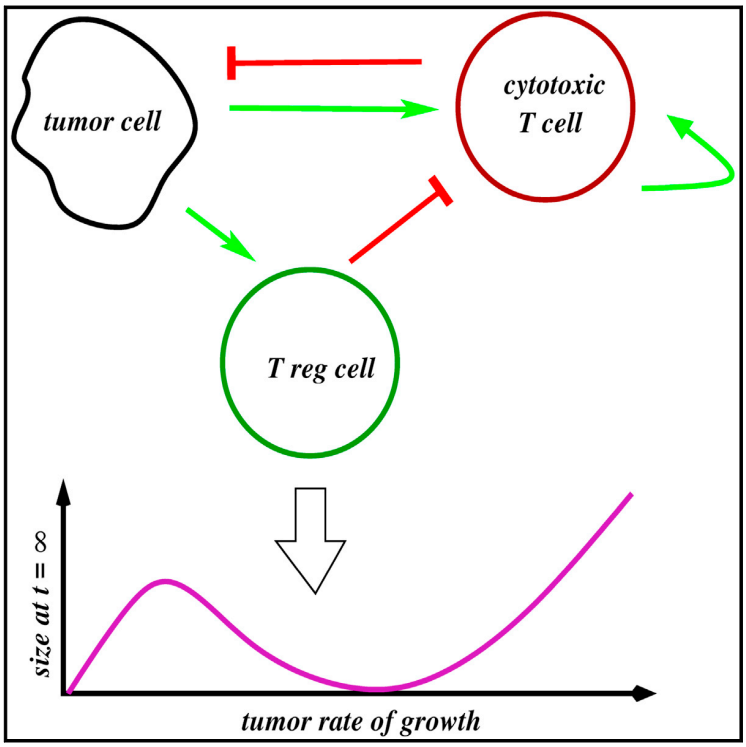
Cell Systems

Math | Bio

A Dynamic Model of Immune Responses to Antigen Presentation Predicts Different Regions of Tumor or Pathogen Elimination

Graphical Abstract



Authors

Eduardo D. Sontag

Correspondence

eduardo.sontag@rutgers.edu

In Brief

The author presents a model of immune recognition that combines three motifs from systems biology, negative feedback, incoherent feedforward loops, and bistability, and captures important dynamic features of antigen presentation. The model recapitulates experimental observations of separate zones of tumor control.

Highlights

- Pathogen and tumor recognition is aided by dynamical features of antigen presentation
- Our model combines negative feedback, incoherent feedforward loops, and bistability
- The model recapitulates experimental observations of separate zones of tumor control

Sontag, 2017, Cell Systems *4*, 231–241
February 22, 2017
http://dx.doi.org/10.1016/j.cels.2016.12.003

CellPress

Cell Systems
Math | Bio

A Dynamic Model of Immune Responses to Antigen Presentation Predicts Different Regions of Tumor or Pathogen Elimination

Eduardo D. Sontag[1,2,*]
[1]Department of Mathematics and Center for Quantitative Biology, Rutgers University, New Brunswick, NJ 08903, USA
[2]Lead Contact
*Correspondence: eduardo.sontag@rutgers.edu
http://dx.doi.org/10.1016/j.cels.2016.12.003

SUMMARY

The immune system must discriminate between agents of disease and an organism's healthy cells. While the identification of an antigen as self/non-self is critically important, the dynamic features of antigen presentation may also determine the immune system's response. Here, we use a simple mathematical model of immune activation to explore the idea of antigen discrimination through dynamics. We propose that antigen presentation is coupled to two nodes, one regulatory and one effecting the immune response, through an incoherent feedforward loop and repressive feedback. This circuit would allow the immune system to effectively estimate the increase of antigens with respect to time, a key determinant of immune reactivity in vivo. Our model makes the prediction that tumors growing at specific rates evade the immune system despite the continuous presence of antigens indicating disease, a phenomenon closely related to clinically observed "two-zone tolerance." Finally, we discuss a plausible biological instantiation of our circuit using combinations of regulatory and effector T cells.

INTRODUCTION

In vertebrates, the innate and adaptive immune systems combine to provide a finely orchestrated multicomponent host defense mechanism that protects against infective pathogens and also helps to identify and eliminate malignantly transformed cells (Kindt et al. 2013; Abbas et al., 2016). A prerequisite for successful immune action is the ability to distinguish agents of disease from an organism's own healthy cells. A central role in this discriminatory ability is played by antigens, which are molecules capable of inducing an immune response. Non-self-antigens appear not only in pathogens but also in malignantly transformed cells because of their overexpression of normal proteins, mutated proteins, or oncogenic viruses. In its normal functioning, the immune system is trained not to react improperly to self-antigens. This paradigm of "self/non-self pattern" recognition traces back to Burnet (1957a) and related ideas in Talmage (1957).

The static view, however, is not entirely consistent with a number of phenomena which hint at a role for discrimination based on the antigen's dynamic features, as suggested by the following examples: (1) The presence of commensal bacteria (microbiome) is stably tolerated by the immune system despite the presence of bacterially derived, non-self molecules that exist in intimate proximity to the host (Pradeu, 2012). (2) Slow-growing tumors are known to evade the immune system (Grossman and Berke, 1980) despite their expression of non-self antigens. (3) Decreased activation of natural killer cells is observed under chronic receptor activation (Pradeu et al., 2013) despite the continued presence of antigens. (4) There is reduced capacity of a host to respond to the pro-inflammatory stimuli of bacterial signatures such as lipopolysaccharides after a first exposure to the same type of stimulus, a phenomenon called endotoxin tolerance in macrophages (also known as deactivation, desensitization, adaptation, or reprogramming) (West and Heagy, 2002). Notably, autoimmune diseases also offer examples of the immune system responding to changes in antigen presentation. For example, Pradeu (2012) argues that many autoimmune diseases appear during puberty, when relatively fast changes occur in the physiology of the host, or due to a sudden exposure to chemical or biological agents, and, conversely, allergy treatment by slow desensitization through antigen exposure leads to tolerance (Burks et al., 2012). These examples suggest that, to understand the immune response, one might want to also consider *dynamic* features of antigen presentation as a complement to discrimination mechanisms that are only based on a static response that depends on the presence or absence of antigen.

During the past 30 or so years, a number of authors, most notably Grossman and Berke (1980) and Pradeu (2012), have proposed the necessity of incorporating dynamics of antigen presentation when attempting to understand the body's decision to initiate the immune response. They support this line of reasoning by experimental observations that (1) lymphocytes mount a sustained response only when faced with a sufficiently steep increase in their level of stimulation (for example, acute antigen presentation, proliferation rates of infected cells or tumors, stress signals) and (2) even when a new motif triggers an immune response, its chronic presence may result in adaptation: downregulation or even complete termination of the inflammatory response. One mathematical formulation was introduced by Grossman and Paul (1992), who postulated the "tunable activation threshold" model for immune responses: effector cells in the innate or adaptive systems should become tolerant to continuously expressed

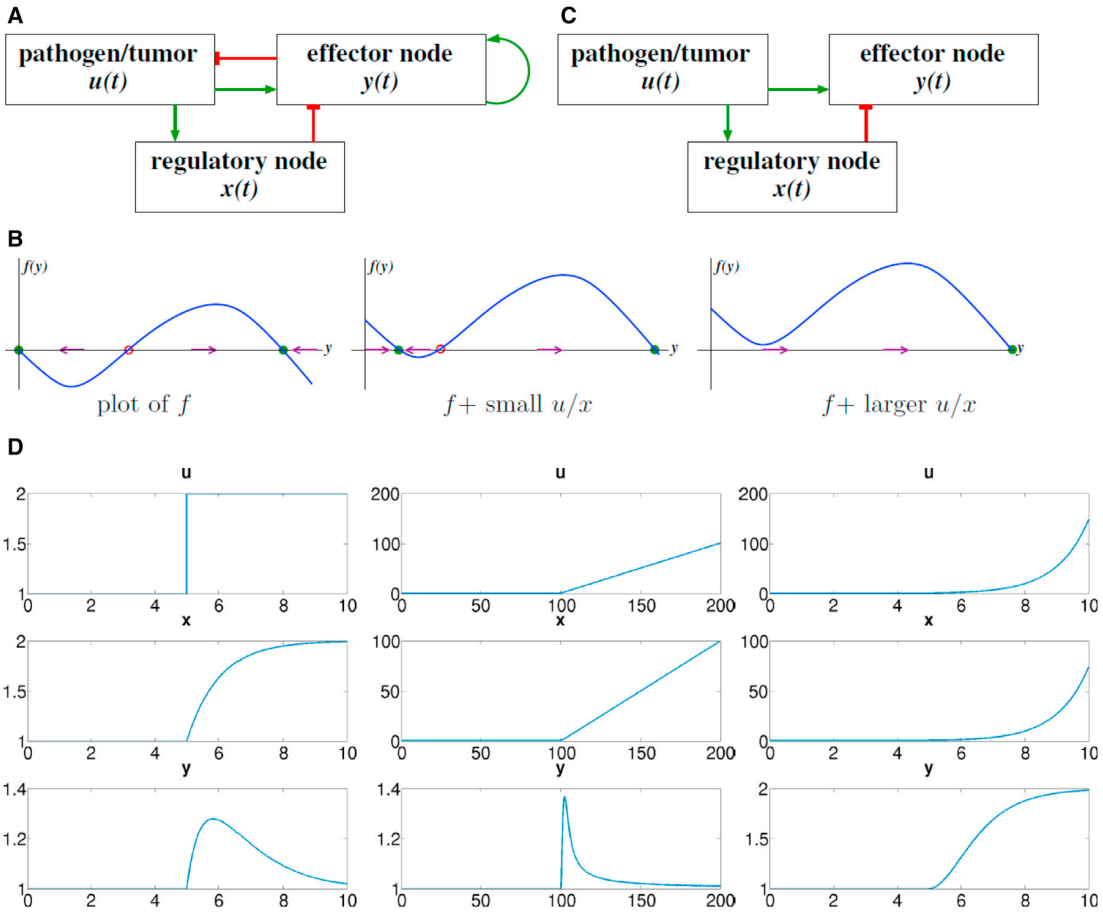


Figure 1. The Model Considered in This Paper, and the Effect of Incoherent Feedforward Loops

(A) The model being considered in this paper. Blunt-end red arrows denote repression, and green arrows denote activation or autocatalytic feedback.

(B) Plot of the nonlinear function *f* that endows the system with bistability (left) and two translates of this plot; open red and filled green circles denote unstable and stable states, respectively.

(C) Simplified model when feedbacks are ignored, an incoherent feedforward loop (IFFL).

(D) Simulations of system, with input u(t) switching in the middle of the interval from u(t) = 1 to a new constant value u(t) = 2, a shifted ramp u(t) = t, and a shifted exponential u(t) = e^t, respectively; also shown are x(t) and y(t). In every case, the horizontal axis is time, *t*. Parameters used were: $\delta_x = \beta = \mu = \delta_y = 1$; initial conditions $x(0) = y(0) = 1$.

motifs, or even gradually increasing ones, but should induce an effector response when a steep change is detected. Among recent variations upon this theme are the "discontinuity theory" postulated by Pradeu et al. (2013) and the "growth threshold conjecture" by Arias et al. (2015) (see STAR Methods section A for details). Here I present an extremely simple, conceptual mathematical model that describes how the immune system may discriminate between immune challenges based on their dynamics. I go on to demonstrate that it can capture a clinically important behavior, "two-zone tolerance," in which tumors growing at specific rates evade the immune system.

RESULTS

My model consists of three ordinary differential equations as I will describe in detail below. They are:

$$\begin{aligned} \dot{u} &= (\lambda - \kappa y)u \\ \dot{x} &= -\delta_x x + \beta u \\ \dot{y} &= h(u/x) + f(y) \end{aligned} \qquad \text{(Equations 1A–1C)}$$

The dots indicate time derivatives (see Figure 1). The constants $\lambda, \kappa, \delta_x, \beta$, are positive, and they represent reaction constants as discussed below. The function *h* is continuous, strictly increasing, and satisfies $h(0) = 0$; for our purposes we could simply take a linear function, $h(p) = \mu p$, where μ is a constant and *p* represents *u*/*x*. Qualitatively, the plot of the function *f* is as in Figure 1B, left.

We view the system described by Equations 1A–1C as a "toy model" that encompasses immune suppression as well as pattern discrimination based on antigen dynamics. The variable *u* represents an immune challenge, such as the volume or number of cells in a tumor, an infection, or the number of antigen-presenting cells (APCs) of a certain type, and the term λu in the

equation for u represents its exponential rate of growth. The variable x represents what I will call an "intermediate regulatory node" because it is driven by u and in turn regulates y. This node has an explicit biological correlate, such as the number of T regulatory (Treg) cells in a defined tumor microenvironment. Within the model, this variable evolves according to a linear activation, proportional to the immune challenge, and decays linearly (inactivation or degradation). The variable y represents an agent that can eliminate the challenge u, such as the number of tumor-specific cytotoxic T cells (CTL) in the same environment. This elimination is represented by the mass-action term $-\kappa yu$ in the equation for u. As I will discuss in detail below, the variable y increases in proportion to the ratio of u to x, which implies that y is driven by the rate of growth of u. The function $f(y)$ combines both degradation of y and a positive autocatalytic feedback between the presence of y and its induction, such as is observed when T cell activation is induced by cytokines; this function is chosen so as to endow the y component with bistable behavior, as explained next.

Initial intuition regarding the system can be obtained by thinking of $p = h(u/x)$ as a parameter in the scalar differential equation (Equation 1C), writing $\dot{y} = p + f(y)$ and temporarily ignoring that, in the full model p is not constant but it depends on y through a feedback repression of u. Note that, in this thought exercise, the plot of $p + f(y)$ is a vertical translation by p of the plot of f. If p is small, then, starting from the initial condition $y(0) = 0$, the solution $y(t)$ approaches asymptotically a low value of y, the leftmost stable green-labeled point in the center plot in Figure 1B. However, once that p has a value large enough that this first equilibrium disappears (in what is known as a "saddle-node bifurcation") then $y(t)$ will converge, instead, to a comparatively large value, the green stable point in Figure 1B, right. This higher equilibrium represents the triggering of a substantially increased immune response.

Further intuition can be gleaned from another simplification. Let us think of u as an input to Equations 1B and 1C, once again ignoring the repression of u by y (see Figure 1C). In addition, let us take $h(u/x) = \mu\, u/x$ and let the function f include only degradation or deactivation but no autocatalytic feedback. Thus, Equations 1B and 1C simplify to

$$\begin{aligned} \dot{x} &= -\delta_x x + \beta u \\ \dot{y} &= \mu u/x - \delta_y y, \end{aligned} \qquad \text{(Equation 2)}$$

where $\beta, \mu, \delta_x, \delta_y$ are some positive constants, and u is viewed as an external stimulus. The resulting system is a type of incoherent feedforward loop (IFFL), a "motif" which is ubiquitous in biological networks (Milo et al., 2002; Alon, 2006) and is statistically enriched in many intracellular networks (metabolic pathways, genetic circuits, kinase-mediated signaling) as well as the intercellular level. IFFLs are characterized by the existence of two antagonistic ("incoherent") alternative pathways from the input to the output; these paths can be direct or indirect (Figure 1C).

For our model, there is an activating, direct path within the IFFL from the stimulus u (given as a step function, for example, in Figure 1D, left) to the effector node y. In addition, there is an inhibitory, indirect path from stimulus u to node y: u activates regulatory node x, which, in turn, represses y. This structure endows IFFLs with powerful signal processing capabilities, studied in detail in Alon's textbook (Alon, 2006). Because the indirect effect requires an accumulation of x over time, there is typically a delay in the downregulation of y, leading to a response that consists of a short activity burst followed by a return to a basal value which is the same regardless of the magnitude of the input, as shown in Figure 1D, left. The return to a basal value independent of the excitation magnitude is a phenomenon called "perfect adaptation" (Alon, 2006; Sontag, 2003) and holds also for ramps (linearly increasing) inputs $u(t)$; see, for example, Figure 1D, center. However, when, instead, we consider an exponentially growing input, such as the solution $u(t) = u(0)e^{\nu t}$ of the differential equation $\dot{u} = \nu u$, then the response $y(t)$ approaches a constant multiple of the rate ν instead of returning to its basal value (Figure 1D, right). Combining these two observations, we see that the system has two interesting properties. First, it effectively estimates the growth constant ν if the input is exponentially growing, and second, it does not respond persistently to subexponential inputs. The STAR Methods includes a formal proof of this property. In this context, if we take the autocatalytic feedback term in f into account, then a large value of u/x, which depends on the steepness of the input u, may trigger an irreversible transition to a higher state for y, which persists even after the excitation goes away. An important property of our system, which will allow us to simplify considerably its mathematical analysis is that it incorporates a subsystem, the IFFL system just discussed, that has the "fold change detection" or "scale invariance" property: if the input is scaled by a positive constant p, then the same response y is obtained, provided that the regulatory variable x is also scaled by p. One uses the term "fold detection" for the initial activity burst, motivated by the response to an input that switches from $u(t) = u_-$ for t, 0 to $u(t) = u_+$ for $t \geq 0$: assuming that $x(0)$ is adapted to u_-, the initial amplitude of u/x, which triggers the initial change in response in y, is $(\delta_x/\beta)u_+/u_-$, and u_+/u_- is the fold change in the input. In the IFFL system Equation 2, an input that is scaled by a positive constant p leads to the same response y, provided that the regulatory variable x is also scaled by p, since one has

$$\dot{(px)} = -\delta_x(px) + \beta(pu)$$

and

$$\dot{y} = \mu u/x - \delta_y y = \mu(pu)/(px) - \delta_y y\,.$$

In mathematical terms, this says that the equations do not change under the one-parameter Lie group of transformations

$$(u, x, y) \to (pu, px, y).$$

In other words, this system is "scale-invariant" (or, in different terminology, it is a "fold change detector") (Shoval et al., 2010). The STAR Methods include a more detailed discussion of scale invariance for this system. Irrespectively of it being a property that is very useful in our mathematical analysis, scale invariance allows a system to detect relative, as opposed to absolute, changes in input signals and to do so robustly even when intermediates in signaling pathways are varied (Shoval et al., 2010).

