (computed from equation 2). Since N^* captures only the average, or mean-field, dynamics, it misses the variability of outcomes in rapidly adapting populations. **Figure 1E** illustrates that the variability of outcomes depends upon the strength of drivers s_d. Higher values of s_d lead to larger stochastic jumps, which leads to larger deviations from mean behavior and more gradual changes in the probability of cancer across N^0. Thus, we formulated and analytically solved a stochastic generalization of equation 1 that incorporates this variability (**SI**). Our solution provides an excellent fit to simulations (**Fig. 1E**) and indicates that N^* and s_d fully describe population dynamics (**SI**).

We can understand how N^* and s_d control cancer progression using a simple random-walk analogy. The population size experiences random jumps, resulting from driver fixation events, which are described by equation 1. These random jumps and declines are effectively a random walk in a one-dimensional effective potential ($U_{eff} = \int (dN/dt\,)\, dN$), **Fig. 1C** and **SI**) with stochastic jumps of frequency f and size ΔN. Similar to chemical reactions activated by thermal energy, cancer progression is a rare event triggered by a quick succession of driver fixations. Below, we show that human tissues operate in a regime where progression is rare and successful lesions are the rare lesions that happen to acquire drivers faster than average. We found that population dynamics depend entirely on two dimensionless parameters: a deterministic mean velocity, dependent only upon N/N^*, and a stochastic step-size that is approximately proportional to s_d. By reducing the complexity of our evolutionary system to two parameters, we were next able to infer their values for real cancers without over-fitting.

Model validation using cancer incidence and genomic data. Our model of cancer progression predicts the presence of an effective barrier to cancer where small lesions are very unlikely to ever progress to cancer. It also predicts a specific distribution in the number of passenger mutations and a specific relationship between drivers and passengers in individual cancer samples. We looked for evidences of these phenomena in age-incidence data (37) and cancer genomics data (28, 38-40). These comparisons also allowed us to estimate some critical parameters of the model: N^0, s_d, and s_p.

Figure 2A presents the incidence rate of breast cancer versus age (37) alongside the predictions from a classic driver-only model (**SI**) and our model. The incidence rate was calculated by considering a process where precancerous lesions arise with a constant rate r beginning at birth. These lesions then progress to cancer in time τ with a probability $P(\tau)$ that we determined from simulations (**Fig. S1**). By convoluting this distribution $P(\tau)$ with the lesion initiation rate r, we can predict the age-incidence rate $I(t)$. Because many lesions go extinct in our model and never progress to cancer, the predicted incidence rate saturates at old-age: $I_{\max} = r \int_0^\infty P\,(\tau)\, d\tau = r P_\infty$, where P_∞ is the probability that a lesion will ever progresses to cancer, determined above.

Both the observed population incidence rates and our driver-passenger model saturate with age. This is a direct result of the probability of progression from a lesion to cancer being low. We estimate a lower bound for the rate of lesion formation r in breast cancer of at least 10 lesions per year that can be arrived at through two separate considerations: first, by considering the quantity of breast epithelial stem cells and the rate at which they can mutate into lesions (**SI**), and second, by considering the number of lesions observed within the breast tissue of normal cadavers (23). By comparing this limit (~ 10 lesions $\cdot$ year^{-1}) to the maximum observed breast cancer incidence rate $I_{\max} \approx 10^{-2}$ cancers $\cdot$ year^{-1}, we find that $P_\infty \approx 10^{-3}$, or only about 1 in 1,000 lesions ever progress. This finding is consistent with a number of clinical studies that have observed that very few lesions ever progress to cancer, while many more regress to undetectable size (41, 42)—another property seen in our model. Thus, good agreement between age-incidence data and our model is obtained when $s_d \approx 0.1 - 0.2$ and N^0/N^* is chosen such that $P_\infty = 10^{-3}$. This suggests that cancer begins at a population size far below N^*, where drivers are most often overpowered by passengers. Indeed, 21 of the 25 most prevalent cancers plateau at old-age suggesting that progression is inefficient in most tumor types (**Fig. S1**). In a driver-only model (see **SI** for details), every lesion progresses to cancer after sufficient time (i.e. $P_\infty = 1$), therefore a plateau in incidence rate can only result from a very low lesion formation rate (0.01 per year), which is inconsistent with abundant pathology data (23, 43).

Recent cancer genomics data offer a new opportunity to validate our model. Specifically, we looked at Somatic Non-synonymous Mutations (SNMs) and Somatic Copy-Number Alterations (SCNAs) derived from over 700 individual cancer-normal sample pairs obtain from the breast (38), colon (28), lung (40), and skin (39) (**Table S2**). We found similar results when analyzing SNMs and SCNAs both separately (**Fig. S2**, **Table S3**) or in aggregation (**Fig. 2BC**). **Figure 2B** shows a wide and asymmetric distribution of the total number of mutations, which is consistent with our model under realistic parameters. A driver-only model yields a narrower and more symmetric distribution that is inconsistent with the data (**Fig. 2B**). The driver-only model can fit the observed distribution only if it assumes that just 1-2 drivers are needed for cancer (**Fig. S3**, **Table S4**)—a value inconsistent with both the extent of recurrent mutations seen in cancer (10) and known biology. This large variance in mutation totals further supports our model and suggests that driver mutations and alterations have a large effect size: $s_d \approx 0.4$. Our estimate of $s_d \approx 0.1 - 0.4$ obtained from age-incidence and mutation histograms is in excellent agreement with experimentally measured changes in the growth rate of mouse intestinal stem cells upon induction of p53, APC or k-RAS mutations where measured values ranged from of 0.16 to 0.58 (24).

