ELSEVIER

Seminars in Cancer Biology 15 (2005) 484–493

seminars in
CANCER BIOLOGY

www.elsevier.com/locate/semcancer

Review

Dynamics of colorectal cancer

Franziska Michor[a,d,*], Yoh Iwasa[b], Christoph Lengauer[c], Martin A. Nowak[a]

[a] *Program for Evolutionary Dynamics, Department of Organismic and Evolutionary Biology, Department of Mathematics, Harvard University, Cambridge, MA 02138, USA*
[b] *Department of Biology, Kyushu University, Fukuoka 812-8581, Japan*
[c] *Novartis Institutes for BioMedical Research, Cambridge, MA 02139, USA*
[d] *Society of Fellows, Harvard University, Cambridge, MA 02138, USA*

Abstract

Colorectal cancer results from an accumulation of mutations in tumor suppressor genes and oncogenes. An additional defining characteristic of colorectal cancer is its genetic instability. Two main types of genetic instability have been identified. Microsatellite instability leads to an increased point mutation rate, whereas chromosomal instability refers to an enhanced rate of accumulating gross chromosomal aberrations. All colon cancer cell lines are genetically unstable. An interesting question is whether genetic instability arises early in tumorigenesis. An early emergence of genetic instability could drive most of the somatic evolution of cancer. Here, we review mathematical models of colorectal tumorigenesis and discuss the role of genetic instability.

Keywords: Colorectal cancer; Genetic instability; Mathematical model

Contents

1. Introduction

'I wish I had the voice of Homer, to sing of rectal carcinoma' are the opening lines of a poem by British mathematical biologist J.B.S. Haldane, who was diagnosed with this disease in 1964 and died a few years later. Haldane, together with Ronald Fisher and Sewell Wright, was one

* Corresponding author. Tel.: +1 617 496 4698; fax: +1 617 496 4629.
E-mail address: michor@fas.harvard.edu (F. Michor).

1044-579X/$ – see front matter © 2005 Elsevier Ltd. All rights reserved.
doi:10.1016/j.semcancer.2005.06.005

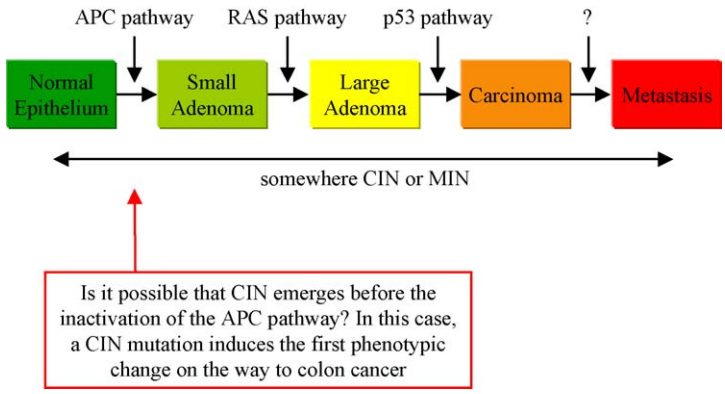

Fig. 1. The genetic model of colorectal tumorigenesis. Colorectal cancers develop over the course of 20–40 years due to genetic disruption of the APC, RAS and p53 pathways. Genetic instability arises somewhere during the process of colorectal tumorigenesis, but whether it is the first event and therefore initiates and drives the neoplastic transformation is still a matter of much debate.

of the founding fathers of population genetics. He unified Darwinian evolution with Mendelian genetics and brought about the current understanding of evolutionary biology. In this review, we apply population genetics to the evolutionary dynamics of colorectal cancer.

Colorectal cancer is the second leading cause of cancer death in the United States. In 2004, there were an estimated 100,000 new cases and 57,000 deaths from this disease [1]. The incidence of colorectal cancer, however, could be reduced dramatically by preventative methods such as colonoscopy and detection of mutations in fecal DNA [2,3].

Colorectal cancer progresses through a series of clinical and histopathological stages ranging from single crypt lesions through small benign tumors (adenomatous polyps) to malignant cancers (carcinomas) [4]. The model of colorectal tumorigenesis includes several genetic changes that are required for cancer initiation and progression [5] (Fig. 1). The earliest and most prevalent genetic event yet identified in colorectal tumorigenesis is genetic disruption of the APC (*adenomatous polyposis coli*) pathway [6]. This pathway is documented to be altered in approximately 95% of colorectal tumors [7]. A critical function of APC is the inhibition of β-catenin/Tcf-mediated transcription [8]. Inactivation of APC leads to increased β-catenin/Tcf-mediated transcription of growth-promoting genes [9]. In tumors with wild-type APC, mutations of β-catenin have been observed that render it resistant to the inhibitory effects of APC [10]. Loss of APC function or gain of β-catenin function leads to clonal expansion of the mutated epithelial cell, giving rise to a small adenoma [11]. Mutations in the RAS/RAF pathway, the p53 pathway, and several other genes/pathways drive tumor progression towards malignancy and metastasis [4]. Fig. 2 shows the age-specific incidence of colorectal cancer.

Two types of genetic instability have been identified in colorectal cancer [12]. In about 15% of all sporadic colorectal cancers, mismatch-repair deficiency leads to microsatellite instability (MIN) at the nucleotide level [13]. The remaining 85% of sporadic colorectal cancers have chromosomal instability (CIN), which refers to an increased rate of losing or gaining whole chromosomes or parts of chromosomes during cell division [14]. CIN can contribute to an increased rate of loss of heterozygosity (LOH), which is considered an important mechanism of tumor suppressor gene inactivation. More than a hundred genes can cause CIN when mutated in yeast cells, many of which have several homologues in humans [15]. These include genes that are involved in chromosome metabolism, spindle assembly and dynamics, cell-cycle regulation and mitotic checkpoint control [16]. The classification of CIN genes is based on the mutational events that are required to trigger CIN [17]. Class I CIN genes, such as MAD2, trigger CIN if one allele of the gene is inactivated or lost. Class II CIN genes, such