In summary, our model combines three central motifs in modern systems biology: an IFFL for estimating the rate of growth of

$u(t)$, which is scale-invariant; multi-stable dynamics for $y(t)$; and feedback control of $y(t)$ that represses the input $u(t)$.

In spite of its extreme simplicity, analysis of the model reveals a key biological implication: an immune challenge, for instance a tumor, can be eliminated in more than one range of growth rates. Specifically, we find three thresholds, λ_1, λ_2, λ_3, so that

- if the per-capita rate of growth λ of the tumor is less than λ_1, then it is eventually eliminated by the immune system;
- when the tumor is more aggressive, $\lambda > \lambda_1$ but $\lambda < \lambda_2$, it cannot be eliminated (it is "tolerated" by the immune system);
- for an even more aggressive challenge, with $\lambda > \lambda_2$ but $\lambda < \lambda_3$, again the tumor is eliminated; and
- if $\lambda > \lambda_3$, again there is no elimination (the tumor "escapes").

Intuitively, in the intermediate range $\lambda_2 < \lambda < \lambda_3$, the immune system goes into "overdrive," engaging additional resources through activation of a positive feedback mechanism, and this level is powerful enough to effectively repress the challenge. As we mentioned earlier, the impact of growth rates on immune responses has been a hot topic of discussion in the immunology literature, and it has been proposed that different rates provide a way to differentiate among threats based on their aggressivity. Non-self and potentially dangerous cells presumably reproduce faster compared to self and also beneficial microorganisms. As we will discuss, the role of *exponential* rates in determining immune response has also been the subject of experimental research, including a recent immunotherapy patent, and the role of T suppressor cells in providing what we may now view as the regulatory node in an incoherent feedforward loop has been well established experimentally as well.

The prediction of the existence of disjoint regions of tumor elimination, depending on rate of growth, remains to be tested. However, this behavior is strongly reminiscent of the well-known phenomenon of two-zone tumor tolerance, which has been observed in the experimental literature since the mid-1960s (Gatenby et al., 1981; Kölsch and Mengersen, 1976; Li et al, 2016). This phenomenon is analogous to our predictions, the only difference being the zones are now determined by the magnitude of an initial tumor inocula in animal subjects instead of tumor growth rates. Moreover, the model is capable of logarithmic sensing and scale invariance. These phenomena, and the origins of two-zone tolerance, will be discussed in detail below.

Mathematical Analysis

A key step in our analysis will be an exact reduction to a certain two-dimensional system, as such a system is far easier to analyze than the full three-variable one. Indeed, scale invariance under the transformations $(u,x,y) \mapsto (pu,px,y)$ suggests performing a change of variables in which x is replaced by $p = u/x$. As $x = u/p$, the original variables (u,x,y) can be recovered from (u,p,y) so the transformation is invertible. Rearranging the order of equations, we will from now on study the system in these new coordinates:

$$\begin{aligned} \dot{y} &= h(p) + f(y) \\ \dot{p} &= p\,(\delta_x + \lambda - \kappa y - \beta p) \\ \dot{u} &= (\lambda - \kappa y)u. \end{aligned} \qquad \text{(Equations 3A–3C)}$$

Writing the equations in these new coordinates has a major advantage for analysis because the equations for y and p are now decoupled from u, and, therefore, we may study the two-dimensional (y,p) system using techniques suitable for planar systems, such as phase planes and nullclines. Information regarding the asymptotic behavior of u can be inferred from that of the (y,p) system. Suppose that we have determined that a solution $(y(t),p(t))$ tends to an equilibrium $(\overline{y},\overline{p})$ with $\overline{p} \neq 0$. This implies that $\lambda - \kappa y \to \lambda - \kappa \overline{y}$ as $t \to \infty$, and, because $\delta_x + \lambda - \kappa \overline{y} - \beta \overline{p} = 0$ at any equilibrium with $\overline{p} \neq 0$, asymptotically $u(t) \propto e^{\nu t}$ with $\nu = \lambda - \kappa \overline{y} = \beta \overline{p} - \delta_x$. This allows us to decide if $u(t)$ converges to zero or infinity (that is, whether the immune challenge is eliminated or grows without limit):

$$\overline{p} < \delta_x/\beta \;\Rightarrow\; \nu < 0: \quad \text{elimination (rejection)}$$

$$\overline{p} > \delta_x/\beta \;\Rightarrow\; \nu > 0: \quad \text{proliferation (escape, tolerance)}.$$

Thus, from now on, we study Equations 3A and 3B. Recall the plot of f is as in Figure 1B and that h is assumed to be strictly increasing with $h(0) = 0$, which we will later specialize to $h(p) = \mu p$. We are only interested in solutions with $y(t) \geq 0$ and $p(t) \geq 0$.

The equilibria of system (Equations 3A and 3B) are obtained by simultaneously solving

$$h(p) + f(y) = 0$$

and

$$p(\delta_x + \lambda - \kappa y - \beta p) = 0.$$

When $p = 0$, the first of these is simply $f(y) = 0$ (because $h(0) = 0$), and when $p > 0$, the second gives $\delta_x + \lambda - \kappa y - \beta p = 0$. The y and p nullclines of this system are the subsets of the first quadrant where $\dot{y} = 0$ and $\dot{p} = 0$; that is,

$$N_1 = \{(y,p) \mid y \geq 0, p \geq 0, p = h^{-1}(-f(y))\}$$

and

$$N_2 = \{p = 0\} \cup \{(y,p) \mid y \geq 0, p \geq 0, p = (1/\beta)(\delta_x + \lambda - \kappa y)\},$$

respectively. Note that (y,p) can only belong to N_1 if $-f(y)$ is in the range of h, and, in particular, this means that $f(y) \leq 0$, thus ruling out the values of y for which the plot of f is positive. Similarly, (y,p) can only belong to N_2 if $y \leq (\delta_x + \lambda)/\kappa$. The equilibria of the system are the points at the intersection of these two sets.

We are interested in analyzing the behavior of the system for different values of the parameter λ, which quantifies the initial growth rate of the immune challenge. The only place where λ plays a role is the equation for the line $p = (1/\beta)(\delta_x + \lambda - \kappa y)$, picking one of its parallel translates (Figure 2).

We assume, for simplicity, there are no more than two positive intersections between any (green) line $p = (1/\beta)(\delta_x + \lambda - \kappa y)$ and the (blue) y nullcline; this means that the slope $-\kappa/\beta$ is not ≈ 0. We also assume for the values of λ that we analyze that there is, at most, one intersection, like $\hat{\xi}$, on the decreasing branch of the blue curve. (If there are such additional intersections, the theoretical analysis is somewhat more complicated

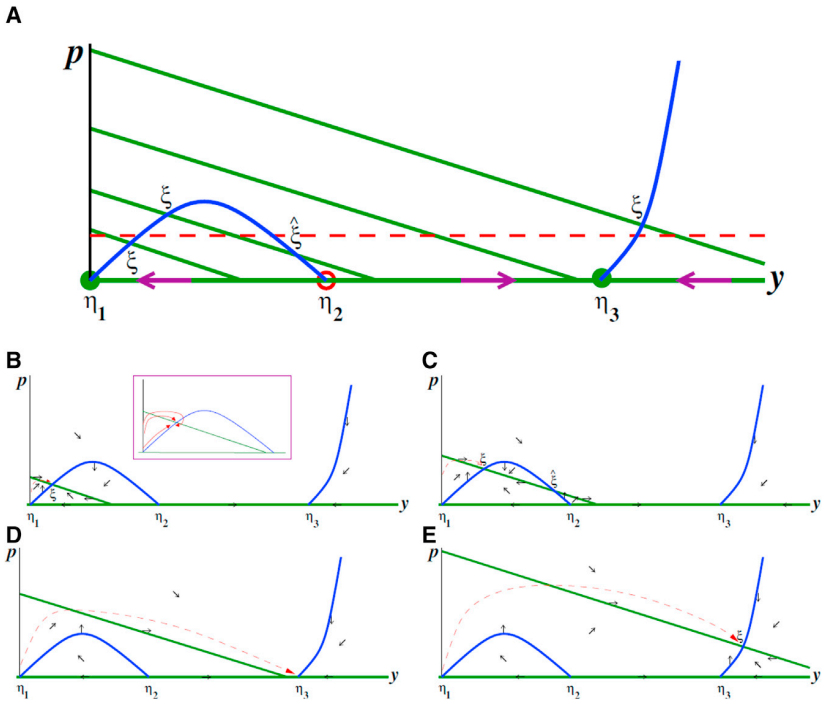


Figure 2. Nullcline Analysis of Model

(A) Nullclines for the system (Equations 3A and 3B) on the quadrant y ≥ 0, p ≥ 0. Blue curve is y-nullcline. Green lines are p-nullclines, shown for four typical values of λ. There are two different types of equilibria: the η's are equilibria with p = 0 and the ξ's with p > 0. The four points labeled ξ are stable for the reduced dynamics, for the respective λ's, as discussed in the text. The intercepts of the green lines with the *p* axis (vertical) are at the locations $p=(1/\beta)(\delta_x+\lambda_i)$, i = 1, 2, 3, 4. Red dashed horizontal line is threshold $p=\delta_x/\beta$, discussed in the text.

(B–E) Directions of movement in each region determined by the nullclines along with typical trajectories, respectively, low to high for each of the values of λ represented in (A). The inset in (B) shows various possibilities for approach of trajectory to the point labeled ξ.

Related to the STAR Methods.

but is analogous.) The dashed line is the threshold that determines the behavior of u: if $(y(t),p(t))$ converges to an equilibrium that has $\overline{p}<\delta_x/\beta$, then $u(t)\to 0$ as $t\to\infty$, but if $\overline{p}>\delta_x/\beta$, then $u(t)\to\infty$.

The Jacobian matrix for the system (Equations 3A and 3B), evaluated at a generic equilibrium $(\overline{y},\overline{p})$, is

$$J = \begin{pmatrix} f'(\overline{y}) & h'(\overline{p}) \\ -\kappa\overline{p} & \delta_x+\lambda-\kappa\overline{y}-2\beta\overline{p} \end{pmatrix}$$

and, therefore,

$$J = \begin{pmatrix} f'(\overline{y}) & h'(0) \\ 0 & \delta_x+\lambda-\kappa\overline{y} \end{pmatrix}$$

if $\overline{p}=0$ and

$$J = \begin{pmatrix} f'(\overline{y}) & h'(\overline{p}) \\ -\kappa\overline{p} & -\beta\overline{p} \end{pmatrix}$$

if $\overline{p}>0$, where we used that

$$\delta_x+\lambda-\kappa y-2\beta\overline{p}=-\beta\overline{p}$$

when $\overline{p}>0$ (since the equilibrium condition is $\delta_x+\lambda-\kappa y-\beta\overline{p}=0$ in that case). Note that these cases correspond, respectively, to points labeled η and ξ in Figure 2.

At points with $\overline{p}=0$, the eigenvalues of the upper triangular matrix J are $f'(\overline{y})$ and $\delta_x+\lambda-\kappa\overline{y}$. Referring to the plot of f in Figure 1B, $f'(\overline{y})>0$ at the point η_2 and $f'(\overline{y})<0$ at the points η_1 and η_3. This is because the blue curve in Figure 2 is $h^{-1}(-f(y))$ and, therefore, is qualitatively like an upside-down version of the plot of f so that positive (respectively, negative) f' corresponds to negative (respectively, positive) slope at the equilibrium. Thus, η_2 is unstable. At the point η_1, the second eigenvalue is $\delta_x+\lambda-\kappa\overline{y}=\delta_x+\lambda$, which is positive, so η_1 is also unstable (a saddle point). The stability of η_3 depends on the sign of $\delta_x+\lambda-\kappa\overline{y}$: η_3 is stable if this sign is negative, that is if

$$\delta_x+\lambda-\kappa\overline{y}<0,$$

and unstable otherwise. Suppose that all parameters are fixed except for λ. We use Figure 2 to determine stability: if λ is such that the corresponding line

$$p=(1/\beta)(\delta_x+\lambda-\kappa y)$$

intersects the y axis at a point y_0 to the left of η_3 (that is, if $y_0<\overline{y}$), then $\delta_x+\lambda-\kappa y_0=0$ implies $\delta_x+\lambda-\kappa\overline{y}<0$, and, hence, η_3 is stable. If, on the other hand, λ is larger and $p=(1/\beta)(\delta_x+\lambda-\kappa y)$ intersects the y axis at a point y_0 to the right of η_3, then $\delta_x+\lambda-\kappa y_0=0$ implies $\delta_x+\lambda-\kappa\overline{y}<0$, so we conclude that η_3 is unstable. In Figure 2, only for the largest λ (top line) is η_3 unstable.

We next analyze stability of the equilibrium points for which $\overline{p}>0$, labeled ξ or $\hat{\xi}$ in Figure 2. Stability is equivalent to the trace of J being negative and its determinant positive (Hirsch and Smale, 1974; Sontag, 1998); that is,

$$-\mathrm{tr}\,J=\beta\overline{p}-f'(\overline{y})>0$$

and

$$(1/\overline{p})\det J=\kappa h'(\overline{p})-\beta f'(\overline{y})>0.$$

Because h is increasing and all constants are positive, these expressions are both positive when $f'(\overline{y})<0$. Remembering again that the blue curve is an "upside-down" version of f, we conclude that all intersections ξ shown in Figure 2 are stable, with the possible exception of the intersection labeled $\hat{\xi}$, which

Box 1. Case Study: A Concrete Mathematical Model Featuring Treg Cells and Cytokines

In aggregate, the results of our model analysis show the role of the underlying IFFL as an estimator of the exponent λ of immune challenge growth and the existence of four regimes, alternating tolerance and rejection, which correspond to different values of λ. In our abstract theoretical arguments, we did not specify the function $f(y)$, except for the requirement that its graph be as in Figure 1B. We view the model as qualitative and phenomenological, merely as an illustration of behaviors that can arise from, and can be easily explainable by, simple motifs and not necessarily instantiated by a specific biological system. Nonetheless, it is fair to ask if there is a plausible biological system that gives rise to an $f(y)$ of this form and for which our conclusions hold true. We address this issue now, using as a guide simplified versions of standard models in immune dynamics such as found in Kuznetsov et al. (1994) and Kirschner and Panetta (1998) and using simulations to verify for this system the theoretically predicted four-regime behavior. We will pick $h(p) = \mu p$, a linear function, and

$$f(y) = \frac{Vy^2}{K+y} - \varepsilon y^2 - \delta_y y.$$

The term $Vy^2/(K+y)$ in this expression can be thought of as representing a cytokine-mediated positive feedback, as discussed in the STAR Methods. Motivated by a similar term in Khailaie et al. (2013), the term $-\varepsilon y^2$ can be thought of as representing cell contact-dependent, activation-induced cell death ("fratricide") through the Fas receptor/FasL ("death ligand") mechanism for T cell homeostasis suggested by Callard et al. (2003). Finally, the term $-\delta_y y$ represents a linear constitutive deactivation and/or degradation. The resulting system is, therefore,

$$\begin{aligned} \dot{u} &= (\lambda - \kappa y)u \\ \dot{x} &= -\delta_x x + \beta u \\ \dot{y} &= \mu \frac{u}{x} + \frac{Vy^2}{K+y} - \varepsilon y^2 - \delta_y y \end{aligned},$$

and the reduced system written in (y,p) coordinates is

$$\begin{aligned} \dot{y} &= \mu p + \frac{Vy^2}{K+y} - \varepsilon y^2 - \delta_y y \\ \dot{p} &= p\,(\delta_x + \lambda - \kappa y - \beta p). \end{aligned}$$

We take the units of time to be days. Cell populations (u,x,y) are in units of 10^6 cells, but $p = u/x$ is non-dimensional. In this model, *u, x, y* represent, respectively, populations of tumor, Treg, and effector T cells. The parameters we used, and corresponding units, are as follows: $\mu = 10$ day^{-1}, $V = 0.25$ day^{-1}, $\delta_x = 0.1$ day^{-1}, $\delta_y = 0.1$ day^{-1}, $\beta = 1$ day^{-1}, $\varepsilon = 10^{-5}$ day^{-1} (10^6 cells)$^{-1}$, $K = 100$ (10^6 cells), and $\kappa = 10^{-5}$ (10^6 cells)$^{-1}$ day^{-1}. The particular choice of algebraic forms and parameters is discussed in the STAR Methods; basic descriptions of the model's behavior can also be found in the STAR Methods.

