RESEARCH ARTICLE

Punctuated mutagenesis promotes multi-step evolutionary adaptation in human cancers

Christopher J Graser[1,2,3], Wenbo Wu[1], Cole Christini[1,4], Mia Peljak[5,6], Franziska Michor[1,2,3,7,8,9]*

[1]Department of Data Science, Dana-Farber Cancer Institute, Boston, United States; [2]Department of Biostatistics, Harvard T.H. Chan School of Public Health, Boston, United States; [3]Department of Stem Cell and Regenerative Biology, Harvard University, Cambridge, United States; [4]School of Computer Science, Carnegie Mellon University, Pittsburgh, United States; [5]Department of Pathology, New York University School of Medicine, New York, United States; [6]Laura and Isaac Perlmutter Cancer Center, New York University, New York, United States; [7]Center for Cancer Evolution, Dana-Farber Cancer Institute, Boston, United States; [8]The Eli and Edythe L. Broad Institute, Cambridge, United States; [9]Ludwig Center at Harvard, Harvard Medical School, Boston, United States

***For correspondence:** michor@jimmy.harvard.edu

Competing interest: See page 14

Funding: See page 14

Preprint posted
04 July 2025
Sent for Review
30 June 2025
Reviewed preprint v1
02 October 2025
Reviewed preprint v2
05 December 2025
Version of Record
09 October 2026

Reviewing Editor: Anna Panchenko, Queen's University, Canada

© Copyright Graser *et al*. This article is distributed under the terms of the Creative Commons Attribution License, which permits unrestricted use and redistribution provided that the original author and source are credited.

eLife Assessment

This **valuable** study presents a theoretical model of how punctuated mutations influence multistep adaptation, supported by empirical evidence from some TCGA cancer cohorts. This **solid** model points to the case of possible punctuated evolution rather than gradual genomic change. There was some disagreement amongst the reviewers in terms of how closely the theoretical results apply to the phenomena examined empirically, and alternative explanations should be considered in the future.

Abstract The rate of acquisition of genomic changes in cancer has been the topic of much discussion, with several recent investigations finding evidence of punctuated evolution instead of gradual accumulation of such changes. Despite forays into the description and quantification of these punctuated events, the effects of such changes on subsequent cancer evolution remain incompletely understood. Here we investigate how non-gradual mutagenesis affects the ability of tumor cells to acquire and retain fitness-enhancing adaptations. We find that punctuated mutagenesis significantly facilitates adaptation in scenarios where adaptation requires crossing a fitness valley, that is when multiple mutations are required which individually are maladaptive but jointly confer a fitness advantage. By increasing the probability that multiple mutations occur in close succession, punctuation increases the chance that mutants in a fitness valley mutate further to reach a fitness peak before going extinct. Analyzing data from The Cancer Genome Atlas, we find that tumors with signatures of APOBEC mutagenesis, which has been shown to proceed in episodic bursts, exhibit patterns consistent with higher rates of crossing fitness valleys. Lastly, we characterize how the interplay between this enhanced ability to cross fitness valleys and adaptation-limiting effects of clonal interference affects overall adaptability in complex fitness landscapes.

Introduction

Tumors evolve via the acquisition of randomly arising genetic and/or epigenetic alterations and their selection (**Merlo et al., 2006**). The rate at which tumor cells accumulate such genomic changes to produce potentially adaptive innovation is thus an important determinant of the capacity of tumor cells to adapt to diverse selection pressures (**Yates and Campbell, 2012**; **Neinavaie et al., 2021**), affecting their propensity to progress locally, to metastasize, or to evolve resistance to therapeutic intervention. Several methods to quantify mutation rates have been developed (**Araten et al., 2005**; **Williams et al., 2018**; **Werner et al., 2020**), and proxies for mutation rates such as tumor mutation burden often feature in prediction models of therapeutic outcomes (**Galuppini et al., 2019**; **Sha et al., 2020**).

Most existing methods deal with constant mutation rates (**Werner et al., 2020**; **Altrock et al., 2015**), assuming gradual evolutionary change over time. However, accumulating evidence suggests that mutagenesis in cancer is fluctuating. Indeed, a recent study **Petljak et al., 2019** demonstrated in vitro that mutations associated with DNA-editing activity of APOBEC cytidine deaminases (**Petljak et al., 2022**) occur in episodic bursts, with more than 100-fold differences in the rate of APOBEC-associated mutations across otherwise identical cell culture replicates (**Figure 1A**). This episodicity observed in vitro aligned with patterns in previously published in vivo data investigating APOBEC mutagenesis in lung cancer (**Jamal-Hanjani et al., 2017**; **Lee et al., 2017**), and data from intestinal crypts (**Wang et al., 2023**) later confirmed explicitly that this episodic pattern also occurs in patients. Other recent studies using genomic data to reconstruct evolutionary histories of tumors have found evidence of punctuated evolution across several cancer types (**Navin et al., 2011**; **Sottoriva et al., 2015**; **Gao et al., 2016**). These studies have demonstrated the existence of distinct phases of mutation bursts (**Gao et al., 2016**; **Minussi et al., 2021**; **Figure 1B**), challenging the prevailing paradigm of gradual emergence of mutant lineages (**Figure 1C**). Various mutagenic processes such as chromothripsis (**Stephens et al., 2011**; **Zhang et al., 2015**), chromoplexy (**Baca et al., 2013**), and other drivers of chromosomal instability **Drews et al., 2022** have been implicated in causing such punctuated patterns (**Davis et al., 2017**).

How punctuated mutagenesis affects the evolutionary dynamics of tumor cell adaptation across complex fitness landscapes remains incompletely understood. Addressing this question with existing experimental data is inconclusive, as only short time horizons are observed (single expansions [**Petljak et al., 2019**]; time to a most recent common ancestor [**Gao et al., 2016**; **Minussi et al., 2021**]), and as fitness differences between observed cell lineages are incompletely characterized. Mathematical modeling of human cancer genomics data, however, can elucidate the dynamics of adaptation during punctuated tumor evolution. Here, we set out to systematically investigate the evolutionary consequences of punctuated mutation acquisition during tumorigenesis using mathematical modeling and analysis of genomic data from The Cancer Genome Atlas (TCGA).

Results

Temporal clustering facilitates multi-step adaptation

We set out to investigate how *temporal clustering* of mutagenic events into distinct episodes affects the ability of a cell population to achieve multi-step evolutionary adaptations. We considered two scenarios: tumor evolution via a single advantageous step, such as an oncogenic adaptation (**Chakravarty et al., 2017**) ('one-step adaptations', **Figure 1D**), and accumulation of multiple mutations with synergistic fitness effects ('multi-step adaptations', **Figure 1E**). Indeed, the two hit-hypothesis (**Knudson, 1971**) for tumor suppressor genes (TGSs) suggests that inactivation of one copy of a TSG can be inconsequential or maladaptive, while bi-allelic inactivation confers a selective advantage (**Iwasa et al., 2004**). While some TSGs exhibit (context-dependent) haploinsufficiency (**Paige, 2003**; **Wang et al., 2018**; **Park et al., 2021**), synergistic fitness effects for successive (epigenetic; **Paige, 2003**; **Wang et al., 2018**; **Issa, 2022**) alterations in the same gene remain the prominent feature of TSGs. Analogously, there are ample examples of successive alterations across different genes acting synergistically (**Gu et al., 2010**; **Hsu et al., 2010**; **Deming et al., 2014**). High mutation rates are known to limit the rate of retained one-step adaptations per emerging one-step adaptation due to clonal interference (**Gerrish and Lenski, 1998**; **Wilke, 2004**; **Fogle et al., 2008**; **Figure 1F**), which has been highlighted as a potential consequence of punctuated cancer evolution (**Sun et al., 2018**).

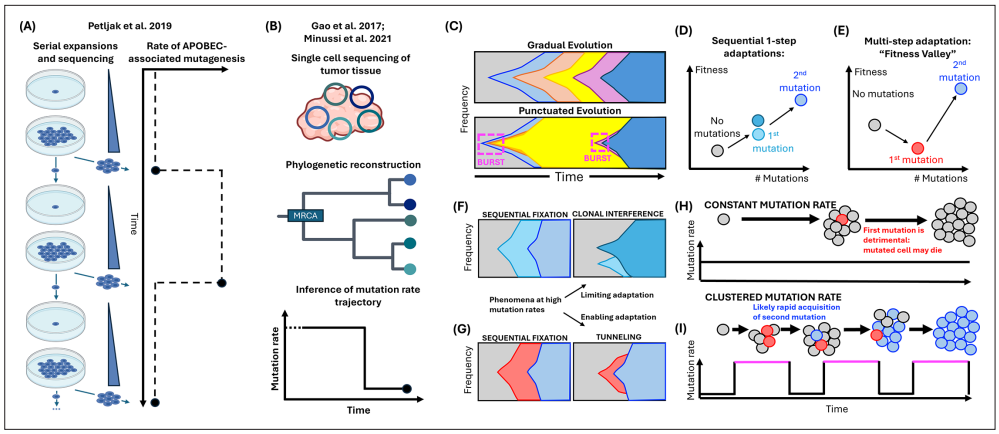


Figure 1. Population dynamics under gradual vs punctuated mutagenesis. (**A**) Evidence for punctuated APOBEC mutagenesis (***Petljak et al., 2019***) from successive cell culture expansions, seeded with single progenitors from the preceding expansion. Sequencing of the expanded populations revealed large fluctuations in APOBEC-associated mutagenic signatures SBS2 and SBS13. (**B**) Evidence for punctuated copy number evolution (***Gao et al., 2016***; ***Minussi et al., 2021***). Patterns in branch lengths of reconstructed phylogenetic trees reveal fluctuations in the rates of copy number alterations. Dynamics prior to the most recent common ancestor (MRCA) are unidentifiable. (**C**) Gradual vs punctuated evolution. In gradual evolution, novelty emerges and spreads at a constant rate over time. In punctuated evolution, novelty emerges and spreads during distinct burst phases. (**D**) Fitness schematic of two one-step adaptations which each independently confer a fitness advantage. (**E**) Fitness schematic for a two-step adaptation, in which carrying one mutation confers a fitness disadvantage, but carrying two mutations confers a fitness advantage. Schematic of possible evolutionary dynamics for two one-step adaptations (**F**) and of modes of valley crossing with or without prior fixation of the first mutant (**G**). Sequential fixation occurs at low mutation rates where emerging mutant lineages are likely to have fixated or gone extinct before the next mutation occurs. At higher mutation rates, the second mutation can occur in a multi-clonal population. (**H**) Sketch of failing two-step adaptation under a uniform (time-invariant) mutation rate. (**I**) Sketch of a successful two-step adaptation via stochastic tunneling under a punctuated mutation process with distinct clusters of high mutation rates.

The dynamics of multi-step adaptation have been studied extensively (***Iwasa et al., 2004***; ***Komarova et al., 2003***; ***Nowak et al., 2004***; ***Weinreich and Chao, 2005***; ***Weissman et al., 2009***; ***Proulx, 2011***; ***Ashcroft et al., 2015***; ***Van Egeren et al., 2018***), elucidating two modes of evolution (***Figure 1G***): 'sequential fixation' refers to the scenario in which the second mutation only emerges after cells harboring the first mutant have taken over the population, while 'stochastic tunneling' refers to situations in which the second mutation arises sooner. In large populations, mutants with a selective disadvantage become vanishingly unlikely to reach fixation, so that stochastic tunneling becomes the dominant mode of evolution for multi-step adaptations.

