Current Biology, Vol. 14, 242–246, February 3, 2004, ©2004 Elsevier Science Ltd. All rights reserved. DOI 10.1016/j.cub.2003.12.026

Age-Specific Acceleration of Cancer

Steven A. Frank*
Department of Ecology and Evolutionary Biology
University of California
Irvine, California 92697

Summary

One of the great challenges of cancer research is to explain the epidemiological patterns of cancer incidence based on the molecular processes that lead to uncontrolled cellular proliferation. The epidemiological data demonstrate that the age-specific incidence of many cancers increases in an approximately linear way with age when plotted on a log-log scale, with different slopes for different cancers [1]. However, those epidemiological data also show that cancers of various tissues depart from log-log linearity in particular ways [2]. Here, I illustrate those departures from log-log linearity by introducing plots of the age-specific acceleration of cancer. I then develop a very general model of cancer progression, which I use to explain the observed differences between tissues in age-specific acceleration. In one application of the model, I show that the spectacular rise and fall in age-specific acceleration observed in prostate cancer may be explained by multiple rounds of clonal expansion. In a second application, I demonstrate that the steady decline in age-specific acceleration of breast cancer may occur because precancerous mutations accumulate in many cellular lineages.

Results and Discussion

The standard log-log plots of age-specific incidence in Figure 1A show the increase in cancer rates with age for four adult-onset epithelial cancers. We get a better view of the differences between tissues by replotting the data. Figures 1B and 1C show the same data, plotting the slope in Figure 1A at each age. The slope of cancer rate (incidence) at each age measures the age-specific acceleration of cancer. For example, breast cancer accelerates most quickly at early ages, and the acceleration declines steadily throughout life. By contrast, acceleration in prostate cancer rises rapidly in early life, peaking at a very high level near 40 years, and then the acceleration drops rapidly through later life at the same time as the number of cases increases greatly. These striking epidemiological differences between tissues remain unexplained in terms of molecular and cellular processes.

I developed a general set of equations to link mutational processes at the cellular level to the epidemiological patterns. The model follows the commonly accepted assumption that cancer arises through the sequential accumulation of mutations within cell lineages [3]. The basic model for mutation accumulation is

$$\dot{x}_0(t) = -u_0(t)\,x_0(t)$$

$$\dot{x}_i(t) = u_{i-1}(t)\,x_{i-1}(t) - u_i(t)\,x_i(t) \quad i = 1,\ldots,n-1$$

$$\dot{x}_n(t) = u_{n-1}(t)\,x_{n-1}(t),$$

where $x_j(t)$ is the number of cell lineages with j mutations at time t, $u_j(t)$ is the rate at which lineages move from having j mutations to having $j + 1$ mutations, and the dot is the derivative with respect to time. I assumed that once an individual had a single lineage with n mutations, that person had cancer. Thus, $x_n(t)$ is the cumulative probability that an individual develops cancer between ages 0 and t.

The simplest models assumed the u_j values do not change with time [1, 4]. Then $x_n(t) \approx Ut^n$, where U depends on the various constant transition (mutation) rates u_j. The first derivative is the age-specific rate (incidence), $\dot{x}_n \approx Unt^{n-1}$, and so $\log(\dot{x}_n) \approx \log(Un) + (n-1)\log(t)$. Thus, with constant mutation rates, the log-log plot of age-specific incidence is approximately linear with time and has a slope of $n - 1$. The derivative of $\log(\dot{x}_n)$ with respect to $\log(t)$ is $n - 1$. This log-log second derivative is used in Figure 1 as a measure of acceleration. With the simple model of constant mutation, the standard approximation yielded a constant log-log acceleration of $n - 1$ that is independent of time.

The goal here is to understand the observed departures from constant log-log acceleration. Two qualitative patterns in the epidemiological data appear in Figure 1. First, a decline in acceleration occurs throughout life for breast cancer and at later ages for colorectal, lung, and prostate cancers. Second, acceleration rises until 40–50 years in colorectal, lung, and prostate cancers, with a particularly strong increase for the prostate.