We then used cancer genomics data to compare the number of drivers and passengers observed in individual cancer samples to our model's predicted relationship. In our model, additional passengers must be counterbalanced by additional drivers for the population to succeed. If a lesion lingers around N^* for a long time, then it must have acquired both many passengers and many counterbalancing drivers; while lesions that quickly progress through the barrier at N^* acquire fewer of each. As a result, we expect a positive linear relationship between the number of drivers and passengers: $n_d \cdot s_d - n_p \cdot s_p =$ constant; this result follows directly from the definition of fitness in our model (**SI**). Our predicted positive linear relationship between drivers and passengers is indeed observed in all tumor types that we studied (**Fig. 2C**, **Table S3**, $p < 0.08 - 10^{-6}$). The slope of this regression line is predicted to be s_p/s_d, which ranged from 1/21 to 1/193 (**Table S3**) for the various subtypes. While there is considerable variation and large margins of error in these numbers,

these slopes ($s_p/s_d \approx 5 \cdot (10^{-2} - 10^{-3})$) correspond to an s_p of $5 \cdot (10^{-3} - 10^{-4})$ when $s_d = 0.1$. These rough values are similar to germ-line SNMs in humans of European descent, where 64% of all mutations exhibit an s_p between 10^{-5} and 10^{-2} (35).

We considered and refuted several alternative explanations for the observed positive linear relationship between drivers and passengers. First, that the strength of SCNAs may differ from SNMs. Hence, we investigated each alteration-type separately and found positive linear relationships in both cases (**Fig. S2**, **Table S3**). Second, that the number of driver alterations might be explained by variation in the tumor stage, or the rate and/or mechanism of mutagenesis. In **Table S3** we show that these factors cannot suppress the correlation between drivers and passengers. Lastly, we considered and refuted the possibility that this relationship between drivers and passengers is non-linear (**Fig. 2C** insert). Because the data disagrees with all of these alternate hypotheses, we believe that it supports our conclusion that cancer progression is a tug-of-war between drivers and passengers.

A critical mutation rate. We next used simulations to investigate the probability of cancer over a broad range of evolutionary parameters (**Fig. S4**) and found that there is a critical mutation rate above which the probability of cancer is exceedingly low (**Fig. 3A**). To explain this phenomenon and to find the parameters that determine this critical mutation rate μ^*, we modified our analytical framework to consider selection against passengers and the effects of unfixed passengers on the accumulation of drivers. The modified framework, described in the SI, explains observed dynamics well (**Fig. 3AB**, **S4**). Previous theoretical work has shown that the number of unfixed passengers per cell is Poisson distributed with mean μ_p/s_p [first described in (44)]. This result assumes an approximate balance between the mutation rate of passengers and the selection against them, otherwise known as mutation-selection balance. The average fitness reduction of a cell due to this mutational load (i.e. the reduction in fitness relative to the fittest cells in the population) is μ_p. A new driver arises in one of these cells at random and must carry the load of passengers residing in its cell along with it to fixation (18) (**Fig. 3C**); this process is often referred to as hitchhiking, so we describe these passengers as 'hitchhikers'. If the reduction in fitness due to the load of passengers ($\mu_p = \mu T_p$) exceeds the benefit of a new driver (s_d), then the driver will not fixate (**Fig. 3C**). Hence, cancer is extremely rare when $\mu_p > s_d$. This suggests a critical mutation rate:

$$\mu^* = s_d/T_p \qquad [3]$$

The critical mutation rate suggests a new mode by which mutational meltdown operates. Prior models of mutational meltdown consider deleterious mutations in isolation (17), whereas our model points at the ability of deleterious mutations to inhibit the accumulation of advantageous mutations as a mechanism of meltdown. While it has been previously shown that deleterious mutations interfere with the fixation of beneficial alleles (18, 26, 45), this phenomenon has never been studied in the context of population survival. We discuss some important implications of this critical mutation rate for cancer treatment below.

We found support for the critical mutation rate in both cancer age-incidence and cancer genomics data. If we constraint cancer progression to develop within the typical timeframe for cancer progression (i.e. when we begin to see a plateau in incidence: $\sim$ 60 years or 10,000 generations), the probability of cancer exhibits an optimum across mutation rates (**Fig. 3D**). Above μ^* population meltdown is very common, while at very low mutation rates progression is too slow. The optimal mutation rate ($10^{-9} - 10^{-8}$mutations $\cdot$ nucleotide$^{-1} \cdot$ generation^{-1}) is similar to mutation rates observed in cancer cell lines with a mutator phenotype (9) and the inferred mutation rate derived from the median number of mutations observed in a pan-cancer study of >3,000 tumors (10). Because μ^* depends only on s_d and T_p, and is independent of other variables, we believe the maximal mutation rate should be the same across tumor subtypes (**SI**, **Fig. S4**). Indeed, a maximum of approximately 100 somatic mutations per Mb (99.4th percentile) was observed in the pan-cancer study mentioned above, which corresponds to our theoretical estimate (if we assume that the most mutagenic cancers still require 1,000 generations to progress).

