MATHEMATICAL BIOSCIENCES
AND ENGINEERING
Volume **10**, Number **3**, June **2013**

doi:10.3934/mbe.2013.10.761

pp. **761–775**

SPATIAL STOCHASTIC MODELS OF CANCER: FITNESS, MIGRATION, INVASION

Natalia L. Komarova

Department of Mathematics, University of California Irvine
Irvine, CA 92697, USA

Abstract. Cancer progression is driven by genetic and epigenetic events giving rise to heterogeneity of cell phenotypes, and by selection forces that shape the changing composition of tumors. The selection forces are dynamic and depend on many factors. The cells favored by selection are said to be more "fit" than others. They tend to leave more viable offspring and spread through the population. What cellular characteristics make certain cells more fit than others? What combinations of the mutant characteristics and "background" characteristics make the mutant cells win the evolutionary competition? In this review we concentrate on two phenotypic characteristics of cells: their reproductive potential and their motility. We show that migration has a direct positive impact on the ability of a single mutant cell to invade a pre-existing colony. Thus, a decrease in the reproductive potential can be compensated by an increase in cell migration. We further demonstrate that the neutral ridges (the set of all types with the invasion probability equal to that of the host cells) remain invariant under the increase of system size (for large system sizes), thus making the invasion probability a universal characteristic of the cells' selection status. We list very general conditions under which the optimal phenotype is just one single strategy (thus leading to a nearly-homogeneous type invading the colony), or a large set of strategies that differ by their reproductive potentials and migration characteristics, but have a nearly-equal fitness. In the latter case the evolutionary competition will result in a highly heterogeneous population.

1. **Introduction.** The idea that cancer is an evolutionary process has been applied successfully by many computational biologists, as it allows them to use methods of theoretical population biology and ecology [34, 5, 18, 30, 44, 43, 27]. Populations of cancerous cells within a tumor are heterogeneous, and cells of different phenotypes compete with one another in a fast-paced evolutionary system. At the molecular level, mutations are introduced into the tumoral genome; these mutations may be caused by inherited deficiencies, loss of mismatch repair systems, downregulation of the proofreading checkpoints, and chromosomal instabilities. Other alterations come about by epigenetic events. At the cellular level, these alterations introduce changes in phenotype, some profound (and often deadly for the cell) but many others more subtle. They generate the flexibility and adaptability of the cancer disease state.

2010 *Mathematics Subject Classification.* 92B05, 92C17, 92C50, 92D25.

Key words and phrases. Cellular automata, lattice models, spatial cell dynamics, agent-based modeling, fitness, cell motility.

In this review we discuss the forces of natural selection that act upon cells in a heterogeneous tumor environment. In particular, we focus on understanding the role of spatial constraints and cellular motility in mutant dynamics. What cellular characteristics make certain cells more fit than others? If a mutant is introduced in a cell colony, what combinations of the mutant characteristics and "background" characteristics make the mutant cells win the evolutionary competition?

Here we focus on two types of phenotypic changes induced by mutations. The first type involves mutations in genes affecting cell proliferation. Activation of some oncogenes, or inactivation of tumor suppressor genes, change the cells' reproductive capacity, and are thought to be early events in the natural history of many cancers [19]. The second type of genetic change influences the cells' ability to migrate/move. Genes of the second type, while commonly associated with metastases, are also affected in primary tumors [37]. These two types of variation are thought to be implicated in malignant transformations for many (if not all) types of solid tumors. How do the two types of change trade-off to create a mutant which is "fitter" than the background?

Questions of this kind are related to the general theory of fitness landscapes, first introduced by [45]. Fitness is viewed as a surface in a multidimensional space, where the dynamics is assumed to be directed toward local fitness maxima. The global maximum corresponds to the evolutionarily stable strategy [40]. In scenarios where fitness of an individual strategy depends on the current composition of the population (frequency-dependent fitness), the formalism of fitness generating functions is used [42].

In this review we concentrate on a specific aspect of the general problem of fitness landscapes. Namely, we describe a qualitative framework to study the forces of selection acting within a spatially distributed, stochastic colony of cells, which can vary with regards to the two above mentioned characteristics. The models we describe for this purpose are a spatial generalization of the well-known Moran process, which was first introduced in [29]. This process has been used recently in cancer modeling (see [31, 25, 33, 28, 21]). The first spatial (1D) generalization of the Moran process was described in [24], where the processes of one-hit and two-hit mutant generation and fixation were considered. The simplicity of the (generalized) Moran process enabled us to study analytically, as well as numerically, the role of space in the processes of loss-of-function and gain-of-function mutations, see also [23]. In this spatial Moran process, the cells were allowed to divide in response to a death of a neighboring cell on a 1D grid. A generalization of the Moran process which explicitly includes cellular migration was introduced more recently in [41].

Results described here add to the large body of recent literature, where spatial cancer dynamics is studied by means of cellular automata or agent-based modeling, see e.g. the reviews [11, 6, 15, 17, 1], and the references therein, and more recent papers [10, 2, 35, 46, 20, 36, 7, 9, 38, 8, 22]. Rather than modeling many complex biological factors, in the present paper we focus on understanding how just the two forces, proliferation and migration, trade-off to influence the overall fitness of cells.

The rest of this review is organized as follows. In Section 2 we present a brief review of the non-spatial Moran process, which is a basis for the class of models used here to study cellular reproduction and motility. In Section 3 we describe the first generalization of the simple model, where cellular interactions acquire a spatial aspect, nut no cell movement is allowed. Section 4 reviews two Moran-based models of cell dynamics that include motility. The first model contains explicit

parameters responsible for cellular movement. The second model includes motility in an implicit form, and has the advantage of being analytically tractable. Section 5 discusses the issue of cellular selection and argues that the invasion probability is an appropriate and convenient measure of cellular fitness. It also develops a general theory of two-component fitness systems subject to external constraints. Section 6 contains discussion.

