

UC San Diego

UC San Diego Previously Published Works

Title

Efficient mathematical methodology to determine multistep mutant burden in spatially growing cell populations

Permalink

<https://escholarship.org/uc/item/5gq252ds>

Journal

PNAS Nexus, 4(9)

ISSN

2752-6542

Authors

Komarova, Natalia L

Pritchard, Justin R

Wodarz, Dominik

Publication Date

2025-08-30

DOI

10.1093/pnasnexus/pgaf271

Copyright Information

This work is made available under the terms of a Creative Commons Attribution-NonCommercial License, available at <https://creativecommons.org/licenses/by-nc/4.0/>

Peer reviewed

Efficient mathematical methodology to determine multistep mutant burden in spatially growing cell populations

Natalia L. Komarova^a, Justin R. Pritchard^b and Dominik Wodarz^{c,*}

^aDepartment of Mathematics, University of California San Diego, La Jolla, CA 92093, USA

^bDepartment of Biomedical Engineering, The Pennsylvania State University, State College, PA 16801, USA

^cDepartment of Ecology, Behavior and Evolution, University of California San Diego, La Jolla, CA 92093, USA

*To whom correspondence should be addressed: Email: dwodarz@ucsd.edu

Edited By Dennis Discher

Abstract

The accurate computational prediction of mutant burden in spatially structured growing cell populations is a major goal both for basic evolutionary science, such as interpreting bacterial evolution studies, and for clinical applications, such as predicting the timing of drug resistance-induced cancer relapse for individual patients. Yet, this is currently not feasible for biologically realistic parameters, due to the inefficiency of computationally simulating stochastic mutant dynamics in large populations. Here, we fill this gap by deriving universal scaling laws that allow the straightforward prediction of the number of single-hit, double-hit, and multihit mutants as a function of wild-type population size in spatially expanding populations, in different spatial geometries, without the need to perform lengthy computer simulations. We demonstrate the applicability of this approach by reconciling different results from experimental evolution studies in bacteria that examine the role of gene amplifications for the rate of evolution.

Significance Statement

This work significantly advances the state of the art of evolutionary/population biology. It presents a derivation of previously unknown scaling laws, which relate the expected number of mutants to the total population size of cells in a spatially restricted expanding/invading population. The laws are derived for advantageous and neutral mutants, which contain an arbitrary number of mutational hits. The knowledge of mutant burden is very important, e.g. in oncology, microbiology, and other evolutionary contexts. Previously this type of estimate required extensive stochastic simulations and was not plausible for realistic population sizes. The laws can be applied to discover causal relationships in evolutionary processes.

Introduction

The accumulation of mutations in expanding cell populations determines the ability of the cells to adapt to changing environments, such as during drug therapy of bacterial infections or tumor cells, or when encountering naturally occurring selection barriers. An extensive theoretical understanding exists about the laws of mutant growth in exponentially expanding populations (1–8), but an equivalent understanding for spatially structured expanding populations remains to be fully worked out. Principles of mutant evolution in spatially structured populations have been elucidated by both computational and experimental approaches (9–24). Yet, challenges exist on the computational side. Agent-based and similar models, which are typically used to simulate spatial evolutionary processes, become computationally challenging at large population sizes (e.g. when studying bacterial or cancer cell populations), even if more advanced algorithms are used, such as the next reaction method (25),

tau-leaping methods (26), or related approaches (27–29). For well-mixed systems, hybrid simulation methods (30) have been used, in which only relatively small mutant colonies are treated stochastically, while larger subpopulations are simulated deterministically. For spatially explicit simulations, however, the development of such an approach would be highly complex and has not been worked out so far. An alternative way is to develop analytical insights, such as scaling laws, describing how the number of mutants in a growing system scales with the total population based on the system's dimensionality. A scaling law that connects the number of neutral, single-hit mutants with the colony size was derived previously for a 2D range expansion (9), and the theory was also generalized to neutral mutants in 3D systems, as well as advantageous mutants in 2D and 3D (19). Other types of insightful scaling laws that deal with smooth vs. rough fronts, allele surfing, and Muller's ratchet have been derived in (9, 10, 19, 31).

Competing Interest: The authors declare no competing interests.

Received: January 24, 2025. **Accepted:** July 28, 2025

© The Author(s) 2025. Published by Oxford University Press on behalf of National Academy of Sciences. This is an Open Access article distributed under the terms of the Creative Commons Attribution-NonCommercial License (<https://creativecommons.org/licenses/by-nc/4.0/>), which permits non-commercial re-use, distribution, and reproduction in any medium, provided the original work is properly cited. For commercial re-use, please contact reprints@oup.com for reprints and translation rights for reprints. All other permissions can be obtained through our RightsLink service via the Permissions link on the article page on our site—for further information please contact journals.permissions@oup.com.

Many biologically relevant evolutionary processes are more complex and involve multiple steps, for which such scaling laws currently do not exist. The aim of this study is to fill this gap. Our work builds on the previous approaches (9, 19) to derive scaling laws for 2D and 3D population growth for those more complex multistep evolutionary processes, where intermediate mutants with varying fitness are taken into account. We apply this methodology to predict the number of cells with a target point mutation at population size N and, in particular, investigate to what extent an intermediate cell type with an accelerated mutation rate can contribute to the acquisition of the target mutation. Specifically, we consider the role of gene amplifications for mutant evolution. On the one hand, the higher gene copy number can translate into a higher rate of point mutation generation in these genes; on the other hand, the generation and emergence of cells with gene amplifications requires extra rate-limiting evolutionary steps. Mathematics can help us resolve this tradeoff and predict the conditions under which gene amplifications enhance the emergence of point mutations in expanding colonies.

Using the growth laws derived here, we find that gene amplifications can only substantially enhance the emergence of target point mutations if population dynamics are characterized by spatial restrictions. This effect is most pronounced for 2D spatially restricted population growth and is still substantial in 3D. Interestingly, during growth of well-mixed populations (exponential growth), gene amplifications are not predicted to contribute to the emergence of cells with the target point mutation, unless the fitness advantage of the amplified cells is excessively large and/or the population size reaches biologically unrealistic levels. We use these insights to interpret experimental evolution studies and reconcile apparently contradictory experimental results. We also adapt this methodology to model the role of mismatch repair (MMR)-deficient cancer cells for mutant generation in expanding populations (32). Similar to cells with gene amplifications, MMR-deficient cells represent an intermediate state that accelerates the generation of point mutations. We again find that spatial population structure is important for MMR-deficient cancer cells to accelerate the emergence of cells with a target point mutation in growing populations.

Scaling laws for the expected mutant load

Consider the contact process where individuals on a lattice experience births and deaths. In the most basic scenario, the wild-type individuals can die at a rate d_0 or attempt division at a rate r_0 , whereby they place the offspring in a randomly chosen location in its neighborhood (for the purposes of this paper, we considered the Moore neighborhood consisting of eight locations). The division is aborted if the location is full, and it is successful if the location is empty, thus reflecting density dependence of the effective division rate. In the case of successful division, the wild-type cell is placed in the chosen spot with probability $1 - u_1$, while with probability u_1 the newly created cell is “mutant,” that is, it may have different kinetic parameters compared with the parent (the division rate r_1 and the death rate d_1). We will assume that the mutant offspring cells are either neutral (that is, $r_1 = r_0$ and $d_1 = d_0$) or advantageous with respect to the parent cells. We will characterize the relative fitness of the mutants by means of a nonnegative selection coefficient, s_1 , assuming that $r_1 = (1 + s_1)r_0$ and $d_1 = d_0$. In a one-hit mutant scenario, mutant cells do not experience further genetic changes. Otherwise, we may consider subsequent mutations, which can be neutral or advantageous with respect to the type that produced them (see the schematic in Fig. 1a). We will always

assume that the fitness advantage is associated with an increase in the division rate and keep the death rate of all types the same.

A typical simulation is started from a single wild-type cell (placed in the center of the grid in the spatial setting). As the simulation proceeds, the population expands according to surface growth (see Fig. 1b for the example of a 2D surface growth with $n = 3$ consecutive mutations). During cell divisions, mutated cells are generated with their corresponding rates, and these undergo range expansions that either continue to grow or become enclosed by competing cells over time.

Computer simulations as well as earlier work suggested the existence of the so-called power laws that guide the emergence and growth of mutants in spatially expanding populations (9, 19). They also clearly demonstrate the need for a more general approach, which would allow for a systematic exploration of the evolutionary dynamics under a wide range of parameters and in the presence of multiple mutations. Therefore, we turned to developing scaling laws that govern mutant growth in expanding populations. Note that in this paper we only focus on mutations that are neutral or advantageous. The scaling laws for disadvantageous mutations are simple: their abundance in the population of size N scales with N (19), because they are generated by divisions in the existing population and form “bubbles” of a finite size (9). Therefore, here we turn to neutral and advantageous mutants that tend to form “wedges.” In the context of this study, we say that the number of mutant cells increases as a power law if it grows proportionally to t^x , or equivalently, R^x , where t is time, R is the whole colony’s radius, and x is the power (which, as we show below, depends on factors such as the number of mutations, system’s dimensionality, and whether mutations are advantageous).

