

# **Passenger mutations can accelerate tumor suppressor gene inactivation in cancer evolution**

Dominik Wodarz<sup>1,2</sup>, Alan C. Newell<sup>3</sup>, and Natalia L. Komarova<sup>2,1</sup>

1: Department of Ecology and Evolutionary Biology, 321 Steinhaus Hall, University of California, Irvine, CA 92697

2: Department of Mathematics, Rowland Hall, University of California, Irvine, CA 92697

3: Department of Mathematics, The University of Arizona, 617 N. Santa Rita Ave, Tucson, AZ 85721

## Abstract

Carcinogenesis is an evolutionary process whereby cells accumulate multiple mutations. Besides the “driver mutations” that cause the disease, cells also accumulate a number of other mutations with seemingly no direct role in this evolutionary process. They are called passenger mutations. While it has been argued that passenger mutations render tumors more fragile due to reduced fitness, the role of passenger mutations remains understudied. Using evolutionary computational models, we demonstrate that in the context of tumor suppressor gene inactivation (and hence fitness valley crossing), the presence of passenger mutations can accelerate the rate of evolution by reducing overall population fitness and increasing the relative fitness of intermediate mutants in the fitness valley crossing pathway. Hence, the baseline rate of tumor suppressor gene inactivation might be faster than previously thought. Conceptually, parallels are found in the field of turbulence and pattern formation, where instabilities can be driven by perturbations that are damped (disadvantageous), but provide a richer set of pathways such that a system can achieve some desired goal more readily. This highlights, through a number of novel parallels, the relevance of physical sciences in oncology.

## Introduction

The development and progression of cancer is an evolutionary process whereby cells accumulate multiple mutations, which enables them to break out of homeostasis and to proliferate out of control. The mutations that enable this process typically confer a selective advantage to cells and have been called driver mutations [1-5]. Tumors, however, are highly heterogeneous and cells also contain a variety of other mutations, called passenger mutations [1-5]. They arise from random mutations in sequences that do not contribute directly to disease, facilitated by exposure to mutagenic processes and lack of repair [6]. While passenger mutations have been thought to have minimal biological consequences on the disease process, the properties and role of passenger mutations remain poorly understood. Recent data indicate that passenger mutations carry a certain fitness cost [7,8], and that they might therefore render tumors more fragile, which could be exploited therapeutically [4,7,9]. One particular evolutionary process that is central to carcinogenesis and cancer progression is the inactivation of tumor suppressor genes (TSG)<sup>9,10</sup>. This typically requires two mutational hits because both copies of the gene need to be inactivated to achieve full loss of function [10]. Different tumor suppressor genes display different characteristics, and in principle, the inactivation of only one copy of the gene either results in no change in the fitness of the cell, or it could entail a certain selective disadvantage. To inform model assumptions, we will specifically consider the tumor suppressor gene APC, which becomes inactivated early in the development of colorectal cancer [11]. In this case, data indicate that heterozygous APC<sup>+/-</sup> cells can experience reduced fitness, which means that a fitness valley has to be crossed for the inactivation of APC to occur. Experiments with

colorectal cancer cell lines revealed that a truncating mutation in APC has a dominant effect resulting in a spindle checkpoint defect, aneuploidy, and a reduced proliferation rate of cells [12,13]. Similar effects have been found in vivo in APC<sup>Min/+</sup> mice [14], which have an APC<sup>+/-</sup> germ line mutation. In general, if a copy of a tumor suppressor gene is lost as a consequence of aneuploidy, the cell is likely to suffer a fitness reduction (see e.g. references [15,16]). Motivated by these studies, our paper investigates the effect of passenger mutations on the evolutionary dynamics of tumor suppressor gene inactivation, assuming that a fitness valley needs to be crossed.

## Results

We consider a computational model for the inactivation of tumor suppressor genes (TSG) [17,18], where cells only acquire an advantage once they have accumulated two separate mutations, but are neutral or disadvantageous in the presence of only one of the mutations. For convenience, cells with both copies of the TSG present are referred to as TSG<sup>+/+</sup>, and cells with one or both copies of the TSG inactivated are referred to as TSG<sup>+/-</sup> and TSG<sup>-/-</sup> respectively. Much evolutionary work has been performed that studied how fast such fitness valleys/plateaus can be crossed, depending on the scenario under consideration [19-30]. To study the role of passenger mutations, we employ a stochastic agent-based model that is also referred to as a contact process. This model assumes the existence of  $N$  spots, which can either be empty or contain a cell. Each time step, the system is sampled  $M$  times, where  $M$  is the number of cells

present. If the chosen spot contains a cell, it can divide and die with defined probabilities. When a cell is chosen to divide, a target spot is chosen randomly from the whole system, and division only proceeds if this target spot is empty. Upon division, mutations can occur that give rise to different cell genotypes (Figure 1A). TSG<sup>+/+</sup> cells without passenger mutations are denoted by  $x$  and attempt division with a probability  $L_x$  per cell per update. TSG<sup>+/-</sup> cells without passenger mutations are denoted by  $y$  and have a fitness cost of  $s_1$  ( $s_1 \leq 1$ ), such that their division probability is  $s_1 L_x$ . TSG<sup>+/+</sup> cells that also contain passenger mutations are denoted by  $z$  and have a fitness cost  $s_2$  ( $s_2 < 1$ ). TSG<sup>+/-</sup> cells that also contain passenger mutations are denoted by  $w$  and have a fitness cost  $s_3$  ( $s_3 < 1$ ). Both TSG<sup>+/-</sup> populations,  $y$  and  $w$ , can give rise to the advantageous TSG<sup>-/-</sup> double mutant. All the mutation processes are defined in Figure 1A. For simplicity, each cell type is assumed to die with the same rate  $D$ . The model was simulated repeatedly, and the fraction of realizations when an advantageous TSG<sup>-/-</sup> mutant had been generated by a defined time threshold was determined. We compared simulations without passenger mutations ( $n=0$ ) with those that did allow the generation of passenger mutations ( $n>0$ ). Two different regimes have been observed [24]: In one regime, the double-hit mutant arises without the intermediate TSG<sup>+/-</sup> mutant reaching fixation, a process called stochastic tunneling. The second regime can be called sequential fixation, where the intermediate TSG<sup>+/-</sup> mutant fixates before the double mutant is created.

For the tunneling regime, we find that the presence of passenger mutations can accelerate the generation of the advantageous double mutant (Figure 2Ai). The

magnitude of this effect increases with higher fitness of the passenger mutants,  $s_2$  (Figure 2Ai). This requires that (i) the number of passenger mutations that can be accumulated,  $n$ , is sufficiently large relative to the inverse of the mutation rate and (ii) the intermediate TSG<sup>+/-</sup> mutation reduces the fitness to a lesser degree in cells with passenger mutations than in cells without passenger mutations, i.e. there are epistatic interactions between drivers and passengers [31,32]. The exact condition is  $s_3 > s_1 s_2$ , see Supplementary materials for computational details. This is a necessary condition for the passenger mutations to accelerate evolution (both in the tunneling regime and in the sequential fixation regime, see below). In the Discussion section, we describe a specific example where the fitness of TSG<sup>+/-</sup> mutants is context dependent, indicating that an assumed occurrence of epistasis in such cells is biologically relevant.

The reason for the accelerated evolution in the presence of passenger mutations is that these mutations increase the relative fitness of the intermediate TSG<sup>+/-</sup> cells through a set of complex interactions. If the wild-type population consists mostly of cells without passenger mutations, the evolutionary dynamics are largely driven by the  $x, y$  system, which is relatively slow due to the more pronounced disadvantage of  $y$ . The larger the proportion of cells with passenger mutations, however, the more the evolutionary dynamics are driven by the  $z, w$  system, where the intermediate mutant suffers an overall lower fitness cost. This allows the total population of intermediate TSG<sup>+/-</sup> cells to persist at a higher selection-mutation balance, making it more likely to generate the double mutant. The average rate of double mutant generation can be calculated from ordinary differential equations (ODEs, see Supplementary Materials),

which is a reasonable model that quantifies population dynamics in the tunneling regime. The rate of double mutant generation is increased by the presence of passenger mutants, with more pronounced effects for larger values of  $s_2$  (Figure 3), thus explaining our observations. If the fitness of passenger mutants crosses a threshold (which depends on the total rate of passenger mutant generation), the cells with passenger mutations,  $z$ , outcompete those without passenger mutations,  $x$  (because  $z$  is generated by  $x$ ). In this regime, the advantageous double mutants are created fastest because the intermediate TSG $^{+/-}$  mutants ( $w$ ) have the highest relative fitness out of all scenarios. This might be a biologically relevant parameter region given the ubiquitous occurrence of passenger mutations in cancer cells, and even in aged non-cancerous tissue [33,34]. We note that in this model, the generation of cells with passenger mutations does not increase the total population size and hence does not provide additional targets for mutation. The accelerating effect of passengers stems from the overall reduction in population fitness and the consequent elevation of the relative fitness of intermediate TSG $^{+/-}$  mutants.

The extent to which passenger mutants accelerate fitness valley crossing further depends on the relative fitness of intermediate TSG $^{+/-}$  mutants without passenger mutations ( $s_1$ , the fitness cost of population  $y$ ). The closer the fitness of the intermediate mutant  $y$ -population is to the fitness of the wild-type  $x$ -population ( $s_1 \rightarrow 1$ ), the less pronounced the accelerating effect (compare Figures 2Ai-iii). This is also seen in the ODE predictions, which show that for larger values of  $s_1$ , the average rate of double mutant generation is accelerated by passenger mutations to a lesser extent (Figure 3).

The reason is that for higher  $s_1$ , the intermediate TSG<sup>+/-</sup> mutants (y) have less of a disadvantage compared to population x, which leaves less room for improvement by the z, w interactions. Thus, if the intermediate TSG<sup>+/-</sup> mutant is almost neutral with respect to the wild-type, passenger mutants are not likely to accelerate evolution in the tunneling regime.

Next, consider the parameter regime where the intermediate mutant fixates prior to the generation of the double mutant (sequential fixation). This tends to occur for parameters where the generation of the double mutant takes a longer period of time due to lower mutation rates or smaller population sizes. In this scenario, the accelerating effect of passenger mutations can be significantly more pronounced than in the tunneling regime (Figure 2B). For a physiologically realistic rate of gene inactivation ( $10^{-7}$  per gene per division), even if the TSG<sup>+/-</sup> mutants without passenger mutations only have a 0.1% fitness cost ( $s_1=0.999$ ), and if the passenger mutations lead to a 1% fitness cost ( $s_2=0.99$ ), the presence of passenger mutations can accelerate the emergence of the double mutant almost 3-fold (Figure 2Bii). If the fitness cost of the intermediate TSG<sup>+/-</sup> mutant is 1%, the acceleration can be up to 35-fold (Figure 2Biii). The reason is that the fixation probability of the TSG<sup>+/-</sup> mutants is markedly higher when the dynamics are governed more by the z,w system compared to the x,y system.

These results remain robust if instead of assuming that all passenger mutants have the same fitness cost, those fitness cost values are taken from a power function



distribution between zero and one, with averages given by  $s_2$  and  $s_3$ . This potentially allows for some significantly deleterious passenger mutants even though many of them can be close to neutral (Supplementary Materials). Results are further shown to remain robust in a spatially explicit model, where dividing cells place their offspring in a randomly chosen spot nearby (Supplementary Materials). Finally, the same patterns are observed in a constant population Moran process, which represents tissues where normal cells are maintained at carrying capacity and their homeostatic turnover is driven by cell death (Supplementary Materials).

While our models have shown that the presence of passenger mutations can accelerate the rate of TSG inactivation, the question arises how significant this acceleration can be. To gauge that, we compare the degree of acceleration that can be observed in our passenger mutation model to the acceleration observed in the context of a different and unrelated process that occurs in colorectal carcinogenesis, and that is known to lead to clinically significant accelerations in evolutionary processes: tumor initiation in Lynch Syndrome patients. It is known that Lynch Syndrome patients develop colorectal tumors with a significantly faster rate than the general population. This is because Lynch Syndrome patients are characterized by a germ line mutation in one copy of a mismatch repair (MMR) gene. Hence a single point mutation can frequently generate MMR-deficient cells that promote mutant accumulation and hence the inactivation of APC. Therefore, we describe tumor formation in the context of Lynch Syndrome (without passenger mutations) in the same kind of computational framework studied so far (Figure 1B), and determine the extent to which evolution is accelerated in

that model. If the degree of acceleration observed in the Lynch Syndrome model is of a similar magnitude as the acceleration observed in the passenger model, then there is indication that passenger-induced acceleration can be clinically highly relevant. If the acceleration in the passenger model is much less than that in the Lynch Syndrome model, then the passenger-induced acceleration is less relevant. In particular, we investigated by how much the mutation rate in the mismatch repair (MMR)-deficient cells has to be increased to obtain a degree of evolution acceleration that is comparable to that observed with passenger mutations. To do so, we assumed that genes are inactivated with a rate of  $10^{-7}$  per division, and that an intermediate TSG+/- cell carries a 1% fitness cost, consistent with data that documented reduced growth of APC+/- cells [12]. As before, passenger mutants were also assumed to carry a 1% fitness cost. If we assume that MMR-deficient cells do not carry a fitness cost, we obtain that MMR-deficient cells need to have a 100-500 fold increase in their mutation rate to accelerate evolution to a similar degree as seen in corresponding passenger mutant simulations (Figure 4a). If MMR-deficient cells have a 1% fitness cost, then the fold increase in the mutation rate has to be 1000-5000 fold to match the acceleration afforded by the presence of passenger mutations (Figure 4b). Because this increase in mutation rate is thought to be typical for mismatch-repair deficient colorectal cells [35], this suggests that passenger mutations can have an accelerating effect that is similar in magnitude to acceleration in Lynch Syndrome, pointing to potentially strong biological relevance. If APC+/- cells are characterized by a significantly lower fitness cost, or if the assumed epistatic interactions between drivers and passengers are significantly weaker, this effect would be reduced.

