Can chromosomal instability initiate tumorigenesis?

Franziska Michor\textsuperscript{a}, Yoh Iwasa\textsuperscript{b}, Bert Vogelstein\textsuperscript{c}, Christoph Lengauer\textsuperscript{c}, Martin A. Nowak\textsuperscript{a,}\textsuperscript{*}

\textsuperscript{a} Program for Evolutionary Dynamics, Department of Organismic and Evolutionary Biology, Department of Mathematics, Harvard University, Cambridge, MA 02138, USA
\textsuperscript{b} Department of Biology, Kyushu University, Fukuoka 812-8581, Japan
\textsuperscript{c} Sidney Kimmel Comprehensive Cancer Centre, Johns Hopkins University, Baltimore, MD 21231, USA

Abstract

Cancers result from the accumulation of inherited and somatic mutations in oncogenes and tumor suppressor genes. These genes encode proteins that function in growth regulatory and differentiation pathways. Mutations in those genes increase the net reproductive rate of cells. Chromosomal instability (CIN) is a feature of most human cancers. Mutations in CIN genes increase the rate at which whole chromosomes or large parts of chromosomes are lost or gained during cell division. CIN causes an imbalance in chromosome number (aneuploidy) and an enhanced rate of loss of heterozygosity, which is an important mechanism of inactivating tumor suppressor genes. A crucial question of cancer biology is whether CIN is an early event and thus a driving force of tumorigenesis. Here we discuss mathematical models of situations where inactivation of one or two tumor suppressor genes is required for tumorigenesis. If two tumor suppressor genes have to be inactivated in rate-limiting steps, then CIN is likely to emerge before the inactivation of the first tumor suppressor gene.

© 2004 Elsevier Ltd. All rights reserved.

Keywords: Somatic evolution of cancer; Mathematical model; Chromosomal instability; Tumor suppressor gene inactivation

Contents

1. Introduction ........................................................................................................................ 43
2. Chromosomal instability before one tumor suppressor gene ........................................ 44
3. Chromosomal instability before two tumor suppressor genes ....................................... 45
4. Conclusions .................................................................................................................... 47
5. Methods .......................................................................................................................... 47
Acknowledgements ............................................................................................................. 48
References ........................................................................................................................... 48

1. Introduction

Tumor suppressor genes (TSGs) are negative regulators of cellular growth [1–5]. Loss of their function contributes to tumorigenesis [6–8]. The concept of a TSG emerged from a statistical analysis of retinoblastoma incidence in children [6]. This study and subsequent work led to the 2-hit hypothesis, which proposes that two hits in the retinoblastoma gene are the rate limiting steps of this particular cancer [5]. In the inherited form, the first mutation is present in the germ line, whereas the second mutation emerges during somatic cell divisions. In the sporadic form, both mutations arise during somatic cell divisions. Inactivation of the first allele confers a fitness advantage to the cell and can promote neoplastic growth. Human cancers result from the accumulation of mutations in several TSGs, oncogenes, and genes that are involved in maintaining genomic stabil-
ity. Here, we discuss mathematical models of cancers that require the inactivation of one or two TSGs in rate limiting steps.

Genetic instability is a defining characteristic of most human cancers [9,10]. Two types of genetic instability have been identified [11]. In a small fraction of colorectal and some other cancers, a defect in mismatch repair results in an elevated mutation rate at the nucleotide level and consequent widespread microsatellite instability (MIN) [12,13]. The majority of colorectal and most other solid cancers, however, have chromosomal instability (CIN). CIN refers to an increased rate of losing or gaining whole chromosomes or large 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 (LOH). An elevated rate of LOH is an important property of CIN, because it accelerates the inactivation of TSGs.

A large number of genes that trigger CIN when mutated have been identified in the yeast *Saccharomyces cerevisiae* [14–16]. These so-called ‘CIN genes’ are involved in chromosome condensation, sister-chromatid cohesion, kinetochoore structure and function, and microtubule formation as well as in cell cycle checkpoints. By comparison with yeast, we expect several hundred human CIN genes, but only few have been identified so far [17]. These genes include hBUB1, MAD2, BRCA1, BRCA2 and hCDC4 [18–22]. The genes hBUB1 and MAD2 encode proteins that are required for function of the spindle assembly checkpoint [18,20]. This checkpoint modulates the timing of anaphase initiation in mitotic cells containing improperly aligned chromosomes and increases the probability of successful delivery of a correct chromosome set to each daughter cell. The genes BRCA1 and BRCA2 encode proteins that are implicated in DNA repair and recombination, checkpoint control of the cell cycle, and transcription [19,21]. hCDC4 is an E3 ubiquitin ligase that is thought to be involved in regulating the G1/S cell-cycle checkpoint by targeting proteins for destruction [22].