Scaling laws for single mutations

Consider the simplest situation described by the selection-mutation diagram: $Y_0 \xrightarrow{u_1} Y_1$, where Y_0 denotes the wild type, Y_1 denotes the population of single mutants (with relative fitness $(1 + s_1)$), and the mutation rate is given by u_1 . Back mutations are not included here and below because they do not make a significant contribution to the scaling laws derived here (Section 1.6 of the Supplementary Material). We will compare three types of spatial growth: a flat front expansion (a “1D” growth), a 2D range expansion, where a colony grows in a circular shape (Fig. 2a), and a 3D range expansion (a spherical colony, not shown). We will use variable R to characterize the colony size. Denote by R the linear colony size (equivalent to the length in the direction of the 1D growth and the radius in 2D and 3D), where distance is measured in terms of cell diameters. If the total number of cells is denoted by Y_0 , we have

$$(1D) \ Y_0 \propto WR, \quad (2D) \ Y_0 \propto \pi R^2, \quad (3D) \ Y_0 \propto \frac{4}{3}\pi R^3, \quad (1)$$

where W stands for the colony width in 1D. In particular, in the absence of cell death, the density of cells is one and the proportionality sign becomes “=.” Assuming so-called surface growth (33), where only cell divisions on the surface contribute to expansion, the mean temporal growth of the population can be described by the following ordinary differential equations (ODEs):

$$(1D) \ \dot{Y}_0 = b_1 W, \quad (2D) \ \dot{Y}_0 = b_2 Y_0^{1/2}, \quad (3D) \ \dot{Y}_0 = b_3 Y_0^{2/3}, \quad (2)$$

where b_i with $1 \leq i \leq 3$ are constants that include information about the growth rate of cells and also depend on the dimensionality of the system (see Eqs. 1 and 2 of Supplementary Material for

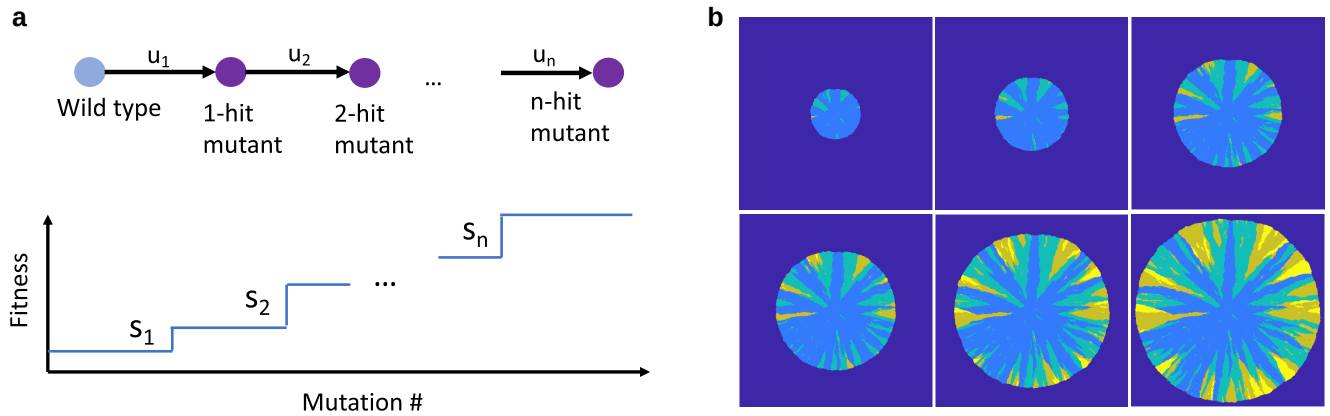


Fig. 1. The contact process with mutations. a) A schematic showing an example of a linear mutation accumulation process. The relative fitness of consecutive mutations is shown as a stepwise function (where each step may be positive or zero). b) Example of a spatial configuration of a 2D expanding and evolving colony of cells, a typical simulation representing an example with $n = 3$ mutations. The individual graphs represent different snapshots in time, going from left to right top row, followed by left to right bottom row. Dark blue represents empty space, and the other colors correspond to different cell types: blue to wild types, green to single-hit mutants, brown to double-hit mutants, and yellow to triple-hit mutants. The parameters are: the basic division rate $r_0 = 0.9$, the death rates of all the types $d_i = 0.0005$, mutation rates $u_i = 2.5 \times 10^{-4}$, and selection coefficients relative to the previous type $s_i = 0.01$ for $i = 1, 2, 3$.

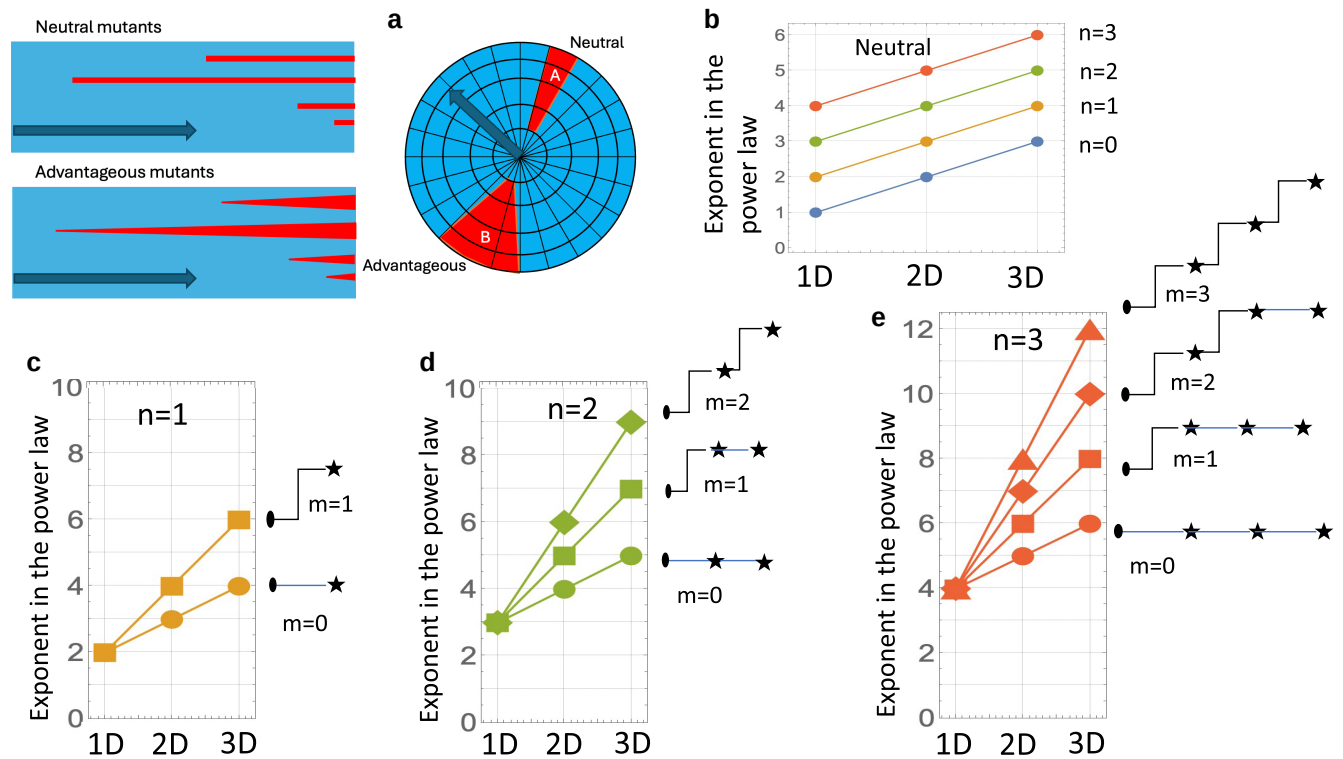


Fig. 2. Mutant scaling laws (population $\sim t^x$). a) Schematics showing generation and spread of neutral and advantageous mutant. Light blue represents the wild-type cells and red the mutant cells. Top left: Flat front (1D) geometry with neutral mutations; the corresponding clones retain a constant width. Bottom left: Flat front with advantageous mutations giving rise to expanding colonies. Right: The 2D range expansion geometry. Mutant A is neutral and its expected domain follows the radial lines. Mutant B is advantageous and its domain boundaries cross radial lines, increasing the arc angle of the mutant region. b-e) Scaling laws for mutant growth (see Eq. 5 of the main text and Eq. 13 of the Supplementary Material). b) Neutral mutants. For each spatial dimensionality (1D, 2D, 3D), the exponent, x , in the power law (t^x) for each type of cells is shown for the wild type (blue), one-hit mutants (yellow), two-hit mutants (green), and three-hit mutants (red). c and d) The scaling laws in the presence of advantageous mutational steps (m advantageous hits out of n total hits). Each panel shows the exponent in the power law of growth of the population of n -hit mutants in 1D, 2D, and 3D geometries. A cell is considered to have an advantage if it has an increase of fitness with respect to the type that generates it (not just the wild type). For the one- (c), two- (d), and three-hit (e) mutant scenarios, all the mutants could be neutral ($m = 0$, circles), one of them could be advantageous ($m = 1$, squares), two could be advantageous ($m = 2$, diamonds), or three could be advantageous ($m = 3$, triangles). Examples of relative fitness changes are shown to the right of each case by means of stepwise shapes, where the initial oval represents the wild type and stars represent mutational hits. According to the theory, the order in which mutations occur does not matter for the scaling laws. In (b-d), only the exponent in the power law is shown; the coefficient depends on the product of all the mutation rates.

details). Consequently, the total number of cells grows according to the following power laws:

$$(1D) Y_0 \propto t^1, \quad (2D) Y_0 \propto t^2, \quad (3D) Y_0 \propto t^3.$$

In our model, the linear colony size always grows linearly with time, $R \propto t$, which can be seen, e.g. by comparing the temporal growth laws above with the volume formulas in Eq. 1. Consequently, we can use t and R interchangeably when talking about the scaling laws.