Two regimes of dynamics. Taken together, our results demonstrate that the tug-of-war between advantageous drivers and deleterious passengers creates two major regimes of population dynamics: an adaptive regime where the probability of progression (cancer) is high ($\sim$ 100 %) and a non-adaptive regime where cancer progression is exceedingly rare (**Fig. 3**, **S4**). Evolving populations that fail to adapt and go extinct may do so because they reside in a non-adaptive region of the phase space. Similarly, normal tissues that avoid cancers may present a tumor microenvironment that is in this non-adaptive regime. By keeping N^* sufficiently high, a tissue or clinician could keep cancerous populations outside of the adaptive regime. This critical population size $N^* = T_p s_p/(T_d s_d^2)$ depends on the evolutionary parameters of the system. For example, if s_p were increased by tuning the response of the immune system to mutation-harboring cells, or if T_d were decreased via a driver-targeted therapy, adaptation would become less likely. Below we demonstrate that a successful treatment must push a cancer back to the non-adaptive regime.

The adaptive barrier and critical mutation rate explain cancer treatment outcomes. We simulated cancer growths and treatments and then monitored the long-term dynamics of these populations. Most treatments used today attempt to reduce tumor size, e.g. by specifically inhibiting key drivers (46) or by simply killing rapidly dividing cells (chemotherapy and radiation). Chemotherapy and radiation also elevate the mutation rate, thus affecting evolutionary dynamics. Previous work on the evolution of resistance to therapy has not considered the barriers to adaptation that we observe, so we re-investigated evolutionary outcomes from standard therapies and identified new potential ways to treat cancer. While real cancers have a varied evolutionary history, our analytical formalism predicts that cancer's future dynamics depend only on their current state, not their history (i.e. cancer dynamics are approximately path-independent, **Fig. S5**). We show below that this assumption can accurately predict outcomes in simulations and the clinic.

In **Figure 4**, we present the evolutionary paths of cancer—from hyperplasia, to cancer, to treatment, and relapse or remission—on top of the phase diagrams described earlier. Our analysis demonstrates that a treatment is successful if it pushes a cancer into the non-adaptive regime of evolutionary dynamics where the probability of adaptation is low. Conversely, therapies fail, and populations re-adapt and remiss, when the therapy does not move cancer far enough to place it in the non-adaptive regime.

Our model suggests that chemotherapies succeed, in part, because they move cancers across the mutational threshold μ^*. Beyond this threshold, the probability that a driver is strong enough to overpower a load of passengers becomes small (see

above, **Fig. 2C**), making it hard for cancer to readapt. Increasing the mutation rate has little effect on the critical population size N^* (see above).

Thus, our model suggest that cancers with a very high load of mutations/alterations are close to the critical mutation rate and should be more susceptible to mutagenic treatments, such as chemotherapy. Several recent studies (13, 47) have noticed that patients survive breast and ovarian cancer most often when their tumors exhibited exceptional high levels of chromosomal alterations. This phenomenon was robust within and between subtypes of breast cancer (47). This finding is paradoxical for all previous models of cancer, where a greater mutation rate always accelerates cancer evolution and adaptation; yet is fully consistent with our model (**Fig. 4B**).

Treatments exploiting the mutational load of cancers (i.e. their accumulated passengers) remain largely unexplored. We show that increasing the deleterious effect of passengers s_p causes tumors to enter remission. Increasing s_p is doubly effective because it exacerbates the deleteriousness of accumulated passengers and also slows down future adaptation. When we simulate such treatment by a 3-5 fold increase of s_p (**Fig. 4C**), we observe an immediate decline in the population size followed by a low probability of replace due to an increased N^*. The phase diagram shows that a mild increase in s_p is sufficient to push a population into an extinction regime and thus induce remission. Below we discuss possible treatment strategies that would increase s_p.

Given the large number of treatment options, finding therapies that work synergistically is a very important problem in cancer research (reviewed in (48)). While synergism is often discussed in the context of pharmacology, our phase diagrams identify *evolutionarily* synergistic treatments. We found that remission was most likely to occur when the mutation rate and the fitness cost of passengers were increased simultaneously, more so than would be expected from simply adding together the effects of the individual therapies (**Fig. S6**). Hence, combinations of mutagenic chemotherapy along with treatments that elevate the cost of a mutational load may be most effective. According to our model, these therapies should also be compatible and complementary to driver-targeted therapies.

Discussion

We present an evolutionary model of rapid adaptation that incorporates rare, strongly advantageous driver mutations and frequent, mildly deleterious passenger mutations. In this process, a population can either succeed and adapt, or fail and go extinct. We found theoretically, and confirmed by simulations, two regimes of dynamics: one where a population almost always adapts, and another where it almost never adapts. Complex stochastic dynamics, which emerge due to a tug-of-war between drivers and passengers, can be faithfully described as diffusion over a potential barrier that separates these two regimes. The potential barrier is located at a critical population size that a population must overcome to adapt. We also found a critical mutation rate, above which populations quickly meltdown. This general framework for adaptive asexual populations appears to be perfectly suited to characterize the dynamics of cancer progression and responses to therapy.