The classification of CIN genes is based on the mutational events required to engage CIN [10]. Class I CIN genes, such as MAD2, trigger CIN if one allele of the gene is mutated or lost. Class II CIN genes, such as hBUB1, trigger CIN if one allele is mutated in a dominant negative fashion. Both class I and class II CIN genes are ‘single hit’ CIN genes. Class III CIN genes, such as BRCA1 and BRCA2, trigger CIN if both alleles are mutated.

The mathematical investigation of cancer began in the 1960s, when Nordling [23], Armitage and Doll [24,25], and Fisher [26] set out to explain the age-dependent incidence curves of human cancers. These seminal studies led to the idea that multiple probabilistic events are required for the somatic evolution of cancer [27,28]. In the early 1980s, Knudson used a statistical analysis of the incidence of retinoblastoma in children to explain the role of tumor suppressor genes in sporadic and inherited cancers [6]. This work was later extended to a two-stage stochastic model for the process of cancer initiation and progression [29], which inspired much subsequent work [30–32]. Later on, specific theories for drug resistance [33,34], angiogenesis [35], immune responses against tumors [36], and genetic instabilities [37–40] were developed.

A major question in cancer genetics is to what extent CIN, or any genetic instability, is an early event and thus a driving force in tumorigenesis [41–47]. Here, we discuss mathematical models that investigate the role of CIN in cancers which require elimination of one or two TSGs [42,46–48].

2. Chromosomal instability before one tumor suppressor gene

Consider a path to cancer where both alleles of a TSG, A, have to be inactivated in a single cell (Fig. 1a) [42,46–48]. Initially, the cell is wild type, A+/+. The cell evolves from A+/+ via A+/− to A−/−. The first allele of a TSG is typically inactivated by a point mutation. The second allele can be inactivated by one of several possible mechanisms (Fig. 2). It can be inactivated by a second point mutation, leading to two distinct point mutations in the two alleles. Alternatively, the chromosome on which the second allele resides can be lost; this chromosome loss might be deleterious or lethal. The two alleles at the TSG locus can recombine mitotically, leading to two identical point mutations on partly homozygous alleles. Finally, the chromosome harboring the wild type allele can be lost and the chromosome harboring the first point mutation duplicated. This mechanism may require two steps, occurring in two distinct cell divisions, but could also occur in a single division through a process broadly related to mitotic recombination.

---

**Fig. 1.** Inactivation dynamics of a tumor suppressor gene (TSG). A, with and without chromosomal instability (CIN). (a) The wild type cell (‘start’) has two unmutated alleles of TSG A, A+/+. The cell evolves from A+/+ via A+/− to A−/−. CIN can arise at any stage of TSG inactivation and cancer very fast LOH. (b) A compartment of cells evolves along the evolutionary trajectory: If an A−/− cell clone with CIN produces an A−/− cell before taking over the compartment, a stochastic tunnel arises (diagonal arrow): the compartment evolves from A−/− without CIN to A−/− with CIN without ever visiting A+/− with CIN.
CIN can emerge at any stage of tumorigenesis. First, assume that CIN is dominant, i.e. triggered by mutation of a class I or II CIN gene [10,42]. The crucial effect of CIN is to increase the rate of LOH, thereby accelerating the transition from $A^{+/−}B^{+/−}$ to $A^{−−}B^{+/−}$. CIN can have a cost for the cell by increasing the rate of accumulating lethal mutations and triggering apoptosis. CIN, however, can also be advantageous by reducing the duration of the cell cycle (if certain checkpoints have been eliminated) or by providing increased evolvability in detrimental environments [49,50]. Therefore, it makes sense to consider possibilities of costly, neutral, and advantageous CIN phenotypes.

