Available online at www.sciencedirect.com
SCIENCE DIRECT®
Journal of Theoretical Biology 241 (2006) 216–222
Journal of Theoretical Biology
www.elsevier.com/locate/yjtbi

Genetic instability and the quasispecies model

Yisroel Brumer[a,*], Franziska Michor[b], Eugene I. Shakhnovich[a]

[a]Department of Chemistry and Chemical Biology, Harvard University, Cambridge, MA 02138, USA
[b]Harvard Society of Fellows, Program for Evolutionary Biology, Harvard University, Cambridge, MA 02138, USA

Received 14 March 2005; received in revised form 9 November 2005; accepted 15 November 2005
Available online 4 January 2006

Abstract

Genetic instability is a defining characteristic of cancers. Microsatellite instability (MIN) leads to by elevated point mutation rates, whereas chromosomal instability (CIN) refers to increased rates of losing or gaining whole chromosomes or parts of chromosomes during cell division. CIN and MIN are, in general, mutually exclusive. The quasispecies model is a very successful theoretical framework for the study of evolution at high mutation rates. It predicts the existence of an experimentally verified error catastrophe. This catastrophe occurs when the mutation rates exceed a threshold value, the error threshold, above which replicative infidelity is incompatible with cell survival. We analyse the semiconservative quasispecies model of both MIN and CIN tumors. We consider the role of post-methylation DNA repair in tumor cells and demonstrate that DNA repair is fundamental to the nature of the error catastrophe in both types of tumors. We find that CIN introduces a plateau in the maximum viable mutation rate for a repair-free model, which does not exist in the case of MIN. This provides a plausible explanation for the mutual exclusivity of CIN and MIN.

Keywords: Genetic instability; Quasispecies model; Mathematical biology

1. Introduction

Genetic instability is a hallmark of human cancers (Lengauer et al., 1998; Loeb, 2001), and two main types have been identified. Microsatellite instability (MIN) refers to subtle sequence changes that alter one or a few base pairs (Kinzler and Vogelstein, 1996; Perucho, 1996). MIN is caused by a deficiency of the mismatch repair (MMR) pathway, and six human genes are known that, when recessively inactivated lead to a MIN phenotype in cancer patients. MIN, however, is fairly uncommon in human cancers and is only found in a small fraction of colorectal, endometrial and gastric cancers.

The majority of human cancers have chromosomal instability (CIN) (Rajagopalan et al., 2003). CIN refers to an increased rate of losing or gaining whole chromosomes or parts of chromosomes during cell division. The consequence of CIN is an imbalance in chromosome number (aneuploidy) and an increased rate of loss of heterozygosity. A large number of genetic alterations can trigger CIN in yeast (Kolodner et al., 2002), but so far, only a few genetic causes of CIN have been identified in humans. These so-called 'CIN genes' include MAD2, hBUB1, hCDC4, and BRCA2 (Michor et al., 2004).

CIN and MIN are generally mutually exclusive (Lengauer et al., 1998). MIN cancers are diploid and exhibit normal rates of gross chromosomal change, whereas CIN cancers are usually aneuploid and exhibit increased rates of chromosomal change, but have normal point mutation rates.

Mathematical modeling of genetic instability has led to considerable insight into human tumorigenesis. Nowak and his group used stochastic processes to determine the role of CIN in colorectal tumor initiation (Michor et al., 2004; Nowak et al., 2002). They found that, in dependence of tissue organization and the number of CIN genes in an organism, CIN is very likely to initiate tumorigenesis (Michor et al., 2003, 2004). Little and Wright (2003) used the multi-stage stochastic model of carcinogenesis (Armitage and Doll, 1954) to describe colorectal tumorigenesis with genetic instability, finding that a model with five

*Corresponding author.
E-mail address: brumer@post.harvard.edu (Y. Brumer).

0022-5193/$ - see front matter © 2005 Published by Elsevier Ltd.
doi:10.1016/j.jtbi.2005.11.018

stages and two levels of genomic destabilization fits colon cancer incidence data. Breivik and Gaudernack (2004) analysed genetic instability from the perspective of molecular evolution and information processing. They presented a mathematical model that predicts loss of genetic stability in environments where the evolutionary cost of DNA repair exceeds the cost of errors.

A particularly useful model for the study of evolution at high mutation rates is the quasispecies model (Eigen, 1971; Eigen et al., 1989). A quasispecies is a "cloud" of genetically related genomes. The quasispecies model is based on a phenomenological description of an explicit population of genomes and incorporates a fitness landscape, i.e. the assignment of reproductive fitnesses to specific genomes. The model has provided an impressive number of experimentally verified predictions, ranging from the existence of an error catastrophe to a quantitatively accurate prediction of human B-cell mutation rates (Kamp and Bornholdt, 2002) and novel anti-viral therapies (Crotty et al., 2001; Loeb et al., 1999).