Progression to cancer is an adaptive process, driven by a few mutations in oncogenes and tumor suppressors. During this process, however, cells acquire tens of thousands of random mutations many of which may be deleterious to cancer cells. While strongly deleterious passenger mutations are weeded out by selection, mildly deleterious can fixate and even accumulate in a cancer by hitchhiking on drivers, as we have shown earlier (11). Passengers may be deleterious by inducing loss-of-function in critical proteins (14), gain-of-function toxicity via proteotoxic/misfolding stress (15, 49), or by triggering an immune response by a mutated epitope (16, 50). While we looked at passengers in cancer exomes and in SCNAs, passengers may also constitute epigenetic modifications, or karyotypic imbalances (15). Hence the number of deleterious passengers may be more than currently observed by genome-wide assays.

Our framework suggests that most normal tissues reside in a regime where cancer progression is exceedingly rare; i.e. most lesions fail to grow above the critical population size and, thus, fail to overcome the adaptive barrier. Clinical cancers, on the contrary, reside above the adaptive barrier in a rapidly adapting state. Therapies must push a cancer below this adaptive barrier to succeed. In our framework, this entails moving the population below N^* or increasing the mutation rate above μ^*. The availability of a broad range of data for cancer allowed us to thoroughly test our framework's applicability.

We tested out model and estimated its parameters using cancer age-incidence curves, cancer exome sequences from almost 1,000 tumors in four cancer subtypes, and data on clinical outcomes. Age-incidence curves support the notion that the vast majority of lesions fail to progress and allow us to estimate the fitness benefit of a driver as $s_d \sim 0.1 - 0.4$. Genomics data suggests that passengers are indeed deleterious and that their deleterious effect is approximately one hundred times weaker than driver's beneficial effect. Moreover, the fitness benefit of a driver estimated from genomic data is roughly $s_d \sim 0.4$. Our range of values are consistent with recently measured 16-58% increases in the mouse intestinal stem cell proliferation rate upon mutations in APC, k-RAS or p53 (24). Taken together these data support the notion of a tug-of-war between rare and large-effect drivers and frequent, but mildly deleterious passengers $s_p \sim 5 \cdot (10^{-4} - 10^{-3})$, which nevertheless have a large collective effect.

Results of our analysis have direct clinical implications. Available clinical data (13, 47, 51) show that cancers with a higher load of chromosomal alterations, i.e. close to μ^*, respond better to treatments. Our study suggests two potentially synergistic therapeutic strategies: to increase the mutation rate above μ^*, and/or to increase the deleterious effect of accumulated passengers. An increase in the fitness cost of passengers would not only magnify the effects of an already accumulated mutational load, but also reduce future adaptation. This may be accomplished by (i) targeting unfolding protein response (UPR) pathways and/or the proteasome (15), (ii) hyperthermia that may further destabilize mutated proteins or clog UPR pathways (52), or (iii) by eliciting an immune response (16). Intriguingly, all these strategies are in clinical trials, yet they are often believed to work for reasons other than by exacerbating passengers' deleterious effects. In contrast, we predict that these therapies will be most effective in cancers with more passengers and an elevated mutation rate. Thus, characterizing the load of mutations/alterations in tumors may offer a new biomarker for predicting treatment outcomes and identify the best candidates for mutational chemotherapies.

While this study focused on asexual innovative evolution in cancer, our model may be generally applicable to other innovating populations. Consider a population in a new environment. The population is often initially small and fluctuating in size and often goes extinct, yet occasionally it expands to a larger stationary size by rapidly acquiring several new traits that are highly advantageous in the new environment. Both the evolutionary parameters (53) and observed phenomena (54) match our model well. Our mathematical framework may explain why these populations sometimes adapt, yet often fail.

Materials and Methods

Please see the **Supplemental Information** for details.

ACKNOWLEDGMENTS. CM and LM were supported by the National Cancer Institute-supported Physical Sciences in Oncology Center at MIT. KK was supported by Pappalardo Fellowship in Physics at MIT. We are grateful to Shamil Sunyaev, Gregory Krykov, Geoffrey Fudenberg, Maxim Imakaev, & Anton Goloborodko for many productive discussions.

