Mutant evolution in spatially structured and fragmented expanding populations

Dominik Wodarz[1,2] and Natalia L Komarova[2]

(1) Department of Population Health and Disease Prevention, Program in Public Health, Susan and Henry Samueli College of Health Sciences, University of California Irvine, Irvine CA 92697

(2) Department of Mathematics, University of California Irvine, Irvine CA 92697

Abstract

Mutant evolution in spatially structured systems is important for a range of biological systems, but aspects of it still require further elucidation. Adding to previous work, we provide a simple derivation of growth laws that characterize the number of mutants of different relative fitness in expanding populations in spatial models of different dimensionalities. These laws are universal and independent of "microscopic" modeling details. We further study the accumulation of mutants and find that with advantageous and neutral mutants, more of them are present in spatially structured, compared to well-mixed colonies of the same size. The behavior of disadvantageous mutants is subtle: if they are disadvantageous through a reduction in division rates, the result is the same, and it is the opposite if the disadvantage is due to a death rate increase. Finally, we show that in all cases, the same results are observed in fragmented, non-spatial patch models. This suggests that the patterns observed are the consequence of population fragmentation, and not spatial restrictions per se. We provide an intuitive explanation for the complex dependence of disadvantageous mutant evolution on spatial restriction, which relies on desynchronized dynamics in different locations/patches, and plays out differently depending on whether the disadvantage is due to a lower division rate or a higher death rate. Implications for specific biological systems, such as the evolution of drug-resistant cell mutants in cancer or bacterial biofilms, are discussed.

Introduction

The dynamics of mutant creation and invasion are relatively well understood under a variety of conditions and assumptions, mostly assuming perfect mixing of individuals. In the context of constant populations, the fixation probability of mutants as well as fixation times have been thoroughly defined under various assumptions in the population genetics literature (KIMURA 1962; PATWA AND WAHL 2008). The emergence of mutants in exponentially growing bacterial populations is also well studied, based on the famous Luria-Delbruck experiments (LURIA AND DELBRUCK 1943) and the resulting rich theoretical framework (ZHENG 1999; KEPLER AND OPREA 2001; DEWANJI *et al.* 2005; KOMAROVA *et al.* 2007). This has been instrumental for understanding the principles according to which antibiotic-resistant microbes emerge (JOHNSON AND LEVIN 2013), and has also been applied to studying the emergence of drug resistance in some cancers (GOLDIE AND COLDMAN 1983; GOLDIE AND COLDMAN 1998; KOMAROVA AND WODARZ 2005). The majority of tumors, however, are characterized by the growth of 2D and 3D spatial structures, and so is the growth of bacteria in biofilms. Recent experimental and theoretical work (FUSCO *et al.* 2016) has extended our understanding of mutant emergence in such spatially structured, expanding populations. An excess of mutational jackpot events was observed in spatial compared to well-mixed systems. Such events result from mutations arising at the surface of expanding, spatially structured populations, surfing at the edge of range expansions, and appearing as mutant “sectors” or “slices”. These jackpot

events can occur relatively late in the expansion process, which is in contrast to well-mixed systems in which mutational jackpot events can only occur early on in the population growth process (FUSCO *et al.* 2016). Hence, overall, the average number of mutants when the total population reaches a given threshold size is significantly larger in spatial compared to non-spatial settings (FUSCO *et al.* 2016). This work was done under the assumption that cells do not die, and theory and computations were mostly developed in the context of neutral mutants. A number of other papers studied the spread of mutants in spatial and fragmented settings. A study by (GRALKA AND HALLATSCHEK 2019) considered the spread of advantageous mutants focusing on the role of habitat fragmentation. In (GRALKA *et al.* 2016) several scaling relationships for individual mutant clones were derived, including advantageous mutants, in the context of expanding 2D colonies, both ”flat” and radial. Spatial dynamics of disadvantageous mutants were studied by (OTWINOWSKI AND KRUG 2014), focusing on Write-Fischer dynamics in constant populations, and by (LAVRENTOVICH *et al.* 2016), who looked at the mutational meltdown. The dynamics of advantageous mutants in a 3D patch model was studied by (WACLAW *et al.* 2015) in the context of tumors.

Here, we build on the existing work and investigate the dynamics of mutant emergence and growth in spatially structured cell populations assuming varying death rates, different mutant fitness, different dimensionalities of space, and different spatial modeling approaches. One of the two main messages of this paper is to report interesting dynamics observed for disadvantageous mutants,

which could apply for example to drug-resistant mutants that emerge before the onset of therapy. If the disadvantage is caused by a larger death rate of the mutant cells, then we find that in contrast to other scenarios, the number of mutants at a given size can be larger in a well-mixed compared to the spatial system. If, on the other hand, the fitness disadvantage arises because of a slower replication rate, then more mutants are found in the spatial compared to the non-spatial system, similar to the results obtained for neutral or advantageous mutants.

The second message is that surprisingly similar results are obtained in explicitly spatial models and in patch models, where local within-patch dynamics are governed by perfect mixing, but individuals migrate to other patches. Interestingly, the results do not depend on the assumption that patches are spatially arranged, with migration of individuals to nearest neighboring patches. The same outcomes are observed if migration can occur to any randomly chosen patch in the system. Therefore, the properties of mutant growth in the spatial agent-based model might not be the direct consequence of spatial dynamics, but the consequence of population fragmentation.

In addition, in this paper we provide a simple and straightforward derivation of scaling growth laws that govern cell expansion in spatially constrained models. The so-called “surface growth” law of homogeneous cell colonies in space has previously been described in experiments (FREYER AND SUTHERLAND 1985; BRU *et al.* 1998; GUNTHER *et al.* 2007) and in the modeling

literature (BRU *et al.* 2003; BLOCK *et al.* 2007; KOMAROVA AND WODARZ 2010; RODRIGUEZ-BRENES *et al.* 2013; TALKINGTON AND DURRETT 2015; MURPHY *et al.* 2016). Here we study the laws of mutant generation, spread, and competition with the wild type individuals, in the context of spatially restricted colony expansion. We derive formulas that relate the expected number of disadvantageous, neutral, and advantageous mutants to the total population size in different spatial dimensions. While some of these laws have been derived previously (such as the growth laws of neutral mutants (FUSCO *et al.* 2016)), others are novel or confirm previous numerical observations (see the conjecture of paper (f) on advantageous mutant growth in 2D expansion),

Materials and methods

Two-dimensional agent-based model

We used a 2-dimensional, agent-based model, where a 2-dimensional square grid is considered. A spot on the grid can be empty or can contain a cell, which is either wild-type or mutant. At each time step, the grid is randomly sampled N times, where N is the total number of cells currently in the system. If the sampled cell is wild-type, the cell attempts division (described below) with a probability L_w or dies with a probability D_w. When reproduction is attempted, a target spot is chosen randomly from the nearest neighbors of the cell (8 neighbors, i.e. the Moore neighborhood, was used unless otherwise noted). If that spot is empty, the offspring cell is placed there. If it is already filled, the division event is

aborted. The offspring cell is assigned wild type with probability 1-*u* and it is a mutant with probability *u*. If the sampled spot contains a mutant cell, the same processes occur. Attempted division occurs with a probability L_m, and the cell dies with a probability D_m. The offspring of a mutant cell is always a mutant in the absence of back mutation. In a different version of the model, a mutant's offspring can be of wild-type with probability *u*. Initial and boundary conditions are determined by the specific geometric setting investigated. For 2D spatial simulations, an *nxn* square domain is considered. At the boundaries of the domain, a spot is assumed to have fewer neighbors, i.e more division events will fail. The simulations start with a single wild-type cell, placed into the center of the grid. Simulations always stop before the boundary of the grid is reached. For 1D cylinder simulations, we use an *nxw* rectangular domain of width *w*. We start with an array of *w* wild type cells at the left boundary of the domain, and impose periodic boundary conditions in the transversal direction. In each simulation, the cell population is allowed to grow to a size M, and the number of mutant cells at this size is recorded. Such simulations are performed repeatedly, and the average number of mutants is calculated. Simulation runs, in which the total cell population goes extinct due to stochastic effects, are ignored.

Analysis of 2D spatial stochastic models is presented in Section 3 of the Supplement. Growth laws for different geometries are derived in Section 4 of the Supplement.

Modeling exponential growth

To compare the spatial dynamics to a well-mixed system, we considered a simple stochastic simulation of exponential growth. Denoting the number of wild-type cells with x_w and the number of mutant cells with x_m, one of the cell types is chosen with a probability given by their proportion in the whole cell population. Wild-types can divide with a probability L_w and can die with a probability D_w. Mutants can divide with a probability l_m and die with a probability d_m. Mutations of wild type cells happen with probability *u*. As in the spatial system, the average number of mutants at population size M was determined.

A patch (island / metapopulation) model

We also considered an alternative modeling approach to capture mutant dynamics in spatially structured populations. Instead of tracking the spatial location of individual cells, we analyze a model that consists of a two-dimensional grid of nxn patches or demes. Deme models to approximate spatial dynamics have been explored before (WACLAW *et al.* 2015), and our approach is conceptually related. Within each patch, local dynamics occur where cells are assumed to mix perfectly. At each time step, cells are allowed to migrate to a different patch with a given rate. In each local patch, Gillespie simulations (GILLESPIE 1977) of the following ordinary differential equation model were run:

$$\frac{dx_i}{dt} = l_w x_i (1-u)\left(1-\frac{x_i+y_i}{k}\right) - d_w x_i - m x_i + \frac{m}{8}\sum_{neighbors,j} x_j$$

$$\frac{dy_i}{dt} = u l_w x_i \left(1-\frac{x_i+y_i}{k}\right) + l_m y\left(1-\frac{x_i+y_i}{k}\right) - d_m y_i - m y_i + \frac{m}{8}\sum_{neighbors,j} y_j$$

Wild-type cells are denoted by x_i, and mutant cells by y_i, where the subscript *i* enumerates the spatial locations in a two-dimensional array. Wild type cells divide with a density-dependent rate $l_w(1-(x+y)/k)$, die with a rate d_w, and migrate out of the patch with a rate *m*. Migration is assumed to occur to one of the eight neighboring patches, chosen randomly. During replication of the wild-type cells, a mutant cell can be generated with a probability u. Mutant cells divide with a density-dependent rate $l_m(1-(x+y)/k)$, die with a rate d_m, and migrate with a rate m.

In an alternative (fragmentation) model, instead of migrating with rate *m*/8 per patch to one of the eight neighboring patches, cells migrate with probability *m*/(n-1) per patch to any other patch regardless of its location. This holds for cells in all the patches in the system, thus removing a spatial component from the migration process. Otherwise, the equations are identical to the ones above.

Simulations were started with a single wild-type cell in the middle patch. The simulations were run until the total cell population size, summed over all patches, reached size M. At this point, the number of mutants summed over all patches was recorded. This was done repeatedly, and the average number of mutants at size M was determined. Instances of the simulation that resulted in population extinction across all patches were ignored.

For comparison, Gillespie simulations were performed in a non-fragmented, well mixed system described by the following equations:

$$\dot{x} = l_w x(1-u)\left(1-\frac{x+y}{n^2 k}\right) - d_w x,$$
$$\dot{y} = u l_w x\left(1-\frac{x+y}{n^2 k}\right) + l_m y\left(1-\frac{x+y}{n^2 k}\right) - d_m y$$

The carrying capacity of the non-fragmented system is taken n^2 times the carrying capacity of the individual patches (n^2 is the total number of patches). The average number of mutants at population size M was determined in the same way as in the patch model.

Deterministic (ODE) versions of these models are presented in Section 2 of the Supplement.

Data availability:

This paper uses mathematical models and does not have new data.

Results

Generation and spread of mutants in spatial and non-spatial models

We used a 2-dimensional agent-based model and a patch model (see Materials and Methods) to explore the spread of mutants in spatial and non-spatial growth processes. Denote by L_w and L_m the division rates of wild type and mutant cells, and by D_w and D_m their respective death rates. Below we report the results for neutral, disadvantageous, and advantageous mutants.

Neutral mutants

First we used the 2-dimensional agent-based model under the same assumptions as used in (FUSCO *et al.* 2016), i.e. with neutral mutants and zero death rates ($L_w=L_m>0$, $D_w=D_m=0$). The same type of dynamics is observed as previously reported, with mutant clones either being engulfed by wild-type cells after creation, or mutant clones establishing growing sectors. The average number of mutants at size M is significantly larger for the spatial compared to the non-spatial system (not shown).

Similar results are observed under the assumption that cells can die ($L_w=L_m>0$, $D_w=D_m>0$). Mutants either grow as expanding sectors or are engulfed by the wild-type cells after temporary expansion (Figure 1, inset in (a)). The

number of mutants at population size M is always larger in the spatial compared to the non-spatial system (Figure 1(a)). The extent of the difference is larger for higher death rates and lower reproduction rates, i.e. for populations with a higher turn-over (Figure 1a). This is because in higher turn-over systems, more generations are required to reach a given size threshold, resulting in the amplification of the observed effect.

Results of this comparison were qualitatively similar to those obtained from the patch model. The number of neutral mutants at population size M (assumed much smaller than the maximum system size) was always higher for the spatial (patch) compared to the well-mixed system (Fig S2(A) of the Supplement). Interestingly, this result holds for different spatial organizations of the patch model. In the most spatially restricted system, individuals can only migrate to and from the eight nearest neighboring patches. In an alternative model, migration is allowed between any patches regardless of their location. In either case, a patch model produces significantly more mutants than the mass-action system. This suggests that it is not the spatial arrangement per se but fragmentation of the system that may be responsible for the observed increased number of mutants. The difference is more pronounced for larger cell death rates (Fig S2(A) of the Supplement).

Advantageous mutants

If the mutant is advantageous, the dynamics are similar as those observed for neutral mutants. First, we assume that the advantage is given by a larger division rate of the mutant cells. The number of mutants at population size M is always larger in the spatial compared to the well-mixed setting (Fig 1(b,c)); higher death rates lead to a larger difference between the number of mutants in spatial and non-spatial settings (Fig S1(a) in the Supplement). Mutants can again either grow as expanding sectors, or show a temporary growth phase before being engulfed by wild-type cells, see inset in Fig 1(c). Similar results are obtained if we assume that the mutant advantage is given through a reduced death rate of mutant cells (Figure 1bii and Fig S1(b) in the Supplement). These conclusions remain robust if we use the patch model (either spatially constrained or with random migration between any two patches) instead of the agent-based model (Fig S2(B) of the Supplement). We note that the graphs cover a range of degrees to which the mutant is advantageous, starting from almost neutral up to a 4-fold advantage (perhaps unrealistically high, where the spatial system is completely invaded by mutants). The aim was to show that results do not change for a large parameter range.

Disadvantageous mutants

Disadvantageous mutants are very unlikely to grow as sustained sectors, especially if the disadvantage is more pronounced (inset in Figure 1(d)). In the

absence of death, after creation, mutants undergo a few cell divisions and are then engulfed by the expanding wild-type cell population; in the presence of cell death, they form mutant “islands”, which can become repeatedly generated by mutations and tend to be outcompeted by wild-type cells.

The average number of mutants when the overall population reaches size M depends on spatial structure in a more complex way, compared to the case of neutral mutants. First, we assume that the fitness difference lies in the division rate of the cells (figure 1(d)). In this case, we observe that the average number of mutants is always larger for the spatial compared to the well-mixed simulations. The extent of the difference, however, becomes very small as the extent of the disadvantage grows. Hence, unless the mutant is almost neutral, the increase in the number of mutants in the spatial compared to the non-spatial system becomes negligible. In addition, the difference is most pronounced for small death rates and diminishes for larger death rates (Fig S1(c) of the Supplement).

A different result is observed if the lower fitness of the mutant strain is brought about by a higher death rate of mutant cells. If the difference in death rates lies above a threshold level, the average number of mutants at size M is observed to be larger in well-mixed compared to spatial simulations (Fig 1(e)), which is the opposite trend compared to the previous cases, and also the opposite result compared to those reported in (Fusco *et al.* 2016). Panel (f) shows more details

318 of this behavior. We observe that the number of mutants in a well-mixed system
319 starts exceeding that of the spatial system when the % death rate increase is
320 about 16% for this parameter set (when measured at size 10^3), but this
321 percentage decreases for larger population sizes: it is about 12% when
322 measured at size 10^5, and we expect that the effect is observed for even smaller
323 degrees of disadvantage at lager sizes. This might be relevant for cancer cell
324 populations, where the number of cells can reach 10^{10}-10^{13}. For smaller degrees
325 of disadvantage, even though there are more mutants in the spatial system, the
326 difference is drastically reduced compared to that reported for neutral systems
327 (Fig 1f). This again indicates that a disadvantage in death counters the potential
328 of spatial structure to increase mutant numbers. Lower reproduction rates result
329 in more pronounced differences between the number of mutants in spatial and
330 non-spatial settings (Fig S1(d) of the Supplement). All in all, the effect reported
331 here is manifested for a wide range of disadvantages.
332
333 An analysis of the spatial stochastic model is developed in Section 3 of the
334 Supplement. Using the pair approximation, we derive a formula for the selection-
335 mutation balance of mutants away from the colony boundary (formula (36)). This
336 theory predicts patterns similar to those described above. An intuitive explanation
337 of disadvantageous mutant behavior under decreased reproduction and
338 increased death is provided in the next section, see Fig 3.
339

To confirm the robustness of these results, we performed simulations with disadvantageous mutants in a patch model. Again, the outcome of the dynamics depends on the parameters upon which the disadvantage is based, see Fig 2 for migration to nearest patches and Fig S12 for global migration. If the mutant has a lower division rate than the wild-type, the number of mutants at population size M is larger for the spatial than for the well-mixed scenario. This difference is largest if cells do not die, and diminishes with increasing cell death rates. If, however, the mutant is characterized by a larger death rate than the wild-type, then the opposite result is obtained: the number of mutants at population size M is smaller in the spatial than in the well-mixed system (as long as the difference in the death rate lies above a threshold). Again, the results are qualitatively similar for the spatially restricted (nearest neighbor) and non-restricted (migration to all patches) models (see yellow symbols in the central panels of Fig 2 and Fig S12).

Deterministic (ODE) versions of the patch models are developed and analyzed in Section 2 of the Supplement. In particular, Section 2.5 of the Supplement provides approximate formulas for the numbers of mutants in a metapopulation model and shows for what division and death parameters the number of disadvantageous mutants is higher (lower) in the deterministic metapopulation model compared to the mass-action model. This confirms the above finding that for mutants with lower division rates, more mutants occur in a deterministic metapopulation model, and for mutants with sufficiently high death rates, there are more mutants in mas-action.

Disadvantageous mutants: an intuitive explanation of growth patterns

An intuitive explanation of this phenomenon can be built by observing the growth patterns of disadvantageous mutants in a single patch, starting from a single wild type cell (see Sections 2.5-2.6 of the Supplement for details). Typically, as the total population increases and reaches its carrying capacity, the mean number of mutants is an increasing function of the total population size, and it eventually on average saturates at the selection mutation balance. The number of mutants however does not grow proportional to the total population size, in fact, in some cases the percentage of mutants increases with size, and in others it decreases with size. It turns out that mutants characterized by decreased division rates, which grow relatively slowly at the initial stages, gradually increase in fraction and are most abundant at carrying capacity (Fig 3, blue line panel (a)). On the other hand, mutants with larger death rates grow relatively fast at the initial stages (because they have the same division rates equal to those of the wild types, and initially behave like neutral mutants). As time goes by, however, the larger death rates of mutants start making a difference. The disadvantageous mutants start being ``weeded out" and decrease in percentage down to the selection-mutation balance, when the system reaches carrying capacity (Fig 3, orange line in panel (a)). In other words, if the mutants are characterized by decreased divisions, we expect to observe the largest fraction of mutants when the patch reaches its maximum population; in contrast,

386 if the mutants are characterized by increased deaths, then the percentage of
387 mutants is larger at intermediate stages of growth compared to patches that
388 reach capacity.
389
390 Next, we note that a well-mixed system can be viewed as a superposition
391 of identical, independent smaller patches that all grow simultaneously (figure
392 3(b)). A (proper) patch model is also a collection of patches, but the growth in
393 different patches does not happen simultaneously; instead, it starts in one patch,
394 after a while a second patch starts growing, etc (Fig 3(c)). Therefore, an
395 important difference between the well-mixed system and a patch system is that in
396 the latter model, different patches are desynchronized, such that at a given point
397 in time some patches are completely filled to capacity while others have not
398 started growing yet.
399
400 Keeping this in mind, we can see whether a synchronized (well-mixed) or
401 a desynchronized (patch) model will contain a larger number of mutants. If the
402 mutants have decreased division rates and their percentage grows with total
403 population size, then we are likely to find more mutants in a desynchronized
404 system (spatial or fragmented) that consists of a number of full patches
405 (maximum size, maximum mutant percentage), plus a number of empty patches
406 that do not contribute. In a synchronized (i.e. mixed) system, populations in all
407 patches will lie below carrying capacity at total size M, resulting in fewer mutants.

On the other hand, if the mutants have increased death rates and their percentage is larger at the intermediate stages of population growth, then we expect to have more mutants in a fully synchronized system (i.e. a mixed system), which is equivalent to a set of identical patches that are all relatively early in their growth and thus contain a relatively large percentage of mutants. In the desynchronized (spatial or fragmented) system, populations in several patches will have reached carrying capacity when the total population size reaches M, and thus will have already experienced a decline in mutant percentages.