In 1989, Nowak and Schuster (1989) investigated error thresholds in finite populations. They determined that, at error rates above the critical value, the quasispecies ceases to be localized in sequence space and starts to drift randomly. Solé and Deisboeck (2004) used the quasispecies model to investigate the error threshold in cancer cells. They demonstrated that, once the threshold is reached, the highly unstable cancer cells become unable to maintain their genetic information, leading to a decrease in the velocity of tumor growth. Recently, Brumer and Shakhnovich (2004) demonstrated that incorporation of semiconservative replication into the quasispecies model (Tannenbaum et al., 2004) presents a paradox in tumor progression, discussed in detail below.

The original quasispecies model assumes that genomes replicate conservatively, i.e. each single-stranded genome replicates by producing a new, possibly error-prone, single-stranded copy without affecting the original. In this form, the quasispecies model predicts the existence of an error catastrophe or "error threshold", a threshold mutation rate above which no viable species can exist. This threshold depends on the replication rate of the fittest sequence, the master sequence. In the commonly used single fitness peak landscape, the threshold mutation rate increases indefinitely with master sequence fitness. Qualitatively, this occurs because no information is lost upon conservative replication. Although a perfect copy is rarely created, the viable genome remains in the population as long as the replication rate is high enough. Cancer cells, for example, replicate very fast, thus allowing for the high mutation rates they exhibit without passing the error threshold (Solé, 2003).

The conservative model, however, is applicable only to RNA genomes. In contrast, DNA genomes replicate semiconservatively: each double-stranded genome replicates by unzipping and producing a complementary copy of each single strand (Fig. 1). Semiconservative replication

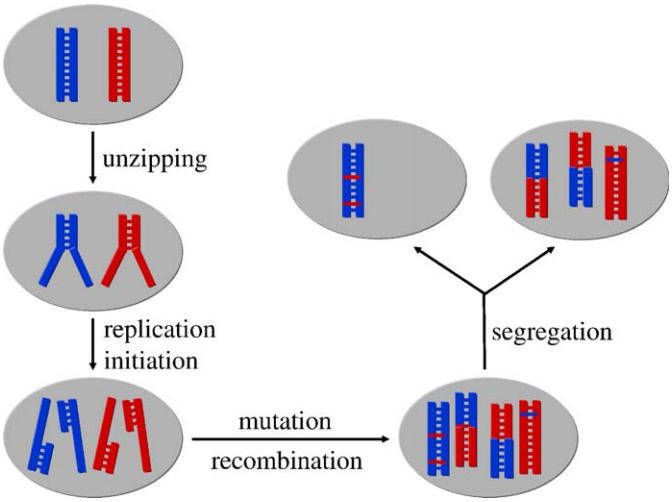

Fig. 1. Quasispecies replication. Double-stranded chromosomes unzip and initiate replication. Mutation and recombination cause deviation of the sequence from the consensus (master) sequence. Here, point mutations and reciprocal translocation are shown. During segregation, sister chromosomes might not be partitioned perfectly into the two daughter cells. The chromosomal instability phenotype increases the probability of segregation errors.

drastically alters the behavior of the system (Tannenbaum et al., 2004). The threshold mutation rate plateaus at a low value of the master sequence fitness and never increases above a low error rate (Fig. 2a). For the conservative system, there exists a master sequence fitness for any given mutation rate such that the quasispecies survives. For the semiconservative system, however, this is not true. Mutation rates above the plateau will cause the error catastrophe *independent of the master sequence fitness*. The existence of this plateau can be understood best by considering the nature of error repair in semiconservative systems. Post-methylation repair yields a non-zero chance that a master sequence will be changed to a sequence of lower fitness upon replication. The chance of this occurring increases with increasing replication rate. Thus, master sequences can be lost through replication. In the conservative model, master sequences can be overwhelmed by the creation of sequences with lower fitness, but replication never affects the original sequence. Thus, in the conservative case, a master sequence can always "out-replicate" the error rate: that is, with a high enough replication rate, the master sequence can produce enough copies that a finite number of new master sequences are produced, no matter what the error rate. However, this is not true in the semiconservative case. For high error rates, a higher replication rate may not lead to any new master sequences, as the original master may be destroyed in the process. This means that there are error rates which are past the error threshold for all values of the master sequence fitness, resulting in the plateau.

This plateau creates a paradox in cancer models (Brumer and Shakhnovich, 2004), as cancer cells routinely display mutation rates that far exceed any reasonable estimate of the error threshold plateau in semiconservative systems.