1. Kawashima Y et al. (2009) Adaptation of HIV-1 to human leukocyte antigen class I. Nature 458:641-5.
2. Zhang Q et al. (2011) Acceleration of emergence of bacterial antibiotic resistance in connected microenvironments. Science 333:1764-7.
3. Wright SI et al. (2005) The effects of artificial selection on the maize genome. Science 308:1310-4.
4. Di Nicolantonio F et al. (2005) Cancer cell adaptation to chemotherapy. BMC Cancer 5:78.
5. Ishii K, Matsuda H, Iwasa Y, Sasaki A (1989) Evolutionarily stable mutation rate in a periodically changing environment. Genetics 121:163-74.
6. Moxon ER, Rainey PB, Nowak M a, Lenski RE (1994) Adaptive evolution of highly mutable loci in pathogenic bacteria. Curr Biol 4:24-33.
7. Cameron TC, O'Sullivan D, Reynolds A, Piertney SB, Benton TG (2013) Eco-evolutionary dynamics in response to selection on life-history. Ecol Lett 16:754-63.
8. Hanahan D, Weinberg R a (2011) Hallmarks of cancer: the next generation. Cell 144:646-74.
9. Jackson a L, Loeb L a (1998) The mutation rate and cancer. Genetics 148:1483-90.
10. Lawrence MS et al. (2013) Mutational heterogeneity in cancer and the search for new cancer-associated genes. Nature 499:214-8.
11. McFarland CD, Korolev KS, Kryukov G V, Sunyaev SR, Mirny L a (2013) Impact of deleterious passenger mutations on cancer progression. Proc Natl Acad Sci U S A 110:2910-5.
12. Williams BR, Amon A (2009) Aneuploidy: cancer's fatal flaw? Cancer Res 69:5289-91.
13. Birkbak NJ et al. (2011) Paradoxical relationship between chromosomal instability and survival outcome in cancer. Cancer Res 71:3447-52.
14. MacArthur DG et al. (2012) A systematic survey of loss-of-function variants in human protein-coding genes. Science 335:823-8.
15. Sheltzer JM, Amon A (2011) The aneuploidy paradox: costs and benefits of an incorrect karyotype. Trends Genet 27:446-53.
16. Segal NH et al. (2008) Epitope landscape in breast and colorectal cancer. Cancer Res 68:889-92.
17. Lynch M (2008) The cellular, developmental and population-genetic determinants of mutation-rate evolution. Genetics 180:933-43.
18. Johnson T, Barton NH (2002) The effect of deleterious alleles on adaptation in asexual populations. Genetics 162:395-411.
19. Good BH, Rouzine IM, Balick DJ, Hallatschek O, Desai MM (2012) Distribution of fixed beneficial mutations and the rate of adaptation in asexual populations. Proc Natl Acad Sci U S A 109:4950-5.
20. Zeyl C, Mizesko M, de Visser J a (2001) Mutational meltdown in laboratory yeast populations. Evolution 55:909-17.
21. Neher R a, Shraiman BI (2012) Fluctuations of fitness distributions and the rate of Muller's ratchet. Genetics 191:1283-93.
22. May, R. Lawton, J. Stork N (1995) Assessing Extinction Rates (Oxford University Press).
23. Wellings SR, Jensen HM, Marcum RG (1975) An atlas of subgross pathology of the human breast with special reference to possible precancerous lesions. J Natl Cancer Inst 55:231-73.
24. Vermeulen L et al. (2013) Defining stem cell dynamics in models of intestinal tumor initiation. Science 342:995-8.
25. Domingues JS (2012) Gompertz Model?: Resolution and Analysis for Tumors. J Math Model Appl 1:70-77.
26. Bachtrog D, Gordo I (2004) Adaptive evolution of asexual populations under Muller's ratchet. Evolution 58:1403-13.
27. Brunet E, Rouzine IM, Wilke CO (2008) The stochastic edge in adaptive evolution. Genetics 179:603-20.
28. Cancer T, Atlas G (2012) Comprehensive molecular characterization of human colon and rectal cancer. Nature 487:330-7.
29. Keightley PD, Kryukov G V, Sunyaev S, Halligan DL, Gaffney DJ (2005) Evolutionary constraints in conserved nongenic sequences of mammals. Genome Res 15:1373-8.
