Dynamics of metastasis suppressor gene inactivation

Franziska Michor[a,*], Yoh Iwasa[b]

[a]Harvard Society of Fellows, Harvard Program for Evolutionary Dynamics, Cambridge, MA 02138, USA
[b]Department of Biology, Kyushu University, Fukuoka 812-8581, Japan

Abstract

For most cancer cell types, the acquisition of metastatic ability leads to clinically incurable disease. Twelve metastasis suppressor genes (MSGs) have been identified that reduce the metastatic propensity of cancer cells. If these genes are inactivated in both alleles, metastatic ability is promoted. Here, we develop a mathematical model of the dynamics of MSG inactivation and calculate the expected number of metastases formed by a tumor. We analyse the effects of increased mutation rates and different fitness values of cells with one or two inactivated alleles on the ability of a tumor to form metastases. We find that mutations that are negatively selected in the main tumor are unlikely to be responsible for the majority of metastases produced by a tumor. Most metastases-causing mutations will be present in all (or most) cells in the main tumor.

© 2006 Elsevier Ltd. All rights reserved.

Keywords: Metastasis; Metastasis suppressor genes; Tumor suppressor genes; Mathematical model; Stochastic dynamics

1. Introduction

Tumor metastasis is a significant contributor to death in cancer patients (Vogelstein and Kinzler, 2002). Metastases arise when cancer cells leave the primary tumor site and form new tumors elsewhere (Weiss, 2000; Chambers et al., 2000, 2001, 2002). Metastasis formation is driven by genetic alteration of many genes, including activation of oncogenes such as RAS and MYC (Pozzatti et al., 1986; Wyllie et al., 1987). The hypothesis that metastasis might also involve loss of gene functions that maintain the normal state of a cell emerged after the identification of RB as a tumor suppressor gene. In 1988, the first metastasis suppressor, NM23, was identified (Steeg et al., 1988). Since then, 12 metastasis suppressor genes (MSGs) have been confirmed (Steeg, 2004) (Table 1).

NM23 was identified by its reduced expression in highly metastatic melanoma cell lines and has been shown to reduce the in vivo metastatic potential of cells when transfected into metastatically competent cell lines (Steeg et al., 1988). However, NM23 expression does not affect proliferation in vitro or primary tumor size in vivo. NM23 is a histidine kinase that phosphorylates the kinase suppressor of RAS (KSR) protein and reduces the metastatic potential of melanoma, breast, colon and oral squamous cell carcinomas (Backer et al., 1993).

MKK4 phosphorylates JNK and p38 and its loss facilitates metastatic colonization, may be by preventing apoptosis in response to the stress of a foreign environment. It is a functional metastasis suppressor when transfected into metastatic prostate and ovarian cancer cell lines (Yoshida et al., 1999). Furthermore, MKK4 has been found to have reduced expression in primary prostate tumors of increasing Gleason Grade and in metastatic ovarian carcinomas (Kim et al., 2001; Yamada et al., 2002).

The breast cancer metastasis suppressor 1 gene (BRMS1) inhibits metastasis in breast carcinoma and melanoma cell lines with no effect on tumorigenicity (Seraj et al., 2000). BRMS1 functions in Gap-junctional communication and its identification as an MSG suggests that increased GAP-junctional communication among metastatic tumor cells might contribute to the inhibition of metastatic outgrowth (Saunders et al., 2001).

In a previous paper, we have developed a model for the situation where cells acquire metastatic ability by one

---

*Corresponding author.
 E-mail address: michor@fas.harvard.edu (F. Michor).

Table 1
Twelve metastasis suppressor genes have been identified so far that are involved in the metastatic ability of diverse types of cancer

Gene	Cancer type	Function
BRMS1	Breast, melanoma	Gap-junctional communication
CLAUDIN4	Pancreas	Tight-junctional constituent
CRSP3	Melanoma	Transcriptional co-activator
DRG1	Prostate	Unknown
KAI1	Prostate, breast	Integrin interaction
KiSS1	Melanoma, breast	G-protein-coupled receptor ligand
MKK4	Prostate, ovarian	Activation of p38 and JNK kinases
NM23	Melanoma, breast, colon, oral squamous cell	Histidine kinase
RKIP	Prostate	Inhibits RAF function
RHOGDI2	Bladder	Regulates RHO and RAC function
SSeCKs	Prostate	Scaffolding protein for protein kinases A and C
VDUP1	Melanoma	Thioredoxin inhibitor

genetic alteration (Michor et al., 2006). In the present paper, we develop a mathematical representation of situations where two genetic alterations are needed to confer metastatic ability to the cell. These two genetic alterations can be the mutations inactivating an MSG, or they can be mutations in two independent genes, both of which are needed for metastatic ability. We calculate the number of metastases generated by a tumor of size N. The number of metastases depends on the two mutation rates, the fitness values of cells with one or two mutations, the population size of the main tumor, and the rate at which cells with both mutations are exported from the main tumor to found distant metastases.

Note that our formulas for the expected number of MSG-deficient cells in the main tumor can also be used to calculate the expected number of cells with two inactivated alleles of a tumor suppressor gene. A tumor suppressor gene contributes to tumorigenesis if both alleles are inactivated (Vogelstein and Kinzler, 2002). A quantitative theory of tumor suppressor gene inactivation is essential for a complete understanding of tumorigenesis (Nowak et al., 2004; Michor et al., 2004a; Iwasa et al., 2005).

Mathematical models of metastasis have provided considerable insights. Bosl et al. (1983) use a multivariate analysis of prognostic variables to study metastatic testicular cancer. Panetta (1996) develops a competition model to describe tumor-normal cell interaction with chemotherapy. Thames et al. (1999) determine the frequency of first metastatic events in breast cancer. Pescarmona et al. (1999) design a non-linear model of cancer growth and metastasis. Delsanto et al. (2000) study a mathematical model of the spatiotemporal evolution of neoplasias. Wodarz and Krakauer (2001) use a mathematical model to examine the role of genetic instability in angiogenesis and metastasis. Suzuma et al. (2001) design a mathematical model of axillary lymph node involvement in breast cancer metastases. Wodarz et al. (2004) study a mathematical model of multifocal tumors, suggesting that the sum of the tumor sizes across all lesions is the best characteristic which correlates with the stage and meta-

static potential of the tumor. These papers contribute to deriving a quantitative understanding of tumorigenesis (Nordling, 1953; Armitage and Doll, 1954, 1957; Fisher, 1959; Goldie and Coldman, 1979; Moolgavkar and Knudson, 1981; Goldie and Coldman, 1983; Sheratt and Nowak, 1992; Taddei et al., 1997; Anderson and Chaplain, 1998; Strauss, 1998; Nunney, 1999; Owen and Sherrat, 1999; Chang et al., 2001; Knudson, 2001; Komarova et al., 2002; Luebeck and Moolgavkar, 2002; Nowak et al., 2002; Tomlinson et al., 2002; Chang et al., 2003; Frank et al., 2003; Gatenby and Maini, 2003; Gatenby and Vincent, 2003; Komarova et al., 2003; Little and Wright, 2003; Michor et al., 2003; Nowak et al., 2003; Iwasa et al., 2004; Michor et al., 2004a, b, 2005a,b, 2006; Wodarz and Komarova, 2005).