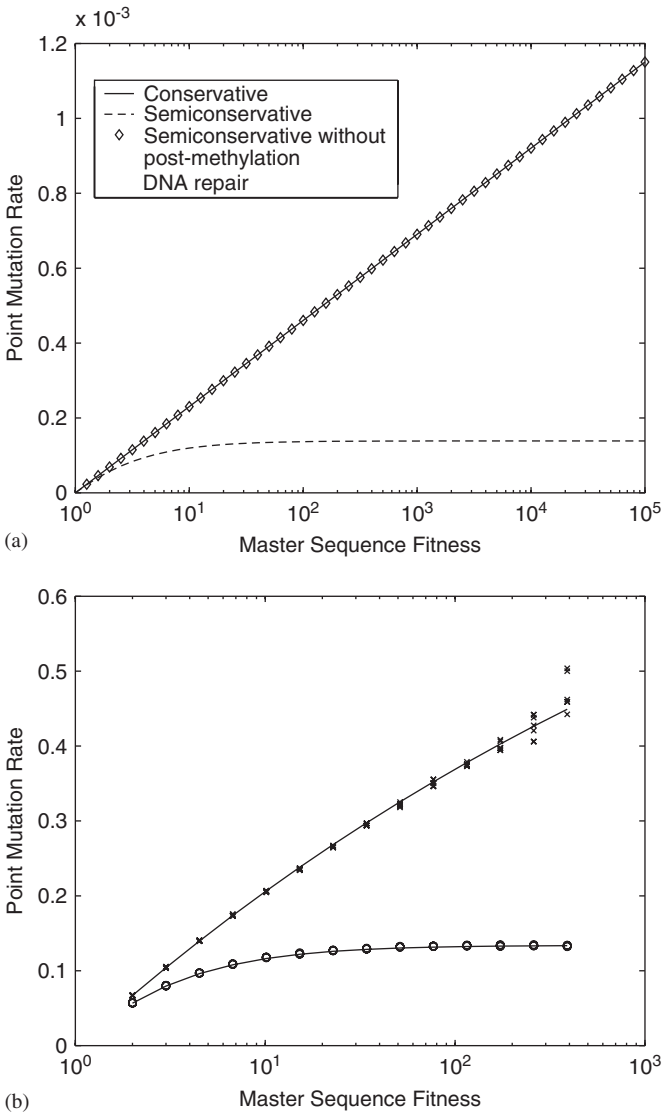

Fig. 2. (a) The value of the error threshold vs. the fitness of the master sequence relative to the population on a single fitness peak landscape. The genome length is $L = 1 \times 10^4$. Conservative, semiconservative and semiconservative systems without post-methylation lesion repair are shown. (b) Analytical solution compared directly to the stochastic simulations. Parameter values are $L = 10$ and $N = 10\,000$. The 'x's correspond to semiconservative runs without DNA repair, and 'o's correspond to perfect DNA repair. Note that, as σ approaches the scales associated with the population size, the stochastic error begins to increase.

The rapid replication rates cannot resolve the paradox, as increasing the replication rate does not increase the maximum viable mutation rate. Failure of post-methylation DNA repair, however, can resolve the paradox (Brumer and Shakhnovich, 2004). For a specific single fitness peak landscape, DNA genomes mimic their RNA counterparts as post-methylation DNA repair begins to fail. Once repair fails completely, the DNA genome becomes identical to a conservatively replicating RNA genome on a transformed landscape. Hence, removing post-methylation DNA repair causes DNA genomes to display, qualitatively, the error catastrophe behavior associated with RNA genomes.

In this paper, we analyse both MIN and CIN tumors in the semiconservative quasispecies model. We use finite population sizes and finite genome lengths and investigate the effect of post-methylation DNA repair on the error catastrophe. We demonstrate that the qualitative behavior of CIN tumors mimics that of their MIN counterparts in that degradation of lesion repair leads to an increase in the error threshold. We demonstrate that novel CIN dynamics can be used to provide a plausible explanation for the mutual exclusivity of CIN and MIN.

2. Microsatellite instability

MIN refers to increased point mutation rates due to MMR deficiency. MIN tumors can conveniently be described by the quasispecies model.