## Discussion and Conclusion

Previous work reported that the presence of passenger mutations can make tumors more fragile in certain circumstances due to a reduction in overall fitness [4,7,9]. Here we have shown that in the context of fitness valley crossing, costly passenger mutations can actually accelerate evolution because they reduce overall population fitness and thereby provide an environment in which intermediate TSG<sup>+/-</sup> mutants enjoy a higher relative fitness. Although passenger mutations are selected against, their accumulation (even at low numbers) provides access to additional pathways to cancer where the fitness valley is shallower and easier to cross. It has been previously suggested that the process of carcinogenesis could be promoted through a reduction of overall population fitness due to aging and other insults, thus providing a more favorable fitness landscape for the evolution of malignant cells [36]. Our passenger mutations model fits well into this concept.

The models further indicate that this can result in an acceleration of evolution that can be comparable to that observed in patients with a predisposition to genetic instability. This suggests that the “baseline” rate of TSG inactivation in the absence of genetic predisposition and genetic instability can be significantly faster than previously thought, which might be conceptually important for understanding the ability of cells to accumulate a number of carcinogenic mutations in a relatively short period of time [37]. This applies not only to tumor progression, but also to cancer initiation in healthy tissue,

which has been shown to contain a significant number of passenger (and driver) mutations [34], especially at advanced age [33].

Our analysis identified possible epistatic interactions between driver and passenger mutations [31,32] to be important for the reported dynamics, and the literature supports this notion. Indeed, it has been pointed out that the classification of mutations into passengers and drivers might be an over-simplification, because the fitness of a given cancer phenotype can be context-dependent [38]. More specifically, we turn again to the tumor suppressor gene APC in colorectal carcinogenesis. There are mouse strains that are heterozygous for the APC<sup>Min</sup> (multiple intestinal neoplasia) mutation, called APC<sup>Min/+</sup> mice. They frequently develop intestinal tumors [39,40]. Significant variation in tumor incidence occurs among APC<sup>Min/+</sup> mice with identical APC mutations and which are kept under identical laboratory conditions [39,40]. This variation is caused by differences in the genetic background of the APC mutation, which in turn depends on variation in “modifier genes” in different mouse strains [39,40] [41]. These are not involved directly in the process of carcinogenesis, but modify the phenotypic properties of APC+/- cells. This indicates that passengers can modulate the fitness of TSG+/- cells, as required by our model to observe accelerated evolution.

Our work adds to the growing literature that investigates the dynamics of fitness valley crossing under various conditions [19-21,23,25-28,42-44]. Beyond this immediate discipline, however, it is also interesting to consider our results in a wider scientific sense. In the presence of passenger mutations, cellular evolvability is predicted to

increase through the introduction of disadvantageous cells. The presence of these disadvantageous cells lowers overall population fitness, allowing intermediate TSG+/- mutants to have an overall higher relative fitness, which promotes faster generation of the TSG-/- double mutant. Studies of the onset of turbulence as well as pattern forming systems have revealed mechanisms of instability that act in a very similar manner. For example, in turbulence it has been shown that three-dimensional perturbations on parallel shear flows are damped more strongly than two-dimensional ones; but because of the slow decay, they provide a new base flow on which a new and richer class of fluctuations can grow more rapidly (details in Supplementary Materials). This provides a fundamental connection between principles in the physical sciences and the particular oncology question under consideration.

**Acknowledgement:** This study was funded in part by NIH grant U01CA187956.

**Figure 1:** Schematic representation of computational models. (A) Model of TSG inactivation in the context of passenger mutations. The cell types and associated fitness values ( $F$ ) are explained in the main text. The arrows show the mutational steps that generate the different cell types. The top row of cell types depicts standard evolutionary processes where the two copies of the TSG are sequentially inactivated. In addition, the model assumes that with a rate  $\nu$ , cells can accumulate passenger mutations. Cells with passenger mutants can also inactivate the TSG, as shown. (B) Model of TSG (APC) inactivation in the context of Lynch Syndrome and mutator phenotypes in colorectal cancer. This model does not contain passenger mutations. The basic evolutionary processes (along the top row of cells) are the same as above. The difference is that unmutated TSG $^{+/+}$  cells can inactivate mismatch repair mechanisms with a rate  $u$ , giving rise to mutator phenotypes that are characterized by an elevated mutation rate  $u_{\text{fast}}$ .

**Figure 2.** Accelerated crossing of fitness valleys (i.e. inactivation of TSGs) in the presence of passenger mutations. The computer simulations were run repeatedly, and the fraction of realizations in which the double mutant was created before a time threshold  $T_{\text{thr}}$  was determined. The number of simulations / sample sizes required by rare events is large and were chosen using reference [45]. For a fixed margin of error, the more rare the events, the larger the sample size. The margins of errors of our study are acceptable, as even in the worst situation ( $N=543449$ , recorded fraction=0.000077),

the margin of error is less than one third of the estimated proportion. Each graph plots the “fold acceleration”, which is the fraction of runs where the double mutant was generated in the presence of passenger mutations with fitness  $s_2$ , divided by the same measure in the absence of passenger mutations. The Z-score for population proportions was used to determine whether the difference in outcome between simulations with and without passenger mutations was statistically significant (the distribution under the null hypothesis, when the two true proportions are the same, is asymptotically normal). (A) Parameter regime where the double mutant evolved through a tunneling pathway. Panels (i-iii) show that a reduced cost of the intermediate TSG+/- mutant (higher value of  $s_1$ ) leads to a reduced effect of passenger mutations on the rate of evolution. Differences between outcomes with and without passenger mutations were statistically significant for (i)  $s_2 \geq 0.93$ , (ii)  $s_2 \geq 0.93$ , and (iii)  $s_2 = 0.99$ . (B) Parameter regime where the double mutant evolves by sequential fixation, in which passenger mutants have a stronger accelerating effect on the rate of evolution. Differences between outcomes with and without passenger mutations were statistically significant ( $p < 0.05$ ) for (i)  $s_2 \geq 0.95$ , (ii)  $s_2 \geq 0.93$ , and (iii)  $s_2 = 0.98$  &  $s_2 = 0.99$ . In panel (iiii), the fold acceleration was only determined for the highest values of  $s_2$  (where the effect is strongest), due to the extensive computational cost associated with this parameter set. Remaining parameters were chosen as follows.  $L_x = 0.15$ ,  $D = 0.01$ ,  $s_3 = s_2$ .  $n = 5 \times 10^{-2} / \mu$ ,  $N = 2500$ . The results do not depend on the assumption  $s_3 = s_2$ . How  $s_3$  needs to depend on  $s_2$  and  $s_1$  for the results to hold is defined in the supplementary materials. The time thresholds are given as follows for the individual graphs. (A) (i)  $T_{thr} = 8,000$ ; (ii)  $T_{thr} = 8,000$ , (iii)  $T_{thr} = 5,000$ ; (B) (i)  $T_{thr} = 120,000$ ; (ii)  $T_{thr} = 1,200,000$ ; (iii)  $T_{thr} = 1,500,000$ .

**Figure 3.** The average behavior of the contact process can be described by ordinary differential equations given in the Supplementary Materials. From these ODEs, equilibrium populations sizes, and hence the average rate of double mutant generation at equilibrium, can be calculated. In the presence of passenger mutations, this is given by  $(s_1 y^* + s_3 w^*) [1 - (x^* + y^* + z^* + w^*)/k]$ , where  $*$  denotes equilibrium population sizes. This is divided by the rate of double mutant generation in the absence of passenger mutations, given by  $s_1 y^* [1 - (x^* + y^*)/k]$ . This yields the fold increase of the double mutant generation rate that is mediated by passenger mutations, and is plotted in the graph for different values of  $s_1$  (fitness cost of TSG+/- cells without passenger mutations,  $y$ ). Parameter values were chosen as follows:  $r=0.1$ ,  $d=0.01$ ,  $\mu=10^{-5}$ ,  $n_1=2$ ,  $n_2=5000$ ,  $k=2500$ .

**Figure 4.** Rate of TSG inactivation in two types of simulations: assuming Lynch Syndrome, which involves the acquisition of microsatellite instability or MSI, i.e. cells with a faster mutation rate (Figure 1B); and assuming the passenger mutant pathway (Figure 1A). The computer simulations were run repeatedly, and the fraction of realizations in which the double mutant was created before a time threshold  $T_{thr}$  was determined. For the Lynch Syndrome model, simulations with different accelerated mutation rates,  $u_{fast}$ , were run. The fraction of runs that resulted in double mutant generation were divided by the fraction obtained without the existence of mutator phenotypes (MSI cells), which is the fold-acceleration depicted by the gray bars. The

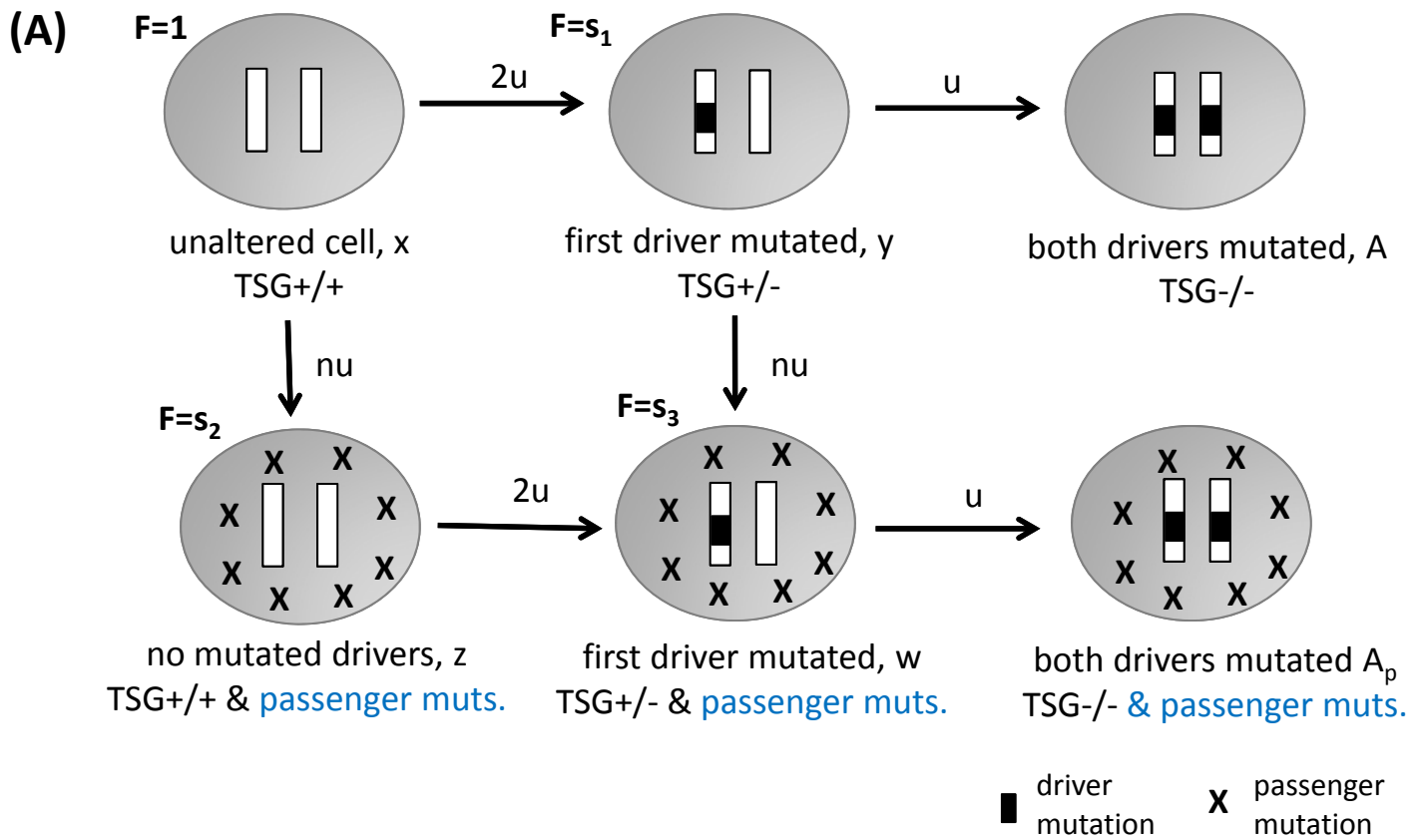


black bar shows the fold-acceleration derived from the passenger mutation model (without mutator cells), but with otherwise identical parameters. This was done in two settings (A) assuming that mutator cells do not suffer from a fitness cost; (B) assuming that mutator cells are characterized by a 1% fitness cost, brought about by the frequent generation of deleterious mutations. The parameters were chosen as follows.  $L_x=0.15$ ,  $D=0.01$ ,  $s_1=0.99$ ,  $s_2=0.99$ ,  $s_3=0.99$   $\mu=10^{-7}$ . For (a)  $s_M=1$ ,  $s_{M2}=1$ . For (b)  $s_M=0.99$ ,  $s_{M2}=0.99$ ,  $n=5 \times 10^{-2}/\mu$ ,  $T_{thr}=1,500,000$ ,  $N=2500$ .

