

# Sustained Coevolution in a Stochastic Model of Cancer–Immune Interaction

Jason T. George<sup>1,2,3</sup> and Herbert Levine<sup>1,2,4</sup>



## ABSTRACT

The dynamic interactions between an evolving malignancy and the adaptive immune system generate diverse evolutionary trajectories that ultimately result in tumor clearance or immune escape. Here, we create a simple mathematical model coupling T-cell recognition with an evolving cancer population that may randomly produce evasive subclones, imparting transient protection against the effector T cells. T-cell turnover declines and evasion rates together explained differences in early incidence data across almost all cancer types. Fitting the model to TRACERx evolutionary data argued in favor of substantial and sustained immune pressure exerted upon a developing tumor, suggesting

that clinically observed incidence is a small proportion of all cancer initiation events. This dynamical model promises to increase our quantitative understanding of many immune escape contexts, including cancer progression and intracellular pathogenic infections.

**Significance:** The early cancer–immune interaction sculpts intratumor heterogeneity through the selection of immune-evasive clones. This study provides a mathematical framework for investigating the coevolution between an immune-evasive cancer population and the adaptive immune system.

## Introduction

T-cell immunotherapy has revolutionized modern cancer therapy, delivering durable remission outcomes with higher rates of overall survival for many cancer subtypes (1, 2). The tumor–immune interaction is quite complex, determined by the recruitment of a variety of immune system effectors, and varies as a function of disease duration, progression, and subtype (3, 4). A majority of observed and therapeutic antitumor immunity is accomplished by the cytotoxic CD8<sup>+</sup> T-cell repertoire, wherein immune cells recognize tumor-associated antigens (TAA) present on the surface of cancer cells (5). Adequate recognition leads to elimination of the target cells and expansion of the recognizing T-cell clone. Consistent with theoretical expectations (6, 7), effective immune repertoire detection of TAAs has been empirically validated during natural cancer progression as well as in the postimmunotherapeutic setting (8, 9).

One critical phenomenon that determines ultimate disease outcome is immunosurveillance, which relates to the degree to which tumor progression (typically in the absence of immunotherapy) is controlled by cancer–immune coevolution. To date, a majority of our understanding of this complex interaction has been experimentally driven (10–13) with immunoeediting and TAA negative selection driving early-stage disease (14, 15). Recognition of TAAs relies on immune system discrimination between self and non-self epitopes, which is largely achieved through host-directed tolerance mechan-

isms. First, thymic negative selection functions by deleting immature T cells, which recognize common self-epitopes, thereby imparting central tolerance to the T-cell repertoire (16). This process is incomplete, however, and necessitates subsequent peripheral tolerance mechanisms (17, 18). Peripheral tolerance, in contrast, relies on T-cell inhibition at a systems-level and is achieved whenever a threat is perceived to be small based on a lower detection limit (19).

The “growth-threshold conjecture” is a key defining feature of systems-level immune detection wherein the total rate of change, not level, of a growing threat is the limiting determinant of immune system recognition (20, 21). Consequently, pathogens with small individual growth rates may slowly expand to large population sizes before risking immune system detection, consistent with the observed indolent course of intracellular viral illnesses (22) and escape of slow-growth tumors (23, 24). The corresponding evolutionary trajectories that arise as a result of immune detection based on threat dynamics are quite diverse and have implications for disease prognosis (15, 25). A mathematical framework that formalizes the underlying probabilistic dynamics governing the tumor–immune interaction promises to shed light on this complex process and could ultimately be implemented to better guide patient-specific immunotherapy.

The idea of modeling tumor progression and, in particular, treatment evasion is not new. Well-established stochastic models have revolutionized our understanding of acquired drug resistance in the setting of targeted therapy, in part, by highlighting the importance of random mutant acquisition and clonal selection on the population level (26, 27). Prior studies have also analyzed tumor escape using ordinary differential equations to model the tumor–immune interaction (28–30). A more recent analysis investigated systems-level cancer-immune dynamics in a deterministic setting without taking into account any evasion (23). Because the immune repertoire and evading cancer population are both evolving systems capable of random and repeated recognition and evasion, stochastic analysis is a useful framework for conceptualizing tumor recognition by the adaptive immune system. We previously considered an extreme version of this problem, namely an adaptive immune analogue susceptible to permanent defeat by a single but durable stochastic evasion event (24). Here we consider a more realistic scenario and use it to interpret data regarding age incidence curves as well as tumor clonal evolution.

<sup>1</sup>Center for Theoretical Biological Physics, Rice University, Houston, Texas.

<sup>2</sup>Department of Bioengineering, Rice University, Houston, Texas <sup>3</sup>Medical Scientist Training Program, Baylor College of Medicine, Houston, Texas.

<sup>4</sup>Department of Physics, Northeastern University, Boston, Massachusetts.

**Note:** Supplementary data for this article are available at Cancer Research Online (<http://cancerres.aacrjournals.org/>).

**Corresponding Authors:** Jason T. George, Rice University, 6100 Main Street, Houston, TX 77005. Phone: 713-348-8121; Fax: 713-348-8125; E-mail: [jason.thomas.george@gmail.com](mailto:jason.thomas.george@gmail.com); and Herbert Levine, Phone: 617-373-2903; E-mail: [herbert.levine@rice.edu](mailto:herbert.levine@rice.edu)

Cancer Res 2020;80:811–9

doi: 10.1158/0008-5472.CAN-19-2732

©2019 American Association for Cancer Research.

Specifically, we introduce an analytically tractable mathematical framework for modeling the interaction between an adaptive T-cell repertoire capable of repeated recognition and a cancer population that may repeatedly evolve mechanisms of immune evasion. We solve for the cancer escape and elimination probabilities analytically and provide a statistical framework to study the mean-variance profiles of observed cancer coevolutionary trajectories. We apply our model to early cancer evolutionary data and predict “common initiation and rare progression,” wherein a majority of cancer-initiating events are controlled by an intact immune system, of which, only a small subset randomly escapes and leads to observed disease.