The original quasispecies model uses a set of differential equations to describe the changing concentrations of all possible genomes in a population (Eigen, 1971; Eigen et al., 1989). Each genome ϕ consists of a set of letters $\phi = s_1 s_2 \ldots s_L$, where each letter is chosen from an alphabet of size S. Here S is four to mimic the nucleotides. Define $A \equiv 1, G \equiv 2, T \equiv 3, C \equiv 4$. Each possible string of letters is assigned a replication rate (fitness) and the population size is held constant, yielding a set of differential equations shown in the Appendix (Eigen et al., 1989). In the semiconservative form, we deal with double-stranded genomes $\{\phi, \phi'\}$ and their corresponding equations (Tannenbaum et al., 2004). To model MIN, we consider varying rates of point mutations and assume that the probability of point mutations is base pair and genome independent. We study the single fitness peak landscape; we define a single-stranded master sequence, ϕ_0, such that the fitness of a double-stranded genome ϕ is $A(\phi) = \sigma \gg 1$ if both strands are equivalent to ϕ_0 or its perfect complement $\overline{\phi_0}$, and $A(\phi) = 1$ otherwise. The single fitness peak model has been shown to accurately capture the local dynamics of genomic evolution about a fitness peak, and yields the same qualitative behavior as some more delocalized landscapes. Insisting that at least one strand be perfect is a reasonable model for essential housekeeping genes. Other landscapes will be treated in future work (Brumer et al., 2004). For more details on the quasispecies model, we refer the interested reader to the excellent review by Eigen (Eigen et al., 1989) and the original work on the semiconservative quasispecies (Tannenbaum et al., 2004).

The system can be treated using stochastic simulations with the following procedure. A set of N double-stranded genomes of length L are initialized. In our case, all are equal to the perfect double-stranded genome $\{\phi_0, \overline{\phi_0}\}$. Below the error threshold, there are two stable equilibrium conditions, a random set of genomes, and a quasispecies surrounding the most probable sequence. Above the error threshold, the second equilibrium condition disappears. By employing an initial population of viable genomes, we

approach the equilibrium from above (i.e. from higher values of the concentration of master sequences) and thus lessen the probability that we will locate the random set of genomes (the first equilibrium condition) below the error threshold by random fluctuation. This is fundamental to allowing an accurate determination of the error threshold. Using a time step small enough to ensure convergence of the computational results (in our case, $\Delta t = 0.33/A(\{\phi_0, \overline{\phi_0}\})$), we propagate the system using the following set of steps. For each time step, each double-stranded genome $\{\phi, \phi'\}$ in the population reproduces with probability $A(\{\phi, \phi'\})\Delta t$. When a double-stranded genome replicates, it unzips and two new complementary strands are synthesized. Each base in the new strand is correctly replicated with probability $1-\varepsilon$, where ε represents the point mutation rate. Methyl-directed MMR would, in nature, repair many of these errors, but this simply rescales the mutation rate ε. In a system with perfect post-methylation DNA repair, global genomic repair and transcription coupled repair find and repair base mismatches caused by transcription errors. However, these enzymes cannot distinguish the new strand from the old strand. Thus, the error is fixed correctly (replacing the erroneous base on the new strand) with 50% probability and incorrectly (replacing the correct base on the old strand to create a base pair) with 50% probability. For more details on this model of semiconservative replication, see Tannenbaum et al. (2004). Since we are interested in the failure of post-methylation DNA repair, we introduce a parameter, λ, representing the efficiency of repair. An error is repaired, either correctly or incorrectly, with probability λ. Hence, a base pair mismatch remains with probability $1-\lambda$. Note that by allowing unrepaired errors, double-stranded genomes no longer have to be perfectly complementary. Lastly, N genomes are chosen to survive to the next step, thus keeping the population size constant; this synchronous updating of generations is known as Wright–Fisher process in population genetics (Wright, 1931; Fisher, 1930). The steps are iterated to create a trajectory in concentration space.

3. Chromosomal instability

CIN refers to increased rates of losing or gaining chromosomes during cell division. We denote the probability that a CIN error occurs per replication by P_C, experimentally determined to be on the order of 0.01 per cell division (Lengauer et al., 1998). Our model describes cells with n chromosomes each. During each cell division with a CIN error, the n chromosomes unzip to form $2n$ single strands. We use the definition of recombination commonly adopted for the quasispecies model (Boerlijst et al., 1996; Barnett, 2003), dubbed uniform crossover. Each base pair in a new strand is chosen with equal probability from one of the $2n$ separated strands. The direction of the strand, however, is kept the same. Thus, each new chromosome will on average contain $1/(2n)$ of the base pairs of any one of the original strands. Lastly, the $2n$ new double-stranded chromosomes are randomly distributed to the daughter cells, producing two cells with x and y chromosomes, respectively; for conservation, $x+y=2n$. This process is shown schematically in Fig. 1. The replication rate of a diploid cell with one or two master chromosomes is $\sigma \gg 1$, the fitness of diploid cells without a master chromosome as well as cells with 1, 3, 4, 5, and 6 chromosomes is 1, and the fitness of cells with 0 or more than 6 chromosomes is 0.