1. Bozic I, Antal T, Ohtsuki H, Carter H, Kim D, et al. (2010) Accumulation of driver and passenger mutations during tumor progression. *Proc Natl Acad Sci U S A* 107: 18545-18550.
2. Greenman C, Stephens P, Smith R, Dalgliesh GL, Hunter C, et al. (2007) Patterns of somatic mutation in human cancer genomes. *Nature* 446: 153-158.
3. Greaves M, Maley CC (2012) Clonal evolution in cancer. *Nature* 481: 306-313.
4. Gatenby RA, Cunningham JJ, Brown JS (2014) Evolutionary triage governs fitness in driver and passenger mutations and suggests targeting never mutations. *Nat Commun* 5: 5499.
5. Haber DA, Settleman J (2007) Cancer: drivers and passengers. *Nature* 446: 145-146.
6. Futreal PA (2007) Backseat drivers take the wheel. *Cancer Cell* 12: 493-494.
7. McFarland CD, Korolev KS, Kryukov GV, Sunyaev SR, Mirny LA (2013) Impact of deleterious passenger mutations on cancer progression. *Proc Natl Acad Sci U S A* 110: 2910-2915.
8. McFarland CD, Yaglom JA, Wojtkowiak JW, Scott JG, Morse DL, et al. (2015) Passenger DNA alterations reduce cancer fitness in cell culture and mouse models. *bioRxiv*: 026302.
9. McFarland CD, Mirny LA, Korolev KS (2014) Tug-of-war between driver and passenger mutations in cancer and other adaptive processes. *Proc Natl Acad Sci U S A* 111: 15138-15143.
10. Knudson AG (2001) Two genetic hits (more or less) to cancer. *Nat Rev Cancer* 1: 157-162.
11. Polakis P (1997) The adenomatous polyposis coli (APC) tumor suppressor. *Biochim Biophys Acta* 1332: F127-147.
12. Tighe A, Johnson VL, Taylor SS (2004) Truncating APC mutations have dominant effects on proliferation, spindle checkpoint control, survival and chromosome stability. *J Cell Sci* 117: 6339-6353.
13. Green RA, Kaplan KB (2003) Chromosome instability in colorectal tumor cells is associated with defects in microtubule plus-end attachments caused by a dominant mutation in APC. *J Cell Biol* 163: 949-961.

14. Caldwell CM, Green RA, Kaplan KB (2007) APC mutations lead to cytokinetic failures in vitro and tetraploid genotypes in Min mice. *J Cell Biol* 178: 1109-1120.
15. Weaver BA, Silk AD, Montagna C, Verdier-Pinard P, Cleveland DW (2007) Aneuploidy acts both oncogenically and as a tumor suppressor. *Cancer Cell* 11: 25-36.
16. Gordon DJ, Resio B, Pellman D (2012) Causes and consequences of aneuploidy in cancer. *Nat Rev Genet* 13: 189-203.
17. Vogelstein B, Kinzler KW (2004) Cancer genes and the pathways they control. *Nat Med* 10: 789-799.
18. Vogelstein B, Lane D, Levine AJ (2000) Surfing the p53 network. *Nature* 408: 307-310.
19. Weinreich DM, Chao L (2005) Rapid evolutionary escape by large populations from local fitness peaks is likely in nature. *Evolution* 59: 1175-1182.
20. Serra MC, Haccou P (2007) Dynamics of escape mutants. *Theor Popul Biol* 72: 167-178.
21. Weissman DB, Desai MM, Fisher DS, Feldman MW (2009) The rate at which asexual populations cross fitness valleys. *Theoretical Population Biology* 75: 286-300.
22. Weissman DB, Feldman MW, Fisher DS (2010) The rate of fitness-valley crossing in sexual populations. *Genetics* 186: 1389-1410.
23. Iwasa Y, Michor F, Nowak MA (2004) Stochastic tunnels in evolutionary dynamics. *Genetics* 166: 1571-1579.
24. Komarova NL, Sengupta A, Nowak MA (2003) Mutation-selection networks of cancer initiation: tumor suppressor genes and chromosomal instability. *J Theor Biol* 223: 433-450.
25. Barton NH, Rouhani S (1987) The Frequency of Shifts between Alternative Equilibria. *Journal of Theoretical Biology* 125: 397-418.
26. Carter AJR, Wagner GP (2002) Evolution of functionally conserved enhancers can be accelerated in large populations: a population-genetic model. *Proceedings of the Royal Society of London Series B-Biological Sciences* 269: 953-960.
27. Kimura M (1985) The role of compensatory neutral mutations in molecular evolution. *Journal of Genetics* 64: 7-19.
28. Durrett R, Schmidt D (2008) Waiting for Two Mutations: With Applications to Regulatory Sequence Evolution and the Limits of Darwinian Evolution. *Genetics* 180: 1501-1509.
29. Durrett R, Foo J, Leder K (2016) Spatial Moran models, II: cancer initiation in spatially structured tissue. *Journal of Mathematical Biology* 72: 1369-1400.
30. Durrett R, Moseley S (2015) Spatial Moran Models I. Stochastic Tunneling in the Neutral Case. *Annals of Applied Probability* 25: 104-115.
31. Bauer B, Siebert R, Traulsen A (2014) Cancer initiation with epistatic interactions between driver and passenger mutations. *J Theor Biol* 358: 52-60.
32. Illingworth CJ, Mustonen V (2011) Distinguishing driver and passenger mutations in an evolutionary history categorized by interference. *Genetics* 189: 989-1000.
33. Holstege H, Pfeiffer W, Sie D, Hulsman M, Nicholas TJ, et al. (2014) Somatic mutations found in the healthy blood compartment of a 115-yr-old woman demonstrate oligoclonal hematopoiesis. *Genome Res* 24: 733-742.
34. Martincorena I, Roshan A, Gerstung M, Ellis P, Van Loo P, et al. (2015) Tumor evolution. High burden and pervasive positive selection of somatic mutations in normal human skin. *Science* 348: 880-886.
35. Komarova NL, Lengauer C, Vogelstein B, Nowak MA (2002) Dynamics of genetic instability in sporadic and familial colorectal cancer. *Cancer Biol Ther* 1: 685-692.
36. DeGregori J (2011) Evolved tumor suppression: why are we so good at not getting cancer? *Cancer Res* 71: 3739-3744.
37. Tomlinson I, Bodmer W (1999) Selection, the mutation rate and cancer: ensuring that the tail does not wag the dog. *Nat Med* 5: 11-12.

38. Marusyk A, Almendro V, Polyak K (2012) Intra-tumour heterogeneity: a looking glass for cancer? *Nat Rev Cancer* 12: 323-334.
39. Montagutelli X (2000) Effect of the genetic background on the phenotype of mouse mutations. *J Am Soc Nephrol* 11 Suppl 16: S101-105.
40. McCart AE, Vickaryous NK, Silver A (2008) Apc mice: models, modifiers and mutants. *Pathol Res Pract* 204: 479-490.
41. Nadeau JH (2001) Modifier genes in mice and humans. *Nat Rev Genet* 2: 165-174.
42. Iwasa Y, Michor F, Nowak MA (2004) Evolutionary dynamics of invasion and escape. *Journal of Theoretical Biology* 226: 205-214.
43. Komarova NL, Urwin E, Wodarz D (2012) Accelerated crossing of fitness valleys through division of labor and cheating in asexual populations. *Sci Rep* 2: 917.
44. Komarova NL, Shahriyari L, Wodarz D (2014) Complex role of space in the crossing of fitness valleys by asexual populations. *J R Soc Interface* 11: 20140014.
45. Hajian-Tilaki K (2011) Sample size estimation in epidemiologic studies. *Caspian J Intern Med* 2: 289-298.