2. **The non-spatial Moran process.** The conventional Moran process is formulated as follows. In a population of N cells, each cell is equipped with a nonnegative replication parameter. The process is a sequence of updates. At each time-step, one cell is chosen randomly for death and is then replaced by a progeny of another cell (note that in this process, the probability to be picked for death is the same for all cells, regardless of their type). To choose which cell reproduces, one weighs in all cells replication parameters, such that the probability for cell i to reproduce is given by $r_i/\sum_{k=1}^{N} r_k$, where r_k is the replication parameter of cell k, and the summation is performed over all cells in the population.

Assume that there are two distinct types in the population, type A with replication parameter r_A and type B with replication parameter r_B. We will refer to parameters $r_A$and r_B as reproductive potentials of the two cell types. The stochastic process can be formulated in terms of only one independent random variable, the number of cells of type B. If there are N_A cell of type A and N_B cells of type B (with $N_A + N_B = N$), then after one update the following transitions are possible:

- $N_B \to N_B + 1$ with probability $P_{N_B \to N_B+1} = N_A/N \times r_B N_B/(r_A N_A + r_B i)$, where the first factor is the probability that a cell of type A dies and the second factor is the probability for a cell of type B to divide;
- $N_B \to N_B - 1$ with probability $P_{N_B \to N_B-1} = N_B/N \times r_A N_A/(r_A N_A + r_B i)$, where the first factor is the probability that a cell of type B dies and the second factor is the probability for a cell of type A to divide;
- $N_B \to N_B$ with probability $1 - P_{N_B \to N_B+1} - P_{N_B \to N_B-1}$.

We will refer to cells of type A as the host, or background, cells, and cells of type B as mutant cells. In the Moran process as it is formulated (that is, in the absence of new mutations), only two outcomes are possible: either type A wins and cells of type B disappear, or type B wins and cells of type A disappear. The probability for mutants to invade starting from one cell can be calculated analytically and is given by

$$\rho = \frac{1 - r_A/r_B}{1 - (r_A/r_B)^N}. \tag{1}$$

In the special case where $r_A = r_B$, we have $\rho = 1/N$. This can be obtained from formula (1) by taking the limit $r_B \to r_A$. Also, this result follows from symmetry considerations: in this simple model, if $r_A = r_B$ then a cell of type B has the same expansion properties as any of the host cells, and the same probability to invade. Since inevitably one of the N cells will invade, the probability of invasion is $1/N$ for every cell, regardless of their type.

3. **A one-dimensional spatial generalization of the Moran process.** In [24] we introduced and analyzed the first spatial generalization of the Moran process. Consider a 1D space, where all the N cells are placed on a regular grid, at locations $1, 2, \ldots, N$. As before, we assume that the total number of cells does not change. Cells are randomly chosen for death. Each cell death is followed by a cell division

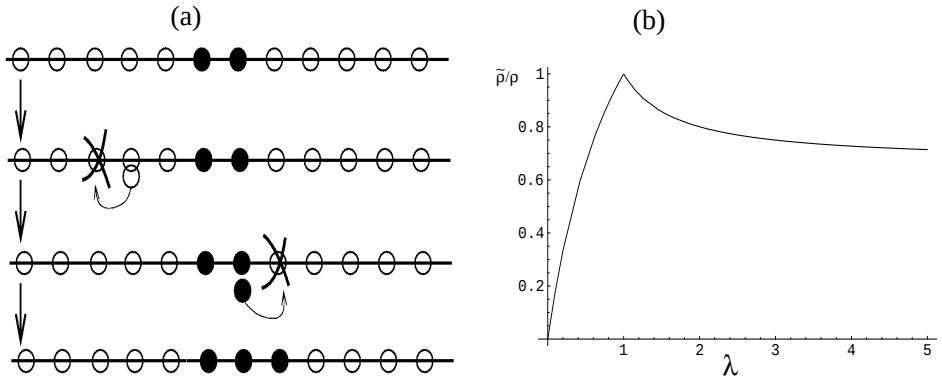

FIGURE 1. (a) The Moran process generalized to the one-dimensional space: a cell is chosen for death at random, and is immediately replaced by a division of one of the two neighboring cells (chosen proportional to their reproductive potential). (b) The quantity $\tilde{\rho}/\rho$, formula (2), as a function of $\lambda = r_B/r_A$ for $N = 1000$. This quantity tells us how much less likely a mutant fixation is in the 1D spatial model compared to the mass-action model.

of one of its two neighboring cells, which places its daughter cell at the empty slot, see figure 1(a). Cell death occurs randomly and division is proportional to the reproductive potential of the cells.

In this spatial model, the probability of mutant invasion in principle depends on the initial position of the mutant cell. However, if we use periodic boundary condition, this dependence disappears and we obtain that starting from one cell, the probability of invasion is

$$\tilde{\rho} = \frac{2r^{N-1}(1-r)}{1+r+r^{N-1}-3r^N}, \tag{2}$$

where $r \equiv r_B/r_A$. In figure 1(b) we can see that $\tilde{\rho} \leq \rho$, with the equality corresponding to the mutants with $r = 1$. For all other mutants, the probability to invade is smaller in a spatial model than it is in the space-free Moran process.

Note that both the space-free Moran process and the spatial generalization described above depend only on the ratio of the replication parameters r_B/r_A.

4. **Two-dimensional spatial generalizations of the Moran process.** Two different 2D generalizations of the Moran process were introduced and analyzed in [41].

