DYNAMICS OF CANCER PROGRESSION

Franziska Michor, Yoh Iwasa‡ and Martin A. Nowak**

Evolutionary concepts such as mutation and selection can be best described when formulated as mathematical equations. Cancer arises as a consequence of somatic evolution. Therefore, a mathematical approach can be used to understand the process of cancer initiation and progression. But what are the fundamental principles that govern the dynamics of activating oncogenes and inactivating tumour-suppressor genes in populations of reproducing cells? Also, how does a quantitative theory of somatic mutation and selection help us to evaluate the role of genetic instability?

Cancer is a genetic disease[1]. Although environmental and other non-genetic factors have roles in many stages of tumorigenesis, it is widely accepted that cancer arises because of mutations in cancer-susceptibility genes. These genes belong to one of three classes[1,2]: gatekeepers, caretakers and landscapers. Gatekeepers directly regulate growth and differentiation pathways of the cell and comprise oncogenes and tumour-suppressor genes (TSGs). Caretakers, by contrast, promote tumorigenesis indirectly[3,4]. They function in maintaining the genomic integrity of the cell. Mutation of caretakers can lead to genetic instability, and the cell rapidly accumulates changes in other genes that directly control cell birth and death. Landscaper defects do not directly affect cellular growth, but generate an abnormal stromal environment that contributes to the neoplastic transformation of cells[5].

The alteration of one gene, however, does not suffice to give rise to full-blown cancer. For progression towards malignancy and invasion, further mutational hits are necessary[6–8]. So, the risk of cancer development depends not only on mutations initiating tumorigenesis, but also on subsequent mutations driving tumour progression.

A quantitative understanding of cancer biology requires a mathematical framework to describe the fundamental principles of population genetics and evolution that govern tumour initiation and progression[9,10]. Mutation, selection and tissue organization determine the dynamics of tumorigenesis[11–14] and should be studied quantitatively, both in terms of experiment and theory[15].

The mathematical investigation of cancer began in the 1950s, when Nordling[16], Armitage and Doll[17,18], and Fisher[19] set out to explain the age-dependent incidence curves of human cancers. These seminal studies led to the idea that several probabilistic events are required for the somatic evolution of cancer[20,21]. In the early 1970s, Knudson used a statistical analysis of the incidence of retinoblastoma in children to explain the role of TSGs in sporadic and inherited cancers[9]. This work was later extended to a two-stage stochastic model for the process of cancer initiation and progression[22], which inspired much subsequent work[23–25]. Later on, considerable effort was devoted to the development of specific theories for drug resistance[26–27], angiogenesis[28], immune responses against tumours[29] and genetic instabilities[30–35].

In this review, we address the following questions: what are the fundamental principles that determine the dynamics of activating oncogenes and inactivating TSGs? How do mutation, selection and tissue architecture influence the rate of tumour initiation and progression? Finally, how do quantitative approaches help us to investigate the role of genetic instability in tumorigenesis?

Oncogenes

Oncogenes can contribute to tumorigenesis if one allele is mutated or inappropriately expressed[1]. Over the past decades, many oncogenes have been discovered that are involved in various stages of human cancers — tumour initiation, progression, angiogenesis and metastasis.

**Program for Evolutionary Dynamics, Harvard University, One Brattle Square, Cambridge, MA 02138, USA.*
‡Department of Biology, Kyushu University, Fukuoka 812-8581, Japan.
Correspondence to M.A.N.
e-mail: martin_nowak@harvard.edu
doi:10.1038/nrc1295

Summary

- Cancer is principally caused by mutations in cancer-susceptibility genes, which include oncogenes, tumor-suppressor genes (TSGs) and genes causing genetic instability. Cancer arises when a single cellular lineage receives multiple mutations.
- Epithelial tissues are subdivided into compartments, and cancer initiation occurs in compartments. Within each compartment, there is a continuous turnover of cells. Each compartment is replenished by division and differentiation of a small number of stem cells. In a healthy tissue, homeostatic mechanisms maintain constant cell numbers.
- Mathematical models describe the process of cancer initiation and progression and provide a quantitative understanding of the dynamics of tumorigenesis with respect to mutation, selection, genetic instability and tissue architecture.
- Mutations that activate oncogenes can confer a selective advantage to the cell. We calculate the time until a cellular lineage with an activated oncogene arises and takes over a population of cells.
- Inactivating both alleles of a TSG also leads to a selective advantage to the cell. The dynamics of TSG inactivation are described by three kinetic laws that depend on the size of the cellular population and the mutation rates. In small, intermediate and large populations, a TSG is inactivated, respectively, by two, one and zero rate-limiting hits.
- Chromosomal instability (CIN) accelerates the rate of TSG inactivation.
- It takes two rate-limiting hits to inactivate a TSG in a small population of cells with or without CIN. Therefore, CIN mutations can occur early in tumorigenesis.
- Knudson's two-hit hypothesis is compatible with the idea that one mutation occurs in the first allele of the TSG and one mutation occurs in a CIN gene. The mutation inactivating the second TSG allele is not rate-limiting in a CIN cell.
- Because of the tremendous acceleration of loss of heterozygosity in CIN cells, it is very likely that most cancers, which require inactivation of at least two TSGs in rate-limiting scenarios, are initiated by CIN mutations, even if CIN has a severe cost in terms of somatic fitness.

Here, we discuss the basic aspects of evolutionary dynamics of oncogene activation and outline concepts such as SELECTION, FIXATION and RANDOM DRIFT.

Tissues of multicellular organisms are subdivided into compartments[36], which contain populations of cells that proliferate to fulfill organ-specific tasks. Compartments are subject to homeostatic mechanisms that ensure that the cell number remains approximately constant over time. Whenever a cell divides, another cell has to die to keep the total population size the same. Cancer results if the equilibrium between cell birth and cell death is shifted towards uncontrolled proliferation. Not all cells of a compartment, however, might be at risk of becoming neoplastic. Differentiated cells, for example, might not have the capacity to divide often enough to accumulate the necessary number of mutations in cancer-susceptibility genes[37]. The effective population size of a compartment describes those cells that are at risk of becoming cancer cells. In the following text, compartment size is used synonymously with effective population size within a compartment.

Consider a compartment of replicating cells. During each cell division, there is a small probability that a mistake will be made during DNA replication; in this case, a mutated daughter cell is produced. The mutation might confer a fitness advantage to the cell by ameliorating an existing function or inducing a new function. Then the cell proliferates more quickly or dies more slowly than its neighbors, and the mutation is advantageous in terms of somatic selection. Alternatively, the mutation might impair an important cellular function and confer a fitness disadvantage to the cell. Then the cell proliferates more slowly or dies more quickly than its neighbours. The net reproductive rate is decreased, and the mutation is deleterious in terms of somatic selection. Finally, the mutation might not change the reproductive rate of the cell. Then the cell proliferates at the same rate as its neighbours and the mutation is neutral in terms of somatic selection.

Now consider the dynamics of a particular mutation within a compartment. Initially, all cells are unmutated. What is the probability that a single mutated cell has arisen by time t? We measure time, t, in cell cycles. If the relevant cells divide once per day, then the unit of time is one day. Denote by N the number of cells in a compartment, and denote by u the mutation rate per gene per cell division. The probability that at least one mutated cell has arisen by time t is given by $P(t) = 1 - e^{-Nut}$ (FIG. 1a).

What is the fate of a single mutated cell? In the simplest scenario, there is a constant probability, q, that this cell will not die, but will initiate a neoplasia. The probability that a compartment has initiated a neoplasia by time t is given by $P(t) = 1 - e^{-Nuqt}$.