**(B)**

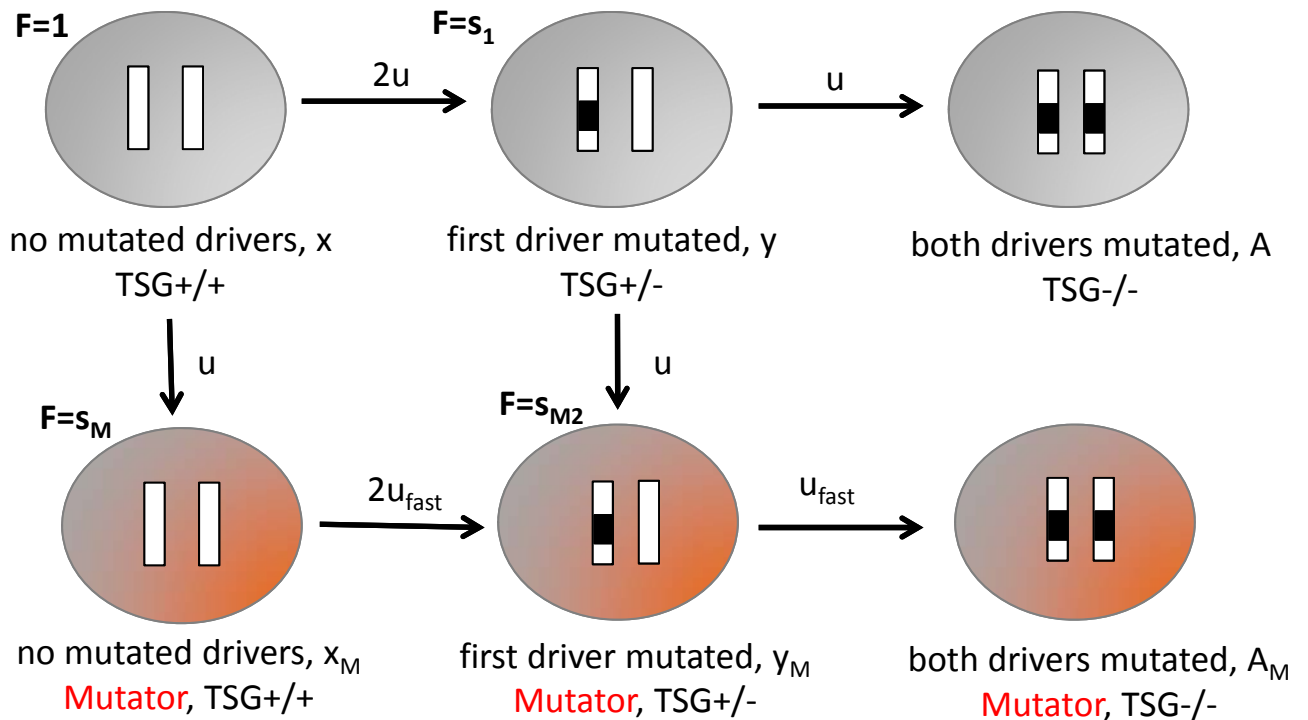


Fig 1

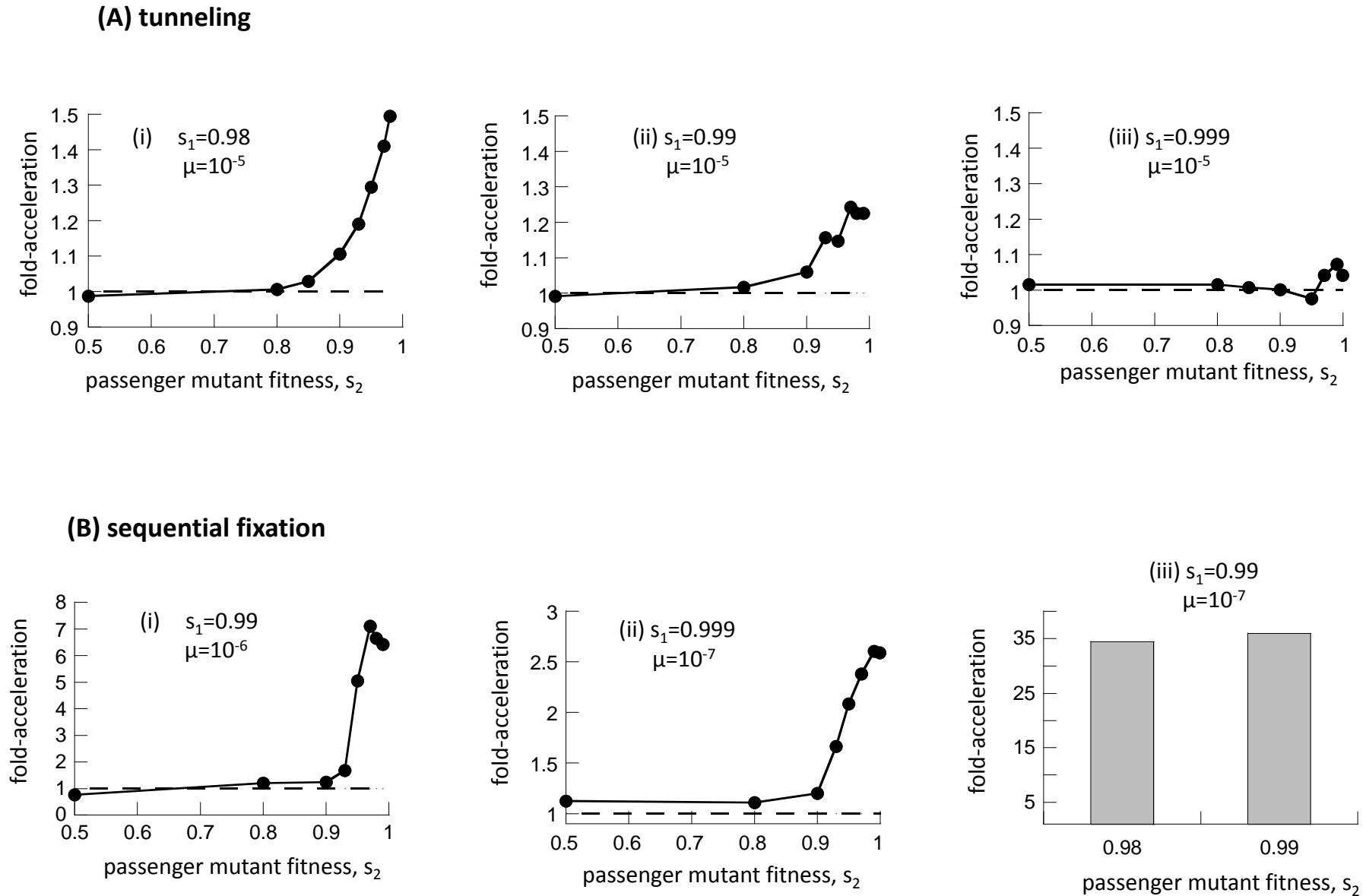


Fig 2

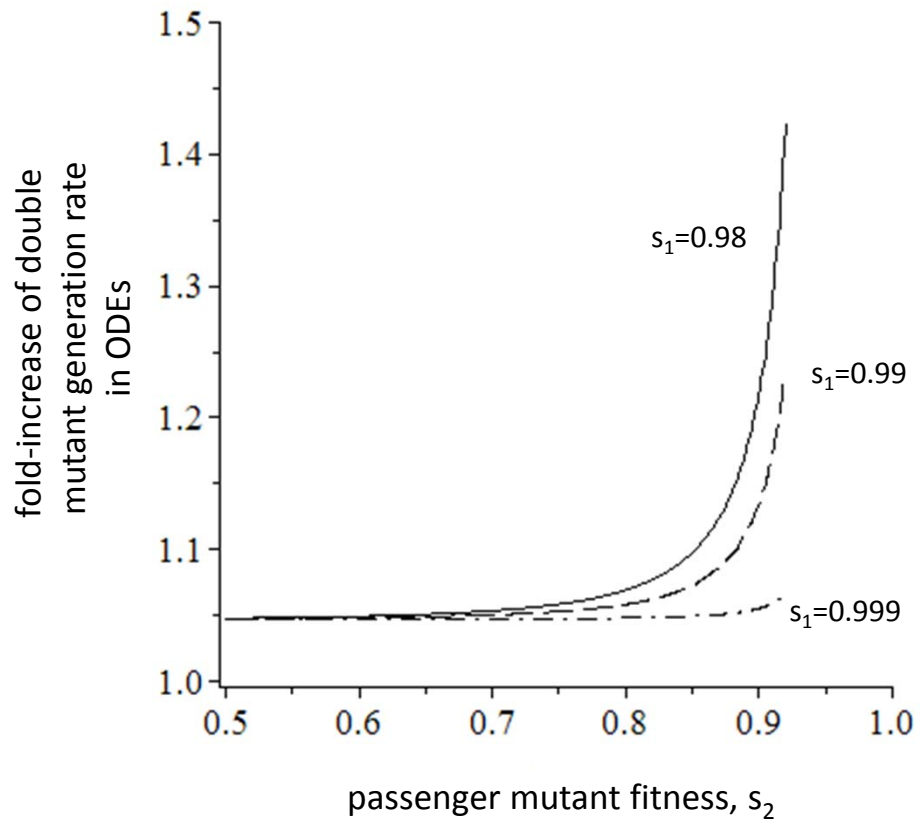


Fig 3

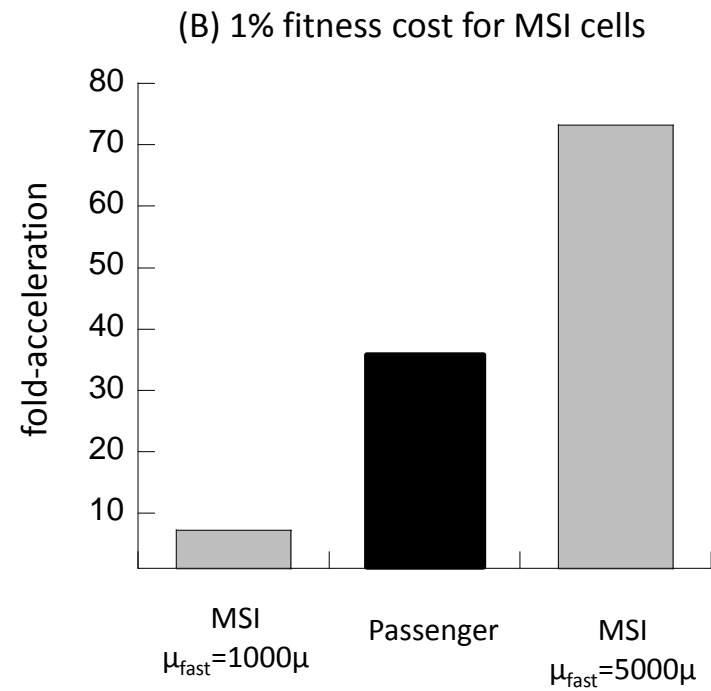
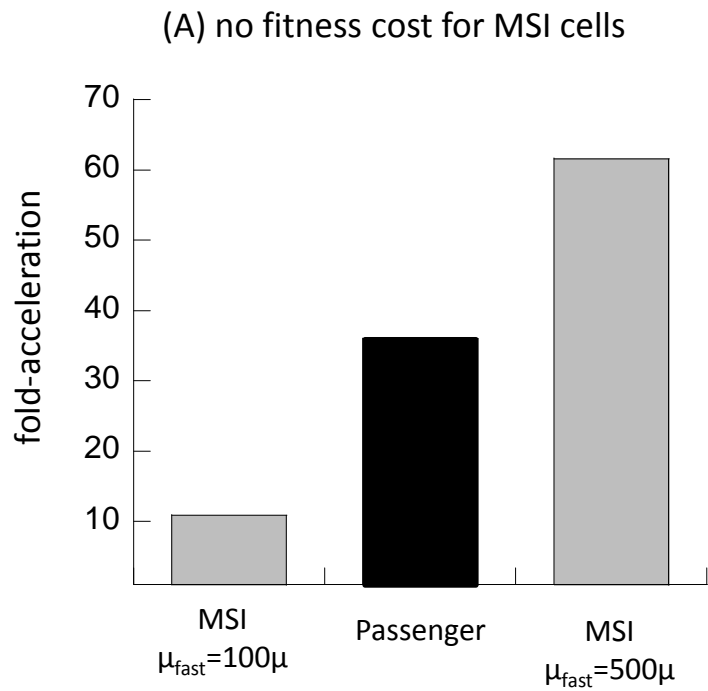


Fig 4

# Passenger mutations can accelerate tumor suppressor gene inactivation in cancer evolution

## Supplementary Materials

### Contents

<b>1</b>	<b>ODE description</b>	<b>2</b>
1.1	Logistic model . . . . .	2
1.2	Quasispecies equations . . . . .	7
<b>2</b>	<b>The contact process: further details</b>	<b>8</b>
2.1	Stochastic fitness of passenger mutants . . . . .	8
2.2	Spatially structured cell populations . . . . .	9
<b>3</b>	<b>Moran process</b>	<b>9</b>
3.1	Numerical simulations . . . . .	11
3.2	No passenger mutations . . . . .	11
3.3	Including the possibility of passenger mutations . . . . .	13
3.4	Comparisons . . . . .	15
<b>4</b>	<b>Connections between the role of 3D perturbations in turbulence/ patterns and passenger mutations in 2-hit mutant generation</b>	<b>16</b>



# 1 ODE description

## 1.1 Logistic model

We used the contact process as a basic model to study the effect of passenger mutations on the rate of fitness valley crossing, i.e. on the emergence of a  $TSG^{-/-}$  mutant. The contact process is a stochastic agent-based model and is described in the main text. Because the cell populations are assumed to be well-mixed (mass action), the average behavior of the contact process can also be studied by means of the ordinary differential equations. Denote by  $x$ ,  $y$ ,  $z$ , and  $w$  the wild type cells,  $TSG^{+/-}$  cells,  $TSG^{+/+}$  cells with passenger mutations, and  $TSG^{+/-}$  cells with passenger mutations respectively. The replication rate of wild type cells is denoted by  $r$  and the death rate by  $d$ . The fitness values for these types of cells are given by 1,  $s_1$ ,  $s_2$ , and  $s_3$  respectively. Let us denote by  $W$  the term describing the competition of the cell types:

$$W = 1 - \frac{x + y + z + w}{K},$$

where  $K$  has the meaning of carrying capacity. Then the ODEs are given by

$$\dot{x} = r(1 - (n_1 + n_2)\mu)xW - dx, \quad (1)$$

$$\dot{y} = rn_1\mu xW + s_1ry(1 - n_2\mu)W - dy, \quad (2)$$

$$\dot{z} = rn_2\mu xW + s_2rzW(1 - n_1\mu)W - dz, \quad (3)$$

$$\dot{w} = rn_1\mu s_2zW + rn_2\mu s_1yW + rs_3wW - dw. \quad (4)$$

By rescaling time we can assume that  $r = 1$  for the purpose of the analysis below (the value of  $d$  is rescaled accordingly). The replicative cost of the various mutants are given by  $s_1$ ,  $s_2$ , and  $s_3$ , which can have values between zero and one. We further assume that

$$s_3 \leq s_2 \leq s_1. \quad (5)$$

The death rate of cells is given by  $d$  and for simplicity is assumed to be identical for all cell populations. The number of different passenger mutations that a cell can accumulate is given by  $n_1$ , and the number of mutations that give rise to the  $TSG^{-/-}$  genotype is given by  $n_2 = 2$ . The mutation rate is denoted by  $\mu$ . Finally, the system is characterized by the carrying capacity  $K$ , which describes the maximum number of cells that can exist in this system. We do not explicitly include the dynamics of the advantageous  $TSG^{+/+}$  cells in this model because the contact process simulations are stopped as soon as the first cell of this type is generated.

System (1-4) cannot be obtained by averaging the dynamics of the contact process, because of the presence of higher moments. The ODE's however provide valuable insights into the system behavior.

In the subsequent analysis, we denote for convenience

$$q = n_2\mu, \quad u = n_1\mu,$$

which are the total mutation rates of the passenger mutation production and TSG inactivation, respectively. There are 5 distinct steady states in this system, only one of which is of interest to us.

**Selection mutation equilibrium.** If  $q < 1 - s_3$ , this model is characterized by a stable equilibrium that describes the coexistence of all cell populations. It is determined by a balance between the rate at which these cells are generated by mutation, and the rate at which they are lost as a result of their selective disadvantage (costs  $s_1$ ,  $s_2$ , and  $s_3$ ). Hence, this level is referred to as the “selection-mutation balance”. The dominant populations are  $x$  and  $z$ , and they are given by

$$x = \bar{x} = \frac{K(1-d-q)(1-s_2-q)}{(1-q)(1-s_2)}, \quad z = \bar{z} = \frac{Kq(1-d-q)}{(1-q)(1-s_2)}, \quad (6)$$

where terms of the order  $u$  and higher are omitted. The other two populations are of the order  $u$ :

$$y = \bar{y} = \frac{Ku(1-d-q)(1-s_2-q)}{(1-q)(1-s_1)(1-s_2)}, \quad w = \bar{w} = \frac{Kqu(1-d-q)(1-q-s_1s_2)}{(1-q)(1-s_1)(1-s_2)(1-s_3-q)^2}, \quad (7)$$

where terms of the order  $u^2$  and higher are omitted. At this equilibrium, the total population is given by

$$\bar{x} + \bar{y} + \bar{z} + \bar{w} = K \left( 1 - \frac{d}{1-q} \right)$$

(again, only the zeroth order in  $u$  is kept).

