Journal of Theoretical Biology 273 (2011) 207–215

Contents lists available at ScienceDirect

Journal of Theoretical Biology

journal homepage: www.elsevier.com/locate/yjtbi

Evolutionary dynamics of BRCA1 alterations in breast tumorigenesis

Laura De Vargas Roditi [a], Franziska Michor [b,*]

[a] Computational Biology Program, Memorial Sloan-Kettering Cancer Center, and Tri-Institutional Training Program in Computational Biology and Medicine, Weill Cornell Medical College, New York, NY 10065, USA

[b] Department of Biostatistics and Computational Biology, Dana-Farber Cancer Institute, and Department of Biostatistics, Harvard School of Public Health, Boston, MA 02115, USA

ARTICLE INFO

Article history:
Received 26 July 2010
Received in revised form
18 October 2010
Accepted 21 December 2010
Available online 29 December 2010

Keywords:
Stochastic modeling
Breast cancer
Moran model

ABSTRACT

Cancer results from the accumulation of alterations in oncogenes and tumor suppressor genes. Tumor suppressors are classically defined as genes which contribute to tumorigenesis if their function is lost. Genetic or epigenetic alterations inactivating such genes may arise during somatic cell divisions or alternatively may be inherited from a parent. One notable exception to this rule is the BRCA1 tumor suppressor that predisposes to hereditary breast cancer when lost. Genetic alterations of this gene are hardly ever observed in sporadic breast cancer, while individuals harboring a germline mutation readily accumulate a second alteration inactivating the remaining allele—a finding which represents a conundrum in cancer genetics. In this paper, we present a novel mathematical framework of sporadic and hereditary breast tumorigenesis. We study the dynamics of genetic alterations driving breast tumorigenesis and explore those scenarios which can explain the absence of somatic BRCA1 alterations while replicating all other disease statistics. Our results support the existence of a heterozygous phenotype of BRCA1 and suggest that the loss of one BRCA1 allele may suppress the fitness advantage caused by the inactivation of other tumor suppressor genes. This paper contributes to the mathematical investigation of breast tumorigenesis.

1. Introduction

Human breast carcinoma is classified into five distinct subtypes: basal-like, HER2+, and normal breast-like types as well as luminal A and B tumors (Potemski et al., 2005). Basal-like tumors comprise 10–20% of all breast cancers and typically present with aggressive phenotypes (Saal et al., 2008). These tumors are highly proliferative, poorly differentiated, genomically unstable, and rarely express any of the receptors most commonly targeted by breast cancer therapies: the estrogen, progesterone, and HER2 receptors (Vogelstein and Kinzler, 2002).

Hereditary breast cancer is characterized by a germline mutation in the susceptibility genes BRCA1 or BRCA2 and usually presents as basal-like tumors. Mutations in BRCA1 are found in a larger percentage of hereditary breast cancers than mutations in BRCA2 (Vogelstein and Kinzler, 2002; Eliot et al., 2003). A single inherited mutation in BRCA1 results in a 47–87% lifetime risk of developing breast cancer (Fackenthal and Olopade, 2007); BRCA1 penetrance is higher in individuals belonging to high-risk families, i.e., those with four or more breast cancer cases, as compared to individuals with no family history (Fackenthal and Olopade, 2007). BRCA1 acts as a tumor suppressor gene (Eliot et al., 2003; Welcsh and King, 2001; Smith et al., 1992; Friedman et al., 1994; Xu et al., 1999; Xu, 1999) and is involved in DNA repair and recombination, checkpoint control of the cell cycle, and transcription (Welcsh and King, 2001; Scully and Livingston, 2000; Hartman and Ford, 2002). Alterations of BRCA1 that impair such functions lead to increased proliferation and chromosomal instability (CIN) (Yasuo et al., 2002). CIN is characterized by an increased risk to loose or gain (parts of) chromosomes during cell divisions, causing elevated rates of loss of heterozygosity (LOH) (Lengauer et al., 1997). The latter is an important mechanism for inactivating tumor suppressor genes. Interestingly, as important as the BRCA1 gene is for the development of hereditary breast cancer, somatic mutations in BRCA1 are found very infrequently in sporadic tumors of the breast (Vogelstein and Kinzler, 2002; Eliot et al., 2003; Yasuo et al., 2002; Elledge and Amon, 2002; Sørlie et al., 1998). This finding contrasts with most tumor suppressor genes involved in familial cancer syndromes; somatic mutations in such genes are often seen in sporadic tumors of the same cancer type (e.g. RB1 in retinoblastoma, APC in Familial Adenomatous Polyposis (FAP), and WT1 in Wilms tumor) (Knudson, 1993). The absence of sporadic BRCA1 mutations despite the widespread loss of the second allele in germline mutation carriers represents a conundrum in cancer genetics (Vogelstein and Kinzler, 2002; Yasuo et al., 2002; Elledge and Amon, 2002).

In this paper, we investigate the mechanisms and dynamics of sporadic and hereditary breast carcinogenesis. We design a

* Correspondence to: Department of Biostatistics and Computational Biology, Dana-Farber Cancer Institute, Boston, MA 02115, USA.
Tel.: +1 617 632 5045.
E-mail address: michor@jimmy.harvard.edu (F. Michor).

0022-5193/$ - see front matter © 2011 Elsevier Ltd. All rights reserved.
doi:10.1016/j.jtbi.2010.12.033

mathematical model of the evolutionary dynamics of BRCA1 mutations arising during tumorigenesis and identify those system parameters that can explain the differential involvement of BRCA1 alterations in sporadic and hereditary breast cancer development. This model is part of an ongoing effort to investigate breast cancer with mathematical techniques (Speer et al., 1984; David, 2001; Berry et al., 1997; Ashkenazi et al., 2007; Clare et al., 2000).

2. The model

Consider a population of cells within the breast tissue proliferating according to a stochastic process known as Moran model (Moran, 1962). This population may consist of self-renewing multi-potent mammary stem cells if these cells are the only ones that live long enough to accumulate the necessary genetic alterations (Gabriela et al., 2003). Alternatively, the population may also contain progenitor cells if some of the (epi)genetic alterations leading to breast cancer confer properties of self-renewal to those cells. Define N to be the number of cells susceptible to such alterations. At each elementary step of the process, a cell is chosen to reproduce proportionally to its fitness (i.e., growth rate) and the daughter cell (possibly harboring a new mutation) replaces another cell randomly chosen for death. Thus, the population size remains strictly constant. A unit of time corresponds to N elementary steps of the stochastic process.