Alternatively, consider a scenario in which the mutated cell has a relative fitness r compared with a wild-type cell with fitness 1. If $r > 1$, the mutation is advantageous; if $r < 1$, the mutation is disadvantageous; if $r = 1$, the mutation is neutral. Normally, we expect mutations in oncogenes to cause increased net growth rates, $r > 1$; however, a mutation in an oncogene could be kept in check by apoptotic defence mechanisms, so r could be less than one.

What is the probability that such a mutation will not die out, but will take over the compartment (FIG. 1b)? To calculate this probability, a specific stochastic process known as the MORAN PROCESS can be considered[38]. At each time step, a cell is chosen for reproduction at random, but proportional to fitness. If there are i mutated cells, then the probability that a mutated cell is chosen for reproduction is $ri/(ri + N - i)$. The chosen cell produces a daughter cell that replaces another randomly chosen cell that dies. The total number of cells remains strictly constant. The probability that a single mutated cell with $r > 1$ or $r < 1$ takes over the whole compartment is given by $\rho = (1 - 1/r)/(1 - 1/r^N)$ (REF. 39). For a neutral mutant, $r = 1$, we have $\rho = 1/N$ (REF. 40). In this stochastic process, a cell can either produce a lineage that dies out or that takes over the whole compartment. Stable coexistence of different cell types is impossible. The quantity ρ is called fixation probability. An advantageous mutation has a higher fixation probability than a neutral mutation, which has a higher fixation probability than a deleterious mutation. The events in a small compartment, however, are dominated by random drift: if N is small, then even a deleterious mutation has a fairly high probability of reaching fixation due to chance events.

The probability that a mutation has been fixed by time t is given by $P(t) = 1 - e^{-Nu\rho t}$. Note that any mutation has a higher fixation probability in a small compartment than in a large compartment, but $P(t)$ is an increasing function

SELECTION
The process of survival of the fittest by which organisms that adapt to their environment survive and those that do not adapt disappear.

FIXATION
A state in which every individual in a population is identical with respect to a particular mutation.

RANDOM DRIFT
Changes in the genetic composition of a population due to probabilistic events.

MORAN PROCESS
Stochastic process that is used to describe the dynamics within a population with strictly constant population size.

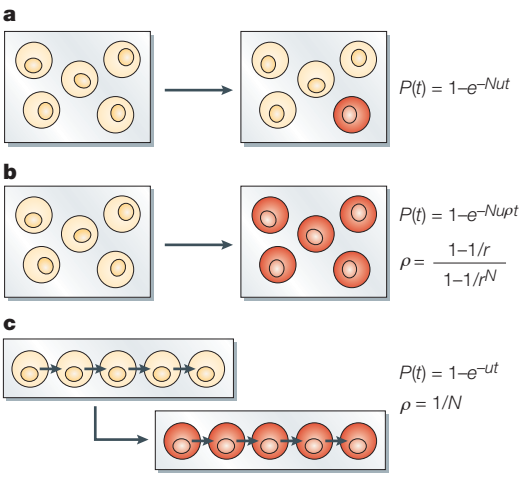

Figure 1 | **Oncogene dynamics. a** | The probability that at least one mutated cell has arisen in a compartment of N cells before time t is given by $P(t) = 1 - e^{-Nut}$. Here, u denotes the mutation rate per gene per cell division, and time is measured in units of cellular generations. **b** | The probability that a compartment of N cells has been taken over by mutated cells by time t is given by $P(t) = 1 - e^{-Nu\rho t}$. The probability that a single mutant cell with relative fitness r reaches fixation is given by $\rho = (1 - 1/r)/(1 - 1/r^N)$. **c** | N cells are arranged in a linear array. Whenever a cell divides, one daughter cell is put to its right, and all cells to the right are shifted by one place. The rightmost cell undergoes apoptosis. Here the fixation probability is $\rho = 1/N$ irrespective of r, because mutations have to arise in the leftmost cell (the stem cell) if they are not to be 'washed out'. The probability that the compartment consists only of mutated cells at time t is given by $P(t) = 1 - e^{-ut}$.

of N for $r > 1$ and a decreasing function of N for $r < 1$. So, large compartments accelerate the accumulation of advantageous mutations, but slow down the accumulation of deleterious mutations. Conversely, small compartments slow down the accumulation of advantageous mutations, but accelerate the accumulation of deleterious mutations. Therefore, the compartment size is important in determining the types of mutations that are likely to occur[41,42].

The linear process. So far, the evolutionary dynamics of a mutation that arises in a well-mixed compartment have been considered. This approach describes a tissue compartment in which all relevant cells are in equivalent positions and in direct reproductive competition with each other — there are no spatial effects. However, we can also envisage theories in which cellular differentiation and spatial structure are explicitly modelled. One simple approach considers N cells in a linear array[43] (FIG. 1c). Again, at each time step, a cell is chosen at random, but proportional to fitness. The cell is replaced by two daughter cells, and all cells to its right are shifted by one place to the right. The cell at the far right undergoes apoptosis; the cell at the far left acts as a STEM CELL. In this approach, the fixation probability of a mutant cell is given by $\rho = 1/N$. The probability of fixation does not depend on its relative fitness r because only a mutation in the far left cell can reach fixation in the compartment. A mutation arising in an offspring cell will eventually be 'washed out' of the compartment by the continuous production of cells and their migration from the stem cell to differentiation and apoptosis. The probability that all cells of the compartment are mutated at time t is given by $P(t) = 1 - e^{-ut}$. Here, time is measured in units of stem-cell divisions. If the stem cell divides more slowly than the other cells, then the accumulation of mutated cells is decelerated.

This 'linear process' of cancer initiation has the important feature of balancing out fitness differences between mutations[43]. Advantageous, deleterious and neutral mutations all have the same fixation probability, $\rho = 1/N$. This is in contrast to a well-mixed compartment, in which the fittest mutation has the highest probability of fixation. In comparison with a well-mixed compartment, a linear compartment delays the development of cancers that are initiated by advantageous mutations, such as mutations in oncogenes and TSGs. However, it can increase the probability of cancer initiation through deleterious mutations, such as mutations in genetic-instability genes[44]. Note also that the linear tissue design does not change the rate of accumulation of neutral mutations.

Numerical examples. First, suppose an organ consists of $M = 10^7$ compartments. This is, for example, the approximate number of colonic crypts. Suppose each compartment consists of $N = 1{,}000$ cells that divide once per day. The mutation rate per base per cell division is approximately 10^{-10} (REF. 45). Assume a particular oncogene can be activated by any one of ten mutations. So, the rate of activating the oncogene per cell division is $u = 10^{-9}$. Suppose the activation of the oncogene confers a 10% growth advantage to the cell, $r = 1.1$. Then the probability of fixation is $\rho = (1 - 1/r)/(1 - 1/r^N) = 0.09$. The probability that a compartment has been taken over by mutated cells at time $t = 70$ years is $P(t) = 1 - e^{-Nu\rho t} \approx 0.0023$. The expected number of mutated compartments at this age is $M \times P(t) \approx 23{,}000$.

Second, assume a linear tissue architecture. Each compartment consists of 1,000 cells, but is fed by one stem cell that divides every ten days[46,47]. Now, the probability that a compartment has been taken over by mutated cells at time $t = 70$ years is reduced to $P(t) \approx 2.6 \times 10^{-6}$. The expected number of mutated compartments at this age is 26.