To investigate how the rate of successful stochastic tunneling events – enabling multi-step adaptation – is affected by punctuated mutagenesis, we simulated selection dynamics of a population of cells which proliferate according to a Wright–Fisher process (***Wright, 1931***; ***Fisher, 1923***) and accumulate mutations according to a constant (***Figure 1H***) or temporally clustered (***Figure 1I***) rate. The Wright–Fisher process models evolution as successive, non-overlapping generations of constant size. Each new generation is populated by drawing with replacement from the cells in the previous generation, with probabilities proportional to their fitness (***Figure 2A***). Cells can accumulate mutations that change their fitness, tha is increase their probability to survive to the next generation. We considered fitness landscapes in which two-step adaptations are the only way by which cells can increase their fitness, with fixed fitness ratios across subsequent two-step adaptations (***Figure 2B, C***). We also confirmed that the findings observed for the Wright–Fisher process, described below, are qualitatively consistent for expanding populations that proliferate according to a branching process model (Methods, ***Figure 2—figure supplement 1***).

Interestingly, for two-step adaptations for which the intermediate mutants that carry only one of the two required mutations have a strong selective disadvantage (***Figure 2B***), we found that the population undergoes substantially (3.55 times) more two-step adaptations when accumulating mutations

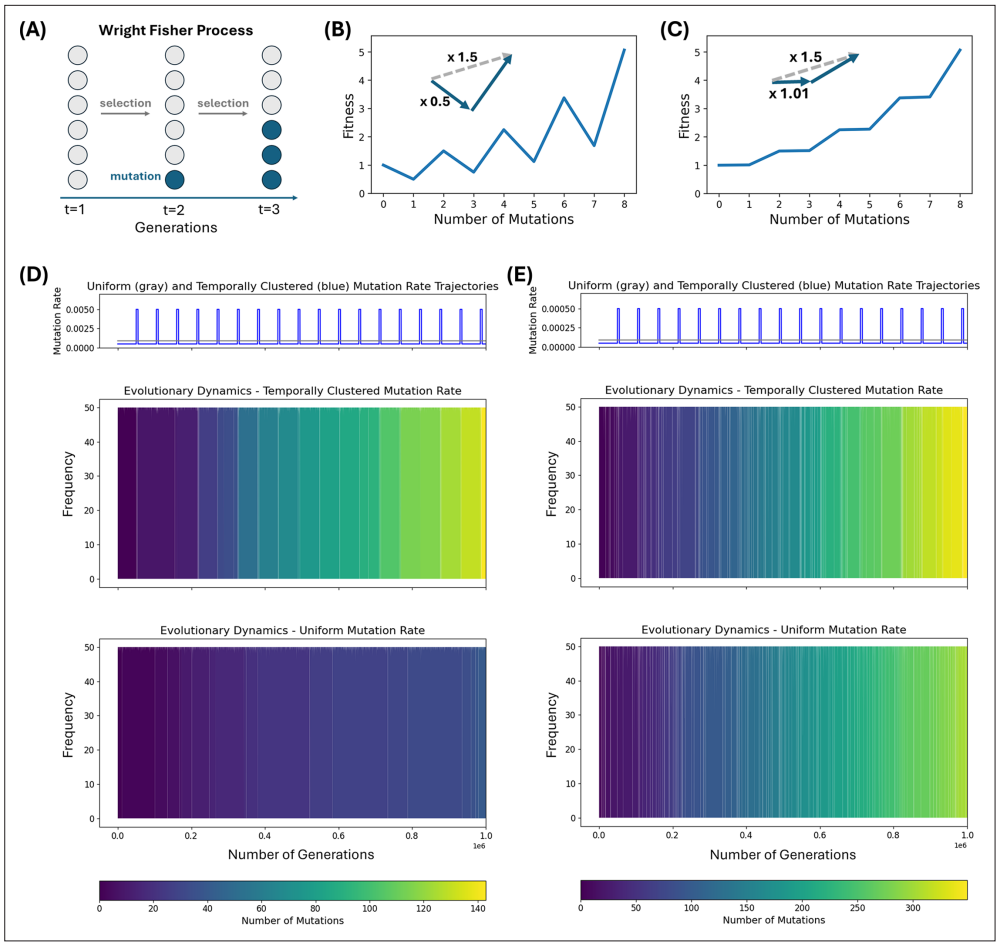


Figure 2. Simulation results: valley crossing under uniform vs temporally clustered mutation rates. (**A**) Schematic of a Wright–Fisher process. (**B, C**) Fitness landscapes used for the simulations in panels (**D**) and (**E**), respectively. Mutations move a cell from left to right through the landscape. Each two-step adaptation corresponds to a fitness increase by a factor of 1.5. Having an odd number of mutations comes at a multiplicative fitness disadvantage of 0.5 in (**B**) and an advantage of 1.01 in panel (**D**). (**D, E**) Heatmaps of simulation results for a Wright–Fisher process with 50 cells. The mutation rate trajectories in each panel are chosen such that the total expected number of mutations under the uniform trajectory is identical to that under the temporally clustered trajectory. Each column of pixels represents a population after a specific number of generations. Homogeneity in pixel-color per column thus indicates that cells with a given number of adaptations have fixated in the population, whereas persistent diversity of pixel colors (see ***Figure 2—figure supplement 1***) indicates coexistence of lineages with different numbers of adaptations.

The online version of this article includes the following figure supplement(s) for figure 2:

Figure supplement 1. Simulation results: valley crossing under uniform vs temporally clustered mutation rates in branching process.

Figure supplement 2. Effect of temporal clustering on stochastic tunneling rates as a function of the mutation rate.

Figure supplement 3. Effect sizes for two-step adaptations in a Galton–Watson process as a function of the mutation rate and the birth–death ratio.

Figure supplement 4. Effect sizes for two-step adaptations in a Galton–Watson process as a function of the distance *d* between subsequent burst phases.

in a temporally clustered rather than in a uniform way (***Figure 2D***). This effect arises because intermediate mutants in 'fitness valleys' have a high chance of going extinct before acquiring the next mutation. During mutation bursts, acquiring the next mutation in time before the disadvantageous mutant goes extinct becomes more likely. Similarly, intermediate mutants are more likely to emerge

in a mutation burst. This temporal clustering of the likelihood of two succeeding mutation events increases the rate of successful two-step adaptations.

This effect also emerges for two-step adaptations that do not constitute proper fitness valleys, that is for which the intermediate mutant is not maladaptive (***Figure 2E***; 1.15-fold increase). Under stochastic selection dynamics, even mutants with a slight selective advantage (***Figure 2C***) have a high chance of going extinct due to random drift. Lineages acquiring sets of synergistic mutations, thus, often do so without prior fixation of each intermediate mutant – via stochastic tunneling (***Iwasa et al., 2004***). In those regimes, temporal clustering of mutations therefore also increases the speed of adaptation, as confirmed with simulations (***Figure 2—figure supplement 2A–D***). Moreover, the proportion of such sets of synergistic mutations acquired via stochastic tunneling rather than sequential fixation also increases with temporal clustering (***Figure 2—figure supplement 2E***).

Our results generalize to multi-step adaptations with arbitrarily many steps (Appendix 1). To demonstrate this result, we compared a uniform mutation process with mutation rate μ to temporally clustered mutation processes in which mutations emerge exclusively during burst phases that constitute a fraction $\frac{1}{k}$ of the time, but at a k-fold increased rate, $k\mu$ (***Figure 2—figure supplement 2A***). We showed analytically that in the limit of the average mutation rate μ going to zero, the rate f_k of successful $(n+1)$-step adaptations is k^n times higher in the temporally clustered mutation process than in the uniform mutation process. We validated this result with simulations of a Wright–Fisher process for two-step adaptations, $n = 1$, and a range of small values for μ (***Figure 2—figure supplement 2F, G***), and with numerical calculations of f_k in a branching process (see Appendices 2 and 3 and ***Figure 2—figure supplements 3 and 4*** for complete characterizations of parameter dependencies). As the mutation rate increases, this fold-increase $\frac{f_k}{f_1}$ becomes smaller. However, since the absolute rates f_k and f_1 are proportional to μ^{n+1} (Appendix 1), the absolute effect of temporal clustering on the rate of fixations of mutants with multi-step adaptations increases in μ (***Figure 2—figure supplement 2F***). Analogously, we found that $\frac{f_k}{f_1}$ increases with the selective disadvantage that cells in a fitness valley have, while $f_k - f_1$ decreases with this selective disadvantage (Appendix 2). In contrast, for low μ, and particularly for parameter choices in biologically relevant ranges ($\mu \leq 10^{-7}$), the fold-change becomes insensitive to the magnitude of the selective disadvantage so that ever smaller fitness differences suffice to reach high fold-changes, $\frac{f_k}{f_1} \approx k^n$ (Appendix 2). For instance, in a population of constant size that evolves according to a branching process with a birth–death ratio of 1, if intermediate mutations in a multi-step adaptation sequence reduce the birth–death ratio to 0.98, we have $\frac{f_{50}}{f_1} = 49.4 = 0.988k$ for two-step adaptations, and $\frac{f_{50}}{f_1} = 2440.6 = 0.976\,k^2$ for three-step adaptations. In light of the large fold-changes in mutation rates during a burst and before or after a burst (***Petljak et al., 2019***; ***Navin et al., 2011***; ***Gao et al., 2016***; ***Minussi et al., 2021***), our findings suggest that the resulting effect on the dynamics of multi-step adaptation is substantial.

Temporal clustering in an exploration–exploitation setting

The increased rate of valley crossing is driven by phases of high mutation rates, when the fold-increase in the chance of valley crossing is larger than the fold-increase in the mutation rate. This effect improves adaptability while the fitness landscape offers scope for multi-step adaptation. However, in fitness landscapes which have global maxima, as the population adapts there is ever less scope for (further) multi-step adaptations, and excessive exploration of the landscape becomes costly.

To investigate the effect of temporal clustering in such an exploration-exploitation setting, we constructed a two-dimensional fitness landscape (***Figure 3A***) which is randomly re-drawn at regular intervals, mimicking exposure to novel physical environments or drugs during cancer evolution and treatment (Methods). Between subsequent re-drawing, the population may move to a global fitness maximum where any further mutation decreases fitness. We considered mutation processes with different average mutation rates μ and different values of the clustering parameter k (***Figure 3B***) and investigated the average fitness of the cell population over long time horizons. As before, the clustering parameter k denotes the fold-increase of the mutation rate in burst phases relative to the average mutation rate μ. However, rather than fixing the out-of-burst mutation rate at zero, the burst duration is now held constant (***Figure 3B***) to control for differences in waiting time until a burst occurs.