4.1. **The explicit motility model.** Suppose that N cells are placed on a rectangular grid in 2D (the algorithm generalizes straightforwardly to 3D), where there is a cell of type A or B at each node. As in the standard Moran process, the first event of an iteration is the random selection of one cell to die. Then, one of the two events are possible: (i) a cell in the immediate proximity of the dead cell divides, placing its offspring in the empty spot, or (ii) one of the cells in the immediate proximity of the dead cell migrates into the empty spot. In all cases, the empty spot created by the initial cell death is filled. If a migration event occurred, a new empty space is created. The division and migration events occur proportionally to parameters r_X

and m_X for $X \in \{A, B\}$ respectively, and are contingent upon the number of each type, A and B, within nearest neighbors distance of the empty space. If a division event occurs, the iteration step is complete and a new cell is selected for death. If a migration event occurs, the migrating cell "trades places" with the empty space and a new event (division or migration) is selected. Thus, at the beginning and end of each iteration, the environment is completely filled; however, there may be many migration steps before a division event ends the step.

The explicit motility model is defined in terms of four cellular characteristics: the reproductive potential, r_X, and the migration potential, m_X, of the two cell types ($X \in \{A, B\}$). A normalization procedure however reduces the number of independent parameters to three:

$$\begin{aligned} P(\text{A divides}) &= \frac{N_A}{\tilde{K}}, \quad P(\text{B divides}) = \frac{N_B \lambda}{\tilde{K}} \\ P(\text{A migrates}) &= \frac{N_A k_A}{\tilde{K}}, \quad P(\text{B migrates}) = \frac{N_B \lambda k_B}{\tilde{K}}, \end{aligned}$$

where $k_X = \frac{m_X}{r_X}$, for $X \in \{A, B\}$, is the ratio of migration over reproductive potentials, $\lambda = \frac{r_B}{r_A}$ is the ratio of invader to background reproductive potentials, and $\tilde{K} = N_A(1 + k_A) + N_B\lambda(1 + k_B)$.

Numerical experiments of [41] demonstrated that an increased migration potential of mutant cells increases their ability to invade. This can be explained intuitively. In the case of small mutant motility, mutant cells tend to concentrate in one region, and the expansion can only occur near the boundary of that region. An increase in mutants' motility increases the degree of mixing in the population, such that mutant cells spread throughout the space. In this case, mutant growth is enhanced as it can occur throughout the bulk of the colony.

One interesting feature of the simulated dynamics has to be emphasized. Namely, when studying the probability of invasion, it is possible to find more than one pair of parameters (k_B, λ) which correspond to the same probability of invasion. In other words, we can see that different strategies–that is, different parameter sets (λ, k_B) relative to a specific background k_A–may have the same invasion probability. It is particularly interesting to investigate these strategies for the probability of invasion $P = \frac{1}{N}$. Prior work (see for example [25]) on the classical Moran process has shown that this invasion level corresponds to cells which have equal fitness, see Section 2. By plotting the pairs (k_B, λ) corresponding to constant probability of invasion given by $P = \frac{1}{N}$, we observe that there is a negative correlation between λ and k_B. In other words, increasing the motility (k_B) of invading cells will require a decrease in their reproductive potential (λ) in order to maintain the same level of invasion probability. An increase in either (or both) of parameters k_B and λ leads to an increase in invasion probability.

In [41] we defined all strategies corresponding to the invasion probability $P = 1/N$ as neutral to the host cells A. We called the sets of all such strategies "neutral ridges".

4.2. **The implicit motility model.** The advantage of the model described in the previous subsection is the explicit representation of motility within the cell population. A disadvantage is that the relationships for determining equality amongst strategies were found to be too complex to study analytically. Thus, we designed a simpler model of motility, in which cell movement is implicitly tied to the survival of new cell progeny [41]. Following the standard Moran process, at each time-step,

one cell is selected randomly to die, to be immediately replaced by the progeny of one of its neighboring cells. As before, the difference between the two phenotypes A and B is reflected in their reproduction parameters. Suppose that as before, cells of type A have reproductive potential r_A and cells of type B have reproductive potential r_B. We further employ the notion of the *division radius*. It measures the distance over which a cell can place its progeny. Suppose that cells of type A have division radius ν_A, and cells of type B have division radius ν_B. From a technical standpoint, a division radius of 1.0 in a 2D square grid corresponds to nearest neighbor interactions, while a radius of 1.5 corresponds to next nearest neighbors. The probability for the empty slot to be filled by either the background (A) type or mutant (B) type depends on the division radii and reproductive potential of each.

Unlike the model of Section 4.1, this implicit model only tracks *net* division and migration events. Further, the model depends only on two parameters: the relative reproductive potential r_B/r_A and the relative division radius ν_B/ν_A.

To explore how changes in both the reproductive potential and division radius of a mutant influence its invasion probability, we place a single mutant cell with reproductive potential r_B and division radius ν_B, into a background of cells with reproductive potential $r_A = 2.0$ and division radius $\nu_A = 2.5$. Results are plotted in figure 2 as contour levels, or level sets, of invasion probability, as a function of the reproduction and motility characteristics of the mutant type. Instead of using the parameter ν_X, the division radius of cells, we found it more convenient to use the number of neighbors, n_X, for each type. By the number of neighbors we mean the total number of slots in the vicinity of the cell which are within the division radius ν of the cell. The level sets of mutant invasion probability are presented in figure 2 as curves in the (n_B, r_B) space.

First of all, we observe that increasing r_B and n_B (separately or together) leads to an increase in the invasion probability. This in itself is not surprising given that larger values of the reproductive potential and division radius will increase the mutants' probability to divide. Thus, if a cell is free to choose any values of r_B and n_B, it will lead to an unrestricted growth of both of these parameters, which corresponds to the cells' climbing up the "hill" in the fitness landscape of figure 2, which corresponds to large r_B and n_B.

