Review

Genetic instability in cancer: Theory and experiment

Robert A. Beckman a,⁎, Lawrence A. Loeb b

a Department of Clinical Research and Development, Hematology/Oncology, Centocor Inc., Malvern, PA 19355-1307, USA
b Department of Pathology, University of Washington School of Medicine, Seattle, WA 98195-7470, USA

Abstract

Epidemiologic data and molecular biology have combined to demonstrate that multiple genetic changes may be required in carcinogenesis. Mutator mutations, defined as genetic changes which increase the rate of genetic change, including both single base changes and chromosomal instability, may accelerate this process. Key questions remain in defining the role of mutator mutations in carcinogenesis as well as in cancer therapy. Theoretical approaches, including deterministic and stochastic models, have played a significant role in hypothesis generation, experimental design, and refinement of conclusions in this field, and are expected to continue to do so in the future.

© 2005 Elsevier Ltd. All rights reserved.

Keywords: Genetic instability, Mutator hypothesis, Cancer, Theory, Mathematical models

Contents

1. Introduction ................................................................. 424
2. Theoretical approaches to genetic instability ........................................... 424
  2.1. Objectives of theory .................................................. 424
  2.2. Theoretical methods ................................................ 425
3. How many genetic changes are required to produce a cancer cell? .............. 426
4. Is genetic instability essential for generating a sufficient number of random mutations to produce key mutations in genes required for carcinogenesis? .................................................. 428
5. Does genetic instability accelerate carcinogenesis? .................................... 428
6. Is the baseline level of genetic instability in wild type stem cells, from which cancers originate, “optimal” for maximizing the rate of tumor evolution? .................................................. 430
7. What kinds of genetic instability may contribute to carcinogenesis? .......... 430
  7.1. DNA polymerases and genetic instability ............................. 430
  7.2. DNA repair enzymes and genetic instability ....................... 431
  7.3. Chromosomal instability ............................................. 431
  7.4. Cell cycle checkpoints and genetic instability ..................... 432
8. What is the role of genetic instability in relation to therapy? ....................... 432
9. Could the clinical appearance of cancer be prevented by decreasing the rate of genetic change? .......................................................... 432
10. Conclusion ........................................................................ 433
Acknowledgements ........................................................................ 433
References .................................................................................. 433

⁎ Corresponding author at: 1045 Hereford Drive, Blue Bell, PA 19422, USA. Tel.: +1 610 651 7579; fax: +1 610 651 6140.
E-mail addresses: eniac1@snip.net (R.A. Beckman), laloeb@u.washington.edu (L.A. Loeb).

1044-579X/$ – see front matter © 2005 Elsevier Ltd. All rights reserved.
doi:10.1016/j.semcancer.2005.06.007
1. Introduction

The increase of cancer incidence with age has long been taken as evidence that there are multiple rate limiting steps in the development of cancer [1–3], and these steps have been assumed to involve genetic and/or epigenetic changes. In order for cells to become tumorigenic in laboratory models, at least six cellular phenotypes must be altered [4], which implies derangement of a corresponding number of metabolic pathways [5]. Selective expansion of altered clones may occur at any time during the process [6]. Thus, carcinogenesis may be viewed as an evolutionary process involving both genetic change and selection.

If multiple genetic changes are required for carcinogenesis, anything which increases the rate of genetic change may accelerate the process [7]. Loeb [8] postulated that mutations leading to greater genetic instability were required to explain the accumulation of the multiple cancer-associated mutations within a cellular lineage at the incidence rates observed within a human lifetime. These mutations, which increase the inherent rate of genetic change, are referred to as mutator mutations, cell lineages harboring them are called mutator clones, and cancer cells are said to exhibit a mutator phenotype [9]. Mutator mutations and genetic instability are generalized concepts, referring not just to mutations which lead to enhanced rate of single nucleotide substitutions, but also to mutations that enhance changes in the number of repeats in repetitive DNA sequences, termed microsatellite instability (MIN) [10,11]; mutations that enhance the rate of chromosome loss, gain, or translocation, termed chromosomal instability (CIN) [12]; and mutations which interfere with cell cycle checkpoints related to DNA damage [13].

In summary, mutations in a mutator gene can result in increased changes in DNA nucleotide sequence throughout the genome. However, other authors have pointed out that if early genetic changes confer a survival advantage on premalignant clones, their progeny will increase in number (clonal expansion of premalignant clones), potentially increasing the frequency of appearance of mutations in the expanded clones and thus accounting for the required number of genetic changes within a cancer without invoking a mutator phenotype [14]. The role of mutator mutations and whether they are essential for carcinogenesis should be amenable to mathematical analysis.

Genetic changes which are required in carcinogenesis are divided into two very broad classes: those which are dominant, requiring alteration of only one gene copy to contribute to a premalignant or malignant phenotype, and those which are recessive, requiring alteration of both gene copies to contribute to a premalignant or malignant phenotype. Genetic changes in dominant cancer-associated genes typically promote cancer through gain of function, whereas genetic changes in recessive cancer-associated genes typically promote carcinogenesis through loss of function, and hence the involved genes are often referred to as tumor suppressors. Dominant oncogenic mutations can occur in tumor suppressor genes if the mutation is dominant negative, i.e. if a loss of function of one copy of the gene product also suppresses the function of the gene product of the wild type copy. Often this is the case when a gene product functions in a multimeric fashion.

The importance of recessive tumor suppressor genes in carcinogenesis was predicted by Knudson, utilizing a statistical model of familial and sporadic retinoblastoma incidence, prior to the molecular identification of any specific recessive oncogene [15]. Now the study of recessive oncogenes is a major field in oncology [16]. This classic paper demonstrates the ability of theory to substantially influence the thinking of experimental cancer researchers.