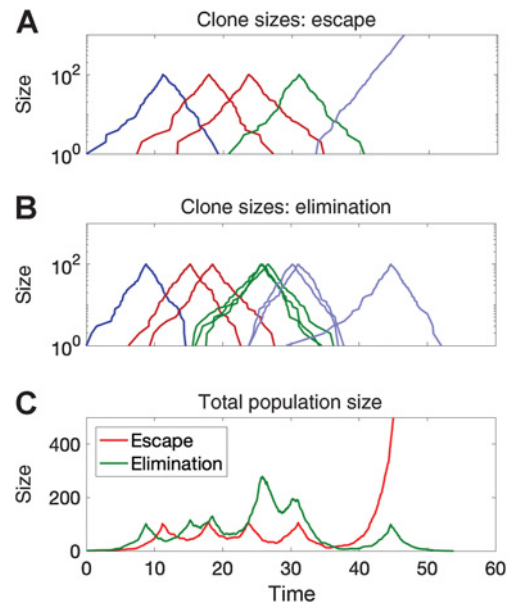
### Quick Guide to Equations and Assumptions

Our basic tumor-immune model consists of clones of cancer cells with pure birth at rate  $r$  until they are detected by the immune system. Recognition may occur once the tumor has reached a minimal detection size  $m_0$  and we work assuming either all clones are detected at size  $m_0$  with a one-time detection probability  $q$  (referred to as deterministic detection) or an ongoing chance of detection at sizes above  $m_0$  (adaptive detection). Once detection occurs, cells are killed at rate  $d > r$ , giving rise to a finite lifetime for that clone. There is also a probability per birth event  $\mu \ll 1$  that the clone will give rise to an evasive tumor subclone. We previously investigated a simpler model that assumed evasion events to be complete and durable against recognition by all current and future T cells (24). Effectively, evading clones were rare but catastrophic cells that were invisible to the immune system, modeling an extreme event such as MHC-I downregulation. In general, it is more reasonable to assume that most evasion events impart transient escape via cancer antigen downregulation, for example, from recognition by the current dominant effector T-cell clone, with future detection of less-frequent or *de novo* TAAs possible by other T-cell clones.

We study here the clonal evolution that results from repeated recognition and evasion events. Elimination of the parent clone follows from its recognition and eventual elimination by the immune system. However, the cancer population evolves if it successfully acquires an immune-evasive daughter clone prior to elimination. This process repeats for each subsequent population until either elimination or escape occurs, and we refer henceforth to the periodic growth followed by detection and elimination as a recognition cycle. To accomplish this, we assume that any evasive clonal population has a chance of being detected upon maturity to detection size. A typical trace of the population dynamics generated by our model is shown in Fig. 1A–C where clones are created by evasion and die by detection over repeated cycles. Note that when the tumor population reaches some higher threshold  $M$  the cancer becomes clinically observable and “escapes” immune surveillance; in this simulation this occurs at around  $t = 50$ . Here the color code reflects subclones created during a specific cycle of the process. In general, a given clone can give rise to multiple evasive lineages during its lifetime. For the deterministic detection scheme, it is straightforward to argue (24) that the number of such daughter clones may be approximated by a Poisson distribution with parameter

$$\lambda = \mu m_0 d / (d - r). \tag{A}$$

Thus, the total evader “intensity,” that is, the parameter governing the total number of clones produced at period  $n$ , is simply  $\gamma_{n,j} = j\lambda$ , where  $j$  is the number of evasive clones at period  $n$ . A generalization of this formula for the adaptive assumption is presented in the Supple-



**Figure 1.** Overview of coevolutionary trajectories. Simulations of the tumor-immune interactions depict clonal population sizes across recognition cycles (distinguished by different adjacent colors) in the event of ultimate escape (A) and elimination (B). The total population sizes are also recorded (C).

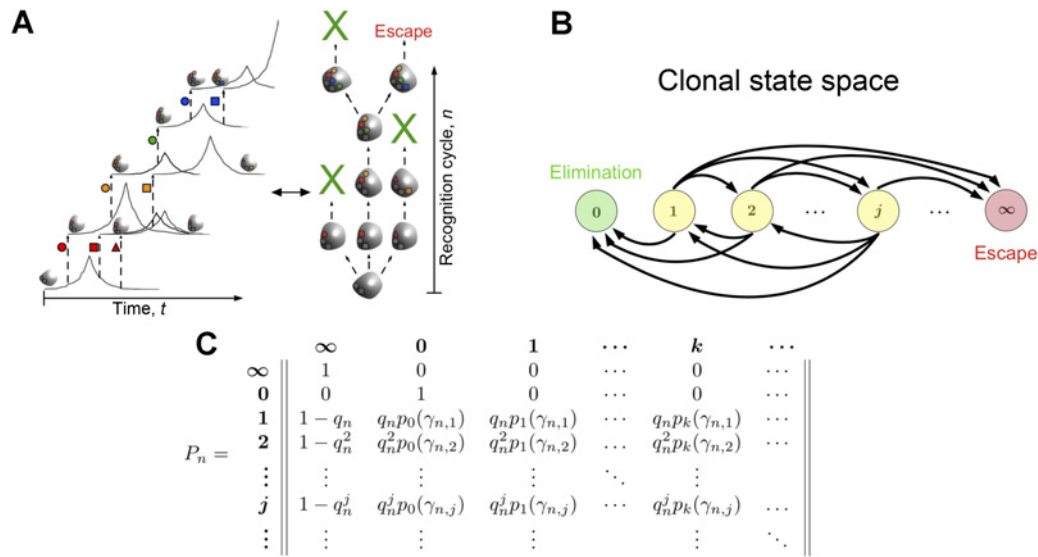
mentary Data. The model can therefore be framed as generalization of the Galton–Watson branching process, well-established in modeling biological systems (31–33), by the addition of possible cancer escape. The process thus eventually culminates in either escape to size  $M$  or complete eradication of the tumor. We shall refer to the expected per-capita progeny as the branching parameter. In the deterministic case, this value is simply  $\lambda$  from Eq. A.

Instead of tracking the population size at a given time  $t$  along with the distribution of arriving clones (Supplementary Fig. S1), we focus instead on the number of evading clones,  $Z_n$ , per recognition cycle,  $n$ , starting from one initial clone (Fig. 2A). The clearance probability, or likelihood that a clone is recognized at cycle  $n$ , is denoted by  $q_n$ . This framework can in general handle any reasonable assumptions on the dependence of detection on cycle number. For example, it may be the case that the immune system becomes less capable over time due to exacerbation by an increasing number of unique clones. One could also generalize the above treatment to allow for the recognition rate to depend explicitly on the number of clones present at cycle  $n$  or even on the individual evasive clone as a function of its appearance order. These assumptions lead to an inhomogeneous Markov process relating the number of clones at cycle  $n$  to the number of clones at cycle  $n + 1$  (Fig. 2B and C). For simplicity, we will mostly assume that  $q_n$  is a constant, which will be denoted as  $q_c$ , and in this case the Markov chain is homogeneous in cycle number. The probability of generating  $k$  clones in the next cycle given the existence of  $j$  clones in the current cycle is

$$p_{n+1,k} = p_k(\gamma_{n,j}) q_n^j, \tag{B}$$

with

$$p_k(\gamma_{n,j}) = e^{-\gamma_{n,j}} \gamma_{n,j}^k / k! \tag{C}$$



**Figure 2.** Cancer-immune coevolutionary dynamics. **A**, A single initial clone grows, becomes recognized, and is eliminated, but not before producing immune evasive clones (distinguished by shapes), which impart transient immunity against the current effector T-cell clone. We assume that each evader has an independent opportunity to either escape or become recognized by the T-cell repertoire and may also give rise to additional evaders (evasion events occurring within the same recognition cycle share the same color and are differentiated by shape). **B**, The state space of this discrete process is the number of distinct evasive clones at each generational period. Escape (state  $\infty$ ) and elimination (state 0) are absorbing states, and all other intermediate states communicate. **C**, This model may be represented in greatest generality by an inhomogeneous Markov process with a transition probability matrix that depends on the immune clearance probability,  $q_n^j$ , of each of the  $j$  clones at time  $n$  and the probability,  $p_k(\gamma_{n,j})$ , of generating  $k$  clones from  $j$  current clones, given by Eq. C.