Before we proceed, we would also like to address the topic of jack-pot mutation events (FUSCO *et al.* 2016). One can think of those events as relatively long lineages of mutants which contribute significantly to the overall expected number of mutants in a growing colony. These lineages are more likely to grow in a spatial system, because a disproportionately large fraction of successful division events happens at the advancing front, thus resulting in the proliferation of individuals that are most separated from the founding individual, and which are more likely to have experienced a mutation. This argument certainly holds for neutral and advantageous mutants, but changes somewhat for disadvantageous mutants. First of all, jack-pot events are less important in the latter case, because mutant clones are unlikely to expand and in any location, mutant levels eventually settle to a low percentage (dictated by selection-mutation balance). Further, if the disadvantage is manifested through an increase in death, relatively

long mutant lineages that are more likely to pop up in spatial and fragmented systems, become progressively diminished by the process of weeding out the mutants, thus making the jack-pot contributions smaller in spatial/fragmented systems compared to the mass-action case. This is explored numerically in section 3.5 of the Supplement.

Growth laws for neutral, advantageous, and disadvantageous mutants in spatial and non-spatial models

Observations presented so far can be generalized by deriving growth laws of mutants in different scenarios, see Table 1 and Section 4 of the Supplement for details.

Consider the type of growth where the population spreads in one direction (examples of such growth can be found in the geometry of colonic crypts (MICHOR *et al.* 2004; LOPEZ-GARCIA *et al.* 2010), or in mitotic zone germ cells in Caenorhabditis elegans (KERSHNER *et al.* 2013)). The mathematical abstraction of this process is the surface of a cylinder, which is a rectangular domain of width W, with the initial cell configuration aligned along one of the boundaries and periodic boundary conditions imposed in the transversal direction. The cell population in this case will engage in a linear growth, such that the mean total population N=W*L grows as N~t. The number of disadvantageous mutants in this

setting will scale with the total population as specified in the first column of Table 1 (2D flat), as these mutants will typically form finite “bubbles” and thus their number will be entirely driven by production. If mutants are neutral, then on average, each newly created mutant will give rise to a clone that grows linearly in time, thus giving a quadratic growth law (uN^2), see Figure 4, curves (a,b). Finally, advantageous mutants, when created, will form expanding clones whose width will grow as the colony proceeds to expand; in other words, advantageous colonies comprise (on average) increasing fractions of the total population size, adding an extra power of N to the growth law (uN^3), see Figure 4, curves (c-g). In the case of a 2D flat front expansion, the number of neutral and advantageous mutants in a colony of a fixed size negatively correlates with the front width: the number of mutants is inversely proportion to the first power of width, W, for neutral, and to the second power of W for advantageous mutants, see Section 4.1 of the Supplement. Note that in the extreme case where W=1, we have a one-dimensional growing array of cells. In this special case (MICHOR *et al.* 2004), in the absence of cell death, all mutants regardless of their fitness properties behave as uN^2.

Next, consider range expansion in 2D (e.g. yeast colony expansion (CHEN *et al.* 2014), 2D melanoma cultures (QIN *et al.* 2016; RODRÍGUEZ *et al.* 2019)), where the population grows outward as an expanding circle. In this case, the total population follows the so-called surface growth: $N \sim t^2$. Mutant cells behave as specified in the second column of Table 1. In particular, disadvantageous mutants are again proportional to the total population; neutral mutations are

479 expected to give rise to colonies whose size does not increase or decrease as a
480 fraction of the total population (the 3/2 law), Figure 5(a,b); advantageous
481 mutations create expanding colonies (the quadratic law), Figure 5(c-f). Note that
482 our theoretical results for the advantageous mutants in a 2D range expansion
483 confirm prior numerical results of (GRALKA AND HALLATSCHEK 2019).
484
485 In a 3D range expansion, which is relevant for example for most solid tumors or
486 biofilms (NADELL *et al.* 2016), the total population engages in a 3D surface growth
487 such that $N \sim t^3$. Mutants are predicted to behave as described in the 4th column
488 of Table 1, and numerical examples confirming the predictions are presented in
489 Figure 6. The 3D flat front expansion is described in the 3rd column of table 1,
490 see also curves (d-f) of Figure 6. The growth of advantageous mutants in a
491 colony with a 3D flat front is characterized by the highest power (the 4th power) of
492 N. Further details are provided in Section 4.2 of the Supplement.
493
494 For comparison, results for non-spatial, exponential growth were derived, for
495 example, by (IWASA *et al.* 2006) and are given in the last column of Table 1. The
496 growth of advantageous mutants in an exponentially expanding population is
497 given by $M(\text{exp, adv}) \sim uN^{(2\alpha-1)/\alpha}$, where α is a parameter that quantifies the mutant
498 advantage ($\alpha=L_m-D_m)/(L_w-D_w)$). Note that as $\alpha \to \infty$, we have at most $M(\text{exp, adv}) \sim uN^2$, and for all finite values of fitness advantage, the power is less than 2.
500

These laws are valid under some restrictions specified in Section 4 of the Supplement. In particular, the laws for advantageous mutants hold for small mutant advantage, and also on the time scales before all the cells in a growing colony's front are replaced by mutants. In the long-term, the replacement of all cells by advantageous mutants is an inevitable outcome in the presence of death, and an approximate outcome as t-> ∞ even in the absence of death. Once this happens, the growth law will be M~N.

The laws of Table 1 are very general and hold in the presence and in the absence of cell death, and also in the presence and in the absence of back mutations (see Materials and Methods). The proportionality coefficients depend on particularities of the underlying dynamics (for example the type of grid used and the number of neighbors, as well as the death to division ratios), but the power laws are universal.

The growth laws derived here have direct consequences for the expected numbers of mutants in equally sized populations growing in different dimensions (and mass-action). The proportion of neutral mutants scales as uN for a flat front (in 2D or 3D), $uN^{1/2}$ for a 2D range expansion, $uN^{1/3}$ in a 3D range expansion, while it is u log N in exponentially growing populations (see also (FUSCO *et al.* 2016)). That is, the number of neutral mutants is always larger in spatial systems compared to the well-mixed system. In space, the proportion of neutral mutants is the largest in low dimensions.

For advantageous mutants, the proportion of mutants is given by uN^3 for a 2D flat front and it is uN^4 for a flat front in 3D, while it is uN^2 for a range expansion (in 2D or 3D); it is $uN^{(\alpha-1)/\alpha}$ in exponentially growing populations. Again, it is the smallest in mass-action.

Finally, for disadvantageous mutants, the power law of mutant growth is the same in all dimensions (and is given by uN). Therefore the results are more subtle and depend on the particular setup. As was shown in the previous section, the behavior depends on whether the disadvantage is manifested through differences in division or death rates.

Discussion

We have used computational models to study mutant evolution in spatially structured and fragmented populations, focusing on the average number of mutants present when the total population size has reached a threshold. Previous work, including (FUSCO *et al.* 2016), established that for neutral mutants, spatial restriction results in a larger number of mutants that are present in a population of a defined size. This was attributed to jackpot events occurring even at relatively large population sizes in spatially structured populations due to

547 the occurrence of range expansion. In contrast, jackpot mutation events can only
548 occur at very early stages of growth in mixed systems. We extended this analysis
549 by considering advantageous and deleterious mutants in greater detail, in the
550 absence and presence of cell death, and assuming that mutant
551 advantage/disadvantage was manifested through either the division or death
552 rate. While the results for advantageous mutants were similar to those for neutral
553 mutants (more mutants in spatial than mixed systems), we found a different trend
554 for disadvantageous mutants. If disadvantage was mediated by an increase of
555 the mutant death rate (rather than a decrease in the division rate), then the
556 difference between the number of mutants in the spatial and well mixed systems
557 becomes dramatically reduced even for very slightly disadvantageous mutants,
558 and as the extent of the disadvantage crosses a threshold, the number of
559 mutants in spatially structured populations becomes *smaller* than in mixed
560 systems.
561
562 The new insights about disadvantageous mutants have important practical
563 implications, for example for understanding the presence of drug resistant
564 mutants prior to the start of treatment in cancers (HORSWELL *et al.* 2013) or
565 bacterial populations that form a biofilm (BANIN *et al.* 2017). According to our
566 results, spatial structure can make it less likely that mutants are present before
567 treatment is started, and if they are present, their average numbers can be lower
568 in spatially structured compared to mixed systems. This requires the
569 disadvantage to be due to a larger death rate and the extent of the disadvantage

570 to lie above a threshold. We have shown that the disadvantage threshold beyond
571 which this effect is observed becomes lower with larger population sizes,
572 indicating that this could be especially relevant for cancer and bacterial
573 populations. Drug resistant mutants have often been shown to be characterized
574 by a fitness cost compared to drug-sensitive cells (GAGNEUX *et al.* 2006; SZAKACS
575 *et al.* 2014), and this is well-documented in the literature for antibiotic resistance
576 in bacteria (ANDERSSON AND HUGHES 2010). The extent of the disadvantage
577 varies depending on the bacterial infection in question, and on the setting in
578 which bacterial growth is measured (ANDERSSON AND HUGHES 2010). In a number
579 of cases substantial fitness costs have been documented for drug resistant
580 bacteria (WICHELHAUS *et al.* 2002; YU *et al.* 2005; NILSSON *et al.* 2006;
581 ANDERSSON AND HUGHES 2010) (in the absence of compensatory mutations), e.g.
582 up to 40% fitness reduction in some rifampin-resistant *Staphylococcus aureus*
583 populations (WICHELHAUS *et al.* 2002), with even larger fitness costs reported in
584 other studies (ANDERSSON AND HUGHES 2010). The fitness cost of resistant
585 mutants is likely to eventually become reduced or eliminated due to the
586 acquisition of compensatory mutations (GAGNEUX *et al.* 2006), but the initial
587 dynamics of mutant evolution before therapy will be significantly determined by
588 the original fitness cost of the mutants. The results reported here about the effect
589 of spatial structure on the evolution of disadvantageous mutations therefore
590 provides valuable information to better understand the emergence of resistant
591 strains before the onset of treatment, and thus for our ability to potentially predict
592 treatment outcomes. This complements other work that has shown an important

593 role of space for the dynamics of drug-resistant cells *during* therapy, through
594 opening up space for the resistant mutants to grow through competitive release
595 (ENDERLING *et al.* 2009; HILLEN *et al.* 2013; FUSCO *et al.* 2016).
596
597 Another interesting result was our finding that qualitatively similar results
598 are obtained if we consider evolutionary dynamics in fragmented rather than
599 spatially structured populations. The same outcomes were obtained for a patch
600 model where individuals in each patch could migrate to any randomly chosen
601 patch in the system (and not just the neighboring patches). Therefore, the
602 properties of mutant evolution described here and also in (FUSCO *et al.* 2016)
603 might not be a particular property of spatial systems, but more generally of
604 fragmented systems. Our intuitive explanation for the differences in the numbers
605 of disadvantageous mutants does not rely on any spatial restrictions in cell
606 migration, but rather on the de-synchronization of mutant dynamics in the
607 patches of a fragmented system.
608
609 In order to understand the intuitive reasons for the observed patterns, it is
610 helpful to consider different ways in which space/fragmentation could influence
611 the dynamics of mutants in expanding populations, see Table 2.
612 (A) How does space/fragmentation affect force of selection? While it is known
613 in the literature that fragmentation may suppress selection (WRIGHT 1931;
614 KOMAROVA 2006; GRALKA AND HALLATSCHEK 2019), this notion has to be
615 applied carefully in each situation. For example, in our setting, we are

comparing an exponentially growing population where no selection at all takes place, to a spatially restricted or fragmented system, where in each patch or spatial location, individuals compete for space, thus leading to a suppression of weaker types and enhancement of stronger types. An exception is found in a patch model in the absence of death, see Section 1.2.2 of Supplement.

(B) How do jack-pot events influence mutant accumulation is space? Since in spatial and fragmented expanding systems, organisms that divide are more likely to be the ones that are further removed from the founding organisms (and thus are more likely to have acquired a mutation), jack-pot events enhance mutant accumulation, although their role is stronger for advantageous than disadvantageous mutants.

(C) How does desynchronization experienced in spatial and fragmented systems affect the accumulation of mutants? This is something that we saw affecting disadvantageous mutants only, because the nature of the affect relies on convergence of the mutants towards selection-mutation balance. Mutants with lower division rates gain in relative abundance, while mutants with higher death rates lose in relative abundance (are weeded out), resulting, respectively, in their enhancement/suppression in spatial and fragmented settings compared to non-structured systems.

Table 2 summarizes the three influences (A-C) listed here. We can see that in the presence of death, there will always be more advantageous mutants in

space (and with no death, there are situations when we can find more in mass-action, Supplemental Section 1.2). For disadvantageous mutants, if they have increased death rates, there could be either more of fewer of them in space compared to well-mixed systems, as reported here. If they have decreased divisions, in the presence of death, this table shows that there could be some variable results, but we have not found a parameter regime with more such mutants in a well-mixed system (although the difference becomes vanishingly small for larger degrees of disadvantage.).

Our analysis of mutant evolution in expanding spatial populations led to a concise derivation of mutant growth laws (see Table 1) in systems of different dimensionalities and under different assumptions on mutant fitness. In particular, our formulas coincide with those previously derived for neutral mutants (FUSCO *et al.* 2016) and confirm numerical predictions for advantageous mutants in 2D range expansions (GRALKA AND HALLATSCHEK 2019).

Our work builds on and complements previous mathematical and computational investigations that explored the differences in mutant dynamics between spatial and non-spatial systems, such as the paper by (FUSCO *et al.* 2016). A variety of other papers dealt with related topics. For example, deleterious mutants were studied in by (LAVRENTOVICH *et al.* 2016) in the context of conversional meltdown, and it was shown that spatial settings enhance the spread and invasion of disadvantageous mutants. A study by (OTWINOWSKI AND KRUG 2014) analyzed the

662 evolutionary dynamics characterized by a large and constant supply of beneficial
663 or deleterious mutations in a one-dimensional spatial habitat, by using the
664 Wright-Fisher (constant population) dynamics. It was found that compared to
665 non-spatial settings, selection is weakened, adaptation is slower, and fitness
666 variation is larger. In paper by (GRALKA *et al.* 2016) it is shown that spatially
667 structured populations with beneficial mutations can give rise to a higher mutant
668 count than well-mixed populations, while in other scenarios (GRALKA AND
669 HALLATSCHEK 2019), fragmentation could reduce selection effects and lead to a
670 lower mutant count. In general, our work fits into the larger literature concerned
671 with spatial mutant evolution (HALLATSCHEK 2018; KAYSER *et al.* 2019; PAULOSE *et*
672 *al.* 2019; PAULOSE AND HALLATSCHEK 2020) and structured populations (FREAN *et*
673 *al.* 2013; HINDERSIN *et al.* 2016; ALLEN *et al.* 2017; GIAIMO *et al.* 2018).
674
675 To conclude, this study has demonstrated complex evolutionary dynamics
676 in populations that are not well-mixed. We demonstrated that evolution can be
677 influenced in different ways by spatial structure or habitat fragmentation,
678 depending on the relative fitness of the mutant and depending on the parameter
679 through which the fitness difference is expressed. These results can also guide
680 future experiments to address some of the computational observations reported
681 here. Experimental results from 2-dimensional spatial growth of cells, such as
682 reported in (FUSCO *et al.* 2016), should be compared to analogous results from a
683 fragmented system, for example where cells are grown in a collection of different
684 wells and periodically transferred to other, randomly chosen wells. This could test

our prediction that the mutant growth in the two scenarios follows similar patterns. On a more complex level, it would be interesting to devise an experimental system where the evolutionary dynamics of disadvantageous mutants is studied, comparing scenarios where the disadvantage is brought about by a difference in cell death versus cell division.

Acknowledgements: This work was funded by NSF grant DMS-1815406.

Competing interests: The authors of this paper report no competing interests.

References

Allen, B., G. Lippner, Y.-T. Chen, B. Fotouhi, N. Momeni *et al.*, 2017 Evolutionary dynamics on any population structure. Nature 544**:** 227-230.

Andersson, D. I., and D. Hughes, 2010 Antibiotic resistance and its cost: is it possible to reverse resistance? Nat Rev Microbiol 8**:** 260-271.

Banin, E., D. Hughes and O. P. Kuipers, 2017 Editorial: Bacterial pathogens, antibiotics and antibiotic resistance. FEMS Microbiol Rev 41**:** 450-452.

Block, M., E. Scholl and D. Drasdo, 2007 Classifying the expansion kinetics and critical surface dynamics of growing cell populations. Phys Rev Lett 99**:** 248101.

Bru, A., S. Albertos, J. Luis Subiza, J. L. Garcia-Asenjo and I. Bru, 2003 The universal dynamics of tumor growth. Biophys J 85**:** 2948-2961.

Bru, A., J. M. Pastor, I. Fernaud, I. Bru, S. Melle *et al.*, 1998 Super-rough dynamics on tumor growth. Physical Review Letters 81**:** 4008.

Chen, L., J. Noorbakhsh, R. M. Adams, J. Samaniego-Evans, G. Agollah *et al.*, 2014 Two-dimensionality of yeast colony expansion accompanied by pattern formation. PLoS Comput Biol 10**:** e1003979.

Dewanji, A., E. G. Luebeck and S. H. Moolgavkar, 2005 A generalized Luria-Delbruck model. Math Biosci 197**:** 140-152.

Enderling, H., A. R. Anderson, M. A. Chaplain, A. Beheshti, L. Hlatky *et al.*, 2009 Paradoxical dependencies of tumor dormancy and progression on basic cell kinetics. Cancer Res 69**:** 8814-8821.

Frean, M., P. B. Rainey and A. Traulsen, 2013 The effect of population structure on the rate of evolution. Proceedings of the Royal Society B: Biological Sciences 280**:** 20130211.

Freyer, J. P., and R. M. Sutherland, 1985 A reduction in the in situ rates of oxygen and glucose consumption of cells in EMT6/Ro spheroids during growth. J Cell Physiol 124**:** 516-524.

Fusco, D., M. Gralka, J. Kayser, A. Anderson and O. Hallatschek, 2016 Excess of mutational jackpot events in expanding populations revealed by spatial Luria-Delbruck experiments. Nat Commun 7**:** 12760.

Gagneux, S., C. D. Long, P. M. Small, T. Van, G. K. Schoolnik *et al.*, 2006 The competitive cost of antibiotic resistance in Mycobacterium tuberculosis. Science 312**:** 1944-1946.

Giaimo, S., J. Arranz and A. Traulsen, 2018 Invasion and effective size of graph-structured populations. PLoS computational biology 14**:** e1006559.

Gillespie, D. T., 1977 Exact Stochastic Simulation of Coupled Chemical-Reactions. Journal of Physical Chemistry 81**:** 2340-2361.

Goldie, J. H., and A. J. Coldman, 1983 A model for resistance of tumor cells to cancer chemotherapeutic agents. Mathematical Biosciences 65**:** 291-307.

Goldie, J. H., and A. J. Coldman, 1998 *Drug resistance in cancer: mechanisms and models*. Cambridge University Press.

Gralka, M., and O. Hallatschek, 2019 Environmental heterogeneity can tip the population genetics of range expansions. Elife 8.

753 Gralka, M., F. Stiewe, F. Farrell, W. Mobius, B. Waclaw *et al.*, 2016 Allele surfing
754 promotes microbial adaptation from standing variation. Ecol Lett 19**:** 889-
755 898.
756 Gunther, S., C. Ruhe, M. G. Derikito, G. Bose, H. Sauer *et al.*, 2007 Polyphenols prevent
757 cell shedding from mouse mammary cancer spheroids and inhibit cancer cell
758 invasion in confrontation cultures derived from embryonic stem cells. Cancer
759 Lett 250**:** 25-35.
760 Hallatschek, O., 2018 Selection-like biases emerge in population models with
761 recurrent jackpot events. Genetics 210**:** 1053-1073.
762 Hillen, T., H. Enderling and P. Hahnfeldt, 2013 The tumor growth paradox and
763 immune system-mediated selection for cancer stem cells. Bull Math Biol 75**:**
764 161-184.
765 Hindersin, L., B. Werner, D. Dingli and A. Traulsen, 2016 Should tissue structure
766 suppress or amplify selection to minimize cancer risk? Biology direct 11**:** 41.
767 Horswell, S., N. Matthews and C. Swanton, 2013 Cancer heterogeneity and "the
768 struggle for existence": diagnostic and analytical challenges. Cancer Lett 340**:**
769 220-226.
770 Iwasa, Y., M. A. Nowak and F. Michor, 2006 Evolution of resistance during clonal
771 expansion. Genetics 172**:** 2557-2566.
772 Johnson, P. J., and B. R. Levin, 2013 Pharmacodynamics, population dynamics, and
773 the evolution of persistence in Staphylococcus aureus. PLoS Genet 9**:**
774 e1003123.
775 Kayser, J., C. F. Schreck, M. Gralka, D. Fusco and O. Hallatschek, 2019 Collective
776 motion conceals fitness differences in crowded cellular populations. Nature
777 ecology & evolution 3**:** 125-134.
778 Kepler, T. B., and M. Oprea, 2001 Improved inference of mutation rates: I. An
779 integral representation for the Luria–Delbrück distribution. Theoretical
780 population biology 59**:** 41-48.
781 Kershner, A., S. L. Crittenden, K. Friend, E. B. Sorensen, D. F. Porter *et al.*, 2013
782 Germline stem cells and their regulation in the nematode Caenorhabditis
783 elegans. Adv Exp Med Biol 786**:** 29-46.
784 Kimura, M., 1962 On the probability of fixation of mutant genes in a population.
785 Genetics 47**:** 713-719.
786 Komarova, N. L., 2006 Spatial stochastic models for cancer initiation and
787 progression. Bull Math Biol 68**:** 1573-1599.
788 Komarova, N. L., and D. Wodarz, 2005 Drug resistance in cancer: principles of
789 emergence and prevention. Proc Natl Acad Sci U S A 102**:** 9714-9719.
790 Komarova, N. L., and D. Wodarz, 2010 ODE models for oncolytic virus dynamics. J
791 Theor Biol 263**:** 530-543.
792 Komarova, N. L., L. Wu and P. Baldi, 2007 The fixed-size Luria–Delbruck model with
793 a nonzero death rate. Mathematical biosciences 210**:** 253-290.
794 Lavrentovich, M. O., M. E. Wahl, D. R. Nelson and A. W. Murray, 2016 Spatially
795 Constrained Growth Enhances Conversional Meltdown. Biophys J 110**:** 2800-
796 2808.
797 Lopez-Garcia, C., A. M. Klein, B. D. Simons and D. J. Winton, 2010 Intestinal stem cell
798 replacement follows a pattern of neutral drift. Science 330**:** 822-825.