In reality though this is hardly possible and there are biological and energetic constraints on how often a cell divides, and how far it can travel upon division. Therefore "allowed" strategy trajectories in the (r_B, n_B) space can be introduced, such that some external constraints do not allow a simultaneous large increase in both parameters. For instance, we can envisage the relationship

$$\alpha_r r_B + \alpha_n n_B - d = 0, \tag{3}$$

where α_r and α_n are nonnegative weights and d is a constant. An example of such a constraint is shown by a dashed line in figure 2. We observe from figure 2 that within such a constraint, the background condition of $r_A = 2.0$ and $\nu_A = 2.5$ is able to resist most invasions from mutant strategies which maximize one trait at the cost of another (high n with low r or vice versa), such that the probability of mutant invasion is less than 25%. Thus, in this case the mixed strategy is more fit than those relying heavily on only one trait. Only for a narrow range of mixed strategies we obtain the probability of invasion greater than 25%.

A special class of level sets of invasion probability is neutral ridges. It was demonstrated numerically that the reproductive potential and the neighborhood

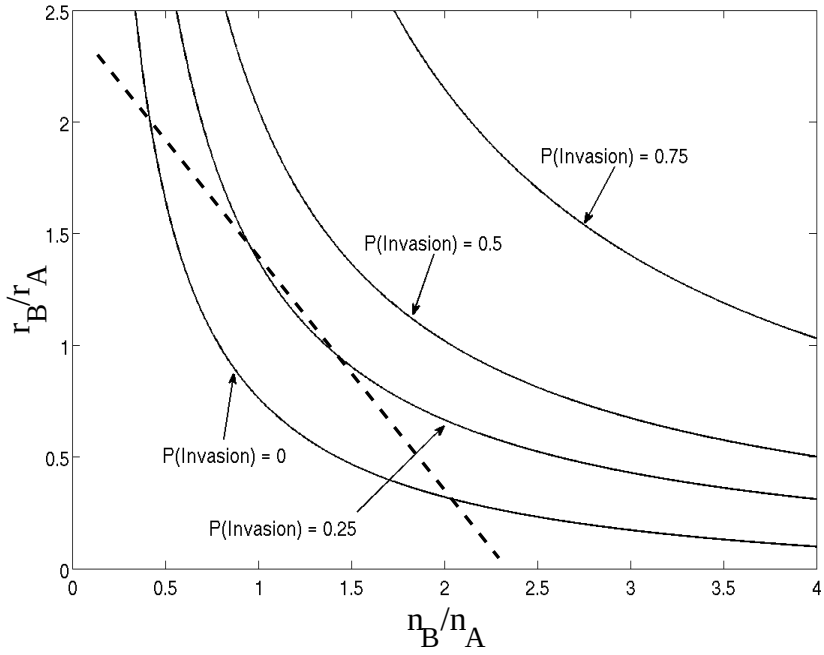

FIGURE 2. The level sets of the mutant invasion probability in the parameter space (n_B, r_B). The background parameters are $r_A = 2$ and $\nu_A = 2.5$. The dashed straight line corresponds to an external constraint of the form $\alpha_r r_B + \alpha_n n_B = d$.

sizes of phenotypes on the neutral ridge satisfies a very simple relationship:

$$r_E = c/n_E, \tag{4}$$

that is, the reproductive potential and number of neighbors (the square of the division radius) of the cells of constant invasion probability are inversely proportional to each other. We deduce that the neutral ridges satisfy the relationship $r_A n_A = r_B n_B$. This relationship makes sense from the point of view of the dimensional analysis, where the probability of reproduction r_X is defined subject to space availability, and this availability is proportional to the number of neighbors, n_X.

5. **Navigating the fitness landscape: Theoretical considerations.**

5.1. **Invasion probability vs sell fitness.** So far we have been investigating the probability of invasion of type B starting from one such mutant (this probability is denoted by ρ). This probability is connected to the *relative fitness* of type B, which we denote as W. Relative fitness is a parameter that is related to the rate of expansion of a phenotype. It is usually defined as the (ensemble-averaged) frequency of the type in the "next" generation divided by that in the current generation: $W = \frac{N_B^{next}}{N_B^{current}}$. A related concept of *invasion fitness* is defined as the exponential growth rate of the mutant type in a host population. The connection of the relative fitness parameter (W) to the probability of invasion (ρ) has the following useful properties:

(i) If $\rho = 1/N$ (see Section 2), then $W = 1$, that is, the cell of type B are neutral, and the number of such neutral cells stays constant on average, from generation to generation.
(ii) If $W \to \infty$, that is, if the mutants are strongly advantageous, then $\rho \to 1$.
(iii) The invasion probability and fitness are positively correlated. The types that have a higher invasion probability ρ will have a higher fitness and vise versa.

The notions of fitness and probability of invasion are both important in theoretical biology, and both have advantages and disadvantages. Arguably, the notion of invasion probability is more informative in our setting.

The fitness parameter is defined for a particular temporal dynamics. In our simple model, we use a discrete-time Markov chain. This is an unrealistic way to treat the mutant dynamics, but it preserves the invasion probabilities. In other words, if one considers the long-term outcomes of a stochastic process, the dynamics become unimportant, and can be simplified in the way implemented here. The expansion rate of cells (equivalent to fitness) is dependent on this simplification and thus is affected by this (unrealistic) aspect of our model.

The probability of invasion does not depend on the exact way we choose the time-step for our updates. Thus in this sense it is a better choice of a competitiveness measure for our system. On the other hand, invasion probability depends on the total number of cells, N. The probability to invade a very small constant-size population is higher than that for a large colony. We have investigated how the probability of invasion depends on N, and found the following.