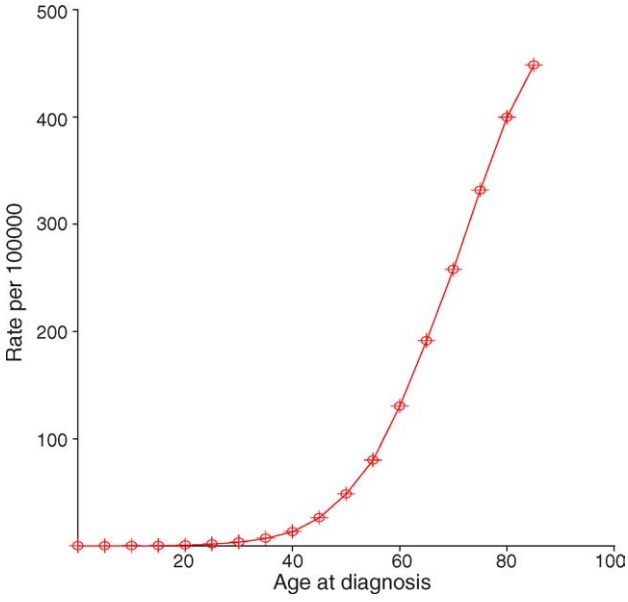

Fig. 2. Age-specific incidence of colorectal cancer in the US. Statistics were generated from malignant cases only, and rates are expressed as cases per 100,000. Data from Surveillance, Epidemiology, and End Results (SEER) Program (www.seer.cancer.gov), National Cancer Institute, released April 2004, based on the November 2003 submission.

as hBUB1, trigger CIN if one allele is mutated, and this mutation exerts an effect. Class III CIN genes, such as BRCA2, trigger CIN only if both alleles are mutated.

The role of genetic instability in colorectal tumorigenesis is an active research area of cancer biology [17,18]. Loeb was the first to propose that cancer cells acquire a 'mutator phenotype' early in tumor progression [19,20]. Several lines of evidence, experimental and theoretical, indicate that genetic instability can indeed arise early in colorectal tumorigenesis and subsequently drive tumor progression [17]. This view is widely, however not universally accepted [18,21].

In this review, we address the following questions: how can the dynamics of colorectal tumorigenesis be described with mathematical models? And what is the role of genetic instability in tumor initiation?

2. Mutations in colorectal cancer cell lines

The genetic model of colorectal tumorigenesis (Fig. 1) is based on decades of investigation into the mutational sequence of colorectal tumorigenesis [5]. The fraction of tumors with mutations in the APC pathway is approximately 95%; the same number holds for the fraction of tumors with mutations in the p53 pathway. Somatic mutations that lead to alteration of the Ras/Raf pathway have so far been found in about 70% of tumors. There are additional genes with somatic mutations in colorectal cancers, but their respective roles and pathways are less well understood and examined.

Mutational studies require both tumor and normal tissue material for comparison. The obtainment of tumor material poses a significant technical difficulty, as contamination with normal tissue renders signal and noise indistinguishable. Only a few colorectal cancers and their corresponding normal tissue counterparts have been analyzed systematically to determine the spectrum of specific mutations in a series of genes, i.e. all APC pathway genes (APC, β-catenin, and axin), p53 pathway genes (p53, MDM2, and BAX), and RAS pathway genes (K-Ras and B-Raf) in the same tumors. Though such data are sparse, profound differences between the mutations found in CIN and MIN tumors have been reported. At least a third of MIN cancers contain mutations in β-catenin instead of APC, whereas β-catenin mutations are extremely rare in non-MIN cancers [22]. This is indirect evidence suggesting that MIN arises before inactivation of the APC pathway. Additionally, the spectrum of APC mutations that occur in most MIN cancers without β-catenin mutations is different from that in non-MIN cancers [23]. In particular, there is a higher frequency of alterations in simple repeat sequences in the MIN than in the non-MIN cancers.

3. Mathematical modeling

Mathematical approaches to understanding tumorigenesis began with a multi-stage stochastic model for the age-specific incidence of cancer. This development was primarily driven by the efforts of Nordling [24], Armitage and Doll [25,26], and Fisher [27], and led to the idea that several probabilistic events are required for the somatic evolution of cancer [28,29]. Nunney also incorporated fitness of mutated cells and tissue architecture into his population genetic model of tumorigenesis [28]. Knudson used a statistical analysis of retinoblastoma to explain the role of tumor suppressor genes in sporadic and inherited cancers [30]. This work was later extended to a two-stage stochastic model of cancer initiation and progression [31], which primed many subsequent studies [32–35]. Later on, specific theories for drug resistance of tumor cells [36–40], angiogenesis [41–45], immune responses against tumors [46], genetic instabilities [47–51], and tissue architecture [52–56] were developed.

Mathematical modeling of colorectal tumorigenesis has provided insights into the dynamics and statistics of the disease. Luebeck and Moolgavkar [32] subdivided the multi-step model into two stages; during the first stage, accumulation of neutral mutations does not lead to any phenotypic changes; during the second phase, advantageous mutations drive clonal expansion of the colorectal tissue. By fitting the age-specific incidence data of colorectal cancer, they found that two rare events (inactivation of both APC alleles) are needed for tumor initiation and only one more event is necessary for malignant transformation. Luebeck and Moolgavkar conclude that genetic instability is not necessary for fitting the incidence curve of colorectal cancer. The difference between their approach and ours is that our model includes specific genetic mutations and the effective population sizes of cells where these mutations have to occur. We discuss possible parameter values for the rates of cell division in various stages of tumor progression and the selective advantages or disadvantages of mutated cells. We argue that age incidence curves of cancers can be fitted by many different models with or without genetic instability. For example, we have shown that two rate limiting hits for inactivation of a tumor suppressor gene can be interpreted as two stochastic events inactivating each of the two alleles, or as one stochastic event inactivating the first allele followed by a CIN mutation followed by very fast inactivation (LOH) of the second allele [50]. Both evolutionary trajectories have two rate limiting steps. Therefore, age incidence curves (Fig. 2) alone do not provide enough information to decide whether genetic instability is early or not.

Herrero-Jimenez et al. [57] provided a statistical analysis of colorectal cancer incidence data accounting for demographic stochasticity, i.e. human heterogeneity for inherited traits and environmental experiences. They fitted the two-stage initiation-promotion model to the incidence data and calculated adenoma growth rates, the number of initiating mutations and the rate of LOH, which was found to be significantly higher than in normal cells. Pinsky [58] adapted the multi-stage model to the problem of development and growth of adenomas, both in sporadic and inherited colorectal cancers. He fitted the model to adenoma data

including adenoma prevalence by age, the size distribution of adenomas, clustering of adenomas within individuals and the correlation between proximal and distal adenomas.