2. The model

Consider a population of N cancer cells proliferating according to the Moran process (Moran, 1962). Initially, all cells are wild type with respect to the MSG. Such cells are called type 0 cells. At each time step, a cell is chosen for reproduction at random, but proportional to fitness. Here a time unit is equal to the mean time of one cell generation. A mutated cell can be neutral ($r = 1$), advantageous ($r > 1$) or deleterious ($r < 1$) as compared to wild type cells. If there are i mutated cells with relative fitness r, then the probability that a mutated cell is chosen for reproduction is $ri/(ri + N - i)$. The chosen cell produces a daughter cell that replaces another randomly chosen cell. The total number of cells remains strictly constant. If the mutated cell has relative fitness $r > 1$ or $r < 1$, then the probability that it will take over the whole population is given by $\rho = (1 - 1/r)/(1 - 1/r^N)$ (Moran, 1962). If the mutated cell has relative fitness $r = 1$, then $\rho = 1/N$. The quantity ρ is called fixation probability. An advantageous mutation has a higher fixation probability than a neutral mutation, which has a higher fixation probability than a deleterious mutation. The events in a small population, however, are dominated by random drift: if N is small, then even a

deleterious mutation has a certain probability of reaching fixation due to chance events. The transition probabilities of the Moran process per time interval Δt are given by

$$P(i, i+1) = \left[\frac{ri}{N-i+ri} N(1-i) \right] \Delta t,$$

$$P(i, i-1) = \left[\frac{N-i}{N-i+ri} Ni \right] \Delta t,$$

$$P(i, i) = 1 - P(i, i+1) - P(i, i-1).$$

The two MSG alleles are inactivated at rates u_1 and u_2 per cell division, respectively. Inactivation of both alleles is necessary to confer metastatic potential to the cell. Let $u_1 < u_2$ because there are more possibilities, such as mitotic recombination and loss of heterozygosity, for the second hit. The mutation inactivating the first MSG allele confers a relative fitness r_1 to the cell. Let us call such a cell type 1 cell. If $r_1 > 1$, then the first mutation is advantageous and a type 1 cell has a higher fitness than a type 0 cell; if $r_1 < 1$, then the first mutation is deleterious and a type 1 cell has a lower fitness than a type 0 cell; if $r_1 = 1$, then the first mutation is neutral and a type 1 cell has the same fitness as a type 0 cell. The mutation inactivating the second allele confers a relative fitness r_2 to the cell. Such a cell is called type 2 cell. The second mutation, too, can be advantageous ($r_2 > 1$), deleterious ($r_2 < 1$), or neutral ($r_2 = 1$) as compared to type 0 cells. The rate of metastasis formation is proportional to the number of type 2 cells in the main tumor,

$$R = q \int_0^T E(t) \, dt. \tag{1}$$

Here $E(t) = N\bar{z}(t)$ denotes the expected number of type 2 cells in the tumor, where $\bar{z}(t)$ is the expected fraction of type 2 cells in the tumor at time t. The rate at which type 2 cells are exported from the main tumor to found distant metastases is denoted by q.

Denote the frequencies of type 0, 1 and 2 cells by x_0, x_1, and x_2, respectively. Let $\bar{r} = x_0 + r_1 x_1 + r_2 x_2$. Then, assuming deterministic dynamics, the frequencies of type 0, 1 and 2 cells change according to

$$\dot{x}_0 = (1 - \bar{r})x_0 - u_1 x_0,$$

$$\dot{x}_1 = (r_1 - \bar{r})x_1 + u_1 x_0 - u_2 x_1,$$

$$\dot{x}_2 = (r_2 - \bar{r})x_2 + u_2 x_1. \tag{2}$$

The expected number of type 2 cells at time t is given by $E(t) = Nx_2(t)$. However, Eq. (2) holds only for large population sizes and small fitness values. In the following, we develop stochastic formulas for the expected number of type 2 cells for several cases differing in the fitness of type 1 and 2 cells.

2.1. Exact stochastic computer simulation

We compare our analytical results with direct stochastic computer simulations of the Moran process. We define three integer variables for the numbers of type 0, 1, and 2

cells, $x_0 \in \{0, 1, \ldots, N\}$, $x_1 \in \{0, 1, \ldots, N\}$ and $x_2 \in \{0, 1, \ldots, N\}$, subject to the constraint $x_0 + x_1 + x_2 = N$. Each process is initiated with N type 0 cells, $x_0 = N$. Let $\Gamma = x_0 + r_1 x_1 + r_2 x_2$. The transition probabilities between states are given by

$$Pr[x_0 \to x_0 + 1] = \frac{x_0(1 - u_1)}{\Gamma} \frac{x_1 + x_2}{N},$$

$$Pr[x_1 \to x_1 + 1] = \frac{x_0 u_1 + r_1 x_1(1 - u_2)}{\Gamma} \frac{x_0 + x_2}{N},$$

$$Pr[x_2 \to x_2 + 1] = \frac{r_1 x_1 u_2 + r_2 x_2}{\Gamma} \frac{x_0 + x_1}{N}.$$

For each parameter choice, we average over many independent runs of the stochastic process to account for random fluctuations. Due to computational restrictions, we scale the parameter values such that we can use small population sizes. The data points generated by the computer simulation are compared with the analytical results in Figs. 1–7.

2.2. Neutral mutations

Let us first discuss the case in which both type 1 and 2 cells are neutral as compared to type 0 cells. In that case, Eq. (2) holds with $r_1 = r_2 = 1$ (see Appendix A). Then the expected number of type 2 cells at time t is given by

$$E(t) = Nx_2(t) = N\left[1 - \frac{u_2 e^{-u_1 t}}{u_2 - u_1} + \frac{u_1 e^{-u_2 t}}{u_2 - u_1} \right]. \tag{3}$$

Fig. 1 shows the results of the exact stochastic computer simulation of the two-hit process (as discussed in Section 2.1) and the numerical simulation of Eq. (3). We plot the expected number of type 2 cells, E, at time t. The mutation rates are $u_1 = 10^{-3}$ and $u_2 = 10^{-2}$ per cell division, and the population size ranges from $N = 10$ to 100, 1000 and 10,000 in Figs. 1a–d.

2.3. Advantageous mutations

Let us now discuss the case in which a type 1 cell has fitness $r_1 \geqslant 1$, and a type 2 cell is advantageous with relative fitness $r_2 > 1$. If population size and mutation rates are large, then the deterministic model, Eq. (2), applies and the expected number of type 2 cells at time t is given by $E(t) = Nx_2(t)$. If the population size is small, however, the stochasticity of the production and spread of mutated cells cannot be neglected. Hence we consider a stochastic approach. The dynamics can be decomposed into two parts: (i) the time of appearance of the first successful mutated cell, and (ii) the growth of this cell clone; the latter will be approximated by a deterministic trajectory.

2.3.1. Fitness values 1, 1, r with r > 1

Assume that the fitness values of type 0, 1 and 2 cells are 1, 1 and $r > 1$, respectively. This case represents genes that are homozygously sufficient (recessive), i.e. one wild type allele is enough to maintain the normal function of the
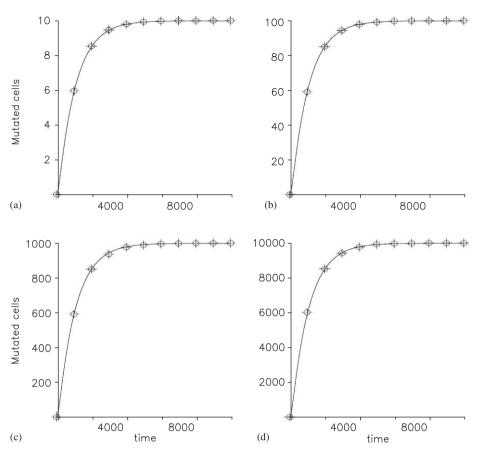
Fig. 1. The figure shows the expected number of type 2 cells if both mutations are neutral as compared to a cell with two wild type alleles of the metastasis suppressor gene, $r_1 = r_2 = 1$. We compare the results of the exact stochastic computer simulation of the two-hit process (circles) with the numerical simulation of Eq. (3) (line) at time t. The mutation rates are $u_1 = 10^{-3}$ and $u_2 = 10^{-2}$ per cell division, and the population size is (a) $N = 10$, (b) $N = 100$, (c) $N = 1000$, and (d) $N = 10,000$.

gene. If the population size of the main tumor is much less than the inverse of the mutation rates, $N \ll 1/u_1$ and $N \ll 1/u_2$, we can assume that the tumor is almost always homogeneous, i.e. consists only of a single cell type. In that limit, a mutation will either go extinct or take over the tumor, and two mutations will not coexist at the same time. Denote the probabilities that the population consists only of type 0, 1 or 2 cells by X_0, X_1, and X_2, respectively. Then we have

$$dX_0/dt = -bX_0 - dX_0,$$
$$dX_1/dt = bX_0 - cX_1,$$
$$dX_2/dt = dX_0 + cX_1. \qquad (4)$$