We found that, in this scenario, relative to the uniform mutation process ($k = 1$) increasing k initially increases the average fitness of the population for any fixed μ (***Figure 3C***). However, at high k, marginal increases of k reduce fitness, since high in-burst mutation rates in small populations increase

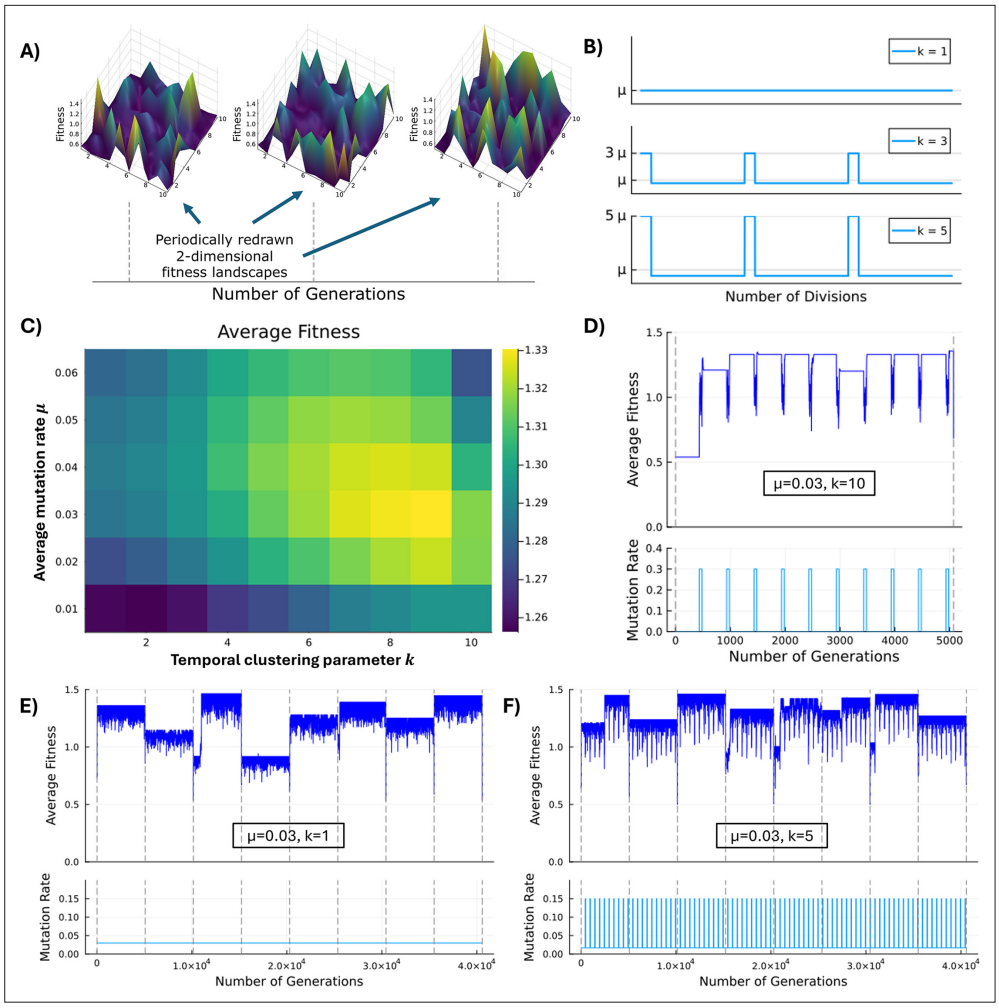


Figure 3. Exploration and exploitation with temporally clustered mutation rates. (**A**) Cells move through two-dimensional fitness landscapes. These fitness landscapes are randomly re-drawn every 50,714 divisions. (**B**) Mutation rate trajectories are parameterized with a mean mutation rate μ, and a clustering parameter k. (**C**) Average fitness in simulations of a population of 20 cells in a Wright–Fisher process. Simulations were run for 10^8 division events. (**D–F**) Average fitness trajectories for representative snippets of the simulations in (**C**).

the chance of leaving fitness peaks due to stochastic drift, and low mutation rates outside of bursts delay the acquisition of one-step adaptations. Similarly, for a fixed k, increasing μ when starting at low μ-values increases the average fitness. However, moving further in either of these directions in the parameter space – toward increasing μ or toward increasing k – eventually brings about a decrease in average fitness reflecting a trade-off between exploration and exploitation (see Appendix 4 and ***Figure 3D–F*** for detailed descriptions and plots of the dynamics). For any fixed k, average fitness peaks at intermediate μ, and the highest average fitness on the k–μ plane is achieved by a $k > 1$. Maximizing average fitness, thus, involves temporal clustering even when there is a trade-off between exploration and exploitation.

Proxies for valley crossing and for temporal clustering found in patient data

We then sought evidence of the modeling prediction that temporal clustering facilitates multi-step adaptation in patient data, leveraging whole-exome sequencing (WES) data from TCGA. We reasoned that among tumors in which similar numbers of mutations had emerged, those tumors with more temporally clustered mutation processes in their evolutionary history would be more likely to have undergone a larger number of multi-step adaptations. Generalizing this idea to

enable comparison of tumors which may differ in the number of mutations they have acquired, we reasoned further that tumors with high levels of temporal clustering would have a high fraction of successful multi-step adaptations per attempted multi-step adaptation, that is per emergence of a first mutant in a multi-step adaptation sequence. We would, thus, expect analogous ratios of detectable quantities that scale with the numbers of successful or attempted multi-step adaptations to correlate with indicators of fluctuations in the mutation rate history of a tumor such as those found for APOBEC mutagenesis. In particular, we considered the frequency of observing two deactivating mutations in a TSG as an indicator for successful multi-step adaptation, and the total number of observed mutations in a tumor as a quantity that scales with attempted multi-step adaptations.

Before considering the TCGA data, to confirm in silico that such a correlation would indeed emerge under biologically realistic parameters, we performed large-scale simulations of tumor growth from a single cell to up to realistically detectable cell numbers of 10^6–10^7 cells (***Fischer et al., 2006***) (Methods). We recorded the fraction of mutations that emerged during burst phases, constructed a TSG deactivation score (***Figure 4A, B***) by counting the number of TSGs (modeled as mutations with synergistic fitness effects) with at least two mutations across the population and divided by the total number of mutations across the genome (Methods). Our simulation results confirmed that these two readouts are correlated (Pearson correlation 0.42, $p < 0.01$; ***Figure 4C***), showing that in silico predictions regarding the likelihood of acquiring two-step adaptations under mutation processes with more vs less temporal clustering are well reflected in our score for TSG deactivation.

Next, to construct similar readouts from the TCGA data, we first performed SBS signature decomposition on each sample and computed the relative contribution of APOBEC-associated mutation signatures (SBS2 and SBS13) vs all other mutation signatures (***Alexandrov et al., 2013***) (Methods). Given the evidence for episodic APOBEC mutagenesis (***Petljak et al., 2019***; ***Wang et al., 2023***), we used this relative contribution as a score for temporal clustering. As a proxy for multi-step adaptations, a list of known TSGs (the TSGene 2.0 database; ***Zhao et al., 2016***) was considered, and for each TCGA sample, we counted the number of TSGs in this list which harbored at least two single-nucleotide substitutions classified as a missense or nonsense mutation (Methods) normalized by the total number of single-nucleotide substitutions in that sample. The TCGA WES data did not enable us to determine whether such mutations indeed deactivated different alleles of a TSG, rather than appearing on the same allele. If the likelihood of appearing on the same allele varied with the contribution of the APOBEC signature, this association could confound our results. Since mutations that are interdependently generated on the same copy of a gene would likely appear in close spatial proximity (***Petljak et al., 2022***; ***Nik-Zainal et al., 2012***; ***Mas-Ponte and Supek, 2020***), we thus verified in a robustness check that our results remain unchanged when filtering out mutations with close-by genomic coordinates (Methods).

We then ranked the different tumor types in TCGA by the average contribution of SBS2 and SBS13 relative to all detected mutation signatures, and for each tumor type, investigated the correlation between our proxies for multi-step adaptations and for temporal clustering. For the top four categories with the highest mean contribution of APOBEC-driven mutagenesis to all mutations, we observed a positive correlation between the two proxies. This correlation was significant at the 5% level (*t*-test on the Pearson correlation coefficient) for three out of four of these top four categories (***Figure 4D***, for results for categories with lower APOBEC signature contribution see ***Figure 4—figure supplement 1***).

Consistent with the hypothesis that this correlation arises because TSG deactivation occurs more readily with APOBEC mutagenesis, we found that mutations which contribute to the TSG deactivation score on average have a strongly increased likelihood of being caused by APOBEC (Methods, ***Figure 4E***). This finding is consistent across cancer types and independent of whether probabilities are averaged over samples or over individual mutations, suggesting that this result is not driven by outlier samples with many TSG inactivations and many APOBEC-associated mutations (Methods, ***Figure 4F***).

An alternative score for multi-step adaptations, in which the list of TSGs is replaced with a list of pairs of genes with synergistic fitness effects (***Gu et al., 2010***) (Methods), again exhibits a positive correlation with the relative contribution of APOBEC-associated mutation signatures in all categories; this correlation is significant for two of these categories (***Figure 4—figure supplement 2***). Additionally, we verified that our results are robust to the choice of correlation metric: using Spearman instead

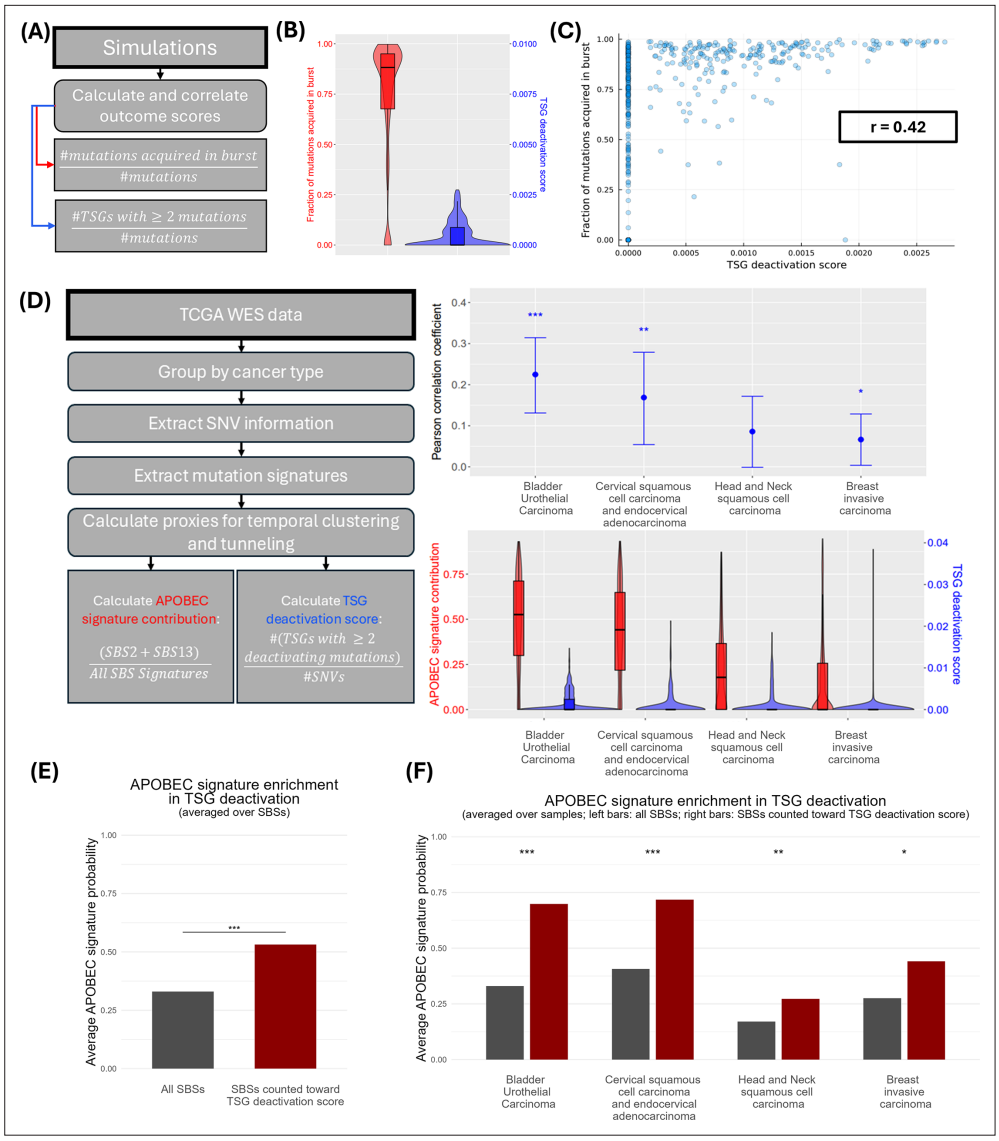


Figure 4. Proxies for valley crossing and temporal clustering in simulations and in TCGA data. (**A**) Sketch of the simulation analysis workflow. (**B**) Distributions of the fraction of mutations acquired during mutation bursts and the TSG deactivation score in simulations. (**C**) Joint distribution of the quantities in (**B**) indicates strong correlation. (**D**) Schematic of analysis workflow for TCGA data (methods) and results for the four cancer types with highest APOBEC signature contribution. (**E**) Probabilities that single base substitutions (SBSs) in samples of the cancer types in (**D**) were caused by APOBEC-associated signatures SBS2 and SBS13. On the left, these probabilities are averaged over all SBSs. On the right, only those SBSs are considered which are classified as 'nonsense' or 'missense', and which appear in a TSG with at least two such deactivating SBSs. The average probabilities between both groups are significantly different (*t*-test, $p < 0.001$). (**F**) Average probabilities analogous to those in panel (**D**) were constructed for each individual sample and then averaged across samples. Samples without deactivated TSGs were excluded. The average probabilities between groups are significantly different for the four cancer types (*, **, and ***, respectively, indicate that the p-value in a *t*-test lies below 0.05, 0.01, and 0.001).