The immune system must recognize every subclone at each period or else the threat escapes. Thus, for  $j$  clones at period  $n$ , the escape probability is

$$p_{n+1,\infty} = 1 - q_n^j. \tag{D}$$

Finally, the probability of tumor elimination is just the probability that no evasive clones are generated and no escape occurs, namely

$$p_{n+1,0} = q_n^j e^{-\gamma_{n,j}}. \tag{E}$$

## Materials and Methods

All simulations, data analysis, and data visualization were performed using MATLAB version R2017b.

### Escape and elimination probabilities

Derivations of the escape and elimination probabilities, together with the branching parameter, were solved using standard probabilistic analysis.

### Simulations

All simulations of the evolutionary process were performed using a modified Gillespie algorithm. Analyses involving escape were simulated for all surviving populations until the emergence of a dominant clone that exceeded the upper clinical escape threshold.

### Cancer age incidence

Age-specific cancer incidence data for breast, bladder, kidney, pancreas, prostate, head and neck, ovary, lung, melanoma, acute

myeloid leukemia, and chronic lymphocytic leukemia were obtained from Cancer Research UK. Age incidence for triple-negative breast cancer was obtained by applying Web Plot Digitizer (34) to Fig. 1 of ref. 35 and taking a weighted average based on the empirical frequency of reported groups (Asian,  $n = 368$ ; Black,  $n = 509$ ; Hispanic,  $n = 920$ ; White,  $n = 2521$ ). Least-squares linear regression was performed on all available incidence data between ages 0 and 40 and in a similar manner between the age at which incidence is first measured and 40. The corresponding least-squares slope parameter is plotted for each cancer subtype against the mutation rate.

### Evasion rates

Estimates of the per-cell mutation rates were obtained by using median rates from the available estimates from ref. 36 for acute myeloid leukemia (mutation rate  $\mu = 3.33 \times 10^{-7}$ ;  $n = 134$ ), bladder ( $\mu = 5.17 \times 10^{-6}$ ;  $n = 35$ ), breast cancer ( $\mu = 9.33 \times 10^{-7}$ ;  $n = 121$ ), chronic lymphocytic leukemia ( $\mu = 7.67 \times 10^{-7}$ ;  $n = 91$ ), head and neck cancer ( $\mu = 3.17 \times 10^{-6}$ ;  $n = 181$ ), lung cancer ( $\mu = 7.20 \times 10^{-6}$ ;  $n = 514$ ), melanoma ( $\mu = 1.32 \times 10^{-5}$ ;  $n = 121$ ), ovarian cancer ( $\mu = 1.65 \times 10^{-6}$ ;  $n = 394$ ), pancreatic cancer ( $\mu = 1.07 \times 10^{-6}$ ;  $n = 13$ ), prostate cancer ( $\mu = 7.33 \times 10^{-7}$ ;  $n = 227$ ), and renal cancer ( $\mu = 1.53 \times 10^{-6}$ ;  $n = 225$ ).

### Statistical analysis

Correlations between the linear slope parameter of early cancer age incidence and per-cell evasion rates were assessed using Pearson correlation coefficient, and reported  $P$  values were obtained. For this, the null hypothesis was that there existed no correlation between the slope of early cancer age incidence and the per-cell evasion rate. Rejection of the null hypothesis and the presence of

significant (nonzero) correlation was assessed at a significance value of  $\alpha = 0.05$ .

**TRACERx timing of evasion events**

The time to most recent common ancestor relative to cancer initiating time for renal clear cell carcinoma was estimated from the TRACERx renal consortium using previously published data (37). The data were separated into two categories: evasion dominated by branching (samples K59, K167, K146, K065, K003, K150, K104, K163, K153, K126, K108, K135, K169, K139, K145, K156, K162, K180) and primarily linear evolution (samples K113, K165, K130, K023, K176, K037, K143, K124, K136, K096, K027, K158, K105) based on samples having their most recent common ancestor earlier or within 10 years of clinical detection, respectively. Mean and variance estimates of the mutational burden in patients with lung cancer were calculated from data reported in (38). Mean-variance frontiers were plotted assuming the quadratic relationship of Eq. J parameterized by some unknown fraction of relevant evasion events  $\alpha$ , where we considered  $\alpha$  up to 10%. Immune-enhanced, immune-compromised, and immune-neutral regions were defined by calculating all allowable mean-variance frontiers based on trajectories generated by the domain of interest.

**Results**

The following sections present the main findings of our analysis. Full mathematical details are provided in the Supplementary Data.

**Escape and elimination probabilities**

Let  $E_n$  (resp.  $F_n$ ) be the event that the threat escapes (resp. is eliminated) at period  $n$ . One can directly derive a backwards discrete-time master equation that determines the long-time limit of these splitting probabilities. Specifically focusing on the elimination and taking  $q = q_c$  to be constant, we have

$$P_k(F_n) = q_c^k \sum_{j=1}^{\infty} P_j(F_{n+1}) e^{-\lambda k} (\lambda k)^j / j!, \tag{F}$$

where  $P_k(F_n)$  is the probability that a tumor trajectory having  $k$  clones at cycle  $n$  eventually leads to complete tumor elimination. Clearly, each clone is independent from its future trajectory and hence  $P_k(F_n) = P_1(F_n)^k$ . Thus, Eq. F becomes a closed-form equation for  $P_1(F_n)$

$$P_1(F_n) = q_c \sum_{j=1}^{\infty} P_1(F_{n+1})^j e^{-\lambda j} / j! \tag{G}$$

Taking the steady-state limit of this equation yields an equation for the asymptotic elimination probability  $p^*$

$$p^* = q_c e^{-\lambda(1-p^*)} \tag{H}$$

One can similarly find the asymptotic value of the escape probability and verify that it is just  $1 - p^*$ ; in other words, the system always chooses one of the absorbing states in the long-time limit. A more complete treatment presented in the Supplementary Data allows this result to be extended to evaluate the full time-dependent value of  $P_1(F_n)$  along with more general assumptions for the  $q_n$ .