For non-neutral mutants, probability of invasion is a monotonically decreasing function of N which saturates for large values of N. In our experiments, we used 2D grid size 21×21. We have also experimented with the size 31×31 (which yields N more than twice the original size), and found no measurable change in the results [41]. More precisely, the calculated values of the mean invasion probability for the larger grid size were within the standard deviation of the mean obtained by the smaller grid size, and vice versa. This is similar in spirit to our earlier analytical results for the invasion probability of the space-free Moran process (see equations (1)), and a 1D Moran process without motility (equation (2)). In that case, as long as $|r-1|N \gg 1$,

$$\rho \approx 1 - 1/r, \quad \tilde{\rho} \approx \frac{2}{3}\frac{1-1/r}{1-1/(3r)}.$$

A more subtle situation arises when the mutant B is neutral. Again using the example of our earlier analytical findings, we can see that if $|r-1|N \gg 1$,

$$\rho \approx \tilde{\rho} \approx 1/N,$$

that is, the invasion probability strongly depends on N. Despite this fact, the notion of invasion probability retains a degree of universality even in the case of neutral mutations. Namely, we found numerically that the expressions for the neutral ridges given by functions $r_B n_B = c$ are N-independent, that is, the proportionality constant c does not change with N. This is consistent with the parabolic dependence of the level sets for the invasion probability.

Ultimately, we are interested in the probability of invasion as long as it is equivalent to the probability of mutants to thrive. When considering cancer, it is irrelevant whether exactly all the host cells in an organ have been replaced by the mutants (which is equivalent to invasion, rigorously speaking). An important notion is whether a mutant colony expands and persists inside an organ for a long

time, which, for large values of N and for advantageous mutants, is very close to the probability of invasion. This notion is more universal than the fitness of mutant cells because it is independent of the time-evolution or the system size. Another meaningful concept is neutrality, which is basically a symmetry property: a neutral mutant behaves just like any host cell, and has the same invasion probability as any other cell ($1/N$). Its fitness is 1 and its expansion rate (invasion fitness) is zero, which is extremely hard to measure. Instead, measuring the invasion probability gives us a useful tool to identify the class of neutral mutants.

In the rest of this section we will discuss the probability of invasion and make some inferences about the fitness of the mutants. As explained above, the two notions are positively correlated, so all the arguments resulting from considering the level sets of the invasion probability hold for the mutant fitness.

5.2. **The optimal mutant strategy.** In Section 4.2 we saw that for cells of a given invasion probability, the simple relationship $r \propto 1/n$ holds, which means that the fitness function ρ (defined as the probability of type B to invade) satisfies

$$\rho(r_B, n_B) = \rho(r_B n_B).$$

Its levels correspond to a family of hyperbolas $r_B n_B = c$. Using the arguments of Section 5.1, we deduce that the same holds for the fitness function:

$$W(r_B, n_B) = W(r_B n_B).$$

Moreover, we know that the function W is a monotonically increasing function of its argument, because the fitness increases both with the reproductive potential and the division radius. Thus we have

$$W' \geq 0.$$

The simple hyperbolic shape of the level sets is a characteristic feature of the implicit motility model, and it does not necessarily hold for other, more realistic models. We can however conjecture that the following, weaker law holds for more general systems. If the fitness function depends on the reproductive potential and the motility parameter, $W = W(r_B, n_B)$, then we have

$$\frac{\partial W}{\partial r_B} > 0, \quad \frac{\partial W}{\partial n_B} > 0,$$

that is, increasing either of these parameters leads to an increase in fitness. In figure 3 level sets of the fitness function are drawn schematically in the (n_B, r_B) space. Using this minimal information we can make some progress in identifying likely evolutionary strategies of mutants which aim to maximize the function W. Let us suppose that external biological and energetic constraints impose a relationship between the allowed values of r_B and n_B, which we can write in the form

$$f(r_B, n_B) = 0$$

(equation (3) is an example of such a relationship). We further assume that this equation can be solved for r_B to obtain

$$r_B = g(n_B),$$

and thus the characteristic levels of the function W are given by

$$W(n_B, g(n_B)) = c.$$

We can further safely assume that $g' \leq 0$, because of the nature of the constraints we consider. Simply speaking, an increase in the reproductive potential must lead

to a somewhat reduced division radius, thus making the function g monotonically decreasing. In the example of constraint (3), we have $g(n_B) = (d - \alpha_n n_B)/\alpha_r$, and $g' = -\alpha_n/\alpha_r < 0$. Note that the absolute value of the derivative in this case is given by the ratio of the weights α_n and α_r.

It further follows that in general, the composite function $W(n_B, g(n_B))$ can take a variety of shapes. Here we restrict ourselves to the cases where it has at most one internal extremum. Multiple minima and maxima could occur, but they are a result of specific biological factors which we cannot specify at the present level of generality. Thus we focus on the simplest types of generic behavior, keeping in mind that other functions can be studied in a similar way. To find the maximum of function $W(n_B, g(n_B))$ we consider the derivative

$$\frac{dW}{dn_B} = \frac{\partial W}{\partial n_B} + g' \frac{\partial W}{\partial r_B}\bigg|_{r_B = g(n_B)}. \tag{5}$$

The first term in this expression is positive and the second is negative. With at most one internal extremum, there can be the following cases:

- The function $W(n_B, g(n_B))$ has an internal maximum. This means that the optimal strategy is to find an intermediate value of reproductive potential and division radius, instead of maximizing them both. This case is demonstrated in figure 2, the dashed line, and is also presented schematically in figure 3(a), where the constraint is shown as a dashed black line, and the most advantageous (the fittest) phenotype is near the letter "A" in the (n_B, r_B) space.
- The function $W(n_B, g(n_B))$ is monotonically decreasing, see figure 3(a), where the medium-grey dashed line indicates the constraint. Then, the maximum of W is achieved for the smallest possible values of n_B and the largest possible values of r_B (both within the constraint; the fittest region is indicated by the letter "B" in the figure). In other words, the optimal strategy is maximizing the reproductive potential (within the allowed range) at the cost of the division radius. This situation could arise in the following scenario: suppose that increasing n is "expensive", that is, the constraint function $f(r_B, n_B)$ depends stronger on n_B than it does on r_B. Then, the value of the derivative, $|g'|$, is relatively large, which shifts the location of the maximum of the function $W(n_B, g(n_B))$ to the left. If the dependence of n_B is sufficiently strong, then this can drive the location of the maximum outside the allowed domain of n_B. As a result, the optimal strategy will be to find the smallest possible n_B (within the constraint).
- The function $W(n_B, g(n_B))$ is monotonically increasing in the allowed domain (relatively small values of $|g'|$, see figure 3(a), where the light-grey dashed line indicates the constraint.). Then, the maximum of W is achieved for largest possible values of n_B and the smallest possible values of r_B (near letter "C" in the figure). Analogous to the previous argument, such scenarios could arise when increasing n_B is "cheaper" than increasing r.

Whether the optimal phenotype (the phenotype of the highest fitness) corresponds to a single point in the (n_B, r_B) space, or can be considered a "quasispecies", or an extended set of many different points in the space of roughly equal fitness, depends on the shapes of the fitness landscape levels, as well as the constraint. For example, the constraint represented by a dashed black line in figure 3 follows very nearly an equal-fitness line near the letter "A", which means that a whole range

of phenotypes will have almost the same high fitness. On the other hand, the constraint represented by a black dotted line points to a clear fitness maximum near the point "A*". In this case, the population of "winners" is relatively homogeneous.

The examples of constraints presented in figure 3(a) and described by equation (5) are all one-dimensional sets. Alternatively, one can envisage constraints which are two-dimensional regions in the (n_B, r_B) space, figure 3(b). In principle the choice of the fittest strategy under such two-dimensional constraints does not differ much from the case of one-dimensional constraints. For the constraint marked by "1" in figure 3, the region near the letter "E" can be considered an extended set of nearly-equally fit phenotypes, because the boundary of the corresponding 2D constraint runs close to a level set of the fitness landscape. On the other hand, the point "D" represents more or less a single winner of the evolutionary competition under constraint "2".

To summarize, we find that the reproductive potential and the division radius are two components of fitness which both play a role in the probability of mutant invasion and the fitness of mutants. The fitness increases with both of these parameters (in the case of the implicit motility model, the fitness landscape has hyperbolic level sets in terms of the reproductive potential and the number of neighbors). A mixed strategy is optimal unless one of the two fitness components (the reproductive potential or the number of neighbors) is relatively more evolutionary "expensive". In the latter case the less expensive characteristic should be maximized at the expense of the other. Depending on the curvature of the levels of the fitness landscape and the shape of the constraint, the optimal phenotype could be just one single strategy or a large set of strategies that differ by their reproductive potentials and migration characteristics, but have a nearly-equal fitness.

6. **Discussion.**

A central idea of this paper is that the cell with the fastest intrinsic growth rate is not always the most fit one. Rather, it is the combination of *ability* to divide with increased *opportunity* to do so that defines the cell's success in the evolutionary competition. The cell's characteristics responsible for divisions and migration must be balanced, depending on the relative costs of both adaptations for a cell.

We have demonstrated how relatively simple, two-component, models can help explore the complicated phenomenon of genetic heterogeneity. With the multiphasic nature of fitness presented here, it becomes possible that many different strategies will have equal or nearly equal fitness. As the number of traits under consideration increases, so too will the possible combinations of strategies which lead to equal fitness. This would correspond to the existence of multidimensional sets equivalent to "neutral ridges" described here.

The genetic heterogeneity of cancer is a primary source of its resistance to modern medicine. Tumors behave as fast evolutionary systems, with many subtly different phenotypes coexisting and competing with each other in the tumor environment. Changes in phenotypes arise from altered gene expression profiles, rapid cell proliferation outpacing DNA repair mechanisms, loss of repair systems or checkpoints, loss of chromosomal integrity and epigenetic events. Addition of therapy (chemical, radiation) to this system changes the selective forces acting upon the various phenotypes, destroying many but selecting a few which soon become dominant as the tumor regrows. This level of adaptability underlies the difficulty in studying, understanding, and treating this complex disease. A better understanding of how

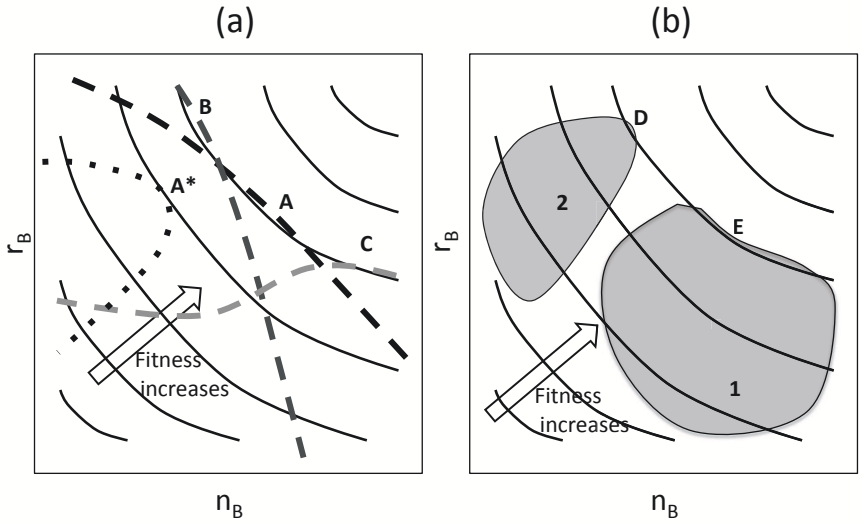

FIGURE 3. A schematic of a fitness landscape in the (n_B, r_B) space, together with several types of constraints. (a) One-dimensional constraints are represented by dashed and dotted lines, and the optimal phenotypes are marked by "A", "A*", "B", and "C" in the four cases shown. (b) Two-dimensional constraints are represented by shaded regions "1" and "2", and the optimal phenotypes are marked by "D" and "E" in the two cases.

mutants penetrate the pre-treatment tumor, how their phenotype survives in the milieu of many thousands of subtle variants, is critical for designing new therapy strategies which minimize the emergence of resistance.