Luria, S. E., and M. Delbruck, 1943 Mutations of bacteria from virus sensitivity to virus resistance. Genetics 28: 491-511.
Michor, F., Y. Iwasa, H. Rajagopalan, C. Lengauer and M. A. Nowak, 2004 Linear Model of Colon Cancer Initiation. Cell Cycle 3: 358-362.
Murphy, H., H. Jaafari and H. M. Dobrovolny, 2016 Differences in predictions of ODE models of tumor growth: a cautionary example. BMC Cancer 16: 163.
Nadell, C. D., K. Drescher and K. R. Foster, 2016 Spatial structure, cooperation and competition in biofilms. Nat Rev Microbiol 14: 589-600.
Nilsson, A. I., A. Zorzet, A. Kanth, S. Dahlstrom, O. G. Berg *et al.*, 2006 Reducing the fitness cost of antibiotic resistance by amplification of initiator tRNA genes. Proc Natl Acad Sci U S A 103: 6976-6981.
Otwinowski, J., and J. Krug, 2014 Clonal interference and Muller's ratchet in spatial habitats. Phys Biol 11: 056003.
Patwa, Z., and L. M. Wahl, 2008 The fixation probability of beneficial mutations. J R Soc Interface 5: 1279-1289.
Paulose, J., and O. Hallatschek, 2020 The impact of long-range dispersal on gene surfing. Proceedings of the National Academy of Sciences 117: 7584-7593.
Paulose, J., J. Hermisson and O. Hallatschek, 2019 Spatial soft sweeps: patterns of adaptation in populations with long-range dispersal. PLoS genetics 15: e1007936.
Qin, Y., J. Roszik, C. Chattopadhyay, Y. Hashimoto, C. Liu *et al.*, 2016 Hypoxia-Driven Mechanism of Vemurafenib Resistance in Melanoma. Mol Cancer Ther 15: 2442-2454.
Rodríguez, N. C., J. Lineros, C. S. Rodríguez, L. M. Martínez and J. A. Rodríguez, 2019 Establishment of Two Dimensional (2D) and Three-Dimensional (3D) Melanoma Primary Cultures as a Tool for In Vitro Drug Resistance Studies. , pp. 119-131 in *Immune Checkpoint Blockade* edited by Y. Pico de Coaña. Humana Press, New York, NY.
Rodriguez-Brenes, I. A., N. L. Komarova and D. Wodarz, 2013 Tumor growth dynamics: insights into evolutionary processes. Trends Ecol Evol 28: 597-604.
Szakacs, G., M. D. Hall, M. M. Gottesman, A. Boumendjel, R. Kachadourian *et al.*, 2014 Targeting the Achilles heel of multidrug-resistant cancer by exploiting the fitness cost of resistance. Chem Rev 114: 5753-5774.
Talkington, A., and R. Durrett, 2015 Estimating Tumor Growth Rates In Vivo. Bull Math Biol 77: 1934-1954.
Waclaw, B., I. Bozic, M. E. Pittman, R. H. Hruban, B. Vogelstein *et al.*, 2015 A spatial model predicts that dispersal and cell turnover limit intratumour heterogeneity. Nature 525: 261-264.
Wichelhaus, T. A., B. Boddinghaus, S. Besier, V. Schafer, V. Brade *et al.*, 2002 Biological cost of rifampin resistance from the perspective of Staphylococcus aureus. Antimicrob Agents Chemother 46: 3381-3385.
Wright, S., 1931 Evolution in Mendelian populations. Genetics 16: 97.
Yu, J., J. Wu, K. P. Francis, T. F. Purchio and J. L. Kadurugamuwa, 2005 Monitoring in vivo fitness of rifampicin-resistant Staphylococcus aureus mutants in a mouse biofilm infection model. J Antimicrob Chemother 55: 528-534.

Zheng, Q., 1999 Progress of a half century in the study of the Luria-Delbruck distribution. Math Biosci 162: 1-32.

Tables

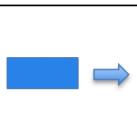
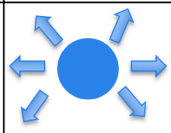
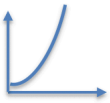
	2D flat	2D range	3D flat	3D range	Exponential
Mutant property					
Disadvantageous	uN	uN	uN	uN	uN
Neutral	uN^2	$uN^{3/2}$	uN^2	$uN^{4/3}$	uN ln N
Advantageous	uN^3	uN^2	uN^4	uN^2	$uN^{(2\alpha-1)/\alpha}$

Table 1. The growth laws of mutants in different spatial and non-spatial growth scenarios, for disadvantageous, neutral, and advantageous scenarios. α is a parameter that quantifies the mutant advantage, $\alpha=(L_m-D_m)/(L_w-D_w)$.

Mutant type	(A)	(B)	(C)
Advantageous, no death	↓	↑	N/A
Advantageous, with death	↑	↑	N/A
Disadvantageous, no death	↑	↑	↑
Disadvantageous by divisions, with death	↓	↑	↑
Disadvantageous by death	↓	↑	↓

Table 2. Summary of the contributions of different mechanisms to mutant accumulation in expanding spatial/fragmeted populations (see text). (A) is the role of fragmentation/spatial restrictions through changing selections strength, (B) is jack-pot events, and (C) is the synchronization phenomenon. Here ↓ means "suppresses mutants in space, compared to exponential growth in well-mixed systems" and ↑ means "enhances mutants in space, compared to exponential growth in well-mixed systems".

Figure legends

Fig 1. Comparison of the number of mutants in 2D spatial agent-based model simulations (red) and a well-mixed system (black). In (a-e), the lines represent the mean numbers of mutants in the spatial and non-spatial systems at equal size, $N=10^3$. (a) Neutral mutants, as a function of the death rate. (b-c)

878 Advantageous mutants, characterized by increased division rates (b) and
879 decreased death rates (c), as a function of the fold-advantage. (d-e)
880 Disadvantageous mutants characterized by decreased division rates (d) and
881 increased death rates (e), as a function of the fold-disadvantage. Typical 2D
882 spatial agent-based simulations of range expansion dynamics are shown in the
883 insets for each mutant type. (f) The ratio of the mean number of mutants in the
884 2D spatial simulations and that for the well-mixed system, for disadvantageous
885 mutants with increased death rates, is shown as a function of fold-disadvantage.
886 The two lines correspond to population sizes of 10^3 (blue) and 10^5 (yellow). The
887 parameters for panels (a-e) (unless otherwise indicated in figure) are: L=0.2,
888 D=0.1, $u=2\times10^{-3}$. For each parameter combination, from 2×10^6 to 3×10^7 repeats
889 were performed; shown are the means; standard errors are too small to see. For
890 panel (f), $L_w=L_m=0.09$, $D_w=0.05$, $u=2\times10^{-3}$. Standard errors are represented by
891 vertical bars and are not visible.
892
893 **Fig. 2.** A systematic study of the number of disadvantageous mutants in mass-
894 action (blue bars in histograms) and in patch model (yellow bars). Data are
895 presented for 8 parameter combinations: 1-3 with mutants of decreased division
896 rates, 4 (the green point) with neutral mutants, and 5-8 with mutants of increased
897 death rate. We observe that the mean number of mutants in the patch models
898 with nearest neighbor migrations (depicted by yellow symbols in the central
899 panel) becomes smaller than that in the mass-action model (blue symbols) if the
900 disadvantage through death is sufficiently large. The numerical probability

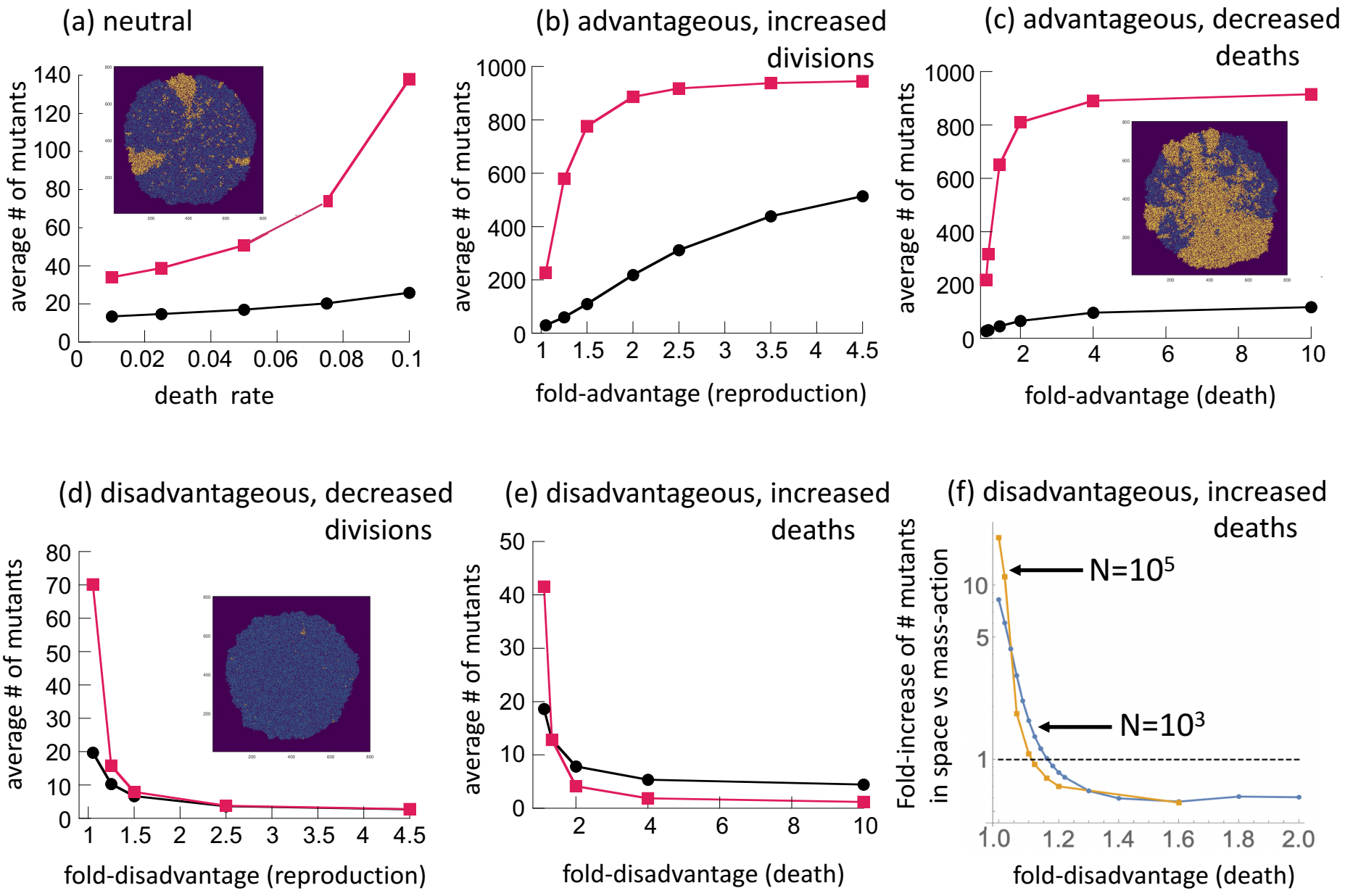
distributions for the numbers of mutants are also presented for well-mixed and patch (with nearest neighbor migration) models; about $2x10^5$ simulations were used for each parameter combination. Please note the logarithmic scale of the histograms. The rest of the parameters are as follows: $u=10^{-3}$, $m=10^{-5}$, k=100, 100 patches; the number of mutants evaluated at total population 10^3. See Supplementary figure S14 for the patch model with global migrations; the results are very similar.

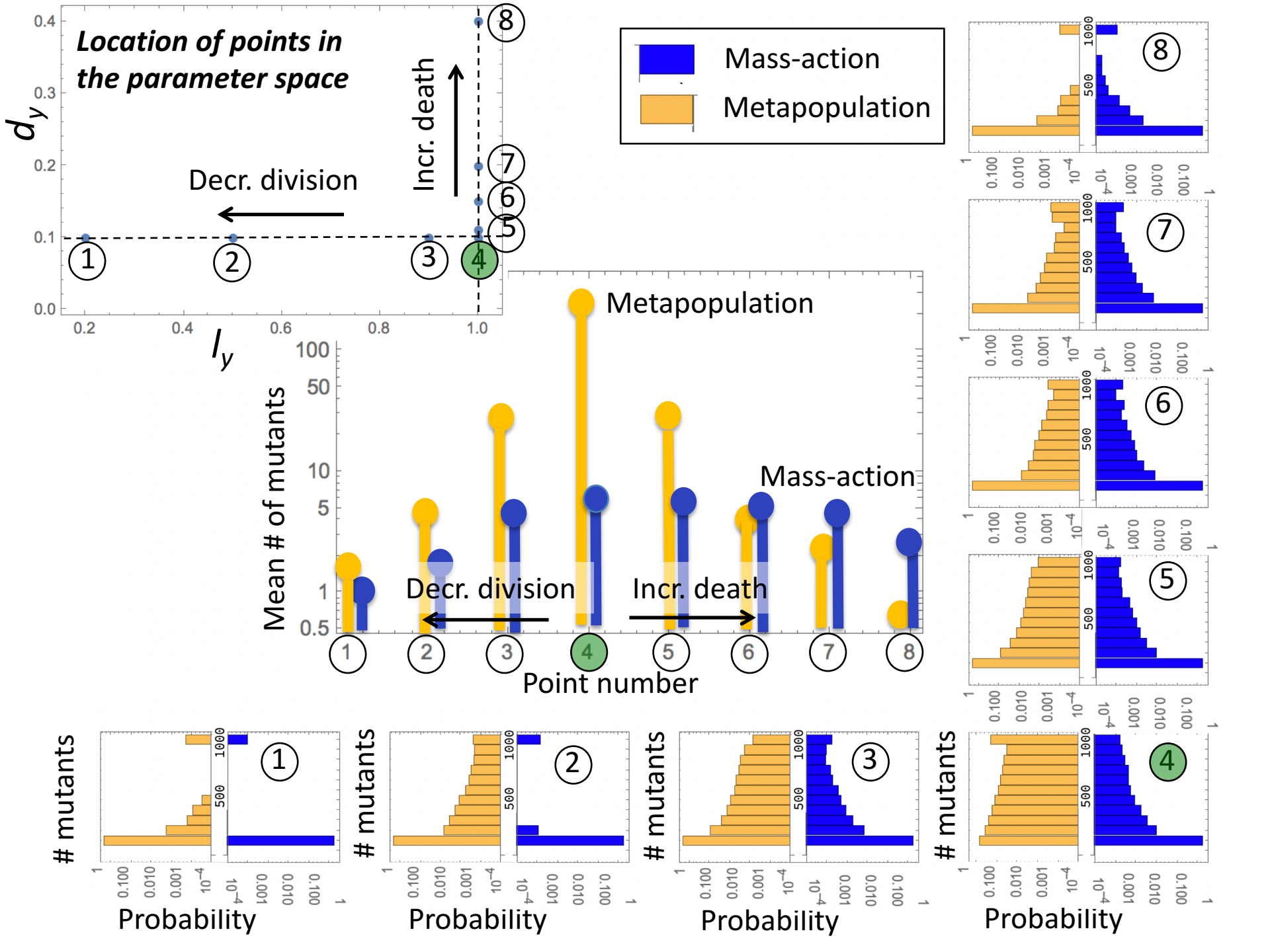
Fig 3. Disadvantageous mutants – an intuitive picture. (a) The fraction of mutants (characterized by increased death and by decreased divisions) as a function of the population size (ODEs, parameters as in Fig S5). Inset: the time-series for the mutant populations with increased (orange) and decreased (blue) death rates, together with the wild type population (scaled by 1000 to fit in the same graph). (b) A schematic representing total population time-series in different patches in a patch model. (c) Same for the well-mixed model represented as a collection of identical, synchronous patches. At the same total population size, in a patch model some populations are at carrying capacity, and some are zero, while in the well-mixed model, all the ``patches" are partially filled. Patches with populations below carrying capacity have more mutants than patches at carrying capacity, if the mutants are characterized by increased death. Patches with populations below carrying capacity have fewer mutants than patches at carrying capacity, if the mutants are characterized by decreased divisions.

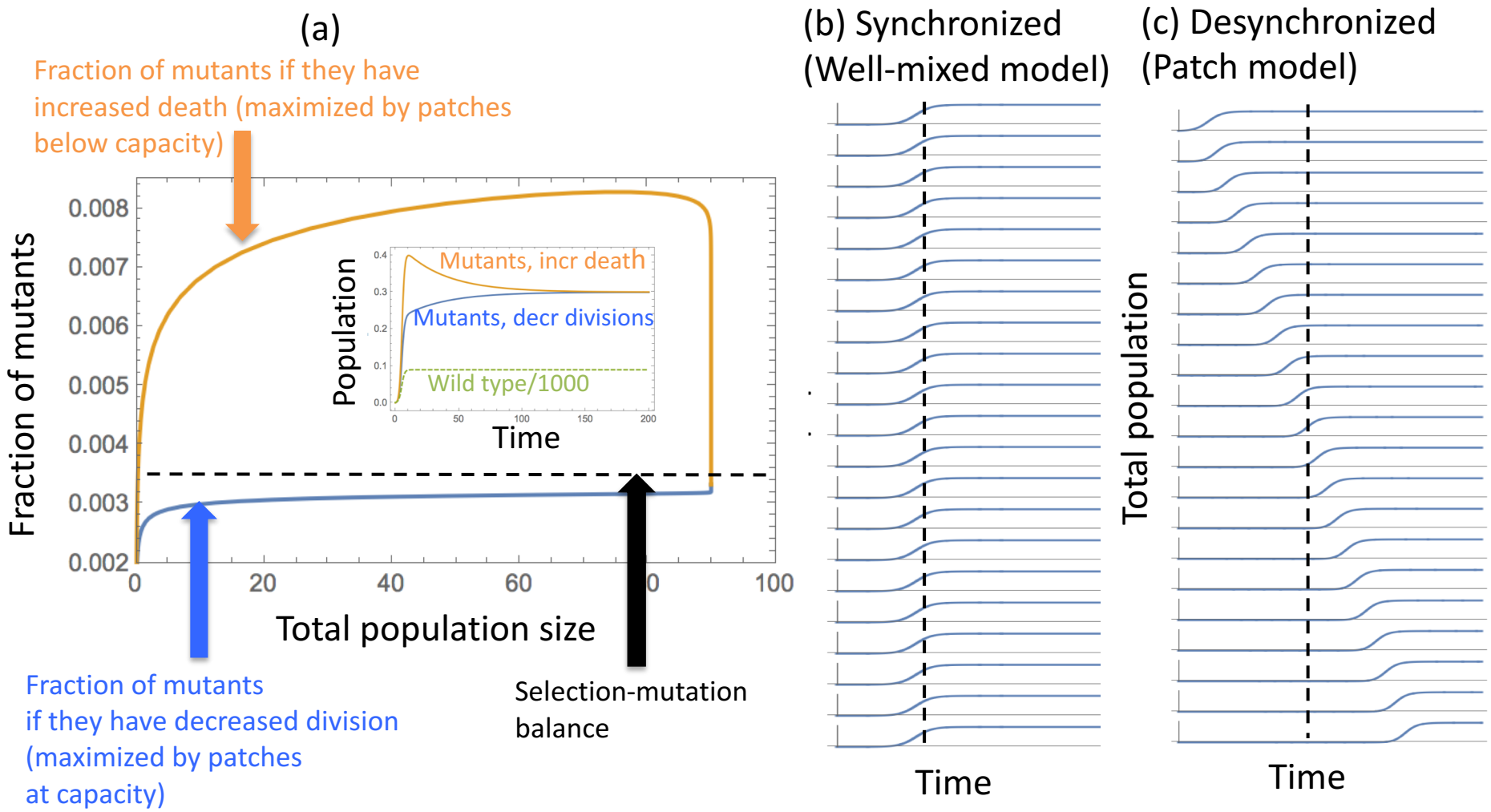
Fig. 4. Mutants in a colony with a 2D flat front: the number of mutants as a function of the total population, averaged over 1000 stochastic runs (standard errors too small to see). Cases (a,b) are neutral, and the corresponding solid black lines are guides to the eye with slope 2 in the log-log plot. Cases (c-g) are advantageous, and the dashed lines are guides to the eye with slope 3. The following parameters are used: (a) Neutral, no death: $L_w=L_m=0.7$, $D_w=D_m=0$. (b) Neutral, with death: $L_w=L_m=0.7$, $D_w=D_m=0.2$. (c) Advantageous, no death: $L_w=0.7$, $L_m=0.9$, $D_w=D_m=0$. (d) Advantageous, no death, larger advantage: $L_w=0.7$, $L_m=1.0$, $D_w=D_m=0$. (e) Advantageous by division, with death: $L_w=0.7$, $L_m=0.8$, $D_w=D_m=0.2$. (f) Advantageous by death: $L_w=L_m=0.7$, $D_w=0.2$, $D_m=0.1$. (g) Same as (f), but with a wider front: W=1000. The rest of the parameters are $u=5\text{x}10^{-5}$, W=100 (except (g)).