I used the model to show two new aspects of how molecular and cellular processes affect epidemiological patterns. The first new aspect was the decline in acceleration at later ages. One possible explanation that has not been discussed concerns how the frequency distribution of lineages with different numbers of mutations changes with time. Let $p_j(t) = x_j(t)/X$ be the frequency of lineages with j mutations at time t, where X is the total number of lineages summed over all x_j. Suppose at birth that all cellular lineages have zero mutations, $p_0(0) = 1$. Then all lineages are n steps from cancer, and the log-log acceleration is $n - 1$. As time passes, it could be that many lineages accumulate a mutations, so that the typical number of steps to cancer is $n - a$ and the log-log acceleration is $n - 1 - a$. As a rises with time, the acceleration declines. We can study this process by solving the model above, using the simplest case of constant mutation rates.

The plots in Figure 2 illustrate the decline in log-log acceleration. The initial number of lineages per individual, $x_0(0)$, strongly affects the nature of the decline. When this value is high, then only very rare lineages become

*Correspondence: safrank@uci.edu

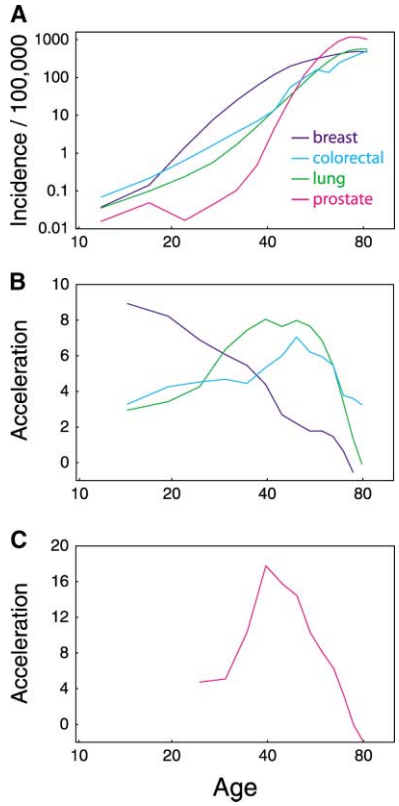

Figure 1. Age-Specific Incidence and Acceleration of Cancer for Different Tissues

(A) Age-specific incidence for four adult-onset epithelial cancers. (B and C) Age-specific log-log acceleration, which are slopes of the plots in (A) at different ages. All data are from the SEER database (www.seer.cancer.gov) using the nine SEER registries, year of diagnosis 1992–2000. Breast cancer data for all females, colorectal cancer for all males, lung cancer for all males, and prostate cancer for white males.

transformed, and the majority of lineages remain in or near the class with zero mutations. When the number of lineages is smaller, then a greater proportion of lineages must progress toward cancer to give the same overall cancer rate. The number of lineages may be relatively lower in some tissues either because there are fewer initial stem lineages or because a precancerous lineage expands and kills off neighboring lineages, lowering the effective number of lineages and decreasing the average number of steps per lineage that remain before cancer. Note that the blue line in Figure 2A shows roughly the same sort of decline as the observed pattern in breast cancer in Figure 1B.

The second new issue concerns an extended model of clonal expansion. Previous models have contributed much to understanding departures from linearity in the log-log incidence plots [5–8]. Most of those models assumed only a single round of clonal expansion. Fisher [9] proposed a model with two rounds of clonal expansion, but under his special assumptions, the log-log plot of incidence versus age remains linear.

In precancerous lineages, a mutation may lead to clonal expansion, followed by another mutation in one cell of the clone. The new mutant may then give rise to another round of clonal expansion. Several rounds of expansion, mutation, and new clonal growth seem possible, given that cancer ultimately arises as uncontrolled clonal growth following the acquisition of several mutations.