Many significant questions remain concerning genetic instability in cancer, each of which may be amenable to theoretical analysis:

1. How many genetic changes are required to produce a cancer cell?
2. Is genetic instability essential for generating a sufficient number of random mutations to produce key mutations in genes required for carcinogenesis?
3. Does genetic instability accelerate carcinogenesis?
4. Is the baseline level of genetic instability in wild type stem cells, from which cancers originate, "optimal" for maximizing the rate of tumor evolution?
5. What kinds of genetic instability may contribute to carcinogenesis?
6. What is the role of genetic instability in relation to therapy?
7. Could the clinical appearance of cancer be prevented by decreasing the rate of genetic change?

Answers to these questions are gradually emerging, in part due to a particularly fruitful interaction between theory and experiment. This article reviews genetic instability in cancer and focuses on the role of theory and experiment in addressing these questions.

2. Theoretical approaches to genetic instability

2.1. Objectives of theory

Albert Einstein summed up several objectives of theory succinctly: "Only theory can tell us what to measure and how to interpret it." It should be noted that theories can be either descriptive or mathematical. However, mathematics is an important tool which enables theoreticians to rigorously state and evaluate theories as well as to confirm or refute theories through analysis of experimental results.

The potential areas in which theory can contribute to scientific inquiry are shown in Fig. 1. Theory can contribute to the generation of hypotheses, as Knudson’s analysis of the incidence of retinoblastoma led to a hypothesis concerning the importance of recessive cancer genes (see Section 3). A detailed theoretical analysis will often lead to the identification of key parameters which frame or answer questions of
interest. For example, Nowak et al. calculated the conditions required for a chromosomal instability mutation to be an initial event in colorectal cancer pathogenesis (see Section 7) [17]. Both the generation of biologically important hypotheses and the identification of key parameters that affect them can delineate the prioritization of experimental efforts. These insights can also affect experimental design. For example, based on a careful theoretical analysis of biochemical mechanisms of ensuring the accuracy of DNA replication, experiments were designed to distinguish among closely related hypotheses for error correction or proofreading. The results of these experiments in conjunction with the theoretical analysis demonstrated that proofreading by *Escherichia coli* DNA polymerase proceeds in two sequential kinetically distinct steps, resulting in even higher accuracy (see Section 7) [18].

Application of a theory may lead to predictions of experimental results in advance, the most stringent test of a theory. For example, Einstein’s theories predicted the extraordinary release of energy upon the annihilation of matter. Because of the complexity of biological systems, such an outcome is extremely difficult to achieve, and hence the dotted line in Fig. 1. Nonetheless, Knudson’s work on recessive oncogenes provides an example in oncology [15].

Experimental verification is the ultimate criterion for establishing the value of any important theory. Theory can then contribute substantially to evaluation and analysis of the experimental data. At times, theory may lead to interpretations of experiment that are not evident from qualitative analysis. For example, Moolgavkar and Knudson [19] explained the observed incidence curves of childhood cancers based on a detailed model involving mutations and clonal expansion of premalignant clones, to be discussed in Section 3. This was far from obvious from inspection of these incidence curves. At other times, a theoretical analysis can lead to rejection of conclusions which appeared to be intuitive. For example, multiple proposed biochemical mechanisms for preserving the accuracy of DNA replication appeared intuitively to be in qualitative agreement with experiment. However, a series of experiments designed with the aid of theory, and subsequent analysis of the results, demonstrated that the majority of these mechanisms were not in detailed agreement with experiment (see Section 3) [18].

### 2.2. Theoretical methods

Theoretical models fall into two very broad classes: deterministic and stochastic models. Deterministic models attempt to model or predict the average behavior of systems according to precise rules. In contrast, stochastic models describe the probability of very specific behaviors of individuals rather than average behavior of the population. Newtonian physics is deterministic; quantum mechanics is stochastic. Deterministic models often use differential equations to exactly describe the changes in average properties of a system over a very brief period of time. A group of interdependent differential equations can be used to describe interacting properties of a complex system. These equations, representing short time intervals, can be integrated, i.e. summed, over longer time periods of interest. For example, one can sum the changes in a cell lineage that may occur in one cell generation over a human lifetime. These techniques can be applied over broad ranges of phenomena since they require only understanding of average system properties. The results can be smooth or continuous because they represent average properties. Knudson [15] applied a deterministic model to his analysis of the incidence of retinoblastoma over time (see Section 3).

Stochastic models evaluate the entire probability distribution of random individual events. This kind of model is potentially more informative in that it considers rare events, not just average properties. Since cancers arise from single cells, many theoreticians posit that rare events may determine the course of a malignancy. Typically one defines a variety of discrete states, for example, cell lineages with given numbers of cancer-associated mutations, and the rates or probabilities of transition between the states. Often the different states of a phenomenon of interest can be represented as a Markov chain. In a Markov chain, the system passes through the defined states in discrete steps with a given set of transition probabilities. The possibilities for where the system will go next, and the chance it will “select” a particular option, depend only on where the system is at the moment (i.e. its present state) rather than on how it got there (its history). This type of analysis can in principle give the chance that the system is in a given state as a function of time or other key variables. However, utilizing this approach often requires a detailed understanding of individual states and transitions, which is not always available. As the system complexity increases, the definition of all the relevant states and the mathematical analysis of all the transitions between them can become daunting. The results will be probabilities of discrete events, rather than average properties. For example, Knudson [15] modeled the probability distribution for having bilateral or unilateral retinoblastoma.
Theoretical modeling frequently uses simplifying assumptions. Simplifying assumptions eliminate complexities which may be peripheral to the issue under consideration, allowing a focus on key features of a complex biological system. For example, most models of genetic instability assume that the rate of genetic change is constant at any location in the genome, even though there is evidence of mutation "hot spots" which violate this assumption [20]. In determining whether this simplifying assumption impacts the results when modeling genetic change in carcinogenesis, one would need to know whether mutation "hot spots" exist at key loci within cancer-associated genes. Simplifying assumptions should be disclosed in detail when a theoretical model is presented. Where feasible, the truth of these assumptions should be compared to what is known and the impact of the simplification on the analysis should be assessed.