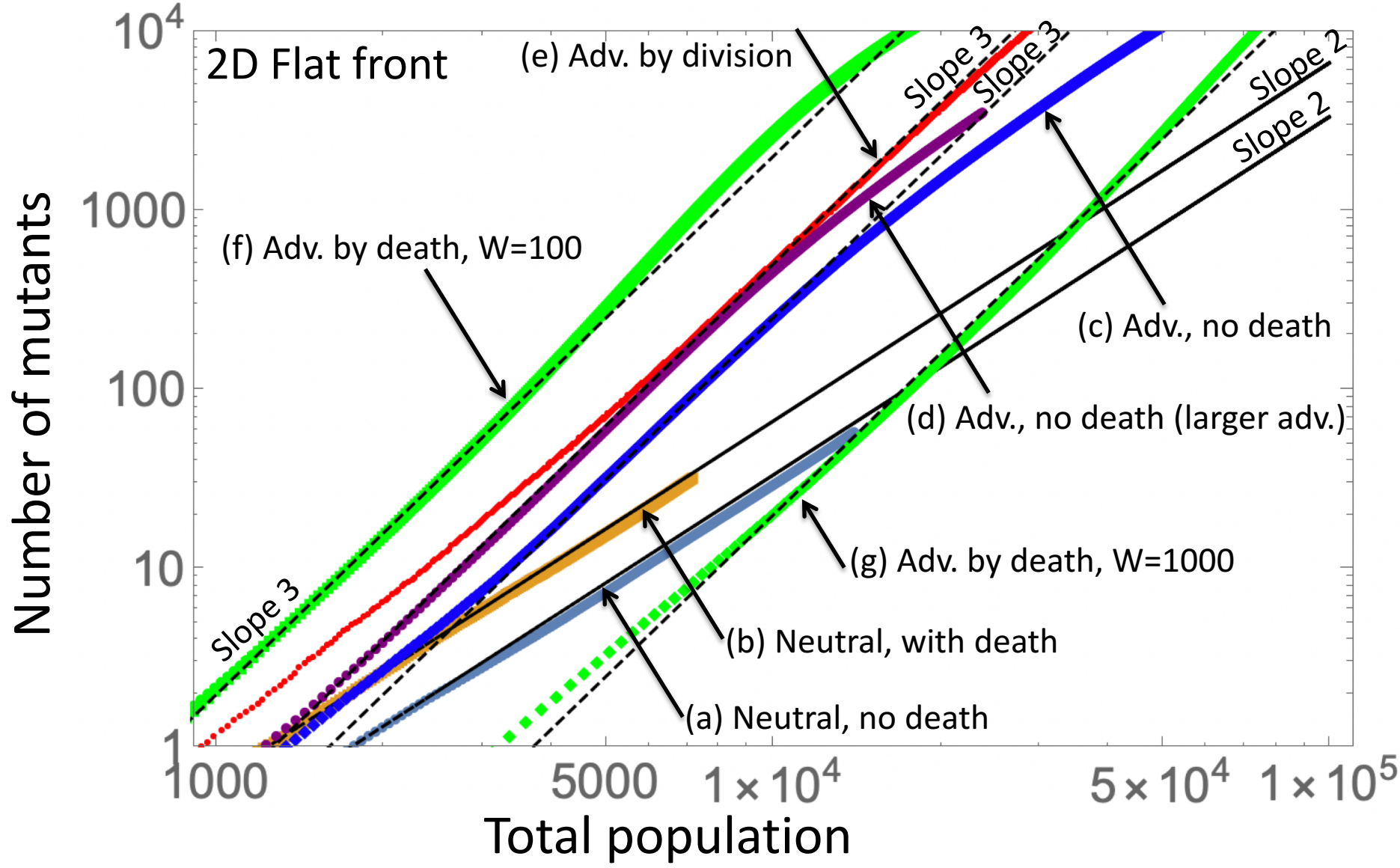
Fig. 5. Mutants in the 2D range expansion: the average number of mutants as a function of the total population (standard errors too small to see). Cases (a,b) are neutral, and the corresponding solid black lines are guides to the eye with slope 3/2 in the log-log plot. Cases (c-f) are advantageous, and the dashed lines are guides to the eye with slope 2. The following parameters are used: (a) Neutral, no death: $L_w=L_m=0.7$, $D_w=D_m=0$, 2000 runs. (b) Neutral, with death: $L_w=L_m=0.7$, $D_w=D_m=0.2$, 1366 runs. (c) Advantageous, no death: $L_w=0.7$, $L_m=0.9$, $D_w=D_m=0$, 2000 runs. (d) Advantageous, no death, larger advantage: $L_w=0.7$, $L_m=1.0$, $D_w=D_m=0$, 2000 runs. (e) Advantageous by division, with death: $L_w=0.7$,

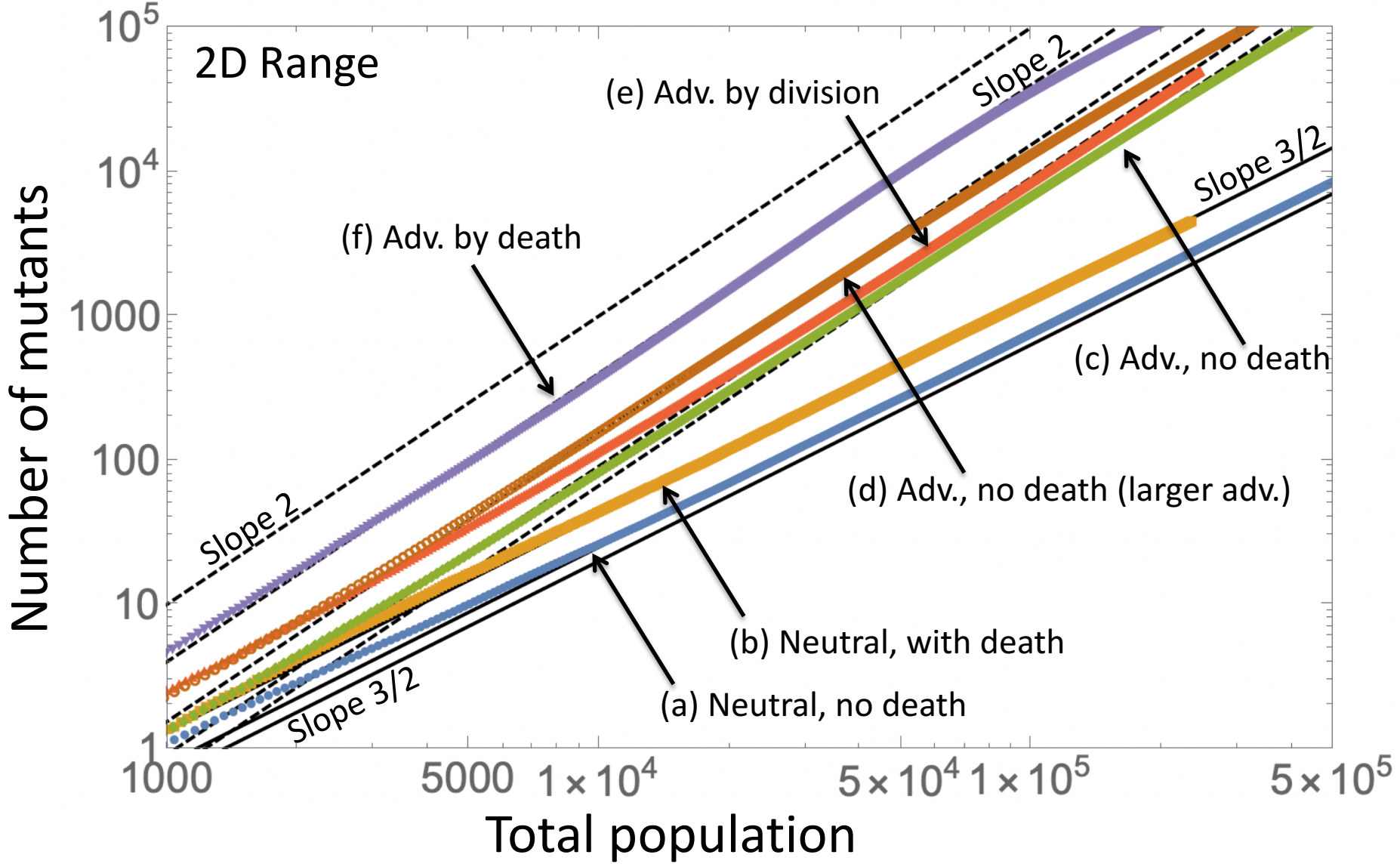
946 L_m=0.8, D_w=D_m=0.2, 1426 runs. (f) Advantageous by death: L_w=L_m=0.7, D_w=0.2,
947 D_m=0.1, 1396 runs. The mutation rate is u=5x10^{-5}.
948
949 **Fig. 6.** Mutants in the 3D expansion: the average number of mutants is plotted as
950 a function of the total population. (a) Neutral mutants in a range expansion, with
951 the corresponding dotted gray guide to the eye with slope 4/3 in the log-log plot.
952 (b,c) Advantageous mutants in a range expansion, and the solid lines are guides
953 to the eye with slope 2. (d) Neutral mutants in a 3D flat front expansion, and the
954 solid guide to the eye has slope 2. (e,f) Advantageous mutants in a colony with a
955 3D flat expansion, the dashed guides to the eye have slope 4. The following
956 parameters are used: (a) Neutral, range: L_w=L_m=0.7, D_w=D_m=0.1, u=2x10^{-5}, 10^7
957 runs. (b) Advantageous by division, range: yellow L_w=0.4, L_m=0.8, D_w=D_m=0.1,
958 u=2x10^{-5}, 4x10^6 runs; red same but u=2x10^{-7}, 66,631 runs. (c) Advantageous by
959 death, range. (d) Neutral, flat: L_w=L_m=0.8, D_w=D_m=0.1, u=2x10^{-7}, 34,967 runs. (e)
960 Advantageous by division, flat: L_w=0.4, L_m=0.8, D_w=D_m=0.1, u=2x10^{-7}, 11,369
961 runs. (f) Advantageous by death, flat: L_w=L_m=0.7, D_w=0.2, D_m=0.1, u=2x10^{-7},
962 53,840 runs.

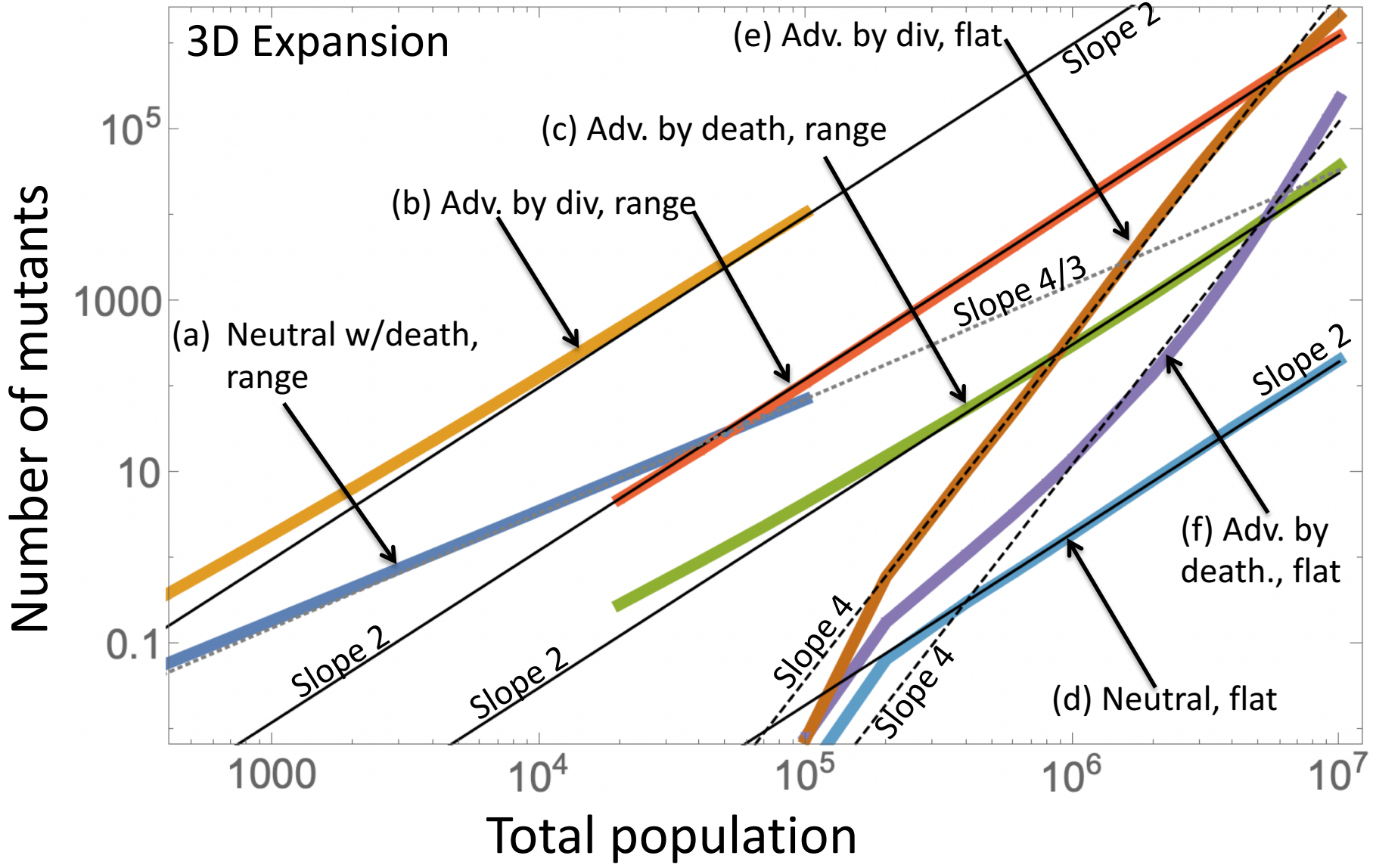

(a) neutral
average # of mutants
death rate
(b) advantageous, increased divisions
average # of mutants
fold-advantage (reproduction)
(c) advantageous, decreased deaths
average # of mutants
fold-advantage (death)
(d) disadvantageous, decreased divisions
average # of mutants
fold-disadvantage (reproduction)
(e) disadvantageous, increased deaths
average # of mutants
fold-disadvantage (death)
(f) disadvantageous, increased deaths
Fold-increase of # mutants in space vs mass-action
N=10^5
N=10^3
fold-disadvantage (death)


Location of points in the parameter space
Decr. division
Incr. death
d_y
I_y
Mass-action
Metapopulation
Mean # of mutants
Point number
Decr. division
Incr. death
mutants
Probability


(a)
Fraction of mutants if they have increased death (maximized by patches below capacity)
Fraction of mutants
0.008
0.007
0.006
0.005
0.004
0.003
0.002
0
20
40
60
0
100
Total population size
Population
Mutants, incr death
Mutants, decr divisions
Wild type/1000
Time
Fraction of mutants if they have decreased division (maximized by patches at capacity)
Selection-mutation balance
(b) Synchronized (Well-mixed model)
Time
(c) Desynchronized (Patch model)
Total population
Time


2D Flat front
(e) Adv. by division
Slope 3
Slope 3
Slope 2
Slope 2
(f) Adv. by death, W=100
(c) Adv., no death
(d) Adv., no death (larger adv.)
(g) Adv. by death, W=1000
(b) Neutral, with death
(a) Neutral, no death
Slope 3
Number of mutants
Total population
10^4
1000
100
10
1
1000
5000
1×10^4
5×10^4
1×10^5


2D Range
Number of mutants
Total population
10^5
10^4
1000
100
10
1
1000
5000
1×10^4
5×10^4
1×10^5
5×10^5
(e) Adv. by division
Slope 2
(f) Adv. by death
Slope 2
Slope 3/2
(c) Adv., no death
(d) Adv., no death (larger adv.)
(b) Neutral, with death
(a) Neutral, no death
Slope 3/2


3D Expansion
(e) Adv. by div, flat
Slope 2
(c) Adv. by death, range
(b) Adv. by div, range
Slope 4/3
(a) Neutral w/death, range
Slope 2
(f) Adv. by death., flat
Slope 2
Slope 2
Slope 4
Slope 4
(d) Neutral, flat
Number of mutants
10^5
1000
10
0.1
1000
10^4
10^5
10^6
10^7
Total population

Mutant evolution in spatially structured and fragmented populations

Supplementary information

Contents

1 Simulations: additional results

1.1 Agent-based model simulations

In the main text we report on the abundances of advantageous, neutral, and disadvantageous mutants in spatial and well-mixed system. Some further results are presented in figure S1. In particular, panels (a) and (b) study advantageous mutants. In (a), the advantage is manifested through an increased division rate, and in (b) through a decreased death rate of mutants. Varied is the rate that is unaffected by the mutation under consideration, which is the death rate in (a) and the reproduction rate in (b). As expected, we observe that the number of advantageous mutants is higher in a spatial system (red) compared to the well-mixed system (black) for both types of mutants. The difference increases with death rate (see panel (a)) and decreases with the reproduction rate (panel (b)); in other words, the difference between well-mixed and spatial models is larger for cells with an overall slower expansion rate.

Next, we turn to disadvantageous mutants. In figure S1(c) we study mutants characterized by a lower reproduction rate, and in figure S1(d) the mutants' death rate is higher compared to that for wild type cells. Again, in both panels, the rate unaffected by mutations is varied, and the mutant abundance in spatial (red) and well-mixed (black) system compared at the same population size. As reported in the main text, we observe that mutants with a lower reproduction rate are more abundant in a spatial model. As seen in panel (c), the differences becomes smaller with an increased death rate of cells.

In panel (d) of figure S1 we turn to mutants characterized by a larger death rates. It is reported in the main text that if the disadvantage is sufficiently pronounced, we expect to find more such mutants in a well-mixed system compared to a spatial system. This is what we see in figure S1(d), when the reproduction rates are lower than a threshold. This trend reverses, however, when the reproduction rates become higher. This provides further information about the phenomenon reported in the main text. The mutant disadvantage must be sufficiently high, for the well-mixed system to accumulate more mutants than the spatial system, and this advantage is measured against the background reproduction rate of the cells. As the reproduction rate gets higher, the difference between mutant and wild type death rates must also become higher, to observe more mutants in a well-mixed system compared to the spatial system.

1.2 Patch model simulations

1.2.1 Dependence on reproduction and death rate

We have run numerical simulations of the stochastic patch model to determine whether spatial arrangement of patches makes a difference for the abundance of mutants at a given system size. Figure S2 shows the results for neutral (a) and advantageous (b) mutants, while figure 2 of the main text contains information about disadvantageous mutants. In figure S2, two types of the patch model are compared. In the spatial 2D patch model (dark green bars), we used a 2D array

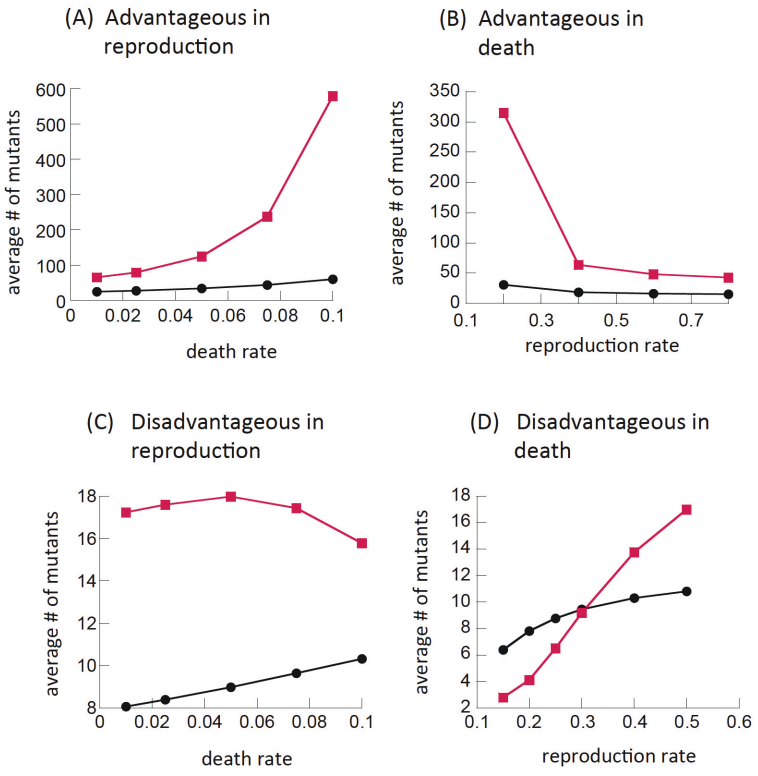


Figure S1: The abundance of mutants in spatial 2D simulations (red) and in a well-mixed system (black), as a function of parameters. (a) Mutants have a larger reproduction rate ($L_w = 0.2, L_m = 0.25$); varied is the death rate (equal for mutants and wild types). (b) Mutants have a smaller death rate ($D_w = 0.1, D_m = 0.09$); varied is the reproduction rate (equal for mutants and wild types). (c) Mutants have a smaller reproduction rate ($L_w = 0.25, L_m = 0.2$); varied is the death rate (equal for mutants and wild types). (d) Mutants have a larger death rate ($D_w = 0.05, D_m = 0.1$); varied is the reproduction rate (equal for mutants and wild types). The rest of the parameters are as in figure 1(a-e) of the main text.

of patches, where cells could migrate only between neighboring patches (with each patch having 8 neighbors as in the Moore model). In the "fragmentation model" (light green bars) migration could happen between any patches regardless of their position. In the latter case, the model cannot be regarded as spatial per se. Nonetheless, similar trends were observed in both models. For advantageous and neutral mutations, more mutants were observed in the patch model (a 2D or a fragmented, non-spatial patch system) compared to the well-mixed system with the same total number of cells; the effect is less pronounced but

still clearly present in the fragmented system.