30. Kryukov G V, Schmidt S, Sunyaev S (2005) Small fitness effect of mutations in highly conserved non-coding regions. Hum Mol Genet 14:2221-9.
31. Haag-Liautard C et al. (2007) Direct estimation of per nucleotide and genomic deleterious mutation rates in Drosophila. Nature 445:82-5.
32. Tong a H et al. (2001) Systematic genetic analysis with ordered arrays of yeast deletion mutants. Science 294:2364-8.
33. Camps M, Herman A, Loh E, Loeb L a (2007) Genetic constraints on protein evolution. Crit Rev Biochem Mol Biol 42:313-26.
34. Geiler-Samerotte KA et al. (2011) Misfolded proteins impose a dosage-dependent fitness cost and trigger a cytosolic unfolded protein response in yeast. Proc Natl Acad Sci U S A 108:680-5.
35. Boyko AR et al. (2008) Assessing the evolutionary impact of amino acid mutations in the human genome. PLoS Genet 4:e1000083.
36. Bozic I et al. (2010) Accumulation of driver and passenger mutations during tumor progression. Proc Natl Acad Sci U S A 107:18545-50.
37. Howlader N et al. (2013) SEER Cancer Statistics Review (CSR) 1975-2010. EER Cancer Stat Rev 1975-2010, Natl Cancer Institute. Available at: http://seer.cancer.gov/csr/1975˙2009˙pops09/results˙single/sect˙01˙table.10˙2pgs.pdf.
38. Stephens PJ et al. (2012) The landscape of cancer genes and mutational processes in breast cancer. Nature 486:400-4.
39. Berger MF et al. (2012) Melanoma genome sequencing reveals frequent PREX2 mutations. Nature:1-18.
40. Ding L et al. (2008) Somatic mutations affect key pathways in lung adenocarcinoma. Nature 455:1069-75.
41. Bota S et al. (2001) Follow-up of bronchial precancerous lesions and carcinoma in situ using fluorescence endoscopy. Am J Respir Crit Care Med 164:1688-93.
42. Tong WWY et al. (2013) Progression to and spontaneous regression of high-grade anal squamous intraepithelial lesions in HIV-infected and uninfected men. AIDS 27:2233-43.
43. Page DL, Dupont WD, Rogers LW, Rados MS (1985) Atypical hyperplastic lesions of the female breast. A long-term follow-up study. Cancer 55:2698-708.
44. Haigh J (1978) The accumulation of deleterious genes in a population—Muller's ratchet. Theor Popul Biol 267:251-267.
45. Desai MM, Nicolaisen LE, Walczak AM, Plotkin JB (2012) The structure of allelic diversity in the presence of purifying selection. Theor Popul Biol 81:144-57.
46. Shaw AT, Engelman J a (2013) ALK in lung cancer: past, present, and future. J Clin Oncol 31:1105-11.
47. Ciriello G et al. (2013) The molecular diversity of Luminal A breast tumors. Breast Cancer Res Treat 141:409-20.
48. Chou T (2006) Theoretical basis, experimental design, and computerized simulation of synergism and antagonism in drug combination studies. Pharmacol Rev 58:621-81.
49. Jordan DM, Ramensky VE, Sunyaev SR (2010) Human allelic variation: perspective from protein function, structure, and evolution. Curr Opin Struct Biol 20:342-50.
50. Vesely MD, Schreiber RD (2013) Cancer immunoediting: antigens, mechanisms, and implications to cancer immunotherapy. Ann N Y Acad Sci 1284:1-5.
51. Wang ZC et al. (2012) Profiles of genomic instability in high-grade serous ovarian cancer predict treatment outcome. Clin Cancer Res 18:5806-15.
52. Rossi-Fanelli A, Cavaliere R, Mondov B, Moricca G (1977) Selective Heat Sensitivity of Cancer Cells (Springer Berlin Heidelberg, Berlin Heidelberg).
53. Goyal S et al. (2012) Dynamic mutation-selection balance as an evolutionary attractor. Genetics 191:1309-19.
54. Bell G, Gonzalez A (2009) Evolutionary rescue can prevent extinction following environmental change. Ecol Lett 12:942-8.
55. Mermel CH et al. (2011) GISTIC2.0 facilitates sensitive and confident localization of the targets of focal somatic copy-number alteration in human cancers. Genome Biol 12:R41.