The above outlines the model for the deterministic case since detection, should it occur, must do so at size  $m_0$ . Adaptive detection incorporates immune decline via an immunomodulatory parameter  $v > 0$  that results in larger average detection sizes and yields a similar elimination probability  $\tilde{p}^*$  (See Supplementary Data for full details). This analytic framework above agrees well with results obtained from simulating the full coevolutionary process (Fig. 3A and B). “Escape by

underwhelming” is an important experimentally observed feature wherein threats of intermediate growth rates are at a survival disadvantage relative to their faster and slower growing counterparts (39). Our model recapitulates this behavior for minimal detection thresholds  $m_0$  limited by the total cancer growth rate (Fig. 3C), in agreement with the growth-threshold conjecture (22). This simple theoretical framework generates a significant degree of diversity in dynamical behavior, governed completely by the clearance probability and the branching parameter (Fig. 3D; Supplementary Figs. S2–S10).

Our foundational model may also be adapted to investigate a variety of assumptions on the nature of immune escape, including declines in immune clearance as a function of elapsed time and total surviving clone numbers (Sec. S4.4, Intertemporal immune decline; Sec. S4.5, Clone frequency-dependent recognition) along with their effects on the likelihood of cancer elimination and escape (Supplementary Figs. S2 and S3). Prior empirical studies have proposed the existence of an equilibrium state, wherein a cancer population is neither fully eliminated nor grows to overwhelm the host (40, 41). Surprisingly, our framework would suggest that ultimate coexistence is not reasonably predicted, occurring only when unreasonable assumptions are imposed on the clearance probabilities (Sec. S4.3).

**Differences in cancer early age incidence correlate with evasion rate**

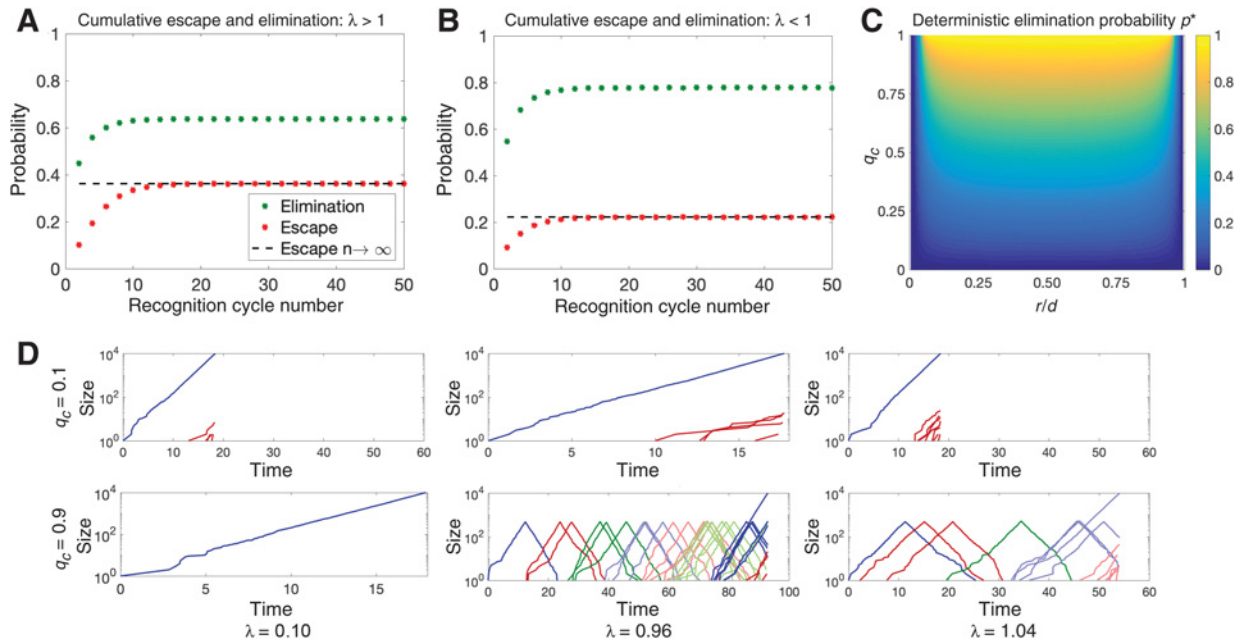
As proof-of-principle that local physiologic changes in immune function may affect observed cancer frequency, even among healthy individuals, we study the effects of physiologic declines in immune turnover on increases in early age incidence for various cancer types using publicly available large population datasets (Supplementary Fig. S6; ref. 42). Previous studies have shown that age-incidence data can be fit to simple models assuming a declining immune system with age (6, 43). Our dynamical model predicts, all else being equal, that tumor escape likelihoods vary directly as a function of evasion rate  $\mu$  implicitly through the branching parameter. Although evasion need not be limited exclusively to genomic mechanisms, we use experimentally derived per-cell mutation rates obtained from a large pancreatic analysis as an indirect measure of evasion (36).

We calculate the slope parameter of linear regression for early age incidence (Fig. 4A; Supplementary Fig. S6) and find a strong correlation between these values and the evasion rate. Aggregate breast cancer incidence appears to be an exception, perhaps a result of early screening detection and the incidence following puberty of hormone-sensitive disease. Restricting the comparison to triple negative breast cancer (35) reveals statistically significant correlation between incidence increases and evasion rate (Fig. 4B; Supplementary Fig. S7). Decreases in immune function modeled via an increased adaptive parameter  $v$  (see Supplementary Data), assumed directly proportional to early age increases, result in increased ultimate escape probability,  $1 - \tilde{p}^*$ . We find excellent agreement in the predicted incidence slope- $\mu$  profile for the best-fit scaled range of immune function parameter  $v$  (Fig. 4C). Moreover, using this normalized range for  $v$  in the model predicts a linear increase in incidence for all observed mutation rates (Fig. 4D). Together, these findings demonstrate that cancer progression and escape are well-explained by variability in immune function across nearly all cancer types.

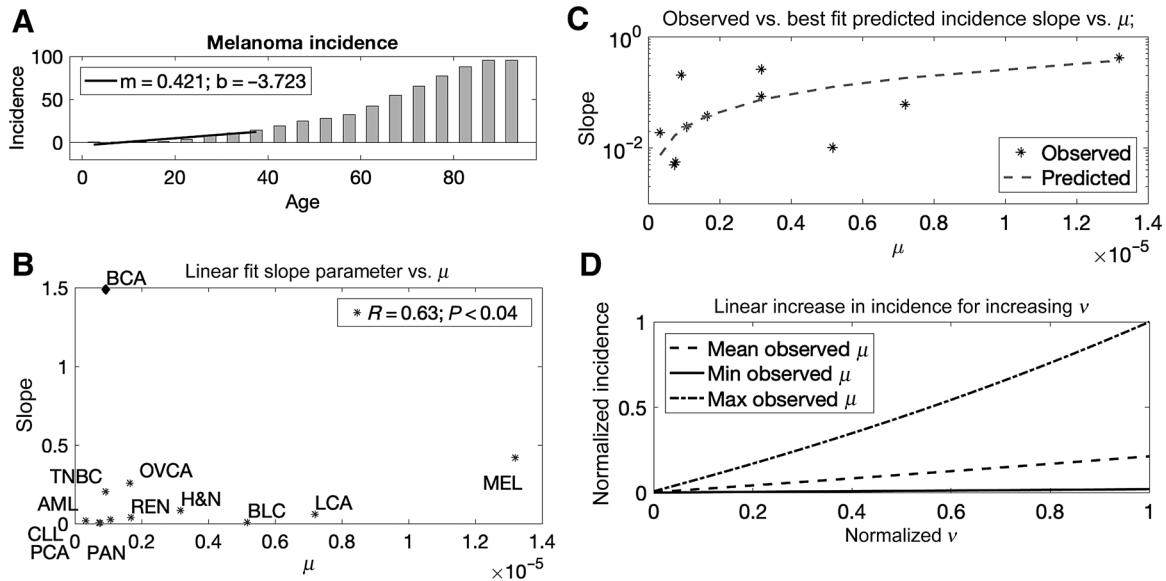
**Evolutionary timing of renal clear cell carcinoma predicts prolonged coevolution with common tumor incidence and rare progression**