The online version of this article includes the following figure supplement(s) for figure 4:

Figure supplement 1. TSG deactivation scores in TCGA.

Figure supplement 2. Synergistic mutation scores in TCGA.

Figure supplement 3. Robustness to choice of correlation coefficient.

Figure supplement 4. Robustness check: TSG deactivation and ROS in TCGA.

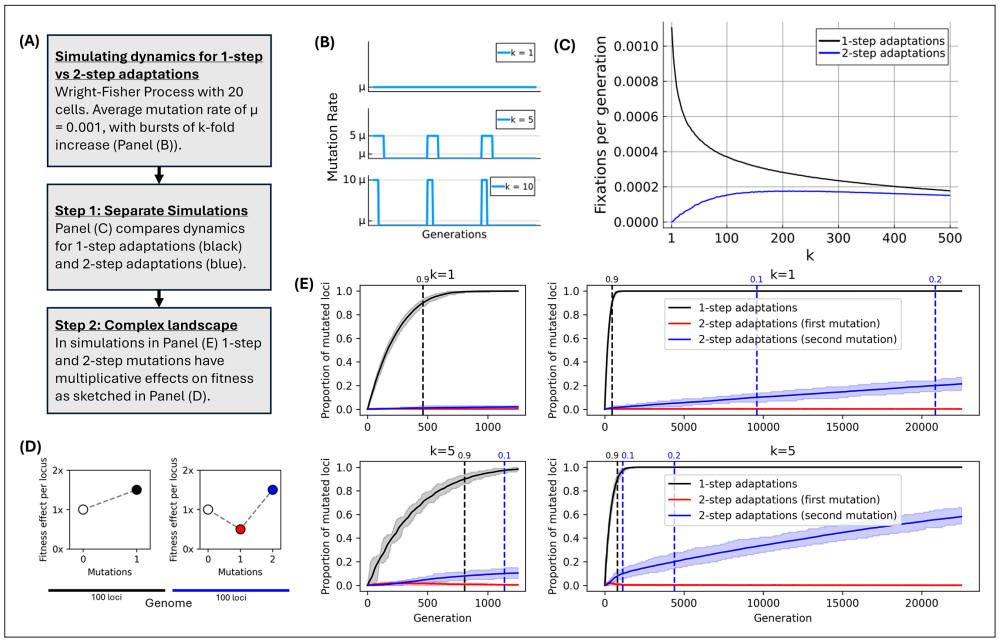


Figure 5. Stochastic tunneling vs clonal interference under temporally clustered mutation rates. (**A**) Schematic of the simulation approach. (**B**) Sketch of the mutation process. (**C**) Simulation results: measuring fixations of higher-fitness mutations per generation as a function of the clustering parameter and of the fitness landscape. In the one-step adaptation setting, fitness is defined as $1.5^{\#mutations}$. Fitness in the two-step adaptation setting is defined as in ***Figure 2B***. Results are averaged over simulation runs with 10^7 generations. (**D**) Sketch of fitness landscape for simulations shown in (**E**). Cells start with an unmutated genome of 200 loci. Mutations on each locus have independent multiplicative effects on cell fitness. In the first 100 loci, a single mutation confers a multiplicative fitness change of 1.5. In the remaining 100, two mutations are required to reach this multiplicative fitness change of 1.5, with the first mutation conferring a multiplicative fitness change of 0.5. (**E**) Simulation results for adaptation with genome sketched in (**D**). Results are averaged over 100 simulation runs per value of *k*. Shaded regions indicate 10th and 90th percentiles. Smaller plots (left) are zoomed-in versions of the first 1200 generations of the larger plots (right), with identical color-coding.

The online version of this article includes the following figure supplement(s) for figure 5:

Figure supplement 1. Temporal clustering and adaptation in fitness landscapes with varying ruggedness.

of Pearson correlations yields significant positive correlations across all top four categories discussed above (***Figure 4D***, ***Figure 4—figure supplement 3***).

Taken together, our analyses of TCGA data show that tumors with a larger share of APOBEC-associated mutations are enriched for pairs of mutations with synergistic fitness effects. While we cannot exclude the influence of unobserved confounders, given the limitations of the available data, we found that the patterns from our analysis are consistent with the theoretical predictions of our model.

Stochastic tunneling vs clonal interference during punctuated evolution

In tumors evolving at high mutation rates, clonal interference may limit the rate at which a cell population manages to acquire and retain one-step adaptations (***Figure 1F***). Previous literature has quantified this limiting effect as a function of the mutation rate and the distribution of fitness values of emerging mutants (***Campos et al., 2004***; ***Campos and de Oliveira, 2004***; ***Park and Krug, 2007***). However, the effects of non-constant mutation rates in this setting have not been elucidated. We thus set out to investigate clonal interference in the setting of temporally clustered mutation rates (***Figure 5A***).

We first considered a scenario in which cells can only acquire one-step adaptations and measured the rate at which novel one-step adaptations reach fixation in the population across simulations with different clustering parameters *k* (***Figure 5B***). We found that fixation rates quickly decrease as *k* increases (***Figure 5C***). This pattern emerges because the extent to which clonal interference reduces

fixation rates disproportionately increases with the mutation rate; for a fixed average mutation rate, temporal clustering thus increases the effects of clonal interference.

These findings stand in contrast to our results for two-step adaptations whose rate of fixation increases when temporal clustering is introduced. However, if in-burst mutation rates are sufficiently large such that multiple clones in the population may acquire a two-step adaptation independently before one of these clones has fixated, effects of clonal interference also play a role for two-step adaptations. Performing simulations for two-step adaptations, we found that fixation rates are non-monotone in k. While at low k increasing k leads to a steep increase in the fixation rate, this trend eventually levels off and becomes negative, with further increases in k leading to a decrease in the fixation rate (***Figure 5C***).

Having observed that temporal clustering increases effects of clonal interference but also facilitates stochastic tunneling, we next set out to investigate the relative contribution of these two effects on overall adaptation rates. We considered a setting in which cells can increase their fitness through both one- and two-step adaptations (***Figure 5D***). We found that one-step adaptations are acquired more quickly under the uniform mutation process ($k = 1$), while two-step adaptations are acquired more quickly under a temporally clustered mutation process ($k = 5$; ***Figure 5E***). The relative magnitude of both effects varies with time. As one-step adaptations spread more readily in the population, differences in the speed at which they fixate manifest early in the dynamics. Over time, cells gradually exhaust the possibilities for one-step adaptations and the difference between the adaptation rate in both processes becomes dominated by differences in the speed of acquiring two-step adaptations. In this latter phase we observe large differences between the two mutation processes in the expected time until any given proportion of the possible two-step adaptations have spread in the population. Analogous time differences for reaching fixed proportions of one-step adaptations in the initial phase of the dynamics are much smaller (***Figure 5E***).

In these analyses, parameters such as mutation rates, fitness effects, and population sizes were chosen so that simulations remained computationally manageable. Decreasing mutation rates toward more biologically realistic ranges decreases the likelihood that multiple lineages with adaptations emerge simultaneously and therefore weaken the effects of clonal interference. In contrast, larger population sizes and smaller fitness differences increase the likelihood of co-existence between competing lineages as each emerging lineage requires more time to fixate in the population, which strengthens the effects of clonal interference.

We also investigated how this pattern depends on the ruggedness of the fitness landscape (***Figure 5—figure supplement 1A***). To this end, we varied the relative proportions of loci with and without fitness valleys and quantified adaptability differences. We found that the duration of the initial phase in which the uniform process yields higher adaptability varies with the ruggedness of the fitness landscape: this phase is roughly three times as long when the proportion of loci with fitness valleys is 5% compared to when it is 95% (***Figure 5—figure supplement 1B–D***). However, over this range of proportions, we consistently observed that this initial phase only accounts for a small fraction of the time it takes to acquire all adaptations.

Discussion

Accumulating evidence suggests that mutagenesis in tumor cell populations proceeds in punctuated bursts rather than gradually; however, the effects of such punctuation on the ability of a tumor cell population to acquire and retain fitness-enhancing adaptations remain incompletely understood. Here, we set out to investigate the evolutionary dynamics of these processes using mathematical modeling and genomics data analysis of human tumors.

We found that when a population acquires multi-step adaptations via stochastic tunneling, the rate of adaptation is substantially enhanced by punctuation. Stochastic tunneling is the predominant mode of evolution in large populations that traverse fitness valleys. However, stochastic tunneling also occurs when the necessary mutation steps to reach a substantial fitness advantage are not individually maladaptive, that is if steps do not constitute a fitness valley but confer a slight fitness advantage. Punctuated evolution therefore facilitates the accumulation of sets of mutations that jointly and synergistically confer a fitness advantage, while the fitness effect of carrying only a subset of these mutations can range from being strongly maladaptive to being moderately adaptive. For the limit of low average mutation rates, we show analytically that the rate of stochastic tunneling under

a temporally clustered process is k^n times larger than under the time-invariant process, where k is the fold-increase in the mutation rate relative to the average mutation rate during mutation bursts in the temporally clustered process, and n is the number of mutation steps that is required to exit the fitness valley. Moving toward higher average mutation rates, this fold-change in the stochastic tunneling rates decreases, but the absolute difference between the rates increases. Investigating these relative and absolute effects in a branching process as a function of the birth–death ratio elucidated a similar pattern: While the fold-change is highest if lineages in a fitness valley have a low birth–death ratio, we observed the largest difference between rates of stochastic tunneling at birth–death ratios close to one.

Applied to cancer evolution where cells' birth–death ratios are influenced by a variety of internal and external factors, these findings highlight that for any given multi-step adaptation sequence, the magnitude of the facilitating effects described here likely varies depending on whether the tumor is in a phase of constant or growing population size.

While temporal clustering facilitates multi-step adaptations, it also impedes one-step adaptations due to clonal interference. We examined the interplay between these contrasting effects in a setting in which cells can acquire both types of adaptations. We showed that the relative importance of both effects varies over time. Since one-step adaptations tend to be acquired more readily, clonal interference matters most in the early phase of the process. As the share of possible two-step adaptations relative to one-step adaptations increases, differences in the tunneling rate become the more relevant determinant for the speed of adaptation. A uniform mutation rate, thus, makes a population more efficient at finding local fitness maxima. However, temporal clustering allows the population to more quickly move between local maxima, thus speeding up the search for a global maximum.

While our theoretical results are easily verified in simulation settings where we can choose a fitness landscape, applying these results to real data remains challenging because of the inherent complexities in determining fitness effects in biological systems. Estimating fitness effects of individual mutations requires either intricate experimental approaches (***Salehi et al., 2021***) or large patient cohorts to control for differences in the (epi-)genetic background against which these mutations emerge; such estimations become considerably harder when considering joint fitness effects of sets of multiple mutations.