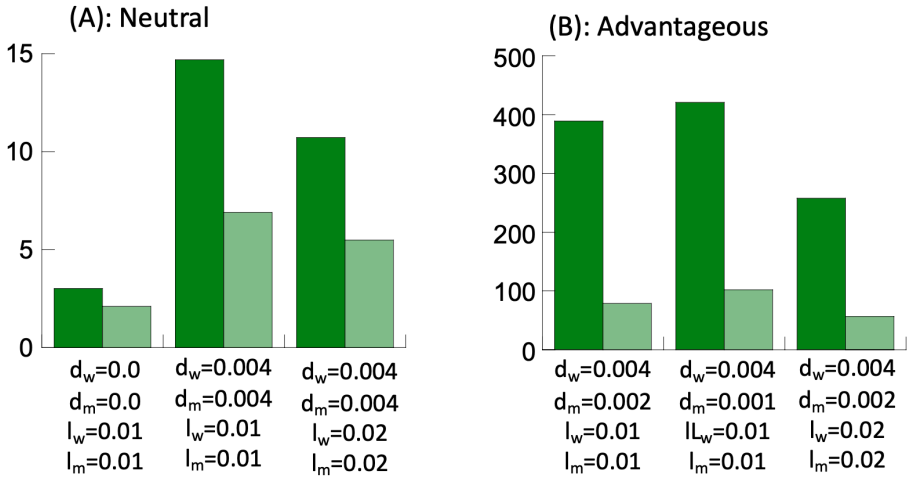


Figure S2: Comparison of the number of mutants in stochastic patch model simulations and a well-mixed system. The bars represent the ratio between the mean number of mutants in the patch model and the mean number of mutants in well-mixed systems at equal size, $N = 10^4$. Dark green bars correspond to the 2D patch model with 8 nearest neighbor migration. Light green bars correspond to the fragmented model where migration happens to all patches. (a) Neutral mutants. (b) Advantageous mutants. Between 10^6 and 4×10^7 runs were performed for each bar. Division and death rates are indicated below each bar. Other parameters are $u = 2 \times 10^{-5}, k = 100, m = 10^{-5}, n = 31 \times 31$ patches.

1.2.2 The case of very small patch size

Next, we explore the behavior of the patch model in the regime of very small patches. It is known that fragmentation/spatial restrictions weaken selection. In the extreme case of a patch model with very small patch size, this effect can significantly influence mutant dynamics and even reverse the results for advantageous mutants. Figure S3 explores this regime.

We observe that in the absence of death, under the patch size of $k = 2$ and slightly advantageous mutants (figure S3(a)), there are more mutants in the mass-action compared to the patch model. This reverses when we increase the path size to $k = 3$. For a larger mutant advantage (panel (b)), it takes $k = 5$ to observe more mutants in the patch model. In the extreme case where mutants enjoy a 10-fold fitness advantage, even for the patch size $k = 100$ we still see more mutants in the mass-action system. This effect disappears if we

include death in the system. For example, panel (d) shows how results of panel (c) change if we increase death rate to a modest $d_w = d_m = 0.01$: in the case, there are significantly more mutants in the patch model. Panel (d) shows the effect of increasing the death rate, where the number of mutants in the patch system with small patch size ($k = 3$) becomes significantly larger than that i the mass-action model even for small (but non-zero) death rates.

The reason for the reversal of the pattern under very small patch size in the absence of death is the complete selection suppression experienced in such system. Increasing patch sizes or death rate allows selection to work, leading to more mutants in the patch model compared to mass action.

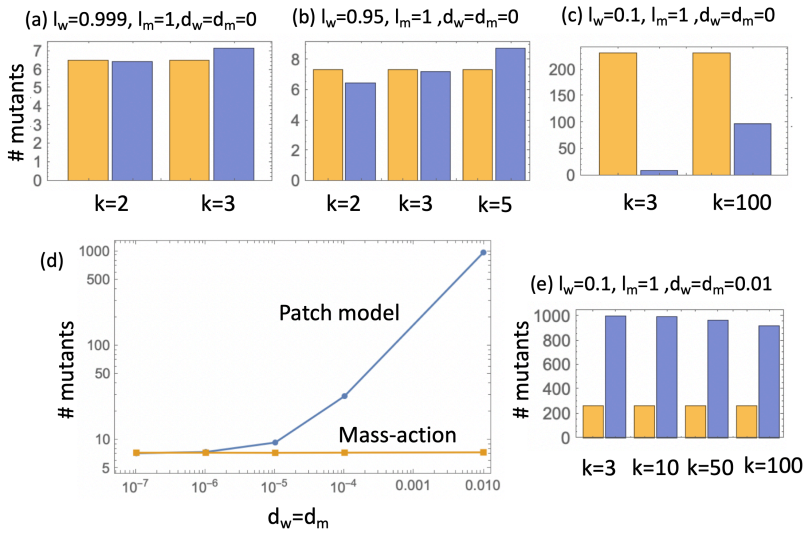


Figure S3: Behavior of the patch model with very small patches. The number of mutants in the patch model (blue bars) and mass-action model (yellow bars) corresponding to the same total population size. Different bars correspond to different patch size, k. (a-c) No death, (e) in the presence of death. (d) The numbers of mutants in the patch model and mass action model as functions of the death rate. Simulation parameters are market at the top of each bar graph. Other parameters are $N = 10^3, u = 10^{-3}, m = 10^{-5}$, global migration in the patch model. For panel (d), $l_w = 0.95, l_m = 1, k = 3$.

1.2.3 Mutant growth curves

Here we present growth curves for the number of advantageous and neutral mutants in the patch model, figure S4. These were obtained by running a large number of simulations and recording the number of mutants at different population sizes. Mean values are presented, and the standard errors are small and not visible. The blue symbols correspond to runs with advantageous mutants and

orange symbols to neutral mutants. For comparison, we also present curves for the number of advantageous and neutral mutants in well-mixed systems, taken far from carrying capacity. For non-neutral mutants this is given by formula (20), and for neutral mutant this simplifies to $l_w u N \ln N/(l_w - d_w)$.

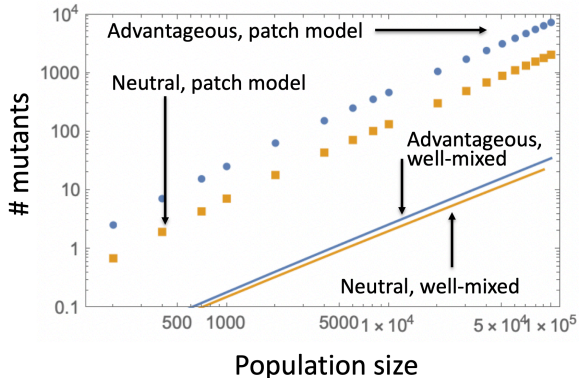


Figure S4: The number of mutants in a patch model as a function of the total population size. Blue circles, advantageous mutants ($l_m = 1.05$), the number of runs for each point is $27,061$. Orange squares, neutral mutants ($l_m = 1$), the number of runs for each point is $61,083$. Blue and orange lines represent advantageous and neutral mutants in a well-mixed system (theoretical, formula (20)). Other parameters are $l_w = 1, d_w = d_m = 0.1, m = 10^{-5}, k = 100, u = 2 \times 10^{-5}$; in the patch model, global migration was implemented.

Fitting a straight line through the data on a log-log scale, we obtain the following powers: advantageous mutants increase as $N^{1.2}$, and $N^{1.19}$. For comparison, a similar procedure yielded powers $N^{1.12}$ and $N^{1.1}$ for the well-mixed system (we did not attempt to distinguish between a power law with as power close to one and an $N \ln N$, which is the "true" growth law in the latter case).

2 Disadvantageous mutants: a deterministic metapopulation model

2.1 Basic formulations and selection mutation balance

Let us denote the wild type population as $x(t)$ and the mutant population as $y(t)$. Denote the rate of mutations as u, the division and death rates of wild type cells as l_w and d_w, and the division and death rates of mutants as l_m and d_m. Then the competition dynamics of cells can be formulated as follows:

$$\dot{x} = l_w x(1-u)\left(1 - \frac{x+y}{\mathcal{K}}\right) - d_w x, \tag{1}$$

$$\dot{y} = (l_w x u + l_m y)\left(1 - \frac{x+y}{\mathcal{K}}\right) - d_m y, \tag{2}$$

where $\mathcal{K}$ is the carrying capacity. If the mutants are disadvantageous, that is, if

$$\nu = \frac{d_m}{d_w} - \frac{l_m}{l_w} > 0, \tag{3}$$

then the selection mutation balance predicts that the equilibrium number of wild type cells is given by

$$\bar{x} = \mathcal{K}\left(1 - \frac{d_w}{l_w}\right), \tag{4}$$

and the number mutants is

$$\bar{y} = \bar{x} u \left(\frac{d_m}{d_w} - \frac{l_m}{l_w}\right)^{-1} \tag{5}$$

(we only took the largest contributions in terms of small u).

In model (1-2), as the population approaches the carrying capacity, the divisions slow down, while deaths remain occurring at a constant rate. We will refer to this model as *division-controlled growth*, which follows the terminology of [**?**]. Alternatively, we can assume that as the population grows, the death rate increases, while the division rate stays constant. In the mass action case, this can be modeled as follows:

$$\dot{x} = l_w x(1-u) - d_w x\left(1 + \frac{x+y}{\mathcal{K}}\right), \tag{6}$$

$$\dot{y} = (l_w x u + l_m y) - d_m y\left(1 + \frac{x+y}{\mathcal{K}}\right). \tag{7}$$

We will refer to this model as *death-controlled growth* [**?**]. In this case, the equilibrium population size is

$$\bar{x} = \mathcal{K}\left(\frac{l_w}{d_w} - 1\right), \tag{8}$$

and the number of mutants in selection mutation balance is again given by formula (5).

The early dynamics of wild type and mutant populations in models (1-2) and (6-7) are identical. Assuming that $x + y \ll \mathcal{K}$, we can solve the resulting linear equations exactly, to obtain

$$x_{lin}(t) = e^{(l_w - d_w)t}, \quad y_{lin}(t) = l_w u \frac{e^{(l_w - d_w)t} - e^{(l_m - d_m)t}}{(l_w - d_w) - (l_m - d_m)}. \tag{9}$$

2.2 Decreased divisions and increased death

A mutant is disadvantageous if inequality (3) holds. If we fix l_w and d_w, the division and the death rates of the wild type cells, this inequality defines a half

plane in the (l_m, d_m) space, where mutants are disadvantageous (more precisely, it is the region above the line $d_m/l_m = d_w/l_w$, see figure S8, the red line). Note that this definition is not equivalent to using the linear growth rate to define fitness, because instead of linear initial expansion of the mutants, it measures their steady state level in the presence of the wild types. For comparison, the line where the linear growth rates of mutant and wild-type cells are equal to each other, is shown in figure S8 in green.

The quantity ν defined in (3) can be seen as a measure of fitness disadvantage. Mutants with equal ν have the same level of disadvantage, which is for the purposes of this paper defined as the same level of the selection mutation balance equilibrium. The definition of ν (equation (3)) with fixed l_w, d_w generates a one-parametric family of types of equal disadvantage. Disadvantage can be achieved by different combinations of division and death rates, l_m, d_m. In particular, if the mutants have the same death rates as the wild types, and their disadvantage is achieved through lowered division rates, then we have a type with **decreased divisions**,

$$l_m = l_w(1-\nu), \quad d_m = d_w; \tag{10}$$

in figure S8, "decreased division" mutants correspond to all the points on the horizontal dashed line (in the disadvantageous region). If on the other hand, the division rate of mutants matches that of the wild types, we have a type with **increased death**:

$$l_m = l_w, \quad d_m = d_w(1+\nu); \tag{11}$$

in figure S8, "decreased death" types correspond to disadvantageous points on the vertical dashed line. Note that if the two types have the same fitness (i.e. converge to the same selection-mutation balance), the percentage decrease in the division rate must be equal to the percentage increase in the death rate. The linear growth rate of the two types is however different, and is given by

$$l_m - d_m = \begin{cases} l_w - d_w - \nu l_w, & \text{decreased divisions,} \\ l_w - d_w - \nu d_w, & \text{increased death.} \end{cases}$$

We can see that since $l_w > d_w$, the "increased death" mutants are always characterized by a faster growth compared to the "decreased divisions" mutants.

Figure S5(a) shows an example of the numbers of mutants under the assumptions of decreased divisions (blue) and increased deaths (yellow); specific parameter values are given in the figure caption. The behavior of the wild types can be seen from the dashed green line that shows $x(t)/1000$. Before the population reaches the carrying capacity, the "increased death" mutants grow faster than the "decreased divisions" mutants. At later times, they both reach the same selection mutation equilibrium.

A useful representation of this information is given in panel (b) of figure S5, where we plot the number of mutants, $\mathcal{M}(x)$, contained in the system of size x. This quantity is presented for both types (decreased divisions and increased death) in figure S5(b).

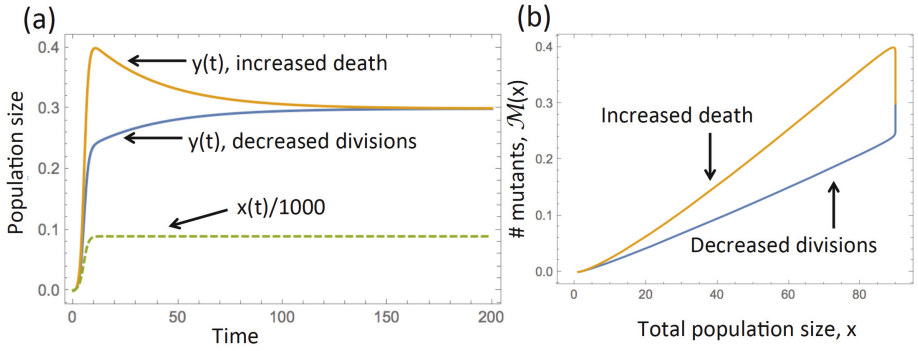


Figure S5: Solutions of system (1-2), where the division and death rates of the wild types are fixed, and the mutant rates are given by equation (10) for the "decreased divisions" type, and equation (11) for the 'increased death" type (only one type of mutant is included at a time). (a) The level of mutants as a function of time for the two systems; the wild types are given by the green dashed line, where the value was divided by 1000 to bring it to the same scale. (b) The number of mutants plotted as a function of the total cell population for the two types. The parameters are $l_w = 1; d_w = 0.1, \nu = 0.3, u = 0.001, K = 100$. The fitness disadvantage is $\nu = 0.3$.

2.3 The patch model

Consider N patches, such that in each patch cells undergo deterministic mass-action dynamics of divisions and deaths, subject to a carrying capacity K. We assume that patches communicate with each other through migrations. Let us denote the migration matrix as $M = \{M_{ij}\}$, where M_{ij} is the probability that, given that a cell from patch i migrates, it is transferred to patch j; we have $\sum_{j \neq i} M_{ij} = 1$ for all i. We explore two types of such matrices:

1. **A 1D spatial model:** a ring of patches where only migration between nearest neighbors is possible:

$$M_{ij} = \begin{cases} 1/2, & |i-j| = 1, \\ 0, & \text{otherwise}, \end{cases} \tag{12}$$

   with the additional nonzero values $M_{1N} = M_{N1} = 1/2$.

2. **A complete graph model**: all patches are connected:

$$M_{ij} = \begin{cases} 1/(N-1), & i \neq j, \\ 0, & i = j. \end{cases}$$

The ODEs in each patch are given by equations similar to (1-2) for division-controlled growth, or (6-7) for death-controlled growth, with $\mathcal{K} = K$ and migration terms included. In the former case, the system at each patch looks like

this:

$$\dot{x}_i = l_w x_i(1-u)\left(1-\frac{x_i+y_i}{K}\right) - d_w x_i - m\left(x_i - \sum_{j=1}^{N} M_{ji}x_j\right), \quad (13)$$

$$\dot{y}_i = (l_w x_i u + l_m y_i)\left(1-\frac{x_i+y_i}{K}\right) - d_m y_i - m\left(y_i - \sum_{j=1}^{N} M_{ji}y_j\right), \quad (14)$$

where x_i and y_i are the numbers of wild type and mutant cells in patch i, respectively. The initial value problem is completed with the following initial condition:

$$x_i(0) = \begin{cases} 1, & i = (N+1)/2, \\ 0, & \text{otherwise}, \end{cases} \qquad y_i(0) = 0 \quad \forall i \quad (15)$$

(we assumed an odd number of patches), that is, there is a single wild type cell in the middle patch, and the rest of the patches are unoccupied.

The results of the N-patch model will be compared with an unfragmented, mass-action system of size (carrying capacity) $\mathcal{K} = NK$:

$$\dot{X} = l_w X(1-u)\left(1-\frac{X+Y}{KN}\right) - d_w X, \quad (16)$$

$$\dot{Y} = (l_w X u + l_m Y)\left(1-\frac{X+Y}{KN}\right) - d_m Y, \quad (17)$$

$$X(0) = 1, \quad Y(0) = 0. \quad (18)$$

For death-controlled growth, all equations are modified accordingly.

2.4 The number of mutants in fragmented and mass-action systems

Let us compare the number of mutants obtained in the fragmented system with migrations and in the mass action system of the same total carrying capacity. Suppose we are interested in measuring the number of mutant at a total population size N_{tot}. Then we define the time t_{fr} and t_{ma} (for "fragmented" and "mass-action" respectively) as follows:

$$\sum_{i=1}^{N}(x_i(t_{fr}) + y_i(t_{fr})) = N_{tot}, \quad X(t_{ma}) + Y(t_{ma}) = N_{tot}.$$

We want to compare the numbers of mutants,

$$M_{fr}(N_{tot}) = \sum_{i=1}^{N} y_i(t_{fr}), \quad M_{ma}(N_{tot}) = Y(t_{ma}).$$

Figures S6(a) and S7(a) show two examples of functions $M_{fr}(N_{tot})$ and $M_{ma}(N_{tot})$,

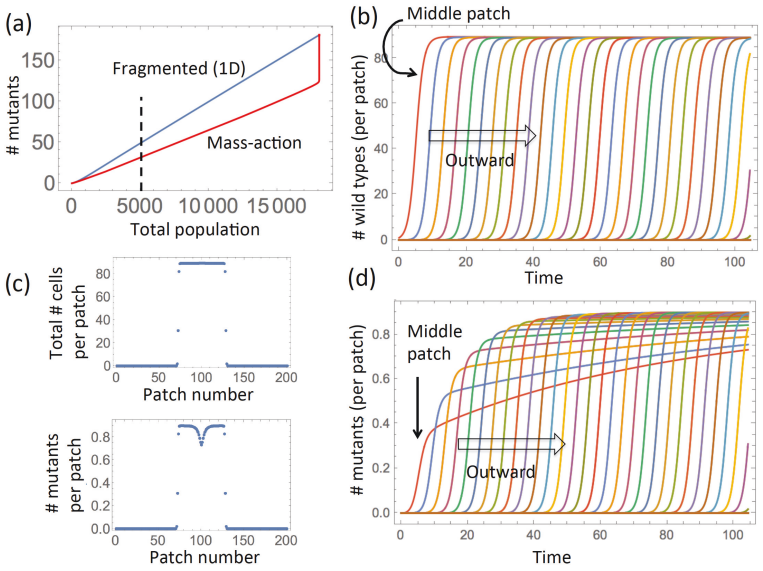


Figure S6: Mutant dynamics in the case of "decreased divisions" mutants. (a) Comparison of the mutant number as a function of the population size for the fragmented (blue) and mass-action (red) systems. A particular value of the total population size, $N_{tot} = 5000$, is indicated by a dashed vertical line. (b) The wild type populations in different patches as functions of time, for $t \in [0, t_{fr}]$ corresponding to $N_{tot} = 5000$. The leftmost line corresponds to the middle patch, and the rest of the lines to consecutive patches moving outward. (d) Same as (b) for the mutant populations. (c) The total populations (top) and the mutant populations (bottom) in all the patches at time t_{fr}. The parameters are $N = 201$ patches, $l_w = 1; d_w = 0.1, l_m = 0.9; d_m = 0.1, u = 0.001, K = 100; m = 0.01$.

for "decreased divisions" and "increased death" mutants respectively. We obtained numerical solutions of systems (13-14) with initial conditions (15) and the migration matrix (12) (a 1D ring of patches), to plot M_{fr} (the blue line) as a function of the total population size. The corresponding number of mutants in the mass action system, M_{ma} (the red line), was obtained from system (16-18). We observe that in figures S6(a) ("decreased divisions"), there are more mutants in the fragmented, spatial system, and in figures S7(a) ("increased deaths"), there are more mutants in the mass action system.

Panels (b,c,d) of figures S6 and S7 elucidate some underlying patterns. Panels (b,d) depict population dynamics in the patches. The initial wild type population in the "middle" patch number 101 (out of 200) grows and "seeds" the neighboring patches, whose population starts increasing, which in turn gives rise to growth in the next patches, etc, see figures S6(b) and S7(b). Because of

the symmetries of the ODEs, patches equidistant from the middle have identical solutions. Figures S6(d) and S7(d) show the numbers of mutants as they evolve in time, in each patch. Again, the leftmost line corresponds to the middle patch that was occupied initially. Individual trajectories of mutants in patches behave differently: "decreased divisions" mutants (figure S6(d)) grow monotonically toward the mutation selection balance (compare the blue line in figure S5(a)); "increased death" mutants (figure S7(d)) "overshoot" and then decrease toward the mutation selection balance (compare the yellow line in figure S5(a)).