Shown in Figures 3A–3D are simulations of the complete closed-loop system, where we now included a carrying capacity term for the immune challenge: $\dot{u} = [\lambda(1 - Bu) - \kappa y]u$. When u is small, the term Bu is dominated by the other terms, but we include this term for numerical convenience to avoid blow-up of solutions in the case when u is unstable; in any event, such a term is biologically realistic. We picked, specifically, $B = 10^{-3}$ (STAR Methods). As predicted, for increasing λ, there are alternating decay and growth behaviors for $u(t)$. Notably, for the third value, $\lambda = 10^{-1}$, the immune challenge load $u(t)$ only starts decreasing after about 120 days. This timescale happens to be, purely coincidentally, the start of remission observed in some patients under ipilimumab (CTLA-4 checkpoint blockade therapy; Wolchok 2010). For this model and parameters, Figure 3E shows the asymptotic value of u, the immune challenge, as a function of the parameter λ, clearly illustrating the four-regime phenomenon. In the STAR Methods, similar results are shown to hold for a model in which the effect of regulatory variables is through an inactivation of a helper cell population. The existence of disjoint regions of tumor elimination depending on rate of growth is reminiscent of "sneaking through." The idea of tumors "sneaking through" from immune control can be traced to the mid-1960s, when Klein (1966) found the "preferential take of tumors after small size inocula to a similar degree with that seen with large size inocula, compared to the rejection of medium sized inocula." Put simply, there is an intermediate region in which tumors can be eliminated. This picture is at least consistent with a larger initial rate of increase in exposure, leading to tumor suppression, as in our model. These ideas were explored experimentally in Gatenby et al. (1981) and Kölsch and Mengersen (1976) and are analogous to behavior that can be seen in our model (Figure 4), but with the important proviso that the numbers represent tumor incidence and that we work with rates of increase instead of initial tumor size (see STAR Methods for more discussion).

Our analysis is merely a phenomenological toy model which does not specify immune components. Nonetheless, one might speculate that, as far as T cell activation and deactivation, Tregs may play a role as a regulating intermediate variable x. Tregs are a type of CD4$^+$ cell that play "an indispensable role in immune homeostasis" (Josefowicz et al., 2012). They arise during maturation in the thymus from autoreactive cells ("natural Tregs") or are induced at the site of an immune response in an antigen-dependent manner

(Continued on next page)

Box 1. *Continued*

("induced Tregs"). They are thought to play a role in limiting cytotoxic T cell responses to pathogens, and Treg$^-$ mice have been shown to suffer from extreme inflammatory reactions. It is known from animal studies that Tregs inhibit the development of autoimmune diseases, such as experimentally induced inflammatory bowel disease, experimental allergic encephalitis, and autoimmune diabetes (Owen et al., 2009). Moreover, the involvement of T suppressor cells (now Tregs) in regulating the immune response to tumor has a long history; see, for instance, Fujimoto et al. (1976).

Our analysis suggests that the immune system might act as an estimator of the rate of growth of an immune challenge and, more specifically, the rate of exponential increase. In recent work, Kundig et al. (2008) and Johansen et al. (2008) emphasized the role of antigen kinetics in determining immune reactivity, with exponentially increasing antigenic stimulation recognized by the immune system as a pattern associated with pathogens and, thus, leading to strong immune responses. They showed that the dynamics of immune stimulation (through dendritic cell vaccination and for T cells stimulated *in vitro*) was by itself a factor in determining the strength of T cell and anti-tumor responses and obtained a patent for the idea of using exponentially increasing antigen stimulation (Kundig et al., 2008). See the STAR Methods for more discussion.

needs further analysis. At the point $\widehat{\xi}$, the slope of the green line is $-\kappa/\beta$, and the slope of the blue line is the derivative of $h^{-1}(-f(y))$ evaluated at $y=\bar{y}$, where $h(\bar{p})+f(\bar{y})=0$, so this slope is $-f'(\bar{y})/h'(\bar{p})$. Since the slope of the green line is larger (less negative) than the slope of the blue line, we have that $-\kappa/\beta > -f'(\bar{y})/h'(\bar{p})$, or, equivalently, $\kappa h'(\bar{p})-\beta f'(\bar{y})<0$, so the determinant of J is negative, which means that $\widehat{\xi}$ is a saddle point and, therefore, unstable.

These theoretical results strongly suggest that, for trajectories that start with small $p=u/x$, convergence will be to one of the points labeled ξ in Figure 2 or η_3 in the case of the line that does not intersect the blue curve. In other words, if λ, the reproduction rate of the immune challenge, such as a tumor, is small, then we will have an intersection below the threshold (Figure 2, dashed line), meaning that the challenge will be eliminated. For a larger, but not too much larger, λ, the intersection is above the threshold, so the challenge will be tolerated (it will grow). For λ's such that there are no intersections with the blue curve, solutions should converge to η_3, which is again under the threshold, implying a second zone of elimination of the challenge. Finally, for very large λs, the immune system is not capable of eradicating the challenge (ξ is over the threshold), so escape again occurs. This two-zone tolerance (at both intermediate and large λ) is reminiscent of analogous experimental findings (Box 1). The analysis based on stability of equilibria can be complemented by numerical simulations, done in Box 1 for an example, as well as a global phase-plane analysis, done below. A simplified case is analyzed further in the STAR Methods.

Observe that $\dot{y}>0$ in those regions of the phase plane where $h(p)+f(y)>0$, that is to say, where

$$p > h^{-1}(-f(y)).$$

Geometrically, this means that the flow of the vector field defining the system will point to the right at points that lie over the blue curve (y nullcline); it will point left under the blue curve and will be exactly vertical (or an equilibrium) on the curve itself.

Similarly for $p>0$, $\dot{p}>0$ in regions of the phase plane where

$$\delta_x+\lambda-\kappa y-\beta p>0,$$

that is to say, where

$$p<(1/\beta)(\delta_x+\lambda-\kappa y).$$

Geometrically, this means that the flow of the vector field defining the system will point up at points that lie under the green lines (p nullcline); it will point down over the green lines and will be exactly horizontal (or an equilibrium) on the lines. At $p=0$, the vector field is horizontal.

In summary, motions are "northeast," (NE) etc., according to these rules in each region delimited by the blue curve and the line $\delta_x+\lambda-\kappa y-\beta p=0$:

1. NE: over blue, under green
2. Southeast (SE): over blue, over green
3. Northwest (NW): under blue, under green
4. Southwest (SW): under blue, over green

Figure 2 shows the directions of movement in each region for each of the sample nullclines shown earlier as well as what a typical trajectory might look like, consistent with these directions of movement and converging to the stable points ξ or η_3. (The precise approach depends on the functional forms of f and h and their parameters. The inset in Figure 2B shows several possibilities.) As λ is increased, the equilibrium shown lies under, above, under, and finally again above the threshold. Box 1 discusses a specific model where these predictions are verified and relates our results to experimental evidence of similar phenomena.

DISCUSSION AND CONCLUSIONS

The study of immune systems and their interactions with tumors has long been the focus of theoretical and mathematical immunology (Jerne, 1974; Bell et al., 1978). Two influential contributions were the paper by Stepanova (1979), in which a set of two ODEs was used to represent tumor and immune system cells, and the paper by Kuznetsov et al. (1994), in which a similarly simple model was used to provide an explanation for the sneaking-through phenomenon, although with escape of small tumors and with no mechanism for detection of rates of change of the immune challenge. It is impossible here to review the literature in this very active area of research; some reviews and textbooks are Bell et al. (1978), Callard and Yates (2005), Andrew et al., (2007), Eftimie et al., (2011), Wodarz and Komarova (2014), de Pillis and Radunskaya (2014), and Vodovotz et al. (2016).

In this paper, we proposed a very simple phenomenological model that recapitulates some of the basic features of interactions

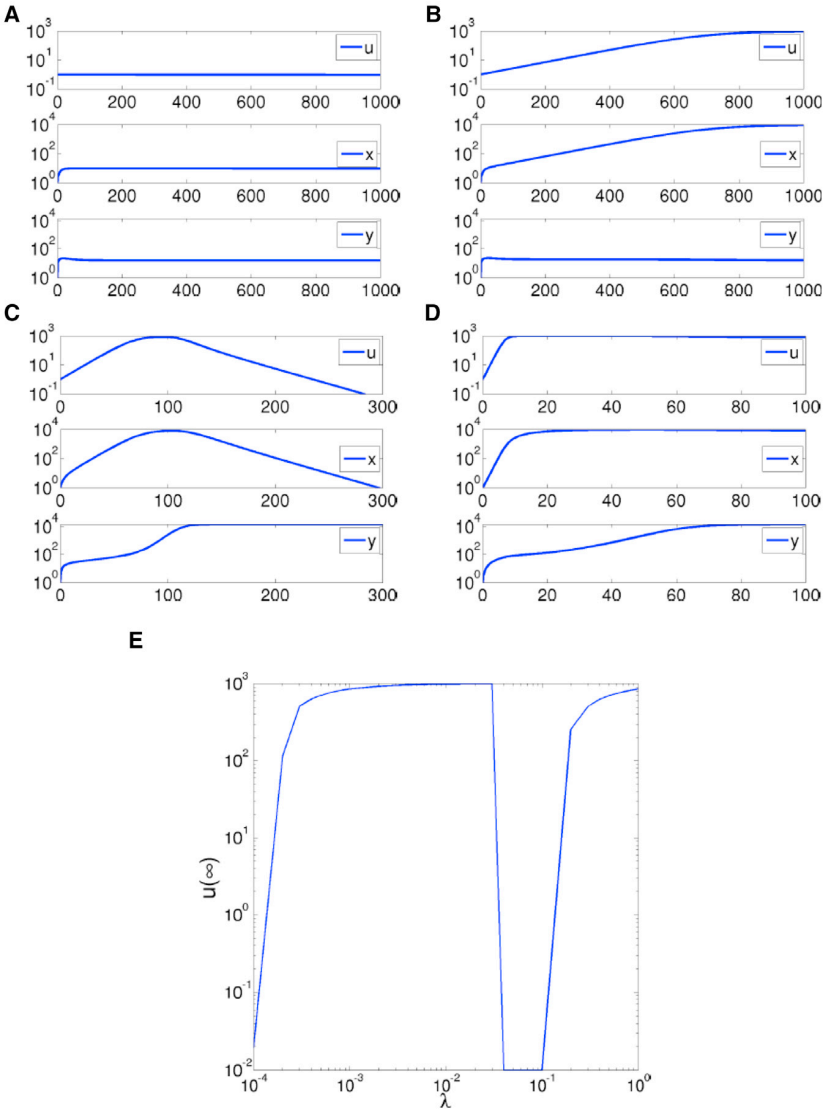


Figure 3. Simulations of Model

(A–D) Simulations of full system as described in text, for $\lambda = 10^{-4}$, 10^{-2}, 10^{-1}, and 1. (In A, u(t) converges to zero as $t \to \infty$, but very slowly.) Parameters as described in text. Initial states are always u = 1, x = 1, y = 0.

(E) Values of $u(\infty)$ plotted against λ, showing four regimes of elimination, tolerance, elimination, and escape (zero values shown as 10^{-2} to fit in log scale).

Related to the STAR Methods.

between the immune systems and tumors (or, more generally at this level of abstraction, other immune challenges) in the context of the estimation of tumor growth rates. The model leads to interesting conclusions regarding transitions between tolerance and elimination and the role of dynamics in self/non-self discrimination and makes contact with several theory and experimental papers. Obviously, our model represents a purely phenomenological, macroscopic, and hugely over-simplified view of a highly complex, intricate, and still poorly understood network of interactions between different components of the immune system as well as immune interactions with pathogens and tumors, Paraphrasing the well known quote, our model is "as simple as possible but not simpler" to illustrate the particular phenomena of interest.

Since Paul Ehrlich's work in 1909, (Ehrlich, 1909) the interplay with the immune system has been a controversial, although recently accepted, aspect of cancer biology. These ideas were formalized in Burnet's immune surveillance hypothesis (Burnet, 1957b) as an interaction between cancers that continuously arise and their repression by immune system, resulting in eventual elimination. Initial interest in this work was soon tempered by early experiments, but eventually new data led to a revival of these ideas in the late 1990s (Dunn et al., 2002). There is little doubt nowadays that immunosurveillance acts as a tumor suppressor, although it is also widely understood that the immune system can facilitate tumor progression by "sculpting" the immunogenic phenotype of tumors as they develop. Indeed, one current paradigm (Dunn et al., 2004) is the so-called "three Es of cancer immunoediting" hypothesis: elimination, equilibrium, and escape. The first of this corresponds with the classical immunosurveillance idea: the immune system successfully eradicates the developing tumor. In the equilibrium or immune-sculpting phase, the host immune system and any tumor cells that have survived the elimination phase enter into a dynamic quasi-steady state equilibrium during which the tumor cell population stays at sub-clinical levels. Ultimately, however, genetic and epigenetic heterogeneity in tumors, coupled by the Darwinian selection pressure exerted by the immune system, lead to the emergence of dominant clones with reduced immunogenicity which expand and become clinically detectable, and this is termed the escape phase. Our model does not directly address the effect of genetic or epigenetic modifications, and expanding it to do so remains a most interesting direction for further work.

Clinically, the interactions between the immune system and tumors are the focus of much current research because of the promise of novel immunotherapies such as checkpoint inhibitors (Pardoll 2012). It is worth pointing out the role of dynamical responses in immunotherapies compared to classical chemotherapy and pathway inhibitors, which is emphasized in the UpToDate physician reference guide, from which we quote from the September 1, 2015 version: "patients may have a transient worsening of disease, manifested either by progression of

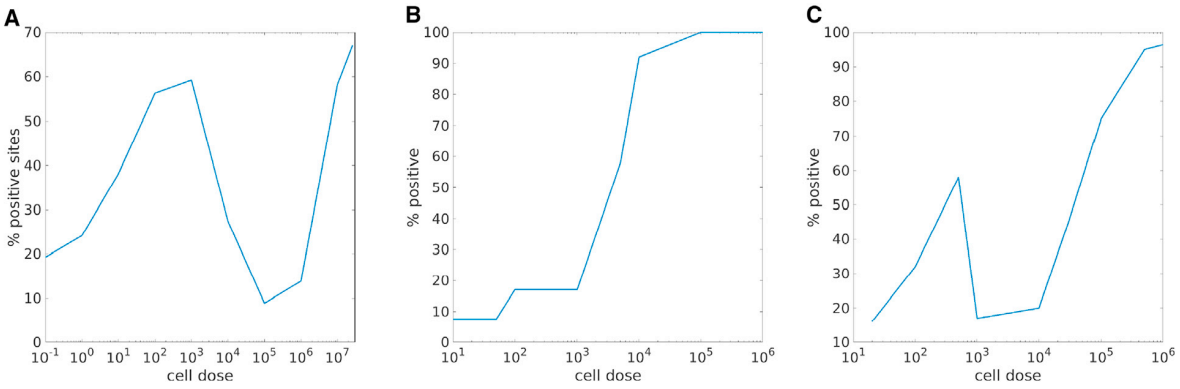


Figure 4. Experimental Plots Showing Four-Region Tumor Behavior

(A and B) Four regions of "sneaking through" (A) and no sneaking-through (B) when eliminating suppressor T cells; the plots were drawn from data in Gatenby et al. (1981). Very low zone (less than 100 cells), low tumor incidence; low zone (10^3–10^5 cells), tolerance of tumor; moderate zone (10–10^5 cells), immunogenic (low tumor); high zone (>10^6), again tolerogenic.

(C) Another illustration of the same general idea of four regions. The plot was drawn from data in Kölsch and Mengersen (1976).

The x axes in these plots measure the initial number of tumor cells and not the rate of growth λ of tumor as they did as in Figure 3. In both cases, the y axes measure the percentages of mice in which tumors were detected by some predetermined point. The STAR Methods discuss the experiments in more detail.

known lesions or the appearance of new lesions, before disease stabilizes or tumor regresses." (Interestingly, the solution in Figure 3B has this behavior). This statement helps justify, in our view, the introduction of dynamic systems concepts into the field of immunotherapy.

STAR★METHODS

Detailed methods are provided in the online version of this paper and include the following:

- KEY RESOURCES TABLE
- METHOD DETAILS
  - A. Comparison to other models
  - Tunable Activation Threshold (TAT)
  - Discontinuity theory of immunity
  - Growth threshold conjecture
  - B. Theory for system without autocatalysis
  - IFFL's responses to various classes of inputs
  - IFFL's as feedback controllers
  - Analysis of the closed-loop system
  - C. Perfect adaptation and scale-invariance
  - D. Details on the model used for simulations
  - Cell number units
  - The autocatalytic term $Vy^2/(K + y)$
  - The fratricide term $-\varepsilon y^2$
  - The decay terms $-\delta_x x$ and $\delta_y y$
  - The term βu
  - The term $\mu u/x$
  - The terms λu and $-\kappa y u$
  - Sensitivity to parameters in the function f
  - E. Nullclines for model and parameters used in text
  - F. A model with an intermediate population
  - G. More details on exponential rate detection and two-zone experimental results

SUPPLEMENTAL INFORMATION

Supplemental Information includes four figures and can be found with this article online at http://dx.doi.org/10.1016/j.cels.2016.12.003.

AUTHOR CONTRIBUTIONS

E.D.S. designed the research, performed mathematical analysis, and wrote the paper.

ACKNOWLEDGMENTS

This work was partially supported by NIH grant 433885 1R01GM100473, AFOSR grant FA9550-14-1-0060, and ONR grant N00014-13-1-0074.

Received: August 11, 2016
Revised: October 24, 2016
Accepted: December 2, 2016
Published: January 25, 2017

REFERENCES

Abbas, A.K., Lichtman, A.H.H., and Pillai, S. (2016). Basic Immunology: Functions and Disorders of the Immune System, Fifth Edition (St. Louis: Elsevier).