Little and Wright [59] extended the multi-stage model to incorporate genomic instability and an arbitrary number of mutational stages. They found that a model with five stages and two levels of genomic destabilization fits colon cancer incidence data. Yatabe et al. [60] investigated stem cells in the human colon by using methylation patterns and statistics. They found that colonic stem cells follow a stochastic rather than deterministic model, i.e. each stem cell division produces two, one, or zero daughter stem cells and zero, one, or two differentiated cells, respectively. This leads to drift in the number of descendents of each stem cell lineage over time. Zahl [61] developed a proportional regression model for colon cancer in Norway. Mehl [62] presented a mathematical computer simulation model to predict the natural history of colon polyp and cancer development. He incorporated different cell types and two kinetic processes, mutation and promotion, into his model.

Most mathematical analyses consist of fitting a stochastic model to colorectal cancer incidence data. The number of steps is then extracted from the fitting instead of from cell-biological and genetic assumptions of the necessary number of mutations, rate constants and fitness values. Quantitative estimates, however, are essential to correctly infer the mechanistic sequence of events. Population genetic studies have shown that the effective number of steps observed from the probability distribution of the time to cancer initiation can be smaller than the actual number [49]. This is demonstrated by stochastic tunneling in which two mutations occur in one rate-limiting process [50,63]. The effective step number also depends on the population size and the possibility of somatic selection [49].

Colorectal cancer incidence data can be explained by many alternative models. Most models are not easily falsifiable, as the assumptions of intermediate steps and rate constants are not explicitly stated. The identification of the number of rate-limiting steps in a particular model cannot justify the rejection of other interpretations of the incidence data. The assumptions about intermediate steps and rate constants must be identified to render mathematical models testable by experiments. A careful discussion of parameter values and their likely ranges is needed for evaluating hypotheses. Hence, the specification of intermediate steps and their parameter values is essential for a full understanding of colorectal tumorigenesis.

Our group has developed a mathematical model of colorectal cancer that investigates the role of genetic instability in tumorigenesis [50,64–66]. We have studied the question whether a CIN mutation induces the first phenotypic change on the way to colon cancer. We have shown in various models that this is indeed highly probably if there are enough ‘CIN genes’ in the human genome. We have also shown that tissue organization tremendously influences the probability of emergence of genetic instability [55]. Furthermore, we have developed a specific model including cell-biological and mechanical assumptions of a colonic crypt to describe colorectal tumorigenesis [56].

In the following, we offer a perspective of colorectal tumorigenesis including all sequential mutations needed for tumor initiation and progression and determine the importance of genetic instabilities in colorectal cancer.

4. The model

Here, we discuss the evolutionary dynamics of a model of colorectal tumorigenesis including alterations of three genes (Fig. 1). Inactivation of the tumor suppressor gene APC initiates tumorigenesis. Subsequent tumor progression is driven by activation of the oncogene RAS. Finally, inactivation of the tumor suppressor gene p53 gives rise to a carcinoma. Somewhere CIN or MIN emerges.

We will analyze the radical hypothesis that CIN precedes inactivation of the APC tumor suppressor gene. Here, we do not discuss MIN because it occurs only in a minority of sporadic colorectal cancers and is rarely found in other cancers. See [65] for a theoretical investigation of MIN in sporadic and inherited colorectal cancer.

The mammalian colon is lined with a rapidly proliferating epithelium. This epithelium is subdivided into 10^7 compartments of cells called crypts. Each crypt contains about 1000–4000 cells. Approximately 1–10 of those cells are stem cells, residing at the base of each crypt [60,67]. The progeny of stem cells migrate up the crypt, continuing to divide until they reach its mid-portion. Then they stop dividing and differentiate into mature cells. When the cells reach the top of the crypt, they undergo apoptosis and are engulfed by stromal cells or shed into the gut lumen. The cell migration from the base to the top of the crypt takes about 3–6 days [68].

In this setting, two scenarios of cellular dynamics within a colonic crypt are possible [50,56]. In the first scenario, only colorectal stem cells are considered to be at risk of becoming cancer cells [50]. Differentiated cells are assumed not to live long enough to accumulate the necessary number of mutations that lead to cancer. Mutated differentiated cells are washed out of the crypt by the continuous production and migration of cells to the top of the crypt where they undergo apoptosis [54]. In this case, the effective population size of each colonic crypt is equal to the number of stem cells. The stem cell compartment is assumed to be well-mixed, i.e. all stem cells reside in equivalent positions and are in direct reproductive competition with each other. There are no spatial effects. In the second scenario, all cells of the colonic crypt are assumed to be at risk of accumulating mutations that lead to cancer [56]. Mutations in specific genes might confer an increased probability to the cell to stick on top of the crypt instead of undergoing apoptosis. In this review, we use the first scenario to discuss the evolutionary dynamics of mutations in APC, RAS, and p53. See [56] for an analysis of the second scenario.

4.1. APC inactivation

We will now describe the dynamics of APC inactivation in a well-mixed compartment of cells [49,50,64–66]. Denote the number of stem cells by N_0; stem cells divide every τ_0 days. Denote the probabilities of inactivating the first and second APC allele per cell division by u_1 and u_2, respectively. We have $u_1 < u_2$ because there are more possibilities, such as LOH and mitotic recombination, for the second hit. If the mutation rate is less than the inverse of the population size, then the approximation of homogeneous compartments holds: at any time, all cells of the compartment will have 0, 1 or 2 inactivated alleles of APC. Denote the probabilities that all cells of the compartment have 0, 1 or 2 inactivated alleles of APC at time t by $X_0(t)$, $X_1(t)$ and $X_2(t)$, respectively (Fig. 3a). The transition rate between states $APC^{+/+}$ and $APC^{+/-}$ is equal to the rate that a mutation occurs in the first APC allele times the probability that the mutant cell takes over the compartment. The mutation occurs at rate $N_0 u_1/\tau_0$, because (i) the mutation can arise in any one of N_0 cells, (ii) the rate at which the first APC allele is inactivated per cell division is u_1, and (iii) the cells divide once every τ_0 days. The probability that the mutated cell takes over the compartment (i.e., reaches fixation) is given by $1/N_0$, if we assume that the mutation is neutral [69]. Thus, the compartment size, N_0, cancels in the product, and we obtain u_1/τ_0 as the transition rate between states $APC^{+/+}$ and $APC^{+/-}$. The transition rate between states $APC^{+/-}$ and $APC^{-/-}$ is equal to the rate that a mutation occurs in the second APC allele times the probability that the mutant cell will take over the compartment. We assume that a cell with two inactivated APC alleles has a large fitness advantage, which means the fixation probability of such a cell is close to one. Therefore, we obtain $N_0 u_2/\tau_0$ as the transition rate between states $APC^{+/-}$ and $APC^{-/-}$ (Fig. 3a). The differential equation describing the stochastic process is outlined in Appendix A. The overall probability that APC has been inactivated by time t is given by $X_2(t) = N_0 u_1 u_2 t^2/(2\tau_0{}^2)$. This probability accumulates as