Figure S8 generalizes these results. It presents the parameter space with coordinates l_m and d_m, where the points where $M_{ma}(N_{tot}) = M_{fr}(N_{tot})$ are shown as a purple line (parameters other than l_m and d_m are fixed). In the region above the purple line (gray shadowing), there are more mutants in the mass action system, and below it there are more mutants in the fragmented system. In particular, the system of figure S6 has coordinates $(l_m, d_m) = (0.9, 0.1)$ and falls outside the gray region, and the system of figure S7 corresponds to $(1, 0.2)$ and is inside the gray region; both points are marked by a "$*$" in the figure. In fact, we can see that all the "decreased divisions" mutants (points that lie on the dashed horizontal line in the disadvantageous region) fall under the purple line, that is, decreasing division rate leads to having more mutants in the fragmented system. It is somewhat less straightforward with "increased death" mutants (disadvantageous mutants on the dashed vertical line): if the increase in death is sufficiently large, then such points are above the purple line (and there are more mutants in mass action). For a small increase in death rates, there are more mutants in the fragmented system.

2.5 A simplified theory of mutant abundance in spatial and mass-action systems

Simulations in panels (b,d) of figures S6 and S7 were run for time t_{fr} that corresponds to the fragmented system reaching a specific total size $N_{tot} = 5000$, denoted by the dashed vertical line in panel (a). At that time, the total numbers of cells per patch and the numbers of mutants per patch are shown in panel (c). We observe that at time t_{fr},

(i) not all patches are occupied,

(ii) in most occupied patches, except for the outmost ones, the total populations have reached the carrying capacity (given by $l_w(1 - d_w/l_w)$), and

(iii) in most patches, the number of mutants is at the selection mutation balance. The exceptions are again the outmost patches that have not stabilized yet, and the middle patches, where the growth of mutants is slower.

Using the observations of patch dynamics listed above, let us approximate the number of mutants in a 1D fragmented system. At the time the total population reaches a size N_{tot}, we assume that there will be a number of patches

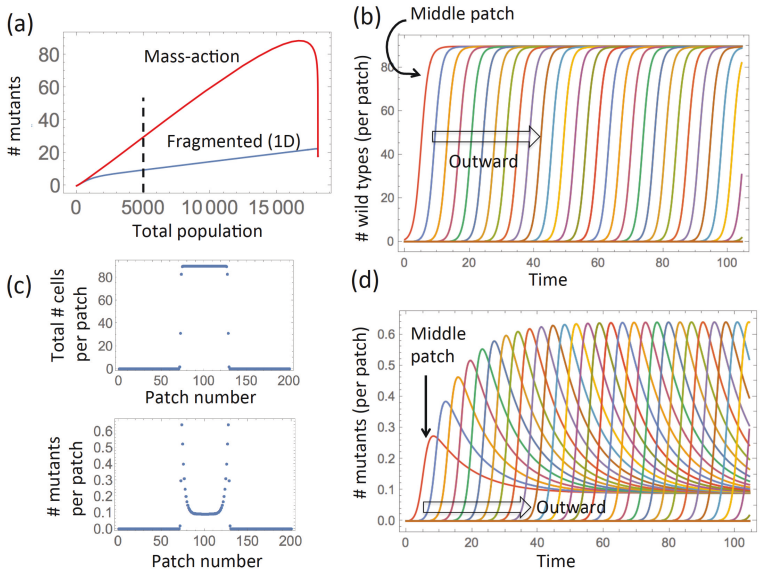


Figure S7: Mutant dynamics in the case of "increased death" mutants. Same as in figure S6, except $l_m = 1, d_m = 0.2$.

that are completely occupied, with population at carrying capacity, $\bar{x} = K(1 - d_w/l_w)$, and a number of patches that have not been reached yet. There are

$$n = \frac{N_{tot}}{K(1 - d_w/l_w)}$$

such occupied patches. Therefore, the total number of mutants at size N_{tot} is given by $n\bar{y}$, where y is given by equation (5), and we have

$$M_{fr}(N_{tot}) \approx N_{tot} u \left(\frac{d_m}{d_w} - \frac{l_m}{l_w} \right)^{-1}. \tag{19}$$

To calculate the number of mutants in the non-fragmented population, we use formula (9). We have, using $x_{lin} = e^{(l_w - d_w)t} \approx N_{tot}$, that

$$M_{ma}(N_{tot}) \approx l_w u \frac{N_{tot} - N_{tot}^{\frac{l_m - d_m}{l_w - d_w}}}{(l_w - d_w) - (l_m - d_m)}; \tag{20}$$

note that this approximation works not only for disadvantageous but for neutral or advantageous mutants. The solution set of the equation $M_{fr} = M_{ma}$ using approximations (19, 20) is shown in figure S8 as a black solid line. Although it is similar to the full solution (purple line) for small values of l_m, it deviates from it as l_m becomes closer to l_w.

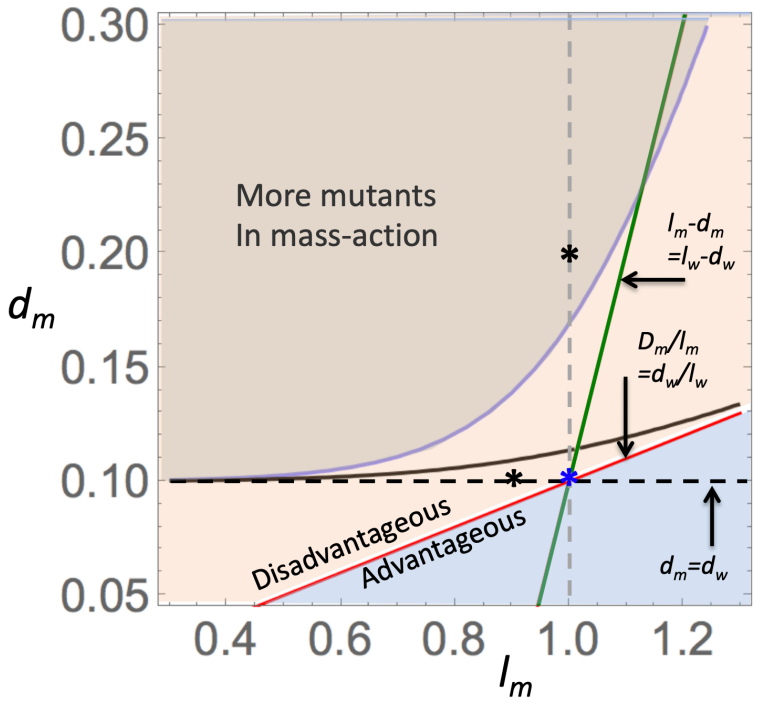


Figure S8: Regions in the parameter space (l_m, d_m), where $M_{ma} > M_{fr}$ (under fixed (l_w, d_w)). The red line indicate sets where mutants and the wild types have equal fitness (equation (3) with $\nu = 0$; mutants are disadvantageous above this line); the green line corresponds to equal linear growth rates. The set (21) is above the dashed horizontal line. The solid black line comes from a numerical solution of $M_{ma} = M_{fr}$ with expressions (19,20). The purple solid line is the solution of the same equation using the values obtained from solving the ODEs. $N_{tot} = 1000$, the rest of the parameters are as in figure S6. The point where $l_m = l_w$ and $d_m = d_w$ is marked by a blue star.

One can further simplify formula (20), assuming that $N_{tot} \gg N_{tot}^{\frac{l_m - d_m}{l_w - d_w}}$. Then, solving $M_{ma} > M_{fr}$, one obtains a condition that does not depend of N_{tot}:

$$\frac{d_m}{d_w} - \frac{l_m}{l_w} > \frac{(l_w - d_w) - (l_m - d_m)}{l_w},$$

which can be solved for d_m to yield:

$$d_m > d_w. \tag{21}$$

In other words, if $d_m > d_w$, there are more mutants in the mass action system, and otherwise there are more mutants in the fragmented system (given that

the total sizes of the two populations are the same). Figure S8 shows this approximation as a horizontal dashed line.

2.6 An intuitive explanation

Here we provide an intuitive explanation for the following result:

- If the mutant disadvantage is due to reduced division rates compared to the wild type, then there tends to me more mutants in the fragmented system compared to the mass action system of the same size.
- If the mutant disadvantage is due to increased division rates, then the result is the opposite, and there tends to be more mutants in the mass-action system. This is true only if the disadvantage is sufficiently strong.

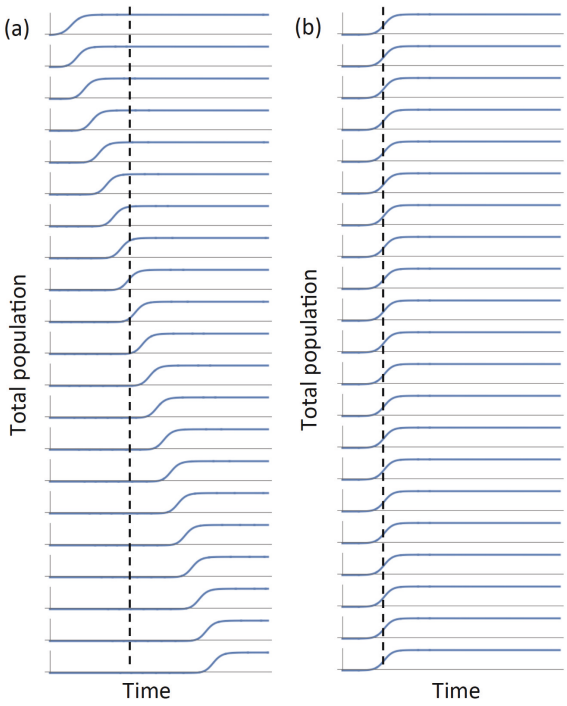


Figure S9: Time series of total populations sizes in patches for (a) a fragmented system with migrations according to a 1D spatial pattern and (b) a number of identical, uncoupled, patches that mathematically are equivalent to the mass action system.

Let us compare two systems: (i) a fragmented system of N patches of carrying capacity K each, where initially there is a single cell in one patch, and popu-

lations spread from patch to patch by means of migration, reaching their carrying capacity at different times, and (ii) a mass action system, which can be mathematically represented an a equivalent system of N decoupled identical systems of carrying capacity K, with identical initial conditions $x_i(0) = 1/N, y_i(0) = 0$. Figure S9 shows the simulated total population size in a number of patches, for (a) a 1D chain of fragmented patches and (b) a mass-action system represented as N identical patches.

If the target population, N_{tot}, is the maximum size of the mass-action system ($N\bar{x}$ with x given by formula (4) or (8) in the case of the division-controlled or death controlled growth, respectively), then the above results do not hold, and the number of mutants in both fragmented and non-fragmented system is simply $N\bar{y}$, formula (5). Let us instead assume that the total population size, N_{tot}, is well below maximum. As a consequence, at size N_{tot}, each of the identical, disconnected populations in figure S9(b) that represent the mass action system, have not reached its maximum size. This is shown by a vertical dashed line that cuts across the growth phases of the N patches in panel (b). A different picture is observed in the case of a fragmented system (panel (a)). There, since the growth in different patches happens at different times, by the time some of the patches have reached their maximum size, others have hardly started growing. As a consequence, at total size N_{tot}, we expect that a number of patches are "full" and others are "empty", see panel (b).

Now, we can formulate the problem of maximization of the number of mutants as size N_{tot} in the following way. We can make up size N_{tot} out of individual (identical) patches of size x (precisely, N_{tot}/x patches). The total number of mutants is then given by

$$M_{tot}(x) = N_{tot}\frac{\mathcal{M}(x)}{x}, \tag{22}$$

where $\mathcal{M}(x)$ is the number of mutants in a single patch of size x. What value of x maximizes the function $M_{tot}(x)$? The answer depends on the function $\mathcal{M}(x)$. If, for example, it is a convex function, then quantity (22) is a growing function of x, and is maximized by a smaller number of patches, each at its maximum size ($x = \bar{x}$). If $\mathcal{M}(x)$ is a concave function, then it M_{tot} is a decreasing function of x, and we find a maximum number of mutants if all the N patches contribute the smallest possible amount into the total.

Functions $\mathcal{M}(x)$ have different shapes for different mutant types, see figure S5(b). For "decreased divisions" mutants, it is always a convex function, and thus M_{tot} is maximized by a fragmented system, where at total size N_{tot}, a number of patches are already at carrying capacity while other have hardly began to grow.

For "increased death" mutants, the situation is slightly more complex. We know that these mutants grow faster than the wild type at the initial stages of growth (figure S5(a), yellow line), but depending on the degree of disadvantage, this growth may result in a monotonically growing function $y(t)$ for small degrees of disadvantage, or in a function $y(t)$ that "overshoots", for larger degrees of disadvantage. In the latter cases, the function $\mathcal{M}(x)$ has the shape depicted

in figure S5(b), yellow line. As a consequence, function (22) will have a larger value for small x (corresponding to the mass-action system where many virtual identical patches contribute a small amount) than for large x (corresponding to the fragmented system).

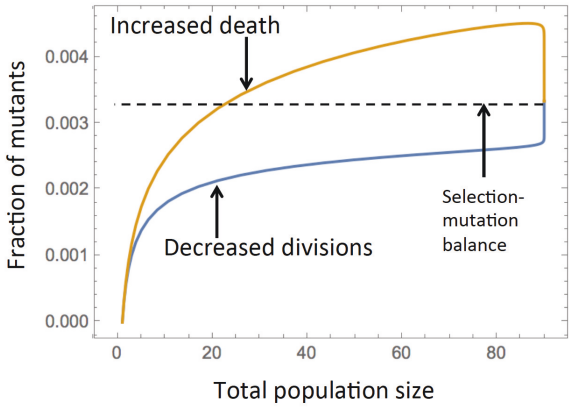


Figure S10: The fraction of mutants as a function of total population size for "decreased divisions" and "increased death" mutants. Parameters are as in figure S5.

Figure S10 shows the fraction of mutants as a function of total population size for "decreased divisions" and "increased death" mutants. For "decreased divisions" mutants, this is an increasing function of the total population size, and therefore to increase the percentage of mutants, one needs to maximize population size. This corresponds to having fewer patches at maximum size. For "increased death" mutants, the fraction of mutants first increases and the decreases. It is larger at an intermediate colony size compared to the maximum size for a large range of sizes, except for the initial stage of growth. Therefore, in most situations, a colony that is still growing will contain a larger percentage of mutants than the colony at maximum size. Therefore, we expect more mutations in the mass action situation where the saturation has not happened yet, and not in a fragmented system, where most colonies are at maximum size.

2.7 Migrations on a complete graph model: deterministic and stochastic cases

Note that the intuitive explanation presented above is not specific for a 1D geometry and holds for any fragmented system where patches are de-synchronized, that is, they grow to their maximum size at different times. This is a necessary condition to be able to identify the stage of growth where a subset of patches is fully grown while the rest are empty. This is why the results described above (that is, a difference in the behavior of decreased divisions mutants and in-

creased death mutants) is observed in stochastic simulations where no specific spatial arrangement of patches is assumed, and migration happens randomly from a patch to any other patch. In such a stochastic system, some patches will grow faster than others, and the general growth pattern similar to figure S9(a) is observed.

This behavior however is not captured by the corresponding system of ODEs. There, in the presence of equal migration rates to any patch, all patches (except for the original patch containing the first cell) are synchronized and grow in an identical manner. In fact, the ODEs in this case can be rewritten as only 2 equations in 2 patches:

$$\begin{aligned}
\dot{x}_* &= l_w x_*(1-u)\left(1-\frac{x_*+y_*}{K}\right) - d_w x_* - m\left(x_* - \frac{1}{N-1}X_*\right), \\
\dot{y}_* &= (l_w x_* u + l_m y_*)\left(1-\frac{x_*+y_*}{K}\right) - d_m y_* - m\left(y_* - \frac{1}{N-1}Y_*\right), \\
\dot{X}_* &= l_w X_*(1-u)\left(1-\frac{X_*+Y_*}{K(N-1)}\right) - d_w X_* - m\left(\frac{1}{N-1}X_* - x_*\right), \\
\dot{Y}_* &= (l_w X_* u + l_m Y_*)\left(1-\frac{X_*+Y_*}{K(N-1)}\right) - d_m Y_* - m\left(\frac{1}{N-1}Y_* - y_*\right), \\
& x_*(0)=1, \quad y_*(0)=0, \quad X_*(0)=0, \quad Y_*(0)=0.
\end{aligned}$$

This system can be thought of as a two-patch model, where the first patch has population (x_*, y_*) and the carrying capacity K, and the second patch population (X_*, Y_*) and the carrying capacity $K(N-1)$. The migration rate from the small to the large patch is m, and back it is $m/(N-1)$. Therefore, because of this artificial symmetry arising from the deterministic description (which is broken in a stochastic model), the behavior of this system does not reflect the patterns described above.

3 Disadvantageous mutants: a 2D spatial stochastic model

If mutants are disadvantageous, the quasi-equilibrium level of mutants is defined by the selection-mutation balance.

Let us calculate equilibrium densities of mutants and wild type cells in a spatially distributed system at steady state. This will also correspond to the densities in the core of an expanding system away from the advancing front.

We restrict our description to a 2D square grid, with the von Neumann neighborhood (that is, each location has 4 nearest neighbors); the methodology is generalizable to the Moore neighborhood (8 neighbors). We use a method similar to that of [**?**]. Two random variables describe the state of the stochastic system at each spatial location, x: ρ_x describes wild type cells, such that

$$\rho_x = \begin{cases} 1, & \text{if a wild type cell is at location } x, \\ 0, & \text{otherwise,} \end{cases}$$

and η_x describes mutant cells, such that

$$\eta_x = \begin{cases} 1, & \text{if a mutant cell is at location } x, \\ 0, & \text{otherwise.} \end{cases}$$

Note that ρ_x and η_x cannot be equal to one simultaneously; an empty spot corresponds to $\rho_x = \eta_x = 0$. We assume that wild type cells have division and death rates l_w and d_w, and mutant cells have division and death rates l_m and d_m. Wild type cells mutate with probability u, and no back mutations are considered.

3.1 Equations for the densities

Denote the expectation of ρ_x and η_x by

$$\langle \rho_x \rangle = \rho, \quad \langle \eta_x \rangle = \eta,$$

where we assumed that the expected values do not depend on spatial location, since we are interested in spatially homogeneous equilibrium solutions. We have

$$\dot{\rho} = \left\langle \frac{l_w}{N_b}(1-u)(1-\rho_x)(1-\eta_x)\sum_k \rho_x^{(k)} - d_w \rho_x \right\rangle, \tag{23}$$

$$\dot{\eta} = \left\langle \frac{1}{N_b}(1-\rho_x)(1-\eta_x)\sum_k \left(l_w u \rho_x^{(k)} + l_m \eta_x^{(k)} \right) - d_m \eta_x \right\rangle, \tag{24}$$

where the product $(1-\rho_x)(1-\eta_x)$ is nonzero only if location x is empty, and the summation goes over all the neighbors of point x, which reproduce into location x at rates l_w/N_b and l_m/N_b if they are wild type of mutant, respectively.

Let us consider the von Neumann neighborhood ($N_b = 4$). In the right hand side of equation (23), the terms is the summation have the form

$$\langle (1-\rho_x)(1-\eta_x)\rho_x^{(k)} \rangle = \langle \rho_x^{(k)} \rangle - \langle \rho_x \rho_x^{(k)} \rangle - \langle \rho_x^{(k)} \eta_x \rangle + \langle \rho_x^{(k)} \rho_x \eta_x \rangle = \rho - W - I, \tag{25}$$

and in equation (24) there are also terms of the form

$$\langle (1-\rho_x)(1-\eta_x)\eta_x^{(k)} \rangle = \langle \eta_x^{(k)} \rangle - \langle \rho_x \eta_x^{(k)} \rangle - \langle \eta_x^{(k)} \eta_x \rangle + \langle \eta_x^{(k)} \rho_x \eta_x \rangle = \eta - I - M. \tag{26}$$

In the expressions above, we have $\langle \rho_x \rho_x^{(k)} \eta_x^{(k)} \rangle = 0$, because either $\eta_x^{(k)}$ or $\rho_x^{(k)}$ is zero at location $x^{(k)}$, and the three types of dyads are defined as follows:

- $W = \langle \rho_x \rho_x^{(k)} \rangle$ is the probability to have two wild type cells at two neighboring locations,
- $I = \langle \rho_x \eta_x^{(k)} \rangle$ is the probability to have a wild type cell and a mutant at two neighboring locations,
- $M = \langle \eta_x \eta_x^{(k)} \rangle$ is the probability to have two mutant cells at two neighboring locations.

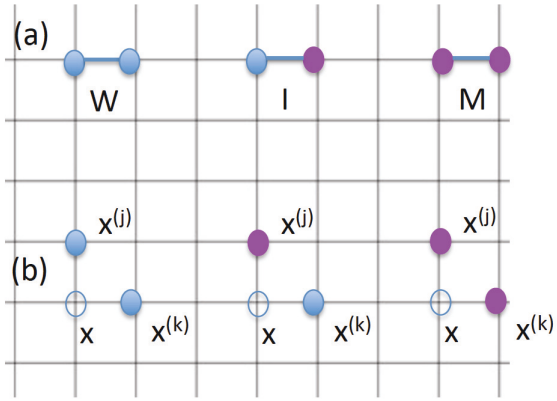


Figure S11: Steps in the derivation of equations for a two-component system of wild type and mutant cells. Blue circles denote wild type, and purple denote mutant cells. (a) Three configurations, whose correlations appear in equations (27) and (28). (b) Three types of correlations needed for equations for W, I, and M.