Given the ability of our mathematical model to link observed cancer evasion and recognition cycle timing to underlying tumor-immune

Downloaded from <http://aacrjournals.org/cancerres/article-pdf/80/4/811/2802293/811.pdf> by guest on 04 August 2024



**Figure 3.** General dynamics of the tumor-immune coevolutionary model. Escape (red) and elimination (green) probabilities are plotted as a function of recognition cycle number and compared to analytic escape (black dashed line), assuming supercritical ( $\lambda = 1.04$ ; **A**) and subcritical ( $\lambda = 0.96$ ; **B**) branching (in each case,  $q_c = 0.95$  and simulations are averaged over  $10^6$  iterations). **C**, The dynamics induced by Eq. H permit a unique limiting elimination probability,  $p^*$ , which is plotted as a function of clearance probability and relative net growth rate ( $\lambda$  calculated using Eq. A, with  $\mu = 10^{-6}$ ,  $d = 0.2$ ,  $R = 10^4$ , and  $m_0 = R/r$ ). **D**, A variety of evolutionary trajectories are plotted conditional on ultimate escape, assuming various clearance and branching parameters (in all cases, deterministic recognition was assumed, with  $m_0 = 500$  and  $r/d = 0.5$ ).



**Figure 4.** Cancer incidence and evasion rate. **A**, The least-squares linear regression slope parameter of cancer early age incidence (between ages 0 and 40 years) is calculated for a variety of cancer subtypes (age at primary disease; melanoma shown); **B**, Rate of change in cancer early age incidence versus evasion rate. For each age incidence curve, linear regression is performed for incidence between ages 0 and 40 years. This parameter is plotted as a function of per-cell mutation rates for each cancer type. **C**, Assuming that the relationship between age and  $v$  is linear,  $\text{age} = kv$ , the slope in cancer age-incidence can be calculated as a function of increasing  $v$ , and are plotted for  $k$  that gives the least-squares error as compared with the experimental data. **D**, This parameterization consequently predicts linear increases in early cancer incidence as a function of diminished immune performance for all observed ranges of evasion rates  $\mu$ .

Downloaded from <http://aacrjournals.org/cancerres/article-pdf/80/4/811/12802293/811.pdf> by guest on 04 August 2024

coevolutionary dynamics, we next assessed the extent of immune evasion and clearance observed in cancer evolutionary data. The TRACERx renal clear cell carcinoma (RCCC) dataset (37) was particularly useful, providing estimates on the timing of landmark evolutionary events. Surprisingly, and consistent with model behavior under sustained control (Fig. 3D,  $q_c = 0.9, \lambda = 0.96$ ), the authors found extended periods of clonal evolution sustained by small population sizes (estimated to be of order  $10^2$  cells). Using the arrival time of the most recent common ancestor (MRCA) as a representative of escape, we estimated the expected number of recognition cycles given the observed intervening time between MRCA detection and disease initiation. If  $t_m$  is the time it takes a clone to grow to minimal detection size  $X(t_m) \sim 10^2$  and  $T_M$  is the times it takes to grow to ultimate detection size  $X(t_m) \sim 10^9$ , then the intervening time, assuming  $n$  recognition cycles occurs, is given by  $\Delta t = t_m - nt_m$ . From this, we can estimate the number of recognition cycles via

$$n = (7nt_m)/2(t_M - nt_m), \quad (\text{I})$$

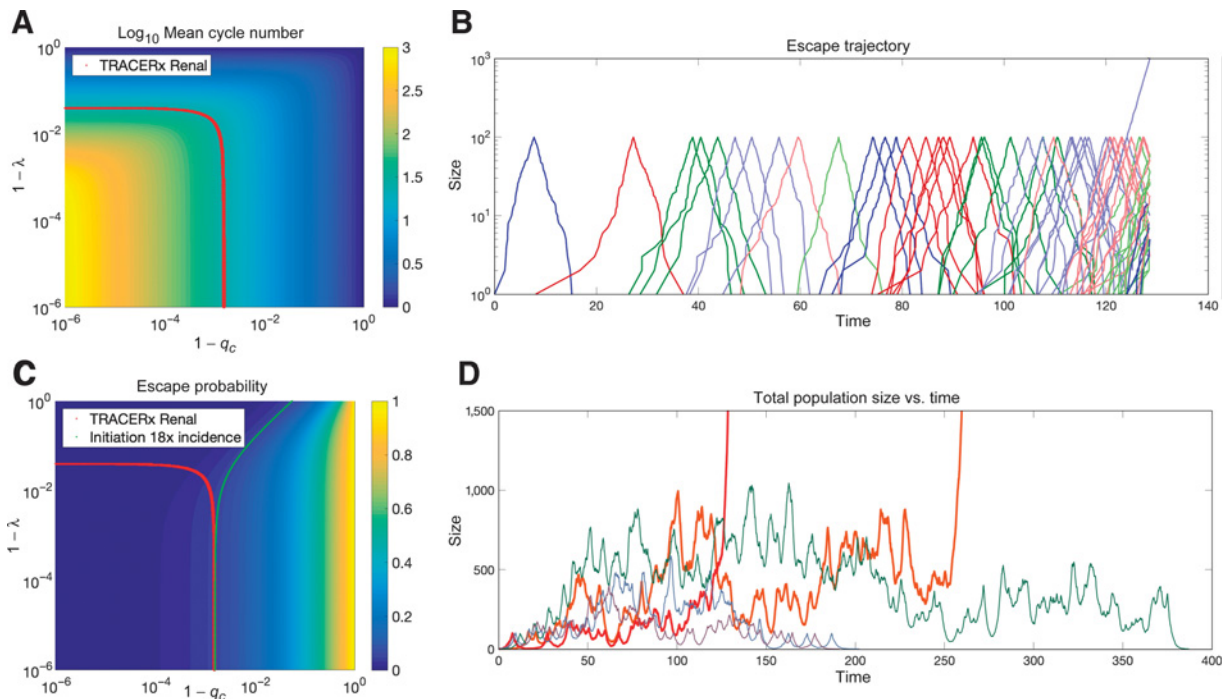
where  $nt_m$  and  $t_M$  are the observed times (see Supplementary Data for details). Our calculations predict that, on average, RCCC undergoes approximately 27 distinct, immunologically relevant recognition cycles prior to escape (Fig. 5A and B; Supplementary Fig. S11A).