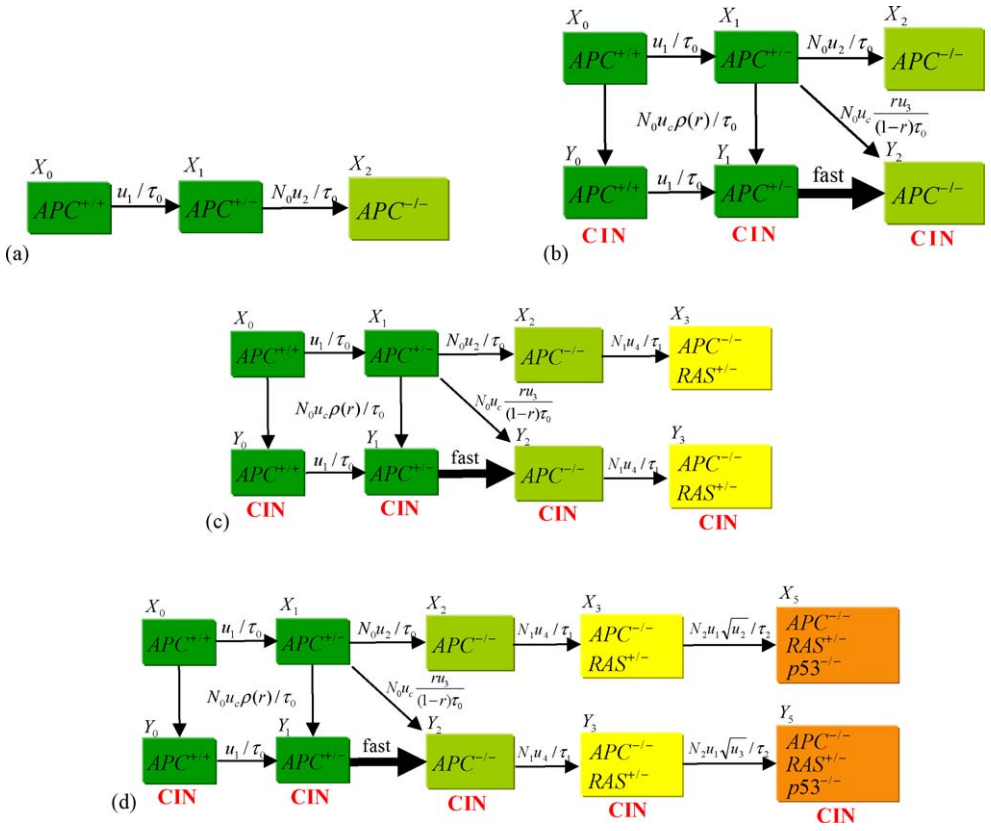

Fig. 3. Dynamics of colorectal tumorigenesis. (a) Inactivation of APC in a well-mixed compartment of colonic stem cells, N_0, that divide once every τ_0 days. The probabilities of inactivating the first and second APC allele per cell division are denoted by u_1 and u_2. (b) Chromosomal instability (CIN) can arise at any stage of tumorigenesis and causes an accelerated rate of inactivation of the second APC allele; denote this rate by u_3. A CIN cell has relative fitness r, arises at rate u_c per cell division, and reaches fixation in the compartment with probability ρ. If an $APC^{+/-}$ cell clone with CIN produces an $APC^{-/-}$ cell with CIN before taking over the compartment, a stochastic tunnel arises: the compartment evolves from state $APC^{+/-}$ without CIN directly to state $APC^{-/-}$ with CIN without ever visiting state $APC^{+/-}$ with CIN (diagonal arrow). (c) APC inactivation leads to a small lesion of N_1 cells dividing every τ_1 days in which activation of RAS might be the next sequential mutation necessary for colorectal tumorigenesis. The rate at which RAS is mutated per cell division is denoted by u_4. (d) RAS activation leads to an adenoma of N_2 cells dividing every τ_2 days in which inactivation of p53 might be the next sequential mutation. Suppose the p53 alleles are inactivated at the same rates as the APC alleles. Due to the large compartment size, it takes only one rate-limiting hit to inactivate both p53 alleles.

a second order of time, which means that it takes two rate-limiting hits to inactivate APC in the compartment of cells [30,50].

4.2. Chromosomal instability (CIN)

CIN could arise at any stage of tumorigenesis. Here we discuss the possibility that CIN emerges before inactivation of APC (Fig. 3b). The crucial effect of CIN is to accelerate the inactivation rate of the second APC allele. Denote the probability of inactivating the second APC allele per CIN cell division by u_3; we have $u_3 \gg u_2$. This increased rate, however, might confer a selective cost to the CIN cell, because such a cell has an increased probability of lethal mutations and apoptosis. Conversely, it is possible that a CIN cell has a faster cell cycling time because of avoiding certain checkpoints. Denote the relative fitness of a CIN cell by r. If $r = 1$, then CIN is neutral; if $r < 1$, it is disadvantageous, and a CIN cell has a lower growth rate than a cell with stable karyotype; if $r > 1$, then it is advantageous, and a CIN cell has a larger growth rate than a cell with stable karyotype. Denote the probabilities that all cells of the compartment have CIN and 0, 1 or 2 inactivated APC alleles at time t by $Y_0(t)$, $Y_1(t)$ and $Y_2(t)$, respectively. The transition rate between states $APC^{+/+}$ without CIN and $APC^{+/+}$ with CIN is the same as the transition rate between states $APC^{+/-}$ without CIN and $APC^{+/-}$ with CIN and equal to the rate at which a mutation triggering CIN occurs, $N_0 u_c/\tau_0$, times the probability that the mutant cell will take over the compartment. The fixation probability of a CIN cell depends on its somatic fitness, r, and the compartment size, N_0, and is given by $\rho = (1 - 1/r)/(1 - 1/r^{N_0})$. This is the standard fixation probability of a Moran process [64]. Therefore, the transition rate is given by $N_0 u_c \rho/\tau_0$. Note that we only consider 'single hit', i.e. class I and II CIN genes here.