To include mutant generation, we will assume that mutations are relatively infrequent, such that a single mutation typically gives rise to a clone that does not merge with other clones, and the total mutant number is significantly smaller than the colony size (see Section 2.1 of the Supplementary Material for model applicability conditions). As a result, it is possible to study mutant clones in isolation. The drawing in Fig. 2a(left) presents a schematic of mutant generation in a colony expanding to the right as a flat front. The wild-type cells are shown in light blue (the background color) and mutants are shown as thin rectangles or wedges in red. Mutants are generated randomly by the dividing cells, and as the colony expands to the right, so do the mutant clones. If the mutants are assumed to be neutral, on average each clone will remain at a constant width (the top panel). If the mutants are advantageous, their clones will on average become increasingly wider (the bottom panel). Figure 2a(right) shows a similar picture for a 2D range expansion. A neutral clone (A) will retain its relative width and grow parallel to radial lines, while an advantageous clone (B) will expand and cross the radial lines.

We assume that mutants are generated by the wild-type cells at the edge of expansion. This is always valid in the absence of cell death. For neutral mutants, this is also true in the presence of cell death. Indeed, in the core of an expanding colony, cells undergo a birth and death process, but births compensate for deaths, resulting in a quasi-equilibrium state with a cell-density fluctuating about a steady state. It is only at the colony's expanding front where births exceed death and lead to an increase of cell numbers. For advantageous mutants, this modeling assumption is a good approximation as long as the advantage, $s \ll 1$ and $d \ll r$. Denote by $n^s(r_1, r_2)$ the expected size of a mutant clone with relative fitness advantage s , that was created at colony size r_1 , and measured at size r_2 . Further, denote by $M_0(r)$ the size of the advancing front, through which the growth occurs, e.g. $M_0(r) \propto 2\pi r$ in 2D, and $M_0(r) \propto 4\pi r^2$ in 3D. Then, the expected number of mutants with fitness advantage s_1 , in a colony of size R can be approximated by: $Y_1^{s_1}(R) = \int_0^R u_1 M_0(r) n^{s_1}(r, R) dr$, where u_1 is the rate at which mutations are generated.

The expression for the clone size depends on the system dimensionality as well as the mutant fitness relative to the wild type, which is given by $(1 + s_1)$ with $s_1 \geq 0$. It is derived in the supplement for dimensionalities $D = 1, 2, 3$. As a result, we obtain a scaling law for the number of mutants in a colony of linear size R , with $s_1 \geq 0$:

$$Y_1^{s_1}(R) = 2^{D-1} \pi u_1 (D-1)! \sum_{k=0}^{D-1} \frac{R^{2D-k} s_1^{D-1-k}}{k!(D-k)!(2D-k)}. \quad (3)$$

In particular, for neutral mutants ($s_1 = 0$), the only nonzero term on the right corresponds to $k = D - 1$, and we have, e.g.

$$(2D) Y_1^0(R) = \frac{2}{3} \pi u_1 R^3, \quad (3D) Y_1^0(R) = \pi u_1 R^4$$

Table 1 shows the power of the leading term for different species in different dimensionalities (in terms of time, t ; the power in terms of R is the same). For completeness, this table also includes

Table 1. The leading power of time in the growth laws of different populations in expanding colonies with single mutations, in 1D, 2D, and 3D.

Single-mutant scaling laws	1D	2D	3D
Wild type or disadvantageous mutant ($s_1 < 0$)	t^1	t^2	t^3
Neutral mutant ($s_1 = 0$)	t^2	t^3	t^4
Advantageous mutant ($s_1 > 0$)	t^2	t^4	t^6

Only the power laws are provided. The coefficients in front of these terms are not included in the table, but they are proportional to the mutation rates (as also follows from the full mathematical form of the scaling laws that are derived here).

disadvantageous mutants. A more detailed discussion of laws for disadvantageous mutants can be found in (19); here, it suffices to say that each disadvantageous mutant is assumed to give rise to a clone of a constant size, $n^s(r_1, r_2) = \text{const}$, which leads to the results presented in Table 1. Note that, for example, in 2D, while the total colony size (Y_0) increases as t^2 , the population of single neutral mutants grows as a third power of time (i.e. faster). This is because both production and subsequent growth of mutant clones are taken into account, and in the case of advantageous mutants, the growth is even faster (resulting in the fourth power of t in 2D). As the population grows, this leads to an eventual breakdown of the laws when the smallness of the mutation rate (which multiplies the mutant colony size) is no longer able to compensate for a larger power of time (see also Section 2.1 of the Supplementary Material).

Figure 3 illustrates the difference between the spatial growth laws derived here and the case of exponential growth by using a 2D example. In this figure, we show a simulation of two colonies: an exponentially growing colony (Fig. 3a) and a 2D spatially growing colony (Fig. 3b); please note that this represents the expected numbers of cells, which can be < 1 . The wild-type cells are shown in blue, and mutants (that are created at the same rate of 10^{-7} per cell division and have the same selection coefficient of 5%) are shown in yellow. The green horizontal line marks the population size (10^8) at which we compare the numbers of mutants. The black dashed vertical lines mark the time when each colony reaches the reference size (one can see that it takes a much longer time for a spatially growing colony to reach that size). We compare the number of mutants (marked by black dashed horizontal lines) when each colony reaches the reference size. For the given parameters, we expect to have (on average) 302 mutant cells in the exponentially growing colony and 1.0×10^7 mutants in the spatial colony.

As previously described (9, 19), spatial constraints modify the growth of colonies and the relative abundance of mutants, which in itself is a counterintuitive and often underappreciated fact. In a mixed culture, the growth is exponential because cells do not “feel” each other's presence and do not engage in direct competition for space. The number of mutants created during expansion is relatively small except for very rare “jackpot” events, when a mutant is created early on and gives rise to a huge colony. Under spatial growth, such “jackpot” events are much more common: a mutant that arises at the surface of an expanding front will block the wild-type cells from expanding in that direction, making a whole region of the expanding colony a mutant clone (seen as a “wedge” in Fig. 2a, the case of 2D growth). As a result, the relative abundance of mutants in a spatially growing colony (given that the rest of the parameters are the same) is massively higher compared with an exponentially growing colony. Next, we generalize and expand these results to more mutations.

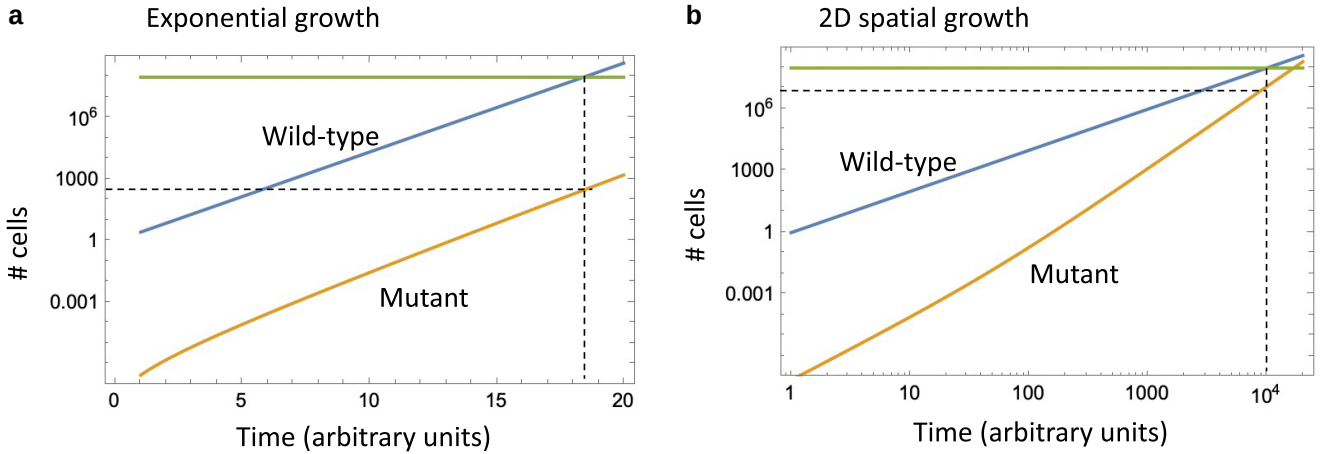


Fig. 3. A comparison of mutant accumulation in an exponentially growing colony (a) and a 2D spatially growing colony (b), a simulation. Parameters are $s_1 = 0.05$, $u_1 = 10^{-7}$. Blue: wild type, orange: mutant, green: the reference size of 10^8 cells. In each panel, the black dashed vertical line marks the time when the colony reaches the reference size. The black dashed vertical line shows the mutant population size when the colony reaches the reference size.