Here $b = u_1$, $c = Nu_2\rho(r)$ and $d = Nu_1\left[\sqrt{u_2\rho(r)} - 1/N\right]_+$ (Iwasa et al., 2004). The notation $[x]_+$ means $\max\{0, x\}$. The process described by d is called stochastic tunneling (Iwasa et al., 2004). With the initial conditions $X_0(0) = 1$ and $X_1(0) = X_2(0) = 0$, the expected fraction of type 2 cells is given by dX_2/dt. The growth of the lineage is described by

$$Z(t) = 1/(1 + (N - 1)e^{-(r-1)t}).$$

Then the number of type 2 cells at time t is given by

$$E(t) = N \int_0^t Z(t - t') \frac{dX_2}{dt}(t') \, dt'$$
$$= N \int_0^t \frac{1}{1 + (N - 1)e^{-(r-1)(t-t')}}$$
$$\times \left(u_1 e^{-2u_1 t'} + c \frac{u_1}{2u_1 - c} \left(e^{-ct'} - e^{-2u_1 t'} \right) \right) dt'. \qquad (5)$$

Fig. 2 compares the results of the exact stochastic computer simulation with the numerical simulation of Eqs. (2) and (5). We plot the expected number of type 2 cells, E, at time t. The mutation rates are $u_1 = 10^{-3}$ and $u_2 = 10^{-2}$ per cell division, and the population size ranges from $N = 10$ to 100, 1000 and 10,000 in Figs. 2a–d.

2.3.2. Fitness values 1, r, r with $r > 1$

Now assume that type 1 and 2 cells are advantageous and have the same relative fitness, $r_1 = r_2 = r > 1$; hence the fitness values of type 0, 1 and 2 cells are 1, r and r. This case represents genes that are homozygously insufficient (dominant), i.e. one wild type allele alone cannot maintain the normal function of the gene. In this case and with
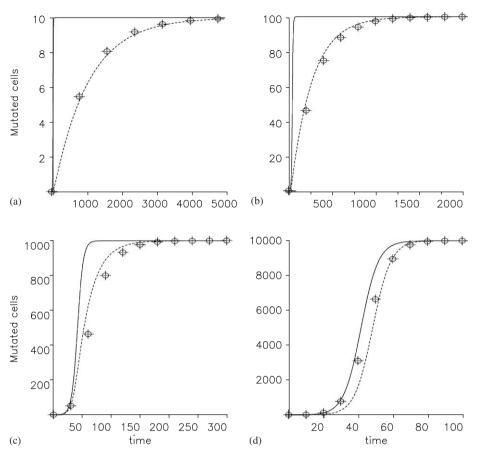
Fig. 2. The figure shows the expected number of type 2 cells if type 1 cells are neutral, $r_1 = 1$, and type 2 cells are advantageous, $r_2 = 1.2$. We compare the results of the exact stochastic computer simulation of the two-hit process (circles) with the numerical simulation of Eq. (2) (solid line) and Eq. (5) (broken line) at time t. The mutation rates are $u_1 = 10^{-3}$ and $u_2 = 10^{-2}$ per cell division, and the population size is (a) $N = 10$, (b) $N = 100$, (c) $N = 1000$, and (d) $N = 10,000$.

$b = Nu_1\rho(r)$, the expected number of type 2 cells at time t is given by

$$E(t) = N \int_0^t e^{-bt'} b \, dt' (1 - e^{-u_2(t-t')})$$
$$= \frac{N}{b - u_2}(b - u_2 - be^{-u_2t} + u_2e^{-bt}). \quad (6)$$

Fig. 3 shows the results of the exact stochastic computer simulation and the numerical simulation of Eqs. (2) and (6). We plot the expected number of type 2 cells, E, at time t. The mutation rates are $u_1 = 10^{-3}$ and $u_2 = 10^{-2}$ per cell division, and the population size ranges from $N = 10$ to 100, 1000 and 10,000 in Figs. 3a–d.

2.3.3. Fitness values $1 < r_1 < r_2$

Assume that both type 1 and 2 cells are advantageous and have relative fitness $1 < r_1 < r_2$, and the population size is small. The probabilities that the tumor consists only of type 0, 1, or 2 cells at time t are again denoted by $X_0(t)$, $X_1(t)$ and $X_2(t)$. Tunneling can be neglected. We have Eq. (4) with $b = Nu_1\rho(r_1)$, $c = Nu_2\rho(r_2/r_1)$ and $d = 0$. With the initial conditions $X_0(0) = 1$ and $X_1(0) = X_2(0) = 0$, the

expected number of type 2 cells at time t is given by

$$E(t) = \int_0^t \frac{dX_2}{dt}(s) \frac{N}{1 + (N - 1)e^{-(r_2-r_1)(t-s)}} \, ds$$
$$= \int_0^t \frac{bc}{b - c}(e^{-cs} - e^{-bs}) \frac{N}{1 + (N - 1)e^{-(r_2-r_1)(t-s)}} \, ds. \quad (7)$$

Fig. 4 shows the results of the exact stochastic computer simulation and the numerical simulation of Eqs. (2) and (7). We plot the expected number of type 2 cells, E, at time t. The mutation rates are $u_1 = 10^{-3}$ and $u_2 = 10^{-2}$ per cell division, and the population size ranges from $N = 10$ to 100, 1000 and 10,000 in Figs. 4a–d.

2.3.4. Summary of the advantageous parameter regimes

The deterministic formula, Eq. (2), predicts a faster spread of mutated cells than the corresponding stochastic formulas, Eqs. (5)–(7). The stochastic computer simulations show that the exact number of cells generally lies between these two predictions. It is closer to the prediction by the deterministic formula when the population size and mutation rates are large and the selective advantage is small. It is closer to the prediction by the stochastic model,
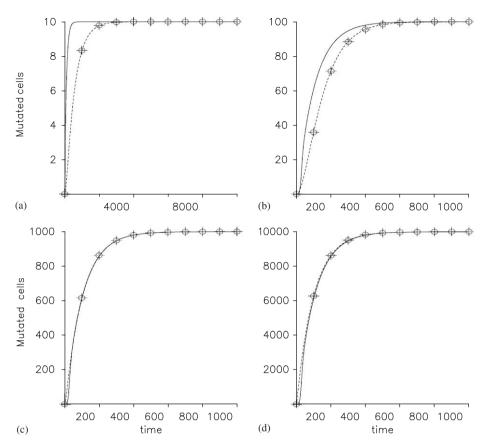
Fig. 3. The figure shows the expected number of type 2 cells if both type 1 and 2 cells are advantageous and have the same fitness, $r_1 = r_2 = 1.2$. We compare the results of the exact stochastic computer simulation of the two-hit process (circles) with the numerical simulation of Eq. (2) (solid line) and Eq. (6) (broken line) at time t. The mutation rates are $u_1 = 10^{-3}$ and $u_2 = 10^{-2}$ per cell division, and the population size is (a) $N = 10$, (b) $N = 100$, (c) $N = 1000$, and (d) $N = 10,000$.

however, when the population size and mutation rates are small and selective advantage is large (see Figs. 2–4). Therefore, both the deterministic and the stochastic approach have to be used to obtain a valid prediction for the number of type 2 cells in a tumor.

2.4. Deleterious mutations

We will now discuss the situations in which type 1 cells have relative fitness $r_1 \leqslant 1$, and type 2 cells have relative fitness $r_2 < 1$.

2.4.1. Fitness values 1, 1, r with r < 1

First assume that type 1 cells are neutral, $r_1 = 1$, and type 2 cells are deleterious with relative fitness $r_2 = r < 1$; hence the fitness values of type 0, 1 and 2 cells are 1, 1, and r. This case again represents genes that are homozygously sufficient (recessive). The dynamics can be decomposed into three different phases. During the first phase, the population is dominated by type 0 cells, but a fraction of type 1 cells is maintained. Type 1 cells produce an even smaller fraction of type 2 cells. The second phase starts when type 1 cells take over the population and type 0 cells go extinct. Then the population consists of type 1 and 2 cells. The third phase begins once type 2 cells reach

fixation. The lengths of these phases depend on the mutation rates, the population size, and the relative fitness of type 2 cells. The formulas are derived in Appendix A.

