Breast Cancer Research and Treatment (2005) 91: 121–124 © Springer 2005

Report

Estimating the number of rate limiting genomic changes for human breast cancer

Xinan Zhang and Richard Simon
Biometric Research Branch, National Cancer Institute, Bethesda, MD, USA

Key words: breast mutations, genomic instability, human breast carcinogenesis, mutator phenotype

Summary

We used multistage models that incorporate the age dependent dynamics of normal breast tissue, clonal expansion of intermediate cells and mutational events to fit data for the age-specific incidence of breast cancers in the surveillance, epidemiology, and end results (SEER) registry. Our results suggest that two or three rate limiting events occurring at rates characteristic of point mutation rates for normal mammalian cells set in motion a sequence of other genomic changes that lead with high probability to breast carcinoma.

Introduction

Cancer of epithelial tissues is generally thought to develop slowly over many years [1]. Hanahan and Weinberg [2] suggest that the vast catalog of cancer cell genotypes is a manifestation of six essential alterations in cell physiology that collectively dictate malignant growth: self-sufficiency in growth signals, insensitivity to growth-inhibitory signals, evasion of programmed cell death (apoptosis), limitless replicative potential, sustained angiogesis, and tissue invasion and metastasis. The analysis of age-incidence curves for human cancers using quantitative models has also resulted in claims that 6–8 independent rate-limiting 'hits' are needed [3].

The mutation rate in somatic human cells has been estimated at approximately 2.0×10^{-7} mutations per gene per cell division [4,5]. Loeb and his colleagues have proposed that cancer cells must exhibit a mutator phenotype in order to account for the large number of mutations found in tumors [5–8]. Tomlinson and Bodmer et al. [9–11] have challenged the mutator phenotype hypothesis. They have questioned the biological accuracy of the multi-stage models referred to by Loeb, and claimed that selection is more important than an increased mutation rate in the development of a tumor. The relationship between mutation and selection in tumors has been the subject of general debate. Our goal is to clarify and help resolve these issues in the context of analyzing the incidence of breast cancers in the SEER database and to thereby obtain insights into the mechanisms of breast carcinogenesis.

Materials and methods

Incidence data for breast cancers were obtained from the SEER registry for the year 1973–1999 [12]. For our analyses, we use the reported incidence of breast cancer by gender, race, age, and calendar year in the nine SEER geographic areas, which together represent an estimated 10% of U.S. population. The population bases were from SEER population files (bases on data from U.S. Census Bureau) by sex and race and were cross-tabulated by calendar year (1973–1999) and 5-year age groups (ages 0–85+). Our analyses addressed combined races for breast cancer in females during the period 1990–1999. Rates are expressed as cases per 100,000 females.

We used quantitative models of two-six stages, according to the findings of Hanahan and Weinberg [2]. The models we used are extensions of two-stage model of Knudson and Moolgavkar et al. [13,14] which permit clonal expansion of intermediate cells that have accumulated some but not all of the mutations needed for full tumorgenicity. The original two-stage model of Armitage and Doll [15] did not permit clonal expansion of intermediate cells and some of the multi-stage models subsequently developed also ignore this important feature [3]. Tomlinson and Bodmer are correct in claiming that such models are seriously inadequate for estimating the number of rate-limiting events needed to explain human age-incidence curves.

Figure 1 depicts a three-hit model schematically. The left-most compartment represents the population of normal tumor progenitor cells. Moolgavkar et al. [14] assumed that the tumor progenitor cells are breast epithelium stem cells and the number of such stem cells increases according to a logistic growth curve from an initial value of 10 cells at birth to a maximum of 10^7 cells by age 20. The existence of breast stem cells and the identity of breast cancer progenitor cells are not unambiguously established, however, and so we have considered various possible levels for the maximum number of progenitor cells. The number of progenitor cells decreases after age 45 at a rate estimated from the data.

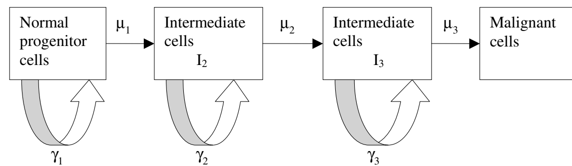

Figure 1. A 3-hit model for carcinogenesis. I_2, I_3 denote the compartments of intermediate cells with 1 and 2 mutations, respectively. μ_1, μ_2, μ_3 are the mutation rates per cell per year of normal progenitor cells, I_2 compartment and I_3 compartment; γ_1, γ_2, γ_3 are the net growth rates per cell per year for the three compartments.