Alon, U. (2006). An Introduction to Systems Biology: Design Principles of Biological Circuits (Chapman & Hall).

Andrew, S.M., Baker, C.T.H., and Bocharov, G.A. (2007). Rival Approaches to Mathematical Modelling in Immunology. J. Comput. Appl. Math. *205*, 669–686.

Arias, C.F., Herrero, M.A., Cuesta, J.A., Acosta, F.J., and Fernández-Arias, C. (2015). The growth threshold conjecture: a theoretical framework for understanding T-cell tolerance. R. Soc. Open Sci. *2*, 150016.

Bell, G.I., Perelson, A.S., and Pimbley, G., Jr., eds. (1978). Theoretical Immunology (NY: Dekker).

Bocharov, G., Ludewig, B., Bertoletti, A., Klenerman, P., Junt, T., Krebs, P., Luzyanina, T., Fraser, C., and Anderson, R.M. (2004). Underwhelming the immune response: effect of slow virus growth on CD8+-T-lymphocyte responses. J. Virol. *78*, 2247–2254.

Burks, A.W., Jones, S.M., Wood, R.A., Fleischer, D.M., Sicherer, S.H., Lindblad, R.W., Stablein, D., Henning, A.K., Vickery, B.P., Liu, A.H., et al.; Consortium of Food Allergy Research (CoFAR) (2012). Oral immunotherapy for treatment of egg allergy in children. N. Engl. J. Med. *367*, 233–243.

Burnet, F.M. (1957a). A modification of Jerne's theory of antibody production using the concept of clonal selection. Aust. J. Sci. *20*, 67–69.

Burnet, M. (1957b). Cancer; a biological approach. I. The processes of control. Br Med J. *1*, 779–786.

Callard, R.E., and Yates, A.J. (2005). Immunology and mathematics: crossing the divide. Immunology *115*, 21–33.

Callard, R.E., Stark, J., and Yates, A.J. (2003). Fratricide: a mechanism for T memory-cell homeostasis. Trends Immunol. *24*, 370–375.

Clarke, F.H., Ledyaev, Y.S., Stern, R.S., and Wolenski, P.R. (1998). Nonsmooth Analysis and Control Theory. Graduate Texts in Mathematics (New York: Springer-Verlag).

Cohen-Saidon, C., Cohen, A.A., Sigal, A., Liron, Y., and Alon, U. (2009). Dynamics and variability of ERK2 response to EGF in individual living cells. Mol. Cell *36*, 885–893.

de Pillis, L.G., and Radunskaya, A. (2014). Modeling Tumor-Immune Dynamics. In Mathematical Models of Tumor-Immune System Dynamics. Proceedings in Mathematics & Statistics 107 (New York: Springer Verlag).

Dranoff, G. (2004). Cytokines in cancer pathogenesis and cancer therapy. Nat. Rev. Cancer *4*, 11–22.

Dunn, G.P., Bruce, A.T., Ikeda, H., Old, L.J., and Schreiber, R.D. (2002). Cancer immunoediting: from immunosurveillance to tumor escape. Nat. Immunol. *3*, 991–998.

Dunn, G.P., Old, L.J., and Schreiber, R.D. (2004). The three Es of cancer immunoediting. Annu. Rev. Immunol. *22*, 329–360.

Eftimie, R., Bramson, J.L., and Earn, D.J. (2011). Interactions between the immune system and cancer: a brief review of non-spatial mathematical models. Bull. Math. Biol. *73*, 2–32.

Ehrlich, P. (1909). Über Den Jetzigen Stand Der Karzinomforschung. Ned. Tijdschr. Geneeskd. *5*, 273–290.

Flaherty, D. (2011). Immunology for Pharmacy, First Edition (Elsevier).

Fujimoto, S., Greene, M.I., and Sehon, A.H. (1976). Regulation of the immune response to tumor antigens. II. The nature of immunosuppressor cells in tumor-bearing hosts. J. Immunol. *116*, 800–806.

Gatenby, P.A., Basten, A., and Creswick, P. (1981). "Sneaking through": a T-cell-dependent phenomenon. Br. J. Cancer *44*, 753–756.

Goentoro, L., and Kirschner, M.W. (2009). Evidence that fold-change, and not absolute level, of beta-catenin dictates Wnt signaling. Mol. Cell *36*, 872–884.

Grossman, Z., and Berke, G. (1980). Tumor escape from immune elimination. J. Theor. Biol. *83*, 267–296.

Grossman, Z., and Paul, W.E. (1992). Adaptive cellular interactions in the immune system: the tunable activation threshold and the significance of subthreshold responses. Proc. Natl. Acad. Sci. USA *89*, 10365–10369.

Hamadeh, A., Ingalls, B., and Sontag, E. (2013). Transient dynamic phenotypes as criteria for model discrimination: fold-change detection in Rhodobacter sphaeroides chemotaxis. J. R. Soc. Interface *10*, 20120935.

Haubeck, H.D., and Kölsch, E. (1982). Regulation of immune responses against the syngeneic ADJ-PC-5 plasmacytoma in BALB-c mice. III. Induction of specific T suppressor cells to the BALB/c plasmacytoma ADJ-PC-5 during early stages of tumorigenesis. Immunology *47*, 503–510.

Hirsch, M.W., and Smale, S. (1974). Differential Equations, Dynamical Systems and Linear Algebra (Academic Press).

Jerne, N.K. (1974). Towards a network theory of the immune system. Ann. Immunol. (Paris) *125C*, 373–389.

Johansen, P., Storni, T., Rettig, L., Qiu, Z., Der-Sarkissian, A., Smith, K.A., Manolova, V., Lang, K.S., Senti, G., Müllhaupt, B., et al. (2008). Antigen kinetics determines immune reactivity. Proc. Natl. Acad. Sci. USA *105*, 5189–5194.

Josefowicz, S.Z., Lu, L.F., and Rudensky, A.Y. (2012). Regulatory T cells: mechanisms of differentiation and function. Annu. Rev. Immunol. *30*, 531–564.

Keener, J., and Sneyd, J. (2009). Mathematical Physiology, Second Edition (New York: Springer-Verlag).

Khailaie, S., Bahrami, F., Janahmadi, M., Milanez-Almeida, P., Huehn, J., and Meyer-Hermann, M. (2013). A mathematical model of immune activation with a unified self-nonself concept. Front. Immunol. *4*, 474.

Kindt, T.J., Goldsby, R.A., Osborne, B.A., and Kuby, J. (2013). Kuby Immunology, Seventh Edition (New York: W.H. Freeman).

Kirschner, D., and Panetta, J.C. (1998). Modeling immunotherapy of the tumor-immune interaction. J. Math. Biol. *37*, 235–252.

Klein, G. (1966). Recent trends in tumor immunology. Isr. J. Med. Sci. *2*, 135–142.

Kölsch, E., and Mengersen, R. (1976). Low numbers of tumor cells suppress the host immune system. Adv. Exp. Med. Biol. *66*, 431–436.

Kundig, T., Bot, A., Smith, K.A., and Qiu, Z. 2008. "A Method for Enhancing T Cell Response." Google Patents. https://www.google.com/patents/CA2678353A1?cl=en.

Kuznetsov, V.A., Makalkin, I.A., Taylor, M.A., and Perelson, A.S. (1994). Nonlinear dynamics of immunogenic tumors: parameter estimation and global bifurcation analysis. Bull. Math. Biol. *56*, 295–321.

Lang, M., and Sontag, E.D. (2016). Scale-Invariant Systems Realize Nonlinear Differential Operators. ACC *2016*, 6676–6682.

Lazova, M.D., Ahmed, T., Bellomo, D., Stocker, R., and Shimizu, T.S. (2011). Response rescaling in bacterial chemotaxis. Proc. Natl. Acad. Sci. USA *108*, 13870–13875.

Leisegang, M., Wilde, S., Spranger, S., Milosevic, S., Frankenberger, B., Uckert, W., and Schendel, D.J. (2010). MHC-restricted fratricide of human lymphocytes expressing survivin-specific transgenic T cell receptors. J. Clin. Invest. *120*, 3869–3877.

Li, Q., Rane, L., Poiret, T., Zou, J., Magalhaes, I., Ahmed, R., Du, Z., Vudattu, N., Meng, Q., Gustafsson-Jernberg, Å., et al. (2016). Both high and low levels of cellular Epstein-Barr virus DNA in blood identify failure after hematologic stem cell transplantation in conjunction with acute GVHD and type of conditioning. Oncotarget *7*, 30230–30240.

Liston, A., and Gray, D.H. (2014). Homeostatic control of regulatory T cell diversity. Nat. Rev. Immunol. *14*, 154–165.

McBride, W.H., and Howie, S.E. (1986). Induction of tolerance to a murine fibrosarcoma in two zones of dosage–the involvement of suppressor cells. Br. J. Cancer *53*, 707–711.

Milanez-Almeida, P., Meyer-Hermann, M., Toker, A., Khailaie, S., and Huehn, J. (2015). Foxp3+ regulatory T-cell homeostasis quantitatively differs in murine peripheral lymph nodes and spleen. Eur. J. Immunol. *45*, 153–166.

Milo, R., Shen-Orr, S., Itzkovitz, S., Kashtan, N., Chklovskii, D., and Alon, U. (2002). Network motifs: simple building blocks of complex networks. Science *298*, 824–827.

Owen, J.A., Punt, J., and Stranford, S.A. (2009). Kuby Immunology, Seventh Edition (New York: W.H. Freeman).

Pardoll, D.M. (2012). The blockade of immune checkpoints in cancer immunotherapy. Nat. Rev. Cancer *12*, 252–264.

Pradeu, T. (2012). The Limits of the Self. Immunology and Biological Identity (Oxford University Press).

Pradeu, T., Jaeger, S., and Vivier, E. (2013). The speed of change: towards a discontinuity theory of immunity? Nat. Rev. Immunol. *13*, 764–769.

Rosenberg, S.A., and Lotze, M.T. (1986). Cancer immunotherapy using interleukin-2 and interleukin-2-activated lymphocytes. Annu. Rev. Immunol. *4*, 681–709.

Shoval, O., Goentoro, L., Hart, Y., Mayo, A., Sontag, E., and Alon, U. (2010). Fold-change detection and scalar symmetry of sensory input fields. Proc. Natl. Acad. Sci. USA *107*, 15995–16000.

Shoval, O., Alon, U., and Sontag, E.D. (2011). Symmetry Invariance for Adapting Biological Systems. SIAM J. Appl. Dyn. Syst. *10*, 857–886.

Sontag, E.D. (1998). Texts in Applied Mathematics, Volume 6, Mathematical Control Theory. Deterministic Finite-Dimensional Systems. Second (New York: Springer-Verlag).

Sontag, E.D. (2003). Adaptation and Regulation with Signal Detection Implies Internal Model. Syst. Control Lett. *50*, 119–126.

Stepanova, N.V. (1979). [Immune response dynamics during malignant tumor development]. Biofizika *24*, 897–902.

Talmage, D.W. (1957). Allergy and immunology. Annu. Rev. Med. *8*, 239–256.

Vignali, D.A., Collison, L.W., and Workman, C.J.L.W. (2008). How regulatory T cells work. Nat. Rev. Immunol. *8*, 523–532.

Vodovotz, Y., Xia, A., Read, E.L., Bassaganya-Riera, J., Hafler, D.A., Sontag, E., Wang, J., Tsang, J.S., Day, J.D., Kleinstein, S.H., et al. (2016). Solving Immunology? Trends Immunol., S1471-4906(16)30202-2.

Wadhams, G.H., and Armitage, J.P. (2004). Making sense of it all: bacterial chemotaxis. Nat. Rev. Mol. Cell Biol. *5*, 1024–1037.

Wang, Q., Klinke, D.J., 2nd, and Wang, Z. (2015). CD8(+) T cell response to adenovirus vaccination and subsequent suppression of tumor growth: modeling, simulation and analysis. BMC Syst. Biol. *9*, 27.

West, M.A., and Heagy, W. (2002). Endotoxin Tolerance: A Review. Crit. Care Med. *30*, S64–S73.

Wodarz, D., and Komarova, N. (2014). Dynamics of Cancer: Mathematical Foundations of Oncology (World Scientific Publishing).

Wolchok, J. 2010. Endogeneous and Exogenous Vaccination in the Context of Immunologic Checkpoint Blockade.

STAR★METHODS

KEY RESOURCES TABLE

REAGENT or RESOURCE	SOURCE	IDENTIFIER
Software and Algorithms		
Ordinary Differential Equation Simulations MATLAB,	Mathworks	
Other		
Data for Figure 4	Gatenby et al. (1981) and Kölsch and Mengersen (1976)	

METHOD DETAILS

A. Comparison to other models

We next briefly discuss our interpretation of some existing models for self/nonself dynamic discrimination in immune systems, and compare them to the IFFL model.

Tunable Activation Threshold (TAT)

This model was suggested by Grossman and Paul (1992), motivated by the realization that "self/nonself discrimination may be much more complex than the simple failure of competent lymphocytes to recognize self-antigens". The authors argued that for a stimulus to cause cell activation, the excitation level must exceed an activation threshold, and when engaged in persistent sub-threshold interactions, cells are protected against chance activation. In the TAT model, an *activation threshold* for an immune cell is dynamically modulated by an environment-dependent recent excitation history. This history is summarized by an *excitation index*, which we will denote as $x(t)$, which computes a sort of weighted average of the cell's past excitation levels. Given temporal excitation events, which we denote by $u(t)$, it is assumed that the cell undergoes perturbations that depend on the difference between $u(t)$ and the memory variable $x(t)$. The key assumption is that such a perturbation, which we write as $y(t)$: = $u(t) - x(t)$, must exceed a fixed critical value, which we denote by θ, in order to cause activation. In other words, it must be the case that $u(t) - x(t) > \theta$, or equivalently, $u(t) > x(t) + \theta$ (this is how we interpret the statement in Grossman and Paul (1992) that "the activation threshold equals the excitation index plus that critical value") for activation to occur. Cells maintained at a high level of excitation $x(t)$ therefore are relatively insensitive to activation, thus being in some sense anergic. The authors deduce from their model that "upon gradual increase in the levels of excitation…a cell is not likely to be activated…it will become progressively anergic" which is intuitively equivalent to our remark about the lack of continued excitation under slow increases in antigen presentation. With our notations, the model suggested in Grossman and Paul (1992) is:

$$\dot{x} = \alpha u(u - x)$$

for some constant α, and the output would be $y = u - x$. (No explicit population-based nor signaling mechanism was given.) Notice that we then can derive a differential equation for y:

$$\dot{y} = u(\dot{u}/u - \alpha y)$$

and this means, roughly, that y should approach $\dot{u}/u$, the logarithmic derivative of the input u, so that a "log sensing" property is satisfied by the output. Moreover, when the input is constant, the output converges to the same value (zero), independently of the actual value of the input, so we have perfect adaptation. Moreover, we expect y to be small (and thus not exceeding the threshold θ) unless u changes fast in the sense that its logarithmic derivative is large. For example, for $u(t)$ increasing linearly, $\dot{u}$ would be a constant, so $\dot{u}/u = 0$ and therefore $y(t) \to 0$ as $t \to \infty$. On the other hand, for an exponentially increasing $u(t)$, $y(t)$ converges to a value proportional to the exponential rate. These properties are analogous to those satisfied by our model.

Discontinuity theory of immunity

This model was suggested by Pradeu et al. (2013) as a "unifying theory of immunity". Their key hypothesis is that effector immune responses are induced by an "antigenic discontinuity" by which they mean a "sudden" modification of molecular motifs with which immune cells interact. The authors present evidence that natural killer (NK) cells and macrophages are activated by transient modifications, but adapt (ceasing to be responsive) to long-lasting modifications in their environment, and then propose to extend this principle to other components of the immune system, such as B cells and T cells. They also argue that although tumors give rise to effector immune responses, "a persistent tumour antigen diminishes the efficacy of the antitumor response". In summary, their criterion of immunogenicity is the phenomenological antigenic discontinuity and not the nature of the antigen, including both "discontinuities" arising from self motifs such as tumors as well as from non-self motifs such as bacterial or viral infections. As examples

of mechanisms for desensitization they mention receptor internalization, degradation or inactivation of signaling proteins. A concrete example of the latter is the dephosphorylation triggered by immunoreceptor tyrosine-based inhibition motif (ITIM)-containing receptors antagonizing kinases triggered by immunoreceptor tyrosine-based activation motif (ITAM)-containing receptors. The authors also mention Treg population dynamics. Using our notations, the model in Pradeu et al. (2013) starts by computing a running average of the absolute value of consecutive differences in inputs presented at discrete times on a sliding window K time units long:

$$\Delta u(t) := \frac{1}{K}\sum_{i=1}^{K} |u(t-i+1) - u(t-i)|$$

and then taking as the output $y = \Theta(\Delta u)$, where Θ is a sigmoidal saturating function. The authors employ

$$\Theta(x) = \frac{\alpha}{1+e^{-\mu(x-\theta)}}$$

but one could equally well (and perhaps easier to justify mechanistically) employ a Hill-type function $\Theta(x) = Vx^n/(K^n + x^n)$. In continuous time, and assuming that the input is differentiable, we could interpret

$$\Delta u(t) \approx \left|\left| \dot{u}|_{[t-K,t]} \right|\right|_1 = \int_{t-K}^{t} |\dot{u}(s)|\, ds,$$

where the right hand term is the total variation of the input on this sliding window. Note the absolute value, which means that, in this model, activation is symmetrically dependent on increases or decreases of the excitation: decreases may help with "missing self" recognition, in which the expression of a "self" marker suddenly decreases, thus triggering a response. As with our model, slow variations in the input will lead to small $y(t)$, with the threshold function resulting in an ultrasensitive, almost binary, response (provided that μ or n are large, in the two suggested functions Θ).