**Double-hit mutant production.** The quantity of interest is the total rate of double-hit mutant ( $TSG^{+/+}$ ) production, given by

$$Q = (s_1\bar{y} + s_3\bar{w})W,$$

where the number of intermediate  $TSG^{+/-}$  cells is taken at the selection-mutation equilibrium. This quantity represent the total division rate of the two populations,  $y$  and  $z$ , capable of producing double-hit mutants; the factor  $W$  represents the effect of finite carrying capacity. We have

$$Q = \frac{Kud(1-d-q)[s_1(1-s_2)(1-s_3) - q(s_1-s_3)]}{(1-q)(1-s_1)(1-s_2)(1-s_3-q)},$$

To determine whether passenger mutations facilitate the production of double-hit mutants, we consider the quantity  $dQ/dq$ . Positivity of this quantity indicates that a higher intensity of passenger mutant production will increase the rate of double-hit mutant production. We have

$$\frac{dQ}{dq} = c_2q^2 + c_1q + c_0, \quad (8)$$

where coefficients  $c_0, c_1, c_2$  depend on  $s_1, s_2, s_3$ , and  $d$ :

$$c_0 = (1-s_3)(s_3 - s_1s_2 - d(s_3 + s_1(1-2s_2 - (1-s_2)s_3))), \quad (9)$$

$$c_1 = 2(1-s_3)(ds_1(1-s_2) + s_1s_2 - s_3), \quad (10)$$

$$c_2 = (s_3 - s_1s_2)(1-s_3) - d(s_1-s_3). \quad (11)$$

Depending on the signs of these coefficients, the behavior of the quadratic function (8) changes. All the cases are mapped out in figure 1 for two different parameter sets. The signs

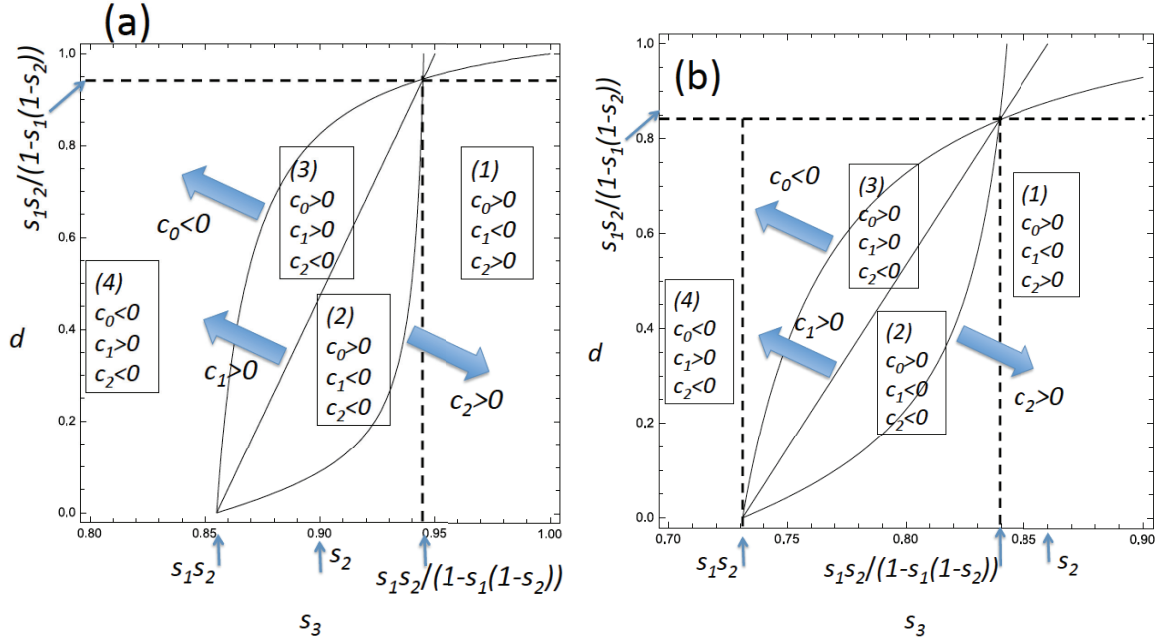


Figure 1: Phase diagrams showing the sign regions of parameters  $c_0$ ,  $c_1$ , and  $c_2$ , equations (8) and (9-11). The axes are  $s_3$  (horizontal) and  $d$  (vertical). The other parameters are fixed at (a)  $s_1 = 0.95$  and  $s_2 = 0.9$ ; (b)  $s_1 = 0.86$  and  $s_2 = 0.85$ .

of the three coefficients are shown as regions in the  $d - s_3$  diagram, with  $s_1$  and  $s_2$  fixed to particular values. The three lines in figure 1 (a) and (b) correspond to  $c_0 = 0$ ,  $c_1 = 0$ , and  $c_2 = 0$ . It can be shown that all three lines intersect in one point, given by

$$s_3 = d = \frac{s_1 s_2}{1 - s_1 (1 - s_2)} \equiv s_*. \quad (12)$$

If  $s_1$  and  $s_2$  are close to 1, then this quantity is also close to 1, and it does not make biological sense to consider values of  $d$  larger than this threshold. Therefore, there are 4 biologically relevant regions in the  $s_3 - d$  space, which are all marked in the figure.

Let us denote by  $q_1$  and  $q_2$  the smaller and the larger roots of the polynomial (8), respectively:

$$\frac{-c_1 \pm \sqrt{c_1^2 - 4c_0 c_2}}{2c_2}.$$

In the four different regions of figure 1, the quantity  $dQ/dq$  is positive in different intervals of the parameter  $q$ . These are listed below. Recall that  $dQ/dq > 0$  is interpreted as ‘‘passenger mutations facilitate double-mutant production’’.

- (1) In this regime, see figure 2(1),  $dQ/dq > 0$  for all values of  $q$  if

$$s_3 > s_*, \quad (13)$$

see definition (12). This regime is depicted by the upper parabola in figure 2(1), and by the horizontal line marked ‘‘(1) upper’’ in figure 3. Note that this regime is not

always relevant. Biologically speaking, the fitness  $s_3$  must be bounded as

$$s_1 s_2 \leq s_3 \leq s_2.$$

It may happen that  $s_* > s_2$ , as in figure 1(a); in this case, this regime is biologically irrelevant. In general, we have

$$s_* < s_2$$

as long as

$$s_2 < s_1 < \frac{1}{2 - s_2}.$$

This inequality holds for the parameters of figure 1(b). The region of values allowed by this inequality for  $s_1$  shrinks as  $s_2$  grows. If inequality (13) is reversed (the lower parabola in figure 2(1)), that is,

$$\frac{1 + d + s_1 s_2 - \sqrt{(1 + d + s_1 s_2)^2 - 4(ds_1 + s_1 s_2)}}{2} < s_3 < s_*,$$

the condition for  $dQ/dq > 0$  becomes

$$q < q_1$$

(note that  $q_2 > 1$  in this regime). This is illustrated by the horizontal line marked “(1) lower” in figure 3(a).

(2) and (3) These regimes are characterized by

$$\frac{s_1(d + s_2 - 2ds_2)}{1 - d + ds_1 - ds_1 s_2} < s_3 < \frac{1 + d + s_1 s_2 - \sqrt{(1 + d + s_1 s_2)^2 - 4(ds_1 + s_1 s_2)}}{2}.$$

In these regimes, see figure 2(2) and (3),  $dQ/dq > 0$  for all values of  $q$  if

$$q < q_2,$$

see also the horizontal line marked “(2) and (3)” in figure 3(a).

(4) Finally, we have the regime where

$$s_1 s_2 < s_3 < \frac{s_1(d + s_2 - 2ds_2)}{1 - d + ds_1 - ds_1 s_2},$$

figure 2(4). In this case, we have

$$q_1 < q < q_2$$

(see the upper parabola in figure 2(4) and the horizontal line marked “(4) upper” in figure 3(a)). In the biologically irrelevant case where  $s_3 < s_1 s_2$ , no values of  $q$  will correspond to  $dQ/dq > 0$  (the lower parabola in figure 2(4) and the horizontal line marked “(4) lower” in figure 3(a)).

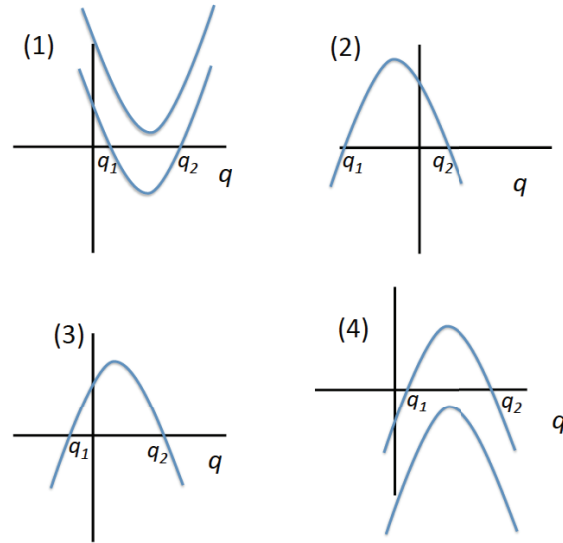


Figure 2: The four regimes identified in figure 1. For each regime, the parabola  $dQ/dq$  (equation (8)) is plotted schematically, to show the position of the relevant interval of  $q$ , depending on the signs of the coefficients  $c_0, c_1, c_2$ .

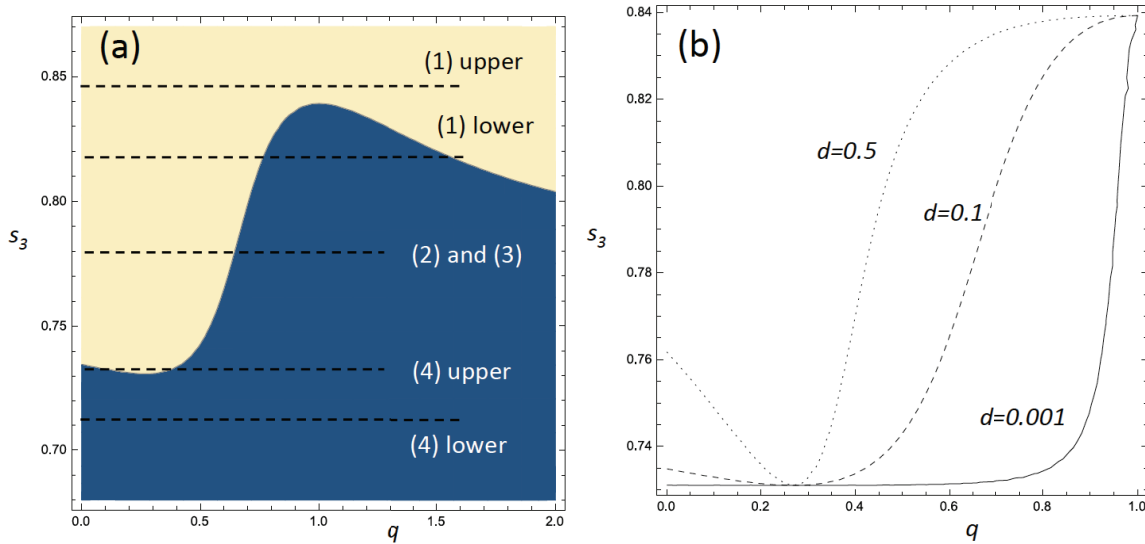


Figure 3: The  $s_3 - q$  diagram showing the regions where  $dQ/dq$  is positive; the parameters correspond to figure 1(b). (a) For a fixed  $d = 0.1$ , the regions where  $dQ/dq > 0$  and  $dQ/dq < 0$  are shown in yellow and blue respectively. Dashed horizontal lines correspond to fixing a particular value of  $s_3$ , in the four regimes of figure 1(b). (b) The contour lines  $dQ/dq = 0$  for three different values of  $d$  ( $dQ/dq$  is positive above the lines).

The region where  $dQ/dq > 0$  shrinks as  $d$  grows, as shown in figure 3(b), and the ODE model predicts that the effect of the passenger mutations weakens for larger  $d$ . On the other hand, the applicability of the ODE approach to explain the behavior of the stochastic system breaks down for large values of  $d$ , because as  $d$  grows relative to the division rates, the population size shrinks, and in the stochastic system, instead of the phenomenon of tunneling, one observes sequential fixation. As a consequence, we find that the passenger mutations play a stronger enhancing role in the creation of double-hit mutants.

## 1.2 Quasispecies equations

An alternative way to model the dynamics of the competing types in a fixed population is to use the quasi species type equations:

$$\dot{x} = x[1 - (n_1 + n_2)\mu] - \phi x, \quad (14)$$

$$\dot{y} = [n_1\mu x + s_1(1 - n_2\mu)y] - \phi y, \quad (15)$$

$$\dot{z} = [n_2\mu x + s_2(1 - n_1\mu)z] - \phi z, \quad (16)$$

$$\dot{w} = [s_1n_2\mu y + s_2n_1\mu z + s_3w] - \phi w, \quad (17)$$

where

$$\phi = \frac{x + s_1y + s_2z + s_3w}{N},$$

and  $N$  is the total, fixed population. The difference between this system and the previous one is the fact that in equations (14-17) the population always remains equal to  $N$ , that is,  $x + y + z + w = N$ , and the number of independent variables is effectively three, and the death rate (given by  $\phi$ ) is non-constant, as it adjusts to exactly balance the divisions. In equations (1-4), the death rate is fixed (to  $d$ ), the population is non-constant, and all populations are independent variables. While equations (1-4) are analogous to the contact process, quasispecies equations can be thought as an ODE equivalent of the Moran process.