Figures & Figure Legends

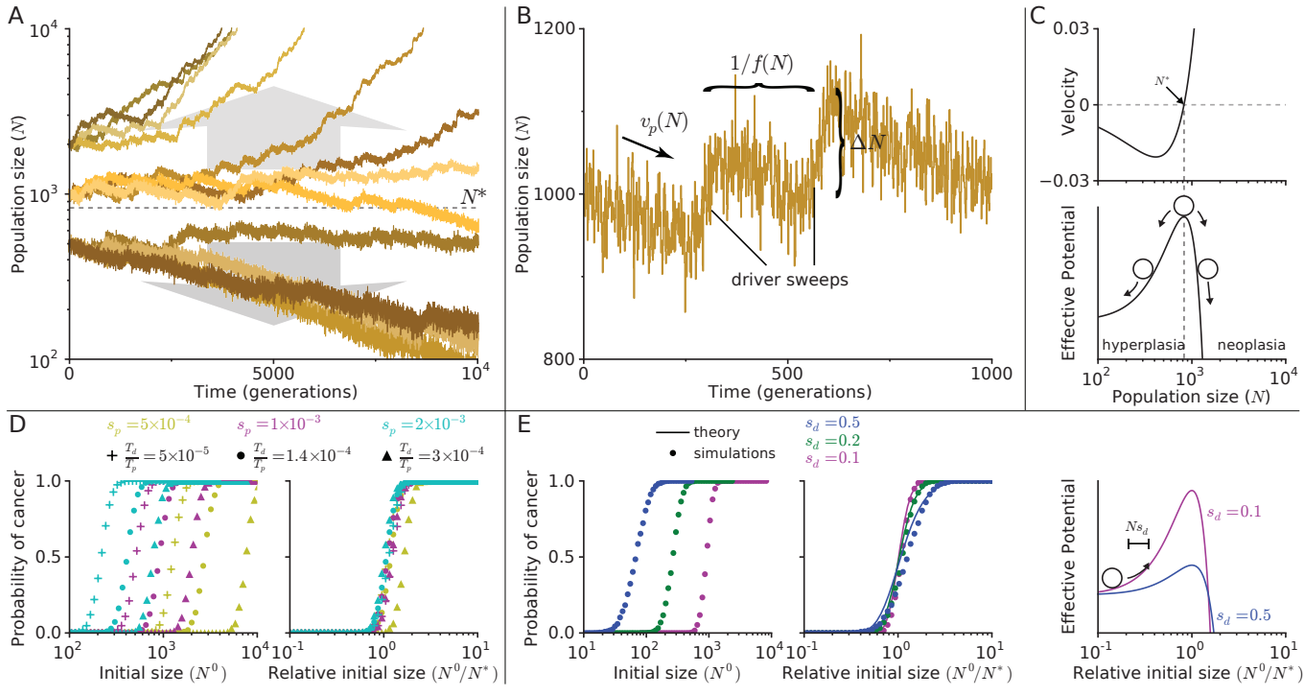

Fig. 1. Tug-of-war between drivers and passengers leads to a critical population size (A) Population size verses time of simulations initiated at various sizes. Populations starting above N^* generally commit to rapid growth (i.e. adaptation) and to extinction below it. (B) A fragment of a trajectory shows periods of rapid growth and gradual decline. New drivers arrive with a frequency $f(N)$ and abruptly increases the population size by an amount ΔN. Meanwhile, passenger accumulation causes populations to gradually decline with rate v_p. (C) Analytically computed mean velocity of population growth (top) and an effective barrier (bottom) as a function of population size N. The velocity is negative below N^* and positive above it. (D) The probability of adaptation (cancer) as a function of initial population size N (left) and a relative initial population size (N/N^*, right) for nine values of evolutionary parameters. Using the relative size N/N^* leads to curve collapse, whereby populations with different evolutionary parameters nevertheless behave similarly. (E) Same as in (D) for simulations and theory but for different values of s_d. Higher values of s_d, leads to more gradual transition from non-adaptive to adaptive regime. Excellent agreement between simulations and theory demonstrates accuracy of the theory. In our formalism, an increase in s_d results in a larger jump size ΔN and lower potential barrier, allowing more populations to overcome the barrier (right).

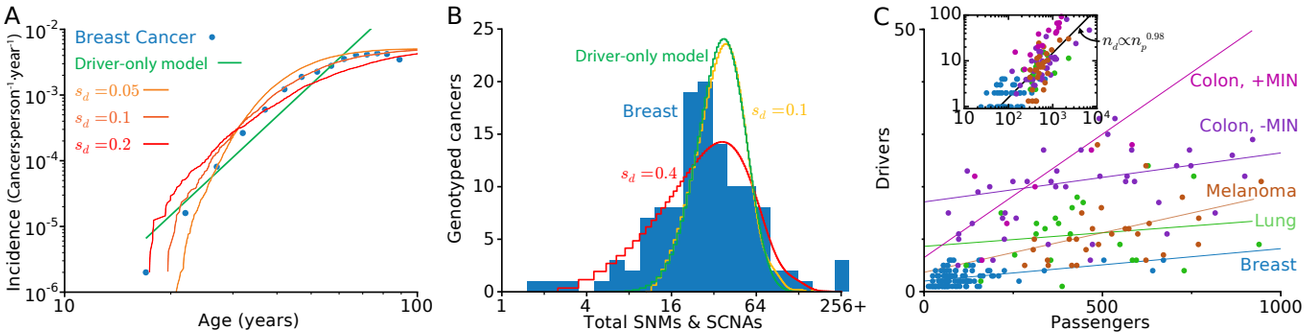

Fig. 2. Signatures of balance between drivers and passengers in incidence and genomics data (A) Predicted and observed breast cancer incidence rates verses age. Incidence rates in our model and the data both plateau at old age, while a traditional driver-only model ($I \propto t^k$) does not. (B) Histogram of the collective number of protein-coding mutations (SNMs) and alterations (SCNAs) in breast cancer alongside predicted distributions. Our model, captures the width and asymmetry of the distribution well for $s_d = 0.4$, while a driver-only model predicts a narrower and symmetric distribution. (C) The total number of driver verses the number of passenger alterations in sequenced tumors for several major subtypes. SCNAs and SNMs were aggregated. As predicted by the model, all subtypes exhibit a positive linear relationship between the number of drivers and passengers ($p < 0.08 - 10^{-5}$). A driver-only model with neutral passengers does not predict this linear relationship. (Insert) The same genomics data plotted on log axes, with the y-intercept from each subtype's linear fit subtracted. A linear relationship on logarithmic axes with a slope of approximately one suggests that the relationship between drivers and passengers is indeed linear.