To circumvent these challenges, we focused on TSGs as a representative set of gene pairs with synergistic effects and on one mutation process known to tend to fluctuate over time – APOBEC-driven mutagenesis. Moreover, we restricted attention to single-nucleotide substitutions as the principal mode of APOBEC-driven mutagenesis and the context in which mutation signatures are best characterized, and we did not consider other classes of mutations such as copy number changes or epigenomic modifications. Consequently, we observed only a subset of groups of mutations with synergistic fitness effects in the data and possibly only a small fraction of the variability in punctuation between different samples. Nevertheless, we found that these scores significantly correlate, which aligns with our model predictions. However, given this narrow scope, the generalization of our model toward a full quantitative estimate of the relevance of punctuation to adaptation dynamics in cancer requires additional investigations.

The observed empirical patterns may in part reflect the influence of unmeasured confounders. Rates of multi-step adaptation are determined not only by temporal clustering, but also by factors such as average mutation rates, fitness landscapes, and population-size trajectories. Structural differences in these factors between tumors with high or low APOBEC signature contributions therefore represent potential sources of confounding (for a review of known APOBEC-related biological mechanisms and clinical associations, see ***Butler and Banday, 2023***). Additionally, we cannot measure rates of multi-step adaptation directly, and therefore need to consider a proxy; to this end, we chose a score that relates deactivating TSG mutations to the absolute mutation count. Structural differences affecting the relative rates of deactivating TSG mutations, such as APOBEC-related enrichment for nonsense mutations (***Adler et al., 2023***), compared to other mutations can therefore also influence our empirical findings.

Our results have implications for both mechanistic and statistical modeling of tumor evolution. A key quantity in mechanistic models is the rate at which different types of adaptations arise; our results show that this rate depends on the temporal dynamics of the mutation process. Accounting for this effect might be particularly important when modeling the emergence of treatment resistance

in models used to optimize treatment schedules (***McDonald et al., 2023***). There is evidence linking APOBEC mutagenesis to resistance in breast cancer (***Gupta et al., 2025***). Our results suggest that APOBEC, and other punctuated mutation processes, are likely to be relevant more generally to settings in which developing resistance requires multi-step adaptation. Analogously, in statistical models, incorporating measures of punctuated mutagenesis may improve the ability to predict treatment outcomes by accounting for how these temporal dynamics shape adaptability.

Methods

Simulations of a Wright–Fisher and a branching process in an unbounded fitness landscape

In the simulations presented in ***Figure 2***, ***Figure 2—figure supplement 1***, mutations occur between selection steps and are independent of divisions. In the Wright–Fisher process simulations (***Figure 2C***) with a temporally clustered mutation rate, the out-of-burst mutation probability per cell and per update step (between two divisions) was set to 0.01 and was increased to 0.1 in bursts. For the corresponding uniform mutation rate trajectory, this probability was simply set equal to the total amount of mutations that occurred in the simulations under the temporally clustered mutation rate, divided by the total number of update steps.

Mutation rates in the branching model simulations (***Figure 2—figure supplement 1***) were chosen analogously. However, to achieve temporally equidistant mutation bursts in the branching process simulations, we scaled the burst duration and the time between bursts by the population size.

Simulations Wright–Fisher process in exploration/exploitation setting

To produce the results shown in ***Figure 3***, we performed agent-based stochastic simulations of a Wright–Fisher process with 20 cells. Cells were characterized by their location on a discrete fitness landscape (a 30 by 30 two-dimensional lattice). The fitness values f associated with the positions on the lattice were drawn independently at random as $f = 0.5 + x^4$ where $x \sim U[0,1]$ is distributed uniformly between zero and one. The lower bound of 0.5 ensures that cells at all positions have a non-negligible probability of dividing.

Raising x to the fourth power creates a landscape with few peak-positions surrounded by many valley-positions with little variation in fitness.

In each selection step, one cell was randomly chosen to divide with probability proportional to its fitness, and replaced a cell which was drawn uniformly at random. Mutations occurred between consecutive selection steps, and the mutating cell was chosen uniformly at random amongst all cells in the population (independently of the preceding selection step). If a cell mutated, it would move to a location in the fitness landscape chosen uniformly at random from the set of (at most 8) locations in the Moore neighborhood of its current location.

The rate at which mutation events occurred was governed by the parameters μ and k. Mutation bursts lasted 10^2 division events and started every 10^3 division events. Fitness landscapes were re-drawn every 50,714 ($\approx 50,000 + \frac{10^3}{7}$) division events, to periodically vary the relative timing of mutation bursts and re-drawings of the fitness landscape. In this manner, we avoid artifacts in our results caused by phase-alignment of bursts and re-drawings.

Larger-scale simulations

We simulated tumor evolution as a branching process, starting from a single unmutated cell, up to a randomly drawn target cell number between 10^6 and 10^7 cells. We assume a constant death rate, so that over the course of a simulation the probability that the next event is a death event rather than a division event remains fixed (at a value of 0.35). In case of a death event, one cell picked uniformly at random is removed from the population. In case of a division event, one cell is chosen to divide with probabilities proportional to its fitness.

We assume that the number of mutations per division follows a binomial distribution $Bin(10^5, \mu)$, and for each simulation we randomly draw a baseline mutation probability μ uniformly between 0 and 3×10^{-5}. In mutation bursts, this probability gets multiplied by a factor of 20 or 50, again chosen randomly for each simulation. The population enters a burst phase with probability $\frac{n}{10}$ per division, where n is the current population size, and exits a burst phase with probability $\frac{n}{3}$, so that in expectation

bursts phases last 3 generations, and start every 10 generations. The first cell in a simulation is initialized to be in a burst with probability $\frac{3}{13}$ consistent with the expected time spent in bursts given those parameters.

The fitness effect of mutations is modeled as follows. The starting cell has a fitness of one. Each new mutation has an additive effect on the cell's fitness. We assume that mutations occur uniformly at random across the genome and that multi-step adaptations can occur in a fraction $\theta = 0.01$ of the genome, roughly aligning with the ratio of the number of TSGs vs the total number of genes in the human genome. We subdivide this part of the genome into 200 TSGs which are all hit by a mutation with equal probability. For each individual TSG, the first mutation reduces a cell's fitness by 0.05. The second mutation increases fitness by 0.15, and all further mutations are fitness neutral. For the remaining $(1 - \theta)$ fraction of the genome, we assume that mutations are fitness-neutral with a probability of 0.9, and that fitness effects are otherwise drawn from a Gaussian with mean equal to −0.005 and standard deviation equal to 0.005.

To produce the results in ***Figure 4A, B***, we keep track of all mutations that arise in a population, and remove all mutations with less than 1% variant allele frequency from the output, as those would be unlikely to be detected in WES. Moreover, we keep track of the fraction of mutations that a sample acquired during a mutation burst. For each simulation, we then count the number of TSGs for which there are at least two mutations found in the population and divide this by the total number of mutations in the sample.

Analysis of TCGA WES data

WES data was acquired from TCGA. We performed mutation signature decomposition using the *cosmic fit*() function in Python from the package SigProfilerAssignment (***Díaz-Gay et al., 2023***) version 0.1.8 with cosmic version 3.4. Moreover, we used the mutation-level signature probabilities generated via SigProfilerAssignment to compute mean probabilities of signatures SBS2 and SBS13 for individual SBSs (***Figure 4D, E***).

To compute our TSG deactivation score, we filtered the SNV data for missense and nonsense mutations in genes belonging to the TSGene 2.0 database (***Zhao et al., 2016***). For each sample we calculated the number of TSGs with at least two mutations and divided by the total number of SNVs.

Analogously, to construct our synergistic mutations score, we filtered for missense and nonsense mutations in frequently co-mutated gene pairs identified by ***Gu et al., 2010***, and for each sample divided the number of pairs in this list with mutations in both genes by the number of SNVs (***Park et al., 2021***). This enrichment for co-occurrences of mutations in both genes in a pair suggests that mutations in the two genes tend to have a synergistic fitness effect. For each sample in the TCGA WES data we thus computed the number of gene pairs from this list for which there is at least one non-synonymous single-nucleotide substitution mutation in each of the two genes, and divided this number by the total number of single-nucleotide substitutions in the sample.

As a robustness check, we investigated whether the results change if we require a minimum distance between the genomic locations of the mutations in TSGs that we count to our TSG deactivation score. APOBEC has been linked to clustered mutagenesis through mutation processes of kataegis and omikli (***Petljak et al., 2022***; ***Nik-Zainal et al., 2012***; ***Mas-Ponte and Supek, 2020***). If samples with higher APOBEC activity have higher rates of having multiple close-by mutations on the same allele, and if such a sample has multiple deactivating mutations in a TSG, these inactivating mutations might be more likely to be localized on the same allele compared to samples with lower APOBEC activity. Since we use appearances of more than one inactivating mutation as a proxy for bi-allelic deactivation of TSGs, such a pattern would bias our results. To account for that, we explored filtering TSG mutations in the data based on different minimum-distance-thresholds ranging from 1 to 100 bp and found no impact on our results. We did not perform any analyses with allele-specific mutation information, as determining allele-specific TSG inactivation relies on informative co-occurring mutations. For any given mutation, tumors with lower mutation burden are less likely to harbor such co-occurring mutations, and preferentially filtering out observations in tumors with lower mutation burden would have biased our findings.

Additional information

Competing interests

Mia Peljak: Shareholder in Vertex Pharmaceuticals and a compensated consultant for the GLG Network. Franziska Michor: Co-founder and consultant of Harbinger Health, a consultant for Zephyr AI, and on the board of directors of Recursion Pharmaceuticals; none of these relationships are directly or indirectly related to the content of this manuscript. The other authors declare that no competing interests exist.

Funding

No external funding was received for this work.

Author contributions

Christopher J Graser, Conceptualization, Data curation, Formal analysis, Supervision, Investigation, Visualization, Methodology, Writing – original draft, Writing – review and editing; Wenbo Wu, Investigation, Visualization, Methodology, Writing – original draft; Cole Christini, Data curation, Investigation, Visualization, Methodology, Writing – original draft; Mia Peljak, Writing – review and editing; Franziska Michor, Conceptualization, Supervision, Writing – original draft, Writing – review and editing

Author ORCIDs

Christopher J Graser https://orcid.org/0009-0003-8552-7767
Franziska Michor https://orcid.org/0000-0003-4869-8842

Ethics

All data used in this study were obtained from The Cancer Genome Atlas (TCGA, https://www.cancer.gov/tcga), which provides de-identified data collected under established ethical guidelines, with informed consent from participants and approval by relevant institutional review boards.

Peer review material

Version of Record:
Reviewer #1 (Public review): https://doi.org/10.7554/eLife.108058.3.sa1
Reviewer #2 (Public review): https://doi.org/10.7554/eLife.108058.3.sa2
Author response https://doi.org/10.7554/eLife.108058.3.sa3

Reviewed preprint v2
Reviewer #1 (Public review): https://doi.org/10.7554/eLife.108058.2.sa2
Reviewer #2 (Public review): https://doi.org/10.7554/eLife.108058.2.sa1
Author response: https://doi.org/10.7554/eLife.108058.2.sa0
eLife Assessment https://doi.org/10.7554/eLife.108058.2.sa3

Reviewed preprint v1
Reviewer #1 (Public review): https://doi.org/10.7554/eLife.108058.1.sa2
Reviewer #2 (Public review): https://doi.org/10.7554/eLife.108058.1.sa1
Author response: https://doi.org/10.7554/eLife.108058.1.sa0
eLife Assessment https://doi.org/10.7554/eLife.108058.1.sa3

Additional files

Supplementary files

MDAR checklist

Source data 1. Data to recreate the figures.