The simplified textbook notion of microevolution, where one mutant type is selected for proliferation by the collective sum of forces within the tumor, offers little detail about how a great variety of mutants manages to survive. To explain the observed heterogeneity of tumors, we speculate that multiple phenotypes expressed within a tumor are more or less equally fit, but they achieve this level of fitness through different mechanisms. We have formulated and presented mathematical models which show how fitness relates to two separate cellular characteristics, and how those characteristics might be combined in different measures to create cells with different phenotypes but approximately equal fitness.

In the implicit migration model considered here we found that the fitness of cells depends on two parameters, the reproductive potential and number of neighbors. Moreover, we demonstrated that it depends on the product of the two. This is not an unusual finding in evolutionary biology. Another example comes from virus dynamics, where the fitness of viruses (measured in terms of their basic reproductive ratio) can be calculated as a function of various parameters in predator-prey type models. It is typically proportional to the product of two quantities: the infectivity of the virus and the inverse of the death rate of infected cells. The former parameter can be roughly related to the "motility" of the virus, or to how efficiently it infects new cells, and in this sense it is roughly analogous to our division radius. The inverse death rate of infected cells is in some sense similar to our reproductive

potential. This is because a decrease in the cell's lifespan (a decrease in the inverse death rate) leads to a decrease in the total number of viruses produced. Similarly, a decrease in r leads to a decrease in the number of offspring produced. Therefore we can see a strong correspondence between our model and the virus dynamics model which is well studied in the literature. Several papers have been devoted to studying optimal viral strategies, especially given that the two components of fitness are not independent [3, 32, 26, 13, 16, 14, 39, 4]. In spirit, our approach is similar to that. While there has to date been little investigation into this area, it seems reasonable to assume that energy availability limits the total growth and migration ability of an invasive cancer cell. Since both growth and migration are energy-intensive processes, an invasive cell will have to 'budget' its energy to each.

In this paper we obtained numerical and analytical results for two particular models of migration, and based on these insights, developed a fairly general description of systems whose fitness is a function of not just one but two factors (reproductive potential and migration characteristics). We showed that depending on the shape of external constraints, the evolutionary system will chose one or more optimal phenotypes, which can outcompete the resident population and form the population of the growing tumor.

Acknowledgments. The support of NIH grant 1R01CA129286-01A1 is gratefully acknowledged.

REFERENCES

[1] A. Anderson, M. Chaplain, K. Rejniak and J. Fozard, *Single-cell based models in biology and medicine*, Math. Med. Biol., **25** (2008), 185–186.

[2] A. Anderson and V. Quaranta, *Integrative mathematical oncology*, Nature Reviews Cancer, **8** (2008), 227–244.

[3] R. M. Anderson and R. M. May, *Coevolution of hosts and parasites*, Parasitology, **85** (1982), 411–426.

[4] M. Boots, P. J. Hudson and A. Sasaki, *Large shifts in pathogen virulence relate to host population structure*, Science, **303** (2004), 842–844.

[5] J. Breivik and G. Gaudernack, *Carcinogenesis and natural selection: A new perspective to the genetics and epigenetics of colorectal cancer*, Adv. Cancer Res., **76** (1999), 187–212.

[6] H. Byrne, T. Alarcón, M. Owen, S. Webb and P. Maini, *Modeling aspects of cancer dynamics: A review*, Phi. Trans. R. Soc. A, **364** (2006), 1563–1578.

[7] A. Chauvière, L. Preziosi and C. Verdier, "Cell Mechanics: From Single Scale-Based Models to Multiscale Modeling," CRC Press, **32**, 2009.

[8] B. Chopard, R. Ouared, A. Deutsch, H. Hatzikirou and D. Wolf-Gladrow, *Lattice-gas cellular automaton models for biology: from fluids to cells*, Acta Biotheoretica, **58** (2010), 329–340.

[9] T. Deisboeck and G. Stamatakos, "Multiscale Cancer Modeling," CRC Press, **34**, 2010.

[10] T. Deisboeck, L. Zhang, J. Yoon and J. Costa, *In silico cancer modeling: is it ready for prime time?*, in press.

[11] A. Deutsch and S. Dormann, "Cellular Automaton Modeling of Biological Pattern Formation," Birkhauser, 2005.

[12] D. Drasdo and S. Höhme, *On the role of physics in the growth and pattern of multicellular systems: What we learn from individual-cell based models?*, J. Stat. Phys., **128** (2007), 287–345.

[13] D. Ebert and E. A. Herre, *The evolution of parasitic diseases*, Parasitol Today, **12** (1996), 96–101.

[14] D. Ebert and K. L. Mangin, *The influence of host demography on the evolution of virulence of a microsporidian gut parasite*, Evolution, **51** (1997), 1828–1837.

[15] A. Fasano, A. Bertuzzi and A. Gandolfi, *Complex systems in biomedicine chapter mathematical modelling of tumour growth and treatment*, Milan: Springer, (2006), 71–108.