The parameter range that generates this behavior is near criticality ( $\lambda < 1$ ) with high clearance probabilities ( $q_c$  approaching 1), suggesting effective immune handling of renal cell threats (Fig. 5C; Supplementary Fig. S11B). To further investigate the extent of immune protection, we used parameter estimates consistent with these findings to assess

the frequency of observed cancer incidence relative to the total number of (unobserved) cancer-initiating events, and predict that initiation is 18 times as frequent as measured incidence. These results suggest that the adaptive immune system actively filters many potential threats manifesting as *de novo* initiating cancer clones, but may occasionally miss due to statistical chance, in support of the “bad-luck hypothesis” (44). The resulting population dynamics closely resemble those observed in the cancer incidence data, with a large disease period consisting of clonal evolution with low tumor burden, followed by either elimination or escape with extensive subclonal evolution (Fig. 5D). When separating samples into categories based on extensive linear or branched evolution and repeating this analysis, we find that patients with disease dominated by linear evolution are predicted to have significantly larger cycle numbers ( $57 \pm 35$ ) with a concomitant increase in the number of predicted initiating events relative to observed incidence. This behavior contrasts with trajectories dominated by branched evolution, which have a reduced cycle number ( $5 \pm 2$ ) with increased escape probability (Supplementary Fig. S12).

#### Fluctuations in lung cancer preceding escape partition disease subtypes based on immune function

Given the predicted implications of immune performance on eliminating early and indolent disease, we wanted to investigate the relationship between disease subtype and predicted immune status. Available data on large cohort patients with lung cancer (38) provided enough samples for each studied subtype to estimate mean and variance profiles for the distribution of clonal mutational events. This analysis was repeated for each lung cancer subtype and then compared with all predicted values



**Figure 5.**

Evolutionary dynamics of renal clear cell carcinoma. **A**, Relative timing of the TRACERx renal time to MRCA versus time to clinical escape implies an average of 26.6 recognition cycles, which restricts the range of allowable values of  $q_c$  and  $\lambda$  in the coevolution model (red line traces the contour line for relevant mean cycle number). **B**, A typical stochastic realization of the clonal branching that occurs in the deterministic case before and after escape (colors distinguish adjacent subclones arising within the same cycle number). **C**, Minimal values of observed clearance probability provide an upper estimate for the observed escape probability at 1/18 (plotted as a function of  $q_c$  and  $\lambda$ ; green line traces the contour line for relevant escape probabilities). **D**, Simulation output of the total population size as a function of time for various realizations of the process.

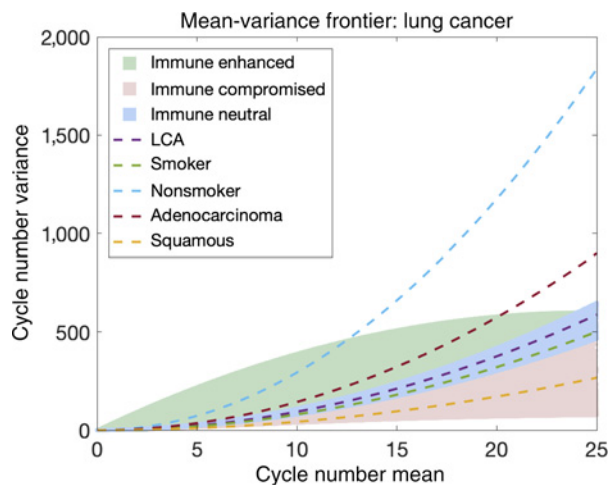
in the model. Because only a small (unknown) fraction  $\alpha \ll 1$  of total mutations are expected to impart immune evasion, we consider the resulting allowable mean-variance frontier assuming a range of small  $\alpha$ . Very briefly, for  $X$  being the total number of clonal nonsynonymous mutational events with respective mean and variance given by  $\mu_X$  and  $\sigma_X^2$ , the mean and variance for the fraction  $\alpha X$  that are relevant to immune evasion relate to the observed parameters  $\mu_X$  and  $\sigma_X^2$  via

$$\begin{aligned} \mathbb{E}[\alpha X] &= \mu_X \alpha; \\ \text{Var}(\alpha X) &= \sigma_X^2 \alpha^2. \end{aligned} \quad (\text{J})$$

Equation J defines a quadratic mean-variance curve parameterized by  $\alpha$  that can be applied to compare the observed curves against all allowable simulated parameters in the domain considered (Fig. 5A and C) to determine model conditions that are consistent with empirical data. Our above model defines a baseline relationship that well characterizes the reported mean-variance profiles across all lung cancer evolutionary data subtypes (Fig. 6; Supplementary Table S1). However, several variations in the mean-variance frontier are only achievable in our generalized model assuming either immune suppression or heightened immune surveillance (see Supplementary Data for full details). Interestingly, squamous cell carcinoma, and to a lesser extent positive smoking status, are more consistent with a process that is significantly immunosuppressed, suggesting that the observed pattern of mutational variability emerges in the setting of reduced immune targeting efficiency. Intriguingly, the predicted mean-variance frontier requires enhancements to immune clearance rates over successive periods to explain the nonsmoker signature (Eq. S123).

## Discussion

The underlying early clonal architecture of an evolving cancer population contains a significant amount of information related to



**Figure 6.**

Mean-variance frontiers for lung cancer subtypes. Empirical mean-variance frontiers are plotted using Eq. J and observed data for various LCA subtypes (dashed lines). These are compared with predicted allowable regions in the mean-variance space based on all parameter combinations studied in Fig. 5A and C (colored regions). Blue region corresponds with allowable frontiers, assuming deterministic recognition ( $v = 0$ ). Red region highlights areas of nonoverlap in adaptive recognition where immune recognition is impaired ( $v = 10$  taken to be 10% of  $m_0$ ). Green regions represent dynamics assuming immune enhancement over time ( $\rho_{c,n+1} = \max\{\rho_{c,\infty}, \rho_{c,n} + (\rho_{c,n} - \rho_{n,\infty})/25\}$  for  $\rho_{c,\infty} = 0.9$ ).

cancer progression and the extent of immunosurveillance. Understanding this complex behavior is essential for better predicting the extent of tumor immunoeediting, measuring the frequency of disease progression relative to incidence, estimating immune targeting efficacy, and proposing optimized treatment strategies conditioned on post-escape clonal distributions. Here, we developed a stochastic population dynamical model of the battle between an evolving cancer population and the CD8<sup>+</sup> T-cell immune repertoire. Prior models of the cancer-immune interaction have explicitly represented and tracked immune effectors as a distinct population (22, 29). While useful for describing population control in a variety of contexts, including nonimmunogenetic tumors where immune evasion is a nondominant contributor to cancer progression, this approach falls short in characterizing the phylogenetic architecture and elimination probabilities of evasive cancer subclones following recognition via the adaptive immune system. Our alternative formulation explicitly models recognition on the population level assuming that the cancer compartment may randomly acquire evasive clones, each of which requires distinct T-cell clones in the immune repertoire for recognition and control. We made the implicit assumptions that cancer adaptivity is represented by an evasion rate  $\mu$ , while immune system adaptivity depends on the repertoire diversity and turnover rates. The resulting tumor escape or elimination behavior is thus reduced to an effective branching parameter,  $\lambda$ . This framework is sufficiently general for studying the coevolution of tumor and immune system, in addition to formulating the key issue of immune escape and elimination with regards to the ability of the immune repertoire to keep pace with a heterogeneous and evading cancer population.