To study clonal expansion, it was useful to rewrite the model above as

$$x_0(t) = x_0(0)\, D_0(t,0)$$

$$x_i(t) = \int_0^t u_{i-1}(s)\, x_{i-1}(s)\, D_i(t,s)\, ds \quad i = 1,\ldots,n-1$$

$$x_n(t) = \int_0^t u_{n-1}(s)\, x_{n-1}(s)\, ds,$$

under the assumption that at $t = 0$, all cells started in the class with zero mutations, so that we could study the changes in each class by measuring the influx and outflux of cells into that class. For example, $u_{i-1}(s)\, x_{i-1}(s)$ is the influx at time s, and

$$D_i(t,s) = e^{-\int_s^t u_i(y)dy}$$

is the decay or outflux of that component over the remaining period from s to t. In this case, the $u_i(t)$ values varied with time, in part because transition rates between classes depended on clonal expansion. I summed over all influx and outflux by integrating from 0 to t.

Let clonal expansion follow the logistic model, which has the well-known solution that starting with a single cell and after a time period of length t, the number of cells in the clone is

$$y_i(t) = \frac{K_i e^{r_i t}}{K_i + e^{r_i t} - 1},$$

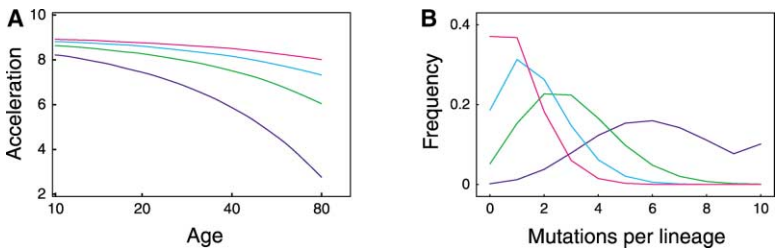

Figure 2. Decline in Age-Specific Acceleration with Age Calculated from the Model

(A) Fewer long-lived cell lineages in a tissue cause a greater decline with age. The numbers of initial lineages, $x_0(0)$, from top to bottom are 10^6 (red), 10^4 (cyan), 10^2 (green), and 10^0 (blue). There are $n = 10$ mutational steps to cancer. The transition (mutation) rate per year was adjusted so that the total incidence of cancer over all ages up to 80 years is 10%, requiring transition rates for the lines from top to bottom of $u_i = 0.0124$, 0.021, 0.037, 0.078 for all i.

(B) Age-specific decline in acceleration is caused by a rise over time in the number of mutations per lineage, leaving fewer steps to cancer. The plot shows the frequency distribution for the number of mutations per cell lineage at age 80.

where K is the carrying capacity and r is the intrinsic rate of increase of the clone [10]. The subscripts describe the class according to the number of mutations so that clones with different numbers of mutations may have different carrying capacities and rates of increase.

Let the total mutational capacity of a cell lineage be the mutation rate per cell, v, multiplied by the clone size, Y. Then we have

$$D_i(t,s) = e^{-\int_s^t v_i Y_i(\alpha)d\alpha} = \left(\frac{K_i}{K_i + e^{r_i(t-s)} - 1}\right)^{v_i K_i / r_i}.$$

The total transition rate from class i to class $i + 1$ is the mutation rate per cell, v_i, multiplied by the average clone size for members of the class, $\overline{Y}_i(t)$. Thus,

$$u_i(t) = v_i \overline{Y}_i(t) =$$
$$v_i \int_0^t u_{i-1}(s)\, x_{i-1}(s)\, D_i(t,s)\, Y_i(t-s)\, ds \,/\, x_i(t).$$

With expressions for u_i and D_i, we can use the model to study multiple rounds of clonal expansion. This model is general enough to fit many different shapes of log-log acceleration. However, the goal here is not to fit but to emphasize that a few general processes can explain the differences between tissues in their log-log acceleration curves.