The mutation-selection equilibrium of this system is stable if  $q < 1 - s_3$ , and is identical to that given by equations (6,7), where we set  $d = 0$  and replace  $K \rightarrow N$ . At this equilibrium we have

$$\phi = 1 - q.$$

Note that although the equilibrium of the quasi species system is given by equations (6,7) with  $d = 0$ , it cannot be obtained from system (1-4) directly by setting  $d = 0$  in the ODEs. Equations (1-4) with  $d = 0$  do not describe the same dynamics as the quasi species system; one obvious difference is the absence of death in the former system, and the death rate given by  $\phi$  in the latter. Aside from the trivial equilibrium, equations (1-4) with  $d = 0$  have a family of neutral equilibria that satisfy  $\bar{x} + \bar{y} + \bar{z} + \bar{w} = K$ , with no further restrictions on the values of the individual populations.

Returning to the quasi species system, consider the function that defines the rate of double-hit mutant production. This is given by

$$Q_{qs} = s_1\bar{y} + s_3\bar{w},$$

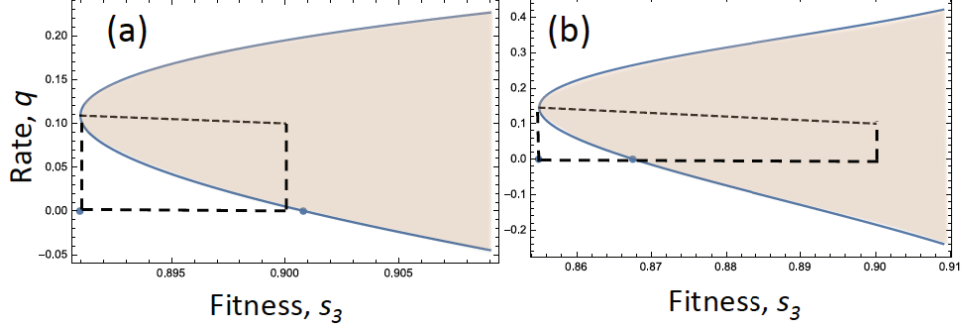


Figure 4: A parametric plot showing the dependence of the production term for the quasi species equations,  $Q_{qs}$ , on the rate of passenger mutations,  $q$ . The horizontal axis is  $s_3$ , with vertical dashed lines restricting this quantity to the range  $s_1 s_2 \leq s_3 \leq s_2$ . The vertical axis is  $q$ , with dashed lines restricting this quantity to the range  $0 \leq q \leq 1 - s_3$ . The shaded regions correspond to the quantity  $Q_{qs}$  increasing with  $q$  (i.e.  $dQ_{qs}/dq > 0$ ); they correspond to condition (18). The parameters  $s_2 = 0.9$ , and (a)  $s_1 = 0.99$ ; (b)  $s_1 = 0.95$ .

where the number of intermediate  $TSG^{+/-}$  cells is taken at the selection-mutation equilibrium. We have

$$Q_{qs} = \frac{Nu(1-q)[s_1(1-s_2)(1-s_3) - q(s_1-s_3)]}{(1-s_1)(1-s_2)(1-s_3-q)},$$

where terms of order  $u^2$  and higher are omitted. It is interesting to investigate how this quantity depends on  $q$ , the rate of passenger mutant production. Considering  $dQ_{qs}/dq$ , we can see that this quantity can be positive only if  $s_3 > s_1 s_2$ , which, together with inequality (5) gives the following bounds for the quantity  $s_3$ :

$$s_1 s_2 < s_3 \leq s_2.$$

Under this restriction, we have  $dQ_{qs}/dq > 0$  as long as

$$1 - s_3 - \sqrt{\frac{s_3(1-s_3)(s_3 - s_1 s_2)}{s_1 - s_3}} < q < 1 - s_3 + \sqrt{\frac{s_3(1-s_3)(s_3 - s_1 s_2)}{s_1 - s_3}}. \quad (18)$$

This is illustrated in figure 4.

## 2 The contact process: further details

### 2.1 Stochastic fitness of passenger mutants

The main analysis has been done by assigning passenger mutants an ‘‘average fitness’’, defined by their division rate. In reality, however, different passenger mutations can bring about different costs to fitness. Therefore, we also constructed a model in which the fitness cost of the passenger mutants was drawn from a distribution with values between 0 (maximum cost) and 1 (no cost, the division rate is equal to that of the  $x$  population), and a mean

of  $s_2$ . Specifically, upon generation, each mutant cell was assigned a fitness value, given by  $g^{1/\alpha}$ , where  $g$  is a uniformly distributed random number in  $[0, 1]$ , and  $\alpha = s_2/(1 - s_2)$ . This corresponds to the beta distribution  $B(\alpha, 1)$ , which is a power function distribution on  $[0, 1]$ , which has mean  $s_2$ . Similar methodology was used to generate passenger mutants with one copy of TSG inactivated (type  $w$ ). As seen in Figure 5, the accelerating effect of passenger mutations remains robust in this model.

## 2.2 Spatially structured cell populations

We also investigated a model version which violated the assumption of mass-action. Thus, upon division, we assumed that the offspring cell could only be placed into the immediate vicinity of the source cell. To do so, we constructed a two-dimensional grid and tracked the location of each cell on the grid. The place for the offspring cell was randomly chosen from the eight nearest neighboring spots. If the chosen spot was empty, the daughter cell was placed there, otherwise, the division was aborted. Apart from that, the model remained the same. As seen in Figure 5, the accelerating effect of passenger mutations remains robust in the spatially structured model, although the extent of the acceleration is less pronounced in the spatial compared to the non-spatial model. The reason is that evolution of the double mutant in the absence of passenger mutations is accelerated in the spatial relative to the non-spatial model, while the rate of evolution in the spatial and non-spatial models is similar in the presence of passenger mutations. The effect of space on the rate of fitness valley crossing in this type of model is complex. Space can accelerate or slow down the rate of fitness valley crossing, depending on the exact parameter combinations. Therefore, in the spatial model, the effect of passenger mutations can in principle be either weaker than in the non-spatial model (as shown here), or it can be stronger. This has not been explored further here, because it is beyond the scope of the current investigation.

## 3 Moran process

As explained in the main text, the assumptions underlying the contact process were also formulated as a Moran process, which assumes a constant population. We analyzed the model assuming that passenger mutants are characterized by an average fitness cost,  $s_2$ , and that cells can mix perfectly (mass action). The Moran process was used for the following reasons: (i) This model is more amenable to mathematical analysis, which is presented below. (ii) The constant population Moran process is more suited to describing the evolutionary dynamics in healthy tissue, which is tightly regulated by homeostatic mechanisms. While the occurrence of passenger mutations is mostly discussed in the context of a cancer, the literature documents an array of seemingly non-functional mutations in the healthy tissue of people, especially with advancing age [6, 11].



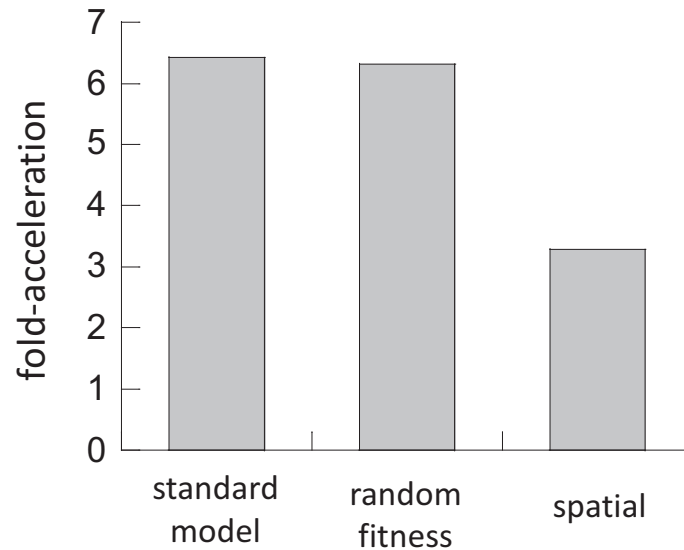


Figure 5: Effect of alternative model assumptions. The fold-acceleration of evolution due to passenger mutants is plotted for different models. The computer simulations were run repeatedly, and the fraction of realizations in which the double mutant was created before a time threshold was determined. The fold acceleration is the fraction of runs where the double mutant was generated in the presence of passenger mutations divided by the fraction in the absence of passenger mutations. This is done for one specific value of passenger mutant fitness,  $s_2/s_3$ . The standard model corresponds to the model explored in the main text. The random fitness models draws passenger mutant fitness from a distribution as explained in the text of the supplementary materials. The spatial model assumes that cell offspring can only be placed into the eight nearest neighboring spots, as explained in the text of the supplementary materials. All three fold increases are statistically significant by the Z-test ( $p < 0.05$ ). Parameters were chosen as follows.  $L_x = 0.15$ ,  $D = 0.01$ ,  $\mu = 10^{-6}$ ,  $s_1 = 0.99$ ,  $s_2 = 0.99$ ,  $s_3 = 0.99$ ,  $n = 5 \times 10^{-2}/\mu$ ,  $T_{thr} = 120,000$ ,  $N = 2500$ .

### 3.1 Numerical simulations

The accelerating effect of passenger mutations is the same as that observed in the contact process, and the parameter dependencies remain qualitatively the same as in the contact process. Figure 6 repeats the analysis and shows the same plots as done for Figure 2 in the main text.

A mathematical analysis of the Moran process is presented next.

### 3.2 No passenger mutations

We envisage a Moran process where a constant population of  $N$  cells is in a steady turnover. At each elementary update, each cell has the same probability to die (given by  $1/N$ ), and then this cell is replaced by an offspring of another cell. Cells are chosen for reproduction with a probability proportional to their fitness value. One unit of time corresponds to  $N$  elementary updates. We will investigate a general process where three types of mutations are possible. The first mutation (which corresponds to the inactivation of the first copy of a TSG) happens with probability  $u_1$  per cell division. The second copy of a TSG is inactivated with probability  $u_2$ . Finally, passenger mutations are acquired with probability  $nu$ .

Denote the number of  $TSG^{+/-}$  mutants of fitness  $s_1$  by  $i$ ,  $i \in \{0, 1, 2, \dots\}$ , and by  $\varphi_i(t)$  the probability to have  $i$  mutants at time  $t$ . State  $E$  corresponds to the creation of the first double-hit mutant, and this is an absorbing state. We have:

$$\begin{aligned} P_{i \rightarrow i+1} &= \frac{N-i}{N} \left( \frac{N-i}{s_1 i + (N-i)} u_1 + \frac{s_1 i}{s_1 i + (N-i)} (1-u_2) \right) \approx u_1 + \frac{s_1 i}{N} (1-u_2), \\ P_{i \rightarrow i-1} &= \frac{i}{N} \frac{N-i}{s_1 i + (N-i)} (1-u_1) \approx \frac{i}{N} (1-u_1), \\ P_{i \rightarrow E} &= \frac{s_1 i}{s_1 i + (N-i)} u_2 \approx \frac{s_1 i u_2}{N}, \end{aligned} \quad (19)$$

where we neglected  $i$  compared to  $N$ . We will rescale time by having  $N$  elementary events during a unit time. The function  $\varphi_i(t)$  satisfies the following Kolmogorov forward equation,

$$\dot{\varphi}_i = \varphi_{i-1} (N u_1 + s_1 (i-1) (1-u_2)) + \varphi_{i+1} (i+1) (1-u_1) - \varphi_i (N u_1 + (s_1 + 1 - u_1) i).$$

Introducing the probability generating function,

$$\Psi(x, t) = \sum_i \varphi_i x^i,$$

we can derive the following 1st order PDE:

$$\frac{\partial \Psi}{\partial t} = \frac{\partial \Psi}{\partial x} (x^2 s_1 (1-u_2) + (1-u_1) - x(s_1 + 1 - u_1)) - N u_1 \Psi(1-x).$$

Suppose  $x(t)$  is the solution of the initial value problem,

$$\dot{x} = x^2 s_1 (1-u_2) + 1 - u_1 - x(s_1 + 1 - u_1), \quad x(0) = 1. \quad (20)$$