Previous empirical studies have proposed the existence of an equilibrium state wherein a growing cancer population can be maintained indefinitely (40, 41). In contrast, our model only predicts ultimate equilibrium for the unlikely case in which recognition rates of new clones increase dramatically as an increasing function of total clone number. More realistic assumptions on the decline in immune recognition, occurring either over time or when overwhelmed by a greater number of competing clones, prevents the cancer population from reaching ultimate equilibrium and implies that the immune recognition process ultimately results in either escape or elimination with certainty, given sufficient time.

The model's underlying dynamical behavior subsequently encodes information on both the likelihood and timing of ultimate cancer escape or elimination. Our framework highlights the importance of immunosurveillance in early disease progression as evidenced by differential increases in cancer incidence that scale with evasion rate across nearly all cancer types. We utilized per-cell mutation rates as an indirect measure of evasion capacity. Our findings were consistent despite changing the age range used as input for the linear regression. Indeed, many mutations are independent of immune evasion events and may increase the antigenic burden of the cancer population. This framework only distinguishes progeny based on an evasion event, which results in alteration or removal of an immune target recognized on the parent clone. Evaluating the competing effects of advantageous antigen down-regulation and undesirable tumor neoantigen generation on overall recognition by the immune system is an important consideration and warrants further investigation.

Using the excellent data available on the timing of early evolutionary events in RCCC, together with the observation that early clonal evolution is estimated to be sustained at small cell sizes, we show that the mean number of recognition cycles is quite large, suggestive of a drawn-out competition between immune and cancer compartments. Our results suggest that cancer progression from early initiation to

appreciable disease only occurs a small fraction of the time, in support of our proposed “common initiation and rare progression” postulate, wherein most cancer initiating events are effectively controlled by host recognition mechanisms. In the minority of cases where the cancer population escapes, our analysis predicts this to be due to random chance, consistent with the “bad-luck” hypothesis (44). Moreover, partitioning of samples into groups based on trajectories dominated by linear versus branched evolution revealed differences in mean recognition cycle and relative escape likelihood, the latter group characterized by punctuated evolution and early disease. We remark that these results are most sensitive to the assumed minimal population size (here taken to be  $m_0 = 10^2$  as estimated empirically); the resulting mean cycle number estimate scales as a decreasing function of  $m_0$  as  $2/\log_{10} m_0$ . Despite this, significant and prolonged coevolution is still predicted (15 recognition cycles in renal cancer patients) assuming an order of magnitude increase in the minimal detection size. Although empirically challenging, further investigation into the early tumor-immune interaction following initiation could refine this estimate and shed light on the underlying evolution of early disease for a variety of cancers. When comparing the mean-variance frontiers in the clonal distributions of TRACERx LCA subtypes, we found that observed evolutionary trajectories for squamous cell carcinoma and positive smoking status were consistent with simulated profiles assuming immune impairment, while nonsmokers and adenocarcinoma followed an opposite trend. Our results demonstrate the utility of further experimental investigation into the benefit of distinct treatment strategies based on cancer subtype and predicted immune function of each patient.

Cancer evolution is complex and highly variable owing to the dynamic interaction that occurs between tumor cells and the adaptive immune system. Improvement in our understanding of cancer progression therefore requires a detailed description of individual disease evolutionary trajectories. Toward this end, we have developed a mathematical framework that models a heterogeneous population of cancer cells subjected to adaptive immunosurveillance. When applied to empirical data, our model is a useful tool for studying the tumor-

immune interaction as it establishes a fundamental relationship between evasion and the observed branching structure that emerges during disease progression. Although we have focused our analysis on describing the control and progression of cancer, our model is broadly applicable in understanding similar phenomena, such as intracellular infection by pathogens that evolve mechanisms of immune evasion. In HIV, for example, the delayed minimal disease burden produced by repeated recognition cycles in our model shares some general features of the latent period (45). Future efforts to apply this model in a broader context may provide insights regarding the observed dynamics of slow-progression intracellular threats.

## Disclosure of Potential Conflicts of Interest

No potential conflicts of interest were disclosed.

## Authors' Contributions

**Conception and design:** J.T. George, H. Levine

**Development of methodology:** J.T. George, H. Levine

**Acquisition of data (provided animals, acquired and managed patients, provided facilities, etc.):** J.T. George

**Analysis and interpretation of data (e.g., statistical analysis, biostatistics, computational analysis):** J.T. George, H. Levine

**Writing, review, and/or revision of the manuscript:** J.T. George, H. Levine

**Study supervision:** J.T. George

## Acknowledgments

We would like to thank Jeffrey Mollidrem for fruitful discussions on immunotherapy and cancer-immune coevolution.

J.T. George is supported by the NCI of the NIH (F30CA213878). H. Levine is supported by Physics Frontiers Center NSF Grant PHY-1427654 and by NSF PHY-1605817.

The costs of publication of this article were defrayed in part by the payment of page charges. This article must therefore be hereby marked *advertisement* in accordance with 18 U.S.C. Section 1734 solely to indicate this fact.

Received August 31, 2019; revised November 15, 2019; accepted December 18, 2019; published first December 20, 2019.