Third, assume a population of $N = 10^7$ cells that divide every day. This population size describes, for example, a lesion that has already accumulated mutations in one or a few cancer-susceptibility genes; it does not, however, describe normal compartments in human tissues, as their population sizes are smaller. The probability that an oncogene leading to a relative fitness of $r = 1.1$ is activated within the next $t = 1$ year within a compartment is given by $P(t) \approx 0.28$. The time until the probability of activating the oncogene is one-half is obtained as $T_{1/2} = 2.1$ years.

STEM CELL
A precursor cell that can self renew and undergo clonal, multilineage differentiation.

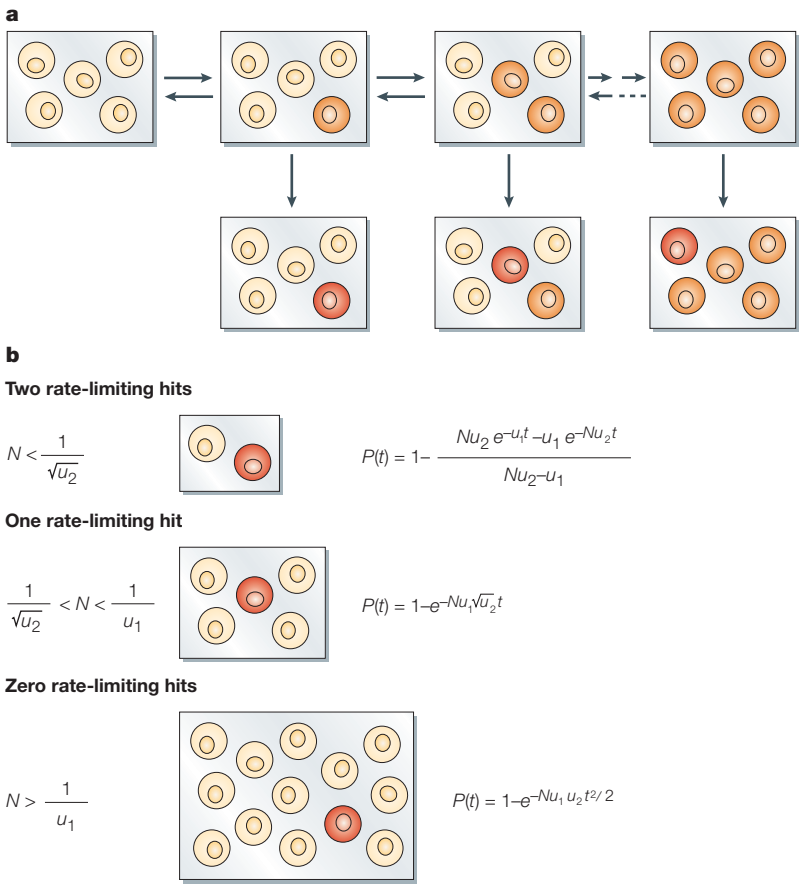

Figure 2 | **Tumour-suppressor gene dynamics. a** | Initially all cells are unmutated (yellow). A cell with one inactivated allele (orange) arises. This cell gives rise to a lineage that can take over the whole compartment or become extinct again. In one of those cells, the second allele of the tumour-suppressor gene (TSG) might become inactivated (red). **b** | TSG inactivation is described by three dynamic laws that depend on the population size, N, and the probabilities of inactivating the first and the second allele, u_1 and u_2, respectively, per cell division. Time, t, is measured in cellular generations. In small, intermediate and large populations, a TSG is inactivated by two, one and zero rate-limiting steps, respectively.

Conclusions. Tissue architecture and the scale of somatic selection greatly influence the rate at which mutations accumulate. Mutations that activate oncogenes are thought to confer a selective advantage to the cell and are best contained by small compartments. The mutant cell is likely to reach fixation in the compartment, but its further spread is at least initially limited by the compartment boundaries. The risk of accumulating such mutations is reduced by adopting a linear tissue architecture in which the compartment is replenished by one or a few stem cells. Once one or several mutations have induced a neoplasia, however, additional mutations accumulate faster because of the increased population size.

The mathematical analysis of the dynamics of oncogene activation indicates several interesting avenues that could be experimentally investigated. What are the relative growth rates of cells bearing specific mutations in particular oncogenes compared with wild-type cells? What is the average time for the emergence and spread of an oncogenic mutation in a culture of dividing cells or in an experimental mouse tumour? How are human tissues organized? How large is the effective population size of a compartment within a particular tissue, and is it replenished asymmetrically by one or a few stem cells?

MITOTIC RECOMBINATION
The exchange — reciprocal or nonreciprocal — of genetic material between one DNA molecule and a homologous region of DNA that occurs during mitotic cell divisions.

Tumour-suppressor genes

So far, the discussion has focused on genes that confer an altered phenotype to the cell when mutated in one allele, and that are dominant at the cellular level. Some genes, however, are recessive at the cellular level, and have to be mutated in both alleles to cause a phenotypic change of the cell. Examples of recessive mutations are those that inactivate some TSGs[1,48,49].

The concept of a TSG emerged from a statistical analysis of retinoblastoma incidence in children[9]. This study and subsequent work[1,10,50] led to Knudson's two-hit hypothesis, which proposes that two hits in the *RB* gene are the rate-limiting steps of retinoblastoma. In the inherited form, the first mutation is present in the germline, whereas the second mutation emerges during somatic cell divisions. In the sporadic form, both mutations arise during somatic cell divisions. A large number of TSGs have since been discovered that function in apoptosis, cell senescence and other signalling pathways[1].

A normal cell has two alleles of a TSG. The mutation that inactivates the first allele can be neutral, disadvantageous or advantageous, and a cell with one inactivated allele correspondingly has a normal, decreased or increased net reproductive rate. The first hit is neutral if the TSG is strictly recessive; that is, if the remaining wild-type allele has sufficient tumour-suppressive function. The first hit is disadvantageous if the TSG is checked by apoptotic defence mechanisms; that is, if as soon as surveillance mechanisms discover an imbalance in the TSG product, apoptosis is triggered. The first hit is advantageous if the TSG is haploinsufficient; that is, if the remaining wild-type allele has insufficient tumour-suppressive function. Here, TSGs with a neutral first hit are considered. The mutation that inactivates the second allele is advantageous, and a cell with two inactivated alleles has an increased net reproductive rate. The mutation rates for the first and the second hit are denoted by u_1 and u_2, respectively. We assume $u_1 < u_2$ because some mutational mechanisms, such as MITOTIC RECOMBINATION, can only constitute the second hit.

What is the probability that a single cell with two inactivated TSG alleles has arisen by time t in a population of N cells[39,51,52]? Interestingly, the answer depends on the population size, N, as compared with the mutation rates that constitute the first and second hit, u_1 and u_2. There are three different cases (FIG. 2).

First, in small populations, $N < 1/\sqrt{u_2}$, a cell with one inactivated allele reaches fixation in the population before a cell with two inactivated alleles arises. The probability that at least one cell with two hits emerges before time t is given by equation 1:

$$P(t) = 1 - \frac{Nu_2\, e^{-u_1 t} - u_1 e^{-Nu_2 t}}{Nu_2 - u_1} \qquad (1)$$

For very short times, $t < 1/Nu_2$, we can approximate $P(t) \approx Nu_1u_2t^2/2$. Therefore, this probability accumulates as a second order of time: it takes two rate-limiting hits to inactivate a TSG in a small population of cells.

Second, in populations of intermediate size, $1/\sqrt{u_2} < N < 1/u_1$, a cell with two inactivated alleles emerges before a cell clone with one inactivated allele has taken over the population. The population 'tunnels' from a wild-type phenotype directly to the second hit without ever having fixed the first hit[39,51]. The probability that at least one cell with two hits has arisen before time t is given by equation 2:

$$P(t) = 1 - e^{-Nu_1\sqrt{u_2}t} \quad (2)$$

This probability accumulates as a first order of time: it takes only one rate-limiting hit to inactivate a TSG in a population of intermediate size.