Figure S11(a) illustrates these three configurations. In terms of these three correlations, equations (23) and (24) can be rewritten as

$$\dot{\rho} = l_w(1-u)(\rho - W - I) - d_w\rho, \tag{27}$$
$$\dot{\eta} = l_w u(\rho - W - I) + l_m(\eta - I - M) - d_m\eta. \tag{28}$$

The correlations for the three dyads that appear in these equations require their own equations to close the system. Let us derive an equation for W. We have

$$\dot{W} = \left\langle 2(1-\rho_x)(1-\eta_x)\rho_x^{(k)} \sum_j \rho_x^{(j)} \frac{l_w}{N_b}(1-u) - 2d_w W \right\rangle,$$

where we assume that one of the points in the dyad contains a wild type cell (term $\rho_x^{(k)}$), while the other point is empty (term $(1-\rho_x)(1-\eta_x)$), and that one of its neighbors (location $x^{(j)}$) contains a wild type cell, which reproduces faithfully into point x at rate $l_w(1-u)/N_b$. Note that either of the two points could be empty, which results in the multiplier 2 in the first term on the right hand side. Similarly, either of the dyad's locations can experience cell death, resulting in the negative rate $2d_w$. In order to calculate the average, we need to consider terms

$$\langle (1-\rho_x)(1-\eta_x)\rho_x^{(k)}\rho_x^{(j)} \rangle. \tag{29}$$

Note that here and below, the operation of averaging makes the expression independent on the actual location x. Further, the superscripts (k) and (j) do not refer to any specific neighbor of x, but to any neighbor of x; in particular,

location $x^{(j)}$ may be the same or different than location $x^{(k)}$. In the case when the two locations are different, correlation (29) is presented in figure S11(b), on the left.

In equations for $\dot{M}$ and $\dot{I}$, the following expressions appear in addition to (29):

$$\langle(1-\rho_x)(1-\eta_x)\rho_x^{(k)}\eta_x^{(j)}\rangle, \quad \langle(1-\rho_x)(1-\eta_x)\eta_x^{(k)}\eta_x^{(j)}\rangle.$$

These correlations are shown in figure S11(b), center and right. Therefore, denoting by a and b either ρ or η, we evaluate the average of the form

$$\langle(1-\rho_x)(1-\eta_x)a_x^{(k)}b_x^{(j)}\rangle, \tag{30}$$

which corresponds to a dyad with one of the locations (location x) empty, and the other (location $x^{(k)}$) containing type "a", while a neighbor of x (location $x^{(j)}$) contains type "b". First let us assume that location $x^{(j)}$ is different from location $x^{(k)}$. Under von Neumann neighborhoods this implies that $x^{(j)}$ are $x^{(k)}$ are not each other's neighbors, because on a square grid, there could not be a non-degenerate triangle with diameter 1 or less. Expression (30) is equal to

$$\begin{aligned}
&P(b_x^{(j)}=1|a_x^{(k)}=1,\rho_x=\eta_x=0)P(a_x^{(k)}=1,\rho_x=\eta_x=0)\approx\\
&P(b_x^{(j)}=1|\rho_x=\eta_x=0)P(a_x^{(k)}=1,\rho_x=\eta_x=0)=\\
&\frac{P(b_x^{(j)}=1,\rho_x=\eta_x=0)P(a_x^{(k)}=1,\rho_x=\eta_x=0)}{P(\rho_x=\eta_x=0)}
\end{aligned} \tag{31}$$

The expression in the denominator is calculated as follows:

$$P(\rho_x=\eta_x=0)=\langle(1-\rho_x)(1-\eta_x)\rangle=\langle 1-\rho_x-\eta_x+\rho_x\eta_x\rangle=1-\rho-\eta.$$

Depending on the types at location x, the expressions in the numerator of (31) can be of two types:

$$P(\rho_x^{(k)}=1,\rho_x=\eta_x=0) \text{ or } P(\eta_x^{(k)}=1,\rho_x=\eta_x=0),$$

and they are calculated in (25) and (26) respectively.

Next, we assume that location $x^{(j)}$ is the same as $x^{(k)}$. Then, if types "a" and "b" in expression (30) are different, then we obtain $\langle(1-\rho_x)(1-\eta_x)\rho_x^{(k)}\eta_x^{(k)}=0$. If the types are the same, then we obtain expression (25) or (26). To summarize, expressions of type (30) are given as follows:

$$\begin{aligned}
\langle(1-\rho_x)(1-\eta_x)\rho_x^{(k)}\rho_x^{(j)}\rangle &= \begin{cases} \frac{(\rho-W-I)^2}{1-\rho-\eta}, & x^{(j)}\neq x^{(k)},\\ \rho-W-I, & x^{(j)}=x^{(k)}, \end{cases}\\
\langle(1-\rho_x)(1-\eta_x)\rho_x^{(k)}\eta_x^{(j)}\rangle &= \begin{cases} \frac{(\rho-W-I)(\eta-I-M)}{1-\rho-\eta}, & x^{(j)}\neq x^{(k)},\\ 0, & x^{(j)}=x^{(k)}, \end{cases}\\
\langle(1-\rho_x)(1-\eta_x)\eta_x^{(k)}\eta_x^{(j)}\rangle &= \begin{cases} \frac{(\eta-I-M)^2}{1-\rho-\eta}, & x^{(j)}\neq x^{(k)},\\ \eta-I-M, & x^{(j)}=x^{(k)}. \end{cases}
\end{aligned}$$

The equation for W is then given by

$$\dot{W} = \frac{l_w}{2}(1-u)\left(\rho - W - I + \frac{3(\rho - W - I)^2}{1-\rho-\eta}\right) - 2d_w W. \tag{32}$$

Similarly, the other two equation can be derived:

$$\begin{aligned}
\dot{I} &= \frac{3}{4}[l_w(1-u)+l_m]\frac{(\rho - W - I)(\eta - I - M)}{1-\rho-\eta} + \frac{l_w u}{4} l_m \left(\rho - W - I + \frac{3(\rho - W - I)^2}{1-\rho-\eta}\right) \\
&- (d_w + d_m) I, \qquad (33) \\
\dot{M} &= \frac{3 l_w u}{2}\frac{(\rho - W - I)(\eta - I - M)}{1-\rho-\eta} + \frac{l_m}{2}\left(\eta - I - M + \frac{3(\eta - I - M)^2}{1-\rho-\eta}\right) - 2d_m M. \qquad (34)
\end{aligned}$$

The closed system of equations for ρ, η, W, I, and M is given by equations (27), (28), (32), (33), and (34).

3.2 Selection mutation balance solution

Solving these equations in steady state exactly is difficult, but if the mutation rate $u \ll 1$, we can find the approximate solution. We start by setting $u = 0$ and obtaining the steady state solution. Apart from the trivial solution and a negative solution, there are two symmetric solutions where only one type survives (competitive exclusion). We will use the one where the wild type excludes mutants:

$$\rho^{(0)} = 1 + \frac{3d_w}{d_w - 3l_w}, \quad W^{(0)} = 1 - \frac{6d_w}{d_w - 3l_w} - \frac{4d_w}{l_w}, \quad \eta^{(0)} = I^{(0)} = M^{(0)} = 0, \tag{35}$$

where the superscript corresponds to the zeroth order in the expansion in terms of small u. Note that the expression for $\rho^{(0)}$ is identical to the steady state density of the one-component system given by equation (**??**) with $\xi = d_w/l_w$. We then look for the first correction by substituting

$$\rho = \rho^{(0)} + u\rho^{(1)}, \quad \eta = u\eta^{(1)}, \quad W = W^{(0)} + uW^{(1)}, \quad I = uI^{(1)}, \quad M = uM^{(1)},$$

inserting in the system of 5 equations, keeping only the first order of expansion in u, and solving for $\rho^{(1)}, \ldots, M^{(1)}$. We obtain

$$\eta = \eta^{(1)} u = \frac{d_w l_w (4d_w - 3l_w)(4d_m^2 + 3d_m l_w + d_w l_m) u}{d_m (d_w - 3l_w)(4d_m + 3l_w)(d_m l_w - d_w l_m)}.$$

This is the equilibrium solution corresponding to mutation-selection balance in the presence of spatial interactions. This approach is valid as long as the wild type is advantageous (inequality (3)). In the opposite scenario, this solution is unstable, and the system converges to the mutants excluding the wild type.

Under selection-mutation balance, of interest is the equilibrium proportion of mutants in the system given by

$$\nu_{vN} = \frac{\eta^{(1)} u}{\rho^{(0)}} = \frac{d_w l_w (4d_m^2 + 3d_m l_w + d_w l_m) u}{d_m (4d_m + 3l_w)(d_m l_w - d_w l_m)}. \tag{36}$$

In the absence of death ($d_w = d_m = 0$), we obtain the limiting value

$$\nu_{vN,D=0} = \frac{(3l_w + l_m)u}{3(l_w - l_m)}.$$

3.3 Comparison with mass-action

We would like to compare the proportion of mutants under selection mutation balance in space and in mass-action. In mass action we have

$$\dot{x} = l_w(1-u) - d_w x, \tag{37}$$
$$\dot{y} = = l_w u x + (l_m - d_m)y, \tag{38}$$

where x and y is the expected number of wild type and mutant cells respectively. The proportion of mutant cells at a given size N is calculated by using $x^{(0)} = e^{(l_w - d_w)t}$ and evaluating the solution $y(t)$ of equation (38) at time $t = \ln N/(l_w - d_w)$. Forming the fraction, we obtain the proportion at mutants at size N:

$$\nu_{ma} = \frac{l_w u \left(1 - N^{-\frac{l_w - d_w - (l_m - d_m)}{l_w - d_w}}\right)}{l_w - d_w - (l_m - d_m)}.$$

One can see that this quantity grows with the size N, while ν_{vN} is independent of the population size.

In figure S12 we study the quantity

$$\ln\left(\frac{\nu_{vN}}{\nu_{ma}}\right),$$

which is greater than 0 of the numerator is larger than the denominator. This quantity is presented as a contour plot in panel (a), as we vary l_m and d_m (for fixed values of l_w and d_w). The red line in the figure indicates the region where the mutants are disadvantageous and we expect to observe selection-mutation balance (this is given by inequality (3)). Positive regions (below the purple dashed line) correspond to having more mutants in the spatial system compared to the mass action system. Negative regions (above the purple dashed line) are regions where there are fewer mutants in the spatial system than in the mass action system.

Figure S12(b) focuses on two particular scenarios. The top panel presents the case where the division rates of mutants and wild type cells are the same, and the disadvantage is manifested through $l_m < l_w$. We can see that in such cases, $\nu_{vN} > \nu_{ma}$, that is, we have more mutants in space. The bottom panel shows the case where $l_m = l_w$, and the mutants have a larger death rate than the wild types. In this case, if the disadvantage is sufficiently large, we have fewer mutants in space than in the mass action system.

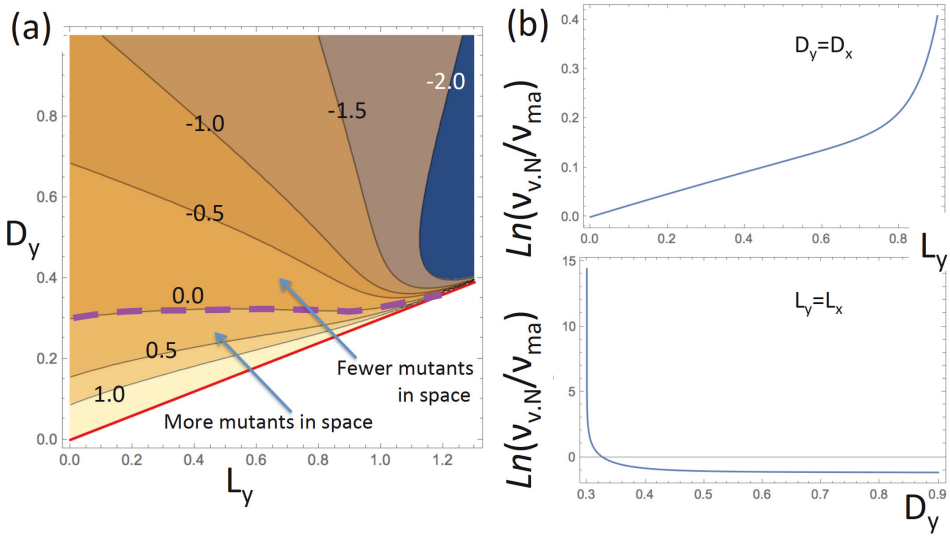


Figure S12: Comparison of the spatial and mass action system. (a) The quantity $\ln\left(\frac{\nu_{vN}}{\nu_{ma}}\right)$ is presented as a contour plot as a function of l_m and d_m, for fixed values of l_w and d_w. The red line indicates the region where mutants are disadvantageous (inequality (3)). The contours's values are indicated, and the zero contour is marked with a dashed purple line. (b) Top panel: the same quantity plotted as a function of l_m, with $d_m = d_w$. Bottom panel: the same, as a function of d_m, with $l_m = l_w$. The rest of the parameters are given by $u = 2 \times 10^{-5}$, $N = 10^5$, $l_w = 1$, and $d_w = 0.3$.

3.4 Comparison with computations

The expected number of mutants predicted theoretically was compared with results of numerical simulations. This was done in the following way. At size N, the number of mutants (in the von Neumann case) is predicted to be $Nu\nu_{vN}$, see equation (36). Solving the equation $Nu\nu_{vN} = const$, we can obtain the pairs (l_m, d_m) of mutant kinetic rates corresponding to a predicted given number of mutants in a system of size N. Figure S13(a) shows the predicted number of mutant as a contour plot. The closer to the "neutrality" line (see inequality (3)) the larger the predicted number of mutants. Solution of equation $Nu\nu_{vN} = const$ is shown in panel (b), and for 5 points from the solution set, the predicted number of mutants (given by 10) is compared with the numerically obtained mean (plotted together with the standard deviation, panel (c)). We can see that for larger mutant division rates, the deviation from the theory becomes significant. Panel (d) shows a histogram of numbers of mutants for the parameters corresponding the 5th point. One can see that the distribution has a long tail and a very large standard deviation. This is the consequence of "slices" that cannot be handled

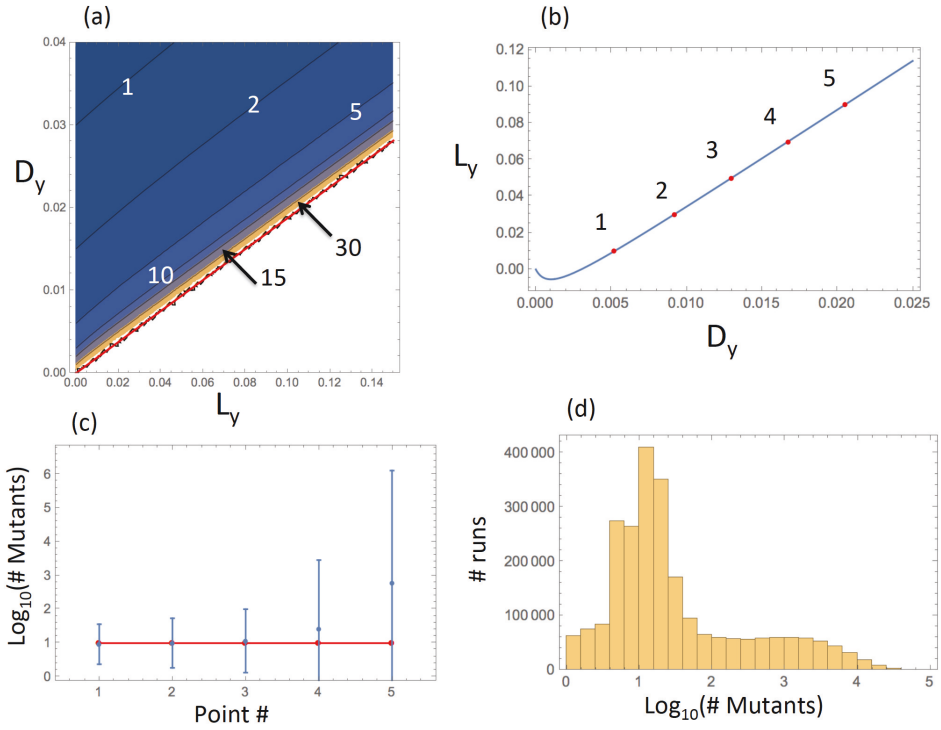


Figure S13: The level of mutants in the spatial (von Neumann) system: analytical approximation and numerical results. (a) The quantity $Nu\nu_{vN}$ is presented as a contour plot as a function of l_m and d_m, for fixed values of l_w and d_w. Mutants are disadvantageous above the red line (inequality (3)). The contours' values are specified. (b) Solution l_m of equation $Nu\nu_{vN} = 10$ as a function of d_m; the 5 points used in panel (c) are marked in red and numbered. (c) The comparison of predicted (10, horizontal red line) and simulated number of mutants in the 5 parameter pairs from panel (b). Simulated means and standard deviations are shown (out of 2.5×10^6 runs). (d) For the 5th parameter combination, the numerically obtained histogram of the number of mutants is shown. The rest of the parameters are $u = 2 \times 10^{-5}$, $N = 10^5$, $l_w = 0.08$, and $d_w = 0.015$.

by the present method.

3.5 Jack-pot events

In Fig. 2 of the main text we studied the mean and the distribution of the number of disadvantageous mutants in mass-action systems and in a metapopulation model, were the demes were arranged in a 1D array, and migration happened be-

tween neighboring demes. Here we present results for a metapopulation model where migration was equally likely among all demes, see figure S14. We can see that regardless of the structure of the deme-to-deme network, the results are very similar.

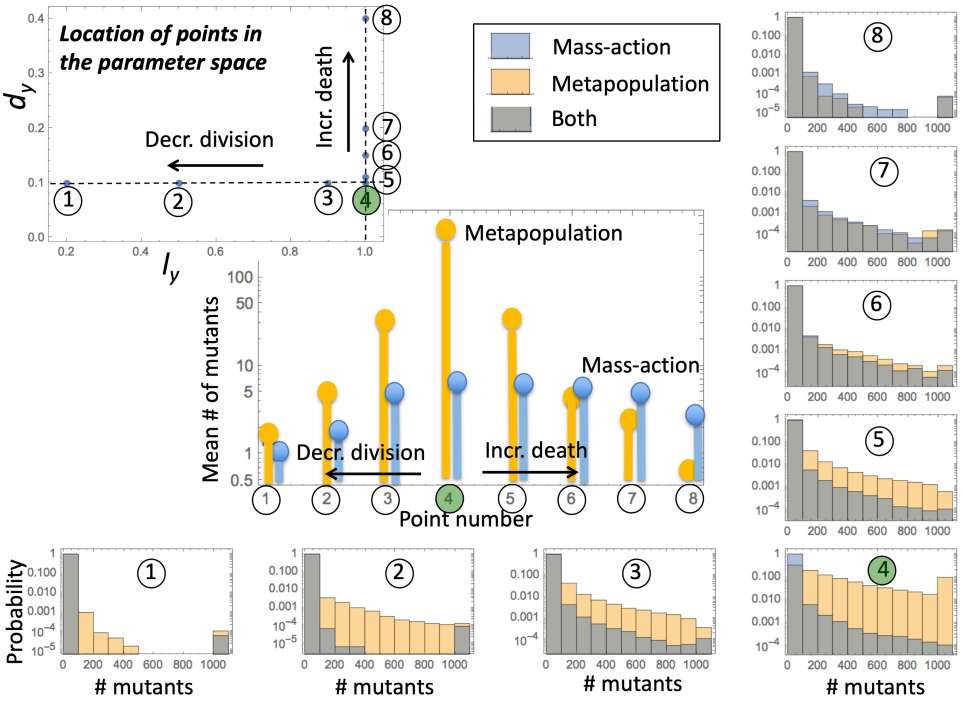


Figure S14: Disadvantageous mutants: probability distributions of mutant numbers in mass-action and metapopulation simulations. Both decreased divisions and increased death mutants are investigated, with division and death rates given in the left upper panel (compared to the rates of the wild type, denoted by the green circle). The bar graphs represent numerical histograms for mass action simulations (blue) and metapopulation simulations (yellow), where the demes were all connected to each other; between 1.8×10^5 and 2.0×10^5 simulations were run for each parameter combination, and the number of mutants recorded when the total population reached 1000. The mean numbers of mutants are shown for all the simulations in the central panel, with yellow markers corresponding to the metapopulation model (migration among all demes), and the blue markers to the mass-action model.

In order to compare the number of rare ("jackpot") events in different models and for different parameters, we designed a function that quantifies the spread of the distribution or the "fatness" of its tail (heavy-tailedness). For each set of data, Y, representing the numbers of mutants in each of the runs performed

for a given parameter set for a given model, we used the quantile function

$$F(Y, q) = Q(Y, 1 - q),$$

defined as follows. Let Y_{qL} and Y_{qR} are two subsets of Y such that $Y = Y_{qL} \cup Y_{qR}$ and all the elements of Y_{qL} are smaller or equal to all the elements of Y_{qR}; we will refer to parameter q as "threshold". We further assume that

$$|Y_{qR}| = \lceil |Y| q \rceil,$$

where $|.|$ denotes the number of elements and $\lceil . \rceil$ denotes ceiling. Then we set

$$Q(Y, 1 - q) = \min(Y_{qR}).$$

It follows from this definition that the function $F(Y_1, q) > F(Y_2, q)$ for all q in a vicinity of 0, as long as set Y_1 has a higher number of large outliers compared to Y_2, or if Y_1 is drawn from a distribution with a heavier tail compared to Y_2.

Figure S15 presents the plots of the heavy-tailedness function $F(Y, q)$ for the 8 cases studied in figure S14, with Y representing the mass-action model (blue), the metapopulation model where all demes are connected (yellow), and the 1D metapopulation model with only neighboring demes connected (green). The bottom row of graphs represents the cases where the disadvantageous mutants are characterized by a decreased division rate (except case 4, which describes neutral mutants). In the top row the disadvantageous mutants are characterized by increased death rates. We can see that the heavy-tailedness generally increases toward the neutral case. We also observe the following patterns:

- In the bottom row, for small thresholds, q, the yellow and green lines are above the blue line, which means that if the mutants have decreased division rates, the fragmented (metapopulation) models are characterized by a higher heavy-tailedness (more jack-pot events) compared to the well-mixed system.

- This trend weakens and reverses in the graphs of the top row, from right to left (away from neutral mutants). In other words, if the mutants have increased death rates, and the disadvantage is sufficiently large, the fragmented (metapopulation) models are characterized by a lower heavy-tailedness (fewer jack-pot events) compared to the well-mixed system.