If the population size is smaller than the inverse of the mutation rate, then the population consists of at most two cell types at any time: cell types i and j. These two cell types differ from each other by only a single (epi)genetic alteration. The parameter r_i denotes the fitness value of cell type i and specifies the growth rate of this cell type relative to the wild type cells. The mutation rate from cell type i to j is given by u_{ij}. If there are A cells of type i and B cells of type j, where $A+B=N$, and cell type i can give rise to a cell of type j at rate u_{ij} per cell division, then the probability that a cell of type j is chosen for reproduction is given by $Br_j/[Ar_i+Br_j]$. Similarly, the probability that a cell of type i reproduces is given by $Ar_i/[Ar_i+Br_j]$. Therefore, the chance that a new cell of type j appears is given by the probability that cell type i is chosen for reproduction and mutates, and the probability that cell type j is chosen for reproduction, $u_{ij}Ar_i/[Ar_i+Br_j]+Br_j/[Ar_i+Br_j]$. The probability that a new cell of type i appears is given by the probability that a cell of type i reproduces and does not mutate, $(1-u_{ij})Ar_i/[Ar_i+Br_j]$. All cell types die at random, so the chance that a cell of type i dies is given by i/N, and the chance that a cell of type j dies is given by j/N. This stochastic process is a Markov process with time step $1/N$ and one absorbing state, $j=N$. The probability that one cell of type j reaches 100% frequency in a population of cells of type i, i.e., the probability that the process is absorbed in state $j=N$ starting from $j=1$, is given by $\rho(r_j/r_i)=[1-1/(r_j/r_i)]/[1-1/(r_j/r_i)^N]$ (Komarova et al., 2003); this quantity is called the fixation probability. The quantity $X_i(t)$ specifies the probability that the population consists of only type i cells at time t. Then, the rate of transition from a population of only i cells to only j cells is given by $Nu_{ij}\rho(r_j/r_i)$. If the (epi)genetic alteration does not confer a fitness change to cells, i.e., if cell types i and j have equal growth rates ($r_i=r_j$), then the rate of transition between these two states is given by the mutation rate, u_{ij}, since in that case, the probability of fixation is $1/N$. When the mutation rates are low, this stochastic process can be approximated by a system of differential equations. With an initial condition $X_i(0)=1$ and $X_j(0)=0$, the evolution of the system is then governed by $\dot{X}_i=-Nu_{ij}\rho(r_j/r_i)X_i$ and $\dot{X}_j=Nu_{ij}\rho(r_j/r_i)X_i$.

Let us now investigate the evolutionary trajectories that lead to sporadic and hereditary breast cancer. We consider two different genes that contribute to breast tumorigenesis: BRCA1 and another tumor suppressor gene, denoted by TSG (Fig. 1). Candidates for this second tumor suppressor include p53, PTEN, and Chk2 (Saal et al., 2008; Gasco et al., 2002), since mutations in those genes are found in both hereditary and sporadic breast cancer (Vogelstein and Kinzler, 2002). For the purpose of the mathematical model, we assume that the TSG is sufficient to initiate sporadic breast tumorigenesis if inactivated in both alleles. In sporadic cases, the population of stem/progenitor cells initially contains only wild type (i.e., unmutated) cells. Cells in this population may acquire mutations in BRCA1 and/or the TSG followed by additional mutations in other genes as

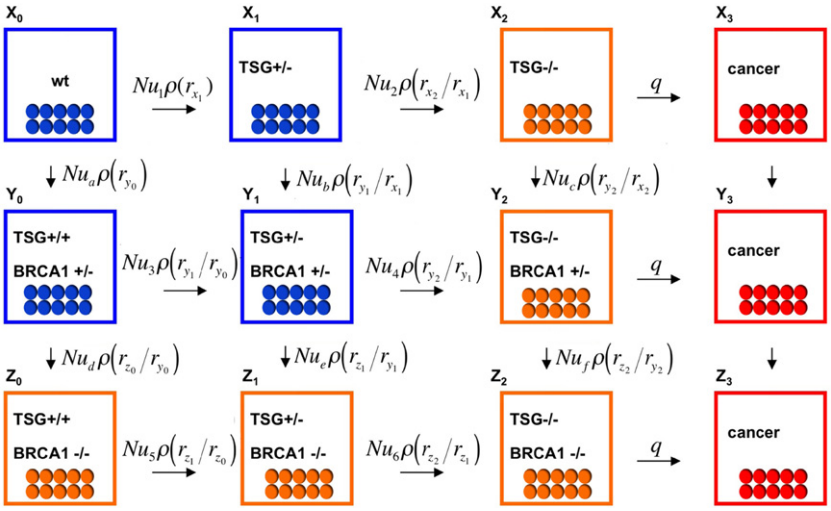

Fig. 1. Evolutionary trajectories leading to breast cancer. A stem/progenitor cell population consisting of N cells may acquire (epi)genetic alterations in two genes which contribute to breast carcinogenesis: BRCA1 and another tumor suppressor gene, TSG. The rates at which these mutations are acquired per allele per cell division are given by u_i. In sporadic breast tumorigenesis, all cells are initially wild type with respect to both genes, while in hereditary breast tumorigenesis, all cells initially harbor one mutated allele of BRCA1. As one of these initial populations accumulate mutations, the cells populate the other states in the mutational network. The relative fitness values of cell types as compared to the wild type cells are given by r_{x_0}, r_{x_1}, r_{x_2}, r_{y_0}, r_{y_1}, r_{y_2}, r_{z_0}, r_{z_1}, and r_{z_2}. The probability that a single cell of type x_2 reaches fixation in a population of $N-1$ cells of type x_1 is given by $\rho(r_{x2}/r_{x1})$. The probability that the population of cells consists only of cells with one particular genotype is given by X_0 for wild type cells, X_1 for TSG$^{+/-}$ cells, X_2 for TSG$^{-/-}$ cells, Y_0, Y_1, and Y_2 for BRCA1$^{+/-}$ cell types, and Z_0, Z_1, and Z_2 for BRCA1$^{-/-}$ cell types. Once a cell has acquired sufficiently many mutations to initiate tumorigenesis, it accumulates further mutations in oncogenes and/or other tumor suppressor genes to transform to invasive cancer (X_3, Y_3, and Z_3); these events occur at rate q. This mutational network is described by a system of differential equations (Table 1).

Table 1
The basic mathematical model. Rows and columns represent different cell types as indicated. Transition rates are given by $a_0=Nu_1\rho(r_{X1})$, $a_1=Nu_a\rho(r_{Y0})$, $a_2=Nu_2\rho(r_{X2}/r_{X1})$, $a_3=Nu_b\rho(r_{Y1}/r_{X1})$, $a_4=Nu_c\rho(r_{Y2}/r_{X2})$, $a_5=q$, $a_6=Nu_3\rho(r_{Y1}/r_{Y0})$, $a_7=Nu_d\rho(r_{Z0}/r_{Y0})$, $a_8=Nu_4\rho(r_{Y2}/r_{Y1})$, $a_9=Nu_e\rho(r_{Z1}/r_{Y1})$, $a_{10}=Nu_f\rho(r_{Z2}/r_{Y2})$, $a_{11}=q$, $a_{12}=Nu_5\rho(r_{Z1}/r_{Z0})$, $a_{13}=Nu_6\rho(r_{Z2}/r_{Z1})$, and $a_{14}=q$.

	$TSG^{+/+}$ cell types	$TSG^{+/-}$ cell types	$TSG^{-/-}$ cell types	$TSG^{-/-}$ fully malignant cells
$BRCA1^{+/+}$ cell types	$\dot{X}_0=-(a_0+a_1)X_0$	$\dot{X}_1=a_0X_0-(a_2+a_3)X_1$	$\dot{X}_2=a_2X_1-(a_4+a_5)X_2$	$\dot{X}_3=a_5X_2$
$BRCA1^{+/-}$ cell types	$\dot{Y}_0=a_1X_0-(a_6+a_7)Y_0$	$\dot{Y}_1=a_6Y_0+a_3X_1-(a_8+a_9)Y_1$	$\dot{Y}_2=a_8Y_1+a_4X_2-(a_{10}+a_{11})Y_2$	$\dot{Y}_3=a_{11}Y_2$
$BRCA1^{-/-}$ cell types	$\dot{Z}_0=a_7Y_0-a_{12}Z_0$	$\dot{Z}_1=a_{12}Z_0+a_9Y_1-a_{13}Z_1$	$\dot{Z}_2=a_{13}Z_1+a_{10}Y_2-a_{14}Z_2$	$\dot{Z}_3=a_{14}Z_2$