Third, in very large populations, $N > 1/u_1$, cells with one inactivated allele arise immediately and the waiting time for a cell with two inactivated alleles dominates the dynamics. The probability that at least one cell with two hits has arisen before time t is given by equation 3:

$$P(t) = 1 - e^{-Nu_1u_2t^2/2} \quad (3)$$

This probability again accumulates as a second order of time. However, eliminating a TSG in a large population of cells is not rate limiting for the overall process of tumorigenesis, as mutated cells are constantly being produced (FIG. 2b).

These three dynamic laws provide a complete description of TSG inactivation. In a normal tissue consisting of small compartments of cells, a TSG is eliminated by two rate-limiting hits. The overall rate of inactivation is proportional to the second order of time. In small neoplasias, only one rate-limiting hit is needed to inactivate a TSG. The rate of inactivation is proportional to the first order of time. In large tumours, it again takes two hits to inactivate a TSG, but neither of them is rate limiting for the overall process of tumorigenesis. Therefore, as the population size increases, a TSG is inactivated in two, one or zero rate-limiting steps.

Numerical examples. Suppose an organ consists of 10^7 compartments. Each compartment contains about 1,000 cells, but is replenished by only $N = 4$ stem cells that divide once per week. Assume a TSG is 10kb long and can be inactivated by a point mutation in any one of 500 bases, occurring with probability 10^{-10} per cell division[45]. Then, the rate of inactivation of the first TSG allele is $u_1 = 10^{-7}$ because the mutation can inactivate either of the two TSG alleles. Suppose that the rate of inactivation of the second TSG allele, including mitotic recombination, chromosome NON-DISJUNCTION and other mechanisms of LOSS OF HETEROZYGOSITY (LOH), is $u_2 = 10^{-6}$. Then, using equation 1, the probability that a cell with two inactivated TSG alleles has arisen in a compartment at time $t = 70$ years is $P(t) \approx 2.7 \times 10^{-6}$. The expected number of compartments containing at least one cell with two hits at time $t = 70$ years is 27.

Now consider a small lesion of $N = 10^4$ cells that divide once per day. What is the probability that a TSG is inactivated within the next twenty years? Again let $u_1 = 10^{-7}$ and $u_2 = 10^{-6}$. Using equation 2 we find $P \approx 0.0073$. What is the time, $T_{1/2}$, until the probability of having produced a cell with an inactivated TSG is one-half? We find $T_{1/2} \approx 1{,}900$ years. However, if a tumour contains $N = 10^9$ cells that divide once per day, the half life of the TSG reduces to $T_{1/2} \approx 120$ days.

Conclusions. The dynamics of TSG inactivation are described by three laws that depend on the population size and the mutation rates that cause the first and the second hit. In a small compartment, a TSG is inactivated by two rate-limiting hits. In a small lesion, it is inactivated by one, and in a large tumour, by zero rate-limiting hits.

Mutations inactivating TSGs, like mutations activating oncogenes, are best contained by small compartments that are replenished by a small number of stem cells. Once a compartment has accumulated one or a few mutations in cancer-susceptibility genes, a neoplasia develops in which other TSGs might have to be inactivated for further tumour progression. Small neoplasias, however, are unlikely to succeed in inactivating a further TSG within a human lifespan, under the assumption of normal mutation rates. Increased mutation rates due to genetic instability might be necessary for the further progression of some small lesions.

The mathematical analysis of TSG inactivation indicates several new experimental studies. How does the inactivation of one or two alleles change the net growth rate of the cell? Is the first step indeed neutral or does it slightly modify the fitness of the cell? Such fitness differences can be measured in cell cultures or animal models. Furthermore, how long does it take for a population of cells to inactivate a TSG? How does the time until inactivation of a TSG depend on the population size and the mutation rates? Can the three dynamical regimens that we predict from the theoretical analysis be verified? What are the relative rates of the various mechanisms that contribute to LOH? Precise kinetic measurements are needed to obtain quantitative insights into cancer progression.

Genetic instability

Genetic instability is a defining characteristic of most human cancers and one of the most active research areas in cancer biology[3]. Two key types of genetic instabilities have been identified[53]. In a small fraction of colorectal and some other cancers, a defect in MISMATCH REPAIR results in an increased rate of point mutations and consequent widespread MICROSATELLITE INSTABILITY. Almost all colorectal and most other cancers, however, have chromosomal instability (CIN), which refers to an increased rate of loss or gain of 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 LOH, which is an important property of CIN because it accelerates the rate of TSG inactivation.

NON-DISJUNCTION
An error in cell division in which the chromosomes fail to disjoin, so that both pass to the same daughter cell.

LOSS OF HETEROZYGOSITY
At a particular locus that is heterozygous for a mutant allele and a normal allele, a deletion or other mutational event within the normal allele renders the cell either hemizygous (one mutant allele and one deleted allele) or homozygous for the mutant allele.

MISMATCH REPAIR
A DNA-repair mechanism that corrects nucleotide sequence errors made during DNA replication by excising the defective sequence and replacing it with the correct sequence.

MICROSATELLITE INSTABILITY
Genetic instability because of mismatch-repair deficiency involving subtle sequence changes that alter one or a few base pairs.

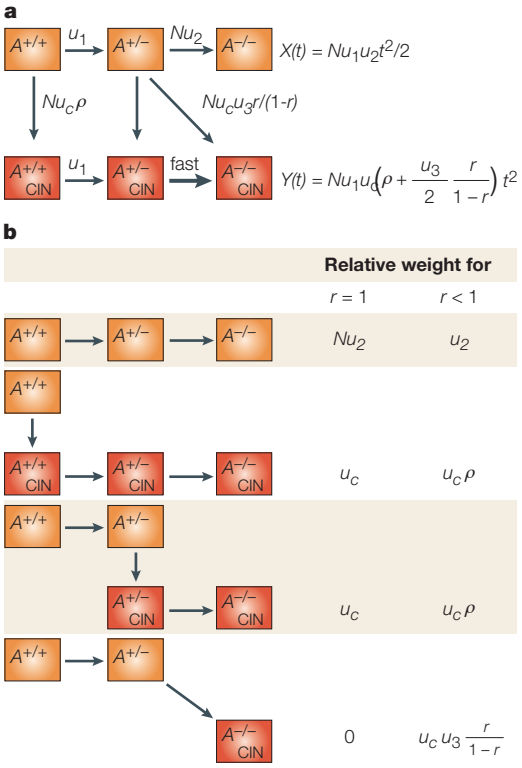

Figure 3 | **Emergence of chromosomal instability during inactivation of one tumour-suppressor gene.** **a** | Inactivation of a tumour-suppressor gene (TSG), *A*, in a compartment of *N* cells requires mutation of the first and the second TSG allele, occurring with probabilities u_1 and u_2, respectively, per cell division. The probability of mutating a chromosomal-instability (CIN) gene is given by u_c. A CIN cell has relative fitness *r* and reaches fixation in the compartment with probability $\rho = (1 - 1/r)\,(1 - 1/r^N)$. CIN increases the rate of inactivating the second TSG allele to u_3. If an $A^{+/-}$ cell clone without CIN produces an $A^{-/-}$ cell with CIN before taking over the compartment, a tunnel arises (diagonal arrow). The probabilities that the compartment is in state $A^{-/-}$ without and with CIN at time *t* are given by *X*(*t*) and *Y*(*t*), respectively. **b** | The compartment can evolve along three evolutionary trajectories, all of which contain two rate-limiting hits. In the first trajectory, the TSG can be inactivated without CIN. In the second trajectory, CIN can either arise before the inactivation of the first TSG allele or it can arise between the inactivation of the first and the second TSG allele. The third trajectory contains a tunnel if CIN has a substantial cost. The relative weights of each trajectory for neutral CIN, r ≈ 1, and costly CIN, r < 1, allow us to calculate their respective contributions to tumorigenesis.