The quantity μ_1 represents the rate of the first event expressed per progenitor cell per year. A first event gives rise to an intermediate cell (I_2). The I_2 compartment expands at the rate γ_2 exponentially. An intermediate cell in I_2 is transformed to an intermediate cell in I_3 by the occurrence of a second event that occurs at rate μ_2. The intermediate cells in I_3 expand at rate γ_3 and are subject to a final transformation event occurring at rate μ_3 per cell per year. For a three hit model, the third event leads to a fully malignant cell. The mathematical specification of the model is presented in the Supplementary materials.

We fit 2–6 hit models to the breast cancer incidence in the SEER registry (1973–1999). The parameters of these models were optimized to fit the data. The growth rate for each stage was permitted to have one value for age less than 45 years and then take on a different value, because of the possible hormonal influence on intermediate cells. The growth rates parameters for age less than 45 years were restricted to be non-negative.

Results

Figure 2 shows the SEER age-specific breast cancer incidence and the fit of 2 and 3 hit models. The curves for the 4–6 hit models are indistinguishable from those for the 3 hit models and are omitted. The 3–6 stage models fit the data very well. The fit for the 2-stage model is not quite as good.

Table 1 shows the fitted values of the mutation rates per cell per year for each model. In order to express the mutation rates per cell division, the values in Table 1 should be divided by the number of cell divisions per year. Direct measurements of the number of divisions per year for these cell populations are not available, however. If the proginator cells were constantly proliferating with cell cycle duration of 48 h, then 180 rounds of DNA replication per stem cell would be possible. Since stem cells are thought to be non-proliferating most of the time, however, the number of divisions per year is probably much less than this upper limit [16].

For the 2-hit model the event rates in Table 1 are 5.4×10^{-8} and 1.1×10^{-5} events per cell per year. If there are 10 cell divisions per year these rates become 5.4×10^{-9} and 1.1×10^{-6} events per cell division. The geometric mean of these two rates is 7.7×10^{-8}. This is somewhat smaller than the point mutation rate per cell division per gene for normal mammalian cells. If there were more cell divisions per year per stem cell, the event rates per cell division become even smaller. If, however, the number of normal proginator cells were only 10^4 instead of 10^7 as assumed in Table 1, then with 10 cell divisions per year per proginator cell, the geometric

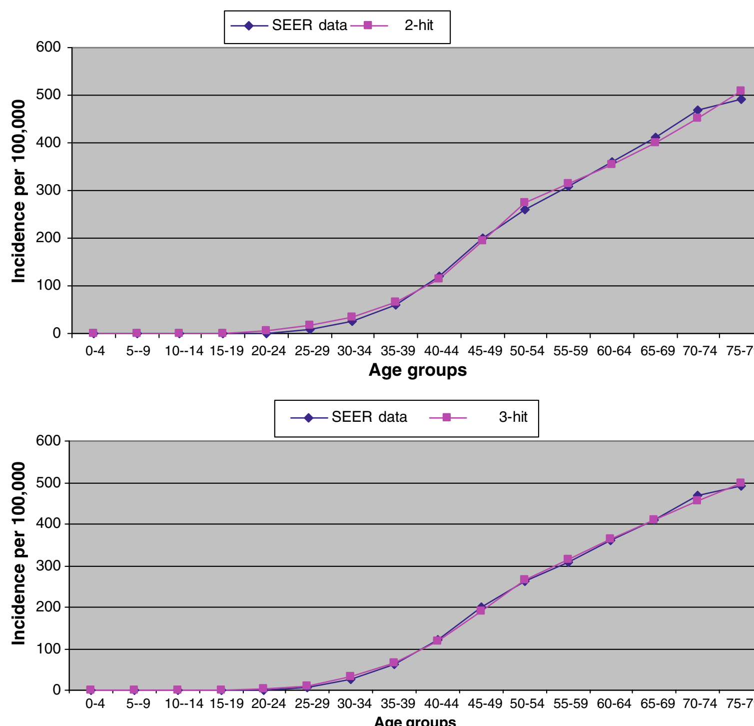

Figure 2. The age-specific incidence data of all races per 100,000 females for breast cancers from SEER registry for the year 1973–1999 (black) and rate predicted by 2-hit and 3-hit models (purple). Results for 4–6 hit models are not distinguishable from those for 3-hit model.