Limiting cases involve a particular type of simplifying assumption. Where a particular aspect of the problem is not determined from prior work, an extreme assumption may be utilized and that case studied. Often if the conclusion is true in this type of extreme or "worst case" scenario, it can be shown to be true for any other reasonable scenario. For example, we have examined the effect of mutations which reduce cellular fitness on tumor evolution of mutator clones. Wherever there was uncertainty regarding aspects of this problem, we made an extreme assumption which maximized the importance of the effect. Therefore, we argue that the true effect cannot be more than what we have estimated (Beckman RA, Loeb LA, Genetics, 2005, in press).

Parameters are the variables which, based on the theoretical analysis, are expected to influence the outcome of interest. In some cases, values of the parameters are known from prior experiment, and therefore these values can be fixed. In other cases, the values of the parameters are unknown or could reasonably be expected to vary over a known range. In this case, the parameters are adjustable. A common error is to use a theoretical model with a large number of adjustable parameters to fit a relatively limited experimental dataset. The fact that the model can fit the data does not guarantee the validity of the model in this case, since nearly any model could be fit to the data by adjusting the parameters. In general, it is important to verify that the number of fitted or predicted experimental data points exceeds the number of adjustable parameters in the model. The greater the excess of independent experimental data points over adjustable parameters, the more valid the experimental confirmation of the theory.

3. How many genetic changes are required to produce a cancer cell?

Over half a century ago, Nordling [1] cited literature dating back to the 1920s linking cancer to genetic change and noting that most carcinogens are mutagens. Analyzing cancer death rates at all tumor sites in the United States, United Kingdom, France, and Norway, he noted that cancer deaths increased approximately as the sixth power of age between the ages of 25 and 74. Thus, the logarithm of cancer deaths plotted versus the logarithm of time yielded a straight line with slope six. He postulated that seven genetic changes would therefore be required to produce a cancer cell: one pre-existing mutation and six subsequent mutations, the latter rate limiting, leading to a dependence on the sixth power of age. We note that other interpretations of this data are possible. For example, Fisher and Holloman [21] suggested that at least six cells must each acquire one mutation to form a sufficient cluster of genetically altered cells to result in a tumor. This hypothesis results in the same time dependence as that preferred by Nordling.

Armitage and Doll [2] analyzed the data by organ specific tumor sites, utilizing incidence figures, which are more pertinent than mortality figures. They noted that cancers in organs influenced by hormones, as well as lung cancer, which is primarily caused by smoking, do not fit the sixth power incidence law as closely. They pointed out therefore that if the factors leading to various genetic changes varied over time, such that the risk of acquiring the next genetic change was not constant, observed incidence would deviate from the sixth power law. The rate of increase of incidence at a time point on a log-log plot reflects the rate of genetic change at some time years or even decades in the past, due to the latency between initiation of carcinogenesis and appearance of a clinical cancer. They also pointed out that fewer than seven steps could still lead to a sixth power law if one or more of those steps increased in greater than linear proportion with age. Fisher [6] considered a model in which premalignant clones of epithelial tumors expanded radially within an epithelial plane, such that the number of premalignant cells increased as the square of time. A mechanism involving three mutations, one pre-existent, and two rounds of radial clonal expansion (after the first and second mutation) was consistent with the sixth power law. It should be noted that each genetic change must result in a phenotypic change leading to clonal expansion; thus alterations in tumor suppressor genes, which require two rate limiting steps, would not fit into a sixth power model in this proposal.

Cook et al. [3] examined more types of cancer and data from more countries. Utilizing this wider dataset, they found that the sixth power relationship held precisely in only a minority of cases, especially if the whole human lifespan was taken into account, where downward curvature in incidence rates on a log-log plot was common. This latter phenomenon might be due to the delay between exposure to a risk factor and development of cancer. Thus, exposure of an elderly individual to a carcinogen might not result in a tumor prior to that individual dying from another cause. Tumors from different organ sites had characteristic curvatures, either upward or downward. Again, time variation in exposure to causative factors was viewed as the most likely explanation, and in some cases such as prostate cancer, the curves were best fit
by assuming that the initiating event occurred later in life, so the data needed to be plotted as a function of time from the initiating event. In addition, different cancers increased as different powers of time, ranging from 1 (melanoma) to 11 (prostate cancer when time is measured from birth), with most cancers falling between 3.5 and 6.5. Pike [22] pointed out that the incidence data could not prove a single unique mechanism for carcinogenesis.

Knudson [15] applied a deterministic model to the incidence of the childhood tumor retinoblastoma as a function of time from birth. He noted that the incidence of the hereditary form was linear as a function of time from birth, whereas the incidence of the sporadic form increased with the square of time. Noting that these two patterns could reflect the need for one and two genetic “hits,” respectively, he suggested that the two hits may not actually be inactivation of two copies of a recessive gene which suppressed cancer. In hereditary retinoblastoma, the individual has inherited the first hit, inactivating one copy of the gene. The retinoblastoma (Rb) gene was subsequently cloned [23] and shown to revert tumorigenic properties of cell lines deficient in its function.