The molecular basis for CIN is just beginning to be understood. A large number of genes that trigger CIN when mutated have been discovered in *Saccharomyces cerevisiae*[54–56]. These so-called 'CIN genes' are involved in chromosome condensation, sister-chromatid cohesion, kinetochore structure and function, microtubule formation and cell-cycle checkpoints. By comparison with yeast, we expect that there are several hundred human CIN genes, but only a few have been identified so far. These genes include *BUB1*, *MAD2*, *BRCA2*, and *CDC4* (REF. 57). The classification of CIN genes is based on the mutational events that are required to trigger CIN[3]. 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 *BUB1*, trigger CIN if one allele is mutated in a DOMINANT-NEGATIVE fashion. Class III CIN genes, such as *BRCA2*, trigger CIN if both alleles are mutated.

Several hereditary syndromes are known that stem from germline mutations in what might be CIN genes. Inherited mutations in the genes that encode the RECQ-like helicases BLM AND WRN give rise to the Bloom and Werner Syndromes, respectively. An inherited deficiency in NUCLEOTIDE-EXCISION REPAIR causes Xeroderma pigmentosum. Germline mutations in *ATM* give rise to ataxia telangiectasia. These syndromes are all characterized by a high incidence of several types of cancer, but the mechanistic connection between these genes and CIN is still somewhat unclear.

An important question in cancer genetics is to what extent CIN, or any genetic instability, is an early event and driving force of tumorigenesis[58–60]. The investigation of the role of genetic instability requires a quantitative theory of how mutation and selection of gatekeeper and caretaker genes contribute to cancer initiation and progression[52,61–63]. Here, we review studies of the role of CIN in cancers that are initiated by inactivation of one or two TSGs[52,64].

CIN before one TSG. Consider a case in which tumorigenesis is initiated by inactivating a TSG, *A*, in a small compartment of cells[52]. An appropriate example is the inactivation of adenomatous polyposis coli (APC) in a colonic crypt[1]. If the mutation rate is less than the inverse of the compartment size, then the compartment almost always consists of a single type of cells — a mutated cell either reaches fixation or becomes extinct before another mutated cell arises. Initially, all cells are wild type, $A^{+/+}$. The compartment evolves from $A^{+/+}$, through $A^{+/-}$, to $A^{-/-}$. A mutation that triggers CIN can arise at any stage of this process. The crucial effect of CIN is to increase the rate of LOH, thereby accelerating the transition from $A^{+/-}$ to $A^{-/-}$. Stochastic tunnelling can lead from $A^{+/-}$ without CIN directly to $A^{-/-}$ with CIN[51].

The rates of evolution of TSG inactivation with or without CIN in a small compartment of cells are shown in FIG. 3a. There are two rate-limiting hits for inactivating the TSG without CIN. Interestingly, there are also two rate-limiting hits for inactivating the TSG with CIN[52]: one rate-limiting step is needed for inactivating the first allele of the TSG, another rate-limiting step is needed for the CIN mutation. The inactivation of the second TSG allele is greatly accelerated in the presence of CIN and is therefore not rate limiting[52].

There are three evolutionary trajectories, all of which contain two rate-limiting steps (FIG. 3b): the TSG can be inactivated without CIN; the CIN mutation can occur first, followed by the inactivation of the two TSG alleles; one TSG allele can be mutated first, followed by a CIN mutation, followed by the inactivation of the second TSG allele. The third trajectory contains a tunnel if there is a significant cost of CIN; that is, if a CIN cell has a high probability of undergoing apoptosis because of deleterious and lethal mutations. Such a cell

BUB1 AND MAD2
Their gene products act cooperatively to prevent unequal sister chromatid separation by inhibiting the anaphase-promoting complex.

BRCA2
Its gene product is implicated in DNA repair and recombination, and checkpoint control of the cell cycle. In mice, its loss might result in chromosomal instability.

DOMINANT-NEGATIVE MUTATION
A mutation whose gene product adversely affects the wild-type gene product within the same cell, often by dimerizing with it.

BLM AND WRN
Their gene products participate in DNA-repair pathways, particularly those that repair double-strand DNA breaks, and their loss results in genetic instability.

NUCLEOTIDE-EXCISION REPAIR
A DNA-repair mechanism that excises and replaces damaged DNA bases.

ATM
Its gene product functions in X-ray-induced DNA-repair mechanisms.

APC
A tumour-suppressor that is thought to initiate colorectal tumorigenesis. Mutation of *APC* leads to increased β-catenin-mediated transcription of growth-promoting genes.

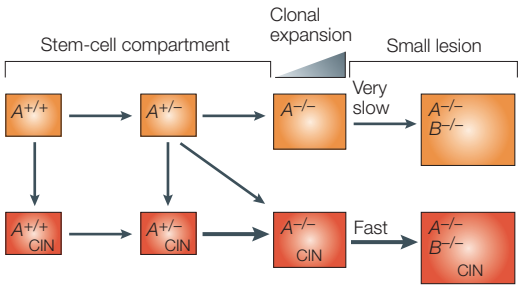

Figure 4 | **Emergence of chromosomal instability during inactivation of two tumour-suppressor genes.** Inactivation of tumour-suppressor gene (TSG) *A* in a compartment of N_0 cells causes a clonal expansion to N_1 cells in which TSG *B* is inactivated. Because of the increased compartment size, the evolutionary trajectory tunnels from $A^{-/-}B^{+/+}$ directly to $A^{-/-}B^{-/-}$. Chromosomal instability (CIN) can arise at any stage of tumorigenesis and increases the rate of TSG inactivation. There are three rate-limiting hits both with and without CIN.

might not reach fixation in the compartment, but might nevertheless produce a cell with two inactivated TSG alleles that reaches fixation. The relative rates of these trajectories depend on the compartment size, *N*, the mutation rates, the number of CIN genes, and the cost of CIN (FIG. 3b). We can use these rates to calculate the fraction of mutated compartments with or without CIN. We can also estimate the minimum number of CIN genes in the genome that are required to ensure that CIN precedes the inactivation of the first or second TSG allele[52].

Numerical examples of CIN before one TSG. Again, consider an organ that consists of 10^7 compartments. Suppose each compartment contains 1,000 cells, but is replenished by a small pool of $N = 4$ stem cells. Assume that the probabilities of inactivating the first and the second allele of a TSG per cell division are $u_1 = 10^{-7}$ and $u_2 = 10^{-6}$, respectively. We assume that the inactivation of the TSG confers a big selective advantage to the cell, such that the probability of fixation is 1.

Assume that there is one class I CIN gene in the genome that does not change the fitness of the cell when mutated, $r = 1$. What is the ratio of the probabilities of inactivating the TSG with CIN versus without CIN? We find that 50% of all mutated compartments have inactivated the TSG in a cell with CIN. In 25%, the CIN mutation occurred before the first TSG mutation, whereas in another 25%, the CIN mutation occurred between the first and the second TSG mutation. In 50% of mutated compartments, there was no CIN mutation before the inactivation of the TSG. These results are largely independent of time.