Phase I: During the first phase, type 1 cells coexist with type 0 cells at an approximately constant proportion. Furthermore, fixation of type 1 cells occurs much later than expected of a neutral mutant. These observations suggest that type 1 cells do not behave like neutral mutants, but have a slight fitness disadvantage. The observed fitness disadvantage is caused by the constant mutation of type 0 cells to give rise to type 1 cells and the mutation of type 1 cells to give rise to type 2 cells. Therefore, type 1 cells can be maintained at the mutation-selection balance expected for slightly deleterious mutants. During the first phase, the fractions of type 1 cells, x_1, and type 2 cells, x_2, are given by

$$x_1 = \frac{e^{(u_2 - u_1)t} - 1}{(u_2/u_1)e^{(u_2 - u_1)t} - 1}, \tag{8a}$$

$$x_2 = \frac{u_2}{1 - r}x_1. \tag{8b}$$

Here we neglect the time delay of $1/(1 - r)$. See Appendix A for derivations.

Phase II: The second phase starts once type 0 cells go extinct and type 1 cells reach fixation. Then the fraction of
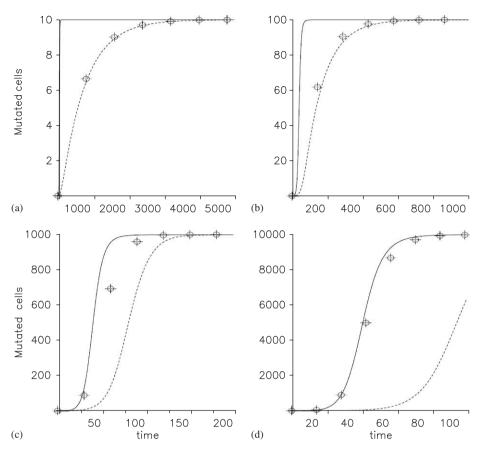
Fig. 4. The figure shows the expected number of type 2 cells if both type 1 and 2 cells are advantageous and have fitness values $r_1 = 1.1$ and $r_2 = 1.2$. We compare the results of the exact stochastic computer simulation of the two-hit process (circles) with the numerical simulation of Eq. (2) (solid line) and Eq. (7) (broken line) at time t. The mutation rates are $u_1 = 10^{-3}$ and $u_2 = 10^{-2}$ per cell division, and the population size is (a) $N = 10$, (b) $N = 100$, (c) $N = 1000$, and (d) $N = 10,000$.

type 2 cells is given by

$$x_2 = \frac{u_2}{1-r}. \tag{9}$$

Eq. (9) exceeds Eq. (8) because $u_1 < u_2$, and the latter converges to a value of u_1/u_2 of the former. The time of fixation of type 1 cells is random, and the average fraction of type 2 cells before and after the fixation of type 1 cells is given by Eqs. (8) and (9). The effective fitness of type 1 cells is given by $(1 - u_2)/(1 - u_1)$, which is less than one. Therefore, the fixation probability of type 1 cells is given by $\rho((1 - u_2)/(1 - u_1)) < 1/N$, and the rate of their fixation in the population is given by

$$a = Nu_1 \rho \left(\frac{1 - u_2}{1 - u_1} \right) = Nu_1 \frac{u_2 - u_1}{\exp[N(u_2 - u_1)] - 1}.$$

The frequency of type 2 cells is given by

$$x_2 = \frac{u_2}{1-r} \frac{e^{(u_2-u_1)t} - 1}{(u_2/u_1)e^{(u_2-u_1)t} - 1} e^{-at} + \frac{u_2}{1-r}(1 - e^{-at}).$$

Phase III: The third phase begins once type 2 cells reach fixation. This transition rate is given by $b = Nu_2(1 - 1/r)/(1 - 1/r^N)$.

Denote by T_1 and T_2 the times at which type 1 and 2 cells reach fixation in the population. There are three possibilities: (i) type 1 cells have not yet reached fixation and the population is dominated by type 0 cells, $t < T_1$; (ii) type 1 cells but not type 2 cells have reached fixation, $T_1 < t < T_2$; and (iii) type 2 cells have reached fixation, $t > T_2$. These probabilities are given by

$$Pr[t < T_1] = e^{-at},$$

$$Pr[T_1 < t < T_2] = \frac{a}{a-b}(e^{-bt} - e^{-at}),$$

$$Pr[t > T_2] = 1 - \frac{a}{a-b}e^{-bt} + \frac{b}{a-b}e^{-at}.$$

The expected fraction of type 2 mutants is given by

$$\begin{aligned} E = {} & N \frac{u_2}{1-r} \frac{e^{(u_2-u_1)t} - 1}{(u_2/u_1)e^{(u_2-u_1)t} - 1} e^{-at} \\ & + N \frac{u_2}{1-r} \frac{a}{a-b}(e^{-bt} - e^{-at}) \\ & + N \left(1 - \frac{a}{a-b}e^{-bt} + \frac{b}{a-b}e^{-at}\right). \end{aligned} \tag{10}$$
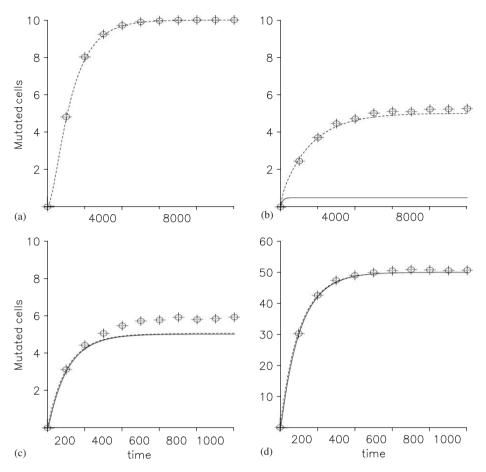
Fig. 5. The figure shows the expected number of type 2 cells if type 1 cells are neutral, $r_1 = 1$, and type 2 cells are deleterious with relative fitness $r_2 = 0.8$. We compare the results of the exact stochastic computer simulation of the two-hit process (circles) with the numerical simulation of Eq. (2) (solid line) and (a,b,d) Eq. (10) and (c) Eq. (A.1) (broken line) at time t. The mutation rates are $u_1 = 10^{-3}$ and $u_2 = 10^{-2}$ per cell division, and the population size is (a) $N = 10$, (b) $N = 100$, (c) $N = 1000$, and (d) $N = 10,000$.

This formula accurately predicts the fraction of type 2 cells in populations of size $N = 10$, 100, and 10,000, but gives a 10% lower estimate if $N = 1000$. A more accurate formula for the latter parameter regime can be obtained by considering the variance of the fraction of type 1 mutants before reaching fixation (see Appendix A, Eq. (A.1)).

Fig. 5 shows the results of the exact stochastic computer simulation of the two-hit process and the numerical simulation of Eqs. (2),(10) and (A.1). We plot the expected number of type 2 cells, E, at time t. The mutation rates are $u_1 = 10^{-3}$ and $u_2 = 10^{-2}$ per cell division, and the population size ranges from $N = 10$ to 100, 1000 and 10,000 in Figs. 5a–d.

2.4.2. Fitness values 1, r, r with r < 1

Now assume that type 1 and 2 cells are deleterious and have the same relative fitness, $r_1 = r_2 = r < 1$; hence the fitness values of type 0, 1 and 2 cells are 1, r, and r, and the gene is homozygously insufficient (recessive deleterious mutation). As explained in Appendix B, the expected

number of type 2 cells at time t is given by

$$
E(t) = N e^{-(b+c)t} \frac{r u_1 u_2}{(1 - r)^2} \left[1 - (1 + (1 - r)t) e^{-(1-r)t} \right]
$$
$$
+ N \left[\frac{b}{b+c} (1 - e^{-(b+c)t}) \right.
$$
$$
\left. - b \left(1 - \frac{r u_2}{1 - r} \right) \frac{e^{-(b+c)t} - e^{-r u_2 t}}{r u_2 - b - c} \right]
$$
$$
+ N \frac{c}{b+c} (1 - e^{-(b+c)t}).
\tag{11}
$$

Here $b = N u_1 \rho(r)$ and $c = [N u_1 / (1 - r)] r u_2 \rho(r)$.