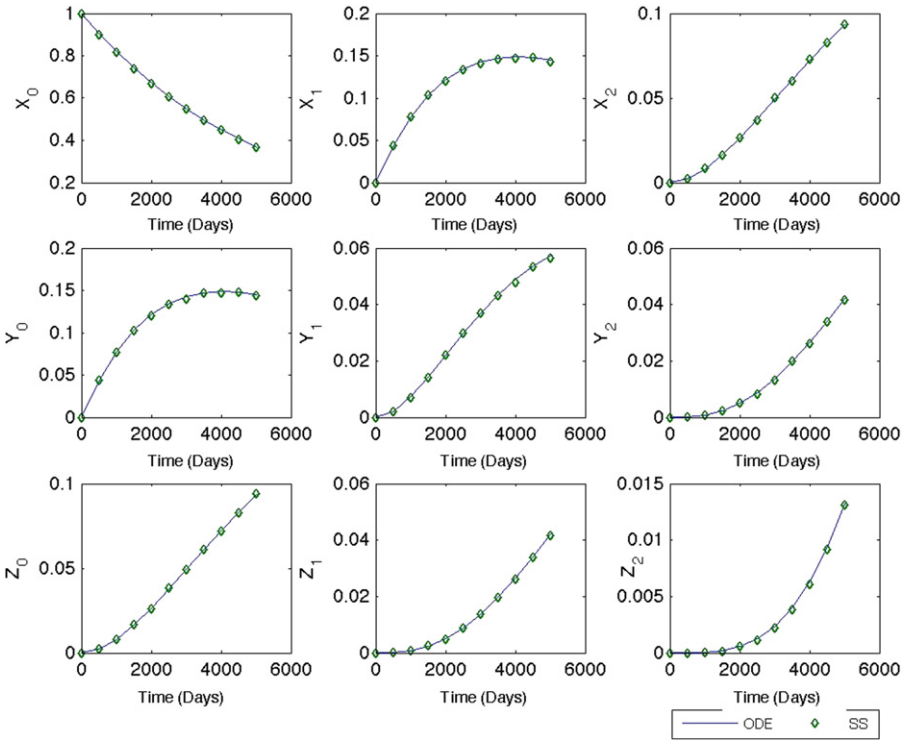

Fig. 2. A deterministic approximation of the stochastic process model. The panels show the fit between the exact computer simulation of the stochastic process (SS) and the predictions of the ordinary differential equation (ODE) system (Table 1) for different cell types. Variables X_0, X_1, X_2, Y_0, Y_1, Y_2, Z_0, Z_1, and Z_2 are denoted as in Fig. 1. Parameter values are: all mutation rates 10^{-4}, population size $N=10$, and fitness values $r_{x_0}=1$, $r_{x_1}=1$, $r_{x_2}=1.2$, $r_{y_0}=1$, $r_{y_1}=1$, $r_{y_2}=1.2$, $r_{z_0}=1.2$, $r_{z_1}=1.2$, and $r_{z_2}=1.2$.

they evolve towards breast carcinoma. The parameter q denotes the rate at which initiated cells, i.e., those that have accumulated mutations inactivating the TSG and possibly also BRCA1, transform into fully malignant cells. This transformation may include accumulation of further genetic and/or epigenetic alterations and other events. In inherited cases, the population initially consists only of cells that are wild type for the TSG, but heterozygous for a BRCA1 mutation. This population again accumulates further mutations driving breast tumorigenesis (Fig. 1). As outlined above, we approximate this stochastic process with a system of differential equations (Table 1); this approximation is accurate as long as the population size of cells at risk of accumulating mutations is small (Nowak et al., 2004) (Fig. 2). We use our mathematical model to study the evolution of the population of cells through this mutational network and identify scenarios that can replicate the clinically observed absence of BRCA1 mutations in sporadic breast cancer as well as the frequent inactivation of the second BRCA1 allele in hereditary cases.

3. Results

The lifetime risk of developing breast cancer for any female individual is about 12% (American Cancer Society, 2009), and 90–95% of breast cancers are sporadic (Fackenthal and Olopade, 2007). Therefore, an individual not harboring an inherited BRCA1 mutation has a 10–11% lifetime chance of developing breast cancer. If an individual harbors an inherited BRCA1 mutation, however, she has a 47–87% lifetime chance of being diagnosed with hereditary breast cancer (Fackenthal and Olopade, 2007).

We performed a comprehensive investigation of the system parameters in our model, selecting those that result in a 10% lifetime risk of developing sporadic breast cancer, while reproducing the scarcity of BRCA1 somatic mutations and the 47–87% penetrance of inherited BRCA1 mutations. To initialize the parameter search, we estimated the values of the mutation rates and fitness values from previously published experimental results. These parameter values for the initialization are shown in Table 2, and the reasoning for their choice is outlined in the following.

Without loss of generality, the fitness values of cells wild type with regard to both BRCA1 and the TSG are set to one. For the initialization, we assume that cells with one inactivated allele of either tumor suppressor have fitness one, i.e., are neutral as compared to wild type cells. Cells with two inactivated TSG alleles or two inactivated BRCA1 alleles are considered to have a 20% increased fitness as compared to wild type cells; therefore, their relative fitness value is 1.2. We consider BRCA1 and the TSG

Table 2
Initializing and new parameter values. Parameters r_i are defined as the fitness values of cell types i. The mutation rate from cell type i to j is given by u_{ij}. Parameters N and τ correspond to the population size and time between cell divisions, respectively. The parameter q denotes the rate at which a population of initiated cells develops into invasive breast cancer.

	r_{x_0}	r_{x_1}	r_{x_2}	r_{y_0}	r_{y_1}	r_{y_2}	r_{z_0}	r_{z_1}	r_{z_2}		N	q
Initializing parameters	1	1	1.2	1	1	1.2	1.2	1.2	1.44		10	10^{-6}
New parameters	1	1	1.35	1	1	0.5	1.5	2.5	3.5		10	10^{-6}
	u_1	u_2	u_3	u_4	u_5	u_6	u_a	u_b	u_c	u_d	u_e	u_f
Initializing parameters	10^{-7}	10^{-6}	10^{-7}	10^{-6}	10^{-7}	10^{-2}	10^{-7}	10^{-7}	10^{-7}	10^{-6}	10^{-6}	10^{-6}
New parameters	10^{-6}	10^{-6}	10^{-6}	10^{-6}	10^{-6}	10^{-2}	10^{-7}	10^{-7}	10^{-7}	10^{-6}	10^{-6}	10^{-6}

to have multiplicative fitness effects on cells, and hence cells with two inactivated alleles of both the TSG and BRCA1 are considered to have a relative fitness value of 1.44. These choices serve as starting point for the global parameter search.

So far, about 500 mutations in BRCA1 have been associated with breast cancer (Human Gene Mutation Database, http://www.hgmd.cf.ac.uk/ac/index.php). Since the baseline mutation rate per base per cell division is about 10^{-10} (Kunkel, 2004), the rate of inactivating the first of two BRCA1 alleles is approximately $2\times 500\times 10^{-10}=10^{-7}$. The second BRCA1 allele may be inactivated either by a second point mutation or by an LOH event; therefore, the rate at which the second allele is inactivated is given by the sum of the rates of point mutation and of LOH. No reliable measurements of the rate of LOH in genetically stable cells are available, but plausible values vary from 10^{-7} to 10^{-5} per allele per cell division (Lengauer et al., 1997; Lengauer et al., 1998). The rate of LOH is likely of the same order of magnitude as the point mutation rate; otherwise, two distinct point mutations in the two alleles would never be observed. Hence, we consider a rate of inactivating the second allele of a tumor suppressor of about 10^{-6} per cell division. Additionally, some tumor suppressor genes (such as BRCA1) cause chromosomal instability when lost, increasing the rate of LOH to about 10^{-2} per cell division (Lengauer et al., 1997). Again, these values are utilized for the initialization of the parameter search, but varied later on to identify parameter regimes replicating the breast cancer statistics.