Knudson also applied a stochastic model to explain the observed distribution of unilateral versus bilateral tumors in hereditary retinoblastoma. The appearance of a tumor (in this case due to the second mutation) is modeled using a Poisson process, which assumes that the probability of the second mutation occurring is small and constant over small time periods. He calculated that on average three tumors would appear, and from this fact and the relative incidence of hereditary and sporadic retinoblastoma, he showed that the two “hits” in sporadic retinoblastoma must each occur at a rate of \(10^{-7}\) per cell per year. It has subsequently been recognized that the two hits may not actually be inactivation of two copies of the Rb gene. Rather the first hit may be inactivation of one copy of the Rb gene, leading to a heterozygous cell with one wild type and one mutated gene, and the second hit may be a mutation which dramatically increases chromosomal instability, leading to rapid loss of heterozygosity (LOH) at the Rb gene locus (i.e. deletion of a chromosomal segment encompassing the second gene copy) [17].

Knudson’s work inspired more detailed and general two-stage carcinogenesis models. Moolgavkar and Knudson [19] developed a model in which after the first hit, there is exponential clonal expansion of premalignant clones, followed by a second hit which need not necessarily be a mutation. This type of model is termed a TSCE or “two-stage clonal expansion” model. The model is used to explain the incidence of all cancers, including childhood cancers, hormonally mediated cancers, lung cancer, and sporadic and hereditary colon cancer. The model is also discussed in terms of the two-stage paradigm for tumor initiation and promotion characteristic of experimental carcinogenesis in rodents. Initiators would influence the mutation steps, whereas promoters would influence the premalignant clonal expansion. The authors have demonstrated that as few as two mutations may be required, but they are careful to point out that the data are also compatible with more mutational steps, especially in the absence of external promoters such as smoking or hormones. Doll [25] points out that the incidence of lung cancer in non-smokers fits very well to a curve increasing as the fourth power of time, something which could be equally well modeled by the need for four new mutations or by a TSCE. Luebeck and Moolgavkar [26] performed a detailed analysis of the incidence of colorectal cancer utilizing the U.S. Surveillance, Epidemiology, and End Results (SEER) registry, and adjusting for birth cohort and calendar year effects that cause cancer incidence to vary independently of mechanism. They studied a family of models with clonal expansion of premalignant clones incorporated, concluding that a model with two rare events followed by clonal expansion and then a third rare event fits the data best. A model with a fourth rare event required fits the data nearly as well however.

Thus, based on the epidemiologic data, it would appear that anywhere from 2 to 7 rate limiting steps are required prior to the clinical appearance of a tumor, and at least some of these steps are likely to involve genetic change. In addition, it appears that the number of steps may depend on the organ site of the primary tumor. Molecular and cell biological advances have clarified what some of these steps may be. Transformation assays in vitro and tumorigenesis studies using rodent xenograft models have defined six cellular control pathways which must be acquired by a cell in order for it to become tumorigenic. These include self-sufficiency with respect to positive growth signals, insensitivity to growth inhibitory signals, evasion of apoptosis, limitless replicative potential, sustained angiogenesis, and tissue invasion and metastasis [4]. More recently, transformation of cells in culture with genes carried by viral vectors has defined six cellular control pathways which must be altered in human fibroblasts to create a tumorigenic phenotype in mouse xenografts and in vitro: p53, Rb, PP2A, telomerase, Raf, and Ral-GEFs. Since p53 and Rb are tumour suppressor genes, up to eight mutations could be required. In contrast, murine fibroblasts require only perturbation of two pathways, p53 and Raf. Thus, experimental carcinogenesis in mice may not be representative of human carcinogenesis. Tissue specificity is observed for different cell types: embryonic kidney cells required perturbation of the PI3 kinase pathway but not the Raf pathway, and mammary epithelial cells required perturbation of all the pathways considered, a total of seven [5]. We note that while this discussion classifies tumors and their carcinogenic mechanism by organ site of origin, a more fundamental classification of tumors is emerging based on gene expression patterns. These molecular classifications can have prognostic value [27], and it is likely that these differing molecular subtypes of cancer will require different numbers of mutations in their pathogenesis. Since mutations occur randomly throughout the genome, a cancer cell is likely to harbor many random mutations in addition to those in the critical sites in cancer-associated genes.
4. Is genetic instability essential for generating a sufficient number of random mutations to produce key mutations in genes required for carcinogenesis?

Loeb et al. [7-9] have suggested that mutator mutations, i.e. mutations which increase the rate of genetic change, may be required for cancer cells to be formed within a human lifetime. The rate of mutation per nucleotide locus is estimated to be as low as $10^{-9}$ per cell generation in somatic cells [28], and may be as low as $10^{-11}$ in the immortal stem cells which likely give rise to tumors [29]. The number of stem cell generations in a human lifetime is unknown and varies among tissues, but estimates range from on the order of $10^7$ for the majority of reproductive stem cells which are usually not in cycle [30,31] to on the order of $10^4$ for colonic stem cells [32]. The total number of cells in the human body is on the order of $10^{14}$ (based on a linear cell dimension of 10 μm, a density of 1 g/cm$^3$, and a human weight on the order of 100 kg), although only a small fraction of these cells are stem cells. If six independent, specific mutations are required for carcinogenesis, and in the absence of selection and clonal expansion of premalignant clones, the probability of any one cell lineage acquiring these independent mutations is the probability of acquiring one mutation raised to the sixth power, where the probability of acquiring each mutation is approximately the mutation rate times the number of cell generations. If any of these events required two mutations in a tumor suppressor gene, then the probability of producing a cancer would even be further reduced. The total probability of any stem cell in the body reaching this state would be approximately the probability for any one stem cell lineage multiplied by the number of stem cells, a number which would be much less than one for any of the parameter values mentioned above. Thus, according to this argument, a mutator mutation would be required to explain the appearance of malignant cells.