Fig. 6 shows the results of the exact stochastic computer simulation of the two-hit process and the numerical simulation of Eqs. (2) and (11). We plot the expected number of type 2 cells, E, at time T. The mutation rates are $u_1 = 10^{-3}$ and $u_2 = 10^{-2}$ per cell division, and the population size ranges from $N = 10$ to 100, 1000 and 10,000 in Figs. 6a–d.

2.4.3. Fitness values $1 > r_1 > r_2$

Finally, assume that both type 1 and 2 cells are deleterious and have fitness values $1 > r_1 > r_2$. In this case,
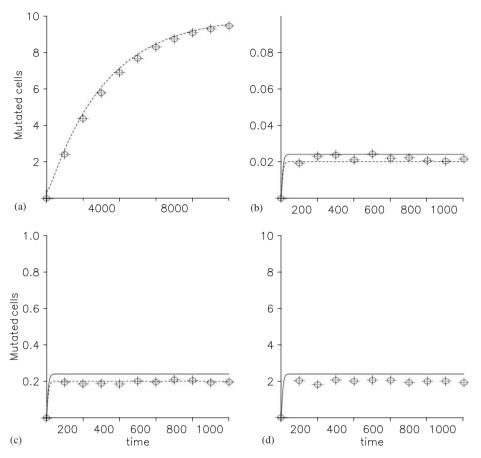
Fig. 6. The figure shows the expected number of type 2 cells if both type 1 and 2 cells are deleterious and have the same relative fitness, $r_1 = r_2 = 0.8$. We compare the results of the exact stochastic computer simulation of the two-hit process (circles) with the numerical simulation of Eq. (2) (solid line) and Eq. (11) (broken line) at time t. The mutation rates are $u_1 = 10^{-3}$ and $u_2 = 10^{-2}$ per cell division, and the population size is (a) $N = 10$, (b) $N = 100$, (c) $N = 1000$, and (d) $N = 10,000$.

we consider tunneling and the mutation-selection balance in the population. Let $b = Nu_1\rho(r_1)$, $c = Nr_1u_2\rho(r_2/r_1)$, and $d = [Nu_1/(1 - r_1)]r_1u_2\rho(r_2)$. Then the expected number of type 2 cells before the fixation of type 1 cells is given by

$$E_0(t) = N\frac{r_1u_2}{1 - r_2}\frac{u_1}{1 - r_1}(1 - e^{-(1-r_1)t})e^{-(b+d)t}. \tag{12a}$$

The expected number of type 2 cells after the fixation of type 1 cells, but before the fixation of type 2 cells, is given by

$$E_1(t) = N\frac{r_1u_2}{r_1 - r_2}\int_0^t e^{-(b+d)s}e^{-c(t-s)}b\,\mathrm{d}s$$

$$= N\frac{r_1u_2}{r_1 - r_2}b\frac{e^{-ct} - e^{-(b+d)t}}{b + d - c}. \tag{12b}$$

The expected number of type 2 cells after the fixation of type 2 cells is given by

$$E_2(t) = N\int_0^t e^{-(b+d)s}d\,\mathrm{d}s + N\int_0^t e^{-(b+d)s}(1 - e^{-c(t-s)})b\,\mathrm{d}s$$

$$= N\frac{d}{b + d}(1 - e^{-(b+d)t})$$

$$+ Nb\left(\frac{1 - e^{-(b+d)t}}{b + d} - \frac{e^{-ct} - e^{-(b+d)t}}{b + d - c}\right). \tag{12c}$$

Here the first term accounts for tunneling and the second for the two-step evolution. In total, the expected number of type 2 cells given by

$$E(t) = E_0(t) + E_1(t) + E_2(t). \tag{12d}$$

Fig. 7 shows the results of the exact stochastic computer simulation of the two-hit process and the numerical simulation of Eqs. (2) and (12). We plot the expected number of type 2 cells, E, at time t. The mutation rates are $u_1 = 10^{-3}$ and $u_2 = 10^{-2}$ per cell division, and the population size ranges from $N = 10$ to 100, 1000 and 10,000 in Figs. 7a–d.

3. Discussion

In a previous paper (Michor et al., 2006), we have discussed the dynamics of metastasis formation if one genetic alteration is sufficient to confer metastatic ability to the cell. In the present paper, we extend the analysis to situations in which both alleles of a MSG need to be inactivated such that a cell can metastasize.

Let us call cells with two wild type alleles of the MSG type 0, cells with one inactivated allele type 1, and cells with two inactivated alleles type 2 cells. In this paper, we calculate the risk of metastasis formation of tumors as a
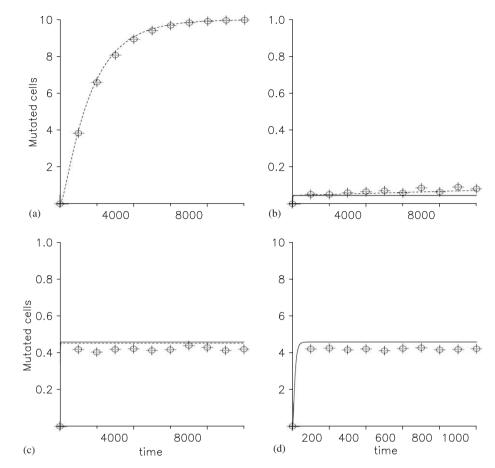
Fig. 7. The figure shows the expected number of type 2 cells if both type 1 and 2 cells are deleterious and have fitness values $r_1 = 0.9$ and $r_2 = 0.8$. We compare the results of the exact stochastic computer simulation of the two-hit process (circles) with the numerical simulation of Eq. (2) (solid line) and Eq. (12) (broken line) at time t. The mutation rates are $u_1 = 10^{-3}$ and $u_2 = 10^{-2}$ per cell division, and the population size is (a) $N = 10$, (b) $N = 100$, (c) $N = 1000$, and (d) $N = 10,000$.

function of their population size, N, the rates at which the two MSG alleles are inactivated per cell division, u_1 and u_2, the fitness values of type 1 and 2 cells, r_1 and r_2, and the rate of export of type 2 cells from the main tumor, q. We obtain different formulas for different parameter regimes.

If MSG-deficient cells have a fitness advantage, then they have a large probability of taking over the entire main tumor. Therefore, if metastases are caused by mutations that are positively selected in the main tumor, then metastatic ability will be the property of all (or the majority of) cells in the tumor. Conversely, if MSG-deficient cells have a fitness disadvantage, then they will be maintained at a quasi-equilibrium determined by the mutation-selection balance in the main tumor. This quasi-equilibrium is temporary and eventually, one cell type will take over the tumor. Nevertheless, the presence of the quasi-equilibrium is very important for calculating the risk of metastasis. If metastases are caused by mutations that are negatively selected in the main tumor, then metastatic ability will be the property of only a small fraction of cells in the tumor (Fig. 8). Hence, most metastases will arise by advantageous mutations as long as both advantageous and deleterious mutants have

comparable rates of being exported from the main tumor and forming metastases elsewhere. Only if the export rate or efficiency at forming metastases is greatly enhanced in deleterious mutants, then they will outperform the advantageous ones (Table 2).

However, we can also imagine scenarios where deleterious mutations are more effective. If the effective population size of the tumor is small, then deleterious mutations are more effective at taking over and hence producing metastases (Fig. 8c). The main cancer can be subdivided into small spatial compartments, again rendering deleterious mutations more important. Also, there could be many more genes that confer metastatic ability together with a fitness disadvantage than metastasis-promoting genes with a fitness advantage.