Scaling laws for double mutations and beyond

Next, we extend the theory to more than one mutation. To develop the growth laws of double-hit mutants, assume the following selection-mutation diagram: $Y_0 \xrightarrow{u_1} Y_1 \xrightarrow{u_2} Y_2$, where, as before, Y_0 denotes the wild type, Y_1 denotes single mutants (with relative fitness $(1 + s_1)$), and Y_2 denotes double mutants (with relative fitness $(1 + s_2)$). For now, we fix the order of events and assume that type Y_2 is generated by a mutation in type Y_1 . Note that here and below, for each consecutive mutation, we measure the resulting fitness with respect to the type that hosts the mutation and not with respect to the wild type. The rates of mutation are given by u_1 and u_2 for the first and the second hit, respectively. As in the case of single mutants, back mutations do not make a significant contribution to the scaling laws (Section 1.6 of the [Supplementary Material](#)).

Using an argument similar to the one for single-hit mutants, we assume that double-hit mutants are generated by the single-hit mutants at the edge of expansion (see [Supplementary Material](#) for details). The expected number of double-hit mutants is given by: $Y_2^{s_1, s_2}(R) = \int_0^R u_2 M_1^{s_1}(r) n^{s_2}(r, R) dr$, where $M_1^{s_1}(r)$ is the number of single mutants of fitness advantage s_1 at the edge of expansion for a colony of linear size r , and $n^{s_2}(r, R)$ is the size of the double-mutant colony of fitness advantage s_2 that is generated at size r and measured at size R . We obtain the following expression for the expected number of double mutants:

$$Y_2^{s_1, s_2}(R) = u_1 u_2 \pi 2^{D-1} ((D-1)!)^2 \sum_{k=0}^{D-1} \sum_{m=0}^{D-1} \frac{R^{3D-k-m} s_1^{D-1-k} s_2^{D-1-m}}{k! m! (2D-k-m)! (3D-k-m)!}. \quad (4)$$

For example, when both mutations are neutral ($s_1 = s_2 = 0$), the only nonzero contribution comes from the term with $k = m = D - 1$, and we obtain

$$(2D) \quad Y_2^{0,0}(R) = \frac{1}{4} \pi u_1 u_2 R^4, \quad (3D) \quad Y_2^{0,0}(R) = \frac{2}{5} \pi u_1 u_2 R^5.$$

In the [Supplementary Material](#), we provide an explicit formula (Eq. 13 of the [Supplementary Material](#)) for any number of consecutive mutations, some or all of which can be advantageous (Fig. 1a). There are several consequences of the scaling laws that are derived, which we list here.

1. For neutral mutants, each hit adds one power of R . A population of n -hit neutral mutants under D dimensional range expansion grows as R^{D+n} , where R is the whole colony radius; the power

scaling with time is the same, t^{D+n} . This is shown in Fig. 2b for one-, two-, and three-hit neutral mutants. The general formula is given in Eq. 14 of the [Supplementary Material](#).

2. Advantageous mutations each add two powers of R , and the order of mutations is not important. Suppose that among the n sequential hits, m hits confer a nonzero advantage with respect to the previous cell type, and the rest of the mutations do not alter the fitness properties of cells that harbor those mutations. Surprisingly, the order in which these advantageous mutations are acquired does not influence the scaling law for the number of n -hit mutants, which grow as

$$M \sim R^{D+n+m(D-1)}; \quad (5)$$

the power scaling with time is the same, $t^{D+n+m(D-1)}$. The exponents of the scaling law for different values of n (the number of hits) and m (the number of advantageous hits) are presented in Fig. 2c–e. The leading term in the expression where m out of n mutants are advantageous and the rest are neutral is given by Eq. 15 of the [Supplementary Material](#).

3. Advantageous mutants behave as if they are neutral initially, and their growth accelerates when R exceeds the inverse selection coefficient. Suppose that individuals in a population of n -hit mutants contain m mutations that confer a selective advantage to cells (relative to the cell type that acquired the mutation) and denote the corresponding selection coefficient values by $s_1, \dots, s_m$ (where they are listed in a nondecreasing order). If they all have roughly the same order of magnitude (s), then the n -hit mutant population behaves as if neutral, until the colony size reaches $R \sim 1/s$, at which time it switches to the growth specified in Eq. 5. In the case where selective advantage coefficients are different, the growth law increases the power by one at sizes $1/s_m, 1/s_{m-1}, \dots, 1/s_1$.

Note that the last point is similar in spirit to the concept that a beneficial mutation may behave neutrally when $N_e s \ll 1$, where N_e is the effective population size and s is the selection coefficient.

Using this methodology, one can investigate the relative contributions of different pathways to the evolutionary processes in growing populations.

Using derived growth laws to obtain insights into evolutionary processes in spatially growing populations

Here, we show how the analytical expressions for mutant numbers derived above can be used to gain insights into complex evolutionary processes in spatially growing populations. In particular, we ask under what conditions an intermediate population with an increased rate of mutation accumulation can enhance the presence of cells carrying a target point mutation. We first consider gene duplications/amplifications and then also apply this theory to MMR-deficient cancer cells as the intermediate cell population.

Spatially structured growth is required for gene duplications to significantly accelerate mutant production

We are interested in the rate of production of a specific mutation (which we will refer to as a “target mutation”). Consider an expanding population that contains the following types (Fig. 4a): (i) wild-type cells (Z_{wt}), which have not yet acquired any mutation; (ii) mutant cells (Z_m), which have acquired the target mutation due to a replication error of the wild-type population; (iii) accelerated cells (Z_a), which have acquired a gene amplification in the gene under consideration, thus effectively having a higher rate at which the target mutation can be generated; cells containing this change are assumed to have selective advantage s compared with the wild type; and (iv) mutated accelerated cells ($Z_{am} + Z_{ma}$), which are cells with the higher mutation rate that have acquired the target mutation. Upon division, wild-type cells are assumed to give rise to the target mutation (and hence to cell type Z_m) with a rate μ . They can also acquire the accelerated phenotype with a rate p , giving rise to type Z_a cells. Accelerated cells (Z_a) are assumed to produce the target mutation with an elevated rate $\mu_A = a\mu$, giving rise to population Z_{am} . For a cell with gene duplication $a = 2$, and larger values of a are assumed for higher copy numbers. In addition, mutated cells (Z_m) can acquire the accelerated phenotype with a rate p , type Z_{ma} . Note that the path

from Z_{wt} to Z_{am} consists of two steps with relative advantage values $s_1 = s, s_2 = 0$, and the path from Z_{wt} to Z_{ma} consists of two steps with relative advantage values $s_1 = 0, s_2 = s$.

To quantify the role of cells with gene duplications for target mutant emergence, we compare the following populations (Fig. 4): on the one hand, population Z_m is the total number of mutants generated in the absence of mutation acceleration; on the other hand, the sum of the populations $Z_{am} + Z_{ma}$ represents the number of mutants that are associated with the possibility of accelerating the mutational process. Here, cells Z_{am} first undergo a mutation that allows for an accelerated mutation rate, and then acquire the target mutation at a higher rate, μ_A . Cells Z_{ma} are created by first acquiring the target mutation and then developing the characteristics of accelerated cells by a second hit (generally, the Z_{ma} population is smaller cells Z_{am} because of the acceleration in the mutation rate that is experienced after the first hit). Figure S1 shows a series of snapshots of a typical simulation, where on the dark background (empty space), blue represents wild-type cells (Z_{wt}), cyan is target mutants (Z_m), brown is amplified cells (Z_a), and yellow amplified mutant cells (Z_{am}).

Figure 4c compares stochastic numerical simulations with the theoretical results for these populations, where for Z_m , we used $Y_1^{s_1}(R)$ in formula (3) with $u_1 = \mu$ and $s_1 = s$; for Z_{ma} , we used $Y_2^{s_1, s_2}(R)$ in formula (4) with $u_1 = \mu, u_2 = p, s_1 = 0, s_2 = s$; and for Z_{am} , we used $Y_2^{s_1, s_2}(R)$ with $u_1 = p, u_2 = \mu_A, s_1 = s, s_2 = 0$. We can see that the theoretical approximation works well (see further information on validation and applicability in the [Supplementary Material](#)).