For each system and for a given set of σ, ε and λ, the system is run until equilibrated, and the error threshold is defined as the mutation rate at which the equilibrium population included at most one master genome. A genome length of $L=10$ is sufficient to nearly converge to the infinite sequence results, and $N=10\,000$ is large enough to obtain excellent convergence for reasonable values of σ (obviously, as σ approaches N, finite size effects become prominent). Larger systems and longer genomes were run to confirm that the qualitative trends are robust.

4. Results and discussion

Fig. 2a shows the analytical solution for the value of the error threshold in a quasispecies model that replicates conservatively, semiconservatively, and semiconservatively without post-methylation DNA-repair as a function of the master sequence fitness. While the threshold in the semiconservative system plateaus at a low value of σ, the conservative threshold, mimicked exactly by the repair-free semiconservative system, increases indefinitely with increasing σ. This has a number of important implications, most pertinently the hypothesis that degradation of post-methylation DNA repair is fundamental to the survival of a MIN tumor (Brumer and Shakhnovich, 2004). The analytical solution leading to this hypothesis requires a number of approximations, including neglecting mutations to the master sequence, and infinite population size and genome length. Fig. 2b presents numerical results for these systems, free of the above approximations. Finite populations and genomes, as well as the inclusion of back mutation, have essentially no effect on the results. Further, these results validate the use of small genome lengths to approximate the very large genomes found in nature.

Fig. 3a shows the results of numerical simulation of the CIN system with the biologically motivated $P_C = 0.01$, and Fig. 3b shows the results for $P_C = 0.1$ to demonstrate the robustness of the results. Both figures include the results of an analytical calculation (appendix) obtained by making a few simple approximations. Note the excellent agreement between theory and simulation. The CIN model retains one fundamental trend found for MIN tumors: the error threshold in the semiconservative system plateaus rapidly, and when post-methylation DNA repair is removed, this threshold is significantly increased (for $P_C = 0.1$, this increase is about a factor of 2, while the more biological $P_C = 0.01$ shows an increase greater than four-fold). Fig. 4

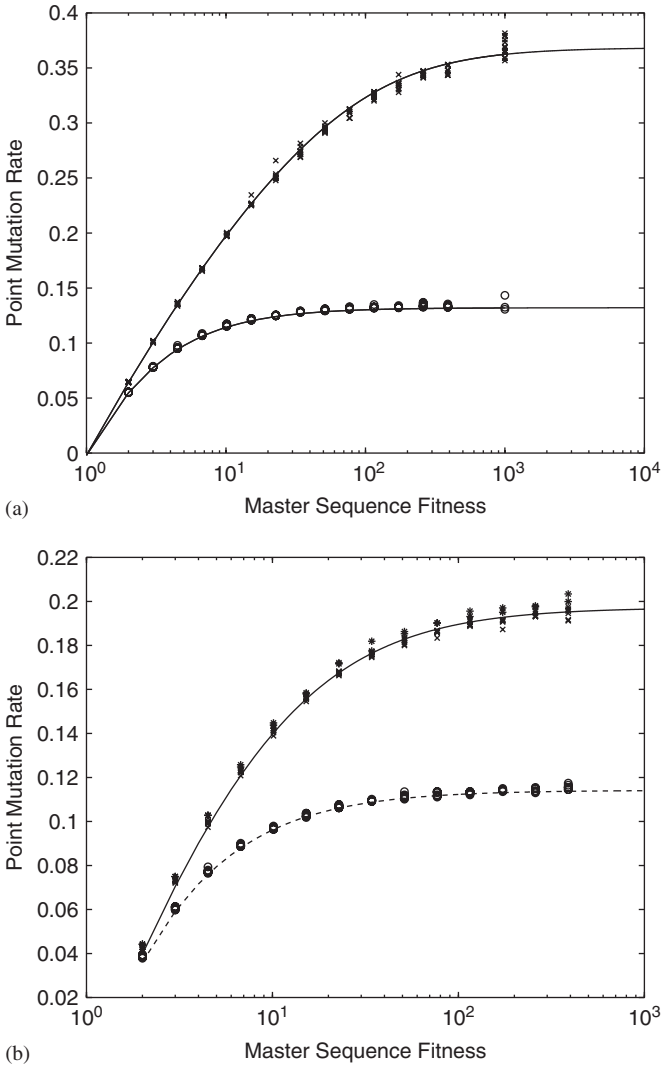

Fig. 3. (a) The error threshold vs. master sequence fitness for the chromosomal instability model. Parameter values are $P_C = 0.01$, $L = 10$, and $N = 10\,000$. As before, the 'x's represent semiconservative runs without DNA repair and the 'o's represent perfect DNA repair. (b) Same as (a), but for $P_C = 0.1$. The '*'s are the same as the 'x's, but with population size $N = 50\,000$.