Data availability

The results shown here are in part based upon data generated by the TCGA Research Network: https://www.cancer.gov/tcga. Simulation code is publicly available at https://github.com/Michorlab/Punctuated-mutagenesis (copy archived at ***Graser, 2026***).

References

Adler N, Bahcheli AT, Cheng KCL, Al-Zahrani KN, Slobodyanyuk M, Pellegrina D, Schramek D, Reimand J. 2023. Mutational processes of tobacco smoking and APOBEC activity generate protein-truncating mutations in cancer genomes. *Science Advances* **9**:eadh3083. DOI: https://doi.org/10.1126/sciadv.adh3083, PMID: 37922356

Alexandrov LB, Nik-Zainal S, Wedge DC, Aparicio SAJR, Behjati S, Biankin AV, Bignell GR, Bolli N, Borg A, Børresen-Dale A-L, Boyault S, Burkhardt B, Butler AP, Caldas C, Davies HR, Desmedt C, Eils R, Eyfjörd JE, Foekens JA, Greaves M, et al. 2013. Signatures of mutational processes in human cancer. *Nature* **500**:415–421. DOI: https://doi.org/10.1038/nature12477, PMID: 23945592

Altrock PM, Liu LL, Michor F. 2015. The mathematics of cancer: integrating quantitative models. *Nature Reviews. Cancer* **15**:730–745. DOI: https://doi.org/10.1038/nrc4029, PMID: 26597528

Araten DJ, Golde DW, Zhang RH, Thaler HT, Gargiulo L, Notaro R, Luzzatto L. 2005. A quantitative measurement of the human somatic mutation rate. *Cancer Research* **65**:8111–8117. DOI: https://doi.org/10.1158/0008-5472.CAN-04-1198, PMID: 16166284

Ashcroft P, Michor F, Galla T. 2015. Stochastic tunneling and metastable states during the somatic evolution of cancer. *Genetics* **199**:1213–1228. DOI: https://doi.org/10.1534/genetics.114.171553, PMID: 25624316

Azevedo RBR, Olofsson P. 2021. A branching process model of evolutionary rescue. *Mathematical Biosciences* **341**:108708. DOI: https://doi.org/10.1016/j.mbs.2021.108708, PMID: 34560091

Baca SC, Prandi D, Lawrence MS, Mosquera JM, Romanel A, Drier Y, Park K, Kitabayashi N, MacDonald TY, Ghandi M, Van Allen E, Kryukov GV, Sboner A, Theurillat J-P, Soong TD, Nickerson E, Auclair D, Tewari A, Beltran H, Onofrio RC, et al. 2013. Punctuated evolution of prostate cancer genomes. *Cell* **153**:666–677. DOI: https://doi.org/10.1016/j.cell.2013.03.021, PMID: 23622249

Butler K, Banday AR. 2023. APOBEC3-mediated mutagenesis in cancer: causes, clinical significance and therapeutic potential. *Journal of Hematology & Oncology* **16**:31. DOI: https://doi.org/10.1186/s13045-023-01425-5, PMID: 36978147

Campos PR, Adami C, Wilke CO. 2004. Modelling stochastic clonal interference. In: Ciobanu G, Rozenberg G (Eds). *Modelling in Molecular Biology*. Springer. p. 21–38. DOI: https://doi.org/10.1007/978-3-642-18734-6_2

Campos PRA, de Oliveira VM. 2004. Mutational effects on the clonal interference phenomenon. *Evolution; International Journal of Organic Evolution* **58**:932–937. DOI: https://doi.org/10.1111/j.0014-3820.2004.tb00427.x, PMID: 15212374

Chakravarty D, Gao J, Phillips SM, Kundra R, Zhang H, Wang J, Rudolph JE, Yaeger R, Soumerai T, Nissan MH, Chang MT, Chandarlapaty S, Traina TA, Paik PK, Ho AL, Hantash FM, Grupe A, Baxi SS, Callahan MK, Snyder A, et al. 2017. OncoKB: a precision oncology knowledge base. *JCO Precision Oncology* **2017**:1–16. DOI: https://doi.org/10.1200/PO.17.00011, PMID: 28890946

Davis A, Gao R, Navin N. 2017. Tumor evolution: Linear, branching, neutral or punctuated? Biochimica et Biophysica Acta (BBA). *Reviews on Cancer* **1867**:151–161. DOI: https://doi.org/10.1016/j.bbcan.2017.01.003, PMID: 28110020

Deming DA, Leystra AA, Nettekoven L, Sievers C, Miller D, Middlebrooks M, Clipson L, Albrecht D, Bacher J, Washington MK, Weichert J, Halberg RB. 2014. PIK3CA and APC mutations are synergistic in the development of intestinal cancers. *Oncogene* **33**:2245–2254. DOI: https://doi.org/10.1038/onc.2013.167, PMID: 23708654

Díaz-Gay M, Vangara R, Barnes M, Wang X, Islam SMA, Vermes I, Duke S, Narasimman NB, Yang T, Jiang Z, Moody S, Senkin S, Brennan P, Stratton MR, Alexandrov LB. 2023. Assigning mutational signatures to individual samples and individual somatic mutations with SigProfilerAssignment. *Bioinformatics* **39**:btad756. DOI: https://doi.org/10.1093/bioinformatics/btad756, PMID: 38096571

Drews RM, Hernando B, Tarabichi M, Haase K, Lesluyes T, Smith PS, Morrill Gavarró L, Couturier D-L, Liu L, Schneider M, Brenton JD, Van Loo P, Macintyre G, Markowetz F. 2022. A pan-cancer compendium of chromosomal instability. *Nature* **606**:976–983. DOI: https://doi.org/10.1038/s41586-022-04789-9, PMID: 35705807

Fischer BM, Olsen MWB, Ley CD. 2006. How few cancer cells can be detected by positron emission tomography? A frequent question addressed by an in vitro study. *European Journal of Nuclear Medicine and Molecular Imaging* **33**:697–702. DOI: https://doi.org/10.1007/S00259-005-0038-6/METRICS, PMID: 16612588

Fisher RA. 1923. XXI.—On the Dominance Ratio. *Proceedings of the Royal Society of Edinburgh* **42**:321–341. DOI: https://doi.org/10.1017/S0370164600023993

Fogle CA, Nagle JL, Desai MM. 2008. Clonal interference, multiple mutations and adaptation in large asexual populations. *Genetics* **180**:2163–2173. DOI: https://doi.org/10.1534/genetics.108.090019, PMID: 18832359

Galuppini F, Dal Pozzo CA, Deckert J, Loupakis F, Fassan M, Baffa R. 2019. Tumor mutation burden: from comprehensive mutational screening to the clinic. *Cancer Cell International* **19**:209. DOI: https://doi.org/10.1186/s12935-019-0929-4, PMID: 31406485

Gao R, Davis A, McDonald TO, Sei E, Shi X, Wang Y, Tsai P-C, Casasent A, Waters J, Zhang H, Meric-Bernstam F, Michor F, Navin NE. 2016. Punctuated copy number evolution and clonal stasis in triple-negative breast cancer. *Nature Genetics* **48**:1119–1130. DOI: https://doi.org/10.1038/ng.3641, PMID: 27526321

Gerrish PJ, Lenski RE. 1998. The fate of competing beneficial mutations in an asexual population. *Genetica* **102–103**:127–144. DOI: https://doi.org/10.1023/A:1017067816551, PMID: 9720276

Graser CJ. 2026. Punctuated-mutagenesis. swh:1:rev:610ba8e16897dbc5216e0718b368da666d09fa23. Software Heritage. https://archive.softwareheritage.org/swh:1:dir:89771b9bb260d6770ee2ce89430a0a96390e9541;

origin=https://github.com/Michorlab/Punctuated-mutagenesis;visit=swh:1:snp:8e8f4b4d78e84c70b3362de2fa58bd603981009e;anchor=swh:1:rev:610ba8e16897dbc5216e0718b368da666d09fa23

Gu Y, Yang D, Zou J, Ma W, Wu R, Zhao W, Zhang Y, Xiao H, Gong X, Zhang M, Zhu J, Guo Z. 2010. Systematic interpretation of comutated genes in large-scale cancer mutation profiles. *Molecular Cancer Therapeutics* **9**:2186–2195. DOI: https://doi.org/10.1158/1535-7163.MCT-10-0022, PMID: 20663929

Gupta A, Gazzo A, Selenica P, Safonov A, Pareja F, da Silva EM, Brown DN, Shao H, Zhu Y, Patel J, Blanco-Heredia J, Stefanovska B, Carpenter MA, Chen Y, Vegas I, Pei X, Frosina D, Jungbluth AA, Ladanyi M, Curigliano G, et al. 2025. APOBEC3 mutagenesis drives therapy resistance in breast cancer. *Nature Genetics* **57**:1452–1462. DOI: https://doi.org/10.1038/s41588-025-02187-1, PMID: 40379787

Hsu MS, Yu JC, Wang HW, Chen ST, Hsiung CN, Ding S, Wu PE, Shen CY, Cheng CW. 2010. Synergistic effects of polymorphisms in DNA repair genes and endogenous estrogen exposure on female breast cancer risk. *Annals of Surgical Oncology* **17**:760–771. DOI: https://doi.org/10.1245/s10434-009-0802-0, PMID: 20183911

Issa JP. 2022. The two-hit hypothesis meets epigenetics. *Cancer Research* **82**:1167–1169. DOI: https://doi.org/10.1158/0008-5472.CAN-22-0405, PMID: 35373289

Iwasa Y, Michor F, Nowak MA. 2004. Stochastic tunnels in evolutionary dynamics. *Genetics* **166**:1571–1579. DOI: https://doi.org/10.1534/genetics.166.3.1571, PMID: 15082570

Jamal-Hanjani M, Wilson GA, McGranahan N, Birkbak NJ, Watkins TBK, Veeriah S, Shafi S, Johnson DH, Mitter R, Rosenthal R, Salm M, Horswell S, Escudero M, Matthews N, Rowan A, Chambers T, Moore DA, Turajlic S, Xu H, Lee S-M, et al. 2017. Tracking the evolution of non-small-cell lung cancer. *The New England Journal of Medicine* **376**:2109–2121. DOI: https://doi.org/10.1056/NEJMoa1616288, PMID: 28445112

Knudson AG. 1971. Mutation and cancer: statistical study of retinoblastoma. *PNAS* **68**:820–823. DOI: https://doi.org/10.1073/pnas.68.4.820, PMID: 5279523

Komarova NL, Sengupta A, Nowak MA. 2003. Mutation-selection networks of cancer initiation: tumor suppressor genes and chromosomal instability. *Journal of Theoretical Biology* **223**:433–450. DOI: https://doi.org/10.1016/s0022-5193(03)00120-6, PMID: 12875822

Lee J-K, Lee J, Kim S, Kim S, Youk J, Park S, An Y, Keam B, Kim D-W, Heo DS, Kim YT, Kim J-S, Kim SH, Lee JS, Lee S-H, Park K, Ku J-L, Jeon YK, Chung DH, Park PJ, et al. 2017. Clonal history and genetic predictors of transformation into small-cell carcinomas from lung adenocarcinomas. *Journal of Clinical Oncology* **35**:3065–3074. DOI: https://doi.org/10.1200/JCO.2016.71.9096, PMID: 28498782

Mas-Ponte D, Supek F. 2020. DNA mismatch repair promotes APOBEC3-mediated diffuse hypermutation in human cancers. *Nature Genetics* **52**:958–968. DOI: https://doi.org/10.1038/s41588-020-0674-6, PMID: 32747826

McDonald TO, Cheng YC, Graser C, Nicol PB, Temko D, Michor F. 2023. Computational approaches to modelling and optimizing cancer treatment. *Nature Reviews Bioengineering* **1**:695–711. DOI: https://doi.org/10.1038/s44222-023-00089-7