Most prior models studied a single round of clonal expansion. The new model developed here allows multiple rounds of clonal expansion. To illustrate the consequences of multiple clonal expansions and to compare with earlier studies, it is useful to begin with a single round of expansion in the $n - 1$ class [8].

Figure 3A illustrates the effect of changing the rate of clonal expansion, r, in the single round of clonal expansion. Slower clonal expansion causes the acceleration in cancer to happen more slowly and to be spread over more years, because slow clonal expansion causes a slow increase in the rate at which a lineage acquires the final mutation that leads to cancer. Figure 3B shows that an increase in maximum clone size raises the peak level of acceleration until the clone becomes large enough that a mutation likely occurs in a short time interval, after which further clonal expansion does not increase the rate of progress through the mutational steps.

Figure 3C makes the key point that multiple rounds of clonal expansion can greatly increase the peak acceleration of cancer. The curves from bottom to top have one, two, or three rounds of clonal expansion. With three rounds of clonal expansion, the red curve looks roughly like the spectacular increase and decline in acceleration observed in prostate cancer (Figure 1C).

The acceleration plots discussed in this paper focus attention on the different epidemiological patterns that occur in epithelial cancers. Based on these observed patterns, I developed two hypotheses about cancer progression in individuals that link mutational and cellular processes to population-wide epidemiology. The match between the theoretical curves from my models and the observed epidemiological patterns establish these ideas as plausible hypotheses that deserve further study.

There are, of course, many other processes within individuals that affect cancer progression. Those processes may be alternatives that could link individual progression to epidemiological pattern. At present, no one has formulated a plausible model to link alternative processes of progression to the observed variation in the age-specific acceleration of different epithelial cancers—indeed, one benefit of the acceleration plots is

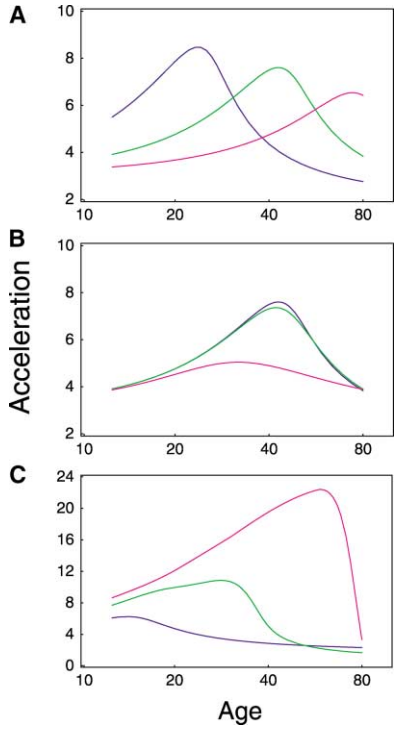

Figure 3. Patterns of Age-Specific Acceleration for Different Parameters of the Model

(A) Slower clonal expansion shifts peak acceleration to later ages. Parameters for all curves are $n = 4$, $x_0(0) = 10^8$, $K_i = 1$ for $i = 0, \ldots, n - 2$, and $K_{n-1} = 10^6$. The curves from left to right have values of $r_{n-1} = 0.4, 0.2, 0.1$ for the blue, green, and red curves, respectively. The mutation rate per year was adjusted so that the total incidence of cancer over all ages up to 80 years is 10%, requiring mutation rates for the curves from left to right of $v_i = 10^{-5}$ multiplied by 3.15, 4.35, and 8.0 for all i. Note that a rapid round of clonal expansion effectively reduces by one the number of steps, n, so that for $n = 4$, one round of rapid clonal expansion yields a nearly constant acceleration of $n - 2 = 2$ over all ages. By contrast, slow clonal expansion often causes a midlife peak in acceleration.