The number of cells at risk of accumulating the mutations driving breast tumorigenesis depends on the types of cells considered capable of becoming malignant. If only mammary stem cells live sufficiently long to accumulate the number of alterations necessary to initiate breast tumorigenesis, then the population at risk consists of approximately 10 cells per niche. This quantity represents the number of stem cells present in small ducts scattered in the breast close to the terminal lobules, amounting to approximately 100,000 cells in total (Gabriela et al., 2003). Alternatively, mammary progenitor cells may be considered in addition to stem cells. In the context of our model, these cells are considered to divide on average once per day; each elementary time step of the Moran model corresponds to a cell division event. The model is then used to determine the probability at time t that a cell with a certain genotype has reached fixation in a population of stem/progenitor cells. Since this quantity does not directly correspond to the statistics available for the lifetime risk of developing sporadic and hereditary breast cancer, we introduced a term q into our model that represents the rate at which initiated cells transform into fully malignant cancer cells. Since this transformation involves an alteration of at least one more gene (Radford et al., 1995), we consider q to range from 10^{-5} to 10^{-8} per cell division. Parameter regimes that were found for $q=10^{-5}$, however, only included scenarios in which mutations in TSG and BRCA1 were disadvantageous to the cell, and were therefore discarded. Regimes found for $q=10^{-7}$ and lower values contained relative fitness values of mutated cells of 20–50; since such a large fitness advantage is unlikely, we used $q=10^{-6}$ for our analyses.

We investigated the behavior of our mathematical model over time and found that using the initial parameter values outlined above, we were not able to replicate the disease statistics seen in the clinic—the magnitude of sporadic breast cancer incidence, the rareness of BRCA1 somatic mutations, and the high probability of developing breast cancer in germline mutation carriers (Fig. 3). Therefore, we performed a comprehensive investigation of the system parameters that revealed mechanisms explaining the differential involvement of BRCA1 in sporadic and hereditary breast tumorigenesis.

Since in hereditary cases, all cells within the body harbor a BRCA1 mutation, the three mutational states with wild type BRCA1 have zero probability of occurring. Hence, the system of differential equations in this case can be reduced to describe six mutational states—cells wild type with respect to the TSG and mutated in one or both TSG alleles, either with one or two inactivated BRCA1 alleles (see Fig. 1). This reduction of the system allowed us to perform the parameter investigation in a space of fewer dimensions. In particular, once the fitness values of BRCA1 wild type cells (with or without mutations inactivating the TSG) were optimized for the case of sporadic breast cancer, these values did not need to be altered to describe hereditary cases. Therefore, we performed our initial parameter search for sporadic cases with just two parameters—the fitness values of $TSG^{+/-}BRCA1^{+/+}$ and $TSG^{-/-}BRCA1^{+/+}$ cells, r_{x_1} and r_{x_2}. Recall that without loss of generality, the relative fitness of cells wild type with respect to both BRCA1 and the TSG, r_{x_0}, is one. When simulating our differential equation system describing sporadic cases for different values of r_{x_1}, we found that the 10% lifetime incidence of sporadic breast cancer could be met by a range of fitness values of $TSG^{+/-}BRCA1^{+/+}$ and $TSG^{-/-}BRCA1^{+/+}$ cells (Fig. 4a). We then turned to an initial parameter search for hereditary cases and found that the mutation rates u_1, u_3, and u_5 had to be set to 10^{-6} to recapitulate the statistics for BRCA1 penetrance (Fig. 5); we have $u_1=u_3=u_5$ since all three rates correspond to the mutation rate for inactivating the first TSG allele (albeit on different mutational backgrounds). This alteration of the mutation rates is necessary, but not sufficient for replicating the disease statistics; additionally, certain fitness values also need to be optimized (see later sections). With this estimate of the mutation rates, we performed another search for sporadic cases and identified fitness values of $r_{x_1}=1$ and $r_{x_2}=1.35$ that, together with the mutation rate estimate, were able to replicate the sporadic breast cancer incidence (Fig. 4b).

In summary, our initial parameter search determined that a 35% increase in fitness of $TSG^{-/-}BRCA1^{+/+}$ cells, and a mutation rate of 10^{-6} per cell division to inactivate the first TSG allele,