Merlo LMF, Pepper JW, Reid BJ, Maley CC. 2006. Cancer as an evolutionary and ecological process. *Nature Reviews. Cancer* **6**:924–935. DOI: https://doi.org/10.1038/nrc2013, PMID: 17109012

Minussi DC, Nicholson MD, Ye H, Davis A, Wang K, Baker T, Tarabichi M, Sei E, Du H, Rabbani M, Peng C, Hu M, Bai S, Lin Y-W, Schalck A, Multani A, Ma J, McDonald TO, Casasent A, Barrera A, et al. 2021. Breast tumours maintain a reservoir of subclonal diversity during expansion. *Nature* **592**:302–308. DOI: https://doi.org/10.1038/s41586-021-03357-x, PMID: 33762732

Navin N, Kendall J, Troge J, Andrews P, Rodgers L, McIndoo J, Cook K, Stepansky A, Levy D, Esposito D, Muthuswamy L, Krasnitz A, McCombie WR, Hicks J, Wigler M. 2011. Tumour evolution inferred by single-cell sequencing. *Nature* **472**:90–94. DOI: https://doi.org/10.1038/nature09807, PMID: 21399628

Neinavaie F, Ibrahim-Hashim A, Kramer AM, Brown JS, Richards CL. 2021. The genomic processes of biological invasions: from invasive species to cancer metastases and back again. *Frontiers in Ecology and Evolution* **9**:681100. DOI: https://doi.org/10.3389/fevo.2021.681100

Nik-Zainal S, Alexandrov LB, Wedge DC, Van Loo P, Greenman CD, Raine K, Jones D, Hinton J, Marshall J, Stebbings LA, Menzies A, Martin S, Leung K, Chen L, Leroy C, Ramakrishna M, Rance R, Lau KW, Mudie LJ, Varela I, et al. 2012. Mutational processes molding the genomes of 21 breast cancers. *Cell* **149**:979–993. DOI: https://doi.org/10.1016/j.cell.2012.04.024, PMID: 22608084

Nowak MA, Michor F, Komarova NL, Iwasa Y. 2004. Evolutionary dynamics of tumor suppressor gene inactivation. *PNAS* **101**:10635–10638. DOI: https://doi.org/10.1073/pnas.0400747101, PMID: 15252197

Paige AJW. 2003. Redefining tumour suppressor genes: exceptions to the two-hit hypothesis. *Cellular and Molecular Life Sciences* **60**:2147–2163. DOI: https://doi.org/10.1007/s00018-003-3027-6, PMID: 14618262

Park SC, Krug J. 2007. Clonal interference in large populations. *PNAS* **104**:18135–18140. DOI: https://doi.org/10.1073/pnas.0705778104, PMID: 17984061

Park S, Supek F, Lehner B. 2021. Higher order genetic interactions switch cancer genes from two-hit to one-hit drivers. *Nature Communications* **12**:1–10. DOI: https://doi.org/10.1038/s41467-021-27242-3, PMID: 34862370

Petljak M, Alexandrov LB, Brammeld JS, Price S, Wedge DC, Grossmann S, Dawson KJ, Ju YS, Iorio F, Tubio JMC, Koh CC, Georgakopoulos-Soares I, Rodríguez-Martín B, Otlu B, O'Meara S, Butler AP, Menzies A, Bhosle SG, Raine K, Jones DR, et al. 2019. Characterizing mutational signatures in human cancer cell lines reveals episodic APOBEC mutagenesis. *Cell* **176**:1282–1294.. DOI: https://doi.org/10.1016/j.cell.2019.02.012, PMID: 30849372

Petljak M, Dananberg A, Chu K, Bergstrom EN, Striepen J, von Morgen P, Chen Y, Shah H, Sale JE, Alexandrov LB, Stratton MR, Maciejowski J. 2022. Mechanisms of APOBEC3 mutagenesis in human cancer cells. *Nature* **607**:799–807. DOI: https://doi.org/10.1038/s41586-022-04972-y, PMID: 35859169

Proulx SR. 2011. The rate of multi-step evolution in Moran and Wright-Fisher populations. *Theoretical Population Biology* **80**:197–207. DOI: https://doi.org/10.1016/j.tpb.2011.07.003, PMID: 21801738

Salehi S, Kabeer F, Ceglia N, Andronescu M, Williams MJ, Campbell KR, Masud T, Wang B, Biele J, Brimhall J, Gee D, Lee H, Ting J, Zhang AW, Tran H, O'Flanagan C, Dorri F, Rusk N, de Algara TR, Lee SR, et al. 2021. Clonal fitness inferred from time-series modelling of single-cell cancer genomes. *Nature* **595**:585–590. DOI: https://doi.org/10.1038/s41586-021-03648-3, PMID: 34163070

Sha D, Jin Z, Budczies J, Kluck K, Stenzinger A, Sinicrope FA. 2020. Tumor mutational burden as a predictive biomarker in solid tumors. *Cancer Discovery* **10**:1808–1825. DOI: https://doi.org/10.1158/2159-8290.CD-20-0522, PMID: 33139244

Sottoriva A, Kang H, Ma Z, Graham TA, Salomon MP, Zhao J, Marjoram P, Siegmund K, Press MF, Shibata D, Curtis C. 2015. A Big Bang model of human colorectal tumor growth. *Nature Genetics* **47**:209–216. DOI: https://doi.org/10.1038/ng.3214, PMID: 25665006

Stephens PJ, Greenman CD, Fu B, Yang F, Bignell GR, Mudie LJ, Pleasance ED, Lau KW, Beare D, Stebbings LA, McLaren S, Lin M-L, McBride DJ, Varela I, Nik-Zainal S, Leroy C, Jia M, Menzies A, Butler AP, Teague JW, et al. 2011. Massive genomic rearrangement acquired in a single catastrophic event during cancer development. *Cell* **144**:27–40. DOI: https://doi.org/10.1016/j.cell.2010.11.055, PMID: 21215367

Sun R, Hu Z, Curtis C. 2018. Big bang tumor growth and clonal evolution. *Cold Spring Harbor Perspectives in Medicine* **8**:a028381. DOI: https://doi.org/10.1101/cshperspect.a028381, PMID: 28710260

Van Egeren D, Madsen T, Michor F. 2018. Fitness variation in isogenic populations leads to a novel evolutionary mechanism for crossing fitness valleys. *Communications Biology* **1**:151. DOI: https://doi.org/10.1038/s42003-018-0160-1, PMID: 30272027

Wang LH, Wu CF, Rajasekaran N, Shin YK. 2018. Loss of tumor suppressor gene function in human cancer: an overview. *Cellular Physiology and Biochemistry* **51**:2647–2693. DOI: https://doi.org/10.1159/000495956, PMID: 30562755

Wang Y, Robinson PS, Coorens THH, Moore L, Lee-Six H, Noorani A, Sanders MA, Jung H, Katainen R, Heuschkel R, Brunton-Sim R, Weston R, Read D, Nobbs B, Fitzgerald RC, Saeb-Parsy K, Martincorena I, Campbell PJ, Rushbrook S, Zilbauer M, et al. 2023. APOBEC mutagenesis is a common process in normal human small intestine. *Nature Genetics* **55**:246–254. DOI: https://doi.org/10.1038/s41588-022-01296-5, PMID: 36702998

Watson HW, Galton F. 1875. On the Probability of the Extinction of Families. *The Journal of the Anthropological Institute of Great Britain and Ireland* **4**:138. DOI: https://doi.org/10.2307/2841222

Weinreich DM, Chao L. 2005. Rapid evolutionary escape by large populations from local fitness peaks is likely in nature. *Evolution; International Journal of Organic Evolution* **59**:1175–1182. DOI: https://doi.org/10.1111/J.0014-3820.2005.TB01769.X, PMID: 16050095

Weissman DB, Desai MM, Fisher DS, Feldman MW. 2009. The rate at which asexual populations cross fitness valleys. *Theoretical Population Biology* **75**:286–300. DOI: https://doi.org/10.1016/j.tpb.2009.02.006, PMID: 19285994

Werner B, Case J, Williams MJ, Chkhaidze K, Temko D, Fernández-Mateos J, Cresswell GD, Nichol D, Cross W, Spiteri I, Huang W, Tomlinson IPM, Barnes CP, Graham TA, Sottoriva A. 2020. Measuring single cell divisions in human tissues from multi-region sequencing data. *Nature Communications* **11**:1035. DOI: https://doi.org/10.1038/s41467-020-14844-6, PMID: 32098957

Wilke CO. 2004. The speed of adaptation in large asexual populations. *Genetics* **167**:2045–2053. DOI: https://doi.org/10.1534/genetics.104.027136, PMID: 15342539

Williams MJ, Werner B, Heide T, Curtis C, Barnes CP, Sottoriva A, Graham TA. 2018. Quantification of subclonal selection in cancer from bulk sequencing data. *Nature Genetics* **50**:895–903. DOI: https://doi.org/10.1038/s41588-018-0128-6, PMID: 29808029

Wright S. 1931. Evolution in mendelian populations. *Genetics* **16**:97–159. DOI: https://doi.org/10.1093/genetics/16.2.97, PMID: 17246615

Yates LR, Campbell PJ. 2012. Evolution of the cancer genome. *Nature Reviews. Genetics* **13**:795–806. DOI: https://doi.org/10.1038/nrg3317, PMID: 23044827

Zhang C-Z, Spektor A, Cornils H, Francis JM, Jackson EK, Liu S, Meyerson M, Pellman D. 2015. Chromothripsis from DNA damage in micronuclei. *Nature* **522**:179–184. DOI: https://doi.org/10.1038/nature14493, PMID: 26017310

Zhao M, Kim P, Mitra R, Zhao J, Zhao Z. 2016. TSGene 2.0: an updated literature-based knowledgebase for tumor suppressor genes. *Nucleic Acids Research* **44**:D1023–D1031. DOI: https://doi.org/10.1093/nar/gkv1268, PMID: 26590405

Appendix 1

Derivations of results on multi-step adaptation rates in the limit of infrequent mutations

In the main text, we discussed differences in the rate of multi-step adaptations f_k as a function of a temporal clustering parameter k. Here, we formalize this discussion and derive the result presented in the main text. We consider processes in which multi-step adaptations occur via stochastic tunneling rather than sequential fixation, that is processes in which mutants with a selective disadvantage are negligibly unlikely to reach fixation. We assume that mutations happen sufficiently infrequently, so that we can neglect valley crossing scenarios in which the same mutation in a multi-step adaptation sequence occurs multiple times. Additionally, we require that valley crossing happens solely via stochastic tunneling, that is that the probability that a maladaptive mutant fixates is vanishingly low.

We denote the average rate at which cells acquire mutations by μ, and compare dynamics under different mutation processes which we index by a temporal clustering parameter $k \geq 1$. Starting every d units of time, a process with parameter k undergoes a mutation burst lasting $\frac{d}{k}$ units of time, during which mutations occur at a rate $k\mu$. Outside of bursts, the mutation rate is 0. We assume that the time $\frac{(k-1)d}{k}$ between subsequent burst phases is sufficiently long such that we can restrict attention to multi-step adaptations that occur during a single burst, and that the duration of each burst $\frac{d}{k}$ is sufficiently long relative to the time that it takes to cross a fitness valley, so that valley crossings, which fail because the burst phase ended but which would have been successful, have a negligible effect on f_k.

As stated in the main text, with this setup we can show that for any population dynamics process for which the above limits can be motivated, the fold-increase in the rate f_k at which $(n+1)$-step adaptations occur in a process with clustering parameter k relative to the uniform process approaches k^n.

$$\lim_{d\to\infty}\lim_{\mu\to\infty}\frac{f_k}{f_1} = \lim_{\mu\to\infty}\lim_{d\to\infty}\frac{f_k}{f_1} = k^n.$$

Deriving this result for any selection process indeed becomes straightforward, if it can be motivated that the mutation rate does not affect the fate of any mutant lineage once the first mutant in this lineage has emerged.