Using the derived growth laws to predict mutant dynamics for population sizes that are too large to track by computer simulation, we can predict at what colony sizes the cells with duplicated genes make a substantial contribution to target mutant emergence, compared with the number of mutations generated directly by a single point mutation hit from the wild type. Let us define the quantity $Q = (Z_{am} + Z_{ma})/Z_m$, as the relative contribution of duplicated cells to the total number of mutants created. This quantity has the following straightforward interpretation. Suppose that in the absence of duplicated cells ($p = 0$), the number of cells

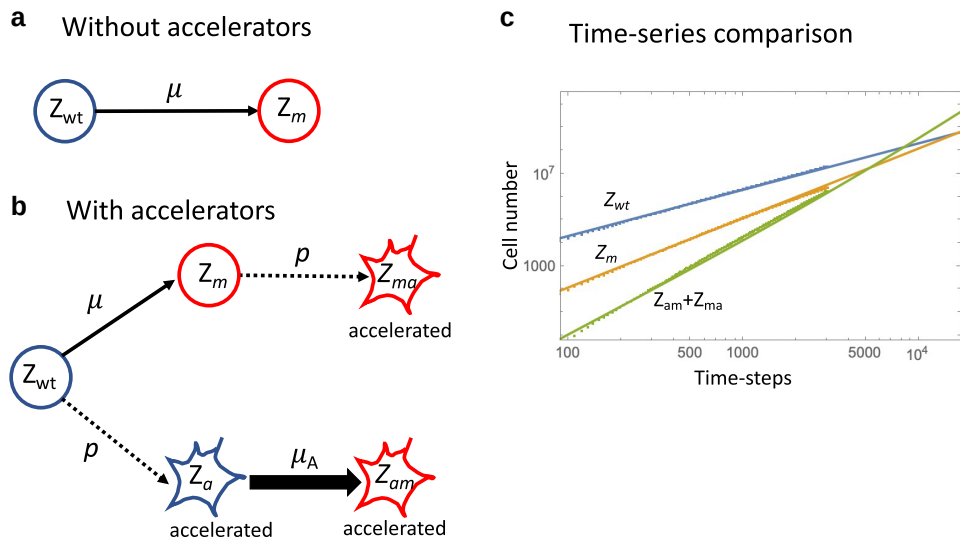


Fig. 4. Comparing two pathways to a mutation. a, b) A schematic showing the pathways leading to the generation of a mutation: a) in the absence and b) in the presence of the accelerated phenotype. The accelerated phenotype is denoted by a star shape, and the mutant of interest is colored red. The accelerated mutation rate is denoted by μ_A . c) The number of mutants generated directly (Z_m , shown in orange) is plotted together with the number of double-hit mutants generated in the presence of an accelerated type ($Z_{am} + Z_{ma}$, green) and the number of wild-type cells (Z_{wt} , blue) in an expanding 2D colony. Simulations (average of 3,612 independent simulations): dots; theory: lines. Parameters are $r_0 = 0.95$, $d_0 = 0$, $s = 0.01$, $\mu = p = 10^{-4}$, $\mu_A = 2\mu$. For details, see [Section 2.2](#) of the [Supplementary Material](#).

that contain the target mutation in a colony of a given size is Z_m . Then, in the presence of duplicated cells ($p > 0$), the colony of the same size will contain $(1 + Q)Z_m$ cells with the target mutation. An alternative measure of the contribution of duplicated cells is $\tilde{Q} = Z_{am}/Z_m$, which simply compares the number of cells that acquired the mutation after duplication with the number of cells that acquired the mutation directly from the wild type. Since $Z_{am} > Z_{ma}$, we have $\frac{1}{2}Q < \tilde{Q} < Q$, which means that the two measures are similar. If the contribution of duplications in a system is predicted to be significant by measure Q , it is also significant by measure $\tilde{Q}$. If it is negligible by measure Q , it is also negligible by measure $\tilde{Q}$.

Assuming that cells with the gene duplication have twice the mutation rate compared with cells without ($\mu_A = 2\mu$), Fig. 5 compares the behavior of spatial mutant colonies in the cases of 2D and 3D range expansion with a mass-action (exponential growth) system; the methodology for the latter case is standard and is provided in the [Supplementary Material](#). Plotted is the relative contribution of the duplicated cells to the total number of mutants created, $Q = (Z_{am} + Z_{ma})/Z_m$, for different colony sizes (the horizontal axes). We can see that the relative contribution increases with the colony size. This happens because the net growth rate of the double mutants is faster; in particular, in the spatial systems, it corresponds to an increased exponent in the power law. We also observe that the relative contribution of the duplicated cells is always the largest in 2D (blue), followed by 3D (orange), and both of these are substantially larger than that for the exponentially growing colony (green).

The striking result is the vast difference in the contribution of the duplicated cells between the spatial and nonspatial (exponentially growing) systems. In the case of the parameters used in Fig. 5, the increase in the number of target mutants due to gene duplication in the nonspatial model never reaches 1%, even for extremely large populations and assuming that gene duplication happens at a rate that exceeds the base mutation rate a hundred-fold and confers a 10% advantage to the cell (see the bottom right panel). In well-mixed populations, gene duplications can only contribute to the accumulation of target point mutations if the fitness of duplicated cells is very large, upwards of 40% (Fig. S3). On the other hand, in spatial systems, the increase in the number of target mutants due to duplications becomes significant for much smaller population sizes and much smaller selective advantages of duplicated cells. If gene duplications occur much more often (100-fold in Fig. 5, bottom) than point mutations (34), duplicated cells make a substantial contribution to target mutant evolution even if the duplicated cells are neutral with respect to the wild-type cells in 2D growth. In the [Supplementary Material](#), we present additional graphs where we vary the base mutation rate over a wide range, revealing the same patterns.

Further amplifications can significantly contribute to mutant production in spatial populations

Beyond duplications, cells can further amplify gene copy number. The schematic of Fig. 6a shows a cascade of amplifications (gene duplications), where each duplication event leads to an increase in the accelerated mutation rate. Here, type Z_1 corresponds to type Z_a of Fig. 1, while types Z_2 , Z_3 , etc. represent cells that have experienced 2, 3, etc. consecutive duplication events. Type Z_{1m} contains cells that have undergone a single duplication event and also acquired a target mutation (corresponding to the sum $Z_{am} + Z_{ma}$ in Fig. 1b), while types Z_{2m} , Z_{3m} , etc. are

types with multiple duplication events that also harbor the target mutation.

It is interesting to determine whether or not further duplication events can significantly increase the production of target mutants, beyond the type Z_{1m} . Applying the scaling laws derived here (see the [Supplementary Material](#) for details), we show that in 2D and 3D systems, further duplication events could make a difference for biologically realistic parameter ranges if each subsequent event induces additional fitness advantage (34) (Fig. S8).

Spatial population structure is required for the inactivation of an MMR gene in somatic cells (a two-hit event) to contribute significantly to mutant production in expanding tumors

MMR deficiency (leading to small-scale genetic instability) is a biologically distinct intermediate cell state that can accelerate the rate of target mutant production in somatic cells, such as colorectal cancer cells in humans (Fig. 6b). In contrast to gene duplications, the generation of the MMR-deficient phenotype in eukaryotic cells requires two mutational events in most cases, because the MMR mechanism needs to be inactivated in cells with diploid genomes. In this setting, type Z_{12} is a cell with both copies of an MMR gene inactivated. It can be produced from the wild type by means of two consecutive inactivation events (through type Z_1 or Z_2) and is characterized by a significantly (orders of magnitude) increased rate at which the target mutation can be generated. Hence, although it takes more steps to generate the MMR-deficient phenotype compared with a gene duplication, the gain in the mutation rate is a lot higher. The question is whether, for realistic population sizes, this mechanism can contribute to mutant production in expanding populations, such as in colorectal tumors, assuming that the MMR-deficient cells are generated in growing tumors (rather than in non-neoplastic tissue as a precursor to tumor initiation; see [Discussion](#)).

Using our scaling laws (see [Supplementary Material](#) for details), we can show that, both in 2D and 3D systems, the contribution of MMR-deficient cells can be similar to or larger than direct target mutations, under the assumption that the MMR-deficient phenotype has some degree of selective advantage. This assumption is supported by experiments showing that MMR-deficient cells have compromised apoptosis responses, which outweighs the fitness cost associated with the accumulation of deleterious mutations (32) (see [Discussion](#) for more details). In contrast, in the nonspatial, exponentially growing systems, the effect of these longer pathways (via MMR deficiency) is negligible (Fig. S7).

Discussion

Spatially explicit stochastic computer simulations that involve evolutionary processes are computationally very “expensive” and are difficult to perform under realistic population sizes. In the case of spatially expanding populations that produce neutral or advantageous mutants, we developed an approach to get around this limitation that is based on newly derived scaling laws for these populations. Using these laws, we can study multi-step evolutionary processes that involve neutral and/or advantageous mutations and quantify the role played by various processes in the evolutionary transformation experienced by populations. In particular, this methodology allows reasoning about the relative likelihood of different pathways and suggests which events may be causal, given the patterns that are observed. This can be done under realistic assumptions about the mutation rates