(B) An increase in the maximum size of a clone raises peak acceleration until the clone becomes sufficiently large that a mutation is almost certain in a relatively short time period. Parameters as in (A), except that $r_{n-1} = 0.2$, and for the blue, green, and red curves, respectively, $K_{n-1} = 10^6, 10^4, 10^2$, and $v_i = 10^{-5}$ multiplied by 4.35, 4.45, 6.8 for all i to keep the total incidence of cancer at 10%.

(C) Multiple rounds of clonal expansion greatly increase peak acceleration and shift peak acceleration to a later age. Parameters are $n = 4$, $x_0(0) = 10^4$, $r_i = 0.5$ for all i, $K_0 = 1$, and $K_{n-1} = 10^6$. For the lower (blue) curve, clonal expansion occurs only in the last round before cancer, so $K_{n-2} = K_{n-3} = 1$. For the middle (green) curve, clonal expansion occurs in the last two rounds before cancer, with $K_{n-2} = 10^6$ and $K_{n-3} = 1$. For the upper (red) curve, clonal expansion occurs in the last three rounds before cancer, with $K_{n-2} = K_{n-3} = 10^6$. The mutation rates for the blue, green, and red curves, respectively, are $v_i = 5.8 \times 10^{-4}$, 9.3×10^{-5}, and 1.55×10^{-6} for all i to keep the total incidence of cancer at 10%.

to stimulate new thought along these lines and new empirical tests to sort out the alternatives. Here, I list a few factors in individual progression that provide more realistic detail for the role of somatic mutation or have been discussed as important complements to somatic mutation and clonal expansion.

Many studies emphasize the importance of environmental factors in progression [11–13], including chronic irritation of tissues and changes in hormonal status. Irritation may, for example, increase tissue renewal and cell division, which would feed back on the rate of mutational accumulation [14]. It would be interesting to look at particular types of chronic irritation, such as prostatitis, to study if there may be a plausible link between the acceleration in the epidemiological patterns and the processes of progression promoted by irritation. Similarly, are there known hormonal syndromes that could explain differences in age-specific acceleration between breast and prostate cancer? The ideas presented here should stimulate new alternative explanations and approaches for empirical study.

I assumed that each tissue has a fixed number of rate-limiting mutational steps required for cancer, but the actual number required is likely to vary from case to case [15, 16]. Although certain genes are frequently mutated in the cancers of particular tissues, no tumor has the same spectrum of mutations. For example, different mutational pathways probably can lead to similar patterns of genetic instability. At present, no one has established even a plausible link between variation in the number of rate-limiting mutations that can lead to cancer progression and the epidemiological patterns of variable age-specific acceleration. Again, the data and ideas presented here should stimulate new work.

Reduced tissue renewal and cell division at later ages may explain a decline in the rate at which mutations accumulate and, thus, a reduction in late-age acceleration. This may be particularly important in the oldest age classes but probably does not explain the steady decline in breast cancer acceleration at all ages or the extreme drop in acceleration in prostate cancer just after midlife. The key here may be to understand more about age-specific fluctuations in cellular turnover and how those fluctuations relate to the epidemiological patterns of acceleration.

Mathematical models have studied how chromosomal instability affects mutation rates and progression within individuals [17], but those dynamics within individuals have not been connected to the aggregate patterns of incidence in the population (epidemiology). It seems likely that an increase in mutation rate caused by chromosomal instability would lead to fewer rate-limiting steps, but the nearly instantaneous change in mutation rate would not lead to the slow rise and fall in acceleration over many years as might happen for slow processes of clonal expansion (see Figure 3).

Somatic mutations during development [18] and inherited genetic predisposition cause heterogeneities between individuals in rates of progression, which will affect the age-specific acceleration curves at early ages. This has not been studied mathematically, but it seems that genetically predisposed individuals are born with fewer steps, n, remaining and, so, contribute lower accelerations to the aggregate population through the early to middle years of life, after which most of those individuals have suffered cancer or died. Loss of the predisposed, low-acceleration group in midlife may lead to a midlife rise in acceleration as cases become dominated by the nonpredisposed part of the population.