Growth threshold conjecture

This model was suggested by Arias et al. (2015) as "a theoretical framework for understanding T-cell tolerance" based on the hypothesis that "T cells tolerate cells whose proliferation rates remain below a permitted threshold". As in the other works, the authors postulate that T cells tolerate cognate antigens (irrespectively of their pathogenicity) as long as their rate of production is low enough, while those antigens that are associated with pathogenic toxins or structural proteins of either infectious agents or aggressive tumor cells are highly proliferative, and therefore will be targeted as foes by T cells. In summary, once again the postulate is that a strong immune response will be mounted against fast-growing populations while slow-growing ones will be tolerated. The model in Arias et al. (2015) is not one of change detection as such, but it is a closed-loop system that includes both detection and a killing effect on pathogens. To compare with our previous models, let us again denote the pathogen population size (or a density in a particular environment) by $u(t)$ and the effector cell population by $y(t)$. The authors give for y a second order equation $\ddot{y} = -\delta y + \alpha u$, modeled on a spring-mass system that balances a "restoring to equilibrium force" to its activation by pathogens. We prefer to write the system as a set of first order ODE's. Thus, we let $x := \dot{y}$, and write:

$$\begin{aligned} \dot{u} &= (\lambda - \kappa y)u \\ \dot{x} &= \alpha u - \delta y \\ \dot{y} &= x\,. \end{aligned}$$

The u equation has an exponential growth term balanced by a kill rate that depends on the effector population. The effector population integrates the amount of x (which we might interpret as an intermediate type of cell); the growth of x is driven by pathogens, with a negative feedback from y (in essence an integral feedback on x), but there is no obvious biological mechanism for this model. Observe that when there is no pathogen, this results in a harmonic oscillator for x and y, with sustained oscillations and even negative values. In any event, the authors computationally obtain a bifurcation-like diagram in the (λ, κ) plane, dividing this plane into two regions, labeled "tolerance" (of infection, hence, failure of the immune system) and "intolerance". These regions show how to trade off the growth rate λ of the pathogen versus the parameter κ, which represents a combination of affinity and clearance rate, and various conclusions regarding evasion strategies and the role of fever and even Treg cells are qualitatively derived from there.

B. Theory for system without autocatalysis

We collect here mathematical results for the system when h is linear and f only contains linear degradation or inactivation terms. This system is easier to study theoretically than the system with feedback, and provides much intuition about the general case, besides it being a local approximation in suitable regimes. The equations are then as follows (we write u as the last variable now, because we will separately study the first two equations):

$$\begin{aligned} \dot{x} &= -\delta_x x + \beta u \\ \dot{y} &= \mu\frac{u}{x} - \delta_y y \\ \dot{u} &= (\lambda - \kappa y)u \end{aligned} \qquad \text{(Equations 4A–4C)}$$

The constants $\delta_x, \beta, \mu, \delta_y, \kappa$ are positive, but λ is allowed to be negative, for completeness, although the interesting case is $\lambda \geq 0$. The scalar functions of time $x = x(t)$, $y = y(t)$, and $u = u(t)$ take positive values. It is easy to verify that, for any positive initial conditions, solutions remain positive for all times.

We will separately study the first two equations (Equations 4A and 4B), viewing $u = u(t)$ as an external input to the IFFL described by (Equations 4A and 4B), and viewing $y = y(t)$ as an output or response of the system. Later, we "close the loop".

Remark. In the system (Equations 4A–4C), and in particular in the system (Equations 4A and 4B), one may assume without loss of generality that $\delta_x = \beta = \mu = 1$. This is because we may eliminate these parameters by rescaling variables. Indeed, substituting

$$x = \frac{\beta}{\delta_x} x^* \ , \ \ y = \frac{\mu}{\beta} y^* \ , \ \ t = \frac{1}{\delta_x} t^* \ , \ \ \delta_y^* = \frac{\delta_y}{\delta_x} \ , \ \ \lambda^* = \frac{\lambda}{\delta_x} \ , \ \ \kappa^* = \frac{\mu\kappa}{\delta_x \beta} \ ,$$

into system (Equations 4A–4C), one obtains:

$$\begin{aligned} \frac{dx^*}{dt^*} &= -x^* + u \\ \frac{dy^*}{dt^*} &= \frac{u}{x^*} - \delta_y^* y^* \\ \frac{du}{dt^*} &= (\lambda - \kappa y^*) u \end{aligned} \qquad \text{(Equations 5A–5C)}$$

IFFL's responses to various classes of inputs

Let us consider the system (Equations 4A and 4B), a differentiable function $u = u(t)$ viewed as an external input or forcing function, and any (positive) solution $(x(t), y(t))$ corresponding to this input. We are interested first in understanding how the growth rate of the input affects the asymptotic values of the output variable y.

We denote the derivative of $\ln u(t)$ with respect to t as follows:

$$v(t) \ := \ \frac{\dot{u}(t)}{u(t)}$$

and its limsup and liminf as $t \to \infty$

$$\underline{\mu} = \liminf_{t \to \infty} v(t) \ , \overline{\mu} = \limsup_{t \to \infty} v(t).$$

We assume that v is bounded, and thus both of these numbers are finite. We also introduce the following function:

$$p(t) \ := \ \frac{u(t)}{x(t)}.$$

Since

$$\dot{p} = \dot{u}/x - u\dot{x}/x^2 = (u/x)[\dot{u}/u - \dot{x}/x] = (u/x)[\dot{u}/u - (-\delta_x x + \beta u)/x] = (u/x)[\dot{u}/u + \delta_x - \beta u/x],$$

we have that p satisfies the following ODE with input v:

$$\dot{p} = p(\delta_x + v - \beta p).$$

Lemma. Let u be a differentiable input to system (Equations 4A and 4B) with $\delta_x = \beta = \mu = 1$. With the above notations,

$$\max\left\{0, 1 + \underline{\mu}\right\} \ \leq \ \liminf_{t \to \infty} p(t) \ \leq \ \limsup_{t \to \infty} p(t) \ \leq \ \max\{0, 1 + \overline{\mu}\}$$

Proof. Since $\delta_x = \beta = \mu = 1$,

$$\dot{p} = p(1 + v - p).$$

To prove the upper bound, we consider two cases, $1 + \overline{\mu} < 0$ and $1 + \overline{\mu} \geq 0$. In the first case, let $\varepsilon := -(1 + \overline{\mu}) > 0$; the definition of $\overline{\mu}$ gives that, for some $T \geq 0$, $1 + v(t) < -\varepsilon/2$ for all $t \geq T$. It follows that $\dot{p} \leq p(-\varepsilon/2 - p)$ for all $t \geq T$. Thus, $\dot{p} < 0$ whenever $p > 0$, from which it follows that $\limsup_{t \to \infty} p(t) = \lim_{t \to \infty} p(t) = 0$. Suppose now that $1 + \overline{\mu} \geq 0$. Pick any $\varepsilon > 0$ and a $T = T(\varepsilon) \geq 0$ such that $v(t) \leq \overline{\mu} + \varepsilon$ for all $t \geq T$. For such t, $\dot{p} = p(1 + v - p) \leq p(1 + \overline{\mu} + \varepsilon - p)$. This implies that $\dot{p} < 0$ whenever $p(t) > 1 + \overline{\mu} + \varepsilon$, which implies that $\limsup_{t \to \infty} p(t) \leq 1 + \overline{\mu} + \varepsilon$. Letting $\varepsilon \to 0$, we conclude that $\limsup_{t \to \infty} p(t) \leq 1 + \overline{\mu}$. We next prove the lower bound. Pick any $\varepsilon > 0$ and a $T = T(\varepsilon) \geq 0$ such that $v(t) \geq \underline{\mu} - \varepsilon$ for all $t \geq T$. Thus $\dot{p} = p(1 + v - p) \geq p(1 + \underline{\mu} - \varepsilon - p)$ for all $t \geq T$. This implies that $\dot{p} > 0$ whenever $p(t) < 1 + \underline{\mu} - \varepsilon$ (recall that $p(t) > 0$ for all t, since by assumption $u(t) > 0$ and $x(t) > 0$ for all t). Therefore $\liminf_{t \to \infty} p(t) \geq 1 + \underline{\mu} - \varepsilon$, and letting $\varepsilon \to 0$ we have $\liminf_{t \to \infty} p(t) \geq 1 + \underline{\mu}$. Since $p(t) \geq 0$ for all t, we also have $\liminf_{t \to \infty} p(t) \geq \max\{0, 1 + \underline{\mu}\}$. This completes the proof. In particular, if $v(t) \to \nu$ as $t \to \infty$ then $\underline{\mu} = \overline{\mu} = \nu$, so we have as follows.

Corollary. If $v(t)\to\nu$ as $t\to\infty$ then $\lim_{t\to\infty}p(t)=\max\{0,1+\nu\}$.

For the original system (Equations 4A and 4B), we have as follows.

Proposition. Consider a solution of (Equations 4A and 4B), with a differentiable $u(t)>0$ as input and $x(t)>0$, $y(t)>0$. Assuming that $v=\dot{u}/u$ is bounded, we have:

$$\frac{\mu}{\beta\delta_y}\max\left\{0,\delta_x+\underline{\mu}\right\}\ \le\ \liminf_{t\to\infty}y(t)\ \le\ \limsup_{t\to\infty}y(t)\ \le\ \frac{\mu}{\beta\delta_y}\max\{0,\delta_x+\overline{\mu}\}$$

Proof. We first assume that $\delta_x=\beta=\mu=1$. Let $\underline{p}:=\liminf_{t\to\infty}p(t)$ and $\overline{p}:=\limsup_{t\to\infty}p(t)$. Equations 4B can be written as $\dot{y}=p-\delta_y y$. This is a linear system forced by the input $p=p(t)$. Pick any $\varepsilon>0$. Then there is some $T=T(\varepsilon)$ such that $\underline{p}-\varepsilon<p(t)<\overline{p}+\varepsilon$ for all $t\geq T$. For such t, $\dot{y}(t)>0$ whenever $y(t)<(1/\delta_y)(\underline{p}-\varepsilon)$ and $\dot{y}(t)<0$ whenever $y(t)>(1/\delta_y)(\overline{p}+\varepsilon)$. It follows that $(1/\delta_y)(\underline{p}-\varepsilon)\leq y(t)\leq(1/\delta_y)(\overline{p}+\varepsilon)$ for all $t\geq T$. Letting $\varepsilon\to 0$ we conclude that

$$\underline{p}/\delta_y\ \le\ \liminf_{t\to\infty}y(t)\ \le\ \limsup_{t\to\infty}y(t)\ \le\ \overline{p}/\delta_y$$

and the inequalities follow when $\delta_x=\beta=\mu=1$. To deal with general parameters, we recall that (Equations 5A and 5B) are obtained with $x=(\beta/\delta_x)x^*$, $y=(\mu/\beta)y^*$, $t=(1/\delta_x)t^*$, and $\delta_y^*=(\delta_y/\delta_x)$. Note that $t^*\to\infty$ if and only if $t\to\infty$. Thus, the inequalities in the last display hold for $p^*=u/x^*=(\beta/\delta_x)p$, y^*, and δ_y^* in place of p, y and δ_y. Similarly, the inequalities in the Lemma hold for $p^*=u/x^*$ and

$$\underline{\mu}^*=\liminf_{t\to\infty}v^*(t^*),\overline{\mu}=\limsup_{t\to\infty}v^*(t^*)$$

where $v^*=((du/dt^*)/u)=(1/\delta_x)v$, so $\underline{\mu}^*=(1/\delta_x)\underline{\mu}$ and $\overline{\mu}^*=(1/\delta_x)\overline{\mu}$. Therefore,

$$\begin{aligned}\liminf_{t\to\infty}\ y(t)\ &=\ \liminf_{t^*\to\infty}\frac{\mu}{\beta}y^*(t^*)\ \geq\ \frac{\mu}{\beta}\frac{\underline{p}^*}{\delta_y^*}\ =\ \frac{\mu}{\beta}\frac{\underline{p}^*}{\delta_y/\delta_x}\ =\ \frac{\delta_x\mu}{\beta\delta_y}\,\underline{p}^*\ =\ \frac{\delta_x\mu}{\beta\delta_y}\max\left\{0,1+\underline{\mu}^*\right\}\\ &=\ \frac{\mu}{\beta\delta_y}\max\left\{\delta_x+\underline{\mu}\right\}.\end{aligned}$$

A similar remark applies to limsup, and the result follows.

Corollary. If $v(t)\to\nu$ as $t\to\infty$ then $\lim_{t\to\infty}y(t)=(\mu/\beta\delta_y)\max\{0,\delta_x+\nu\}$. Three particular cases are:

- When $u(t)$ has sub-exponential growth, meaning that $d\ln u/dt\leq 0$, then $\limsup_{t\to\infty}y(t)\leq\delta_x\mu/\beta\delta_y$.
- In particular, if $u(t)=K_0+C_0t$ is linear, then $\nu=0$ and thus $\lim_{t\to\infty}y(t)=\delta_x\mu/\beta\delta_y$.
- If $u(t)=C_0e^{\nu t}$ is exponential, then $\lim_{t\to\infty}y(t)=(\mu/\beta\delta_y)\max\{0,\delta_x+\nu\}$.

In conclusion, when u is constant, or even with linear growth, the value of the output $y(t)$ converges to a constant, which does not depend on the actual constant value, or even the growth rate, of the input. For constant inputs, this is called the ''perfect adaptation'' property. If, instead, u grows exponentially, then $y(t)$ converges to a steady state value that is a linear function of the logarithmic growth rate.

IFFL's as feedback controllers

As we remarked, in the case of exponential inputs $u(t)=e^{\nu t}$, $\lim_{t\to\infty}y(t)=\overline{y}=(c/\beta\delta_y)\max\{0,\delta_x+\nu\}$. Now suppose that, in turn, $u(t)$ satisfies Equation 4C, which means that $v(t)=\lambda-\kappa y(t)$, and therefore $\nu=\lim_{t\to\infty}v(t)=\lambda-\kappa\overline{y}$. This gives an implicit equation for the rate v:

$$\nu=\lambda-\kappa\overline{y}=\lambda-\frac{\mu\kappa}{\beta\delta_y}\max\{0,\delta_x+\nu\}.$$

We now solve this equation.

Denote

$$F(\lambda)=\frac{\lambda\beta\delta_y-\mu\kappa\delta_x}{\beta\delta_y+\mu\kappa}.$$

Suppose first that $\lambda\leq\delta_x$. Then, since $\delta_x+F(\lambda)=(\delta_x+\lambda)\theta$ (where $\theta=\beta\delta_y/(\beta\delta_y+\mu\kappa)$), $\mu=F(\lambda)$ satisfies $\delta_x+\mu\geq 0$ and also, rewriting $\mu=F(\lambda)$, μ is the unique solution of the implicit equation with $\delta_x+\mu\geq 0$. There are no solutions with $\delta_x+\mu<0$, because such a solution would have to satisfy $\mu=\lambda$, but $\delta_x+\lambda\geq 0$. Suppose instead that $\lambda>\delta_x$. Then $\mu=\lambda$ is the unique solution of the implicit equation with $\delta_x+\mu<0$. There are no solutions with $\delta_x+\mu\geq 0$, because such a solution would have to satisfy $\mu=F(\lambda)$ and therefore have $\delta_x+\mu=\delta_x+F(\lambda)=(\delta_x+\lambda)\theta<0$, a contradiction. In summary, when $\lambda\geq-\delta_x$, the unique solution of the implicit equation is $\mu=F(\lambda)$, and when $\lambda<-\delta_x$ it is $\mu=\lambda$.

Note that when

$$\mu\delta_x\kappa\ >\ \beta\delta_y\lambda$$

(which happens automatically when $\lambda < 0$) the formula $\nu = F(\lambda)$ gives that $\nu < 0$, that is, $u(t) \to 0$ as $t \to +\infty$. Conversely, if $\mu\delta_x\kappa < \beta\delta_y\lambda$, then $\nu > 0$ and so $u(t) \to \infty$ as $t \to +\infty$. Qualitatively, this makes sense: a large feedback gain κ, or a small growth rate λ in the absence of feedback, leads to the asymptotic vanishing of the u variable.

In addition, from the formula $\overline{y} = (\mu/\beta\delta_y)\max\{0, \delta_x + \nu\}$ we conclude the following piecewise linear formula for the dependence of the limit of the output on the parameter λ that gives the growth rate of exponentially growing u when there is no feedback:

$$\overline{y} = \begin{cases} 0 & \text{if} \quad \lambda < -\delta_x \\ \dfrac{\mu(\delta_x + \lambda)}{\beta\delta_y + \mu\kappa} & \text{if} \quad \lambda \geq -\delta_x \end{cases}.$$

These considerations provide helpful intuition about the closed-loop system, but they do not prove that the above inequality is necessary and sufficient for stability, nor do they show the validity of this growth rate for the closed-loop system. The reason that the argument is incomplete is that there is no *a priori* reason for $u(t)$ to have the exponential form $u(t) = C_0 e^{\nu t}$. We next provide a rigorous argument.