However, as we have seen above, the number of genetic changes required to produce a cancer cell could be as few as two, and that would significantly affect the calculation above. Moreover, it is likely that premalignant clones have some degree of survival advantage and some degree of expansion of these clones occurs, enhancing the probability of mutation for subsequent steps. Depending on the population size, for example, stochastic modeling has shown that in the presence of clonal expansion, two mutations can occur with two, one, or zero rate limiting steps, and can occur proportionately to either the first power or the square of time [33]. Selection of premalignant clones and subsequent expansion of them has been argued to be the basis of malignancies occurring without mutator mutations [6,14].

Waves of mutation and clonal expansion can be sequential, leading to a succession of pure populations with increasing numbers of mutations [34]. Alternately after a given mutation, there can be partial expansion of the clone followed by mutation of one of its members to the next state on the way to malignancy. In this case, mosaic populations will be present, and a cell population may ultimately acquire a full complement of cancer-associated mutations without ever having had a pure population with an intermediate number of mutations. The latter phenomenon, termed "tunneling", has been studied in detail using stochastic models, and the conditions for it occurring have been defined [35,36]. Thus, the occurrence of cancer without mutator mutations must be at least considered possible. Indeed, Moolgavkar and Knudson [19] have modeled cancer incidence rates successfully for a large variety of cancers without invoking mutator mutations. However, several caveats need to be borne in mind. First, while the epidemiologic data could be consistent with anywhere from 2-6 new mutations, the number of phenotypes which must be acquired [4] and signaling pathways which must be affected [5] suggests that the number of mutations might be six or more. The more mutations are required, the less of the slope of six on the log-log plots of cancer incidence can be due to clonal expansion. Second, the mutation rates calculated by Moolgavkar and Knudson [19] and others are potentially consistent with wild type mutation rates in somatic cells [28], but as cancers are thought to arise from stem cells, and these may have significantly lower mutation rates [29], mutator mutations may actually be required even to generate the mutation rates calculated by Moolgavkar and Knudson [19]. Third, the effect of clonal expansion is greatest if it occurs early, requiring a significant fitness advantage of premalignant clones after only one or a small number of mutations for a maximal effect. Fourth, the modeling of incidence rates without a mutator mutation assumes that a cancer cell once formed has a very high probability of becoming a clinical cancer [19]. This assumption may be false, due to host immune defenses, as well as the need for small cancer deposits to successfully induce angiogenesis. In that case, a higher rate of generation of cancer cells would be required to explain observed incidence rates.

Given these considerations, we conclude that it has not been conclusively demonstrated that carcinogenesis requires a mutator mutation. Whether or not a mutator mutation is necessary depends on the wild type stem cell mutation rate, the number of mutations actually required for carcinogenesis, the degree and timing of selection for and clonal expansion of premalignant clones, and the probability that a cell with the genetic prerequisites for cancer develops into a clinical cancer.

5. Does genetic instability accelerate carcinogenesis?

Based on epidemiological data, it can be inferred that there is a 20 year interval between initial exposure of an individual to a carcinogen and the clinical detection of a tumor. Whether or not mutator mutations are required for carcinogenesis, they are likely to accelerate this process. Given a large variety of potential competing mechanisms of carcinogenesis, those which are more efficient are more likely to be represented in clinical cancers. If we further assume that the majority of cancer cells created do not result in clinical
cancers due to host immune defenses and failure to stimulate angiogenesis when the tumor deposit exceeds a critical size, highly efficient mechanisms of carcinogenesis may be needed to explain observed rates of cancer (Fig. 2).

In considering the relative efficiency of carcinogenesis in the presence and absence of mutator mutations, several factors which will potentially work against mutator clones must be considered. The first of these is the increased acquisition of potentially deleterious mutations reducing the fitness of mutator clones. This could lead to preferential extinction of mutator clones, a process termed negative clonal selection (NCS).

Beckman and Loeb (Genetics, 2005, in press) utilized a deterministic model to evaluate a limiting case maximizing the impact of negative clonal selection due to accumulation of "reduced fitness" (RF) mutations. Several factors make this model a limiting case. First, NCS was considered in isolation for its impact on survival of mutator clones, neglecting clonal expansion due to cancer-associated mutations that would increase fitness. In fact, oncogenic mutations that increase fitness could be very significant and oppose the effects of NCS. Second, it was assumed that any clone with reduced fitness will become extinct with absolute certainty. In fact, stochastic modeling demonstrates that clones with reduced fitness can still survive in small populations such as might be expected in intestinal crypts, favoring mutator mutations and their spread by random drift [37,38]. While organization of tissues into small cellular compartments may serve to limit the spread of premalignant cells with a survival advantage [39], ironically it also favors genetic instability. Third, the model assumed a constant level of vulnerability to mutations reducing fitness. In fact, it is likely that during carcinogenesis cell lineages will acquire mutations which reduce their vulnerability to apoptosis, mitigating the effect of NCS [40]. Fourth, the model assumed a relatively high level of vulnerability of the cell to mutation, basing its maximum assumed vulnerability on studies of random mutations in proteins [41], and thus not allowing for redundancy of cellular pathways. Both dominant and recessive RF mutations were considered. The model identified key parameters which determine the influence of NCS and defined the conditions under which it might have a substantial effect. Despite the fact that the model examined limiting, maximal cases for NCS, the analysis indicated that NCS did not limit the growth of mutator clones significantly in most cases.