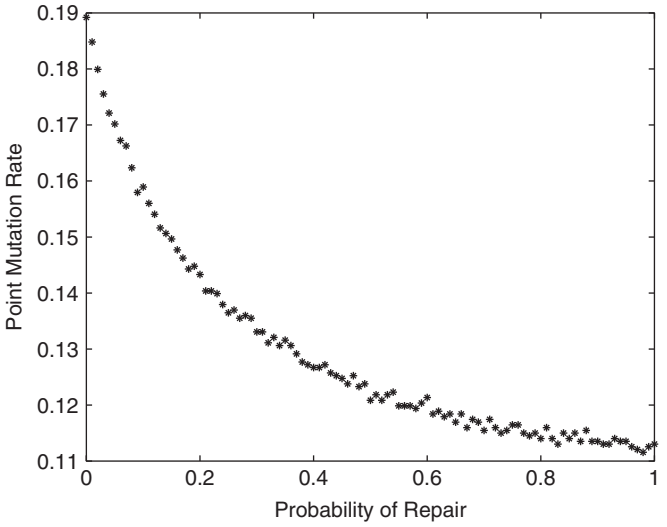

Fig. 4. The error threshold vs. the probability of repair for master sequence fitness $\sigma = 100$. Parameter values are $P_C = 0.1$, $L = 10$, and $N = 10\,000$.

demonstrates that the error threshold increases smoothly from the semiconservative to the conservative threshold as repair is removed (i.e. as the probability of repair goes from 1 to 0). Similar behavior was postulated to be fundamental for MIN tumorigenesis, as slight damage to the repair mechanism increases the error threshold slightly, allowing for higher mutation rates, which makes further damage to repair more likely (Brumer and Shakhnovich, 2004). For a given point mutation rate, the removal of lesion repair allows a higher level of CIN. Hence, while failure of post-methylation DNA repair is postulated to be a pre-requisite for MIN tumor progression, it is shown here to greatly facilitate CIN tumor progression, but without the restriction that tumor progression cannot possibly occur without repair failure. This additional requirement may partially explain why MIN tumors are so much rarer than CIN in nature.

A fundamental difference between MIN and CIN tumors sheds light on an important biological problem. The main difference between the CIN and MIN tumors lies in the existence of a plateau in the maximum viable error catastrophe for CIN tumors *for all values of* λ. As discussed above, this plateau is caused by the possible loss of master sequences in replication. In MIN tumors, the plateau disappears as $\lambda \to 0$. The agreement between the analytical results as well as two different system sizes (shown in Fig. 2b) confirm that this plateau is not a numerical artifact or finite size effect. The behavior is also qualitatively explicable, as the source of the plateau in the semiconservative system is the fact that replication can lead to destruction of any initially perfect strand, a trait that is absent in conservative replication and is lost when lesion repair becomes inactive. However, the introduction of CIN reinstates this property, independent of the existence of lesion repair, as a chromosomal error can destroy perfect strands upon replication.

These results can be used to provide a plausible explanation for the mutual exclusivity of MIN and CIN tumors in nature. While removal of lesion repair allows for arbitrarily high levels of point mutations without crossing the error threshold in a MIN tumor, the plateau present in CIN tumors makes these high point mutation rates incompatible with cell viability. Thus, a MIN cell can survive with extremely high rates of point mutations (up to 1000 times that of normal cells), as long as the cell replicates fast enough, i.e. as long as the master sequence fitness is sufficiently high. However, any quantitative estimate of the error threshold would state that these MIN point mutation rates are far above the error threshold for CIN tumors, *no matter how fast the cell replicates*. For

example, when $P_C = 0.01$, the results in the Appendix suggest a maximum viable mutation rate of 4.595 errors/genome/replication. MIN cells can display error rates two or three orders of magnitude greater than this. Thus, our results suggest that, while CIN and MIN tumors are independently stable, cells with both CIN and MIN instability are inviable. This is in agreement with experimental evidence, which suggests that CIN and MIN are generally mutually exclusive in vivo (Lengauer et al., 1998).

Lastly, we suggest that these theoretical results can be tested directly by experiment. Recently, it was shown that CIN and MIN can be introduced in cancerous cell lines through specific mutagenesis (Bardelli et al., 2001). While an in-depth study of dynamics above the error threshold along with careful consideration of the enzymatic interactions, both subjects of future research, would be necessary to rigorously quantify this statement, preliminary results would suggest that the introduction of CIN into MIN cell lines would select, for a given level of CIN, for the MIN cells with the lowest levels of point mutations and post-methylation DNA repair. This is one example of a quantifiable and testable hypothesis that can be used to experimentally test our theoretical work.