These trends go hand in hand with the results on the expected number of mutants (figure S14, middle graph). For mutants with decreased division rates or slightly increased death rates, there are more jack-pot events in the metapopulation models, and the expected number of mutants is also higher, compared to the well-mixed model. On the other hand, if the mutants are characterized by significantly higher death rates, the trend reverses, and there are more jack-pot events and more mutants on average in the mass-action system.

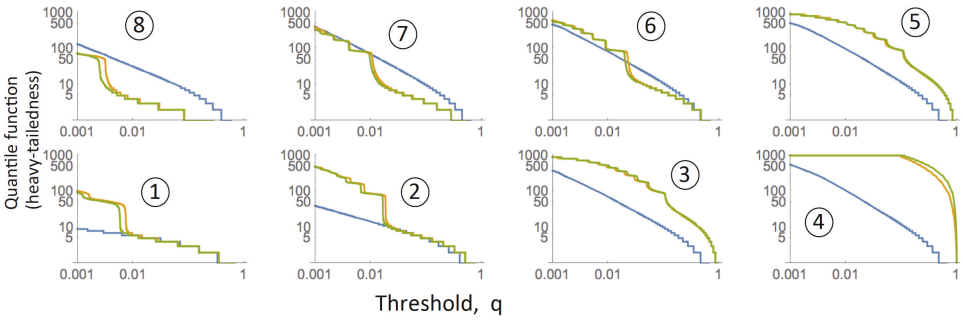


Figure S15: The prevalence of jack-pot events in different models. The heavy-tailedness (quantile) function $F(Y, q)$ is plotted vs q for the mass-action model (blue), the metapopulation model where all demes are connected (yellow), and the 1D metapopulation model with only neighboring demes connected (green). The calculations are based on the simulations of figure S14 and figure 2 of the main text. Case numbering and parameters are as in figure S14.

4 Neutral and advantageous mutants in a range expansion

4.1 Derivation of the growth laws

To derive the laws of mutant growth reported in the main text, we can use the following simple calculations. Let us assume that the death rate of cells is equal to zero, and consider cells growing in different geometries.

2D flat front. Assume that cells grow along the surface of a cylinder of width W. This represents a one-directional growth process, where during each generation, we assume that W new cells appear, and the the total population is given by $N = LW$, where L represents the number of layers. The value of L is proportional to the number of generations, and thus to the physical time, t:

$$L \propto t.$$

The following calculation estimates the growth law of mutants. Every time a new layer (of width W) is added, the mean number of new mutations is given by Wu. Suppose that mutants are neutral. Then, each such mutation will give rise to an array of daughter mutant cells of width 1, see figure S16. The length of this array is given by $L - i$, if i is the layer at which the mutation occurred. Therefore, the total expected number of neutral mutants is a cylinder of length L is given by

$$M^{neut}_{\text{2D flat}} = \sum_{i=1}^{L} Wu \sum_{j=i+1}^{L} 1 = \frac{uWL(L-1)}{2} \approx \frac{uWL^2}{2} = \frac{u}{2W} N^2, \qquad (39)$$

where we assumed $L \gg 1$. Note that in this derivation we assumed that the number of mutants is small compared to the total population, and individual mutant clones do not interact. In a more precise calculation, the number of wild type cells in each layer is smaller than W because of the existence of mutants, and thus the ate of new mutant production is smaller than Wu. We however assume that $uLW \ll 1$, such that the number of mutants is relatively small.

Note that the number of neutral mutants decreases with W, see figure S17; the largest number of mutants is achieved in the case of $W = 1$, a one-dimensional expanding array of cells.

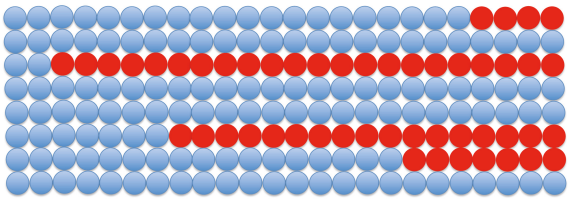


Figure S16: The conceptual model for mutant number calculations, the case of neutral mutants in a colony growing along the surface of a cylinder (2D flat front).

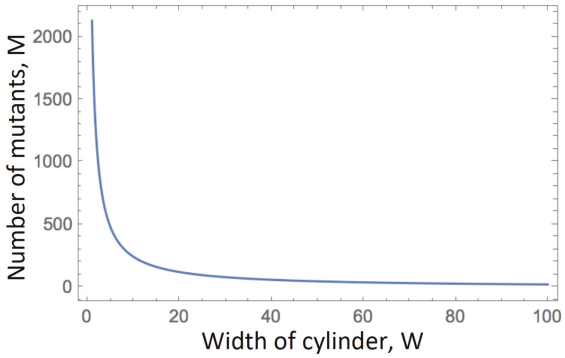


Figure S17: The number of mutants during a 2D flat front expansion decays with the front width. Formula (39) is presented with $N = 10000$ and $u = 5 \times 10^{-5}$.

Next, let us consider advantageous mutants. In this case, each new mutant gives rise to a triangular clone. In the first layer, the width of the clone is 1, in the next layer it is $1+s$, and in the kth layer it is $1+(k-1)s$, where parameter $s \geq 0$ measures the advantage of the mutants (with $s = 0$ corresponding to

neutral mutants). Therefore, we have

$$M^{adv}_{\text{2D flat}} = \sum_{i=1}^{L} Wu \sum_{j=i+1}^{L} (1 + (j - (i+1))s) = uWL(L-1)\left(\frac{1}{2} + \frac{s(L-2)}{6}\right) \quad (40)$$
$$\approx \frac{uWsL^3}{6} = \frac{us}{6W^2}N^3,$$

where for the approximation, we assumed that $Ls \gg 1$. Also, for this simple calculation to be valid, we need to assume that the wedges created by mutants do not come close to the cylinder's width, W, that is, $Ls \ll W$. In particular, formula (40) can be valid for small values of $W > 1$, but only for mutants that are neutral for practical purposes ($s \ll 1$).

Note that when N is fixed, the total number of cell divisions that the system has undergone is also fixed. The number of mutants however is vastly different depending on the spatial configuration. It is the highest for $W = 1$ (one row of cells) and decreases drastically with the width of the cylinder. This is consistent with the notion that spatial restrictions result in a heightened number of mutants, the 1D space ($W = 1$) being the most spatially restrictive system. The reason for this is that in 1D, a mutant, once created, blocks the whole range of expansion and prevents wild type cells from reproducing. The wider the front, the weaker this effect. Further, we note that in the special case where $W = 1$, mutant advantage does not play a role, and the number of advantageous, neutral, and even disadvantageous mutants is given by the same formula, equation (39).

2D: circular range expansion. Next we turn to the dynamics of neutral mutants on a circle. Let us suppose that the radius of the circle is R and $N = \pi R^2$. The size of the colony increases via surface growth with $N \propto t^2$ and

$$R \propto t.$$

As the range expansion proceeds, the circular layer of radius r will on average give rise to $2\pi ru$ new mutations. Each mutation will result in a wedge expanding outwards. If the new mutation occurred in layer with radius r, the number of mutating cells in layer r is 1. The number of mutants in the next layer is given by $\frac{r+1}{r}$, because under the assumption of mutant neutrality, the fraction of mutants in each new layer of radius $j > r$ (with surface $2\pi j$) should stay constant and equal to $\frac{1}{2\pi r}$. For layer j, the number of mutants is then given by j/r. This gives rise to the following calculation:

$$M^{neut}_{\text{2D range}} = \sum_{r=1}^{R} 2\pi ru \sum_{j=r+1}^{R} \frac{j}{r} = \frac{2}{3}\pi R(R^2-1)u \approx \frac{2\pi R^3 u}{3} = \frac{2u}{3\pi^{1/2}}N^{3/2}$$

(the approximation is valid for $R \gg 1$).

For advantageous mutants in a growing 2D circle, the fraction of mutants will grow with each layer:

$$\begin{aligned} M^{adv}_{\text{2D range}} &= \sum_{r=1}^{R} 2\pi r u \sum_{j=r+1}^{R} (1+(j-(r+1))s)\frac{j}{r} = \pi R(R^2-1)u\left(\frac{2}{3}+\frac{1}{4}s(R-2)\right) \\ &\approx \frac{\pi R^4 su}{4} = \frac{su}{4\pi}N^2, \end{aligned} \tag{41}$$

where we assumed $Rs \gg 1$. For this approximation to be valid, the mutant wedges should not exceed the circumference of the colony. Strictly speaking, this results in the condition $Rs << 2\pi R$, that is, $s \ll 1$. For larger values of s, the events where the mutant covers the whole surface of the colony are no longer negligible.

3D flat front. In a 3D space, let us first consider a solid cylinder of constant radius R_0, where initially the cells are situated as a layer at the bottom of the cylinder, and proceed to grow by adding layers of size πR_0^2. Each generation contributes $\pi R_0^2 u$ new mutants, and as the colony grows to length L (and volume $2\pi R_0^2 L$), we have in the neutral case:

$$M^{neut}_{\text{3D flat}} = \sum_{i=1}^{L} 2\pi R_0^2 u \sum_{j=i+1}^{L} 1 = \pi R_0^2 uL(L-1) \approx \pi R_0^2 uL^2 = \frac{u}{\pi R_0^2}N^2,$$

which is similar to the 2D flat front expansion. If the mutants are advantageous, then their number will increase from layer to layer, giving rise to conical wedges. This gives rise to the following calculation:

$$\begin{aligned} M^{adv}_{\text{3D flat}} &= \sum_{i=1}^{L} 2\pi R_0^2 u \sum_{j=i+1}^{L} (1+(j-(i+1))s)^2 \\ &= \frac{L(L-1)\pi R_0^2 u}{6}\left[(L^2-3L+2)s^2+4(L-2)s+6\right] \\ &\approx \frac{\pi R_0^2 s^2 uL^4}{6} = \frac{s^2 u}{6\pi^3 R_0^6}N^4, \end{aligned}$$

where $Ls \gg 1$ for the approximation, and the approach is valid as long as the wedge radius is smaller than that of the cylinder, $sL \ll R_0$.

3D range expansion. Next we consider a 3D expanding sphere. For a sphere of radius R, we have $N = 4/3\pi R^3$ and the surface is given by $4\pi R^2$. The size of the colony increases via 3D surface growth with $N \propto t^3$. Each spherical layer of radius r will on average give rise to $4\pi r^2 u$ new mutations. Each mutation will result in a conical wedge expanding outwards. If the new mutation occurred in layer with radius r, the number of mutating cells in layer r is 1. The number of mutants in a layer of radius $j > r$ is given by $(j/r)^2$, because under the

assumption of mutant neutrality, the fraction of mutants in each new layer should stay constant (and equal to $\frac{1}{4\pi r^2}$). Therefore, we write:

$$M^{neut}_{\text{3D range}} = \sum_{r=1}^{R} 4\pi r^2 u \sum_{j=r+1}^{R} \left(\frac{j}{r}\right)^2 = \pi R(R^2-1)(R+2/3)u \approx \pi R^4 u = \frac{3^{4/3}u}{\pi^{1/3}4^{4/3}} N^{4/3}$$

(the approximation is again valid for $R \gg 1$).

If the mutant in a growing 3D sphere is advantageous, the fraction in each layer will increase according to the fitness advantage s and stretch from layer to layer in the same way as for the neutral mutants. We therefore have,

$$\begin{aligned} M^{adv}_{\text{3D range}} &= \sum_{r=1}^{R} 4\pi r^2 u \sum_{j=r+1}^{R} (1+(j-(r+1))s)^2 \left(\frac{j}{r}\right)^2 \\ &= \frac{\pi u R(R^2-1)}{90}\left[(20R^3-48R^2-5R+42)s^2+(72R^2-90R-108)s+90R+60\right] \\ &\approx \frac{2}{9}\pi s^2 u R^6 = \frac{s^2 u}{8\pi} N^2. \end{aligned}$$

As before, the approximation holds if $Rs \gg 1$. The method assumes that the mutant colony's size in each layer does not come close to the surface area, which amounts to the inequality $s \ll 1$.

Exponential (non-spatial, mass-action) growth. Finally, for exponentially growing population, similar formulas could be derived. In particular, for neutral mutants, we have

$$M^{neut}_{\text{exp}} = Nu \ln N,$$

and for advantageous mutants with advantage α (which is the ratio of the net growth rate of mutants and the net growth rate of wild type cells), we have

$$M^{adv}_{\text{exp}} = \frac{\alpha}{(\alpha-1)2^{\frac{\alpha-1}{\alpha}}} N^{\frac{2\alpha-1}{\alpha}},$$

see [**?**], equation (14c), see also equation (13) for a more general formula.

A summary of some of the results is presented in table 1 of the main text.

4.2 Comparison with numerical simulations

We have run numerical simulations to check the results derived in the previous section. Below we comment on the applicability and limitations of the formulas derived.

4.2.1 Roughness considerations

If we use formula (39) to approximate simulation results for neutral mutants, in the absence of death, in a colony growing on the surface of a cylinder, we

notice that it gives a slight systematic underestimation of the number of mutants corresponding to N cells. Here we investigate the source of this error.

To understand the source of this inconsistency, we note that the idealized model of figure S16 is not realistic. The real front propagates to the right as a jagged line. For example, under the Moore neighborhood, cells can divide into 8 nearby spots, including the diagonal spots and positions up or down. There are two important differences between the idealized model of figure S16 and the real picture. (i) The real, jagged, front is longer (the idealized vertical front is the shortest), and (ii) cells do not always divide to the right, and as a consequence, cell divisions are less efficient: sometimes two neighboring cells "decide" to divide into the same spot, thus preventing some divisions from happening. Let us compare the number of successful divisions in the idealized and in the natural model. In the idealized model, the number of successful divisions per update is exactly W. For the natural model, the first factor above increases the expected number of successful divisions, while the second factor reduces the number of successful divisions. The overall effect turns out to be negative, that is, fewer than W successful divisions are performed. This suggests that the effective width of the cylinder is less than W, and therefore the expected number of mutants corresponding to the same population size is larger than that predicted for the idealized model.

Roughness considerations have been investigated thoroughly in [**?** **?**]. In particular, the authors studied the statistics of the mutant "bubbles" and "wedges". It turns out that their shape is affected, in a predictable way, by the front roughness, which was shown by using the previous theoretical and numerical results of [**?**] and [**?**]. Further, these considerations allow for calculating the scaling laws of the probability distribution of mutant clone sizes [**?**], and influence the probability distribution of the number of mutants, but not its mean. This can be seen, for example, if we use formula (5) of the Methods in [**?**] with the scaling exponents α and β estimated for flat and rough fronts, see Table 1: the resulting dependence of mutant clone size on the total population N remains the same. The proportionality constant, however, is roughness dependent, as follows from formula (4) in the Methods in [**?**].

This is consistent with our findings. Even though our idealized model used to derive mutant growth laws underestimates the "constant", it correctly predicts the power laws in different growth geometries and for different mutant types (disadvantageous, neutral, and advantageous).

4.2.2 2D flat front expansion

For the 2D flat front geometry (the surface of a cylinder), we set up the initial condition for the agent-based simulation to be a one layer (a circle) of wild type cells that coincides with the circumference of one of the cylinder's bases (and has length W). Simulations are run repeatedly for a fixed number of time steps, and the mean trajectory (that is, the mean number of wild type cells and the mean number of mutants, as functions of time-step) is then calculated. Finally, the number of mutants is plotted as a function of the total population size.

Examples for neutral mutants (that is, mutants that have the same division and death rates as the wild type cells) are presented in figure 4 of the main text, curves (a) and (b). We can see that both in the absence (a) and in the presence (b) of death, the mutant population as a function of the total population approaches a power law with the exponent 2 (the black dashed lines corresponding to cases (a) and (b) of figure 4 of the main text have slope 2 in the log-log plot, see table 1). There are more mutants in the presence (b) than in the absence (a) of death.

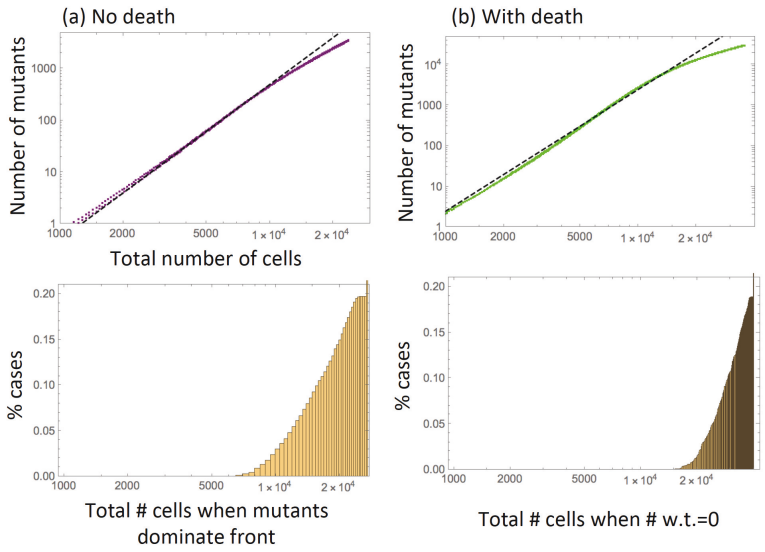


Figure S18: Advantageous mutants under a 2D flat front expansion: deviation from the cubic law for large sizes. (a) In the absence of death ($L_x = 0.7, L_y = 1.0$, $D_x = D_y = 0$, see curve (d) of figure 4 of the main text) and (b) in the presence of death ($L_x = L_y = 0.7$, $D_x = 0.2, D_y = 0.1$, see curve (f) of figure 4). Top panels: the number of mutants as a function of the total number of cells; the dashed straight lines have slope 3. Bottom panels: (a) a numerically obtained CDF of the probability for a given colony to have its front completely dominated by mutants by a given size; (b) a numerically obtained CDF of the probability for a given colony to have its wild type population extinct by a given size.

Advantageous mutants are presented by curves (c-g) of figure 4 of the main text; note that the dashed lines drawn through these curves all have slope 3 in the log-log coordinates (table 1). Curves (c) and (d) correspond to systems without death, and mutants having a larger division rates compared to wild type cells. The advantage is larger in case (d) compared to case (c) (and thus there are more mutants observed at the same population size). In both cases, we can see that the curves have slope 3 up to a certain population size, after

which they deviate from the cubic law. For those larger sizes, mutants grow slower as a function of size (quadratically). The reason for this deviation from the cubic law is as follows. As the colony grows and reaches larger sizes, advantageous mutant clones that grow out and increase in size start reaching the width W, that is, take up the whole width of the cylinder. After that, the mutant colony can no longer expand in width, but instead it grows linearly. This is illustrated in figure S18(a), where the top panel replots the purple line (case (d)) of figure 4 of the main text, and the bottom panel studies the statistics of mutant invasion. For each run, we recorded the total colony size at which the population of wild type cells stopped increasing (given that this happened in the time-span of the simulation). This indicates that, in the absence of death, all the front positions are taken up by mutants and wild type cells can no longer divide. The numerically obtained CDF is presented in the panel. We can see that the probability for the mutants to dominate the front becomes significant around size 10^4, which coincides with the size where the number of mutants starts deviating from the cubic law (the upper panel).

Next, we turn out attention to curves (e) and (f) of figure 4. They represent systems in the presence of death, where mutant advantage is manifested through increased division rate (e) and decreased death rate (f). We can see again that the curves follow a cubic law. A deviation from this law (and a slow-down of the growth as a function of total size) is also observed for larger sizes. The mechanism of this deviation is however somewhat different from the case of zero death. Figure S18(b, top panel) replots curve (f) of figure 4. Since cells die, the eventual outcome of all the simulations is the extinction of the (disadvantageous) wild type. The CDF of the colony size by which the wild type goes extinct is presented in the bottom panel of figure S18(b). The probability of wild type extinction becomes significant around size 2×10^4, where the number of mutants starts deviating from the cubic growth (see the top panel).

Finally, we compare curves (f) and (g) of figure 4. They represent systems with the same parameters (where the mutant advantage is manifested through a lowered death rate), except the cylinder width is $W = 100$ for curve (f) and it is $W = 1000$ for curve (g). Notice that the dashed black lines for the two curves differ by a factor of 100, representing the inverse square dependence of the number of mutants on the cylinder width, see formula (40) (there are 100 times fewer cells in the cylinder that is 10 times wider).

4.2.3 Growth on a circle (2D range expansion)

For simulations studying 2D range expansion, we started with 1 wild cell, and let the colony expand on a 2D grid for 800 time-steps, for. 2000 runs for each parameter combination. In each case, the numbers of wild type and mutant cells were averaged over all the runs that did not result in population extinction[1]. The results for several representative cases are presented in figure 6 of the main

[1]Extinction was not a problem in the cylindrical geometry, because the initial condition contained 100 or more cells, and very few runs resulted in extinction

text, where for each parameter combination the number of non-extinct runs is given in the figure caption.

There are 6 curves plotted in figure 5 of the main text. As in figure 4, cases (a,b) correspond to neutral mutants, and cases (c-f) to advantageous mutants; please note that all the division and death rates in curves (a-f) of figure 5 are identical to curves (a-f) of figure 4.

The black dashed lines for curves (a,b) in figure 5 have slope 3/2, as predicted for neutral mutant growth in a circle. As in the case of the cylinder, there are more mutants in the presence (b) than in the absence (a) of death.

The slope for curves (c-f) in figure 5 is 2, as predicted for the growth of advantageous mutants in a circle, see table 1. Again, we observe deviation fro the predicted power law for large system sizes. In the absence of cell death (cases (c,d)) this happens as the mutants become more likely to spread and occupy all the surface locations, blocking the wild type cells from divisions. In the presence of death, this deviation is associated with the increased likelihood of wild type extinction.