If an $APC^{+/-}$ cell clone with CIN produces an $APC^{-/-}$ cell with CIN before taking over the compartment, a stochastic tunnel arises [50,63]: the compartment evolves from state $APC^{+/-}$ without CIN directly to state $APC^{-/-}$ with CIN without ever visiting state $APC^{+/-}$ with CIN. This tunnel occurs at rate $N_0 u_c r u_3/[(1 - r)\tau_0]$ (Fig. 3b). If CIN is neutral, $r = 1$, then the tunnel does not occur. The transition rate between states $APC^{+/+}$ with CIN and $APC^{+/-}$ with CIN is u_1/τ_0 as before, because CIN is not assumed to cause increased point mutation rates. The transition rate between states $APC^{+/-}$ with CIN and $APC^{-/-}$ with CIN is given by $N_0 u_3/\tau_0$; this rate is fast in the sense that $N_0 u_3/\tau_0 \gg 1$, and hence is not rate-limiting.

The overall probability that APC has been inactivated with CIN by time t is given by $Y_2(t) = N_0 u_1 u_c [2\rho + r u_3/(1 - r)]t^2/(2\tau_0^2)$. This probability again accumulates as a second order of time, so it also takes two rate-limiting hits to inactivate APC in a compartment with CIN. Hence, Knudson's two-hit hypothesis [30] is compatible with tumor suppressor gene inactivation both without and with CIN [50].

4.3. RAS activation

Inactivation of APC leads to clonal expansion of the colonic stem cell compartment to a small lesion of N_1 cells dividing every τ_1 days. It is in this lesion that other genes must be genetically altered for further tumor progression. The next genetic event in the sequence might be activation of the oncogene RAS [4].

Let us discuss the activation dynamics of RAS [49]. Denote the probability that a RAS mutation occurs per cell division by u_4. Denote the probabilities that a lesion consists only of $APC^{-/-}RAS^{+/-}$ cells without and with CIN at time t by $X_3(t)$ and $Y_3(t)$, respectively (Fig. 3c). The transition rate between states $APC^{-/-}RAS^{+/+}$ and $APC^{-/-}RAS^{+/-}$ without CIN is given by $N_1 u_4/\tau_1$, because (i) the mutation can arise in any one of N_1 cells, (ii) the rate at which the RAS allele is mutated per cell division is u_4, (iii) the cells divide once every τ_1 days, and (iv) a cell with a RAS mutation is assumed to have a large fitness advantage, such that its fixation probability is close to 1. The transition rate between states $APC^{-/-}RAS^{+/+}$ and $APC^{-/-}RAS^{+/-}$ with CIN is the same as the transition rate between states $APC^{-/-}RAS^{+/+}$ and $APC^{-/-}RAS^{+/-}$ without CIN, because the CIN phenotype does not alter the point mutation rate.

4.4. p53 inactivation

Activation of RAS leads to clonal expansion of the compartment to a late adenoma of N_2 cells dividing every τ_2 days. In this adenoma, other genes must be altered for progression to a carcinoma. The next genetic event in the sequence might be inactivation of the tumor suppressor gene p53 [4].

Let us now discuss the inactivation dynamics of p53 in the context of our model [49]. Denote the probability of inactivating the first p53 allele per cell division by u_1 and the probabilities of inactivating the second p53 allele per normal and CIN cell division by u_2 and u_3, respectively. Again we have $u_3 \gg u_2 > u_1$. Here, we assume that the p53 alleles are inactivated at the same rates as the APC alleles. Denote the probabilities that a lesion consists only of $APC^{-/-}RAS^{+/-}p53^{-/-}$ cells without and with CIN at time t by $X_5(t)$ and $Y_5(t)$, respectively (Fig. 3d). The transition rate between states $APC^{-/-}RAS^{+/-}p53^{+/-}$ and $APC^{-/-}RAS^{+/-}p53^{-/-}$ without CIN is $N_2 u_1 \sqrt{u_2}/\tau_2$, because (i) the mutation can arise in any of N_2 cells, (ii) the rate at which the two p53 alleles are inactivated per cell division is $u_1\sqrt{u_2}$ due to the increased compartment size [63], and (iii) the cells divide once every τ_2 days. The transition rate between states $APC^{-/-}RAS^{+/-}p53^{+/-}$ and $APC^{-/-}RAS^{+/-}p53^{-/-}$ with CIN is $N_2 u_1 \sqrt{u_3}/\tau_2$, because the rate at which the two p53 alleles are inactivated per CIN cell division is $u_1\sqrt{u_3}$ [63]. The differential equation is given in Appendix A and can be used to calculate the probability as a function of time to develop a carcinoma. We can also determine the role of CIN in tumorigenesis. In the following, we will provide numerical examples of our model.

4.5. The cost of CIN

Once a CIN phenotype has reached fixation in a pre-cancer lesion, the question remains how to quantify the further cost of CIN. Comparing two lesions, one without CIN and one with CIN, we might be tempted to argue that the rate of evolution of the CIN lesion has to be multiplied by a factor (given by a number between 0 and 1) which takes into account the increased death rate caused by CIN. Therefore, the rate of evolution of the CIN lesion is reduced by a factor describing the cost of CIN.

On the other hand, it is conceivable that the increased rate of cell death in the CIN lesion allows for an increased number of cell divisions (to make up for the cells that have died) and therefore even accelerates the rate of evolution. A variety of simple mathematical models can be constructed that demonstrate that the 'cost of CIN' can be a net selective advantage when comparing the competition between CIN and non-CIN lesions.

Since both scenarios are possible, in this paper, we assume that CIN lesions and non-CIN lesions are neutral with respect to the cost of CIN during further tumor progression. The cost of CIN is only considered when calculating the probability that a CIN cell reaches fixation in the compartment where it first arises.

4.6. Numerical examples

The human colon consists of approximately $Z = 10^7$ crypts, each of which is replenished by about 1–10 stem cells [60,66]. Let us choose $N_0 = 4$. Later, we will discuss numerical examples with $N_0 = 1$, 4 and 10 stem cells. The mutation rate per base per cell division is approximately 10^{-10} [70]. Suppose the APC tumor suppressor gene can be inactivated by point mutations in any of about 500 bases [71]; hence the rate of inactivating the first of two alleles is $u_1 = 2 \times 500 \times 10^{-10} = 10^{-7}$ per cell division both in normal and CIN cells. The rate of inactivating the second allele per cell division, u_2, is the sum of the rate of LOH and the rate of a point mutation. There are no reliable measurements of the rate of LOH in non-CIN cells. Plausible values range from 10^{-7} to 10^{-5}. In our opinion, the most likely scenario is that the rate of LOH has the same order of magnitude as u_1 or is slightly larger. 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 the rate of LOH is much larger than u_1, then two distinct point mutations should never be observed in the two TSG alleles. Assume the rate of inactivating the second allele is about $u_2 = 10^{-6}$ per cell division in normal cells. The rate of inactivating the second allele in CIN cells has been measured to be $u_3 = 10^{-2}$ [14].