Instead, assume that there are two class I CIN genes in the genome, each of which is sufficient to trigger CIN when mutated. Mutation of either gene reduces the fitness of the cell to $r = 0.7$. In this case, 54% of all mutated compartments have inactivated the TSG in a cell with CIN. In 25%, the CIN mutation occurred before the first TSG mutation. In another 25%, the CIN mutation occurred between the first and the second TSG mutation without a tunnel, and in 4% with a tunnel. In 46% of mutated compartments, there was no CIN mutation before the inactivation of the TSG.

Alternatively, we can calculate the crucial number of CIN genes in the genome that are needed to ensure that CIN arises before the inactivation of the TSG. If CIN is neutral, $r = 1$, then 2 class I CIN gene or 12 class II CIN genes in the genome are needed. If CIN has a selective disadvantage of $r = 0.7$, then two class I CIN genes or 21 class II CIN genes in the genome are needed to make sure that CIN arises first. Note again that in yeast, more than one hundred CIN genes are known, and it is possible that an even larger number exists in the human genome.

CIN in inherited cancers. In some genetic diseases, one allele of a TSG is mutated in the germline. For example, patients with familial adenomatous polyposis inherit a mutation in one allele of the TSG *APC*. By their teens, they harbour hundreds to thousands of colorectal polyps. In terms of the model, the mutational path starts with a compartment that consists only of $A^{+/-}$ cells without CIN. There are two evolutionary trajectories: the somatic mutation that inactivates the second TSG allele occurs in a cell without CIN, or a CIN mutation precedes the inactivation of the second TSG allele. Both trajectories require one rate-limiting step. Again, we can calculate the relative rates of these two possibilities. We can also calculate the crucial number of CIN genes in the genome that are needed to ensure that CIN arises before the inactivation of the second allele. If CIN has a negligible cost, then twice as many CIN genes as before are needed to make sure that CIN comes first. The factor two comes from the fact that during the inactivation of two TSG alleles, CIN can arise either before the first or before the second hit. If the first hit is already present in the germline, CIN can arise only before the second hit. This reduces the probability that CIN arises before the inactivation of the TSG by a factor of one-half. If CIN has a substantial cost, then the same number of CIN genes is needed as in the sporadic case with neutral CIN. So, there is only a small difference in the relative importance of CIN in those cases in which both TSG alleles have to be inactivated somatically and in which one TSG allele is already inactivated in the germline.

CIN before two TSGs. Consider a path to cancer in which two TSGs, *A* and *B*, have to be eliminated in rate-limiting steps. Initially, the compartment consists of N_0 wild-type cells, $A^{+/+}B^{+/+}$, that divide every τ_0 days. Suppose gene *A* has to be inactivated first, so the evolutionary pathway proceeds from $A^{+/+}B^{+/+}$ via $A^{+/-}B^{+/+}$ to $A^{-/-}B^{+/+}$, and subsequently to $A^{-/-}B^{+/-}$ and $A^{-/-}B^{-/-}$ (FIG. 4). CIN can emerge at any stage of this pathway; once arisen, CIN accelerates the transitions from $A^{+/-}$ to $A^{-/-}$ and from $B^{+/-}$ to $B^{-/-}$. Inactivation of the first TSG induces neoplastic growth. We assume that the $A^{-/-}$ compartment gives rise to a small lesion of N_1 cells that divide every τ_1 days. In this lesion, the second TSG has to be inactivated for further

CIN before one TSG				CIN before two TSGs					
Independent of N_1				$N_1 = 10^4$			$N_1 = 10^5$		
r	n_1	n_2	n_3	n_1	n_2	n_3	n_1	n_2	n_3
1.0	2	12	89	1	1	16	1	2	37
0.9	2	14	96	1	1	17	1	3	41
0.7	2	21	119	1	1	23	1	4	51
0.5	4	42	169	1	2	35	1	8	74

Figure 5 | **Minimum number of chromosomal-instability (CIN) genes needed to ensure that CIN arises before the inactivation of one or two tumour-suppressor genes**. If chromosomal instability (CIN) emerges before inactivation of two tumour-suppressor genes (TSGs), it must arise before inactivation of the first TSG. The selective fitness of CIN cells is denoted by r. 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. Class III CIN genes, n_3, trigger CIN if both alleles are mutated. Parameter values are $u_1 = 10^{-7}$, $u_2 = 10^{-6}$, $u_3 = 10^{-2}$, $N_0 = 4$, and $t = 80$ years.

tumour progression. Because of the increased compartment size, the evolutionary trajectory tunnels[51] from $A^{-/-}B^{+/+}$ directly to $A^{-/-}B^{-/-}$ (FIG. 4).

Without CIN, inactivation of two TSGs requires three rate-limiting hits. It takes two hits to inactivate the first TSG — if this leads to a moderate clonal expansion, then the second TSG can be inactivated in one rate-limiting step (equation 2). If the inactivation of the first TSG leads to a vast clonal expansion, then the inactivation of a further TSG is not rate limiting and CIN becomes less important. With CIN, inactivation of two TSGs also requires three rate-limiting hits. It takes two hits to inactivate the first TSG (one in the first allele and one in a CIN gene), and one further hit to inactivate the second allele.

Again, we can calculate the minimum number of CIN genes in the genome that are needed to ensure that CIN arises before the inactivation of the first TSG in paths to cancer in which two TSGs have to be eliminated in rate-limiting steps. We find that very few CIN genes in the genome are necessary to make sure that CIN arises early (FIG. 5). The cost of CIN is compensated by an acceleration of every successive TSG inactivation. It is possible that the first TSG, *A*, is predominantly inactivated in cells without CIN. So, most lesions that are caused by inactivation of TSG *A* would not have CIN, but only the small fraction of lesions with CIN will eliminate TSG *B* within the timescale of a human life. In such a situation, all (or almost all) cancers will derive from lesions in which a CIN mutation preceded inactivation of the first TSG.

Conclusions. The role of CIN in tumorigenesis is one of the most interesting topics of all of biology. To evaluate hypotheses and interpret experimental observations, a quantitative model is required for understanding how tissue organization, mutation rates and selection determine the role of CIN in tumorigenesis. Here we have outlined a mathematical approach for evaluating the importance of early CIN in tumorigenesis that is initiated by inactivation of 'half', one or two TSGs in rate-limiting situations. We find that one or a few neutral CIN genes in the genome are sufficient to ensure the emergence of CIN before the inactivation of one TSG. One or a few costly CIN genes in the genome are sufficient to ensure the emergence of CIN before the inactivation of the first TSG if two TSGs have to be eliminated in rate-limiting steps.

The theoretical investigation of CIN indicates several new experiments. Can one demonstrate in mouse that the two rate-limiting steps consist of inactivating one allele of the TSG and one CIN mutation? What is the selective cost of a CIN mutation? Can CIN be neutral or even advantageous[65]? Is it possible to find CIN mutations in dysplastic crypts, which represent the first stage of colon cancer[57]? What is the number of CIN genes in the human genome? How many of them fall into classes I, II or III? Is it possible to show that CIN is essential for inactivating two TSGs in small populations of cells?

Implications and future directions

In this article, we have outlined some fundamental principles of somatic evolutionary dynamics. We have discussed the rates at which mutations in oncogenes and TSGs accumulate in small compartments of cells, early lesions and large cancers. We have explained how the relevant timescales depend on population size, mutation rates and fitness differences. We have studied the role of tissue architecture and genetic instability.