Analysis of the closed-loop system

Theorem. Suppose that $(x(t), y(t), u(t))$ is a (positive) solution of Equations 4A–4C, and define

$$v(t) \;:=\; \dot{u}(t)/u(t) = \lambda - \kappa y(t),$$

$$p(t) \;:=\; u(t)/x(t),$$

$\overline{y}$ by the following formula here:

$$\overline{y} \;=\; \begin{cases} 0 & \text{if} \quad \delta_x + \lambda < 0 \\ \dfrac{\mu(\delta_x + \lambda)}{\beta\delta_y + \mu\kappa} & \text{if} \quad \delta_x + \lambda \geq 0 \end{cases}$$

$$\overline{p} = (\delta_y/\mu)\overline{y}, \text{ and}$$

$$\overline{v} \;=\; \begin{cases} \lambda & \text{if} \quad \delta_x + \lambda < 0 \\ \lambda - \kappa\dfrac{\mu(\delta_x + \lambda)}{\beta\delta_y + \mu\kappa} & \text{if} \quad \delta_x + \lambda \geq 0. \end{cases}$$

Then:

$$\begin{array}{rcl} \lim\limits_{t\to\infty} y(t) & = & \overline{y} \\ \lim\limits_{t\to\infty} p(t) & = & \overline{p} \\ \lim\limits_{t\to\infty} v(t) & = & \overline{v}. \end{array}$$

and

$$\lim_{t\to\infty} u(t) = \begin{cases} 0 & \text{if} \quad \delta_x\mu\kappa > \beta\delta_y\lambda \\ \infty & \delta_x\mu\kappa < \beta\delta_y\lambda \end{cases}$$

Proof. Substituting $v(t) = \lambda - \kappa y(t)$ into, we have the surprising and very useful fact that there is a closed system of just two differential equations for p and y:

$$\begin{array}{rcl} \dot{p} & = & p(\delta_x + \lambda - \kappa y - \beta p) \\ \dot{y} & = & \mu p - \delta_y y. \end{array} \qquad \text{(Equations 6A and 6B)}$$

(This system could be viewed as a non-standard predator-prey of system, where y behaves as a predator and p as a prey.) In all of the real plane, there are two equilibria of this system, one at $p = y = 0$ and the other at $p = \delta_y(\delta_x + \lambda)/(\beta\delta_y + c\kappa)$, $y = c(\delta_x + \lambda)/(\beta\delta_y + c\kappa)$. The second equilibrium point is in the interior of first quadrant if and only if $\delta_x + \lambda > 0$.

We start by evaluating the Jacobian matrix of the linearized system. This is:

$$J = \begin{pmatrix} \delta_x + \lambda - \kappa y - 2\beta p & -p\kappa \\ \mu & -\delta_y \end{pmatrix}$$

which, when evaluated at $p = y = 0$, has determinant $-\delta_y(\delta_x + \lambda)$ and trace $\delta_x + \lambda - \delta_y$, and when evaluated at $(\overline{p}, \overline{y})$ has trace

$$\frac{-\beta\delta_y(\delta_x + \lambda)}{c\kappa + \beta\delta_y} - \delta_y$$

and determinant $\delta_y(\delta_x+\lambda)$. Thus, when $\delta_x+\lambda>0$, the trace is negative and the determinant is positive, so the equilibrium $(\overline{p},\overline{y})$ is stable, and $(0,0)$ is a saddle because the determinant of the Jacobian is negative at that point. When instead $\delta_x+\lambda\leq 0$, the only equilibrium with non-negative coordinates is $(0,0)$, and the determinant of the Jacobian is positive there, while the trace is negative, so this equilibrium is stable.

We note that, in general, if have shown that there is a limit $v(t)\to\overline{v}$ as $t\to\infty$ then $u(t)\to 0$ as $t\to\infty$ if $\overline{v}<0$ and $u(t)\to\infty$ as $t\to\infty$ if $\overline{v}>0$. Indeed, in the first case there is some $T\geq 0$ so that for $t\geq T$, $v=\dot{u}/u<\overline{v}/2$, meaning that $d(e^{-\overline{v}t/2}u(t))/dt\leq 0$, and hence $e^{-\overline{v}t/2}u(t)\leq e^{-\overline{v}T/2}u(T)$, so $u(t)\leq e^{\overline{v}(t-T)/2}u(T)\to 0$ (since $\overline{v}<0$). Similarly, in the second case we use that there is some $T\geq 0$ so that for $t\geq T$, $v=\dot{u}/u>\overline{v}/2$, meaning that $d(e^{-\overline{v}t/2}u(t))/dt\geq 0$, and hence $e^{-\overline{v}t/2}u(t)\geq e^{-\overline{v}T/2}u(T)$, so $u(t)\geq e^{\overline{v}(t-T)/2}u(T)\to\infty$ (since $\overline{v}>0$).

Consider first the case $\delta_x+\lambda\leq 0$. Then $\dot{p}=p(\delta_x+\lambda-\kappa y-\beta p)\leq p(-\kappa y-\beta p)<0$ for all $p>0$, and therefore $p(t)\to\overline{p}=0$ as $t\to\infty$. We may now view the linear system $\dot{y}=\mu p-\delta_y y$ as a one-dimensional system with input $p(t)\to 0$, which implies that also $y(t)\to\overline{y}=0$. In turn, this implies that $v=\lambda-\kappa y\to\overline{v}=\lambda<0$. By the general fact proved earlier about limits for $u(t)$, we know that $u(t)\to 0$ as $t\to\infty$. This completes the proof when $\delta_x+\lambda\leq 0$.

So we assume from now on that $\delta_x+\lambda>0$. We will show that, in this case, all solutions with $p(t)>0$ and $y(t)>0$ globally converge to the unique equilibrium $(\overline{p},\overline{y})$. Once that this is proved, it will follow that $v(t)\to\overline{v}=\lambda-\kappa\overline{y}$. Now, this value of $\overline{v}$, for $\overline{y}$ picked as in (case $\delta_x+\lambda\geq 0$), coincides with $\nu=F(\lambda)=(\lambda\beta\delta_y-\mu\kappa\delta_x)/(\beta\delta_y+\mu\kappa)$. So $\overline{v}<0$ if $\mu\kappa\delta_x>\lambda\beta\delta_y$ and $\overline{v}>0$ if $\lambda\beta\delta_y>\mu\kappa\delta_x$, and this provides the limit statement for $u(t)$, completing the proof.

We next show global convergence. A sketch of nullclines (see Figure S1 for a numerical example) makes convergence clear, and helps guide the proof. Consider any $P\geq(\delta_x+\lambda)/\beta$ and any $Y\geq\mu P/\delta_y$ and the rectangle $[0,P]\times[0,Y]$.

On the sides of this rectangle, the following properties hold:

- On the set $\{0\}\times(0,Y)$, $\dot{p}\geq 0$, because $\dot{p}=0$.
- On the set $\{P\}\times(0,Y)$, $\dot{p}\leq 0$, because $\dot{p}=p(\delta_x+\lambda-\beta P)\leq 0$, by the choice of P.
- On the set $(0,P)\times\{0\}$, $\dot{y}\geq 0$, because $\dot{y}=\mu p>0$.
- On the set $(0,P)\times\{Y\}$, $\dot{y}\leq 0$. because $\dot{y}=\mu p-\delta_y Y\leq\mu P-\delta_y Y\leq 0$ by the choice of Y.
- At the corner point $(0,0)$, $\dot{p}\geq 0$, $\dot{y}\geq 0$, because $\dot{p}=\dot{y}=0$.
- At the corner point $(0,Y)$, $\dot{p}\geq 0$, $\dot{y}\leq 0$, because $\dot{p}=0$, $\dot{y}=-\delta_y Y<0$.
- At the corner point $(P,0)$, $\dot{p}\leq 0$, $\dot{y}\geq 0$, because $\dot{p}=p(\delta_x+\lambda-\beta P)\leq 0$, $\dot{y}=\mu P>0$.
- At the corner point (P,Y), $\dot{p}\leq 0$, $\dot{y}\leq 0$, because $\dot{p}=p(\delta_x+\lambda-\beta P-\kappa Y)<p(\delta_x+\lambda-\beta P)\leq 0$, $\dot{y}=\mu P-\kappa y\geq 0$.

These properties imply that the vector field points inside the set at every boundary point and therefore it is forward-invariant, meaning that every trajectory that starts in this set remains there for all positive times (Clarke et al., 1998). The rest of the proof of stability uses the Poincaré-Bendixson Theorem together with the Dulac-Bendixson criterion. Note that, for any initial condition $\xi=(p(0),y(0))$ one can always pick a large enough value of P and Y so that $(p(0),y(0))\in[0,P]\times[0,Y]$. The invariance property guarantees that the omega limit set $\omega^+(\xi)$ is a nonempty compact connected set, and the Poincaré-Bendixson Theorem insures that such a set is one of the following: (a) the equilibrium $(0,0)$, (b) a periodic orbit in the interior of the square, or (c) the equilibrium $(\overline{p},\overline{y})$ (Hirsch and Smale, 1974). Note that a homoclinic orbit around $(0,0)$ cannot exist, because the unstable manifold of this equilibrium is the entire y axis. For the same reason, if ξ has positive coordinates, $\omega^+(\xi)\neq(0,0)$. Therefore, all that we need to do is rule out periodic orbits. Consider the function $\varphi(p,y)=1/p$. The divergence of the vector field

$$\begin{pmatrix}\frac{1}{p}(p(\delta_x+\lambda-\kappa y-\beta p))\\ \frac{1}{p}(\mu p-\delta_y y)\end{pmatrix}=\begin{pmatrix}\delta_x+\lambda-\kappa y-\beta p\\ \mu-\delta_y y/p\end{pmatrix}$$

is

$$\frac{\partial\delta_x+\lambda-\kappa y-\beta p}{\partial p}+\frac{\partial\mu-\delta_y y}{\partial y}=-\beta-\delta_y/p\,,$$

which has a constant sign (negative). The Dulac-Bendixson criterion (Hirsch and Smale, 1974) then guarantees that no periodic orbits can exist, and the proof is complete.

C. Perfect adaptation and scale-invariance

A system is said to be *perfectly adapting* provided that its response returns asymptotically to a pre-stimulus value under constant stimulation. This property is typically exhibited by sensory systems processing light, chemical, and other signals, and it has been extensively investigated both experimentally and mathematically (Alon, 2006; Keener and Sneyd, 2009). In particular, when subjecting a perfectly adapting system to a step-wise input signal, as shown in Figure S2A, the output of the system settles, after a transient response, to a basal value which does not depend on the magnitude of the stimulus. The response amplitude and timing, on the other

hand, typically depends on the input magnitude. This notion can be refined as follows. Suppose that every step has the same relative or "fold" change, $u_{i+1}/u_i =$ constant, as shown in the figure. For *scale-invariant* systems, the responses to such steps have the exact same shape, amplitude, and duration.

The alternative term "fold change detection" is sometimes used for this property, to emphasize the fact that such systems can only react differently if the fold changes are not the same. To put it in another way, such systems can give different responses if the difference $\log u_{i+1} - \log u_i$ is nonzero (log sensing) as opposed to $u_{i+1} - u_i$. The precise mathematical definition of scale-invariance involves arbitrary input signals: responses to arbitrary scaled inputs as in Figure S2B, and not only piecewise constant ones, should be the same, provided that the internal state starts from a preadapted value. We refer the reader to Shoval et al. (2011) for technical details.

Scale invariance or fold change detection (FCD) is a strengthening of the Weber-Fechner "log sensing" property, which is sometimes defined as the requirement that the maximum amplitude of responses to two scaled inputs should be the same, but not necessarily their exact shape or even timing. Recent interest in the FCD property was largely triggered by the papers Goentoro and Kirschner (2009) and Cohen-Saidon et al. (2009), in which fold-change detection behavior was experimentally observed in a Wnt signaling pathway and an EGF pathway, respectively; these are highly conserved eukaryotic signaling pathways that play roles in embryonic patterning, stem cell homeostasis, cell division, and other central processes. Later, the paper Shoval et al. (2010) predicted scale invariant behavior in *E. coli* chemotaxis, a prediction which was subsequently experimentally verified (Lazova et al. 2011). Similar results are available for other bacterial species, for example *R. sphaeroides*, for which theoretical predictions made in Hamadeh et al. (2013) were experimentally confirmed in Wadhams and Armitage (2004). A mathematical study of scale invariance, together with a necessary and sufficient characterization in terms of solutions of a partial differential equation, can be found in (Shoval et al. (2011). It has been recently shown that all scale invariant systems compute a certain type of differentiation operator, such as logarithmic derivatives (Lang and Sontag, 2016).

One example of a scale invariant system is the IFFL that underlies our model, which we repeat here for ease of reference:

$$\begin{aligned} \dot{x} &= -\delta_x x + \beta u \\ \dot{y} &= \mu \frac{u}{x} - \delta_y y \end{aligned}$$

where $\beta, \mu, \delta_x, \delta_y$, are some positive constants and $u(t)$ is viewed as an external stimulus. For any given input function $u(t)$ and initial values $x(0)$ and $y(0)$, the solution of this system can be found by first solving the scalar linear ordinary differential equation for $x(t)$, and then plugging this result together with $u(t)$ into the y equation, which is also a linear ODE. For a constant input $u(t) \equiv u_0 > 0$, there is a globally asymptotically stable steady state, given by

$$x = \frac{\beta u_0}{\delta_x}, \quad y = \frac{\delta_x \mu}{\beta \delta_y}.$$

At steady state, the output y is independent of the particular value of the constant input u_0, meaning that the system is perfectly adapting. Suppose next that $(x(t), y(t))$ is any solution of the system corresponding to an input $u(t)$, now not necessarily a step function. It is then immediate to verify that $(px(t), y(t))$ is a solution corresponding to the input $pu(t)$, $t \geq 0$, for any nonzero constant scaling factor p:

$$\begin{aligned} \dot{x} &= -\delta_x x + \beta u \\ \dot{y} &= \mu \frac{u}{x} - \delta_y y \end{aligned}$$

Implies

$$\begin{aligned} (\dot{px}) &= -\delta_x (px) + \beta (pu) \\ \dot{y} &= \mu \frac{pu}{px} - \delta_y y = \mu \frac{u}{x} - \delta_y y. \end{aligned}$$

Thus, this system responds with the same output signal $y(t)$ to two inputs which differ only in scale, provided that the initial state $x(t)$ had already adapted to the input at time $t < 0$. In other words, given a step input that jumps from $u(t) = u_0$ for $t < 0$ at time $t = 0$ and an initial state at time $t = 0$ that has been pre-adapted to the input $u(t)$ for $t < 0$, $x(0) = \beta u_0/\delta_x$, the solution is the same as if, instead, the input would have been $pu(t)$ for $t > 0$, but starting from the respective pre-adapted state $p\beta u_0/\delta_x$. This means that our IFFL subsystem is scale-invariant.

It would be very interesting to test experimentally the response to scaled versions of antigen presentation, to verify if such scale invariance holds, even in an approximate fashion.

D. Details on the model used for simulations

In this section, we explain the terms in the differential equations used in simulations, including the parameters used. Of course, our model is only a cartoon of a hugely complicated system of interlocking processes. Moreover, even if the model were mechanistic, which it is not, numbers would depend on the specific tumor or infection tissue being modeled. Thus, these algebraic forms and numbers are offered only as a plausible scenario.

As explained in the main text, u represents an immune challenge, specifically a tumor in this case, while x and y might represent populations of activated and specific T suppressor ($CD4^+$ $CD25^+$ Treg) and cytotoxic T cells ($CD8^+$ cells) respectively. We use as a guide in our modeling the paper by Kirschner and Panetta (1998), which has become a classic reference for tumor-immune interactions in the presence of cytokines (no regulatory T cells in that model), together with the more recent paper by Khailaie et al. (2013) which described a model of immune activation in the presence of both chemokines and also regulatory T cells (no tumor dynamics in that model).

Treg cells play a central role in cytotoxic T cell regulation. The various Treg mechanisms can be arranged into four groups centered around four basic modes of action (Vignali et al., 2008): (1) inhibitory cytokines, including IL-10, IL-35) and TGF-β, (2) cytolysis through granzyme-A- and granzyme-B-dependent and perforin-dependent killing mechanisms, (3) metabolic disruption through CD25-dependent cytokine-deprivation-mediated apoptosis, cAMP-mediated inhibition, and adenosine–purinergic adenosine receptor (A2A)-mediated immunosuppression, and (4) targeting dendritic cells through mechanisms that modulate DC maturation and/or function.

Cell number units

Since we use parameters from both Kirschner and Panetta (1998) and Khailaie et al. (2013), it is thus important to clarify the units used in these sources.

Kirschner and Panetta's paper gives "volume" as the unit for cell populations. Since many of these parameters were in turn obtained from the foundational paper by Kuznetsov et al. (1994), which provided one of the first differential equation models for interactions between tumors and the immune system, one can compare the two papers, to map their unit to cell numbers. For this purpose, we can compare the value of the carrying capacity of tumors ("B" in the simulations that we provided) in both papers. In Kirschner and Panetta (1998) $B = 10^{-9}$, and in Kuznetsov et al. (1994) $B = 2\times10^{-9}$. Ignoring the factor of 2, this means that "volume" = number of cells. This is confirmed by comparing the Michaelis-Menten constant for IL-2 activation g_1 (g in the second paper): 2×10^{-7} volume units and 2.019×10^{-7} T cells respectively. Therefore, we will be interpreting cell units in Kirschner and Panetta (1998) as numbers of cells. In our simulations, we use $B = 10^{-3}$, because we prefer to switch to units of 10^6 cells. In Khailaie et al. (2013), "cell" means nondimensional units, cells/ C_0, where C_0 is an unspecified reference quantity of cells. Now, Figure 5 in Khailaie et al. (2013) shows stable branches of equilibria under antigen stimulation in ranges of 2 to 30 nondimensionalized T cells, while in their companion experimental paper (Milanez-Almeida et al., 2015), the same authors provide estimates of T cells in various tissues in mice in the range 10^6 to 8×10^6. Thus approximately $C_0 = 10^6$ cells is consistent with the analysis in Khailaie et al. (2013), and so we will interpret the numbers in that reference in units of 10^6 cells.

