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

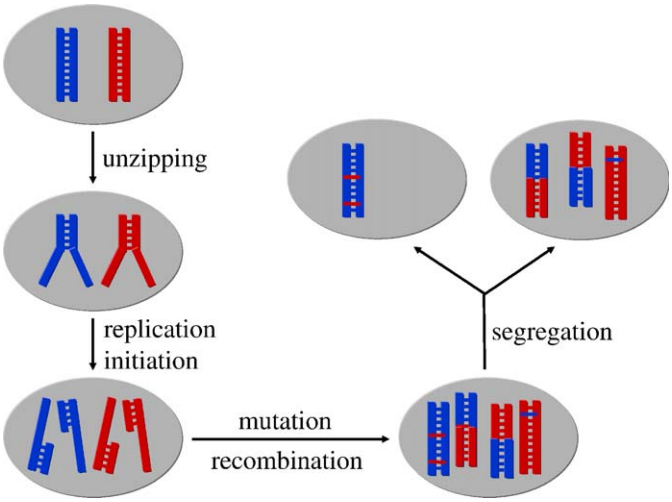

Fig. 1. Quasispecies replication. Double-stranded chromosomes unzip and initiate replication. Mutation and recombination cause deviation of the sequence from the consensus (master) sequence. Here, point mutations and reciprocal translocation are shown. During segregation, sister chromosomes might not 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.