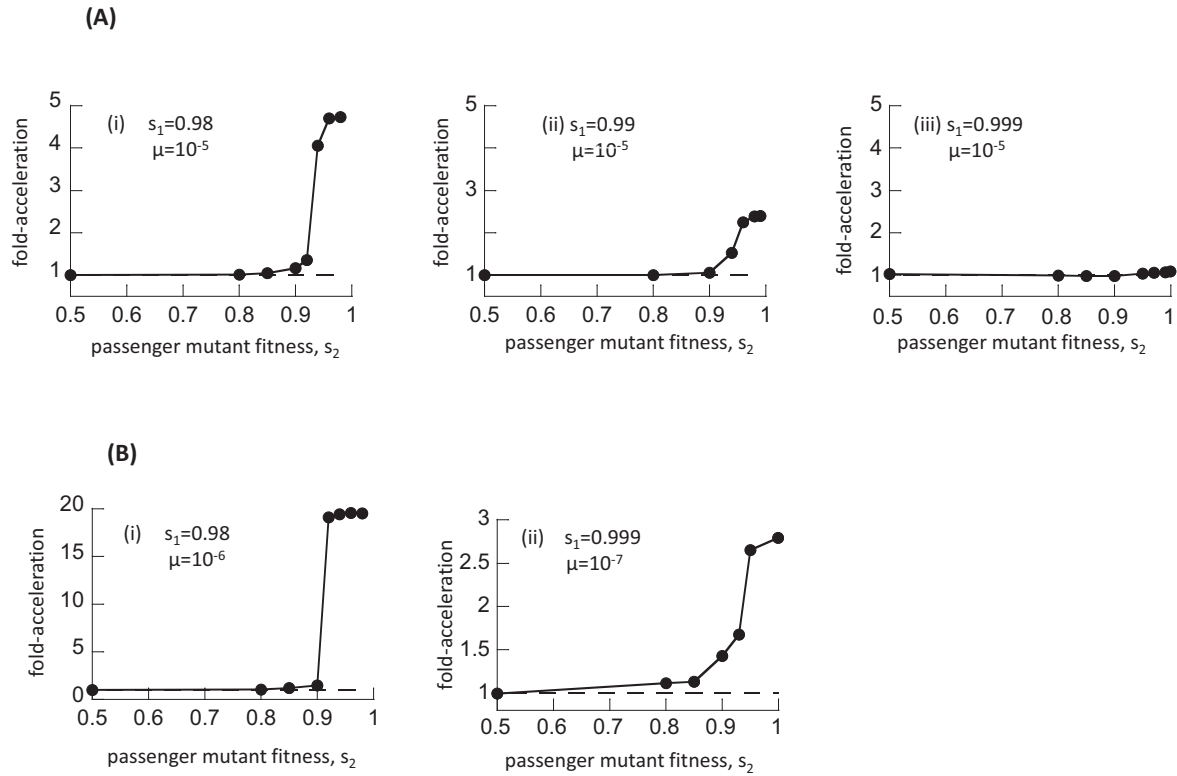


Figure 6: Accelerated evolution through passenger mutations in the Moran process model. The computer simulation was run repeatedly, and the fraction of realizations in which the double mutant was created before a time threshold  $T_{thr}$  was determined. Each graph plots the “fold acceleration”, which is the fraction of runs where the double mutant was generated in the presence of passenger mutations with fitness  $s_2$ , divided by the same measure in the absence of passenger mutations. (A) Parameter regime where the double mutant evolved through a tunneling pathway. Panels (i-iii) show that a reduced cost of the intermediate TSG<sup>+/-</sup> mutant (higher value of  $s_1$ ) leads to a reduced effect of passenger mutations on the rate of evolution. (B) Parameter regime where the double mutant evolves by sequential fixation, in which passenger mutants have a stronger accelerating effect on the rate of evolution. Parameters were chosen as follows. We assumed  $s_3 = s_2$ , although results do not depend on this constraint, see above;  $n = 5 \times 10^{-2}/\mu$ . The total population size was  $N = 2500$ . The time threshold are given as follows. (A) (i)  $T_{thr} = 5,000$ ; (ii)  $T_{thr} = 500$ ; (iii)  $T_{thr} = 200$ . (B) (i)  $T_{thr} = 50,000$ ; (ii)  $T_{thr} = 50,000$ . Differences between outcomes with and without passenger mutations were statistically significant ( $p < 0.05$ ) for A(i)  $s_2 \geq 0.85$ , A(ii)  $s_2 \geq 0.9$ , A(iii)  $s_2 \geq 0.95$ , B(i)  $s_2 \geq 0.85$ , B(ii)  $s_2 \geq 0.8$ .

The probability to create a double-hit mutant by time  $T$  is given by

$$P(T) = 1 - \exp\left(-\int_0^T R(t) dt\right), \quad (21)$$

where the tunneling rate is given by

$$R(t) = Nu_1(1 - x(t)). \quad (22)$$

### 3.3 Including the possibility of passenger mutations

Consider a Moran process. We assume that  $n \gg 1$ , such that the rate of passenger mutations is much larger than the TSG mutation rate. Let us ignore the TSG mutations, and find the population equilibrium. This can be done by considering the quasi-species equations:

$$\dot{X}_1 = r_1 X_1(1 - \mu) - \phi X_1, \quad \dot{X}_2 = r_1 X_1 \mu + r_2 X_2 - \phi X_2, \quad \phi \equiv r_1 X_1 + r_2 X_2, \quad X_2 = 1 - X_1, \quad (23)$$

where  $\mu$  is a one-way mutation rate and  $r_1 > r_2$  are the fitness values of the two populations,  $X_1$  and  $X_2$ , respectively. We have

$$X_1 = 1 - \frac{r_1 \mu}{r_1 - r_2}, \quad X_2 = \frac{r_1 \mu}{r_1 - r_2}, \quad (24)$$

as long as  $\mu < 1 - \frac{r_2}{r_1}$ ; otherwise we have  $X_1 = 0$  and  $X_2 = 1$ .

Populations  $x$  and  $w$  have the fitness values of 1 and  $s_2$  respectively, and the mutation rate is given by  $nu$ . Therefore, we can assume that at a quasi-steady state, we have a population of wild type cells of size  $x = N - j_0$ , and the population of cells with passenger mutations,  $w = j_0$ , given by

$$j_0 = \frac{Nnu}{1 - s_2}.$$

The mean fitness of this “resident” population is given by

$$\frac{(N - j_0) + s_2 j_0}{N} = 1 - nu.$$

We further assume that the following processes happen in the system:

- population  $w$  acquires passenger mutations, population  $z$ , at rate  $2u$ .
- population  $x$  creates mutants  $y$  at rate  $2u$ ,
- population  $z$  creates double-hit mutants at rate  $u_2$ ,
- population  $y$  creates double-hit mutants at rate  $(1 - nu)u_2$ .

By following the same method as before, we denote the number of mutants of fitness  $s_1$  by  $i$ , and the number of mutants of fitness  $s_3$  by  $j$ , with  $i, j \in \{0, 1, 2, \dots\}$ . The function  $\varphi_{i,j}(t)$  is the probability to have  $i$  and  $j$  mutants of each type at time  $t$ . As before, state

$E$  corresponds to the creation of the first double-hit mutant, and this is an absorbing state. We have:

$$\begin{aligned}
 P_{i,j \rightarrow i+1,j} &= \frac{N-i-j}{N} \left( \frac{(N-j_0)(1-nu)u_1 + s_1i(1-nu)(1-u_2)}{s_1i + s_3j + (1-nu)(N-i-j)} \right) \\
 &\approx \frac{(N-j_0)(1-nu)u_1 + s_1i(1-nu)(1-u_2)}{N(1-nu)}, \\
 P_{i,j \rightarrow i,j+1} &= \frac{N-i-j}{N} \left( \frac{s_2j_0u_1 + s_3j(1-u_2) + s_1inu(1-u_2)}{s_1i + 3_2j + (1-nu)(N-i-j)} \right) \\
 &\approx \frac{s_2j_0u_1 + s_3j(1-u_2) + s_1inu(1-u_2)}{N(1-nu)}, \\
 P_{i,j \rightarrow i-1,j} &= \frac{i}{N} \frac{(N-i-j)(1-nu)}{s_1i + s_3j + (1-nu)(N-i-j)} (1-u_1) \approx \frac{i}{N} (1-u_1), \\
 P_{i,j \rightarrow i,j-1} &= \frac{j}{N} \frac{(N-i-j)(1-nu)}{s_1i + s_3j + (1-nu)(N-i-j)} (1-u_1) \approx \frac{j}{N} (1-u_1), \\
 P_{i,j \rightarrow i-1,j+1} &= \frac{i}{N} \frac{s_2j_0u_1 + s_3j(1-u_2) + s_1inu(1-u_2)}{s_1i + s_3j + (1-nu)(N-i-j)} \\
 &\approx \frac{i}{N} \frac{s_2j_0u_1 + s_3j(1-u_2) + s_1inu(1-u_2)}{N(1-nu)}, \\
 P_{i,j \rightarrow i+1,j-1} &= \frac{j}{N} \frac{(N-j_0)(1-nu)u_1 + s_1i(1-nu)(1-u_2)}{s_1i + s_3j + (1-nu)(N-i-j)} \\
 &\approx \frac{j}{N} \frac{(N-j_0)(1-nu)u_1 + s_1i(1-nu)(1-u_2)}{N(1-nu)}, \\
 P_{i,j \rightarrow E} &= \frac{(s_1i + s_3j)u_2}{s_1i + s_3j + (1-nu)(N-i-j)} \approx \frac{(s_1i + s_3j)u_2}{N(1-nu)},
 \end{aligned}$$

where we neglected  $i$  and  $j$  compared to  $N$ . We will rescale time by having  $N$  elementary events during a unit time. The function  $\varphi_{i,j}(t)$  satisfies the following Kolmogorov forward equation,

$$\begin{aligned}
 (1-nu)\dot{\varphi}_{i,j} &= \\
 &\varphi_{i-1,j} [(N-j_0)(1-nu)u_1 + s_1(i-1)(1-nu)(1-u_2)] \\
 &+ \varphi_{i+1,j} (i+1)(1-nu)(1-u_1) \\
 &+ \varphi_{i,j-1} [s_2j_0u_1 + s_3(j-1)(1-u_2) + s_1inu(1-u_2)] \\
 &+ \varphi_{i,j+1} (j+1)(1-nu)(1-u_1) \\
 &+ \varphi_{i+1,j-1} \frac{j_0}{N} (i+1)s_2u_1 \\
 &+ \varphi_{i-1,j+1} \frac{N-j_0}{N} (j+1)(1-nu)u_1 \\
 &- \varphi_{i,j} \left( [(N-j_0)(1-nu) + s_2j_0]u_1 + (i+j)(1-u_1)(1-nu) + s_1i + s_3j \right. \\
 &\left. + \frac{u_1}{N} [s_2j_0i + (N-j_0)j(1-nu)] \right). \tag{25}
 \end{aligned}$$

In the above equation, higher order terms (nonlinear in  $i/N, j/N$ ) are omitted. Introducing the probability generating function,

$$\Psi(x, y, t) = \sum_i \varphi_{i,j} x^i y^j,$$

we can derive the following 1st order PDE:

$$(1 - nu) \frac{\partial \Psi}{\partial t} = f_x(x, y) \frac{\partial \Psi}{\partial x} + f_y(x, y) \frac{\partial \Psi}{\partial y} - g(x, y) \Psi, \quad (26)$$

where

$$\begin{aligned} f_x(x, y) &= x^2 s_1 (1 - nu)(1 - u_2) + (1 - u_1)(1 - nu) + x y s_1 n u (1 - u_2) \\ &\quad - x \left[ s_1 + (1 - nu)(1 - u_1) + \frac{u_1}{N} s_2 j_0 \right] + y \frac{j_0}{N} s_2 u_1, \\ f_y(x, y) &= y^2 s_3 (1 - u_2) + (1 - u_1)(1 - nu) \\ &\quad - y \left[ s_3 + (1 - nu)(1 - u_1) + \frac{u_1}{N} (N - j_0)(1 - nu) \right] + x \frac{N - j_0}{N} u_1 (1 - nu), \\ g(x, y) &= u_1 [(N - j_0)(1 - x)(1 - nu) + s_2 j_0 (1 - y)]. \end{aligned} \quad (27)$$

Suppose that  $x(t)$  and  $y(t)$  are the solutions of the initial value problem,

$$\dot{x} = f_x(x, y)/(1 - nu), \quad x(0) = 1, \quad (28)$$

$$\dot{y} = f_y(x, y)/(1 - nu), \quad y(0) = 1. \quad (29)$$

Then the probability to create a double-hit mutant by time  $T$  is calculated as

$$P_{pass}(T) = 1 - \exp \left( - \int_0^T R_{pass}(t) dt \right), \quad (30)$$

where the tunneling rate,  $R(t)$ , is given by

$$R_{pass}(t) = \frac{u_1 [(N - j_0)(1 - x(t))(1 - nu) + s_2 j_0 (1 - y(t))]}{1 - nu}. \quad (31)$$

We can see that if  $n = 0$  (no passenger mutations), we have  $j_0 = 0$  and  $R(t) = R_{pass}(t)$ , that is, equations (22) and (31) coincide, as they should.