## References

- Couzin-Frankel J. Cancer immunotherapy. *Science* 2013;342:1432–3.
- Luke JJ, Flaherty KT, Ribas A, Long GV. Targeted agents and immunotherapies: optimizing outcomes in melanoma. *Nat Rev Clin Oncol* 2017;14:463.
- De Visser KE, Eichten A, Coussens LM. Paradoxical roles of the immune system during cancer development. *Nat Rev Cancer* 2006;6:24.
- Finn O. Immuno-oncology: understanding the function and dysfunction of the immune system in cancer. *Ann Oncol* 2012;23:viii6–9.
- Coulie PG, Van den Eynde BJ, Van Der Bruggen P, Boon T. Tumour antigens recognized by t lymphocytes: at the core of cancer immunotherapy. *Nat Rev Cancer* 2014;14:135.
- George JT, Kessler DA, Levine H. Effects of thymic selection on t cell recognition of foreign and tumor antigenic peptides. *Proc Natl Acad Sci U S A* 2017;114:E7875–81.
- Mayer A, Balasubramanian V, Walczak AM, Mora T. How a well-adapting immune system remembers. *Proc Natl Acad Sci U S A* 2019;116:8815–23.
- Yadav M, Jhunjhunwala S, Phung QT, Lupardus P, Tanguay J, Bumbaca S, et al. Predicting immunogenic tumour mutations by combining mass spectrometry and exome sequencing. *Nature* 2014;515:572.
- Zamora AE, Crawford JC, Thomas PG. Hitting the target: how t cells detect and eliminate tumors. *J Immunol* 2018;200:392–9.
- Ott PA, Hu Z, Keskin DB, Shukla SA, Sun J, Bozym DJ, et al. An immunogenic personal neoantigen vaccine for patients with melanoma. *Nature* 2017;547:217–21.
- Sahin U, Derhovanessian E, Miller M, Kloke BP, Simon P, Löwer M, et al. Personalized rna mutanome vaccines mobilize poly-specific therapeutic immunity against cancer. *Nature* 2017;547:222–6.
- Leach DR, Krummel MF, Allison JP. Enhancement of antitumor immunity by cta-4 blockade. *Science* 1996;271:1734–6.
- Sadelain M, Riviere I, Riddell S. Therapeutic T cell engineering. *Nature* 2017;545:423–31.
- Fridman WH, Mlecnik B, Bindea G, Pagès F, Galon J. Immunosurveillance in human non-viral cancers. *Curr Opin Immunol* 2011;23:272–8.
- Rosenthal R, Cadieux EL, Salgado R, Al Bakir M, Moore DA, Hiley CT, et al. Neoantigen-directed immune escape in lung cancer evolution. *Nature* 2019;567:479–85.
- Nossal G. Negative selection of lymphocytes. *Cell* 1994;76:229–39.
- Davis MM. Not-so-negative selection. *Immunity* 2015;43:833–5.
- Legoux FP, Lim JB, Cauley AW, Dikiy S, Ertelt J, Mariani TJ, et al. Cd4+ t cell tolerance to tissue-restricted self antigens is mediated by antigen-specific regulatory t cells rather than deletion. *Immunity* 2015;43:896–908.
- Grossman Z, Paul WE. Adaptive cellular interactions in the immune system: the tunable activation threshold and the significance of subthreshold responses. *Proc Natl Acad Sci U S A* 1992;89:10365–9.
- Pradeu T, Jaeger S, Vivier E. The speed of change: towards a discontinuity theory of immunity? *Nat Rev Immunol* 2013;13:764–9.
- Johansen P, Storni T, Rettig L, Qiu Z, Der-Sarkissian A, Smith KA, et al. Antigen kinetics determines immune reactivity. *Proc Natl Acad Sci U S A* 2008;105:5189–94.
- Arias CF, Herrero MA, Cuesta JA, Acosta FJ, Fernandez-Arias C. The growth threshold conjecture: a theoretical framework for understanding T-cell tolerance. *R Soc Open Sci* 2015;2:50016.



23. Sontag ED. A dynamic model of immune responses to antigen presentation predicts different regions of tumor or pathogen elimination. *Cell Syst* 2017;4: 231–41.
24. George JT, Levine H. Stochastic modeling of tumor progression and immune evasion. *J Theor Biol* 2018;458:148–55.
25. Dunn GP, Bruce AT, Ikeda H, Old LJ, Schreiber RD. Cancer immunoediting: from immunosurveillance to tumor escape. *Nat Immunol* 2002;3:991.
26. Iwasa Y, Nowak MA, Michor F. Evolution of resistance during clonal expansion. *Genetics* 2006;172:2557–66.
27. Michor F, Iwasa Y, Nowak MA. Dynamics of cancer progression. *Nat Rev Cancer* 2004;4:197–205.
28. Al-Tameemi M, Chaplain M, dOnofrio A. Evasion of tumours from the control of the immune system: consequences of brief encounters. *Biol Direct* 2012;7:31.
29. de Pillis LG, Radunskaya AE, Wiseman CL. A validated mathematical model of cell-mediated immune response to tumor growth. *Cancer Res* 2005;65:7950–8.
30. Grossman Z, Berke G. Tumor escape from immune elimination. *J Theor Biol* 1980;83:267–96.
31. Athreya KB, Ney PE. *Branching Processes*. New York: Springer Science & Business Media, 2012.
32. Karlin S. *A first course in stochastic processes*. San Diego: Academic Press, 2014.
33. Grimmett G, Grimmett GR, Stirzaker D, et al. *Probability and random processes*. New York: Oxford University Press, 2001.
34. Rohatgi A. *Webplotdigitizer*. 2011. <https://automeris.io/WebPlotDigitizer/>.
35. Kurian AW, Fish K, Shema SJ, Clarke CA. Lifetime risks of specific breast cancer subtypes among women in four racial/ethnic groups. *Breast Cancer Res* 2010;12: R99.
36. Lawrence MS, Stojanov P, Polak P, Kryukov GV, Cibulskis K, Sivachenko A, et al. Mutational heterogeneity in cancer and the search for new cancer-associated genes. *Nature* 2013;499:214.
37. Mitchell TJ, Turajlic S, Rowan A, Nicol D, Farmery JH, O'Brien T, et al. Timing the landmark events in the evolution of clear cell renal cell cancer: TRACERx renal. *Cell* 2018;173:611–23.
38. Jamal-Hanjani M, Wilson GA, McGranahan N, Birkbak NJ, Watkins TB, Veeriah S, et al. Tracking the evolution of non-small-cell lung cancer. *N Engl J Med* 2017;376:2109–21.
39. Bocharov G, Ludewig B, Bertolotti A, Klenerman P, Junt T, Krebs P, et al. Underwhelming the immune response: effect of slow virus growth on CD8+ T-lymphocyte responses. *J Virol* 2004;78:2247–54.
40. Dunn GP, Old LJ, Schreiber RD. The three es of cancer immunoediting. *Annu Rev Immunol* 2004;22:329–60.
41. Teng MW, Swann JB, Koebel CM, Schreiber RD, Smyth MJ. Immune-mediated dormancy: an equilibrium with cancer. *J Leukoc Biol* 2008;84: 988–93.
42. Acute myeloid leukaemia (aml) incidence by age (2013). Available from: <http://www.cancerresearchuk.org/health-professional/cancer-statistics>.
43. Palmer S, Albergante L, Blackburn CC, Newman T. Thymic involution and rising disease incidence with age. *Proc Natl Acad Sci U S A* 2018;115: 1883–8.
44. Tomasetti C, Vogelstein B. Variation in cancer risk among tissues can be explained by the number of stem cell divisions. *Science* 2015;347:78–81.
45. Siliciano RF, Greene WC. HIV latency. *Cold Spring Harb Perspect Med* 2011;1: a007096.