Tissues are organized into small compartments of cells [51–54]. Not all cells of a compartment might be at risk of becoming cancer cells, however, because certain mutations might only have an effect if they occur in stem cells [55]. Suppose $N_0$ cells are at risk in any one compartment. If the mutation rate is smaller than the inverse of $N_0$, then the approximation of homogenous compartments holds: a mutated cell will either take over the compartment or go extinct before the next mutation arises. If an $A^{+/−}$ cell clone with CIN produces an $A^{−−}$ cell before taking over the compartment, then the phenomenon of 'stochastic tunneling' arises [42,56]: the compartment moves from $A^{+/−}$ without CIN to $A^{−−}$ with CIN without ever visiting $A^{+/−}$ with CIN. If CIN is neutral or advantageous, then the tunnel does not arise. Mathematical procedures are outlined in Section 5.

3. Chromosomal instability before two tumor suppressor genes

Consider a path to cancer where two TSGs, $A$ and $B$, have to be inactivated in a single cell (Fig. 3a) [42,46–48]. Initially, the cell is wild type, $A^{+/−}B^{+/−}$. Suppose gene $A$ has to be inactivated first. Hence the cell evolves from $A^{+/−}B^{+/−}$ via $A^{+/−}B^{+/−}$ to $A^{+/−}B^{+/−}$, and subsequently to $A^{−−}B^{+/−}$ and $A^{−−}B^{+/−}$ CIN can emerge at any stage of tumorigenesis due to mutations of class I, II or III CIN genes. Once CIN has emerged, it accelerates the transitions from $A^{+/−}$ to $A^{−−}$ and from $B^{+/−}$ to $B^{+/−}$.

Evolutionary dynamics within a compartment of $N_0$ cells are illustrated in Fig. 3b. CIN might emerge before gene $A$ has been inactivated. A stochastic tunnel arises if an $A^{+/−}B^{+/−}$ cell clone with CIN produces an $A^{−−}B^{+/−}$ cell before taking over the compartment [42,56]. Then the compartment moves from $A^{+/−}B^{+/−}$ without CIN to $A^{−−}B^{+/−}$ with CIN without ever visiting $A^{+/−}B^{+/−}$ with CIN.

Inactivation of the first TSG can induce neoplastic growth. We assume that the $A^{−−}$ compartment gives rise to a small lesion of $N_1$ cells. In this lesion the second TSG has to be inactivated for further tumor progression. Due to the increased compartment size, the evolutionary pathway might tunnel from $A^{−−}B^{+/−}$ directly to $A^{+/−}B^{−−}$. This means that the $A^{+/−}B^{−−}$ cell does not reach fixation before the $A^{−−}B^{+/−}$ cell arises. Mathematical procedures are outlined in Section 5.