### 3.4 Comparisons

Figure 7 summarizes the theoretical results and compares them with the results of stochastic simulations. For simplicity we assumed that the fitness of  $TSP^{+/-}$  cells is  $s_2$ , irrespective of their passenger mutation status; similar results will hold for a range of assumptions on fitness  $s_3$ . The fitness of  $TSG^{+/-}$  cells in the absence of passenger mutants,  $s_1$ , was fixed, and several values of parameter  $s_2 = s_3$  were tested. For each value of  $s_2$ , a large number of simulations were run. Each simulation was stopped when either a  $TSG^{-/-}$  mutant was created, or a stopping time  $T$  was reached. For each simulation, we recorded whether a

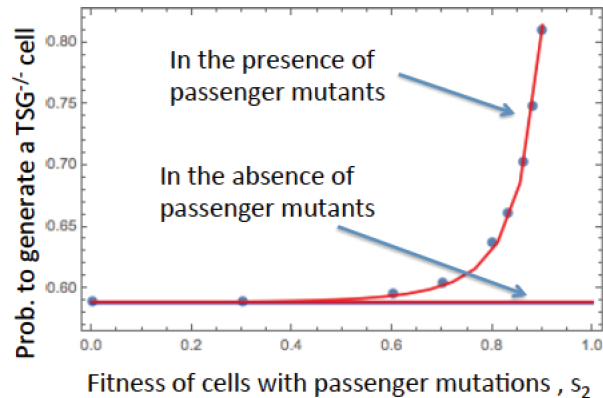


Figure 7: The probability to create a  $TSG^{-/-}$  mutant by time  $T$ . Numerical simulations are depicted in blue, and analytical results (formulas (21) and (30)) in red. The parameters are  $s_1 = 0.9$ ,  $u = 10^{-4}$ ,  $u_1 = 2u$ ,  $u_2 = u$ ,  $n = 500$ ,  $T = 2000$ . The number of simulation was more than 500,000 for each value of  $s_2$ .

$TSG^{-/-}$  mutant was created, and then the probability of  $TSG^{-/-}$  creation by time  $T$  was calculated. This is shown by blue dots in figure 7. These numerical results were compared with quantity  $P_{pass}$ , equation (30), which is plotted as a red curve in figure 7. The horizontal lines in this figure correspond to the numerically calculated (blue) and analytically obtained (red, formula (21)) values  $P(T)$ , the probability to generate a  $TSG^{-/-}$  mutant in the absence of passenger mutations.

First of all, we observe that numerical and analytical results are in a very good agreement. Secondly, we can see that the probability of  $TSG^{-/-}$  mutant creation is higher in the presence of passenger mutation, as long as the fitness of the corresponding cells is not too low (but lower than the base fitness of single hit mutants,  $s_1$ ).

## 4 Connections between the role of 3D perturbations in turbulence/ patterns and passenger mutations in 2-hit mutant generation

The idea that initially disadvantageous or passenger mutations might allow cells make a more rapid transition to carcinogenesis by providing more effective pathways for normal cells to reach that state stems from analogies with the onset of turbulence in fluids and the nucleation of dislocation defects in natural patterns.

No one area of science stands alone. Analogies can provide useful insights. Behavior seen in one context can be seen in another where it initially it might not have been expected. The common denominator is often mathematics. For example, freak waves, about which ocean navigators have many terrifying stories and experiences, also occur, surprisingly perhaps, in optical fibers. Their common origin is modulated wave-trains whose behaviors are captured in water waves, in optics, in plasmas by non-integrable corrections (in the water wave context,

the full Euler equations) to the canonical and universal nonlinear Schroedinger equation for wave envelopes. In the right circumstances, the slowly modulated wave-train can evolve into a signal in which one sees a burst of very large amplitude waves over very few wavelengths. The phenomenon is robust and universal.

The analogies which motivated the present studies had their origins in the onset of turbulence and in pattern formation. The study of the onset of turbulence in shear flows goes back to the classical experiments of Osborne Reynolds in 1883 on pipe flow in which water flows along a pipe [14, 15]. He found that typically the flow will be in the turbulent state at values of the Reynolds' number,  $UD/\nu$ , of about 2000 and laminar below that.  $U$  is the flow velocity say on the center line,  $D$  the pipe diameter and  $\nu$  the kinematic viscosity. If he was very careful to avoid creating disturbances at the inflow point (he could use a funnel to minimize these), then he, and others later, showed that one could retain the laminar character of the flow up to Reynolds numbers of over 10,000. The conclusion is that turbulence onset in pipes is a nonlinear and three dimensional phenomenon and requires finite amplitude perturbations to trigger the turbulent state.

That conclusion is supported by the negative results emanating from linear stability theory which, for shear flows such as pipe flow, plane Poiseuille flow (pressure driven flow in a channel), Couette flow (flow between oppositely moving plates), boundary layer flow (flow over a flat plate) does not predict what is observed. For both pipe and Couette flow, the flows are linearly stable for all Reynolds numbers. For Poiseuille flow, linear instability sets in at  $R \sim 6000$ . In all cases, the observed onset of turbulence typically occurs at values of  $R$  much less than what linear stability predicts. Moreover, the observed turbulent fields have a richer character than do the original shear flows. The latter usually have the form  $(U(y), 0, 0)$ , namely a velocity in the  $x$  or streamwise direction depending on one of the spanwise directions (the radial coordinate in pipes, the vertical coordinate in Couette, Poiseuille and boundary layer flows). On the other hand, turbulent flow depends on all three coordinates,  $x$ ,  $y$  and  $z$ . Linear stability theory also gave rise to another misleading result. A theorem, due to Squire [16], says that of all the disturbances, the two dimensional ones are the least damped and, in the case of Poiseuille flow for  $R > 6000$ , would be the most amplified. So surely, if they were the least damped, they would also be the most important. Clearly that is not what is observed. Turbulence is strongly three dimensional. So linear stability theory does not necessarily identify the right set of shapes from which a turbulent field grows.

Such a conclusion was very evident in the early experiments of [7] (and earlier papers and colleagues) who demonstrated that turbulence is clearly a three dimensional phenomenon. Squire's theorem pointed in the wrong direction. A theoretical breakthrough came with the work of [1] who discovered that three dimensional spanwise oscillations could give rise to shapes which grew algebraically in time. Such behavior could lead to a redistribution of streamwise velocity via what Marten Landahl called "up-lift". Start with a two dimensional  $x, y$  perturbation with its vorticity (the swirl of the fluid) in the  $z$  direction. Think that the fluid rotates around the vortex line in a clockwise direction, down in front, up in back. Now perturb that vortex line, think  $\cos(bz)$ , so that near the front of this loop at  $z = 0$ , there is a flow which lifts fluid parcels up. The slower moving fluid which is carried away from the boundary into the faster moving flow outside can have a significant effect on the velocity profile  $U(y)$  and even can produce points of inflexion where  $U''(y) = 0$ . In boundary layers,



such behavior can lead to very fast instabilities and turbulent bursts. Variations on this theme, namely the presence of algebraic in time growth of disturbances which are very weakly exponentially damped (have a  $t$  dependence  $F(R)t * \exp(-at)$ ,  $F$  depends on the Reynolds number  $R$ ,  $a$  is very small) which can lead to amplitudes at which linear and nonlinear terms in the flow equations are of equal importance, have been discussed in many papers in the past forty years. We list some notable ones: [1, 4, 9, 2, 13, 3, 17, 18, 19, 20, 5]. Trefethen et al identified the source of  $t$  growth multiplied by weak exponential decay behavior as generic if the linear stability operator  $L$  is non-normal. That is the case for shear flows. In short, the message is that comparing the relative and very slow exponential damping rates of different structures and then choosing the least damped or most amplified state as the one which controls the subsequent dynamics is not always correct.

In an altogether different context, similar behavior occurs when a pattern adjusts from one wavelength to another. Let  $w = A \cos(kx + f) \sin z$ ,  $A \exp(if)$  is called the complex amplitude, be the vertical velocity field of a set of roll-like cells whose axes are along the  $y$  direction. If the wavenumber  $k$  is too much larger than some preferred value  $k_0$ , then an instability, called the Eckhaus instability, occurs, the net result of which is that the pattern will remove roll pairs until the new wavenumber is inside the Eckhaus stability boundary. The fastest growing linear instability mode has the envelope  $A$  remain a function of  $x$  and  $t$  only. Along the axis, that is  $y$  dependent, perturbations have a lower growth rate. Now in order to remove a roll pair with an envelope  $A$  only depending on  $x$ , the amplitude  $A$  has to enter its nonlinear regime and become zero along the full length of one pair of rolls. And, if one is very careful to control and suppress  $y$  dependent perturbations, that is indeed what happens [10]. But in most cases, and especially near or just inside the Eckhaus stability boundary,  $y$  dependent perturbations, which initially grow more slowly or decay slightly faster, overtake the  $y$  independent perturbation and dominate the outcome. Instead of having to remove a roll pair by making the real amplitude zero along the length of the roll, the  $y$  dependent perturbation produces a pair of dislocations at which positions the complex field  $A \exp(if)$  is zero. The dislocations repel each other and travel to the boundaries at the end of the roll axes where they are absorbed, thereby removing an excess roll pair. It is easy to see that this scenario is a much more efficient means of wavenumber adjustment. The three dimensional disturbance, although less favorable according to linear stability analysis, is the one which eventually prevails. In short, allowing for three dimensional perturbations, the flow is deformed so that it can find a more efficient pathway to remove the roll pair.

The lesson we draw from these examples is this. Just because a structure is deemed disadvantageous with respect to others based on linear stability analysis does not mean that when the full dynamics is in play, it does not play a significant role. Whereas the structure may initially appear to be less fit, it may provide for a richer set of pathways in order that a system can achieve some desired goal more readily. Motivated by these ideas, we ask in this paper whether similar ideas might apply in enabling cells to use slightly disadvantageous mutations, here called passenger mutations, in order to provide a higher probability in reaching some genuinely malignant mutation by providing the deformed cell with a richer set of pathways to reach that goal.

## References

- [1] Benney, D.J. and Lin, C.C., 1960. On the secondary motion induced by oscillations in a shear flow, *Phys. of Fluids*, 3, 656–657.
- [2] Benney, D.J. and Gustavsson, L.H., 1981. A new mechanism for linear and nonlinear hydrodynamic instability. *Stud Appl. Math.* 64, 185–209.
- [3] Benney, D.J., 1984. The evolution of disturbances in shear flows at high Reynolds numbers. *Stud. Appl. Math.* 70, 1–19.
- [4] Ellingsen, T. and Palm, E. , 1975. Stability of linear flow, *Phys. of Fluids* 18, 487–488.
- [5] Grossman, S., 2000. The onset of shear flow turbulence. *Rev. of Modern Phys.* 72, 603–618.
- [6] Holstege, H., Pfeiffer, W., Sie, D., Hulsman, M., Nicholas, T.J., Lee, C.C., Ross, T., Lin, J., Miller, M.A., Ylstra, B. and Meijers-Heijboer, H., 2014. Somatic mutations found in the healthy blood compartment of a 115-yr-old woman demonstrate oligoclonal hematopoiesis. *Genome research*, 24(5), pp.733-742.
- [7] Klebanoff, P.S. and Tidstrom, K. D. and Sargent, L. M., 1962. The three-dimensional nature of boundary-layer instability, *J. Fluid Mech.*,12, Pt. 1 (1962).
- [8] Komarova, N.L., Sengupta, A. and Nowak, M.A., 2003. Mutation-selection networks of cancer initiation: tumor suppressor genes and chromosomal instability. *Journal of theoretical biology*, 223(4), pp.433-450.
- [9] Landahl, M.T., 1980. A note on an algebraic instability of inviscid parallel shear flows. *J. Fluid Mech* 98, 243–251.
- [10] Lowe, M. and Gollub, J.P., 1985. Pattern Selection near onset of convection. *Phys. Rev. Lett.* 55, 2575-2578.
- [11] Martincorena, I., Roshan, A., Gerstung, M., Ellis, P., Van Loo, P., McLaren, S., Wedge, D.C., Fullam, A., Alexandrov, L.B., Tubio, J.M. and Stebbings, L., 2015. High burden and pervasive positive selection of somatic mutations in normal human skin. *Science*, 348(6237), pp.880-886.
- [12] Nowak, M.A., Komarova, N.L., Sengupta, A., Jallepalli, P.V., Shih, I.M., Vogelstein, B. and Lengauer, C., 2002. The role of chromosomal instability in tumor initiation. *Proceedings of the National Academy of Sciences*, 99(25), pp.16226-16231.
- [13] Orszag, S.A. and Patera, A.T., 1983. Secondary instability of wall bounded shear flows. *J. Fluid Mech.* 128, 347–385.
- [14] Reynolds, O., 1883. An experimental investigation of the circumstances which determine whether the motion of water shall be direct or sinuous, and of the law of resistance in parallel channels. *Proc. Roy. Soc. London* 35, 84–99.

- [15] Reynolds, O., 1883. An experimental investigation of the circumstances which determine whether the motion of water shall be direct or sinuous, and of the law of resistance in parallel channels. *Phil. Trans. R. Soc. London* 174, 935–982.
- [16] Squire, H.B., 1933. On the Stability for Three-Dimensional Disturbances of Viscous Fluid Flow between Parallel Walls. *Proc. Roy. Soc. London Ser. A* 142, 621–628.
- [17] Trefetnen, N., Trefetnen, A., Reddy, S., and Driscoll, T., 1993. Hydrodynamic stability without eigenvalues, *Science* 261, 578–584.
- [18] Waleffe, F., 1995. Beyond transients to a self-sustaining process. *Stud. Appl. at.* 95, 319–343.
- [19] Waleffe, F., 1995. Transition in shear flows: Nonlinear normality versus non-normal linearity. *Phys. Fluids* 7, 3060–3066.
- [20] Waleffe, F., 1998. Three-dimensional coherent states in plane shear flows. *Whys. Rev. Lett.* 81, 4140–4148.