Tissue architecture and compartment size determine the rates at which different types of mutations accumulate. Cells with mutations in oncogenes or TSGs can have an increased somatic fitness. Such mutations are best contained when the tissue is organized into small compartments and each compartment is fed by one (or a few) stem cell, as in the linear process[43]. Cells with mutations in genetic-instability genes are likely to have a reduced somatic fitness. Such mutations accumulate faster in small compartments than in large compartments. Therefore, the optimum compartment size is a trade-off between preventing mutations in oncogenes, TSGs and genetic-instability genes[41].

CIN confers an increased probability of gaining or losing whole chromosomes or arms of chromosomes. CIN accelerates the rate of inactivating TSGs. A key debate in cancer genetics is to what extent CIN is an early event and, therefore, a driving force of cancer progression. We have shown that in a small compartment of cells, it takes two hits to inactivate a TSG both with and without CIN. Whether or not CIN emerges before the first TSG depends on the cost of CIN, the mutation rate, the rate of LOH, the number of CIN genes in the genome and the population size. For a wide range of parameters, one or a few neutral CIN genes in the genome are enough to ensure that CIN initiates tumour formation in a pathway in which one TSG needs to be eliminated in a rate-limiting situation. One or a few costly CIN genes in

the genome are enough to ensure that CIN initiates tumour formation in a pathway in which two TSGs need to be eliminated in rate-limiting situations.

We have reviewed models of somatic evolution with constant selection, which are based on the assumption that the fitness of mutants neither depends on their own relative abundance nor on the relative abundance of other mutants. Cancer progression, however, might often include a complex interplay of various types of cells. Mutations in landscaper genes, alterations promoting angiogenesis and mutations enabling the cell to fight off the immune system might be examples. These complex dependencies require a thorough theoretical analysis[66], and we are still only at the beginning of a long endeavor.

Mathematical models of cancer evolution are essential for quantifying the effects of mutation, selection and tissue architecture. The laws of dynamics of TSG inactivation, oncogene activation and emergence of genetic instability provide fundamental new insights into tumorigenesis. Theory can help to clarify concepts, interpret experimental data and indicate new experiments. Mathematical models describing tumour initiation, progression, invasion and metastasis should continue to contribute towards a quantitative understanding of cancer biology.

1. Vogelstein, B. & Kinzler, K. W. *The Genetic Basis of Human Cancer* (McGraw–Hill, 1998).
2. Kinzler, K. W. & Vogelstein, B. Gatekeepers and caretakers. *Nature* **386**, 761–763 (1997).
3. Rajagopalan, H., Nowak, M. A., Vogelstein, B. & Lengauer, C. The significance of unstable chromosomes in colorectal cancer. *Nature Rev. Cancer* **3**, 695–701 (2003).
4. Sieber, O. M., Heinimann, K. & Tomlinson, I. P. Genomic instability — the engine of tumorigenesis? *Nature Rev. Cancer* **3**, 701–708 (2003).
5. Bissell, M. J. & Radisky, D. Putting tumors in context. *Nature Rev. Cancer* **1**, 46–54 (2001).
6. Boveri, T. *Zur Frage der Entstehung maligner Tumoren* (Gustav Fischer, 1914). English translation: The origin of malignant tumors by Boveri, M. (Williams and Wilkins, 1929).
7. Muller, H. J. Artificial transmutation of the gene. *Science* **46**, 84–87 (1927).
8. Knudson, A. G. Two genetic hits to cancer. *Nature Rev. Cancer* **1**, 157–162 (2001).
9. Knudson, A. G. Mutation and cancer: statistical study of retinoblastoma. *Proc. Natl Acad. Sci. USA* **68**, 820–823 (1971).
10. Moolgavkar, S. H. & Knudson, A. G. Mutation and cancer: a model for human carcinogenesis. *J. Natl Cancer Inst.* **66**, 1037–1052 (1981).
11. Cairns, J. Mutation selection and the natural history of cancer. *Nature* **255**, 197–200 (1975).
12. Cairns, J. Somatic stem cells and the kinetics of mutagenesis and carcinogenesis. *Proc. Natl Acad. Sci. USA* **99**, 10567–10570 (2002).
13. Frank, S. A., Iwasa, Y. & Nowak, M. A. Patterns of cell division and the risk of cancer. *Genetics* **163**, 1527–1532 (2003).
14. Frank, S. A. & Nowak, M. A. Cell biology: developmental predisposition to cancer. *Nature* **422**, 494 (2003).
15. Gatenby, R. A & Maini, P. K. Mathematical oncology: cancer summed up. *Nature* **421**, 321 (2003).
16. Nordling, C. O. A new theory on cancer-inducing mechanism. *Br. J. Cancer* **7**, 68–72 (1953).
17. Armitage, P. & Doll, R. The age distribution of cancer and a multi-stage theory of carcinogenesis. *Br. J. Cancer* **8**, 1–12 (1954).
18. Armitage, P. & Doll, R. A two-stage theory of carcinogenesis in relation to the age distribution of human cancer. *Br. J. Cancer* **11**, 161–169 (1957).
19. Fisher, J. C. Multiple-mutation theory of carcinogenesis. *Nature* **181**, 651–652 (1959).
20. Nunney L. Lineage selection and the evolution of multistage carcinogenesis. *Proc. Roy. Soc. Lond. B Biol. Sci.* **266**, 493–498 (1999).
21. Tomlinson, I., Sasieni, P. & Bodmer, W. How many mutations in a cancer? *Am. J. Pathol.* **160**, 755–758 (2002).
22. Moolgavkar, S. H. & Knudson, A. G. Mutation and cancer: a model for human carcinogenesis. *J. Natl Cancer Inst.* **66**, 1037–1052 (1981).
23. Luebeck, E. G. & Moolgavkar, S. H. Multistage carcinogenesis and the incidence of colorectal cancer. *Proc. Natl Acad. Sci. USA* **99**, 15095–15100 (2002).
24. Grist, S. A., McCarron, M., Kutlaca, A., Turner, D. R. & Morley, A. A. *In vivo* human somatic mutation: frequency and spectrum with age. *Mutat. Res.* **266**, 189–196 (1992).
25. Gatenby, R. A. & Vincent, T. L. An evolutionary model of carcinogenesis. *Cancer Res.* **63**, 6212–6220 (2003).
26. Goldie, J. H. & Coldman, A. J. A mathematic model for relating the drug sensitivity of tumors to their spontaneous mutation rate. *Cancer Treat. Rep.* **63**, 1727–1733 (1979).
27. Goldie, J. H & Coldman, A. J. Quantitative model for multiple levels of drug resistance in clinical tumors. *Cancer Treat. Rep.* **67**, 923–931 (1983).
28. Anderson, A. R. & Chaplain, M. A. Continuous and discrete mathematical models of tumor-induced angiogenesis. *Bull. Math. Biol.* **60**, 857–899 (1998).
29. Owen, M. R. & Sherratt, J. A. Mathematical modeling of macrophage dynamics in tumors. *Math. Models Methods Appl. Biol. Chem.* **377**, 675–684 (1999).
30. Chang, S., Khoo, C. & DePinho, R. A. Modeling chromosomal instability and epithelial carcinogenesis in the telomerase-deficient mouse. *Semin. Cancer Biol.* **11**, 227–239 (2001).
31. Chang, S., Khoo, C., Naylor, M. L., Maser, R. S. & DePinho, R. A. Telomere-based crisis: functional differences between telomerase activation and ALT in tumor progression. *Genes Dev.* **17**, 88–100 (2003).
32. Maser, R. S. & DePinho, R. A. Connecting chromosomes, crisis, and cancer. *Science* **297**, 565–569 (2002).
33. Taddei, F. *et al.* Role of mutator alleles in adaptive evolution. *Nature* **387**, 700–702 (1997).
34. Otsuka, K. *et al.* Analysis of the human *APC* mutation spectrum in a *saccharomyces cerevisiae* strain with a mismatch repair defect. *Int. J. Cancer* **103**, 624–630 (2003).
35. Strauss, B. S. Hypermutability in carcinogenesis. *Genetics* **148**, 1619–1626 (1998).
36. Mintz, B. Clonal basis of mammalian differentiation. *Symp. Soc. Exp. Biol.* **25**, 345–370 (1971).
37. Michor, F., Nowak, M. A., Frank, S. A. & Iwasa, Y. Stochastic elimination of cancer cells. *Proc. Roy. Soc. Lond. B Biol. Sci.* **270**, 2017–2024 (2003).
38. Moran, P. *The Statistical Processes of Evolutionary Theory* (Clarendon Press, 1962).
39. Komarova, N. L., Sengupta, A. & Nowak, M. A. Mutation-selection networks of cancer initiation: tumor suppressor genes and chromosomal instability. *Jour. Theor Biol.* **223**, 433–450 (2003).
40. Kimura, M. The role of compensatory neutral mutations in molecular evolution. *J. Genet.* **64**, 7–19 (1985).
41. Michor, F., Iwasa, Y., Komarova, N. L. & Nowak, M. A. Local regulation of homeostasis favors chromosomal instability. *Curr. Biol.* **13**, 581–584 (2003).
42. Michor, F., Frank, S. A., May, R. M., Iwasa, Y. & Nowak, M. A. Somatic selection for and against cancer. *J. Theor. Biol.* **225**, 337–382 (2003).
43. Nowak, M. A., Michor, F. & Iwasa, Y. The linear process of somatic evolution. *Proc. Natl Acad. Sci. USA* **100**, 14966–14969 (2003).
44. Michor, F., Iwasa, Y., Rajagopalan, H., Lengauer, C. & Nowak, M. A. Linear model of colon cancer initiation. *Cell Cycle* (in the press).
45. Kunkel, T. A. & Bebenek, K. DNA replication fidelity. *Annu. Rev. Biochem.* **69**, 497–529 (2000).
46. Bach, S. P., Renehan, A. G. & Potten, C. S. Stem cells: the intestinal stem cell as a paradigm. *Carcinogenesis* **21**, 469–476 (2000).
47. Yatabe, Y., Tavare, S. & Shibata, D. Investigating stem cells in human colon by using methylation patterns. *Proc. Natl Acad. Sci. USA* **98**, 10839–10844 (2001).
48. Weinberg, R. A. Tumor suppressor genes. *Science* **254**, 1138–1146 (1991).
49. Levine, A. J. The tumor suppressor genes. *Annu. Rev. Biochem.* **62**, 623–651 (1993).
50. Friend, S. H. *et al.* A human DNA segment with properties of the gene that predisposes to retinoblastoma and osteosarcoma. *Nature* **323**, 643–646 (1986).
51. Iwasa, Y., Michor, F. & Nowak, M. A. Stochastic tunnels in evolutionary dynamics. *Genetics*, (in the press).
52. Nowak, M. A. *et al.* The role of chromosomal instability in tumor initiation. *Proc. Natl Acad. Sci. USA* **99**, 16226–16231 (2002).
53. Lengauer, C., Kinzler, K. W. & Vogelstein, B. Genetic instabilities in human cancers. *Nature* **396**, 623–649 (1998).
54. Nasmyth, K. Segregating sister genomes: the molecular biology of chromosome separation. *Science* **297**, 559–565 (2002).
55. Kolodner, R. D., Putnam, C. D. & Myung, K. Maintenance of genome stability in *Saccharomyces cerevisiae*. *Science* **297**, 552–557 (2002).
56. Shonn, M. A., McCarroll, R. & Murray, A. W. Requirement of the spindle checkpoint for proper chromosome segregation in budding yeast meiosis. *Science* **289**, 300–303 (2000).
57. Rajagopalan, H. *et al.* Inactivation of *hCDC4* can cause chromosomal instability. *Nature* (in the press).
58. Loeb, L. A. Mutator phenotype may be required for multistage carcinogenesis. *Cancer Res.* **51**, 3075–3079 (1991).
59. Sieber, O. M. *et al.* Analysis of chromosomal instability in human colorectal adenomas with two mutational hits at *APC*. *Proc. Natl Acad. Sci. USA* **99**, 16910–16915 (2002).
60. Tomlinson, I. P., Novelli, M. R. & Bodmer, W. F. The mutation rate and cancer. *Proc. Natl Acad. Sci. USA* **93**, 14800–14803 (1996).
61. Little, M. P. & Wright, E. G. A stochastic carcinogenesis model incorporating genomic instability fitted to colon cancer data. *Math. Biosciences* **183**, 111–134 (2003).
62. Wodarz, D. & Krakauer, D. C. Genetic instability and the evolution of angiogenic tumor cell lines. *Oncol. Rep.* **8**, 1195–1201 (2001).
63. Plotkin, J. B. & Nowak, M. A. The different effects of apoptosis and DNA repair on tumorigenesis. *J. Theor. Biol.* **214**, 452–467 (2002).
64. Komarova, N. L., Lengauer, C., Vogelstein, B. & Nowak, M. A. Dynamics of genetic instability in sporadic and familial colorectal cancer. *Cancer Biol. Ther.* **1**, 685–692 (2002).
65. Bardelli, A. *et al.* Carcinogen-specific induction of genetic instability. *Proc. Natl Acad. Sci. USA* **98**, 5770–5775 (2001).
66. Nowak, M. A. & Sigmund, K. Evolutionary dynamic of biological games. *Science* **303**, 793–799 (2004).