Assume clonal expansion of the stem cell compartment to $N_1 = 10^3$ cells after APC inactivation and to $N_2 = 10^5$ cells after RAS activation. Suppose normal stem cells divide once every $\tau_0 = 10$ days [60]. Assume that $APC^{-/-}$ cells divide every $\tau_1 = 5$ days and $APC^{-/-}RAS^{+/-}$ cells every $\tau_2 = 1$ day. The most substantial cost CIN can possibly have is $r = (1 - u_3)^{45} \approx 0.6$; this means that loss of any chromosome other than the one containing the tumor suppressor gene locus is lethal.

Table 1
Percentage of carcinomas that were initiated by chromosomal instability (CIN), if there is $n_1 = 1$ class I CIN gene or $n_2 = 10$ class II CIN genes in the human genome

r	$n_1 = 1$ (%)	$n_2 = 10$ (%)
(a) $N_0 = 1$, $N_1 = 10^4$, $N_2 = 10^6$		
1.0	89	81
0.8	90	81
0.6	90	81
(b) $N_0 = 4$, $N_1 = 10^4$, $N_2 = 10^6$		
1.0	70	54
0.8	64	47
0.6	50	33
(c) $N_0 = 4$, $N_1 = 10^3$, $N_2 = 10^5$		
1.0	93	88
0.8	92	85
0.6	86	75
(d) $N_0 = 10$, $N_1 = 10^3$, $N_2 = 10^5$		
1.0	85	74
0.8	74	59
0.6	40	25

Class I CIN genes, n_1, trigger CIN if one allele is mutated or lost. Class II CIN genes, n_2, trigger CIN if one allele is mutated in a dominant negative fashion The relative fitness of CIN cells is denoted by r. Parameter values are $u_1 = 10^{-7}$, $u_2 = 10^{-6}$, $u_3 = 10^{-2}$, $u_4 = 10^{-9}$, $\tau_0 = 10$, $\tau_1 = 5$, $\tau_2 = 1$, $t = 70$ years, and (a) $N_0 = 1$, $N_1 = 10^4$ and $N_2 = 10^6$, (b) $N_0 = 4$, $N_1 = 10^4$ and $N_2 = 10^6$, (c) $N_0 = 4$, $N_1 = 10^3$ and $N_2 = 10^5$, and (d) $N_0 = 10$, $N_1 = 10^3$ and $N_2 = 10^5$.

Assume there is only one class I CIN gene that triggers CIN when mutated; then $u_c = 2u_2 = 2 \times 10^{-6}$, because either allele can be inactivated. The expected number of carcinomas that were initiated by inactivation of APC in a $t = 70$ year-old is 0.0014. The expected number of carcinomas that were initiated by CIN in a $t = 70$ year-old is 0.0082 (see Appendix A). Hence, 85% of carcinomas were initiated by CIN. Among 1000 people, there are nine colorectal cancer cases, eight of which were initiated by CIN. However, all of them might have CIN in the end.

Assume there is one class II CIN gene that triggers CIN when mutated; then $u_c = u_1 = 10^{-7}$. The expected number of carcinomas that were initiated by inactivation of APC in a $t = 70$ year-old is 0.0014. The expected number of carcinomas that were initiated by CIN is 0.0014 (see Appendix A). Hence, 22% of carcinomas were initiated by CIN.

Alternatively, we can calculate the minimum number of CIN genes in the human genome needed to make sure CIN arises before the inactivation of APC. Given the parameter values above, one class I CIN gene or four class II CIN genes are necessary to ensure that CIN arises early in colorectal tumorigenesis. Note that there are hundreds of CIN genes in yeast and even larger numbers are expected in humans.

Table 1 provides further examples of parameter values.

5. Discussion

In this review, we have discussed the mutational sequence that leads to colorectal cancer. The traditional genetic model of colorectal tumorigenesis [5] describes several genetic changes that are required for cancer initiation and progression (Fig. 1). However, no large scale sequencing of genes like APC or p53 has been applied. An important goal of the field should be detailed experimental analyses of the mutations found in colorectal tumors.

We have provided a review of the literature of mathematical and statistical analyses of colon cancer. Several excellent statistical studies have been performed to analyze colorectal cancer incidence data. Stochastic multi-stage and population genetic models have been developed and been fitted to incidence data. Although this effort determined the number of stages needed for initiation and promotion of colon cancer, the biological implications are unclear; biologically and clinically useful investigation of colorectal tumorigenesis must include the exact specification of individual steps. Experimental input of parameter values such as mutation rates, fitness values and rate of population growth are required for a quantitative understanding of cancer progression.

We have reviewed our mathematical approach to colon cancer by means of a specific model of colorectal tumorigenesis. We have assumed that colorectal tumorigenesis is driven by subsequent genetic alterations of APC, RAS, and p53. Naturally, the model describes only a subset of cancers and does not attempt to provide a universally found sequence of mutations leading to colorectal cancer. There is the possibility of a second tumor suppressor gene that needs to be genetically altered after inactivation of APC and before activation of RAS. In that case and if both tumor suppressor genes are inactivated in rate-limiting situations, i.e. if inactivation of APC alone does not lead to a large tumor, then chromosomal instability is very likely to initiate the mutational sequence that leads to colorectal cancer [71].

On the basis of this case study, we developed a mathematical representation of the evolutionary dynamics of colorectal tumorigenesis. We specified the dependence of the transition rates between stages on mutation rates, population sizes, fitness values and cell division times. We introduced mutation of chromosomal instability (CIN) genes that lead to an increased rate of loss of heterozygosity (LOH), thereby accelerating the inactivation of tumor suppressor genes. We investigated the role of CIN in initiating and driving tumor progression in dependence of the number of genes that can initiate CIN when mutated in colorectal tissue. Given the experimental input of rate constants and parameter values, we found that one or two CIN genes in the genome are enough to make sure CIN emerges early. Therefore, it is very likely that a CIN mutation initiates colorectal tumorigenesis.

Acknowledgements

The Program for Evolutionary Dynamics is supported by Jeffrey E. Epstein. The authors thank Bert Vogelstein for advice.