Table 1. Estimated optimal values of mutation rates per cell per year for each cell compartment based on normal proginator compartment of 10^7 cells

Cell Compartment	2-Hit	3-Hit	4-Hit	5-Hit	6-Hit
Normal	5.4×10^{-8}	3.6×10^{-5}	0.92×10^{-6}	3.9×10^{-6}	5.3×10^{-6}
I_2	1.1×10^{-5}	4.4×10^{-5}	7.4×10^{-6}	1.4×10^{-4}	8.9×10^{-6}
I_3		8.2×10^{-5}	8.9×10^{-2}	8.3×10^{-4}	1.0×10^{-2}
I_4			6.7×10^{-2}	4.1×10^{-1}	2.3×10^{-1}
I_5				1.6×10^{-1}	4.0×10^{-1}
I_6					7.3×10^{-1}

mean average mutation rate per cell division for the two hit model becomes 2.1×10^{-6} instead of 7.7×10^{-8} and the two hit model fits the data equally well (See Supplementary material). The range 2.1×10^{-6} to 7.7×10^{-8} bounds the normal point mutation rate of 2×10^{-7} and so a model of two hits occurring at normal point mutation events is consistent with the data.

For the 3-hit model the event rates shown in Table 1 are 3.6×10^{-5}, 4.4×10^{-5}, and 8.2×10^{-5} per cell per year respectively for the three events. With 10 cell divisions per year, these rates correspond to an average mutation rate of 5.1×10^{-6} events per cell division. With 100 cell divisions per year, the rates correspond to an average mutation rate of 5.1×10^{-7} events per cell division. The 3-hit model provides an excellent fit to the data. Hence, the data is consistent with a model of 3 mutational events occurring at normal point mutation rates.

Table 1 indicates that the best fitting 4-hit, 5-hit and 6-hit models have only 2 or 3 rate limiting events which occur at rates similar to the normal point mutation rate. The events beyond the third occur at very high rates.

Tumor progression during the silent interval

In fitting the multi-hit models, we have assumed there is an interval of 5-years between the emergence of a neoplastic cell and clinical detection of a tumor [17,18]. During this 'silent interval,' the tumor acquires numerous additional genomic changes. If tumor detection occurs when there are about 10^9 tumor cells and all tumor cells stay clonogenic, the silent interval consists of 30 generations of tumor growth and 1.07×10^9 cell divisions (see Supplementary material). If only 55% of tumor cells stay clonogenic, then the silent interval consists of 218 generations of tumor growth and 1.06×10^{10} cell divisions.

The normal rate of point mutation is approximately 2×10^{-7} per gene per cell division or approximately 10^{-9} per base pair. Consequently, even at the normal mutation rate, the number of total tumor cell divisions is sufficiently large that almost every nucleotide will have a substantial probability of being mutated in at least one cell of the tumor at the time of diagnosis. Many genomic changes, however, will occur late in the process and thus will not be represented in many tumor cells. The developing tumor will consist of a quasi-species distribution of a vast number of mutant types [19]. This quasi-species distribution provides the basis for selection of phenotypes proficient in angiogenesis, invasion, metastasis, and resistance to therapeutics.

Discussion

Age specific cancer incidence data has previously been used to infer the number of independent rate-limiting events in tumorigenesis. Unfortunately, many of the mathematical models that have been used have not accounted for the age dependent dynamics of the susceptible tissue target cell population or for clonal expansion of intermediate cells and this is a serious limitation. We have utilized a model that accounts for these features for breast carcinoma. This model accounts for the increased incidence of breast cancer associated with early menarche or delayed menopause. Early menarche leads to early breast development and thus to an extended period with substantial numbers of susceptible progenitor cells. Because menopause may reduce the number of intermediate cells as well as number of progenitor cells, late menopause can result in an increase in the incidence of breast cancer.

Previous studies fitting models to the age-specific breast cancer incidence rate have not critically examined the mutation rates associated with the best fitting models. Using biologically plausible models, we have found the age-specific incidence data is best explained by a models requiring two or three hits occurring at normal point mutation rates followed possibly by additional events occurring at higher rates.

One biological scenario consistent with our findings has the initial mutational events occurring at normal point mutation rates causing disregulation of cell cycle gatekeeper genes leading to a dramatically increased number of cell divisions per year for the intermediate cells. Although the intermediate cells are proliferating, the fitted models indicate that size of the intermediate populations is not rapidly increasing as the net growth rates are approximately zero (see Supplementary Material). This may be the result of disregulation of cell cycle control in the presence of intact apoptotic path-

ways for regulating the number of cells. Clonal expansion may take place only after the occurrence of subsequent events. With the number of cell divisions per year dramatically increased for intermediate cells, however, subsequent rate limiting events occur at a greatly elevated mutation rate expressed in terms of mutations per intermediate cell per year.