Acknowledgements

We thank C. Lengauer for discussion. Support from J. Epstein is gratefully acknowledged.

Competing interests statement

The authors declare that they have no competing financial interests.

Online links

DATABASES

The following terms in this article are linked online to:

LocusLink: http://www.ncbi.nlm.nih.gov/LocusLink/

CDC4 | *RB*

OMIM: http://www.ncbi.nlm.nih.gov/omim/

ataxia telangiectasia | Bloom syndrome | familial adenomatous polyposis | Werner syndrome

Access to this interactive links box is free online.

Biographies

Franziska Michor, Yoh Iwasa and Martin Nowak are gatekeeper, caretaker and landscaper in an attempt to study mathematical biology of cancer. Michor prevents analysis of un-biological ideas, Iwasa takes care of precise calculations, while Nowak landscapes their progress. Michor studied molecular biology and mathematics at the University of Vienna and thinks she is very close to finishing a PhD at Harvard. Iwasa studied in Kyoto and Stanford and is currently Professor of Theoretical Biology at Kyushu University. He is distinguished President of the Society of Evolution of Japan. Nowak studied in Vienna, worked at Oxford and Princeton and is now Professor of Mathematics and Biology at Harvard University.

LocusLink: http://www.ncbi.nlm.nih.gov/LocusLink/

CDC4
http://www.ncbi.nlm.nih.gov/LocusLink/LocRpt.cgi?l=55294

RB
http://www.ncbi.nlm.nih.gov/LocusLink/LocRpt.cgi?l=5925

OMIM: http://www.ncbi.nlm.nih.gov/omim/

ataxia telangiectasia
http://www.ncbi.nlm.nih.gov/htbin-post/Omim/dispmim?208900

Bloom syndrome
http://www.ncbi.nlm.nih.gov/htbin-post/Omim/dispmim?210900

familial adenomatous polyposis
http://www.ncbi.nlm.nih.gov/htbin-post/Omim/dispmim?175100

Werner syndrome
http://www.ncbi.nlm.nih.gov/htbin-post/Omim/dispmim?277700