Somewhat more formally, we index the steps in an n-step adaptation sequence by $i \in \{1, \ldots, n\}$, and denote the expected size of the lineage descending from mutant i by $Y_i(t)$. This quantity $Y_i(t)$ of course depends on the specificities of the selection process, but for our derivations it suffices to only require that $Y_i(t)$ does not depend on the mutation rate.

Since all intermediate mutants have a selective disadvantage, and since we assume that we are in a parameter regime in which the likelihood that a disadvantageous mutant fixates is negligibly low, it follows that the integral $\int_0^\infty Y_i\left(t\right) dt$ converges.

$$\int_0^\infty Y_i\left(t\right) dt < \infty.$$

In the limit of a low (and for now time-invariant) mutation rate μ, the probability that the i'th mutant produces a further mutant $P(i \to i+1)$ is simply the product of the mutation rate and this integral.

$$P\left(i \to i+1\right) = \mu \int_0^\infty Y_i\left(t\right) dt.$$

The probability that the first mutant spawns a sequence of $n+1$ mutants can, thus, be written as follows:

$$\begin{aligned} P\left(1 \to n+1\right) &= \prod_{i=1}^{n}\left(\mu \int_0^\infty Y_i\left(t\right) dt\right) \\ &= \mu^n \prod_{i=1}^{n}\left(\int_0^\infty Y_i\left(t\right) dt\right). \end{aligned}$$

Lastly, we use $\tilde{\mu}$ to denote the rate at which new mutants with only one mutation emerge. For a given selection process, this rate may vary with the population size. For the Wright–Fisher models of constant population size N, this expression simplifies to $\tilde{\mu} = N\mu$. We arrive at the following formulation for the rate at which new advantageous mutants emerge f_1.

$$\begin{aligned} f_1 &= \tilde{\mu} \cdot P\left(1 \to n+1\right) \\ &= \tilde{\mu} \cdot \mu^n \cdot \prod_{i=1}^{n} \left(\int_0^\infty Y_i\left(t\right) dt\right). \end{aligned}$$

Finally, we can introduce our clustering parameter k to scale both μ and $\tilde{\mu}$, and multiply by a factor of $\frac{1}{k}$ to arrive at the average rate of valley crossing in the temporally clustered process with parameter k.

$$\begin{aligned} f_k &= \frac{1}{k} \cdot \left(k \cdot \tilde{\mu}\right) \cdot \left(k \cdot \mu\right)^n \cdot \prod_{i=1}^{n} \left(\int_0^\infty Y_i\left(t\right) dt\right) \\ &= k^n \cdot \tilde{\mu} \cdot \mu^n \cdot \prod_{i=1}^{n} \left(\int_0^\infty Y_i\left(t\right) dt\right) \\ &= k^n \cdot f_1 \end{aligned} .$$

One consequence of our assumption that none of the maladaptive intermediate mutants reaches fixation is that the composition of the population once the final mutant emerges in the limit of $\mu \to 0$ does not depend on the mutation rate. For population dynamics models with a constant population size in which the probability that an emerging mutant (absent further mutation events) reaches fixation only depends on this composition, such as the Wright–Fisher process, we can interpret $\frac{f_k}{f_1}$ therefore also as the ratio of the rates at which mutants with $n + 1$ mutations fixate.

Analogously, in certain models of branching evolution in which the prospects of one branch do not depend on other co-evolving branches, such as the Galton–Watson process (***Watson and Galton, 1875***), whether the lineage of an emerging mutant with $(n + 1)$ mutations survives is independent of the mutation rate in the limit of $\mu \to 0$. The ratio $\frac{f_k}{f_1}$ therefore also reflects the ratio of the rates at which surviving lineages with $(n + 1)$ adaptations arise in such models. We provide thorough treatment of dynamics in such processes below (Appendix 2).

The above result suggests that the effect of temporal clustering on relative rates of valley crossing is substantial, and gets exponentially stronger the wider the fitness valley is (n). We validated this result with simulations of a Wright–Fisher process for a range of small values for μ (***Figure 2—figure supplement 2B, E***). As we move away from the limit of rare mutations, the relative effect of temporal clustering $\frac{f_k}{f_1}$ gets smaller, as will be discussed below. However, as the absolute rates f_k and f_1 are proportional to μ^{n+1}, the absolute effect of temporal clustering on the rate of fixations of mutants with multi-step adaptations initially increases in μ (***Figure 2—figure supplement 2D***).

Appendix 2

Dynamics in a branching process

In the Galton–Watson process, a branching process with constant birth and death rates, it is possible to derive explicit formulas for the probability that a lineage will acquire a mutation before going extinct (often termed 'evolutionary rescue' ***Azevedo and Olofsson, 2021***). Formulating this probability for the lineage of the first mutant in a two-step adaptation sequence thus allows us to explore numerically how the magnitude of the effects that we describe behaves across a wide range of parameter combinations, which is especially attractive for parameter regions for which simulations become prohibitively expensive.

We thus investigated this probability of evolutionary rescue, which translates to f_k/μ in our model, as a function of k, μ and the birth/death ratio (***Figure 2—figure supplement 3A–C***). Note that the birth/death ratio that we considered here is that of the first mutant, not that of the wild-type cell. Unsurprisingly, we found that as μ decreases, rescue probabilities for any k converge to zero for subcritical processes (birth/death ratios < 1). For supercritical processes (birth/death ratios > 1), rescue probabilities converge to one minus the extinction probability of the lineage – rescue only happens if the process expands indefinitely. Throughout, rescue probabilities of course increase in the expected time that the lineage persists, that is they increase with the birth/death ratio. However, this increase is most steep for large μ and flattens for the subcritical range as μ decreases.

For the ratio of these escape probabilities for different k relative to that of the uniform process, that is for $\frac{\frac{f_k}{\mu}}{\frac{f_1}{\mu}} = \frac{f_k}{f_1}$, we found that in the subcritical regime, as μ decreases, $\frac{f_k}{f_1}$ approaches k, consistent with our derivations in the previous section (Appendix 1). In the supercritical regime, the condition $\int_0^\infty Y(t)\,dt < \infty$ used in our derivations above is violated – lineages have a non-zero chance of branching out indefinitely. Here we saw that as μ decreases toward zero, $\frac{f_k}{f_1}$ approaches 1 (***Figure 2—figure supplement 3D–F***).

Moreover, we found that $\frac{f_k}{f_1}$ consistently decreases with increasing birth–death ratio. As a result of the interplay between this decrease in $\frac{f_k}{f_1}$ and increase in the absolute rescue probability, we found that the absolute effect, $f_k - f_1$, is highest for birth/death ratios close to one (***Figure 2—figure supplement 3G***). These results remain as μ decreases (***Figure 2—figure supplement 3H, I***) but the magnitude of the absolute effect of course shrinks, as rescue of lineages that would otherwise go extinct (all subcritical lineages, and those lineages in the supercritical regime that go extinct by chance) generally becomes unlikely at small values of μ.

Appendix 3

Dynamics as a function of time between bursts

In our derivations, we assumed that bursts are sufficiently spaced out in time, and last sufficiently long so that we can restrict attention to mutant lineages that emerge and go extinct within the duration of a single burst ($d \to \infty$). To investigate how effect sizes behave away from this limit, as a function of the time between bursts, we numerically approximated rescue probabilities for finite d.

Specifically, we approximated the Galton–Watson process by a Markov process with finitely many states $s \in \{0, 1, \ldots, N, rescue\}$, starting with $s(0) = 1$, where $s \in \{0, 1, \ldots, N\}$ corresponds to the number of viable cells. We considered 0, N and $rescue$ as absorbing states, and otherwise modeled transitions to states $s-1$ or $s+1$ as occurring at rates $s \cdot r_d$ and $s \cdot r_b$, respectively, with constant birth and death rates r_d and r_b, and transitions to the $rescue$ state as occurring at rate $s \cdot k \cdot \mu$ within bursts and at rate zero otherwise. We iteratively solved the corresponding forward equations of this continuous time Markov process numerically for subsequent cycles of in-burst and out-of-burst periods, until the summed mass on states $\{1, \ldots, N-1\}$ fell below a tolerance threshold, and we considered the joint mass on states N and $rescue$ as approximation for the rescue probability. For our results in ***Figure 2—figure supplement 4***, we used $N = 300$ and a tolerance threshold of 10^{-10}.

Generally, we showed that differences in multi-step adaptation rates emerge because lineages with a first mutation experience higher mutation rates under a temporally clustered mutation process compared to a uniform mutation process (Appendix 1). As the burst duration and the time-interval between subsequent bursts decrease, this expected difference in mutation rates for mutant lineages diminishes because of the increased likelihood that a mutant lineage that emerged in a burst also experiences a non-burst phase. Consequently, we found that lowering d diminishes the effect of temporal clustering on the probability of producing a second mutant (***Figure 2—figure supplement 4A–C***). This pattern manifests both in the relative effects, f_k/f_1 (***Figure 2—figure supplement 4D–F***) and in the absolute effects $f_k - f_1$ (***Figure 2—figure supplement 4H–J***). However, for the absolute differences $f_k - f_1$, we found that the strength of this dependency varies non-monotonously with the birth–death ratio. While absolute differences under large d peak close to a birth–death ratio of one (***Figure 2—figure supplements 3–4J***), we also observed the strongest decrease in absolute differences at these values for the birth–death ratio, so that under low d this peak turns into a local minimum of the absolute-difference curve (***Figure 2—figure supplement 4H, I***). This effect arises because the expected lifetime of a mutant lineage before going extinct increases steeply when moving from a subcritical process (birth–death ratios <1) to a critical process (birth–death ratio = 1), making such lineages with higher birth–death ratios more sensitive to changes in mutation rates that occur long after the lineage first emerged.

Appendix 4

Dynamics in exploration–exploitation setting

Our simulations show that for a given μ the effect of increasing the clustering parameter k on the average fitness becomes negative at some k (***Figure 3C***). In the simulations in this section, this effect is driven by two main factors. First, after a re-drawing, the population often is not in a local maximum of the fitness landscape, and hence may be able to increase its fitness by single mutations (one-step adaptations), without tunneling. Waiting for a burst to occur and thereby delaying such local explorations decreases average fitness. Second, once the population has found a peak in the fitness landscape, having very high mutation rates in bursts temporarily scatters cells to lower points in the landscape, and may even cause drift to points of lower fitness. Both of these effects become apparent when considering a representative snapshot of the simulations (***Figure 3D***). The misalignment of re-drawing and burst causes the population to remain in a valley for several divisions until the first burst occurs. Moreover, relative to similar snapshots of simulations with lower k (***Figure 3E, F***), once the population has reached a peak, the scattering during bursts at higher k leads to sharper decreases in fitness, and the population may even remain in a lower fitness point after a burst (***Figure 3D***).

This increased scattering can also be seen when going from $k = 1$ to $k = 5$ (***Figure 3E, F***). However, at $k = 1$, the population spends much time in local maxima, as large jumps in fitness only occur right after re-drawings, whereas at $k = 5$, jumps occur also long after the landscape was re-drawn and the population found a local maximum. These later jumps to higher points in the landscape correspond to tunneling events.

Appendix 5

TSG deactivation and ROS-associated mutagenesis

Following a reviewer suggestion, we also explored the contribution of mutation signatures attributed to reactive oxygen species (ROS; ***Figure 4—figure supplement 4***). However, we found that these signatures account for only small fractions of the total observed mutation signatures, suggesting that these mutagenic processes would not account for much of the fluctuation behavior of overall mutation rates, and we saw no consistent correlations with TSG inactivation.