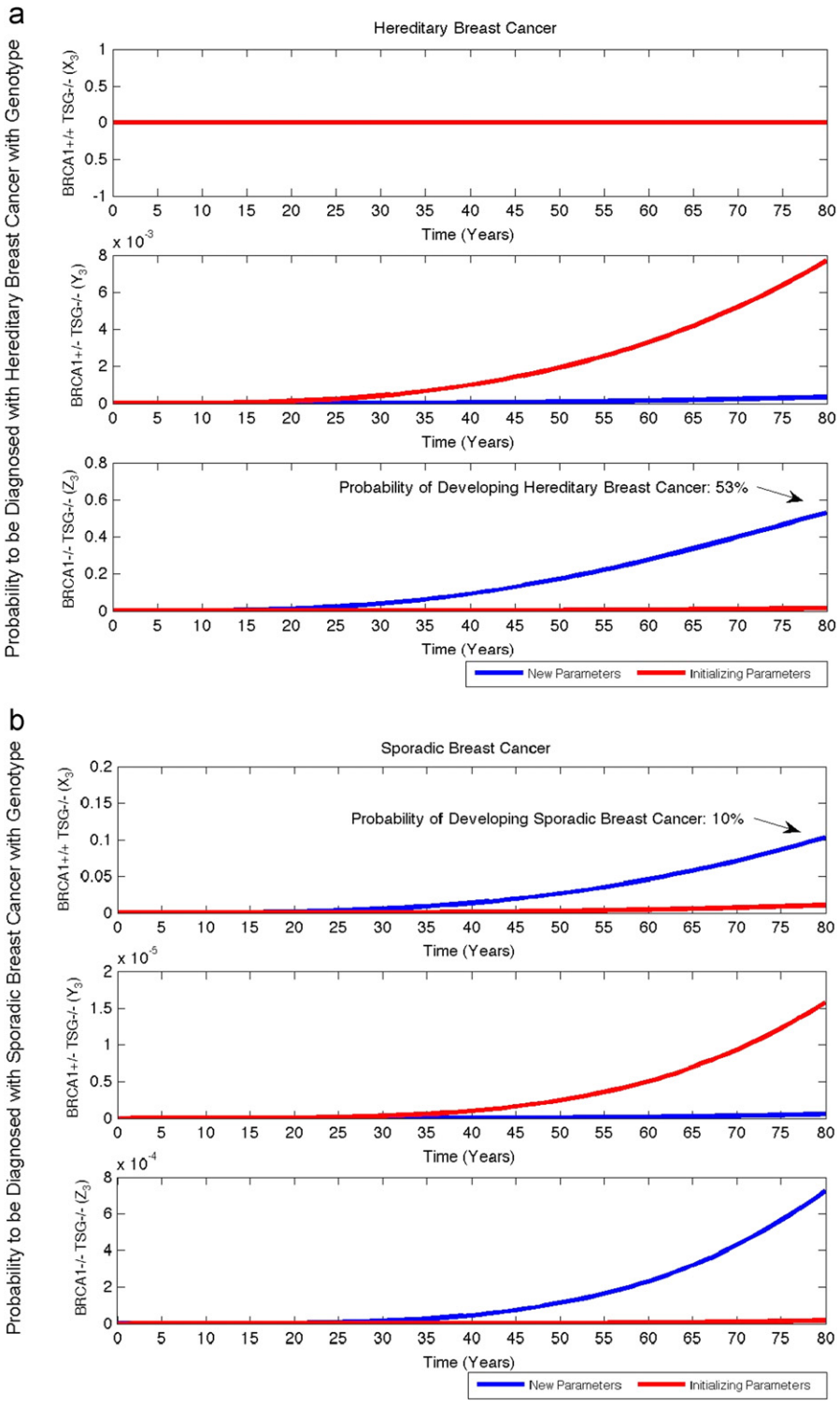

Fig. 3. The probability of breast cancer initiation. The panels display the probability of cancer-initiating cells arising from different genotypes as a function of time. (a) The simulation of the differential equation model is shown for the hereditary initial condition for both the set of parameters used to initialize the parameter search and a new set of parameters found to better replicate breast cancer statistics (see Table 2). The probability of developing hereditary breast cancer (i.e., with genotype $TSG^{-/-}BRCA1^{-/-}$) is approximately 53% for an 80-year old patient. (b) The simulation of the differential equation model is shown for the sporadic initial condition for both the set of parameters used to initialize the parameter search and a new set of parameters found to better replicate breast cancer statistics. Both parameter sets are shown in Table 2. The probability of developing sporadic breast cancer (i.e., with genotype $TSG^{-/-}$ $BRCA1^{+/+}$) is approximately 10% for an 80-year old patient. The results in panels a and b are comparable to the clinically observed values of 47–87% probability for hereditary breast cancer and 10% probability for sporadic breast cancer.

replicated the incidence of sporadic breast cancer and the rareness of somatic BRCA1 mutations. A magnitude of the mutation rate of 10^{-6} could stem from the fact that the TSG represents a mutational hotspot or alternatively that there is a pool of such genes, each of which could trigger the cascade of mutational events leading to breast cancer. For example, if inactivation of any of ten different TSGs (each occurring at a rate of 10^{-7}) is sufficient to initiate sporadic breast cancer, then a mutation inactivating an

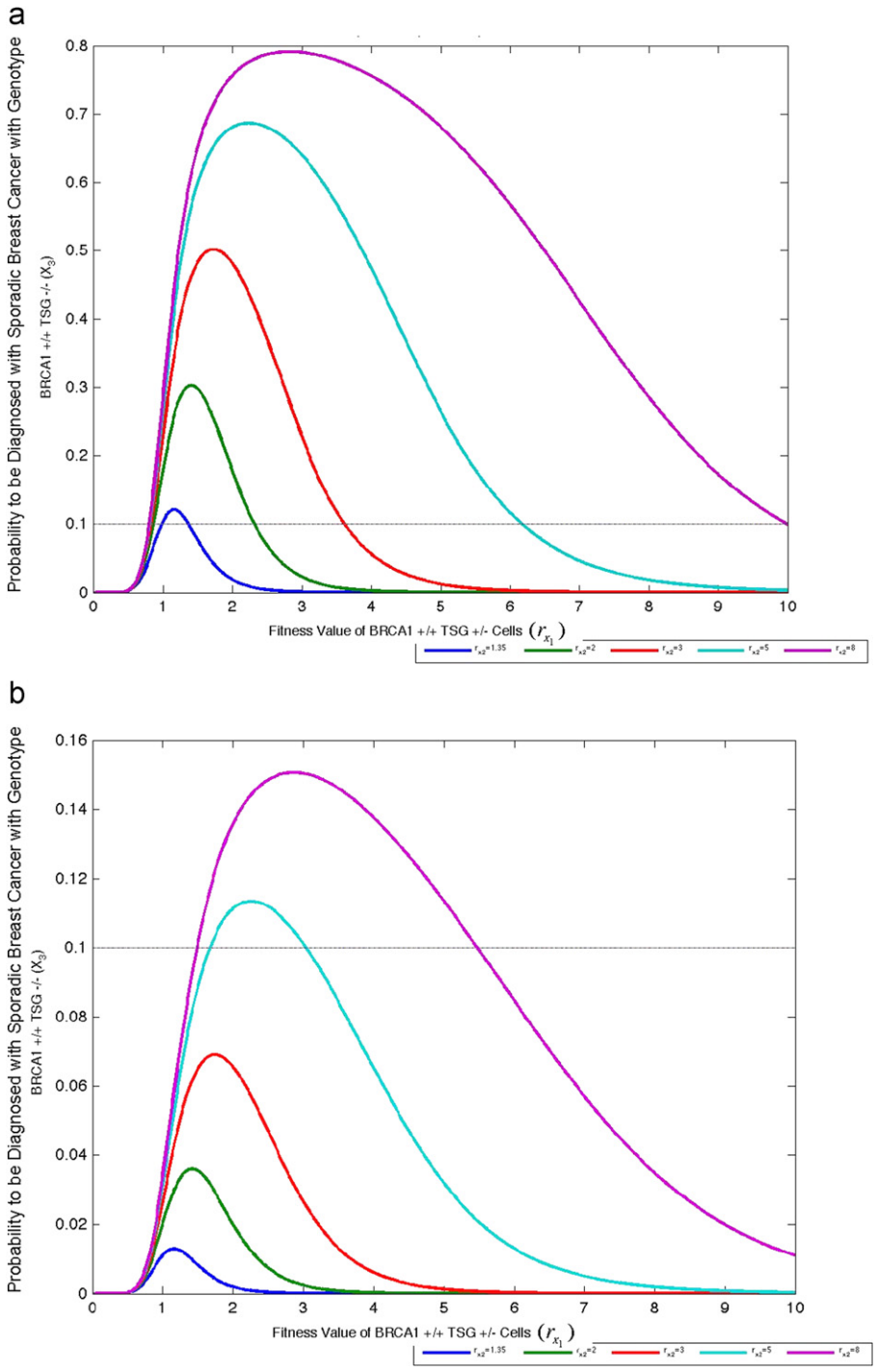

Fig. 4. The effects of cellular fitness on sporadic breast tumorigenesis. The panels show the probability of sporadic breast cancer arising with a $TSG^{-/-}$ $BRCA1^{+/+}$ genotype, X_3, on the vertical axis for different fitness values for $TSG^{+/-}BRCA1^{+/+}$ cells, r_{x_1}, on the horizontal axis. Different colored curves represent different fitness values of $TSG^{-/-}BRCA1^{+/+}$ cells, r_{x_2}. Parameter values are the same as the initializing parameters (Table 2), except for r_{x_1}, r_{x_2}, u_1, u_3, and u_5. (a) The value of mutation rate for the loss of the first TSG allele ($u_1=u_3=u_5$) is 10^{-6}. (b) The value of mutation rate for the loss of the first TSG allele is $u_1=u_3=u_5=10^{-7}$.

allele of any of those genes arises at rate $10\times10^{-7}=10^{-6}$. Although these optimized parameters recapitulate the incidence of sporadic breast cancer and the rareness of somatic BRCA1 mutations, they could not replicate the high penetrance of BRCA1 germline mutations. We proceeded by searching the entire parameter space of fitness values and found that the system, together with our assumptions about reasonable ranges of parameter values, could not replicate all disease statistics under the assumption of multiplicative fitness effects. We therefore analyzed the fitness values individually without the assumption of multiplicativity.