The initial events occurring at normal mutation rates may have effects other than increasing the number of cell divisions per year. The initial events may disrupt caretaker genes leading to genomic instability and thereby increased rates of subsequent events at either the nucleotide or chromosomal level [20]. Struewing et al. [21] and Armstrong et al. [22] reported that patients carrying a BRCA1 mutation have a 60–85% lifetime risk of breast carcinoma. This high penetrance suggests that loss of function of the BRCA1 protein may serve to increase the effective mutation rate for other rate-limiting events. The BRCA1 protein has been connected to DNA repair [23].

It is important to distinguish genomic changes required for tumorigenesis from changes that occur in proliferating malignant tumors that facilitate tumor progression. Most tumors will contain a small clone of cells that contain any single genetic alteration that one can imagine. Such clones can thus be selected for expansion based on environmental conditions in the tissue giving rise to the tumor. It is not possible to infer the number of genomic changes necessary to achieve an invasive metastatic tumor. Our analyses here suggest, however, that two or three genomic events occurring at approximately normal point mutation rates set the stage for a sequence of other events occurring either at higher rates per cell division or in the context of disregulated cell division that leads to carcinoma of the breast.

References

1. Loeb LA, Loeb KR, Anderson JP: Multiple mutations and cancer. Proc Natl Acad Sci USA 100: 776–781, 2003
2. Hanahan D, Weinberg RA: The hallmarks of cancer. Cell 2003; 100: 57–70
3. Renan MJ: How many mutations are required for tumorigenesis? Impliocations from cancer data. Mol Carcinogenesis 7: 139–146, 1993
4. Demars AL, Held KR: The spontaneous azaguanine-resistent mutations of diploid human fibroblasts. Humangenetik 16: 87–110, 1972
5. Loeb LA: A mutator phenotype in cancer. Cancer Research 61: 3230–3239, 2001
6. Loeb LA: Errors in DNA replication as a basis of malignant change. Cancer Research 34: 2311–2321, 1974
7. Loeb LA: Mutator phenotype may be required for multistage carsinogenesis. Cancer Research 51: 3075–3079, 1991
8. Loeb LA: Microsatellite instability: marker of a mutator phenotype in cancer. Cancer Research 54: 5059–5062, 1994
9. Tomlinson IP, Sasieni P, Bodmer W: How many mutations in a cancer. Am J Pathol 160: 755–758, 2002
10. Tomlinson IP, Bodmer W: Selection, the mutation rate and cancer: Ensuring that the tail does not wag the dog. Nat Med 5: 11–12, 1999
11. Tomlinson IP, Novelli MR, Bodmer W: The mutation rate and cancer. Proc Natl Acad Sci USA 93: 14800–14803, 1996
12. Ries LAG, Eisner MP, Kosary CL, Hankey BF, Miller BA, Clegg L, Edwards BK, (eds): SEER Cancer Statistics Review, 1973-1999. National Cancer Institute, Bethesda, MD, 2002
13. Knudson AG: Mutation and cancer: Statistical study of retinoblastoma. Proc Natl Acad Sci USA 68: 820–823, 1971
14. Moolgavkar SH, Day NE, Stevens RG: Two-stage model for carcinogenesis: Epidemiology of breast cancer in females. J Natl Cancer Inst 65: 559–569, 1980
15. Armitage P, Doll R: The age distribution of cancer and a multistage theory of carcinogenesis. Br J Cancer 8: 1–12, 1954
16. Steel GG, editor. Growth Kinetics of Tumours . Clarendon Press: Oxford; 1977
17. Norton L, Simon R: Growth curve of an experimental solid tumor following radiotherapy. J Natl Cancer Inst 58: 1735–1741, 1977
18. Radmacher MD, Simon R: Estimation of Tamoxifen's efficiency for preventing the formation and growth of breast tumors. J Natl Cancer Inst 92: 48–53, 2000
19. Eigen M, McCaskill J, Schuster P: The molecular quasispecies. Advances in Chem. Phys. 75: 149–263, 1989
20. Lengauer C, Kinzler KM, Vogelstein B: Genetic instability in colorectal cancers. Nature 386: 623–627, 1997
21. Struewing JP, Hartge P, Wacholder S, et al. The risk of cancer associated with specific mutation of BRCA1 and BRCA2 among Ashkenazi Jews. N Engl J Med 336: 1401–1408, 1997
22. Armstrong K, Eisen A, Weber B: Assessing the risk of breast cancer. N Engl J Med 342: 564–571, 2000
23. Gowen LC, Avrutskaya AV, Latour AM, et al. BRCA1 required for transcription-coupled repair of oxidative DNA damage. Science 281: 1009–1012, 1998

Address for offprints and correspondence: Richard Simon, National Cancer Institute, 9000 Rockville Pike, Bethesda, MD 20892-7434, USA; *Tel.:* +301-496-0975; *Fax:* +301-402-0560; *E-mail:* rsimon@nih.gov