Even if type 1 cells are perfectly neutral as compared to type 0 cells, they behave as slightly deleterious mutants if the rate inactivating the second MSG allele is larger than the rate inactivating the first MSG allele. In that case, type 1 cells are mutating to become type 2 cells more quickly than they are being replenished by mutating type 0 cells. This effect represents a fitness loss for type 1 cells. This observation is not important in conventional population
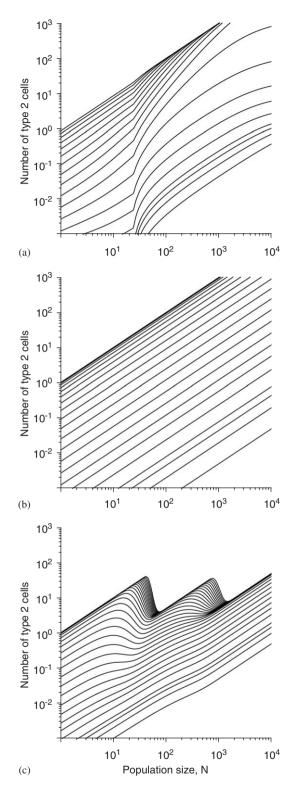
(a)

(b)

(c)

Population size, N

Fig. 8. The figure shows the dependence of the number of type 2 cells on the population size of the cancer. (a) Type 0, 1 and 2 cells have fitness values 1, 1 and r with $r > 1$. We show the numerical simulation of Eq. (5) with $r = 1.2$. (b) Type 0, 1 and 2 cells all have fitness value 1. We show Eq. (3). (c) Type 0, 1 and 2 cells have fitness values 1, 1 and r with $r < 1$. We show Eq. (10) with $r = 0.8$. The lines represent different times, with time flowing upwards in the graph. The mutation rates are $u_1 = 10^{-3}$ and $u_2 = 10^{-2}$ per cell division.

Table 2
The expected number of metastases if type 1 and type 2 cells are deleterious ($r_1 = r_2 = 0.9$), neutral ($r_1 = r_2 = 1$), or advantageous ($r_1 = r_2 = 1.1$)

	$r_1 = r_2 = 0.9$	$r_1 = r_2 = 1.0$	$r_1 = r_2 = 1.1$
$q = 100$	37	> 1000	> 1000
$q = 10$	4	813	> 1000
$q = 1$	0	81	> 1000
$q = 0.1$	0	8	> 1000
$q = 0.01$	0	1	> 1000
$q = 0.001$	0	0	> 1000
$q = 0.0001$	0	0	> 1000
$q = 0.00001$	0	0	130
$q = 0.000001$	0	0	13

Mutated cells are exported at rate q from the main tumor to form metastases elsewhere. The main tumor consists of $N = 10^8$ cells. Advantageous mutations are likely to reach fixation in the main tumor and hence are much more successful in establishing metastases. Neutral and deleterious mutations, however, are maintained at low levels in the main tumor. Deleterious cells are as successful as advantageous cells only if their rate of export is 10^8-fold higher (grey cells). It is therefore very unlikely that mutations that are negatively selected in the main tumor are responsible for most metastases produces by the tumor. We use Eqs. (2) and (3), and parameter values $N = 10^8$, $u_1 = 10^{-7}$ and $u_2 = 10^{-6}$.

genetics where the mutation rate is assumed to be much smaller than the fitness difference between cell types. For carcinogenesis, however, the mutation rates can be as large as or even larger than the selection coefficient, especially if the tumor has already evolved some kind of genetic instability (Lengauer et al., 1998; Nowak et al., 2002).

The results presented in this paper do not only apply to the inactivation of a MSG in a tumor. We can also calculate the number of cells harboring mutations in two independent genes, both of which are necessary for metastatic ability. For example, mutation of two onco-genes could be needed to confer metastatic propensity to a cell. Furthermore, our results are relevant to the dynamics of tumor suppressor gene inactivation. In previous work, we have calculated the probability of inactivating a tumor suppressor gene in a population of cancer cells (Nowak et al., 2004; Iwasa et al., 2005). In this paper, we extend the analysis to calculate the expected number of cells with an inactivated tumor suppressor gene in a cancer. Hence the present paper contributes not only to a quantitative understanding of metastasis, but also to (early) tumorigenesis.

In this paper, we assume that the population size of the tumor is approximately constant over time. This assumption applies for tumors that expand very slowly, and for tumors that cannot grow further until accumulating another mutation. It is also possible to assume exponential growth of the cancer. The dynamics of cells that are generated by a single mutation in an exponentially growing population has been investigated in a previous paper (Iwasa et al., 2006). The dynamics of cells that are generated by two mutations in an exponentially growing population is the topic of ongoing investigation.

A quantitative understanding of metastasis dynamics crucially relies on a knowledge of the fitness values of cells harboring metastasis-promoting mutations. Only if these fitness values are known, a prediction can be made about the dynamics of different mutations. Therefore, it should be an important goal of the field to investigate fitness effects of metastasis-promoting alterations.

Appendix A

Let x_0, x_1, and x_2 be the fraction of type 0, 1 and 2 cells, respectively. We have $x_0 + x_1 + x_2 = 1$. Their fitness values are 1, 1, and $r \leqslant 1$. Type 0 cells produce type 1 cells at rate u_1 per cell division, and type 1 cells produce type 2 cells at rate u_2 per cell division. Initially, the population is dominated by type 0 cells. The change in the fraction of type 1 cells per generation is the sum of the effects of selection, mutation, and drift. Even though type 1 cells are neutral, the selection term is positive, because they constantly produce type 2 cells which have fitness $r < 1$. Therefore, the mean population fitness is less than one, and there is positive selection favoring both type 0 and 1 cells. The mean fitness is given by $\bar{w} = x_0 + x_1 + rx_2 = 1 - (1-r)x_2$. Hence the selection term for type 1 cells is given by

$$(\Delta x_1)_{\text{sel}} = \left(\frac{1}{1-(1-r)x_2} - 1 \right) x_1 = \frac{(1-r)x_2}{1-(1-r)x_2} x_1.$$

Therefore, we have

$$\Delta x_1 = u_1(1 - x_1 - x_2) - u_2 x_1 + \frac{(1-r)x_2}{1-(1-r)x_2} x_1 + \text{drift},$$

$$\Delta x_2 = u_2 x_1 - \frac{(1-r)(1-x_2)}{1-(1-r)x_2} x_2 + \text{drift}.$$

Let us first consider the neutral case ($r = 1$). In this case, the selection terms vanishes. We then introduce the arithmetic average of the gene frequency. Since all the remaining terms are either the first order or the zeroth order with respect to gene frequencies x_1 and x_2, we can derive Eq. (2) without selection terms. This implies that the deterministic model Eq. (2) holds exactly if we regard x_1 and x_2 as arithmetic averages.

For non-neutral cases ($r \neq 1$), we cannot apply the average without considering higher order terms, and we need a different way to simplify the dynamics. Note that the mutation rates, u_1 and u_2, are both small ($O(u) \ll O(1)$), and $1-r$ is not small ($O(1)$). If we consider the situation in which neither type 1 nor type 2 cells reach fixation, then x_2 is maintained at a low level ($O(u)$), but x_1 is not small ($O(1)$). Calculating the averages removes the drift terms and the approximation of difference equations by differential equations gives

$$dx_1/dt = u_1 - (u_1 + u_2)x_1 + (1-r)x_1 x_2 + O(u),$$

$$dx_2/dt = u_2 x_1 - (1-r)x_2 + O(u^2).$$

The average fraction of type 2 cells quickly converges to $x_2 = x_1 u_2/(1-r)$. Therefore, we have

$$dx_1/dt = u_1 - (u_1 + u_2)x_1 + u_2 x_1^2.$$

We approximate the difference equation by a differential equation and solve by integration by parts to obtain Eq. (8a) in the text. Note that

$$\lim_{t \to \infty} x_1 = \begin{cases} 1 & \text{if } u_1 > u_2, \\ u_1/u_2 & \text{if } u_1 < u_2. \end{cases}$$

A.1. The effect of variance

Let us now derive a formula for the quasi-equilibrium distribution of type 1 cells before reaching fixation. It follows a diffusion process. The stochastic differential equation and the diffusion equation (above) are respectively given by

$$dZ = (u_2 Z - u_1)(Z - 1)\,dt + \sqrt{\frac{2}{N}Z(1-Z)}\,dW,$$

$$\frac{\partial p}{\partial t} = -\frac{\partial}{\partial z}\{(u_2 z - u_1)(z-1)p\} + \frac{1}{2}\frac{\partial^2}{\partial z^2}\left\{\frac{2}{N}z(1-z)p\right\}.$$

We neglect the third order and higher moments in order to close the dynamics. With $f(z) = (u_2 z - u_1)(z-1)$ and $g(z) = (2/N)z(1-z)$, we have

$$\frac{d}{dt}\bar{z} = u_1 - (u_1 + u_2)\bar{z} + u_2\bar{z}^2 + u_2 v,$$

$$\frac{d}{dt}v = 2(-(u_1 + u_2) + 2u_2\bar{z})v + \frac{2}{N}\bar{z}(1-\bar{z}) + \frac{1}{2}\left(\frac{-4}{N}\right)v.$$

Based on these equations, Eq. (8) gives a greater value of $\bar{z}$ than it would without the variance.