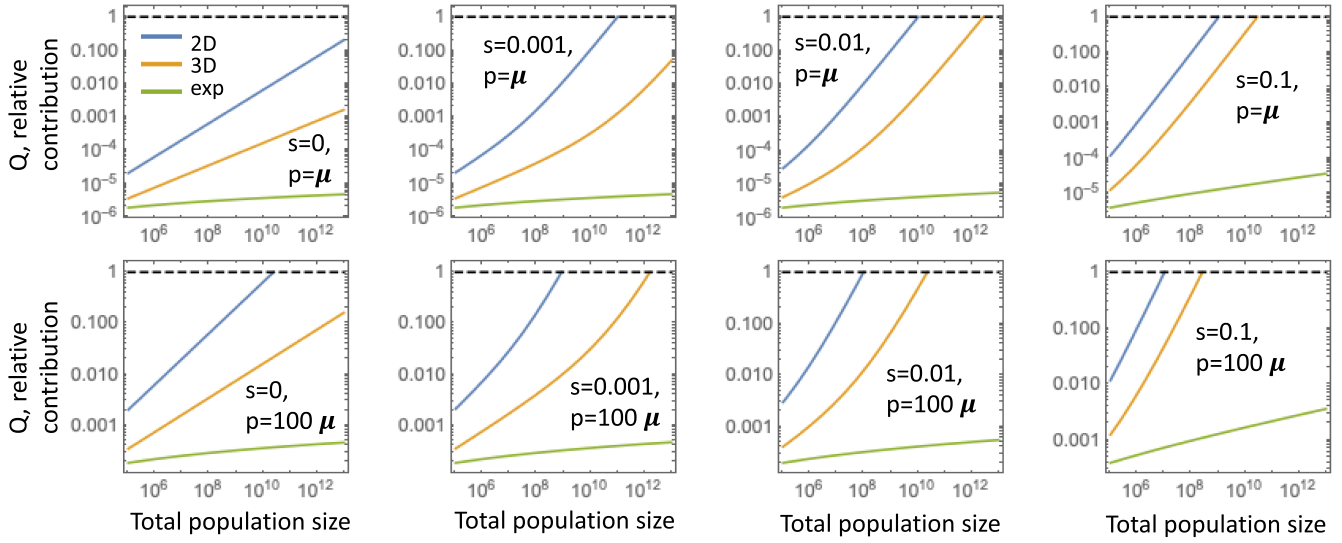


Fig. 5. Gene duplication: the relative contribution of the accelerated type to the total number of mutants, $Q = (Z_{am} + Z_{ma})/Z_m$ (as predicted by the model), is plotted as a function of the total population size. The three lines in each panel correspond to the 2D (blue), 3D (orange), and mass-action (green) systems. We assume that the accelerated mutation rate is $\mu_A = 2\mu$ (duplications). In the top row, the duplications are generated at rate $p = \mu$, and in the bottom row $p = 100\mu$. The fitness parameter, s , increases from left to right as $s = 0$, $s = 0.001$, $s = 0.01$, and $s = 0.1$. The dashed horizontal line represents $Q = 1$. $\mu_A = 10^{-7}$.

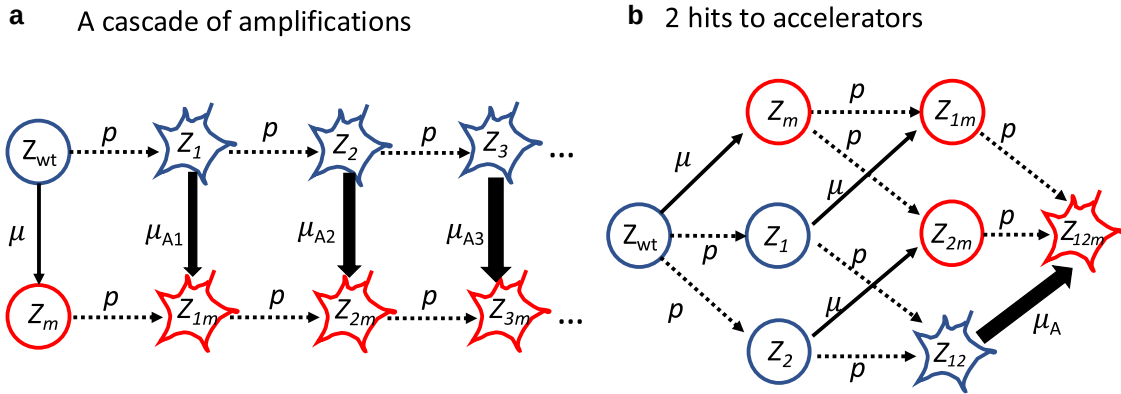


Fig. 6. Other selection-mutation networks with mutation accelerations. The accelerated phenotype is denoted by a star shape, and the mutant of interest is colored red. a) A cascade of amplifications; type Z_k with $k = 1, 2, \dots$ has experienced k gene duplication events in the target gene; type Z_{km} contains cells with a target mutation that have undergone k gene duplication events in the target gene. The increasingly accelerated mutation rates are denoted by μ_{A1} , μ_{A2} , $\dots$. b) Two hits to acceleration. Types Z_1 and Z_2 have the first and the second copy of the MMR gene inactivated, respectively, while the accelerated type Z_{12} has both copies inactivated. Types Z_{1m} , Z_{2m} , and Z_{12m} are the same as types Z_1 and Z_2 , and Z_{12} except they also contain the target mutation.

and population sizes, which is difficult to do using traditional methods of stochastic agent-based modeling.

The scaling laws derived here have the form of

$$(\text{Mutant population size}) \sim R^x$$

(or equivalently, $\sim t^x$), where R is the linear size of the whole colony (measured in cell diameters), t is time, and x is the power that we calculated in this paper for each case. The scaling laws predict the expected size of a mutant population created by an expanding colony of cells/organisms, both in 2D and 3D range expansion. Results include one-hit mutants, as well as mutants that are created by a longer chain of mutation events, which may include both neutral and advantageous hits. The simple rule of thumb is that each neutral mutation results in a multiplier (uR), i.e. adds a power of R (or time) to the basic growth law of the original population, and each advantageous mutation results in a multiplier (uR^2), i.e. adds two powers of R (or time). Here, we denoted by u a typical mutation rate. These laws hold roughly on spatial scales,

R , with $uR \ll 1$ for neutral mutants and with $uR^2 \ll 1$ for advantageous mutants.

These theoretical results were applied to study the role of gene duplications/amplifications for target mutant emergence in growing populations and further expanded to describe the contribution of MMR-deficient cancer cells to mutant production. An emerging theme was that the spatial structure of the growing population might be crucial for intermediate cell genotypes with elevated mutation rates to contribute substantially to target mutant generation. The effect was strongest for 2D spatial population structures, still substantial for 3D spatial population structures, but largely negligible for well-mixed populations. The main finding is that intermediate states with elevated point mutation rates (e.g. cells with gene amplifications or MMR-deficient cells) require spatial structure to substantially contribute to the emergence of point mutations. This is a direct consequence of the mutant growth laws derived here. For cells with elevated mutation rates to contribute to the emergence of target point mutations, the fast mutators first

have to be generated and then gain in abundance relative to the wild-type cells. This is most challenging in nonspatial (exponential) growth, even if the mutation is advantageous. The reason is that no competition for space occurs between cells during exponential growth. Hence, for cells with elevated mutation rates to gain in abundance relative to wild-type cells, they either need to have a massive fitness advantage, or one has to wait until the population size has grown to levels that are biologically not realistic, even for cell and microbial populations. Absent those unrealistic scenarios, it is more likely that the wild type directly gives rise to the cells with the target point mutation. In contrast, with spatially structured population growth, competition for space occurs between wild-type and mutant cells. Mutants arising at the edge of an expanding colony grow outwards and block the wild types, preventing them from expansion. Now, a fitness advantage of an intermediate cell with elevated mutation rates will allow this cell type to substantially gain in abundance relative to the wild-type cells within a much shorter growth period. Mutants' competitive advantage enables them to utilize spatial constraints and expand, even if the selective advantage is relatively modest. This leads to the scenario that cells with elevated mutation rates contribute substantially to the target mutant generation, compared with direct generation by wild-type cells.

As with all mathematical and computational modeling work, results depend on the underlying assumptions, which can be characterized by some uncertainty. In the algorithm implemented here and elsewhere (9, 19), a cell that is marked for division randomly picks a spot for its offspring among all of its neighbors, but the division is aborted if the chosen spot is full. This algorithm reflects local density dependence. Alternatively, less stringent local density dependence can be assumed. This may happen when, for example, offspring cells can squeeze in among other cells and push the existing neighbors away. This comprises an interesting (and underexplored) mathematical problem. In this case, the growth laws are expected to change and fall somewhere between "mass action" (exponential) and surface laws derived here. Even the simple spread/colonization dynamics of a single type (which in our model grows as t^D) would change (accelerate). The specific functional form will depend on the exact model assumptions. For example, if the reproducing cell chooses a random location among empty spots, one needs to specify the neighborhood size where these spots can be located (only direct neighbors or within a certain radius of the reproducing cell) and also what happens if there are no empty spots in the neighborhood (abort replication or look further away). These types of dynamics are a subject of our future work.

Our results add to an extensive and expanding literature on evolutionary dynamics in spatially structured populations. In recent years, spatial evolutionary dynamics have garnered significant mathematical interest, both computationally and theoretically. Agent-based models have been used to show that the number of mutants at a given threshold size in spatially growing populations is larger than at the same size in a well-mixed system, due to the emergence of jackpot mutations (9). Further work has shown that this pattern can be more complex for disadvantageous mutants and depend on the parameter in which the disadvantage is expressed (19). These studies also defined mutant scaling laws for single-hit mutants, which formed the basis for our work; see also (10–12). Other theoretical work (35) examined a spatially distributed population where mutations arose and spread to neighboring locations, calculating the time distribution required for an individual to acquire a certain number of mutations. Deterministic equations for stochastic spatial games were

developed under the assumption of long-range interactions (36). Another study compared a death–birth process on a lattice with the dynamics of the replicator equation, revealing that spatial factors significantly altered the evolutionary outcomes (37). These studies build on a rich tradition of mathematical research into spatial evolutionary stochastic processes, utilizing both analytical approaches and partial differential equation modeling to study trait evolution in spatial contexts (38–42). Spatial evolutionary processes have also been investigated specifically in the context of cancer, with implications for tumor progression and expansion (22), for spatial heterogeneity and tumor sampling (43), and for the evolution and management of drug resistance (15, 44).