Ascertainment may change in age-specific ways, affecting the age-specific accelerations in the epidemiology curves. For example, increased testing for prostate cancer in midlife could cause a midlife rise in acceleration. However, the observed peak in acceleration for prostate cancer occurs at about age 40, just as testing generally begins to increase, suggesting that ascertainment alone may not be a sufficient explanation.

Conclusions

In summary, various tissues show different patterns of age-specific acceleration in cancer. These epidemiological patterns of acceleration have not been emphasized before and may provide important clues about the differences between tissues in molecular and cellular processes that determine cancer progression.

The models here emphasize two new points. First, a rise in the number of mutations per cellular lineage with age may explain a decline in age-specific acceleration, as seen in breast cancer (compare the blue curves in Figures 1B and 2A). This can be tested by measuring the age-specific accumulation of mutations in cellular lineages and calculating, at different ages, empirical frequency distributions of the sort shown in Figure 2B. Such tests will become easier to do as more is learned about which genes are mutated in cancers and as rapid methods for genotyping cells improve in efficiency.

The second new point concerns multiple rounds of clonal expansion. Multiple expansions may explain high age-specific accelerations in midlife, followed by a rapid drop in acceleration in later life, as observed in prostate cancer (compare the red curves Figures 1C and 3C). The susceptibility of different tissues to multiple clonal expansions can be tested empirically. One could, for example, study the frequency and size of precancerous clonal expansions by genotyping cells and measuring the spatial spread that follows from particular mutational events. Multiple clonal expansions may also be studied by modifying the techniques of phylogenetic analysis and molecular evolution that have been used to study multiple demographic expansions in populations [19, 20].

The rough matches between observation and theory shown here establish new hypotheses for cancer progression. The models should stimulate new empirical tests and help to foster closer ties between epidemiology, the molecular processes that drive cancer progression, mathematical models of cellular dynamics, and statistical analyses of DNA sequences to infer the cellular evolution of cancers.

Experimental Procedures

Cell Lineages

The intestine provides the clearest example of how to define a cell lineage. The intestine is renewed from basal stem cells [21]. Stem cell division gives rise to one stem cell that maintains the lineage and one transit cell. The transit cell divides a number of times, pushing cells above up to the surface of the tissue. Cells at the

surface slough off. The net effect is renewal of cells from below and loss from the surface. The model presented here requires that, at any time, cell lineages be taken as those cells that will be progenitors of the tissue at a future time. How to turn this into a practical approach for measurement will depend on the experimental methods and the particular tissue. In the intestine, any cell can be measured to assess the distribution of mutations over time because each cell is either a progenitor or a recent descendant from a progenitor. Some other epithelial tissues, such as the skin, also have clear stem-transit architectures [22]. In other tissues, such as the prostate, the nature of tissue architecture and cell lineages remains controversial; some have argued for a stem-transit division [23], but more work is needed.

Acknowledgments

National Science Foundation grant DEB-0089741 and National Institutes of Health grant AI24424 support my research.

Received: November 3, 2003
Revised: November 25, 2003
Accepted: November 25, 2003
Published: February 3, 2004