The autocatalytic term $Vy^2/(K + y)$

This term is intended to model a cytokine-mediated positive feedback loop on effector T cells. Cytokines are molecules that act as immunomodulating agents and mediate communication among immune systems components and their environment. Their concentrations can increase up to 1,000-fold during inflammatory conditions. Examples of cytokines include interleukins such as IL-2 and IL-6, interferons, and TNF. The role of cytokines in anti-tumor responses, and in particular IL-2, has been the subject of much study (Dranoff, 2004) and of mathematical modeling since at least the work of Kirschner and Panetta (1998), who proposed a simple differential equation model that includes variables for tumor load, effector immune cells, and cytokines. In their model, activated T cells produce cytokines, specifically IL-2, which in turn enhance lymphocyte activation, growth and differentiation, in particular of the cytotoxic T cell (CTL) population. The effect is through a positive feedback that is both autocrine, that is, acting on the cells that produce it, and paracrine, acting on nearby cells. This role of IL-2 in enhancing T-cell proliferation and differentiation is one reason that IL-2 was originally named "T-cell growth factor," although by now many other immunoregulatory functions of IL-2 are known.

The term that represents the effect of the cytokine (IL-2) on y in Kirschner and Panetta (1998) is $p_1yz/(g_1+z)$, where the cytokine z satisfies the differential equation $z' = p_2uy/(g_3+u) - \mu_3 z$. This equation models IL-2 secretion by activated effector T cells, with a Michaelis-Menten kinetics to account for self-limiting production of IL-2, together with a degradation rate. To obtain z as a function of y, we assume that this variable is at equilibrium; on the saturation regime of antigen load u we obtain $z = (p_2/\mu_3)y$. Now substituting this expression into the differential equation for y, we have the autocatalytic term

$$\frac{(p_1p_2/\mu_3)y^2}{g_1+(p_2/\mu_3)y} = \frac{Vy^2}{K+y}$$

where $V = p_1$ and $K = \mu_3 g_1/p_2$. If we start, instead, from Khailaie et al. (2013), the corresponding term in the differential equation for $\dot{y}$ is ayz, where z now satisfies a different equation, $\dot{z} = dy - eyz - fz$ and the term eyz represents IL-2 consumption rate by T cells. Nonetheless, under the same equilibrium assumptions we obtain $z = dy/(f+ey)$, which when substituted into ayz gives

$$\frac{(ad/e)y^2}{(f/e)+y} = \frac{Vy^2}{K+y}$$

where $V = ad/e$ and $K = f/e$. In other words, we derived the same functional form as when starting from Kirschner and Panetta (1998).

Plausible parameters values can be obtained from Kirschner and Panetta (1998) or from Khailaie et al. (2013). The parameters used in Kirschner and Panetta (1998) were $p_1 = 0.1245$, $p_2 = 5$, $g_1 = 2\times10^7$, and μ_3 was arbitrarily picked as 10 from the range 8.31 to 33.27

using a half-life for IL-2 of 30 to 120 minutes given in Rosenberg and Lotze (1986). Plugging these into the formulas given above, we obtain $V = 0.1245$ and $K = 10^6 K_0$, where K_0 ranges from 33 to 133. As discussed earlier, we are reading the units in the paper Kirschner and Panetta (1998) as individual cell counts. When translating to our units of 10^6 cells, we obtain that K in their model ranges between 33 and 133. (The argument is: if we rescale variables letting $\eta = y/10^6$, then the corresponding term in $\dot{\eta}$ is $10^{-6} V(10^6\eta)^2/(10^6 K_0 + 10^6\eta) = V\eta^2/(K_0 + \eta)$, which means that $K = K_0$ when writing the equation in terms of η.) Using parameters from Khailaie et al. (2013) gives similar results. As discussed earlier, we are reading the units in that paper as 10^6 cells. These parameters are picked in Khailaie et al. (2013) as follows: $a = 0.4$, $d = 0.01$, $e = 0.01$, $f = 1$. Plugging these into the formulas given above, these lead to $V = 0.4$ and $K = 100$. In summary, one paper gives K between 33 and 133 and $V = 0.1245$, and the other paper uses $K = 100$ and $V = 0.4$. We therefore take $K = 100$ and for V pick an average, $V = 0.25$ of the two values. Note that the units of K are 10^6 cells, and the units of V are day^{-1}.

The fratricide term $-\varepsilon y^2$

Following the T cell model in Khailaie et al. (2013), we include the term $-\varepsilon y^2$ for cell-contact-dependent activation-induced cell death in activated T cells, a process known as "fratricide". Activated T cells express the receptor FasR, also known as apoptosis antigen 1 (APO-1 or APT), cluster of differentiation 95 (CD95) or tumor necrosis factor receptor superfamily member 6 (TNFRSF6), as well as the ligand for this molecule, FasL; fratricide can result from direct cell contact or from cleavage of FasL ("death ligand"), and the ligation of FasR by soluble FasL results in apoptotic cell death, mediated by caspase activation (Flaherty, 2011). It is believed that the exposure to tumor antigens in T cells might mediate fratricide (Leisegang et al., 2010). Callard et al. (2003) modeled the fratricide mechanism by a nonlinear death term $-\varepsilon y^2$ and speculate that Fas-mediated apoptosis results in a density-dependent death rate for T cell homeostasis that does not require competition for resources nor quorum-sensing mechanisms for density estimation. From Khailaie et al. (2013), we pick $\varepsilon = 10^{-5}$, in units of day^{-1} (10^6 cells)$^{-1}$.

The decay terms $-\delta_x x$ and $\delta_y y$

These represent linear degradation of activated T and Treg cells. The values $\delta_x x = \delta_y y = 0.1$ are from (Khailaie et al. 2013). Units of both are day^{-1}.

The term βu

Stimulation of regulatory cells is a very complex process that involves a wide variety of antigen presenting cells and other mediators. TRegs are exported from the thymus and recirculate through secondary lymphoid tissues as "central" TReg cells, and get activated through T cell receptor (TCR) ligation, CD28 co-stimulation and/or interleukin-2 (IL-2), which induce upregulation of expression of interferon regulatory factor 4 (IRF4), which then orchestrates their differentiation into "effector" TReg cells (Liston and Gray, 2014). We make the simplest possible assumption: the rate of activation is proportional to the immune challenge such as a tumor population, that is, we postulate a term βu in $\dot{x}$. It is virtually impossible to give a numerical value for the parameter β, since this value depends on the nature of the immune challenge, spatial relations between antigen presenting cells and T cells, and so forth. Khailaie et al. (2013) simply use a term $+k(t)$ to represent this stimulation (where $k(t)$ is the product of antigen stimulation "β" and the supply N of naive T cells or resting Treg cells, and introducing an unspecified multiplier to model possibly different effects on T cells compared to Tregs). This additive input is naturally modeled by $\beta = 1$, and we take that simplest possible value. Units are day^{-1}.

The term $\mu u/x$

There are various ways to justify this term. We picked a mathematical form for the effect of the immune challenge u and regulatory elements x on effector cells y that is the simplest possible to model activation by u and repression by x. Let us discuss why this choice is reasonable phenomenologically. The term "regulatory T cell" (Treg) actually encompasses several subclasses of cells that help in peripheral tolerance, preventing autoimmune diseases, and down-modulating immune responses. These cells they affect many other immune components, from B cells to helper cells (Th1, Th2, Th17) and cytotoxic T cells, through both direct and indirect interactions. These interactions form an extremely complicated and poorly understood network that includes inhibitory molecules such as CTLA4 and messaging by cytokines (TGF-β, IL-10, IL-35, and others) which result in the suppression of helper cell differentiation and in indirect down-regulation of MHC and costimulatory molecules on antigen-presenting cells, thereby reducing T cell activation. The repression of T cell activation through TCR-MHC is one way to see the negative effect of x on y. Another is the indirect effect through inhibitory cytokines such as IL-10, TGF-β, and IL-35 that can suppress T cell activation. The simplest mass-action kinetics model would assume independent effects: activation by u and repression by x, leading to a term of the form $h_1(u)h_2(x)$ driving y activation. For the effect of u, let us take $h_1(u) = c_1 u$, for some constant c_1. If we assume that x cells (or messenger molecules) repress through binding to a certain type of receptor, and R represents the number (or fraction, or concentration, depending on units) of free receptors, then at equilibrium we would have $kRx = R_0$, where R_0 quantifies occupied receptors, and from a conservation $R + R_0 = R_T$ assuming a constant total number of receptors, we would have that $R = R_T/(1 + kx)$ is the number of free (unbound) receptors, so unless $k \ll 1$ we may take $h_2(x) = c_2/x$ for some constant c_2. These arguments will result in the algebraic form $h(x,u) = Mu/x$. A different justification is as follows. Let us assume that there is an intermediate variable z, which might represent for example a population of helper T cells (Th cells or CD4$^+$ T cells) which helps activate the cytotoxic T population y and is itself activated by the immune challenge u and inactivated by the regulatory variable x. The simplest equation would be $\dot{z} = -\mu_0 xz + \beta u$, where we are assuming that helper cells are also being activated in a manner proportional to the magnitude of the immune challenge, and

$\mu_0 x$ represents the x-dependent degradation of z. We assume that $\dot{y}$ has a term z corresponding to activation by helper cells. Assuming that this equation is at equilibrium, we may substitute $z=(\beta/\mu_0)u/x$ into the $\dot{y}$ equation, giving a term $\mu u/x$, where $\mu=\beta/\mu_0$. (If helper and T cell activations are at similar timescales and the equilibrium assumption is not made, one add may the z differential equation explicitly. We prefer to keep the model simpler, but see Supplement Section F for simulations using that model.) Khailaie et al. (2013), include in T cell dynamics a similar mass-action degradation or inactivation term, using a rate constant 0.1. Following this, we pick the value $\mu_0=0.1$, so that, together with $\beta=1$ we have $\mu=10$. As u and x are both in units of 10^6 cells, μ has units day^{-1}.

The terms λu and $-\kappa yu$

The term λu is a standard exponential growth term. We view λ as a varying parameter, which quantifies the initial exponential growth of the immune challenge.

The killing term $-\kappa uy$ in the $\dot{u}$ equation represents a simple mass-action suppression of the immune challenge, such as cytotoxic T cells killing tumor cells. The constant κ depends on many factors, such as the type of tumor, size and geometry of tumor microenvironment, accessibility of tumor cells to vasculature, and so forth. In the original paper by Kuznetsov et al. (1994), one finds $\kappa=1.101\times10^{-7}$ in units of day^{-1} cells^{-1} which when normalized to units of 10^6 cells would give the value $\kappa=1.101\times10^{-1}$, This value seems to be too large for most cancers. For example, based on fits to experimental data, the recent paper Wang et al. (2015) obtains a number which is many orders of magnitude smaller. That paper analyzes the killing by cytotoxic CD8+ T cells of MHCI$^+$ tumor cells in a B16 mouse metastatic melanoma model, and determines a killing term for such cells of the following form (with different notations here): $-[c/(\varepsilon+U)]Yu$, where Y is the concentration of effector CD8$^+$ T cells in the tumor microenvironment, using units of cells/mm^3, c is a constant that quantifies MHCI positive tumor death rate due to T effectors, and has the value 2.49×10^{-13} in units mm^3 day^{-1}, U is the total number of tumor cells, ε is a "small number" to account for other cells, and u is the number of major histocompatibility complex class I positive tumor cells. Since $[c/(\varepsilon+U)]Y$ has units day^{-1}, if we convert to y in units of 10^6 cells, we obtain $cY=\kappa y$ where $k=2.49\times10^{-7}/(\varepsilon+U)$ has units (10^6 cells^{-1} day)$^{-1}$. Depending on the number of cells U in the tumor, this number κ could be very small, and it is certainly less than 2.49×10^{-7}. To take another example, Kirschner and Panetta (1998) employ a Michaelis-Menten killing term $-auy/(g_2+y)$, with $a=1$ and $g_2=10^5$. Given these wide ranges, we pick $\kappa=10^{-5}$ for our simulations. Units are (10^6 cells)$^{-1}$ day^{-1}. (A two-zone behavior of tumor elimination can also be found with $\kappa=10^{-4}$, $\kappa=10^{-3}$, $\kappa=10^{-2}$, and $\kappa=10^{-1}$, but shifting the range of λ's at which different behaviors arise.)

Sensitivity to parameters in the function f

We recall the definition of the function f:

$$f(y)=\frac{Vy^2}{K+y}-\varepsilon y^2-\delta_y y\ .$$

The main requirement for the theoretical analysis in the main text is that f have a cubic form as illustrated in Figure 1, so that then the nullcline analysis in Figure 2 applies. In other words, f should have one zero at $\eta_1=0$ and two positive zeros η_2, η_3 so that $f(y)<0$ for $\eta_1<y<\eta_2$, $f(y)>0$ for $\eta_2<y<\eta_3$, and $f(y)<0$ for $\eta_3<y$. (Observe that signs gets reversed in the nullclines in Figure 2, because of the negative sign in the formula $p=h^{-1}(-f(y))$.) Writing $-f(y)=(y/K+y)g(y)$, where

$$g(y)=\varepsilon y^2+(\delta_y-V+K\varepsilon)y+K\delta_y$$

and using that $y/(K+y)$ is positive for $y>0$ and zero at $y=0$, the requirements on f translate into the requirement that the parabola $g(y)$ have two positive zeros η_2, η_3 (and be negative in between them), which is equivalent to:

$$(\delta_y-V+K\varepsilon)^2>4\varepsilon\delta_y K\quad\text{and}\quad\delta_y-V+K\varepsilon<0$$

For our parameters, $V=0.25$, $K=100$, $\delta_y=0.1$, $\varepsilon=10^{-5}$, we have $\delta_y-V+K\varepsilon\approx-0.1490$, $(\delta_y-V+K\varepsilon)^2\approx0.0222$, and $4\varepsilon\delta_y K=4\times10^{-4}$, so that these conditions are satisfied. These requirements imply that the maximal autocatalytic strength V should be large, and the degradation constant δ_y and the fratricide constant ε should be small.

E. Nullclines for model and parameters used in text

Figure S3 shows the nullclines for this system for various increasing values of λ, as well as some typical solution trajectories, showing their convergence to values under, over, under, and finally again over the threshold which determines tolerance or rejection of the immune challenge. This is perfectly consistent with our theoretical predictions.

F. A model with an intermediate population

We consider here that a slightly different model, in which u and x affect the effector variable y only indirectly, through production and repression respectively of a "helper cell" population. Figure S4 plots simulation results (all parameters exactly the same as in earlier model), showing that this model leads to similar results as those for the simpler model.

$$\begin{aligned}\dot{u} &= [\lambda(1-Bu)-\kappa y]u \\ \dot{x} &= -\delta_x x + \beta u \\ \dot{y} &= z + \frac{Vy^2}{K+y} - \varepsilon y^2 - \delta_y y \\ \dot{z} &= u - (1/\mu)\, xz\end{aligned}$$

G. More details on exponential rate detection and two-zone experimental results

In our model, an embedded IFFL acts as an estimator of the rate of exponential increase of the immune challenge. I briefly mentioned the work of Johansen et al. (2008). Let me discuss here some more relations to that work. The authors state that ''antigenic stimulation increasing exponentially over days was a stronger stimulus for CD8 T cells and antiviral immunity than a single dose or multiple dosing with daily equal doses'' and concluded that ''at a clonal level, T cells are capable of decoding the kinetics of antigen exposure.'' They found that IL-2 activation at constant dosage of antigen is almost zero, at linearly increasing dose is higher, and at exponential doses is highest, and concluded (Figure 7, caption) that ''exponential in vitro stimulation of CD8 T cells enhances IL-2 production and cytotoxicity.'' These experimental observations are all roughly consistent with activation of the autocatalytic loop in our model under higher exponential rates. In 2008, Kündig and collaborators, based on this work, obtained a patent (Kundig et al., 2008) for ''A method for enhancing T cell response'' based on the principle that immunogenicity is enhanced by ''exponentially increasing antigenic stimulation of class I MHC CD8+ T cell response …in a manner independent of the dose of the antigen.''