The importance of early CIN in tumorigenesis depends on the number of possible CIN mutations, which in turn depends on the number of CIN genes in the human genome. We can calculate the minimum number of CIN genes in the genome that are needed to ensure that a CIN mutation precedes the inactivation of one or two TSGs. Let us discuss some plausible parameter choices (Table 1). The mutation rate per base per cell division is $10^{−10}$ to $10^{−11}$ [57]. If an average TSG can be inactivated by any one of 1000 point mutations, the mutation rate per gene per cell division is $u = 10^{−7}$. Estimates of the rate of LOH in non-CIN cells range from $p_{B} = 10^{−7}$ to $p_{B} = 10^{−3}$. In our opinion, the most likely scenario is that $p_{B}$ has the same order of magnitude as $u$. In this case and in the absence of CIN, the hit inactivating the second TSG allele is sometimes LOH and sometimes a point mutation. If $p_{B} > u$, then two distinct point mutations should never be observed in the two TSG alleles. The rate of LOH in CIN cells is $p = 10^{−2}$ [58]. Consider a compartment size of $N_0 = 4$. This choice is motivated by the geometry of the colonic crypt [54]. Suppose the size of the lesion after clonal expansion of an $A^{−−}$ cell clone is of the order of $N_1 = 10^4$ actively dividing cells. CIN can be disadvantageous, neutral, or advantageous [17,49]. The most substantial cost CIN can possibly have is $r = (1 − p_{B})^{10} \approx 0.6$; this means that loss of any chromosome other than the one containing the CIN locus is lethal. CIN is neutral if $r = 1$ and advantageous if $r > 1$. Given these parameter choices and $r = 0.6$, we find that 3 class I CIN genes, or 29 class II CIN genes, or more than 100 class III CIN genes are required to ensure that CIN precedes inactivation of the first TSG in a pathway where no other TSG has to be eliminated in a rate limiting step. If two TSGs have to be inactivated in rate limiting steps, then 1 class I CIN gene, or 1 class II CIN gene, or 30 class III CIN genes are sufficient for CIN to precede inactivation of the first
Fig. 3. Mutational network of inactivating two tumor suppressor genes (TSGs), A and B, with and without chromosomal instability (CIN). (a) The wild type cell ('start') has two unmutated alleles of both TSGs, \( A^{++}B^{++} \). Mutations lead via \( A^{+-}B^{++} \), \( A^{--}B^{++} \) and \( A^{--}B^{--} \) to \( A^{--}B^{--} \). Without CIN, the cell maintains a stable genotype (second row). CIN can arise at any stage of tumorigenesis due to mutations in a class I, II, or III CIN gene. Classes I and II CIN genes require one hit to trigger CIN (first row). Class III CIN genes require two hits to trigger CIN, one in each allele (third and forth row). Once arisen, CIN accelerates the inactivation of the second allele of each following TSG. CIN, however, has a cost for the cell by increasing the chance of lethal mutations and apoptosis. (b) Inactivation dynamics starting with a population of \( N_0 \) wild type cells. CIN can emerge at any stage of tumorigenesis. If an \( A^{+-}B^{++} \) cell clone without CIN produces an \( A^{--}B^{++} \) cell with CIN before reaching fixation, the compartment tunnels to \( A^{--}B^{--} \) with CIN (red arrows). Inactivation of TSG A leads to a small lesion of \( N_1 \) cells in which TSG B is eliminated. Due to the increased compartment size, the evolutionary pathway tunnels directly to \( A^{--}B^{--} \) (red arrows).

Table 1 provides further examples for other parameter choices.

The parameters have different effects on the importance of early CIN. An increase in \( N_0 \), \( N_1 \) and \( p_0 \) as well as a decrease in \( r \) and \( p \) make the early emergence of CIN less likely. Instead of calculating the inactivation dynamics of one specific TSG, we can also study the possibility that one gene out of a family of TSGs can be inactivated. Suppose gene A is any one of ten TSGs and likewise gene B is any one of ten TSGs. This is an extreme example. In this scenario and using the same parameter values as above, the same number of CIN genes in the human genome as before is required to ensure that CIN precedes inactivation of the first TSG in a pathway where no other TSG has to be eliminated in a rate limiting step. If two TSGs have to be inactivated in rate limiting steps, then 1 class I CIN gene, or 5 class II CIN gene, or 62 class
clonal expansion by in (c); and p

Class III CIN genes, expansion (is especially strong if inactivation of the first TSG leads to CIN genes suffice for CIN to precede two TSGs. This effect sure that costly CIN emerges before one TSG, only very few

4. Conclusions

III CIN genes are sufficient for CIN to precede inactivation of the first TSG. Hence, a large number of alternative TSGs increases the number of classes II and III CIN genes needed, but does not significantly alter the number of class I CIN genes required for CIN to arise early.

5. Methods

The stochastic process illustrated in Figs. 1b and 3b can be described by differential equations. Denote by X0, X1, X2, and X4 the probabilities that a compartment is in state A+/+, B+/+, A+/−B+/+, A−/−B−/−, respectively, without chromosomal instability (CIN). Denote by Y0, Y1, Y2, and Y4 the probabilities that a compartment is in state A+/−B+/+, A−/−B+/+, A+/−B−/−, and A−/−B−/−, respectively, with CIN. The differential equations are given by

\[
\begin{align*}
X_0 &= X_0(\lambda r_1 - \lambda c_0) \\
Y_0 &= c_0 X_0 - \lambda_0 Y_0 \\
X_1 &= r_1 X_0 - \lambda (r_1 + c_1) + c_1 Y_1 \\
Y_1 &= \lambda_0 Y_0 + c_0 X_1 - s_1 Y_1 \\
X_2 &= r_2 X_1 - r_1 X_2 \\
Y_2 &= s_1 Y_1 + c_1 X_1 - s_2 Y_2 \\
X_4 &= r_3 X_2 \\
Y_4 &= s_2 Y_2
\end{align*}
\]