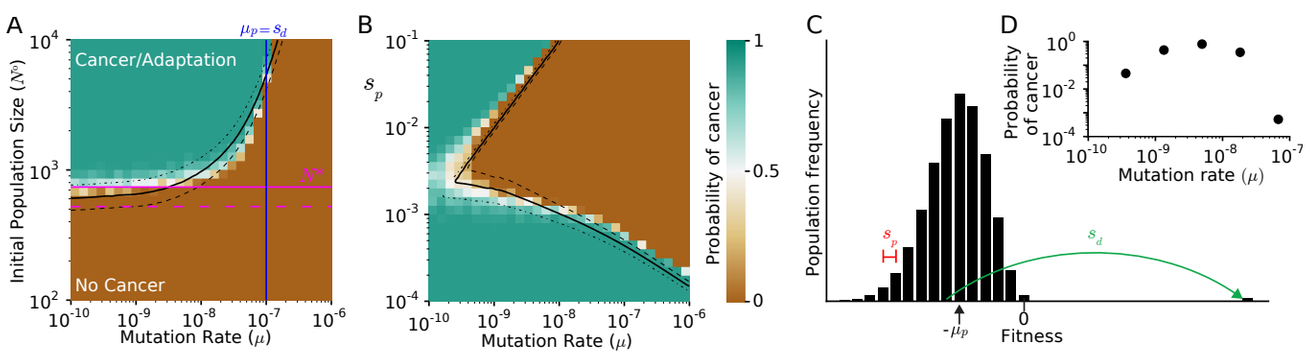

Fig. 3. Effect of mutation rate on cancer dynamics (A) The probability of cancer (adaptation) computed by simulations as a function of the initial population size and mutation rate. Evolutionary parameters roughly partition into a regime where cancer (adaption) is almost certain, and a regime where it is exceedingly rare. Estimates of N^* from our theory (magenta, solid) accurately predict the transition observed at low mutation rates. Another transition is observed as mutation rate exceeds the critical mutation rate, also predicted by theory (μ^*, blue line). Transition between these two regimes is better described by our theory when we incorporated (i) passenger interference with driver sweeps, and (ii) selection against passengers (black lines; dashed, solid, and dotted-dash predict 10%, 50%, and 90% probability of adaptation). (B) Cancer (adaption) probability (color) obtained by simulations as a function of mutation rate and s_p. The theory (black lines) accurately reproduces the complex transition between both regimes. (C) Diagram illustrating how the load of passengers influences the probability of fixation of a driver. The distribution of fitness is due to a distribution of the number of passengers per cell, which follows a Poisson distribution with a mean reduction in fitness of μ_p. Hence $s_p > \mu_p$ for a typical driver to outweigh the load of passengers. Segregating passengers not only reduce a driver's probability of fixation, but also its fitness benefit (26). (D) Probability of cancer for various mutation rates constrained to grow within 10,000 generations. We observe an optimum mutation rate.

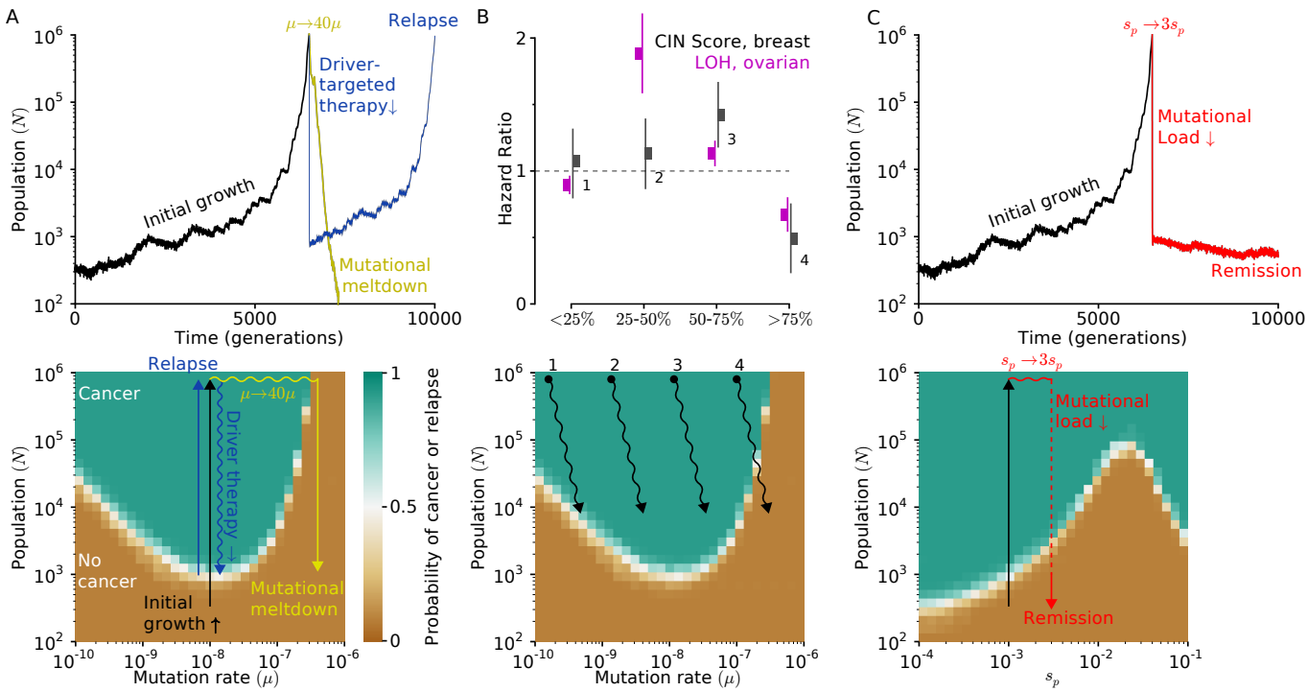

Fig. 4. Mapping and interpreting treatment outcomes (A) An adapted population (grown cancer) can be reverted to extinction by increasing the mutation rate (mutagenic chemotherapy) or by decreasing the population size (e.g. surgery or cytotoxic chemotherapy). Our phase diagrams explains therapeutic outcomes: therapies that reduce cancer size or increase mutation rate enough to push it outside of an adaptive regime cause continued population collapse; those that do not experience relapse. (B) (Top) Comparison of the model to clinical data. Cancers with intermediate mutational loads are the most aggressive (4, 13). Patients with intermediate Chromosomal INstability (CIN) and Loss Of Heterozygosity (LOH) scores are the least likely to survive. Patients with very high CIN are most effectively treated. (Bottom) Our phase diagrams shows these clinical outcomes: traditional therapies which work by decreasing the population size and/or increase the mutation rate work best for cancers with the highest mutation rate. (C) Three fold increase in the effect of passenger mutations leads to rapid population meltdown below N^*, thus relapse is unlikely.