A second factor limiting mutator clones is that cells must first acquire a mutator mutation to become a mutator clone, adding at least one extra step to the process of carcinogenesis. Beckman and Loeb (manuscripts in preparation) utilize a deterministic model to define the conditions under which, despite this additional mutation step, mutator pathways are favored over non-mutator pathways to carcinogenesis, both in the presence and absence of NCS. Mutator pathways are favored to a greater extent if more mutation steps are required for carcinogenesis, and the mutator mutation occurs early.

Nowak et al. [17] utilize a stochastic model to consider a similar question in the case of chromosomal instability mutations in colorectal cancer carcinogenesis, and also conclude that genetic instability can be favored in early stages of carcinogenesis. Komarova and Wodarz [42] have examined chromosomal instability with both stochastic and deterministic models, considering both the extra time it takes to acquire a chromosomal instability mutation, and the possible deleterious consequences of chromosomal instability. They conclude that chromosomal instability may not accelerate carcinogenesis, but if it occurs due to environmental factors, the degree
of it is tuned to optimally balance accelerated carcinogenesis with deleterious consequences on fitness of the clone.

6. Is the baseline level of genetic instability in wild type stem cells, from which cancers originate, “optimal” for maximizing the rate of tumor evolution?

In so far as carcinogenesis can be viewed as an evolutionary process, parallels are drawn between the evolution of tumor clones and species evolution [14]. Deterministic models have found an optimal mutation rate for the evolution of viral subspecies that allows for the maximum rate of variation without creating too many deleterious mutants. This mutation rate is approximately equal to the reciprocal of the genome length [43,44].

This approach can be adapted to derive an optimal mutation rate for carcinogenesis, and define key parameters which affect this optimum (Beckman and Loeb, manuscript in preparation). Komarova and Wodarz [42] calculate an optimum (Beckman and Loeb, manuscript in preparation). Komarova and Wodarz [42] calculate an optimum for maximizing the rate of tumor evolution without creating too many deleterious mutants. This mutation rate is approximately equal to the reciprocal of the genome length [43,44].

In general, the optimal genetic instability rate for tumor evolution depends on the type of instability and may be different from that for species evolution.

7. What kinds of genetic instability may contribute to carcinogenesis?

Genetic instability may affect the enzymes which replicate DNA (DNA polymerases), enzymes which repair DNA (mismatch repair enzymes, nucleotide-excision repair enzymes), proteins which affect chromosomal stability (chromatin structure and condensation proteins, kinetochore proteins, spindle proteins), and proteins which control apoptosis and cell cycle regulation in response to DNA damage (p53 and pRb). Mutations in each of these pathways have been related to the pathogenesis of cancer, either in humans or in animals. The association of these forms of genetic instability with cancer can be viewed as experimental support for the mutator hypothesis.

7.1. DNA polymerases and genetic instability

The first line of defense against genetic instability is the remarkable fidelity of DNA polymerases and associated multienzyme complexes that replicate DNA. Purified replicative DNA polymerase can copy DNA with a fidelity as high as one misincorporation per 10 million nucleotides polymerized. This is five orders of magnitude higher than would be expected theoretically on the basis of Watson–Crick base pairing alone [45], or on the basis of the experimentally observed fidelity of non-enzymatic polymerization [46].

The high fidelity of DNA polymerization is due to two factors: first, the improved selection of the complementary nucleoside triphosphate mediated in part by solvent exclusion [47] and conformational changes at each nucleotide addition step [48], and second, by editing or “proofreading” of errors immediately after they are inserted by the enzyme [18].

While mutations at the polymerization site are likely to have a greater effect on fidelity of polymerization with purified DNA polymerases, mutations that diminish proofreading are more likely to result in enhanced mutagenesis in cells. The latter mutations might not diminish the rate of chain elongation (although elongation of the nascent DNA chain past an error is slower relative to elongation past a correct insertion [18]) and therefore might not result in a proliferative disadvantage. DNA polymerase proofreading has been analyzed extensively both theoretically and experimentally. For E. coli DNA polymerase I, the first DNA polymerase to be studied in depth, an enzymatic activity which could proofread in principle, the 3' → 5' exonuclease activity, was discovered experimentally [49], leading to a variety of subtly different theoretical treatments of proofreading [18,50–55].

Based on a deterministic analysis of branching reaction pathways, Beckman, Loeb and colleagues [18,56–58] designed a series of experiments to distinguish among these models. These experiments utilized the addition of two byproducts and one substrate of the DNA polymerization reaction which could cause mutagenesis by affecting proofreading (“probes of proofreading”), and determined their effects singly and in combination. The results were consistent with only one of the theoretical models, and demonstrated that E. coli DNA polymerase I proofreads a given new base twice before moving on to the next base [18]. The two sequential editing steps give a high fidelity at a relatively low energy cost [59]. The two editing steps may correspond to distinct sites on polymerases which have been seen in crystallographic studies [60,61].

To date, evidence for carcinogenesis mediated by the production of mutator mutants in mammalian DNA polymerases is limited. DNA polymerase-β is the major DNA polymerase in mammalian cells; it is involved in multiple DNA synthesizing processes. A mutator mutation in the proofreading activity of DNA polymerase-β increases the incidence of lymphomas and epithelial tumors in mice [62]. It is possible that similar relevant mutations will be discovered in humans. The theoretical work on proofreading suggests that some proofreading activities which would not be apparent in more conventional assays could be discovered by assaying for the mutagenic effect of an added proofreading probe, the pyrophosphate anion [18].