The calculation of the above equation generates the mean frequency of type 1 cells and its variance. The effective relative fitness of type 1 cells over type 2 cells is

$$\frac{1-u_2}{1-u_2(u_1/u_2)-u_1(1-u_1/u_2)} = \frac{1-u_2}{1-2u_1+u_1^2/u_2}.$$

Let the fixation rate of type 1 cells be

$$a = u_1 \frac{N((1-2u_1+u_1^2/u_2)/(1-u_2)-1)}{((1-2u_1+u_1^2/u_2)/(1-u_2))^N - 1}.$$

This rate is slower than the rate of fixation of a neutral mutant, u_1. Let the rate of fixation of type 2 cells be $b = Nu_2(1-1/r)/(1-1/r^N)$. With these assumptions, we

can calculate the expected number of type 2 cells as

$$E(t) = N \frac{u_2}{1-r} \bar{z} e^{-at} + N \frac{u_2}{1-r} \frac{a}{a-b} (e^{-bt} - e^{-at})$$
$$+ N \left(1 - \frac{a}{a-b} e^{-bt} + \frac{b}{a-b} e^{-at} \right). \quad \text{(A.1)}$$

Appendix B

Consider the situation in which type 1 and 2 cells are deleterious and have the same relative fitness, $r_1 = r_2 = r < 1$. In this case, the frequencies of type 1 and 2 cells are given by

$$\frac{\mathrm{d}x_1}{\mathrm{d}t} = u_1(1 - x_1 - x_2) - ru_2x_1 + \frac{(r-1)(1-x_1-x_2)}{1+(r-1)(x_1+x_2)}x_1$$
$$+ [\text{random drift}],$$
$$\frac{\mathrm{d}x_2}{\mathrm{d}t} = ru_2x_1 + \frac{(r-1)(1-x_1-x_2)}{1+(r-1)(x_1+x_2)}x_2 + [\text{random drift}].$$

By taking the arithmetic average with respect to the runs, $m_1 = E[x_1]$ and $m_2 = E[x_2]$, and if $x_1 \ll 1$ and $x_2 \ll 1$, we have

$$\frac{\mathrm{d}m_1}{\mathrm{d}t} = u_1 - (1-r)m_1,$$
$$\frac{\mathrm{d}m_2}{\mathrm{d}t} = ru_2m_1 - (1-r)m_2.$$

The initial conditions are given by $m_1(0) = m_2(0) = 0$. From these, we have

$$m_2(t) = \frac{ru_1u_2}{(1-r)^2}[1 - (1 + (1-r)t)e^{-(1-r)t}].$$

This holds before the fixation of type 1. After the fixation of type 1 cells, which occurs at t', we have

$$m_2(t) = 1 - \left(1 - \frac{ru_2}{1-r}\right)e^{-ru_2(t-t')}.$$

Here the initial condition is given by the mutation selection balance.

Let $b = Nu_1\rho(r)$ and $c = [Nu_1/(1-r)]ru_2\rho(r)$. At time t, there are three different possibilities: (i) neither type 1 nor type 2 cells are fixed, (ii) type 1 cells are fixed before time t, and (iii) type 2 cells are fixed before time t. Then we have the following estimate for the fraction of type 2 cells:

$$E[x_2(t)] = e^{-(b+c)t} \frac{ru_1u_2}{(1-r)^2}[1 - (1 + (1-r)t)e^{-(1-r)t}]$$
$$+ \int_0^t e^{-(b+c)t'} b \, \mathrm{d}t' \left(1 - \left(1 - \frac{ru_2}{1-r}\right)e^{-ru_2(t-t')}\right)$$
$$+ \int_0^t e^{-(b+c)t'} c \, \mathrm{d}t' \cdot 1.$$

Therefore, the expected number of type 2 cells at time t is given by Eq. (11) in text.

References

Anderson, A.R., Chaplain, M.A., 1998. Continuous and discrete mathematical models of tumor-induced angiogenesis. Bull. Math. Biol. 60, 857–899.

Armitage, P., Doll, R., 1954. The age distribution of cancer and a multi-stage theory of carcinogenesis. Br. J. Cancer 8, 1–12.

Armitage, P., Doll, R., 1957. A two-stage theory of carcinogenesis in relation to the age distribution of human cancer. Br. J. Cancer 11, 161–169.

Backer, J.M., Mendola, C.E., Kovesdi, I., Fairhurst, J.L., O'Hara, B., Eddy, R.L., Shows, T.B., Mathew, S., Murty, V.V., Chaganti, R.S., 1993. Chromosomal localization and nucleoside diphosphate kinase activity of human metastasis-suppressor genes NM23-1 and NM23-2. Oncogene 8, 497–502.

Bosl, G.J., Geller, N.L., Cirrincione, C., Vogelzang, N.J., Kennedy, B.J., Whitmore, W.F., Vugrin, D., Scher, H., Nisselbaum, J., Golbey, R.B., 1983. Multivariate analysis of prognostic variables in patients with metastatic testicular cancer. Cancer Res. 43, 3403–3407.

Chambers, A., Naumov, G., Vantyghem, S., Tuck, A., 2000. Molecular biology of breast cancer biology: clinical implications of experimental studies on metastatic inefficiency. Breast Cancer Res. 2, 400–407.

Chambers, A., Naumov, G.N., Varghese, H.J., Nadkarni, K.V., MacDonald, I.C., Groom, A.C., 2001. Critical steps in hematogenous metastasis: an overview. Surg. Oncol. Clin. North Am. 10, 243–255.

Chambers, A., Groom, A., MacDonald, I., 2002. Dissemination and growth of cancer cells in metastatic sites. Nat. Rev. Cancer 2, 563–572.

Chang, S., Khoo, C., DePinho, R.A., 2001. Modeling chromosomal instability and epithelial carcinogenesis in the telomerase-deficient mouse. Semin. Cancer Biol. 11, 227–239.

Chang, S., Khoo, C., Naylor, M.L., Maser, R.S., DePinho, R.A., 2003. Telomere-based crisis: functional differences between telomerase activation and ALT in tumor progression. Genes Dev. 17, 88–100.

Delsanto, P.P., Romano, A., Scalerandi, M., Pescarmona, G.P., 2000. Analysis of a "phase transition" from tumor growth to latency. Phys. Rev. E 62, 2547–2554.

Fisher, J.C., 1959. Multiple-mutation theory of carcinogenesis. Nature 181, 651–652.

Frank, S.A., Iwasa, Y., Nowak, M.A., 2003. Patterns of cell division and the risk of cancer. Genetics 163, 1527–1532.

Gatenby, R.A., Maini, P.K., 2003. Mathematical oncology: cancer summed up. Nature 421, 321.

Gatenby, R.A., Vincent, T.L., 2003. An evolutionary model of carcinogenesis. Cancer Res. 63, 6212–6220.

Goldie, J.H., Coldman, A.J., 1979. A mathematic model for relating the drug sensitivity of tumors to their spontaneous mutation rate. Cancer Treat. Rep. 63, 1727–1733.

Goldie, J.H., Coldman, A.J., 1983. Quantitative model for multiple levels of drug resistance in clinical tumors. Cancer Treat. Rep. 67, 923–931.

Iwasa, Y., Michor, F., Nowak, M.A., 2004. Stochastic tunnels in evolutionary dynamics. Genetics 166, 1571–1579.