Modeling results help interpret experimental evolution data on the role of copy number variation

Gene amplifications are known to be important drivers of evolutionary processes (34, 45). While amplifications by themselves can be costly to cells, they can help to increase fitness under stressful conditions. We focus our analysis on the dynamics under such conditions. An extensively studied and debated system is *Escherichia coli* cells with a mutation in the lac operon (*lac*[−]) (46–49), which renders them almost unable to grow on lactose as the carbon and energy source (only about 2% enzyme activity compared with wild type). When a large bacterial population is plated on lactose, populations of revertant cells (*lac*⁺) emerge. There has been debate about the mechanisms underlying these evolutionary dynamics (46, 49–52). Experiments have indicated that stress-induced mutagenesis can play a role in this adaptation and that it can happen even in the absence of significant population growth. In addition, it has been shown that gene amplifications can play an important role (which might also be stress-induced) and that amplifications can promote bacterial population growth and subsequent reversion through mutations. It is thought that the originally plated cells contain small subpopulations with a lac operon duplication or that these duplications are induced by the environmental stress. The duplications allow the cells to grow slowly on lactose, and subsequent further gene amplifications allow for more efficient growth. At the same time, the increased gene copy number results in a higher chance to generate a revertant mutant that allows the bacteria to grow on lactose even without the gene amplification. Since gene amplifications by themselves are costly to cells, the increased copy numbers are subsequently lost. While we do not aim to discuss the broader debate about the evolutionary mechanisms underlying the emergence of revertant cells, the amplification-based mechanism observed in this experimental system represents the type of evolutionary process that we studied mathematically here, except that we did not explicitly take into account loss of amplifications, because the target mutation could be any mutation in the amplified sequence, not specifically a mutation that functionally replaces the gene amplification. The conclusion about the role of spatial population structure, however, would not change if we assumed that the target mutation achieves the same fitness advantage as the gene amplification and if we included loss of amplified genes in this scenario.

The same kind of patterns has also been documented for *Salmonella enterica* (53). In all these experiments, the bacteria grew as lawns, which represent 2D spatial growth dynamics. According to our model, this is the population structure in which it is most likely to detect cells with point mutations that were

generated in cells with gene amplifications. Similar evolutionary processes are thought to occur under other conditions of stress, including antibiotic resistance evolution, heat stress, or salt stress (54, 55).

The role of gene duplications and amplifications for the rate of mutant evolution in *E. coli* has also been investigated in experiments in which bacteria grew in liquid medium (56). This represents a more mixed and less spatial system. A synthetic genetic reporter system in *E. coli* was used to distinguish between copy number and point mutations and focused on the early phase of the evolutionary dynamics. The system involved an intact endogenous *galK* gene with a random promoter system. When *E. coli* is grown in the presence of galactose, there is selection pressure for increased *galK* expression, which can occur either through gene amplification or through point mutation. Among the experimental conditions investigated in this study, the low galactose environment (56) is the most applicable to our model, because the point mutations generated in amplified cells did not increase the bacterial fitness beyond the advantage already attained through amplification. In this system, it was found that bacteria carried either one or the other type of alteration, but not both together. The explanation given for this observation was based on clonal interference between adaptive amplifications and point mutations (56). It is interesting to point out, however, that our model also predicts the lack of co-occurrence of gene amplifications and point mutations in this system because the bacteria were grown in liquid medium. In a mixed system, our model suggests that to observe the co-occurrence of point mutations and gene amplifications in cells, the population size would have to grow to unrealistically large sizes, or the fitness advantage of the amplified cells would have to be extremely large. Hence, it is possible that the lack of point mutations arising in amplified cells in these experiments is the result of the basic growth laws of amplified cells in a mixed system (as opposed to a spatial system), without the need to invoke further complexities. It would be interesting to repeat this study in a setting where bacterial growth is more spatially restricted.

The result that spatial population structure greatly promotes the ability of gene amplifications to contribute to the generation of point mutations is important for a detailed interpretation and comparison of experimental results across different systems. It also has important implications for cancer. Since the growth of tumors (especially solid tumors) is characterized by strong spatial restrictions, gene amplifications likely accelerate the accumulation of point mutations in the amplified genes, which can in turn drive the adaptation of tumor cells to overcome natural selective barriers or to become resistant to therapies. While gene amplifications readily occur in cancer genomes (57), this might be especially important for extrachromosomal DNA (58, 59) (ecDNA). Oncogene amplifications occur on ecDNA, and this can promote drug resistance and disease progression through further evolutionary processes, including the acquisition of point mutations in amplified genes.

Apart from gene amplifications, we also considered MMR-deficient cells as an intermediate state with an elevated mutation rate. While MMR deficiency occurs in bacteria, a more relevant system in the context of our model is MMR deficiency in tumor cells, such as in colorectal cancer cells. This has been shown to accelerate colorectal cancer evolution through microsatellite instability. The MMR-deficient phenotype in this system typically requires the inactivation of the MMR gene, which involves two separate mutational events (because somatic cells are diploid). While this slows down the evolutionary pathway via MMR

deficiency, our model predicts that in the context of spatially structured cell growth, the MMR-deficient cells that are generated in a growing population can still substantially contribute to the generation of target mutations, but only if the MMR-deficient cells have a selective advantage. This is likely to be true for colorectal cancer. Interestingly, the gene inactivation leading to MMR deficiency in colorectal cancer cells also leads to a reduction in the rate at which cells undergo apoptotic cell death, which provides these cells with an immediate selective advantage that can outweigh fitness reduction through the accelerated generation of deleterious mutants (32). According to our model, this intrinsic fitness advantage of MMR-deficient cells, coupled with the spatial growth of colorectal tumors, allows MMR-deficient cells to drive tumor evolution (and this would not occur if these conditions were violated). It is unclear at what stage MMR-deficient cells in colorectal cancer are generated. It is possible that they are generated in non-neoplastic tissue before tumor initiation and that these MMR-deficient cells then initiate tumor growth. This would not be a mechanism to which our current model applies. This might be more likely to occur in Lynch syndrome, in which one copy of an MMR gene is already mutated at birth, and hence, only a single mutational hit is required to generate an MMR-deficient cell. Our calculations, however, show that even if the tumor is initiated in MMR-proficient cells, and MMR deficiency only evolves during the course of tumor growth, these MMR-deficient cells can still reach sufficient levels to accelerate tumor evolution due to the spatial structure of colorectal cancer.

Conclusions

Through the derivation of mutant scaling laws in spatially structured and expanding populations, we have provided a tool to study various multistep evolutionary processes during spatial population growth without the restrictions posed by computationally intensive agent-based models. Hence, we can answer evolutionary questions in the context of large populations and low mutation rates, which is relevant for bacterial and cancer cell populations. We have applied this methodology to study the role of intermediate cell types with accelerated mutation rates for the acquisition of target point mutations, particularly focusing on the role of gene duplications/amplifications, and further expanding this to MMR deficiency in cancer cells. This methodology, however, is more broadly applicable and can be used to answer questions concerned with the evolution of complex traits in growing populations. This has not only implications for evolutionary biology but also for biomedical questions such as the evolution of therapy resistance, immune escape, and aspects of disease development.

Supplementary Material

Supplementary material is available at [PNAS Nexus](https://www.pnas.org) online.

Funding

The authors acknowledge funding from National Science Foundation grants MCB 2438986 (D.W., N.L.K., and J.R.P.) and DMS 2435484 (N.L.K. and D.W.).

Preprints

This manuscript was posted on a preprint at <https://www.biorxiv.org/content/10.1101/2024.06.24.600326v2>.

Data Availability

There are no data underlying this work.