We first searched through the space of the following four parameters: the fitness values of $TSG^{-/-}BRCA1^{+/-}$ cells (r_{y_2}), $TSG^{+/+}BRCA1^{-/-}$ cells (r_{z_0}), $TSG^{+/-}BRCA1^{-/-}$ cells (r_{z_1}), and $TSG^{-/-}BRCA1^{-/-}$ cells (r_{z_2}), for different assumptions of the fitness value of $TSG^{+/-}$ $BRCA1^{+/-}$ cells, r_{y_1}. We considered values of r_{y_1} that account for the three scenarios in which $r_{y_1}<r_{y_0}$, $r_{y_1}=r_{y_0}$, and $r_{y_1}>r_{y_0}$, respectively. The parameter r_{y_2} was varied in

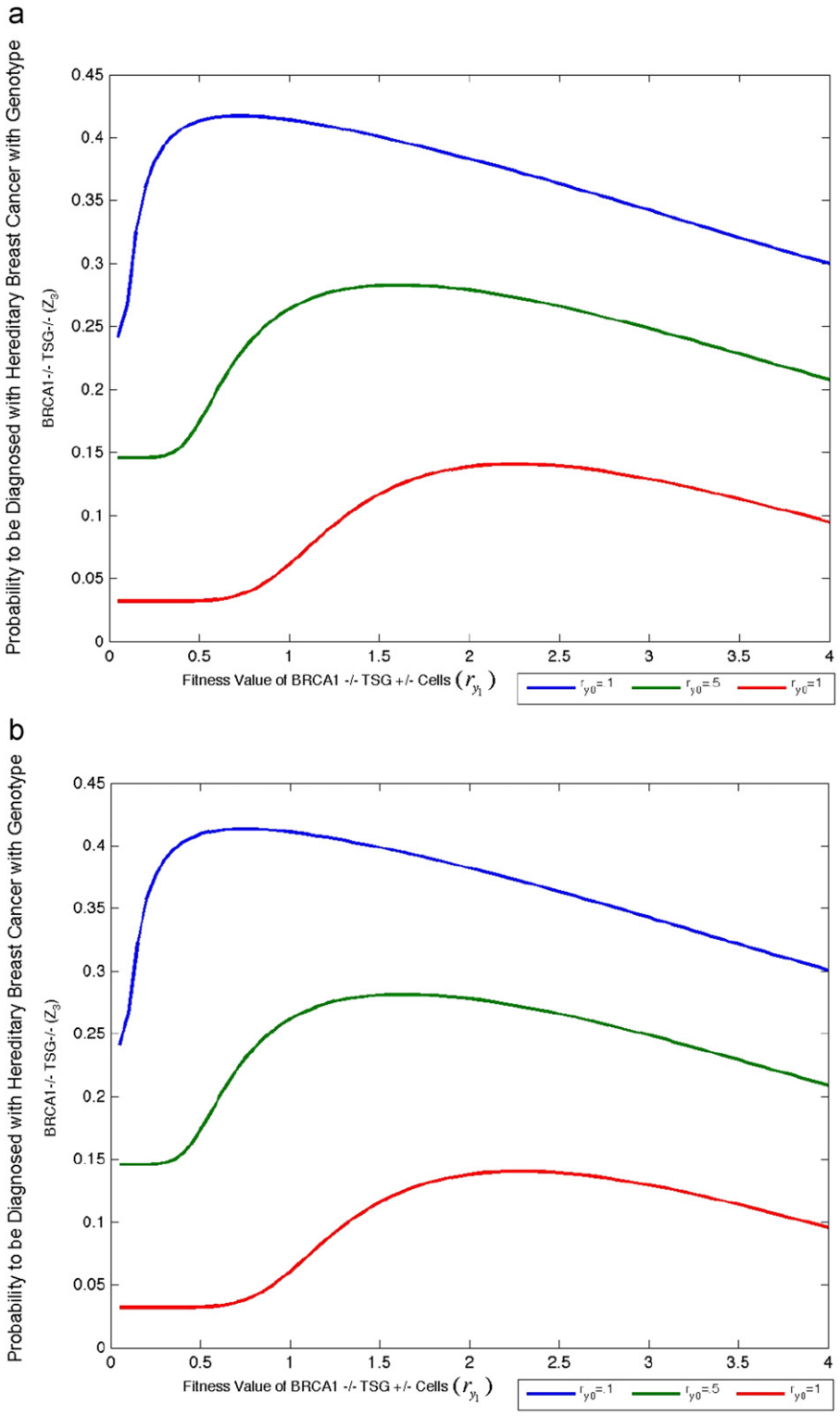

Fig. 5. The effects of cellular fitness on hereditary breast tumorigenesis. The panels show the probability of developing hereditary breast cancer (i.e., for a germline mutation carrier) with a BRCA1$^{-/-}$ TSG$^{-/-}$ genotype, Z_3, after 80 years for different fitness values of BRCA1$^{+/-}$ TSG$^{+/-}$ cells (r_{y_1}). The different colored curves correspond to different fitness values for BRCA1$^{+/-}$ TSG$^{+/+}$ cells (r_{y_0}). Parameter values are $q=10^{-6}$, $N=10$, $\tau=1$, $r_{x_0}=1$, $r_{x_1}=2$, $r_{x_2}=5$, $r_{z_0}=1$, $r_{z_1}=5$, and $r_{z_2}=10$. The fitness value of BRCA1$^{+/-}$ TSG$^{-/-}$ cells (r_{y_2}) is 0.5 in (a) and 5 in (b).

the range 0.2–2, and the values r_{z_0}, r_{z_1}, and r_{z_2} were varied from 1 to 5; we then solved our ordinary differential equation model over time searching for the parameter combinations that resulted in the correct disease statistics. We found that there were no results to our search when $r_{y_2} > r_{y_1}$, irrespective of whether $r_{y_1} < r_{y_0}$, $r_{y_1} = r_{y_0}$, or $r_{y_1} > r_{y_0}$. In particular, this finding implies that there are no solutions for $r_{y_2} = r_{x_2} = 1.35$, which was expected under the assumption that neither BRCA1 nor the TSG has a heterozygous phenotype. This observation suggests the existence of a heterozygous phenotype for BRCA1 that only manifests itself after the complete inactivation of another tumor suppressor gene. It also indicates that this heterozygous phenotype is

disadvantageous to the cell, since all fitness values found for $TSG^{-/-}BRCA1^{+/-}$ cells (r_{y_2}) were below 0.9. Furthermore, this parameter search over r_{y_2}, r_{z_0}, r_{z_1}, and r_{z_2} demonstrated that the disease statistics could only be replicated for values of $r_{z_1} \geq r_{z_0}$.

The results of the global parameter search were projected from the 4-dimensional surface onto three different planes: the plane of fitness values of $TSG^{+/+}BRCA1^{-/-}$ and $TSG^{+/-}BRCA1^{-/-}$ cells (r_{z_0}, r_{z_1}) (Fig. 6a), the plane of fitness values of $TSG^{+/+}BRCA1^{-/-}$ and $TSG^{-/-}BRCA1^{-/-}$ cells (r_{z_0}, r_{z_2}) (Fig. 6b), and the plane of fitness values of $TSG^{+/-}BRCA1^{-/-}$ and $TSG^{-/-}BRCA1^{-/-}$ cells (r_{z_1}, r_{z_2}) (Fig. 6c). The coordinates of each point in the plane correspond to the fitness values of the two cell types for which the solution of the model fits all constraints. The absence of solutions in the r_{z_0}/r_{z_1} plane below the diagonal line $r_{z_0}=r_{z_1}$ is noted in Fig. 6a. Therefore, for all values of r_{y_2}, r_{z_0}, r_{z_1}, and r_{z_2} that represent a solution, there are no values for which, where r_{z_1} is less than r_{z_0}. This observation indicates that there must be a fitness increase of $TSG^{+/-}BRCA1^{-/-}$ cells relative to $TSG^{+/+}BRCA1^{-/-}$ cells.