Iwasa, Y., Michor, F., Nowak, M.A., 2006. Evolution of resistance in clonal expansion. Genetics, in press.

Iwasa, Y., Michor, F., Komarova, N.L., Nowak, M.A., 2005. Population genetics of tumor suppressor genes. J. Theor. Biol. 233, 15–23.

Kim, H., et al., 2001. Mitogen-activated protein kinase kinase 4 metastasis suppressor gene expression is inversely related to histologic pattern in advancing human prostatic cancer. Cancer Res. 61, 2833–2837.

Knudson, A.G., 2001. Two genetic hits to cancer. Nat. Rev. Cancer 1, 157–162.

Komarova, N.L., Lengauer, C., Vogelstein, B., Nowak, M.A., 2002. Dynamics of genetic instability in sporadic and familial colorectal cancer. Cancer Biol. Ther. 1, 685–692.

Komarova, N.L., Sengupta, A., Nowak, M.A., 2003. Mutation-selection networks of cancer initiation: tumor suppressor genes and chromosomal instability. J. Theor. Biol. 223, 433–450.

Lengauer, C., Kinzler, K.W., Vogelstein, B., 1998. Genetic instabilities in human cancers. Nature 396, 643–649.

Little, M.P., Wright, E.G., 2003. A stochastic carcinogenesis model incorporating genomic instability fitted to colon cancer data. Math. Biosci. 183, 111–134.

Luebeck, E.G., Moolgavkar, S.H., 2002. Multistage carcinogenesis and the incidence of colorectal cancer. Proc. Natl Acad. Sci. USA 99, 15095–15100.

Michor, F., Iwasa, Y., Komarova, N.L., Nowak, M.A., 2003. Local regulation of homeostasis favors chromosomal instability. Curr. Biol. 13, 581–584.

Michor, F., Iwasa, Y., Nowak, M.A., 2004a. Dynamics of cancer progression. Nat. Rev. Can. 4, 197–205.

Michor, F., Iwasa, Y., Rajagopalan, H., Lengauer, C., Nowak, M.A., 2004b. Linear model of colon cancer initiation. Cell Cycle 3, 358–362.

Michor, F., Nowak, M.A., Iwasa, Y., 2006. Stochastic dynamics of metastasis formation. J. Theor. Biol. in press.

Michor, F., Iwasa, Y., Hughes, T.P., Branford, S., Shah, N.P., Sawyers, C.L., Nowak, M.A., 2005a. Dynamics of chronic myeloid leukemia. Nature 435, 1267–1270.

Michor, F., Iwasa, Y., Lengauer, C., Nowak, M.A., 2005b. Dynamics of colorectal cancer. Semin. Cancer Biol. 15, 484–493.

Moolgavkar, S.H., Knudson, A.G., 1981. Mutation and cancer: a model for human carcinogenesis. J. Natl. Can. Inst. 66, 1037–1052.

Moran, P., 1962. The statistical processes of evolutionary theory. Clarendon Press, Oxford.

Nordling, C.O., 1953. A new theory on cancer-inducing mechanism. Br. J. Cancer 7, 68–72.

Nowak, M.A., et al., 2002. The role of chromosomal instability in tumor initiation. Proc. Natl Acad. Sci. USA 99, 16226–16231.

Nowak, M.A., Michor, F., Iwasa, Y., 2003. The linear process of somatic evolution. Proc. Natl Acad. Sci. USA 100, 14966–14969.

Nowak, M.A., Michor, F., Komarova, N.L., Iwasa, Y., 2004. Dynamics of tumor suppressor gene inactivation. Proc. Natl Acad. Sci. USA 101, 10635–10638.

Nunney, L., 1999. Lineage selection and the evolution of multistage carcinogenesis. Proc. R. Soc. London B 266, 493–498.

Owen, M.R., Sherrat, J.A., 1999. Mathematical modeling of macrophage dynamics in tumors. Math. Models Methods Appl. Biol. Chem. 377, 675–684.

Panetta, J.C., 1996. A mathematical model of periodically pulsed chemotherapy: tumor recurrence and metastasis in a competitive environment. Bull. Math. Biol. 58, 425–447.

Pescarmona, G.P., Scalerandi, M., Delsanto, P.P., Condat, C.A., 1999. Non-linear model of cancer growth and metastasis: a limiting nutrient as a major determinant of tumor shape and diffusion. Med. Hypotheses 53, 497–503.

Pozzatti, R., Muschel, R., Willians, J., et al., 1986. Primary rat embryo cells transformed by one or two oncogenes show different metastatic potentials. Science 232, 223–227.

Saunders, M., et al., 2001. Breast cancer metastatic potential correlates with the breakdown in homospecific and heterospecific gap junctional intercellular communication. Cancer Res. 61, 1765–1767.

Seraj, M.J., Samant, R.S., Verderame, M.F., Welch, D.R., 2000. Functional evidence for a novel human breast carcinoma metastasis suppressor, BRMS1, encoded at chromosome 11q13. Cancer Res. 60, 2764–2769.

Sheratt, J.A., Nowak, M.A., 1992. Oncogenes, anti-oncogenes and the immune response to cancer: a mathematical model. Proc. R. Soc. London B 248, 261–271.

Steeg, P.S., 2004. Metastasis suppressor genes. J. Natl. Cancer Inst. 96, E4.

Steeg, P.S., Bevilacqua, G., Kopper, L., Thorgerisson, U.P., Talmadge, J.E., Liotta, L.A., Sobel, M.E., 1988. Evidence for a novel gene associated with low tumor metastatic potential. J. Natl. Cancer Inst. 80, 200–204.

Strauss, B.S., 1998. Hypermutability in carcinogenesis. Genetics 148, 1619–1626.

Suzuma, T., Sakurai, T., Yoshimura, G., Umemura, T., Tamaki, T., Naito, Y., 2001. A mathematical model of axillary lymph node involvement considering lymph node size in patients with breast cancer. Breast Cancer 8, 206–212.

Taddei, F., Radman, M., Maynard-Smith, J., Toupance, B., Gouyon, P.H., Godelle, B., 1997. Role of mutator alleles in adaptive evolution. Nature 387, 700–702.

Thames, H.D., Buchholz, T.A., Smith, C.D., 1999. Frequency of first metastatic events in breast cancer: implications for sequencing of systemic and local-regional treatment. J. Clin. Oncol. 17, 2649–2658.

Tomlinson, I., Sasieni, P., Bodmer, W., 2002. How many mutations in a cancer? Am. J. Pathol. 160, 755–758.

Vogelstein, B., Kinzler, K.W., 2002. The genetic basis of human cancer. McGraw-Hill, New York.

Weiss, L., 2000. Metastasis of cancer: a conceptual history from antiquity to the 1990s. Cancer Metastasis Rev. 19, 193–383.

Wodarz, D., Komarova, N., 2005. Computational Biology Of Cancer: Lecture Notes And Mathematical Modeling. World Scientific Publishing Company, Singapore.

Wodarz, D., Krakauer, D.C., 2001. Genetic instability and the evolution of angiogenic tumor cell lines. Oncol. Rep. 8, 1195–1201.

Wodarz, D., Iwasa, Y., Komarova, N.L., 2004. On the emergence of multifocal cancers. J. Carcinogen 3, 13.

Wyllie, A.H., Rose, K.A., Morris, R.G., Steel, C.M., Foster, E., Spandidos, D.A., 1987. Rodent fibroblast tumors expressing human myc and ras genes: growth, metastasis and endogenous oncogene expression. Br. J. Cancer 56, 251–259.

Yamada, S., et al., 2002. Mitogen-activated protein kinase kinase 4 (MKK4) acts as a metastasis suppressor gene in human ovarian carcinoma. Cancer Res. 62, 6717–6723.

Yoshida, B.A., Dubauskas, Z., Chekmareva, M.A., Christiano, T.R., Stadler, W.M., Rinker-Schaeffer, C.W., 1999. Mitogen-activated protein kinase kinase 4/stress-activated protein/ERK kinase 1 (MKK4/SEK1), a prostate cancer metastasis suppressor gene encoded by human chromosome 17. Cancer Res. 59, 5483–5487.