Appendix A

The stochastic process illustrated in Fig. 3 can be described by differential equations. Denote the probabilities that a compartment is in state $APC^{+/+}$, $APC^{+/-}$, $APC^{-/-}$, $APC^{-/-}\mathrm{RAS}^{+/-}$, and $APC^{-/-}\mathrm{RAS}^{+/-}p53^{-/-}$ without CIN by X_0, X_1, X_2, X_3, and X_5, respectively. Denote the probabilities that a compartment is in the corresponding states with CIN by Y_0, Y_1, Y_2, Y_3, and Y_5, respectively. The differential equations are given by

$$\begin{aligned}
\dot{X}_0 &= -(a+b)X_0 & \dot{Y}_0 &= bX_0 - aY_0 \\
\dot{X}_1 &= aX_0 - (b+c+d)X_1 & \dot{Y}_1 &= aY_0 + bX_1 - gY_1 \\
\dot{X}_2 &= cX_1 - eX_2 & \dot{Y}_2 &= gY_1 + dX_1 - eY_2 \\
\dot{X}_3 &= eX_2 - fX_3 & \dot{Y}_3 &= eY_2 - hY_3 \\
\dot{X}_5 &= fX_3 & \dot{Y}_5 &= hY_3
\end{aligned} \quad (1)$$

Here $a = u_1/r_0$, $b = N_0 u_c \rho(r)/\tau_0$, $c = N_0 u_2/\tau_0$, $d = N_0 u_c r u_3/[(1-r)\tau_0]$, $e = N_1 u_4/\tau_1$, $f = N_2 u_1 \sqrt{u_2}/\tau_2$, $g = N_0 u_3/\tau_0$, and $h = N_2 u_1 \sqrt{u_3}/\tau_2$. See the main text for an explanation of the parameter values. Eq. (1) is a system of linear differential equations, which can be solved analytically using standard techniques. The numerical examples in Section 4.6 contain numerical simulation of system (1).

References

[1] Cancer Facts & Figures 2004. Atlanta: American Cancer Society; 2004.
[2] Sidransky D, Tokino T, Hamilton SR, Kinzler KW, Levin B, Frost P, et al. Identification of ras oncogene mutations in the stool of patients with curable colorectal tumors. Science 1992;256:102–5.
[3] Traverso G, Shuber A, Levin B, Johnson C, Olsson L, Schoetz Jr DJ, et al. Detection of APC mutations in fecal DNA from patients with colorectal tumors. N Engl J Med 2002;346:311–20.
[4] Vogelstein B, Kinzler KW. The genetic basis for human cancer. 2nd ed. Toronto: McGraw-Hill; 2001.
[5] Fearon ER, Vogelstein B. A genetic model for colorectal carcinogenesis. Cell 1990;61:759–67.
[6] Kinzler KW, et al. Identification of a gene located at chromosome 5q21 that is mutated in colorectal cancers. Science 1991;251:1366–70.
[7] Powell SM, Zilz N, Beazer-Barclay Y, Bryan TM, Hamilton SR, Thibodeau SN, et al. APC mutations occur early during colorectal tumorigenesis. Nature 1992;359:235–7.
[8] Su LK, Vogelstein B, Kinzler KW. Association of the APC tumor suppressor protein with catenins. Science 1993;262:1734–7.
[9] He TC, Sparks AB, Rago C, Hermeking H, Zawel L, da Costa LT, et al. Identification of c-MYC as a target of the APC pathway. Science 1998;281:1509–12.
[10] Morin PJ, Sparks AB, Korinek V, Barker N, Clevers H, Vogelstein B, et al. Activation of beta-catenin-Tcf signaling in colon can-