The results of the parameter search revealed two conditions that are required for the predictions of the model to replicate both the sporadic and hereditary breast cancer statistics. The first finding was that to meet the disease statistics, alteration of one BRCA1 allele in combination with the homozygous loss of the TSG must have a deleterious effect on the cell as compared to alterations solely in the two TSG alleles. This finding may imply that hemizygous loss of the BRCA1 allele suppresses the fitness advantage conferred by the homozygous loss of the other TSG. The results show, however, that this disadvantageous phenotype can be rescued by acquiring the second BRCA1 mutation before an inactivation of the first TSG allele. This observation was supported by the second finding suggesting that cells with two mutated BRCA1 alleles and hemizygous loss of the TSG alleles have a fitness advantage as compared to $BRCA1^{-/-}TSG^{+/+}$ cells.

4. Discussion

In this paper, we have designed and analyzed a novel mathematical framework to investigate the dynamics of sporadic and hereditary breast tumorigenesis. We considered two genes contributing to breast cancer development: BRCA1, a tumor suppressor associated with hereditary breast cancer, and another tumor suppressor gene (TSG), which is assumed to be sufficient to initiate sporadic breast tumorigenesis if lost. Candidates for the TSG include p53, PTEN, and Chk2 (Saal et al., 2008; Gasco et al., 2002). Using this mathematical framework, we studied the scenarios necessary to replicate the statistics associated with breast cancer – the absence of somatic alterations in BRCA1 in sporadic cases as well as the high penetrance of BRCA1 mutations observed in hereditary cases. We performed a comprehensive investigation of the parameter space and identified regimes that recapitulate these clinical findings. To meet the disease statistics, $BRCA1^{+/-}TSG^{-/-}$ cells must have a fitness disadvantage as compared to $BRCA1^{+/-}TSG^{+/-}$ cells; furthermore, $BRCA1^{-/-}TSG^{+/-}$ cells must have a fitness advantage relative to $BRCA1^{-/-}TSG^{+/+}$ cells. Finally, we found that the rate at which the first TSG is inactivated is of the order of 10^{-6} per gene per cell division, which may point to the fact that the TSG is a mutational

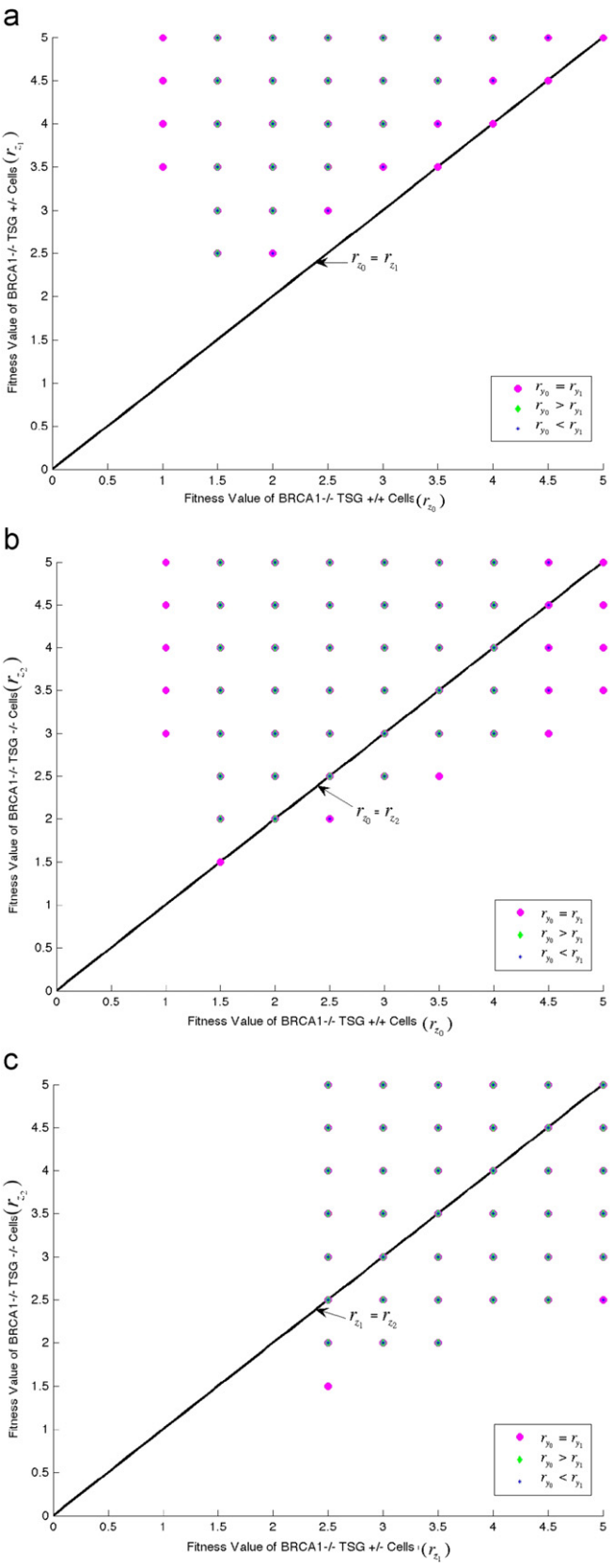

Fig. 6. A global parameter search of cellular fitness values. The plot represents the projection of the 4-dimensional surface resulting from the parameter search over the fitness values of $TSG^{-/-}BRCA1^{+/-}$ cells (r_{y_2}), $TSG^{+/+}BRCA1^{-/-}$ cells (r_{z_0}), $TSG^{+/-}BRCA1^{-/-}$ cells (r_{z_1}) and $TSG^{-/-}BRCA1^{-/-}$ (r_{z_2}) cells on the r_{z_0}/r_{z_1} plane (a), the r_{z_0}/r_{z_2} plane (b), and the r_{z_1}/r_{z_2} plane (c). The coordinates of each point on each plot correspond to the two fitness values that satisfy the parameter search both for sporadic and hereditary initial conditions. The 4-dimensional search was done for three different scenarios, where one TSG mutation is neutral ($r_{y_0}=r_{y_1}$), disadvantageous ($r_{y_1} < r_{y_0}$), and advantageous ($r_{y_1} > r_{y_0}$) with respect to a TSG wild type cell. The three different scenarios are represented by different markers in the plot.

hotspot or that there is a pool of alternative tumor suppressor genes, each of which can initiate breast tumorigenesis if inactivated. Taken together, our results point to the possible existence of a heterozygous phenotype of BRCA1, indicating that the loss of one BRCA1 allele may affect tumorigenesis after another tumor suppressor gene has been inactivated by suppressing some of the phenotypic advantages caused by the inactivation of the TSG. This finding is consistent with experimental evidence (Jeng et al., 2007; Kote-Jarai et al., 2005; Meric-Bernstam, 2007) and contributes to the question about the cellular effects of BRCA1 inactivation. We have, however, not incorporated the possibility that BRCA1 could be inactivated in sporadic breast tumors in ways that have not yet been discovered; inclusion of such effects into our model may alter the dynamics of the system and the results we found.