References

1. Armitage, P., and Doll, R. (1954). The age distribution of cancer and a multistage theory of carcinogenesis. Br. J. Cancer *8*, 1–12.
2. Cook, P.J., Doll, R., and Fellingham, S.A. (1969). A mathematical model for the age distribution of cancer in man. Int. J. Cancer *4*, 93–112.
3. Vogelstein, B., and Kinzler, K.W. (2002). The Genetic Basis of Human Cancer, Second Edition (New York: McGraw-Hill).
4. Kaldor, J.M., and Day, N.E. (1996). Mathematical models in cancer epidemiology. In Cancer Epidemiology and Prevention, Second Edition, D. Schottenfeld and J. F. Fraumeni, eds. (New York: Oxford University Press), pp. 127–137.
5. Armitage, P., and Doll, R. (1957). A two-stage theory of carcinogenesis in relation to the age distribution of human cancer. Br. J. Cancer *11*, 161–169.
6. Moolgavkar, S.H., and Venzon, D.J. (1979). Two-event models for carcinogenesis: incidence curves for childhood and adult tumors. Math. Biosci. *47*, 55–77.
7. Moolgavkar, S.H., and Knudson, A.G. (1981). Mutation and cancer: a model for human carcinogenesis. J. Natl. Cancer Inst. *66*, 1037–1052.
8. Luebeck, E.G., and Moolgavkar, S.H. (2002). Multistage carcinogenesis and the incidence of colorectal cancer. Proc. Natl. Acad. Sci. USA *99*, 15095–15100.
9. Fisher, J.C. (1958). Multiple-mutation theory of carcinogenesis. Nature *181*, 651–652.
10. Murray, J.D. (1989). Mathematical Biology (New York: Springer-Verlag).
11. Roskelley, C.D., and Bissell, M.J. (2002). The dominance of the microenvironment in breast and ovarian cancer. Semin. Cancer Biol. *12*, 97–104.
12. Park, C.C., Bissell, M.J., and Barcellos-Hoff, M.H. (2000). The influence of the microenvironment on the malignant phenotype. Mol. Med. Today *6*, 324–329.
13. Itaya, T., Judde, J.G., Hunt, B., and Frost, P. (1989). Genotypic and phenotypic evidence of clonal interactions in murine tumor cells. J. Natl. Cancer Inst. *81*, 664–668.
14. Cairns, J. (1998). Mutation and cancer: the antecedents to our studies of adaptive mutation. Genetics *148*, 1433–1440.
15. Kerangueven, F., Noguchi, T., Coulier, F., Allione, F., Wargniez, V., Simony-Lafontaine, J., Longy, M., Jacquemier, J., Sobol, H., Eisinger, F., et al. (1997). Genome-wide search for loss of heterozygosity shows extensive genetic diversity of human breast carcinomas. Cancer Res. *57*, 5469–5474.
16. Jiang, F., Desper, R., Papadimitriou, C.H., Schaffer, A.A., Kallioniemi, O.P., Richter, J., Schraml, P., Sauter, G., Mihatsch, M.J., and Moch, H. (2000). Construction of evolutionary tree models for renal cell carcinoma from comparative genomic hybridization data. Cancer Res. *60*, 6503–6509.
17. Nowak, M.A., Komarova, N.L., Sengupta, A., Jallepalli, P.V., Shih, Ie-M., Vogelstein, B., and Lengauer, C. (2002). The role of chromosomal instability in tumor initiation. Proc. Natl. Acad. Sci. USA *99*, 16226–16231.
18. Frank, S.A., and Nowak, M.A. (2003). Developmental predisposition to cancer. Nature *422*, 494.
19. Slatkin, M. (2001). Simulating genealogies of selected alleles in a population of variable size. Genet. Res. *78*, 49–57.
20. Shibata, D. (2002). Molecular tumour clocks and colorectal cancer: seeing the unseen. Pathology *34*, 534–540.
21. Bach, S.P., Renehan, A.G., and Potten, C.S. (2000). Stem cells: the intestinal stem cell as a paradigm. Carcinogenesis *21*, 469–476.
22. Janes, S.M., Lowell, S., and Hutter, C. (2002). Epidermal stem cells. J. Pathol. *197*, 479–491.
23. De Marzo, A.M., Meeker, A.K., Epstein, J.I., and Coffey, D.S. (1998). Prostate stem cells: expression of the cell cycle inhibitor $p27^{Kip1}$ in normal, hyperplastic, and neoplastic cells. Am. J. Pathol. *153*, 911–919.