The transition rates between the states are given by: 

- \(r_0 = 2a/t_0\), \(r_1 = N_0(a + p_0)/t_0\), \(r_2 = N_1 2a\sqrt{\frac{p_0}{t_1}} t_1\), \(c_0 = N_0 a, c_0 = (N_0 a c_1 (1-r) p_0)/t_0, k_0 = r_0, k_1 = N_0 p_0/t_0, \) and \(r_2 = r N_1 2a\sqrt{p_0/t_1} [42,46-48]\). Cells with genotype A+/+ B+/+ and A+/−B+/+ divide every \(t_0\) days, while cells with genotype A+/−B+/+ and A+/−B−/− divide every \(t_1\) days. Let us first discuss the mutation-selection network leading from X0 via X1 to X2 to X4. The rate of change, \(r_0\), is equal to the rate that a mutation occurs in the first allele of TSG A times the probability that the mutant cell will take over the compartment. The mutation occurs at rate 2N0u/t0, because (i) the mutation can arise in any one of \(N_0\) cells; (ii) the cells divide once every \(t_0\) days; (iii) the mutation rate per gene per cell division is given by \(a\); and (iv) there are two alleles. The probability that the mutated cell takes over the compartment (i.e. reaches fixation) is given by \(1/N_0\), unless we assume that the mutation is neutral [59]. Thus, the population size, \(N_0\), cancels in the product, and we simply obtain 2a/t0. Once the first allele of TSG A has been inactivated, the second allele can be inactivated either by another point mutation or an LOH event. This process occurs at rate \(r_1 = N_0(a + p_0)/t_0\), where \(p_0\) is the rate of LOH in normal cells. We assume that A−/− cells have a large fitness advantage, which means the probability of fixation of such a cell is close to one. Inactivation of both alleles of TSG A leads to a small lesion of \(N_1\) cells dividing every \(t_1\) days. In this lesion, TSG B is inactivated at rate \(r_2 = N_1 2a\sqrt{p_0/t_1} [42,46-48]\). CIN can arise at any stage of tumorigenesis and causes very fast LOH [58]. Denote the rate of LOH in CIN
cells by $p$; we have $p \gg p_0$. The rate of change, $c_2$, is equal to the rate at which a mutation triggering CIN occurs, $N_{u_2}p_0$, times the probability that the mutant cell will take over the compartment. The mutation rate triggering CIN, $u_2$, depends on the number of available classes I, II and class III CIN genes in the human genome, $n_1$, $n_2$ and $n_3$, respectively. The mutation rates of classes I and II CIN genes are $u_1 = 2e^{-1}u_1p_0$ and $u_2 = 2e^{-1}u_2p_0$, respectively. The mutation rate of the first allele of class III CIN genes is $2u_3$, and the mutation rate of the second allele of class III CIN genes is $\gamma p_0$. The probability of fixation of a CIN cell depends on the somatic fitness, $r$, and is given by $p = (1 - 1/r)/(1 - 1/r)^{N_0}$. This is the standard fixation probability of a Moran process [60]. Therefore, the rate of change of the second allele of TSG is inactivated is given by $c_3 = [N_{u_2}r/(1 - r)]p_0r$. Let us now discuss the mutation–selection network leading from $Y_0$ via $Y_1$ to $Y_2$. The rate of change, $c_3$, is equal to the rate $r_1$, because CIN does not change the probability of point mutations. The rate at which the second allele of CIN genes reaches fixation before an $A^{+/+}$ cell does. The rate of mutation is given by $c_1 = [N_{u_2}r/(1 - r)]p_0r$. This is the standard fixation probability of a Moran process [60]. Therefore, the rate of change of the second allele of TSG is inactivated is given by $s_1 = N_{u_2}p_0r_1$. The rate at which $A^{+/+}$ is inactivated is given by $s_2 = rN_{u_2}r/2Y_1$. Eq. (1) is a system of linear differential equations which can be solved analytically using standard techniques.

Acknowledgements

The Program for Evolutionary Dynamics is supported by Jeffrey E. Epstein.

References

[38] Little MP, Wright EG. A stochastic carcinogenesis model incorporating genomic instability fitted to colon cancer data. Math Biosci 2003;183:111–34.