DNA polymerase-β is responsible for resynthesis of DNA after removal of altered bases by glycosylases including bases altered by oxygen reactive species [63]. Based on scattered reports from multiple laboratories, it has been estimated that 30% of human tumors examined to date express DNA polymerase-β variant proteins [64]. One of these variants,
K289M, which was identified in human adenocarcinoma of the colon, confers a 16-fold increase in mutation frequency in copying repetitive sequences in DNA [65]. Recently, five new error-prone DNA polymerases have been identified in human cells [66]. These error-prone DNA polymerases can copy past altered non-repaired lesions in DNA. In addition, they frequently incorporate non-complementary nucleotides when copying normal DNA templates. DNA polymerase η (Pol η) copies past UV-dimers in DNA templates; deletion of Pol η in patients with a variant form of xeroderma pigmentosum, XP-V, results in a 1000-fold increase in incidence of skin cancers after UV-exposure [67,68]. Presumably, in the absence of Pol η, another error-prone polymerase copies past the thymidine dimers resulting in enhanced mutagenesis. It has been postulated that non-inherited skin cancers are also associated with mutation or diminished production of Pol η [69]. These bypass DNA polymerases may exhibit specificity as to the types of lesions that can be accommodated within the active site. Pol κ is able to copy benzyopyrene adducts in DNA and even incorporate benzyopyrene adducted dGTP; however, it is ineffective in bypassing thymidine dimers [70]. In contrast, Pol η fails to copy benzyopyrene adducts in DNA. The expression levels of bypass DNA polymerases have been quantitated in several human cancers. Pol κ is upregulated in lung cancers [71], and over-expression of at least one of these error-prone DNA polymerases was reported in 45% of the 68 human tumors samples studied [72]. It has not been established whether or not these polymerases are mutated in cancer cells.

7.2. DNA repair enzymes and genetic instability

The second line of defense is the removal of damaged nucleotides in DNA that have mutagenic potential. The nucleotide-excision repair pathway removes and replaces DNA nucleotides damaged by a variety of exogenous agents, including ultraviolet light [73]. Patients with the autosomal recessive disorder xeroderma pigmentosum are sensitive to ultraviolet light, and develop cutaneous tumors in exposed areas at a rate of up to 1000 times that of the general population [74]. These patients have defective nucleotide-excision repair [75]. Mutations in other DNA synthesizing enzymes are also associated with cancer. Amongst the multiple DNA helicases in human cells are members of the RecQ family. Mutations in the genes that encode these helicases are found in Bloom’s and Werner’s syndromes, rare inherited diseases with a high incidence of cancer [76,77].

Mismatch repair genes correct single base substitutions in DNA [78]. A subset (13%) of individuals with sporadic colon cancer and the vast majority of those with hereditary non-polyposis colorectal cancer (HNPPC), an inherited syndrome with increased colon cancer incidence) exhibit microsatellite instability (MIN), i.e. instability in the length of repetitive DNA sequences termed microsatellites [10,11,79]. In addition to colon cancer, many tumor cell lines exhibit deficits in mismatch repair [80]. It has been assumed that DNA polymerase slippage and a lack of mismatch repair contribute to the occurrence of MIN [80].

7.3. Chromosomal instability

Aneuploidy (alterations in chromosome number), gross chromosomal translocations, and molecular loss of heterozygosity (LOH) without grossly visible karyotypic changes are all common features of tumors. The majority of colon cancers exhibit aneuploidy and other chromosomal alterations, comprising the phenotype of chromosomal instability (CIN). They display a extraordinary rate of alternation in chromosome number determined by fluorescence in situ hybridization (FISH) at approximately 10⁻² per chromosome per cell generation [81]. Cell fusion studies indicate that the CIN genotype is genetically dominant. Generally, a given colorectal cell line will exhibit MIN or CIN but not both.

Greater than 100 genes leading to chromosomal instability have been identified in yeast, including those affecting microtubules, kinetochores, chromosome condensation, sister chromatid cohesion, and cell cycle checkpoints for mitosis [82]. Thus far, the molecular basis of CIN has not been identified in the majority of human cancers. Some tumors with colon cancer exhibit CIN on the basis of a heterozygous mutation in the mitotic spindle checkpoint gene hBUB1 [83,84].

The Adenomatous Polyposis Coli (APC) gene is a recessive cancer-associated gene, mutation of which is believed to be one of the earliest steps in the pathogenesis of colorectal cancer. One copy of the gene is already mutated in familial adenomatous polyposis (FAP), a disorder resulting in the growth of hundreds to thousands of colorectal polyps by the teenage years. Invariably, one or more of these polyps progresses to carcinoma. Given that this is a recessive tumor suppressor gene, it can be inactivated either through mutation of both copies, or through mutation in one copy followed by deletion of chromosomal segments encompassing the second copy (LOH). The latter mechanism would be accelerated by a CIN mutation. Utilizing stochastic modeling techniques to simulate small cell populations within colonic crypts, Nowak et al. [17] determine under what conditions the CIN mutation would likely be the initial event in colorectal carcinogenesis, or might be the second event following mutation of the first copy of APC, as a function of various key parameters. The results indicate that only a small number of dominant CIN genes need to exist for CIN to predominate as an initial step under a variety of conditions. Similar results obtain upon consideration of a large cell population using a deterministic model.