Another conclusion of the analysis is the existence of intermediate regions of challenge (e.g., tumor) growth in which the challenge will be eliminated by the immune system, with challenges in lower as well as in larger regions not being eliminated. This existence of disjoint regions of tumor elimination depending on rate of growth is strongly reminiscent of two related phenomena, ''sneaking through'' and ''two-zone'' tumor tolerance, which have been much discussed since the mid-1960s. The idea of tumors ''sneaking through'' from immune control originated with the findings in Klein (1966) of intermediate regions in which tumors can be eliminated. Further, Gatenby et al. (1981) argued that this four-region phenomenon specifically depends on T-cell repression (just as in our model through the regulatory x variable), and framed this role of suppressor T cells on regulating tumor immune response in the more general idea of low zone tolerance (tolerance to antigens under repeated exposure to small antigen doses). This work was, in turn, motivated by seminal work by Haubeck and Kölsch (1982), who injected exponentially increasing numbers of irradiated syngeneic ADJ-PC-5 plasmacytoma cells into BALB/c mice, starting with 2 cells at day 1, 4 at day two, and doubling subsequent doses for 15 days until about 10^5 were received, and proposed the induction of T suppressor cells (what one now calls Treg cells) as an early event in tumorigenesis that regulated CTL activity. To test their ideas, Gatenby et al. (1981) carried out experiments that show sneaking-through behavior as well as the failure of this behavior when ''suppressor T cells'' are eliminated, see Figures 4A and 4B respectively. Murine sarcoma Meth A was administered in varying doses to BALB/c mice, and incidence of tumors was measured in each group of 12-42 mice, at two weeks after the last mouse died from tumor. Similar results on sneaking-through had been reported by Kölsch and Mengersen (1976) in previous work in which mastocytoma BM3 injected cells were injected into BALB/c mice, see Figure 4C. Care must be taken when interpreting these experimental numbers in terms of a model. The numbers reported are for ''tumor incidence,'' meaning percentages of mice in which tumors were detected by some predetermined point. If we assume that survival (until mouse sacrifice, or indirect death from the tumor) depends probabilistically on tumor size, then we could think of tumor incidence as a proxy for size. Another difference is that, in these works, the different regions correspond to the magnitude of an initial tumor inocula in animal subjects, rather than growth rates. Nonetheless, there is a surprisingly strong resemblance between our plots and the experimental ones. This picture is at least consistent with a larger initial rate of increase in exposure leading to tumor suppression, as in our model. Also related to these general ideas are the papers McBride and Howie (1986) and Bocharov et al. (2004).

Cell Systems, Volume *4*

Supplemental Information

A Dynamic Model of Immune Responses to Antigen Presentation Predicts Different Regions of Tumor or Pathogen Elimination

Eduardo D. Sontag

Supplemental figures and captions

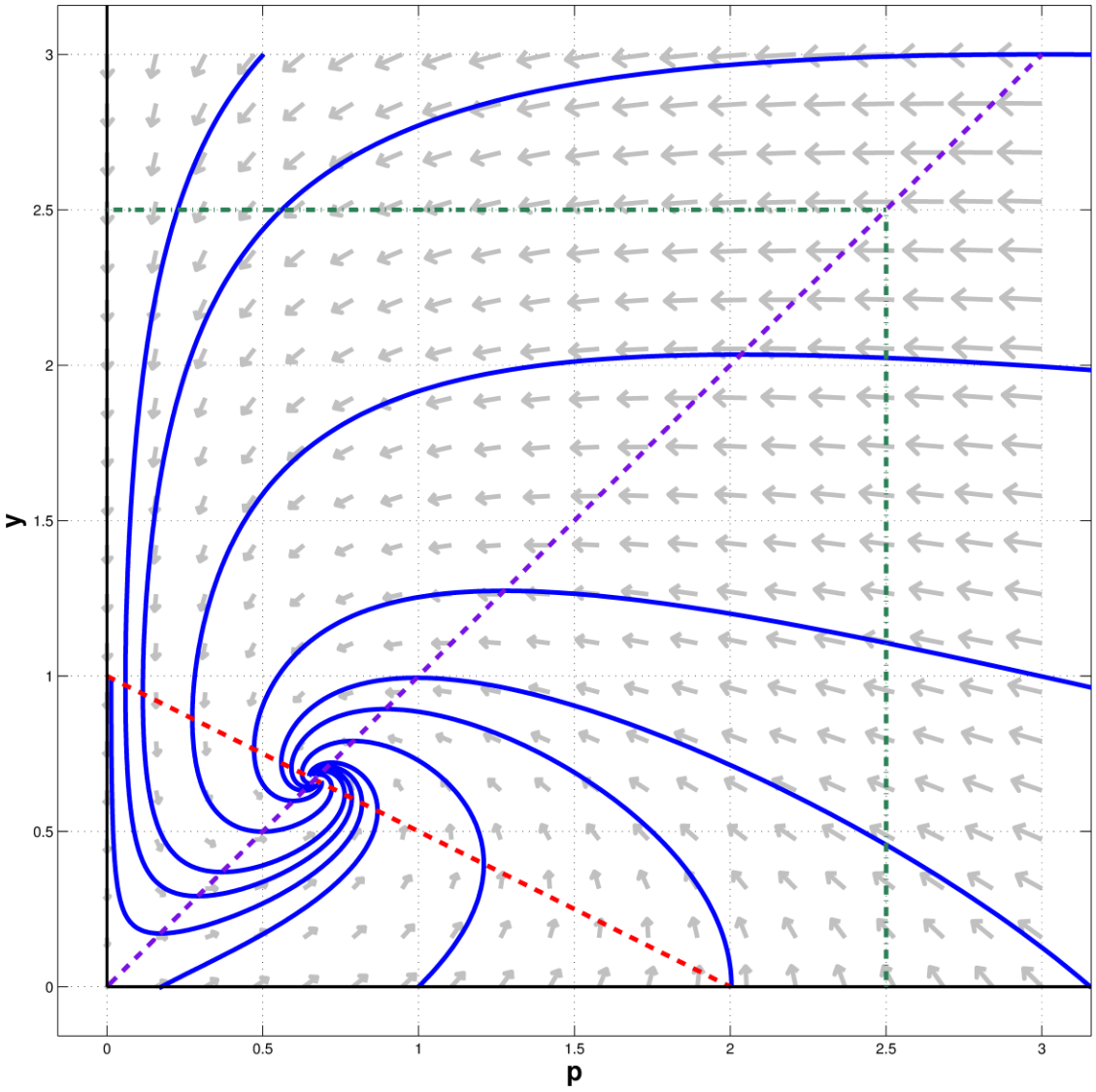


Figure S1. Related to the STAR Methods.

Phase plane for numerical example, with several representative trajectories plotted. Nullclines are the y axis, corresponding to the stable manifold of $(0,0)$, and the lines given by $y = (\delta_x + \lambda - \beta p)/\kappa$ (dashed red line) and $y = \mu p/\delta_y$ (dashed magenta line). In this plot, we picked $\delta_x = \beta = \mu = \lambda = \delta_y = 1$ and $\kappa = 2$, but the qualitative picture is similar for all valid parameter values. With these values, trajectories converge to the equilibrium $(\overline{p}, \overline{y}) = (2/3, 2/3)$. Shown also is an invariant region $[0, P] \times [0, Y]$ with $P = Y = 2.5$ (green dash-dotted lines and axes).

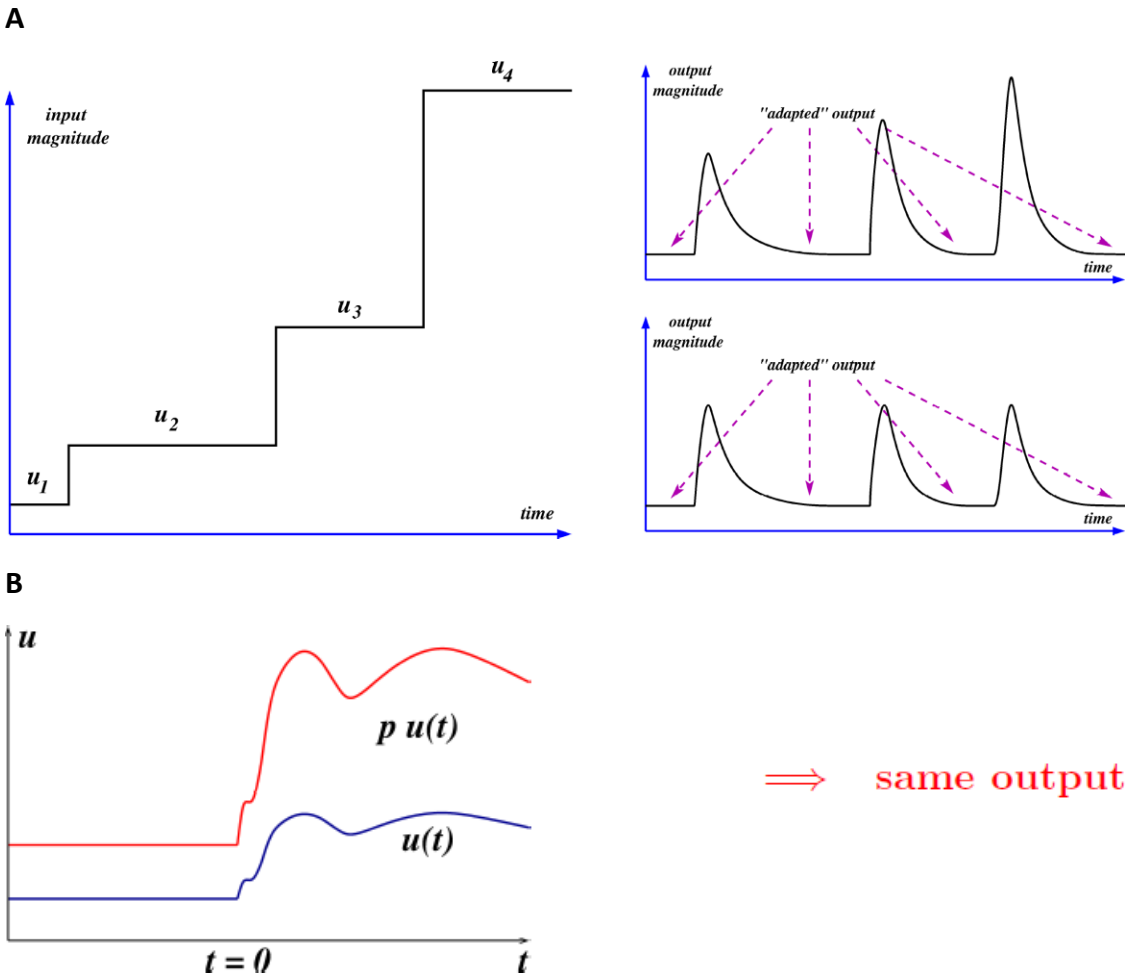


Figure S2. Related to the STAR Methods.

A. *In a perfectly adapting system, a step-wise input (left) gives rise to different responses that settle to the same basal level (top right). If the system has the scale invariance property, these responses are identical (bottom right).* ***B.*** *Scale invariance means that scaled signals should result in the same output, provided that the initial state is preadapted to the respective constant value for* $t < 0$

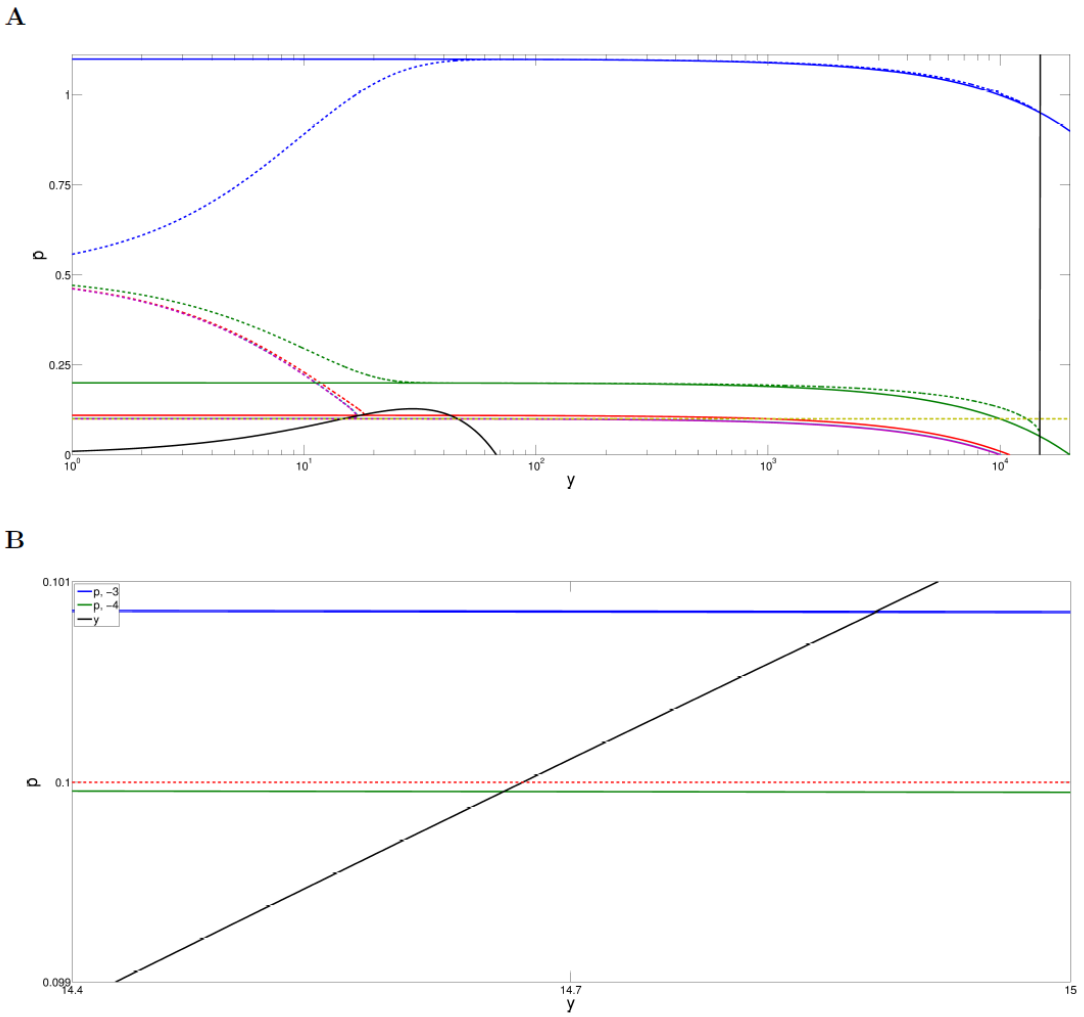


Figure S3. Related to the STAR Methods.

A. *Phase plane for the system described in the text. Shown are nullclines for several increasing values of* $\lambda = 10^i$, $i = -4, -3, -2, -1, 0$ *(bottom to top). Log scale in* y *is used in order to visualize behavior for different orders of magnitude in* λ. *Black curve is* y*-nullcline* $p = -(1/\mu)(Vy^2/(K+y) - \varepsilon y^2 - \delta_y y)$ *Solid color curves are* p*-nullclines* $p = (1/\beta)(d + \lambda - \kappa y)$ *(which look curved because of log scale), and* $p \equiv 0$. *Horizontal dashed line is threshold* $p = \delta_x/\beta$ *that determines if* $\lambda - \kappa y$ *is positive or negative. Dashed curves are trajectories, in same color as the respective nullclines. Initial state for simulation is in every case* $p(0) = 0.5$, $y(0) = 0$, *but only portion of plot for* $y \geq 10^0 = 1$ *is shown.* ***B.*** *Zoomed-in view of two nullclines, for* $\lambda = 10^{-4}$ *and* $\lambda = 10^{-3}$, *to show how steady state falls under/below threshold.*

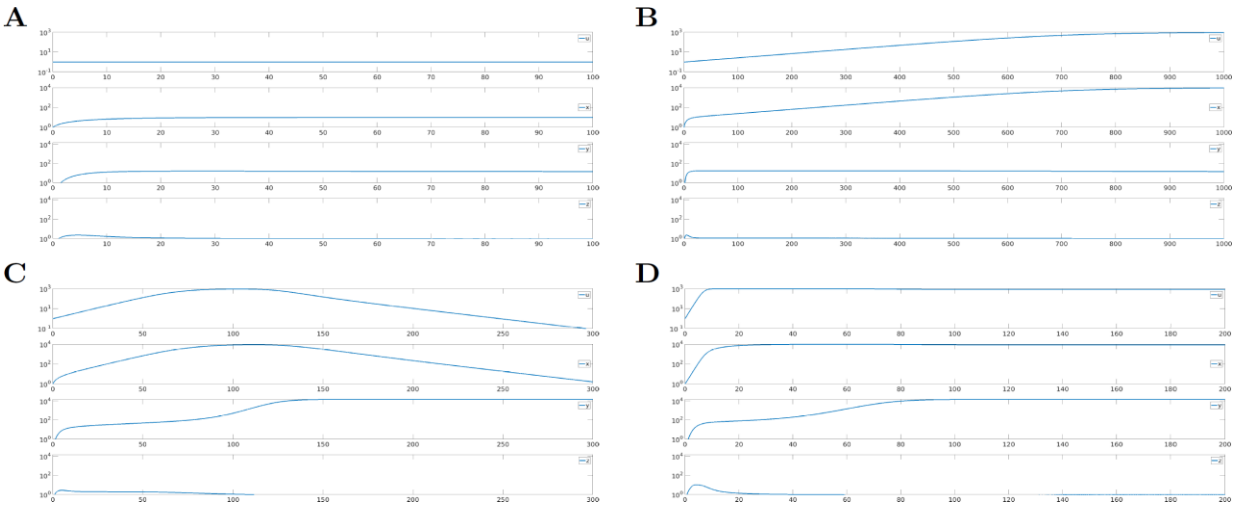


Figure S4. Related to the STAR Methods.

A-D. *Simulations of system with "helper" intermediate, for* $\lambda = 10^{-4}, 10^{-2}, 10^{-1}$, *and* 1. *(In* ***A****,* $u(t)$ *converges to zero as* $t \to \infty$*, but very slowly.) Parameters as described in text for Figure 3. Initial states are always* $u = 1$, $x = 1$, $y = 0$, $z = 0$. *Only noticeable difference with simpler model is a slight delay in activation of* y.