[16] S. A. Frank, *Models of parasite virulence*, Q. Rev. Biol., **71** (1996), 37–78.

[17] J. Galle, G. Aust, G. Schaller, T. Beyer and D. Drasdo, *Individual cell-based models of the spatial temporal organization of multicellular systems- achievements and limitations*, Cytometry, **69A** (2006), 704–710.

[18] R. Gatenby and P. Maini, *Mathematical oncology: Cancer summed up*, Nature, **421** (2003), 321.

[19] D. Hanahan and R. Weinberg, *The hallmarks of cancer*, CELL, **100** (2000), 57–70.

[20] P. Hinow,, P. Gerlee, L. McCawley, V. Quaranta, M. Ciobanu, S. Wang,, J. Graham, B. Ayati, J. Claridge, K. Swanson, et al., *A spatial model of tumor-host interaction: Application of chemotherapy*, Mathematical Biosciences and Engineering: MBE, **6** (2009), 521.

[21] Y. Iwasa, F. Michor and M. A. Nowak, *Stochastic tunnels in evolutionary dynamics*, Genetics, **166** (2004), 1571–1579.

[22] Y. Jiao and S. Torquato, *A cellular automaton model for tumor growth in heterogeneous environment*, Bulletin of the American Physical Society, **56** (2011).

[23] N. Komarova, *Loss- and gain-of-function mutations in cancer: Mass-action, spatial and hierarchical models*, Jour. Stat. Phys., **128** (2007), 413–446.

[24] N. L. Komarova, *Spatial stochastic models for cancer initiation and progression*, Bull. Math. Biol., **68** (2006), 1573–1599.

[25] N. L. Komarova, A. Sengupta and M. A. Nowak, *Mutation-selection networks of cancer initiation: Tumor suppressor genes and chromosomal instability*, J. Theor. Biol., **223** (2003), 433–450.

[26] B. R. Levin, *The evolution and maintenance of virulence in microparasites*, Emerg. Infect. Dis., **2** (1996), 93–102.

[27] L. Merlo, J. Pepper, B. Reid and C. Maley, *Cancer as an evolutionary and ecological process*, Nat. Rev. Cancer, **6** (2006), 924–935.

[28] F. Michor, Y. Iwasa, H. Rajagopalan, C. Lengauer and M. A. Nowak, *Linear model of colon cancer initiation*, Cell Cycle, **3** (2004), 358–362.

[29] P. Moran, "The Statistical Processes of Evolutionary Theory," Clarendon, Oxford, 1962.

[30] M. Nowak and K. Sigmund, *Evolutionary dynamics of biological games*, Science, **303** (2004), 793–799.

[31] M. A. Nowak, N. L. Komarova, A. Sengupta, P. V. Jallepalli, I.-M. Shih, B. Vogelstein and C. Lengauer, *The role of chromosomal instability in tumor initiation*, Proc. Natl. Acad. Sci. U S A, **99** (2002), 16226–16231.

[32] M. A. Nowak and R. M. May, *Superinfection and the evolution of parasite virulence*, Proc. Biol. Sci., **255** (1994), 81–89.

[33] M. A. Nowak, F. Michor, N. L. Komarova and Y. Iwasa, *Evolutionary dynamics of tumor suppressor gene inactivation*, Proc. Natl. Acad. Sci. U S A, **101** (2004), 10635–10638.

[34] P. Nowell, *The clonal evolution of tumor cell populations*, Science, **194** (1976), 23–28.

[35] V. Quaranta, K. Rejniak, P. Gerlee and A. Anderson, *Invasion emerges from cancer cell adaptation to competitive microenvironments: Quantitative predictions from multiscale mathematical models*, Sem. Cancer Biol., in press, (2008).

[36] K. Rejniak and A. Anderson, *Hybrid models of tumor growth*, Wiley Interdisciplinary Reviews: Systems Biology and Medicine, **3** (2011), 115–125.

[37] C. W. Rinker-Schaeffer, J. P. O'Keefe, D. R. Welch and D. Theodorescu, *Metastasis suppressor proteins: Discovery, molecular mechanisms, and clinical application*, CLINICAL CANCER RESEARCH, **12** (2006), 3882–3889.

[38] J. Sagotsky and T. Deisboeck, *Simulating cancer growth with agent-based models*, Multiscale Cancer Modeling, **34** (2010), 173.

[39] A. Sasaki and M. Boots, *Parasite evolution and extinctions*, Ecology Letters, **6** (2003), 176.

[40] J. M. Smith, "Evolution and the Theory of Games," Cambridge University Press, 1982.

[41] C. Thalhauser, J. Lowengrub, D. Stupack and N. Komarova, *Research selection in spatial stochastic models of cancer: Migration as a key modulator of fitness*, Biology Direct, **5** (2010), 21.

[42] T. L. Vincent and J. S. Brown, "Evolutionary Game Theory, Natural Selection, and Darwinian Dynamics," Cambridge University Press, 2005.

[43] P. Vineis and M. Berwick, *The population dynamics of cancer: A Darwinian perspective*, Int. J. Epidemiol, **35** (2006), 1151–1159.

[44] D. Wodarz and N. Komarova, "Computational Biology of Cancer: Lecture Notes and Mathematical Modeling," World Scientific, 2005.

[45] S. Wright, *The roles of mutation, inbreeding, crossbreeding, and selection in evolution*, in "Proceedings of the Sixth International Congress on Genetics", (1932), 355–366.

[46] L. Zhang, Z. Wang, J. Sagotsky and T. Deisboeck, *Multiscale agent-based cancer modeling*, Journal of Mathematical Biology, **58** (2009), 545–559.

Received July 23, 2012; Accepted November 17, 2012.

E-mail address: komarova@uci.edu