cer by mutations in beta-catenin or APC. Science 1997;275:1787–90.
[11] Su LK, Kinzler KW, Vogelstein B, Preisinger AC, Moser AR, Luongo C, et al. Dove WF Multiple intestinal neoplasia caused by a mutation in the murine homolog of the APC gene. Science 1992;256:668–70.
[12] Lengauer C, Kinzler KW, Vogelstein B. Genetic instabilities in human cancers. Nature 1998;396:643–9.
[13] Aaltonen LA, Peltomaki P, Leach FS, Sistonen P, Pylkkanen L, Mecklin JP, et al. Clues to the pathogenesis of familial colorectal cancer. Science 1993;260:812–6.
[14] Lengauer C, Kinzler KW, Vogelstein B. Genetic instability in colorectal cancer. Nature 1997;386:623–7.
[15] Jallepalli PV, Lengauer C. Chromosome segregation and cancer: cutting through the mystery. Nat Rev Can 2001;1:109–17.
[16] Cahill DP, Lengauer C, Yu J, Riggins GJ, Willson JK, Markowitz SD, et al. Mutations of mitotic checkpoint genes in human cancers. Nature 1998;392:300–3.
[17] Rajagopalan H, Nowak MA, Vogelstein B, Lengauer C. The significance of unstable chromosomes in colorectal cancer. Nat Rev Can 2003;3:695–701.
[18] Sieber OM, Heinimann K, Tomlinson IPM. Genomic instability–the engine of tumorigenesis. Nat Rev Can 2003;3:701–8.
[19] Loeb LA, Springgate CF, Battula N. Errors in DNA replication as a basis for malignant change. Cancer Res 1974;34:2311–21.
[20] Loeb LA. A mutator phenotype in cancer. Cancer Res 2001;61:3230–9.
[21] Duesberg PH. Are cancers dependent on oncogenes or on aneuploidy. Cancer Genet Cytogenet 2003;143:89–91.
[22] Sparks AB, Morin PJ, Vogelstein B, Kinzler KW. Mutational analysis of the APC/β-catenin/Tcf pathway in colorectal cancer. Cancer Res 1998;58:1130–4.
[23] Huang J, et al. APC mutations in colorectal tumors with mismatch repair deficiency. Proc Natl Acad Sci USA 1996;93:9049–54.
[24] Nordling NO. A new theory on the cancer-inducing mechanism. Br J Cancer 1953;7:68–72.
[25] Armitage P, Doll R. The age distribution of cancer and a multi-stage theory of carcinogenesis. Br J Cancer 1954;8:1–12.
[26] Armitage P, Doll R. A two-stage theory of carcinogenesis in relation to the age distribution of human cancer. Br J Cancer 1957;9: 161–9.
[27] Fisher JC. Multiple-mutation theory of carcinogenesis. Nature 1959;181:651–2.
[28] Nunney L. Lineage selection and the evolution of multistage carcinogenesis. Proc R Soc Lond B 1999;266:493–8.
[29] Tomlinson IPM, Sasieni P, Bodmer W. How many mutations in cancer. Am J Pathol 2002;160:755–8.
[30] Knudson AG. Mutation and cancer: statistical study of retinoblastoma. Proc Natl Acad Sci USA 1971;68:820–3.
[31] Moolgavkar SH, Knudson AG. Mutation and cancer: a model for human carcinogenesis. JNCI 1981;66:1037–52.
[32] Luebeck EG, Moolgavkar SH. Multistage carcinogenesis and the incidence of colorectal cancer. Proc Natl Acad Sci USA 2002;99:15095–100.
[33] Grist SA, McCarron M, Kutlaca A, Turner DR, Morley AA. In vivo human somatic mutation: frequency and spectrum with age. Mutat Res 1992;266:189–96.
[34] Gatenby RA, Vincent TL. An evolutionary model of carcinogenesis. Cancer Res 2003;63:6212–20.
[35] Frank SA. Age-specific incidence of inherited versus sporadic cancers: a test of the multistage theory of carcinogenesis. Proc Natl Acad Sci USA 2005;102:1071–5.
[36] Goldie JH, Coldman AJ. A mathematical model for relating the drug sensitivity of tumors to their spontaneous mutation rate. Cancer Treat Rep 1979;63:1727–33.
[37] Goldie JH, Coldman AJ. Quantitative model for multiple levels of drug resistance in clinical tumors. Cancer Treat Rep 1983;67:923–31.
[38] Kimmel M, Axelrod DE. Mathematical models of gene amplification with applications to cellular drug resistance and tumorigenicity. Genetics 1990;125:633–44.
[39] Panetta JC. A mathematical model of drug resistance: heterogeneous tumors. Math Biosci 1998;147:41–61.
[40] Jackson TL, Byrne HM. A mathematical model to study the effects of drug resistance and vasculature on the response of solid tumors to chemotherapy. Math Biosci 2000;164:17–38.
[41] Chaplain MA, Giles SM, Sleeman BD, Jarvis RJ. A mathematical analysis of a model for tumour angiogenesis. J Math Biol 1995;33:744–70.
[42] Olsen L, Sherratt JA, Maini PK, Arnold F. A mathematical model for the capillary endothelial cell-extracellular matrix interactions in wound-healing angiogenesis. IMA J Math Appl Med Biol 1997;14:261–81.
[43] Anderson AR, Chaplain MA. Continuous and discrete mathematical models of tumor-induced angiogenesis. Bull Math Biol 1998;60:857–99.
[44] Levine HA, Tucker AL, Nielsen-Hamilton M. A mathematical model for the role of cell signal transduction in the initiation and inhibition of angiogenesis. Growth Factors 2002;20:155–75.
[45] Stroll BR, Migliorini C, Kadambi A, Munn LL, Jain RK. A mathematical model of the contribution of endothelial progenitor cells to angiogenesis in tumors: implications for antiangiogenic therapy. Blood 2003;102:2555–61.
[46] Owen MR, Sherratt JA. Mathematical modeling of macrophage dynamics in tumors. Math Models Methods Appl Biol Chem 1999;377:675–84.
[47] Maser RS, DePinho RA. Connecting chromosomes, crisis, and cancer. Science 2002;297:565–9.
[48] Strauss BS. Hypermutability in carcinogenesis. Genetics 1998;148:1619–26.
[49] Michor F, Iwasa Y, Nowak MA. Dynamics of cancer progression. Nat Rev Can 2004;4:197–205.
[50] Nowak MA, Komarova NL, Sengupta A, Jallepalli PV, Shih I-M, Vogelstein B, et al. The role of chromosomal instability in tumor initiation. Proc Natl Acad Sci USA 2002;99:16226–31.
[51] Breivik J, Gaudernack G. Genomic instability, DNA methylation, and natural selection in colorectal carcinogenesis. Semin Cancer Biol 1999;9:245–54.
[52] Frank SA. Somatic mutation: early cancer steps depend on tissue architecture. Current Biol 2003;13:261–3.
[53] Frank SA, Iwasa Y, Nowak MA. Patterns of cell division and the risk of cancer. Genetics 2003;163:1527–32.
[54] Michor F, Nowak MA, Frank SA, Iwasa Y. Stochastic elimination of cancer cells. Proc R Soc Lond B 2003;270:2017–24.
[55] Michor F, Iwasa Y, Komarova NL, Nowak MA. Local regulation of homeostasis favors chromosomal instability. Curr Biol 2003;13:581–4.
[56] Michor F, Iwasa Y, Rajagopalan H, Lengauer C, Nowak MA. Linear model of colon cancer initiation. Cell Cycle 2004;3:358–62.
[57] Herrero-Jimenez P, Tomita-Mitchell A, Furth EE, Morgenthaler S, Thilly WG. Population risk and physiological rate parameters for colon cancer. The union of an explicit model for carcinogenesis with the public health records of the United States. Mut Res 2000;447:73–116.
[58] Pinsky PF. A multi-stage model of adenoma development. J Theor Biol 2000;207:129–43.
[59] Little MP, Wright EG. A stochastic carcinogenesis model incorporating genomic instability fitted to colon cancer data. Math Biosci 2003;183:111–34.
[60] Yatabe Y, Tavare S, Shibata D. Investigating stem cells in human colon by using methylation patterns. Proc Natl Acad Sci USA 2001;98:10839–44.

[61] Zahl PH. A proportional regression model for 20 year survival of colon cancer in Norway. Stat Med 1995;14:1249–61.
[62] Mehl LE. A mathematical computer simulation model for the development of colonic polyps and colon cancer. J Surg Oncol 1991;47:243–52.
[63] Iwasa Y, Michor F, Nowak MA. Stochastic tunnels in evolutionary dynamics. Genetics 2004;166:1571–9.
[64] Komarova NL, Sengupta A, Nowak MA. Mutation-selection networks of cancer initiation: tumor suppressor genes and chromosomal instability. J Theor Biol 2003;223:433–50.
[65] Komarova NL, Lengauer C, Vogelstein B, Nowak MA. Dynamics of genetic instability in sporadic and familial colorectal cancer. Can Biol Ther 2002;1:685–92.
[66] Michor F, Iwasa Y, Lengauer C, Vogelstein B, Nowak MA. Can chromosomal instability initiate tumorigenesis. Semin Cancer Biol 2005;15:43–9.
[67] Bach SP, Renehan AG, Potten CS. Stem cells: the intestinal stem cell as a paradigm. Carcinogenesis 2000;21:469–76.
[68] Lipkin M, Sherlock PJ, Bell B. Generation time of epithelial cells in the human colon. Nature 1962;195:175–7.
[69] Kimura M. The role of compensatory neutral mutations in molecular evolution. J Genet 1985;64:7–19.
[70] Kunkel TA, Bebenek K. DNA replication fidelity. Annu Rev Biochem 2000;69:497–529.
[71] Nagase H, Nakamura Y. Mutations of the APC gene. Hum Mut 1993;2:425–34.