References

- Luria SE, Delbruck M. 1943. Mutations of bacteria from virus sensitivity to virus resistance. *Genetics*. 28:491–511.
- Zheng Q. 1999. Progress of a half century in the study of the Luria-Delbruck distribution. *Math Biosci*. 162:1–32.
- Dewanji A, Luebeck EG, Moolgavkar SH. 2005. A generalized Luria-Delbruck model. *Math Biosci*. 197:140–152.
- Komarova NL, Wu L, Baldi P. 2007. The fixed-size Luria-Delbruck model with a nonzero death rate. *Math Biosci*. 210:253–290.
- Kepler TB, Oprea M. 2001. Improved inference of mutation rates: I. An integral representation for the Luria-Delbrück distribution. *Theor Popul Biol*. 59:41–48.
- Goldie JH, Coldman AJ. 1983. A model for resistance of tumor cells to cancer chemotherapeutic agents. *Math Biosci*. 65:291–307.
- Goldie JH, Coldman AJ. *Drug resistance in cancer: mechanisms and models*. Cambridge University Press, 1998.
- Komarova NL, Wodarz D. 2005. Drug resistance in cancer: principles of emergence and prevention. *Proc Natl Acad Sci U S A*. 102:9714–9719.
- Fusco D, Gralka M, Kayser J, Anderson A, Hallatschek O. 2016. Excess of mutational jackpot events in expanding populations revealed by spatial Luria-Delbruck experiments. *Nat Commun*. 7:12760.
- Gralka M, Hallatschek O. 2019. Environmental heterogeneity can tip the population genetics of range expansions. *Elife*. 8:e44359.
- Gralka M, et al. 2016. Allele surfing promotes microbial adaptation from standing variation. *Ecol Lett*. 19:889–898.
- Otwinowski J, Krug J. 2014. Clonal interference and Muller's ratchet in spatial habitats. *Phys Biol*. 11:056003.
- Lavrentovich MO, Wahl ME, Nelson DR, Murray AW. 2016. Spatially constrained growth enhances conversional meltdown. *Biophys J*. 110:2800–2808.
- Waclaw B, et al. 2015. A spatial model predicts that dispersal and cell turnover limit intratumour heterogeneity. *Nature*. 525:261–264.
- Aif S, Appold N, Kampman L, Hallatschek O, Kayser J. 2022. Evolutionary rescue of resistant mutants is governed by a balance between radial expansion and selection in compact populations. *Nat Commun*. 13:7916.
- Yu Q, et al. 2021. Mutability of demographic noise in microbial range expansions. *ISME J*. 15:2643–2654.
- Paulose J, Hallatschek O. 2020. The impact of long-range dispersal on gene surfing. *Proc Natl Acad Sci U S A*. 117:7584–7593.
- Birzu G, Matin S, Hallatschek O, Korolev KS. 2019. Genetic drift in range expansions is very sensitive to density dependence in dispersal and growth. *Ecol Lett*. 22:1817–1827.
- Wodarz D, Komarova NL. 2020. Mutant evolution in spatially structured and fragmented expanding populations. *Genetics*. 216:191–203.
- Schreck CF, et al. 2023. Impact of crowding on the diversity of expanding populations. *Proc Natl Acad Sci U S A*. 120:e2208361120.
- Kayser J, Schreck CF, Gralka M, Fusco D, Hallatschek O. 2019. Collective motion conceals fitness differences in crowded cellular populations. *Nat Ecol Evol*. 3:125–134.
- Noble R, et al. 2022. Spatial structure governs the mode of tumour evolution. *Nat Ecol Evol*. 6:207–217.
- Colyer B, Bak M, Basanta D, Noble R. 2024. A seven-step guide to spatial, agent-based modelling of tumour evolution. *Evol Appl*. 17:e13687.
- West J, Schenck RO, Gatenbee C, Robertson-Tessi M, Anderson AR. 2021. Normal tissue architecture determines the evolutionary course of cancer. *Nat Commun*. 12:2060.
- Gibson MA, Bruck J. 2000. Efficient exact stochastic simulation of chemical systems with many species and many channels. *J Phys Chem A*. 104:1876–1889.
- Cao Y, Gillespie DT, Petzold LR. 2006. Efficient step size selection for the tau-leaping simulation method. *J Chem Phys*. 124:044109.
- Gillespie DT. 2007. Stochastic simulation of chemical kinetics. *Annu Rev Phys Chem*. 58:35–55.
- Salis H, Kaznessis Y. 2005. Accurate hybrid stochastic simulation of a system of coupled chemical or biochemical reactions. *J Chem Phys*. 122:54103.
- Wilkinson DJ. *Stochastic modelling for systems biology*. CRC Press, 2011.
- Rodriguez-Brenes IA, Komarova NL, Wodarz D. 2019. The role of telomere shortening in carcinogenesis: a hybrid stochastic-deterministic approach. *J Theor Biol*. 460:144–152.
- Granelli-Piperno A, Shimeliovich I, Pack M, Trumpfheller C, Steinman RM. 2006. HIV-1 selectively infects a subset of nonmaturing BDCA1-positive dendritic cells in human blood. *J Immunol*. 176:991–998.
- Madden-Hennessey K, et al. 2022. Loss of mismatch repair promotes a direct selective advantage in human stem cells. *Stem Cell Reports*. 17:2661–2673.
- Rodriguez-Brenes IA, Komarova NL, Wodarz D. 2013. Tumor growth dynamics: insights into evolutionary processes. *Trends Ecol Evol*. 28:597–604.
- Andersson DI, Hughes D. 2009. Gene amplification and adaptive evolution in bacteria. *Annu Rev Genet*. 43:167–195.
- Foo J, Leder K, Schweinsberg J. 2020. Mutation timing in a spatial model of evolution. *Stoch Process Their Appl*. 130:6388–6413.
- Hwang SH, Katsoulakis M, Rey-Bellet L. 2013. Deterministic equations for stochastic spatial evolutionary games. *Theoretical Economics*. 8:829–874.
- Evilsizor S, Lanchier N. 2016. Evolutionary games on the lattice: death-birth updating process. *Electron J Probab*. 21:1–29.
- Champagnat N, Méléard S. 2007. Invasion and adaptive evolution for individual-based spatially structured populations. *J Math Biol*. 55:147–188.
- Roca CP, Cuesta JA, Sánchez A. 2009. Effect of spatial structure on the evolution of cooperation. *Phys Rev E*. 80:046106.
- Ovaskainen O, et al. 2014. A general mathematical framework for the analysis of spatiotemporal point processes. *Theor Ecol*. 7:101–113.
- Ibsen-Jensen R, Chatterjee K, Nowak MA. 2015. Computational complexity of ecological and evolutionary spatial dynamics. *Proc Natl Acad Sci U S A*. 112:15636–15641.
- Chen Y-T. 2018. Wright-Fisher diffusions in stochastic spatial evolutionary games with death-birth updating. *Ann Appl Probab*. 28(6):3418–3490.
- Opasic L, Zhou D, Werner B, Dingli D, Traulsen A. 2019. How many samples are needed to infer truly clonal mutations from heterogeneous tumours? *BMC Cancer*. 19:1–11.
- Strobl MA, et al. 2022. Spatial structure impacts adaptive therapy by shaping intra-tumoral competition. *Commun Med (Lond)*. 2:46.
- Elliott KT, Cuff LE, Neidle EL. 2013. Copy number change: evolving views on gene amplification. *Future Microbiol*. 8:887–899.
- Hastings P, Bull HJ, Klump JR, Rosenberg SM. 2000. Adaptive amplification: an inducible chromosomal instability mechanism. *Cell*. 103:723–731.

-
- 47 Hersh MN, Ponder RG, Hastings P, Rosenberg SM. 2004. Adaptive mutation and amplification in *Escherichia coli*: two pathways of genome adaptation under stress. *Res Microbiol.* 155:352–359.
- 48 Hendrickson H, Slechta ES, Bergthorsson U, Andersson DI, Roth JR. 2002. Amplification–mutagenesis: evidence that “directed” adaptive mutation and general hypermutability result from growth with a selected gene amplification. *Proc Natl Acad Sci U S A.* 99:2164–2169.
- 49 Kugelberg E, Kofoed E, Reams AB, Andersson DI, Roth JR. 2006. Multiple pathways of selected gene amplification during adaptive mutation. *Proc Natl Acad Sci U S A.* 103:17319–17324.
- 50 Foster PL. 1994. Population dynamics of a Lac-strain of *Escherichia coli* during selection for lactose utilization. *Genetics.* 138:253–261.
- 51 Stumpf JD, Poteete AR, Foster PL. 2007. Amplification of lac cannot account for adaptive mutation to Lac⁺ in *Escherichia coli*. *J Bacteriol.* 189:2291–2299.
- 52 Slack A, Thornton P, Magner DB, Rosenberg SM, Hastings P. 2006. On the mechanism of gene amplification induced under stress in *Escherichia coli*. *PLoS Genet.* 2:e48.
- 53 Slechta ES, et al. 2003. Adaptive mutation: general mutagenesis is not a programmed response to stress but results from rare coamplification of *dinB* with *lac*. *Proc Natl Acad Sci U S A.* 100:12847–12852.
- 54 Yona AH, et al. 2012. Chromosomal duplication is a transient evolutionary solution to stress. *Proc Natl Acad Sci U S A.* 109:21010–21015.
- 55 Kondrashov FA. 2012. Gene duplication as a mechanism of genomic adaptation to a changing environment. *Proc R Soc Lond B Biol Sci.* 279:5048–5057.
- 56 Tomanek I, Guet CC. 2022. Adaptation dynamics between copy-number and point mutations. *Elife.* 11:e82240.
- 57 Albertson DG. 2006. Gene amplification in cancer. *Trends Genet.* 22:447–455.
- 58 Lange JT, et al. 2022. The evolutionary dynamics of extrachromosomal DNA in human cancers. *Nat Genet.* 54:1527–1533.
- 59 Wu S, Bafna V, Chang HY, Mischel PS. 2022. Extrachromosomal DNA: an emerging hallmark in human cancer. *Annu Rev Pathol.* 17:367–386.