Komarova et al. [85] utilize a stochastic model to evaluate the rate of formation of dysplastic crypts by CIN and MIN mechanisms in sporadic colon cancer, HNPPC, and FAP, obtaining broad qualitative agreement with the relative importance of CIN and MIN and the number of polyps generated in these conditions.
7.4. Cell cycle checkpoints and genetic instability

Cell cycle checkpoints prevent progression through the cell cycle, allowing time for DNA repair and accelerating apoptosis in the presence of unrepaired DNA damage. Thus, abrogation of these checkpoints promotes genetic instability. The two most extensively studied checkpoint genes, p53 and pRb, are among the most frequently mutated in human cancers [16,86]. Other checkpoint genes include the ATM gene, which is mutated in ataxia telangiectasia, another rare inherited disorder with increased susceptibility to cancer [87]. There is also evidence that BRCA1 and BRCA2, inherited breast and ovarian cancer susceptibility genes, are involved in cell cycle checkpoints [88].

8. What is the role of genetic instability in relation to therapy?

Genetic instability may play a role in the mechanism of action of chemotherapy. As has been discussed above, genetic instability increases the rate of acquiring mutations which may be essential for carcinogenesis, but also accelerates the acquisition of deleterious mutations which reduces the fitness of clones potentially leading to their extinction. Beyond a critical value, further increases in genetic instability are likely to exceed the optimum and lead to extinction of clones due to NCS. The majority of chemotherapeutic agents interfere with aspects of DNA metabolism and are mutagens. If the mechanism of action of chemotherapy selects for genetically stable cells, as has been elucidated in detail by Komarova and Wodarz [42] utilizing a deterministic model.

A common side effect of chemotherapy is the induction of second malignancies. Given that cytotoxic chemotherapy is mutagenic, it is not surprising it would be carcinogenic. Some authors have suggested that in addition to directly causing mutations, chemotherapy may select for genetically stable cells, as has been elucidated in detail by Komarova and Wodarz [42] utilizing a deterministic model.

Genetic instability may also play a role in resistance to cancer therapies. The probability that a tumor will contain no cells with pre-existing resistance to a given therapy goes down exponentially both with the size of the tumor and with the underlying mutation rate [89]. A mutator phenotype would not only generate the mutations that are required for tumor progression but would in parallel generate an even larger number of mutations that would render cells resistant to chemotherapeutic agents. By the time tumors are clinically apparent, they contain $10^6$ to $10^9$ cells. Thus, if each cancer cell contains thousands of random mutations, there could be as many as $10^{15}$ random mutations within the tumor cell population. Thus, within a tumor there would be multiple cells that contained mutations in genes or regulatory sequences that rendered those cells resistant to specific chemotherapies. In the presence of chemotherapy, the mutant cells would have a growth advantage. Accordingly, treatment with chemotherapy select resistant mutants, and genetic instability exacerbates this problem. It is advisable to treat with non-cross resistant therapeutic combinations to reduce the probability of mutation to resistance to all of the applied therapies [90]. These combinations can either be applied together (at reduced dose to keep the overall toxicity tolerable) or sequentially at high dose, as has been argued based on growth kinetic arguments [91].

9. Could the clinical appearance of cancer be prevented by decreasing the rate of genetic change?

Cancer typically occurs late in life and evolves over decades. In that regard, even a modest reduction in the rate of carcinogenesis could delay the onset of cancer by decades, perhaps even placing the projected time of cancer development outside the human lifespan. Given the role of mutation in carcinogenesis, enhanced genetic stability could slow carcinogenesis in a meaningful way, leading to "prevention by delay".

Even a two-fold reduction in the rate of accumulation of mutations in cancer cells might have profound effects on the age at which patients succumb to certain adult cancers. Prevention by delay may be particularly applicable to cancers associated with prolonged chronic inflammation due to either bacterial infection (gastric cancer) [92], viral infection (hepatitis B or C) [93], immunological insufficiencies (ulcerative colitis) [94], or unknown etiology (chronic pancreatitis) [95]. In the case of primary hepatoma resulting from hepatitis virus B, infection occurs in early infancy and in some individuals persists chronically. It takes 40 years from the start of infection to tumor detection and hepatitis increases the risk of subsequent hepatoma by more than 200-fold. A reduction in the rate of mutation accumulation by only two-fold could delay the clinical appearance of the tumor from age 50 to age 90. A similar analysis pertains to hepatoma associated with persistent infection with hepatitis C virus. Machida et al. comment on the induction of a mutator phenotype after infection with hepatitis C virus (HCV) both in vitro and in vivo [96]. They observe a 5- to 10-fold elevation in the mutation frequency at different loci, including immunoglobulin heavy chain, BCL-6, p53, and β-catenin. It has been postulated that the persistence of viral hepatitis results in an inflammatory reaction with generation of oxygen reactive species which could lead to DNA damage and mutagenesis. Thus, drugs that scavenge oxygen free radicals might have a place in the prevention of primary hepatoma.
Caloric restriction reduces tumor incidence in laboratory animals [97], and also reduces the mutation rate, possibly by reducing the concentration of the pyrophosphate anion, might enhance proofreading by DNA polymerases [18], leading to a decreased mutation rate, an idea that could be tested in cell culture.

10. Conclusion

Genetic instability has a likely role in cancer pathogenesis as well as a possible role in sensitivity and resistance to therapy. The field has witnessed synergistic contributions of theory and experiment supporting the role of genetic instability at multiple levels including epidemiology, genetics, molecular biology, and clinical oncology. This trend is likely to continue in the future as the remaining questions are addressed.

Acknowledgements

We thank Ms. Jessica Hsu for assistance in preparing figures and Ms. Mary Whiting for quality control assistance. This work was supported in part by grants to LAL from the U.S. National Cancer Institute CA 102029.

References