5. Conclusion

In this paper, we use the semiconservative quasispecies model to analyse tumors with chromosomal instability (CIN) and microsatellite instability (MIN). We demonstrate that MIN and CIN tumors both display increased error thresholds upon removal of post-methylation DNA repair. However, the CIN tumors show a plateau in the maximum viable mutation rates for all values of repair efficiency. MIN tumors, in contrast, show a disappearing plateau as the repair efficiency goes to zero. This is a fundamental difference, as the lack of a plateau allows MIN tumors to retain viability with enormously high point mutation rates, as is found in nature. Any CIN tumor with such a high mutation rate, however, will immediately cross the error threshold and become inviable. We conclude that, while CIN and MIN tumors are individually viable, a cell cannot contain both CIN and MIN and survive. This finding agrees well with experimental work in the area, and provides a plausible explanation for a fundamental phenomenon.

Acknowledgements

The authors would like to thank Martin Nowak for helpful discussions and for looking over the manuscript. This research was supported by an NIH postdoctoral fellowship to Y.B.

Appendix. Quasispecies model with chromosomal instability

The original quasispecies model employs a set of differential equations to model the dynamic concentrations of all possible genomes in a population. Each genome ϕ is explicitly described by a string of letters $\phi = s_1 s_2 \cdots s_L$, where each letter is chosen from an alphabet of size S. We choose S to be four to mimic the nucleotides (with the arbitrary definitions $A \equiv 1, G \equiv 2, T \equiv 3, C \equiv 4$). Each possible string of letters is assigned a replication rate (fitness) and the population size is held constant, yielding the following set of differential equations (Eigen et al., 1989)

$$\frac{\mathrm{d}x_\phi}{\mathrm{d}t} = \sum_{\phi'} A(\phi')W(\phi,\phi')x_{\phi'} - f(t)x_\phi, \tag{A.1}$$

where x_ϕ denotes the population fraction of genome ϕ, $A(\phi)$ represents the fitness of sequence ϕ, and $W(\phi,\phi')$ is the likelihood of creating sequence ϕ from ϕ' by mutations. The average fitness of the population is given by $f(t) = \sum_\phi A(\phi)x_\phi$. This term holds the population size constant, thereby introducing competition. In the semi-conservative form, we deal with double-stranded genomes $\{\phi,\phi'\}$ and their corresponding equations (Tannenbaum et al., 2004). Although a full analytical treatment of the chromosomal instability quasispecies model, with recombination, aneuploidy and a cellular fitness landscape is beyond the scope of this work, we present a simple set of approximations that can be used to yield an approximate solution that is in excellent agreement with the numerical results of Section 4. We track the concentration of perfect chromosomes (rather than cells), defined as x_0. Further, we assume that the error catastrophe occurs when x_0, rather than the number of viable cells, goes to zero (this approximation becomes exact as the probability that a perfect chromosome is found in an aneuploid cell goes to zero). We also assume that any CIN error leads to inviable chromosomes, and neglect all point mutations and CIN errors that lead from an inviable to viable cell (which becomes exact in the limit of an infinite genome length, and is an excellent approximation for reasonable lengths). Although this may seem like a significant number of approximations, the end result is a tractable set of equations that are rigorously exact in certain limits, and very nearly exact for reasonable parameters. This yields the set of differential equations, when lesion repair is absent,

$$\frac{\mathrm{d}x_0}{\mathrm{d}t} = \sigma q^L(1-P_C)x_0 - f(t)x_0 - P_C x_0, \tag{A.2}$$

$$\frac{\mathrm{d}x_1}{\mathrm{d}t} = \sigma(1-q^L)x_0 + \sigma q^L P_C x_0 + x_1 - f(t)x_0, \tag{A.3}$$

$$f(t) = \sigma x_0 + 1 - x_0, \tag{A.4}$$

where $x_1 = 1 - x_0$ and $q = 1 - \varepsilon$. This yields two equilibrium solutions, $x_0 = 0$ or

$$x_0 = \frac{1 + P_C\sigma - \sigma q^L(1-P_C)}{\sigma - 1}. \tag{A.5}$$

The error catastrophe occurs when these solutions meet, when

$$\varepsilon = -\left(\frac{1+P_C\sigma}{\sigma(1-P_C)}\right)^{1/L} + 1. \tag{A.6}$$