Our mathematical framework represents only one possibility of modeling the processes of breast tumorigenesis. We considered the breast tissue to be subdivided into multiple small niches of mammary stem cells, which proliferate to regenerate their population as well as to produce the differentiated progeny maintaining breast functions. We modeled these stem cell niches as independent cell populations. For simplicity, the population sizes were considered to be constant, even though the mathematical framework can also be applied to situations in which clonal expansion to a new carrying capacity occurs after the emergence of all or some genetic alterations. Similarly, we have not included many complexities contributing to breast tumorigenesis, such as interactions between cancer cells and the immune system, the effects of spatial organization as well as diffusion of oxygen, nutrients, growth factors, and other molecules which may shape the processes of angiogenesis, growth, and invasion. These factors, although important, are neglected in the current model to be able to focus on the most important aspects of breast tumorigenesis and to identify some contributors to the disparate findings about the involvement of BRCA1 in sporadic and hereditary breast cancer. Our mathematical framework contributes to the investigation of breast cancer dynamics and points to a heterozygous phenotype of BRCA1 as a driver of differential involvement of this gene in sporadic and hereditary tumorigenesis.

Acknowledgement

The authors would like to thank the Michor laboratory for helpful comments and discussions. This work is supported by NCI Grant U54CA143798.

References

American Cancer Society 2009. Cancer Facts and Figures 2009, vol. 2009. American Cancer Society, Atlanta.

Ashkenazi, R., Jackson, T., Dontu, G., Wicha, M., 2007. Breast cancer stem cells-research opportunities utilizing mathematical modeling. Stem Cell Rev. Rep. 3, 176–182.

Berry, D.A., Parmigiani, G., Sanchez, J., Schildkraut, J., Winer, E., 1997. Probability of carrying a mutation of breast-ovarian cancer gene BRCA1 based on family history. J. Natl. Cancer Inst. 89, 227–238.

Clare, S.E., Nakhlis, F., Panetta, J.C., 2000. Molecular biology of breast metastasis: the use of mathematical models to determine relapse and to predict response to chemotherapy in breast cancer. Breast Cancer Res. 2, 430–435.

David, M.E., 2001. Understanding mathematical models for breast cancer risk assessment and counseling. Breast J. 7, 224–232.

Eliot, M.R., Saijun, F., Richard, G.P., Itzhak, D.G., 2003. BRCA1 gene in breast cancer. J. Cell. Physiol. 196, 19–41.

Elledge, S.J., Amon, A., 2002. The BRCA1 suppressor hypothesis: an explanation for the tissue-specific tumor development in BRCA1 patients. Cancer Cell 1, 129–132.

Fackenthal, J.D., Olopade, O.I., 2007. Breast cancer risk associated with BRCA1 and BRCA2 in diverse populations. Nat. Rev. Cancer 7, 937–948.

Friedman, L.S., Ostermeyer, E.A., Szabo, C.I., Dowd, P., Lynch, E.D., Rowell, S.E., King, M.-C., 1994. Confirmation of BRCA1 by analysis of germline mutations linked to breast and ovarian cancer in ten families. Nat. Genet. 8, 399–404.

Gabriela, D., Muhammad, A.-H., Wissam, M.A., Michael, F.C., Max, S.W., 2003. Stem cells in normal breast development and breast cancer. Cell Proliferation 36, 59–72.

Gasco, M., Shami, S., Crook, T., 2002. The p53 pathway in breast cancer. Breast Cancer Res. 4, 70–76.

Hartman, A.-R., Ford, J.M., 2002. BRCA1 induces DNA damage recognition factors and enhances nucleotide excision repair. Nat. Genet. 32, 180–184.

Jeng, Y.M., Cai-Ng, S., Li, A., Furuta, S., Chew, H., Chen, P.L., Lee, E.Y.H., Lee, W.H., 2007. BRCA1 heterozygous mice have shortened life span and are prone to ovarian tumorigenesis with haploinsufficiency upon ionizing irradiation. Oncogene 26, 6160–6166.

Knudson, A.G., 1993. Antioncogenes and human cancer. Proc. Natl. Acad. Sci. United States of America 90, 10914–10921.

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.

Kunkel, T.A., 2004. DNA replication fidelity. J. Biol. Chem. 279, 16895–16898.

Kote-Jarai, Z., Salmon, A., Mengitsu, T., Copeland, M., Ardern-Jones, A., Locke, I., Shanley, S., Summersgill, B., Lu, Y.J., Shipley, J., et al., 2005. Increased level of chromosomal damage after irradiation of lymphocytes from BRCA1 mutation carriers. Br. J. Cancer 94, 308–310.

Lengauer, C., Kinzler, K.W., Vogelstein, B., 1997. Genetic instability in colorectal cancers. Nature 386, 623–627.

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

Moran, P.A.P., 1962. The Statistical Processes of Evolutionary Theory. Patrick A.P. Moran, Oxford.

Meric-Bernstam, F., 2007. Heterogenic loss of BRCA in breast cancer: the "Two-Hit" hypothesis takes a hit. Ann. Surg. Oncol. 14, 2428–2429.

Nowak, M.A., Michor, F., Komarova, N.L., Iwasa, Y., 2004. Evolutionary dynamics of tumor suppressor gene inactivation. Proc. Natl. Acad. Sci. United States of America 101, 10635–10638.

Potemski, P., Kusinska, R., Watala, C., Pluciennik, E., Bednarek, A.K., Kordek, R., 2005. Prognostic relevance of basal cytokeratin expression in operable breast cancer. Oncology 69, 478–485.

Radford, D.M., Fair, K.L., Phillips, N.J., Ritter, J.H., Steinbrueck, T., Holt, M.S., Donis-Keller, H., 1995. Allelotyping of ductal carcinoma in situ of the breast: deletion of loci on 8p, 13q, 16q, 17p and 17q. Cancer Res. 55, 3399–3405.

Saal, L.H., Gruvberger-Saal, S.K., Persson, C., Lovgren, K., Jumppanen, M., Staaf, J., Jonsson, G., Pires, M.M., Maurer, M., Holm, K., et al., 2008. Recurrent gross mutations of the PTEN tumor suppressor gene in breast cancers with deficient DSB repair. Nat. Genet. 40, 102–107.

Smith, S.A., Easton, D.F., Evans, D.G.R., Ponder, B.A.J., 1992. Allele losses in the region 17q12-21 in familial breast and ovarian cancer involve the wild-type chromosome. Nat. Genet. 2, 128–131.

Scully, R., Livingston, D.M., 2000. In search of the tumour-suppressor functions of BRCA1 and BRCA2. Nature 408, 429–432.

Sørlie, T., Andersen, T.I., Bukholm, I., Børresen-Dale, A.-L., 1998. Mutation screening of BRCA1 using PTT and LOH analysis at 17q21 in breast carcinomas from familial and non-familial cases. Breast Cancer Res. Treat. 48, 259–264.

Speer, J.F., Petrosky, V.E., Retsky, M.W., Wardwell, R.H., 1984. A stochastic numerical model of breast cancer growth that simulates clinical data. Cancer Res. 44, 4124–4130.

Vogelstein, B., Kinzler, K.W., 2002. The Genetic Basis of Human Cancer, second ed. Mcgraw-Hill, New York.

Welcsh, P.L., King, M.-C., 2001. BRCA1 and BRCA2 and the genetics of breast and ovarian cancer. Hum. Mol. Genet. 10, 705–713.

Xu, X., Wagner, K.-U., Larson, D., Weaver, Z., Li, C., Ried, T., Hennighausen, L., Wynshaw-Boris, A., Deng, C.-X., 1999. Conditional mutation of BRCA1 in mammary epithelial cells results in blunted ductal morphogenesis and tumour formation. Nat. Genet. 22, 37–43.

Xu, X., 1999. Conditional mutation of BRCA1 in mammary epithelial cells results in blunted ductal morphogenesis and tumour formation. Nat. Genet. 22, 7.

Yasuo, M., Kyoko, I., Noriko, I., Chiyomi, E., Shinzaburo, N., 2002. Acceleration of chromosomal instability of *BRCA1*-associated hereditary breast cancers by p53 abnormality. Breast J. 8, 77–80.