This solution is plotted in Fig. 3. The plateau behavior can be seen by taking the limit $\sigma \to \infty$, which yields

$$q^L \equiv (1-\varepsilon)^L = \frac{P_C}{1-P_C}. \tag{A.7}$$

Thus, a plateau exists for *all* values of P_C. As well, when $P_C > 0.5$, the system is inviable for all values of the mutation rate, as replication makes the loss of information more likely than the gain thereof. When lesion repair remains intact, we obtain the equations

$$\frac{\mathrm{d}x_0}{\mathrm{d}t} = 2\left(1-\frac{\varepsilon}{2}\right)^L \sigma(1-P_C)x_0 - \sigma x_0 - f(t)x_0, \tag{A.8}$$

$$x_1 = 1 - x_0, \tag{A.9}$$

$$f(t) = \sigma x_0 + 1 - x_0 \tag{A.10}$$

which yields the solutions $x_0 = 0$ and

$$x_0 = \frac{1+\sigma-2(1-\varepsilon/2)^L\sigma(1-P_C)}{1-\sigma} \tag{A.11}$$

which meet when

$$\varepsilon = -2\left(\left(\frac{1+\sigma}{2\sigma(1-P_C)}\right)^{1/L} - 1\right) \tag{A.12}$$

yielding, in the limit $\sigma \to \infty$,

$$\varepsilon = -2\left(\left(\frac{1}{2(1-P_C)}\right)^{1/L} - 1\right). \tag{A.13}$$

References

Armitage, P., Doll, R., 1954. Br. J. Cancer 8, 1.

Bardelli, A., Cahill, D.P., Lederer, G., Speicher, M.R., Kinzler, K.W., Vogelstein, B., Lengauer, C., 2001. Proc. Natl Aacd. Sci. 98, 5770.

Barnett, L., 2003. Ph.D. Thesis, School of Cognitive and Computing Sciences, University of Sussex.

Boerlijst, M.C., Bonhoeffer, S., Nowak, M.A., 1996. Proc. R. Soc. London B 263, 1577.

Breivik, J., Gaudernack, G., 2004. FEBS Lett. 563, 7.

Brumer, Y., Shakhnovich, E.I., 2004. Phys. Rev. E 70, 061912.

Brumer, Y., Tannenbaum, E., Shakhnovich, E.I., 2004. q-bio.BM/0406011.

Crotty, S., Cameron, C.E., Andino, R., 2001. Proc. Natl Acad. Sci. 98, 6895.

Eigen, M., 1971. Naturwissenschaften 58, 465.

Eigen, M., McCaskill, J., Schuster, P., 1989. Adv. Chem. Phys. 75, 149.

Fisher, R.A., 1930. The Genetical Theory of Natural Selection. Clarendon Press, Oxford.

Kamp, C., Bornholdt, S., 2002. Phys. Rev. Lett. 88, 068104.

Kinzler, K.W., Vogelstein, B., 1996. Cell 87, 159.

Kolodner, R.D., Putnam, C.D., Myung, K., 2002. Science 297, 552.

Lengauer, C., Kinzler, K.W., Vogelstein, B., 1998. Nature 396, 643.

Loeb, L.A., 2001. Cancer Res. 61, 3230.

Little, M.P., Wright, E.G., 2003. Math. Biosci. 183, 111.

Loeb, L.A., Essigmann, J.M., Kazazi, F., Zhang, J., Rose, K.D., Mullins, J.I., 1999. Proc. Natl Acad. Sci. 96, 1492.

Michor, F., Iwasa, Y., Komarova, N.L., Nowak, M.A., 2003. Curr. Biol. 13, 581.

Michor, F., Iwasa, Y., Nowak, M.A., 2004. Natl Rev. 4, 197.

Michor, F., Iwasa, Y., Rajagopalan, H., Lengauer, C., Nowak, M.A., 2004. Cell Cycle 3, 358.

Nowak, M.A., Komarova, N.L., Sengupta, A., Jallepalli, P.V., Shih, L., Vogelstein, B., Lengauer, C., 2002. Proc. Natl Acad. Sci. 99, 16226.

Nowak, M., Schuster, P., 1989. J. Theor. Biol. 137, 375.

Perucho, M., 1996. Biol. Chem. 377, 675.

Solé, R.V., 2003. Eur. Phys. J. B 35, 117.

Solé, R.V., Deisboeck, T.S., 2004. J. Theor. Biol. 228, 47.

Rajagopalan, H., Nowak, M.A., Vogelstein, B., Lengauer, C., 2003. Natl Rev. Can. 3, 695.

Tannenbaum, E., Deeds, E.J., Shakhnovich, E.I., 2004. Phys. Rev. E 69, 011902.

Wright, S., 1931. Genetics 16, 97.