Duration of infection is a determinant of both individual health and public health. The number of days, weeks, months, or even years that an infection persists in a host has important health implications on multiple biological scales. For an infected individual, for example, the time it takes to clear an infectious agent can affect the likelihood that the infection will become lethal (e.g., for pneumonia [1]), the cumulative severity of symptoms such as anemia (e.g., during chronic malaria [2]), and the amount of tissue damage that must be repaired if the individual is to recover (e.g., from viral hepatitis [3] or gastrointestinal helminthiasis [4]). At the population scale, the longer individuals are infected, the longer they tend to be infectious to others, whether by persistent coughing (e.g., for tuberculosis; [5]), spreading of fecal matter (e.g., by Typhoid Mary [6] and modern analogues [7]), or availability of transmissible propagules to biting vectors (e.g., for malaria [8]). Accordingly, infection duration and its inverse, clearance rate, are canonical parameters in models of dynamic epidemiology [9-11]. For example, the mean and variance of infection duration in a population predict epidemic outbreak risk (e.g., for norovirus [12]) and ease of control by public health interventions (e.g., for sexually-transmitted [13] and vector-borne [14] infections).

Yet determinants of infection duration are poorly understood. Duration arises from a complex, dynamic interplay of host and parasite genetics with environment. In general, “resistant” host genotypes clear parasites more rapidly than “susceptible” genotypes do, via powerful immune responses that mobilize appropriate effector mechanisms (e.g., secretion of mucus by goblet cells, accelerated epithelial turnover, and peristalsis to expel gastrointestinal nematodes [15]). At the same time, parasite genotypes that best immunosuppress (e.g., [16]) or co-opt (e.g., [17]) the host often generate infections of longest duration. Infection duration thus varies with host genotype [18-21], parasite genotype [22-24], and sometimes both (e.g., GHxGP interactions for duration [25, 26]). Considerable effort has been expended to understand the molecular and cellular mechanisms behind these genetic effects, yet genetics (even GHxGP interactions) are not the only drivers of variation. Indeed, infection duration also varies strongly with environmental factors such as doses or rates of exposure [27-30]. Controlled experiments also reveal variation in infection duration among hosts of a given genotype exposed to the same dose [18, 31, 32], even when hosts are exposed to clonal isolates of parasites such as malaria [33] or streptococci [34].

For example, “resistant” mouse genotypes are able to quickly clear high doses of the gastrointestinal parasite *Trichuris muris*  by mounting a strong T-helper 2 (Th2) immune response that mobilizes effector mechanisms, such as secretion of mucus by goblet cells, accelerated epithelial turnover, and peristalsis [28, 34-36]. “Susceptible” genotypes, on the other hand, mount an inappropriate T-helper 1 (Th1) response and become chronically infected. Counterintuitively, if the inoculating dose is reduced, “resistant” strains can become chronically infected, too. This led researchers to suggest that a threshold dose may be required to initiate a strong Th2 response (with thresholds likely to differ among host strains) [35]. The suggestion of thresholds governing infection outcomes is also suggested by experimental infections of fruit flies [37] and flour beetles [38]. In both systems, the duration of infection was acute in some insects, and chronic in others, despite stringent controls, with acute versus chronic outcomes being determined by subtle differences in the initial rates of immune response induction and parasite replication.

We argue that the existence of thresholds in duration and the subtle dependence of duration on initial conditions and early events are indicative of Allee effects driven by strong feedback mechanisms. Allee effects are a relatively common phenomenon in ecological systems. They arise when positive feedback loops generate a positive relationship between per-capita growth rate and population density. The key dynamical signatures of Allee effects are persistence thresholds (e.g., [39, 40]): when density is below the threshold, the population declines to extinction; above it, the population persists. Near the threshold, subtle differences in system state can produce strikingly different persistence times. Two recent mathematical models suggest alternative feedback mechanisms that could produce Allee effects. Van Leeuwen et al. [] showed that parasites can avoid clearance if they can force hosts to reallocate resources away from immunity and towards parasite growth. Ellner et al. [] showed that parasites can avoid clearance by directly interfering with the immune response by sequestering, inactivating, and inhibiting immune effectors. In both models, the key to producing duration thresholds was that host manipulation was dependent on parasite biomass, generating a positive relationship between parasite per-capita growth rate and parasite biomass leading to Allee effects.

We suggest that infection duration thresholds are likely an intrinsic feature of the dynamics of the mammalian immune response. Dynamical systems theory shows that multistability (the existence of multiple, stable equilibria for a single set of parameter values) is critically dependent on the existence of positive feedbacks between variables in the system. (Mutual inhibitory feedback loops function similarly.) Such positive feedback loops are ubiquitous within the mammalian immune response. In particular, polarization of the T cell populations towards either a T-helper 1 (Th1) phenotype (which coordinates the response against intracellular microparasites) or a T-helper 2 (Th2) phenotype (which coordinates the response against macroparasites and thus promotes nematode expulsion) is driven by positive feedback between cytokine production and T cell activation, such that the per-cell growth rate of a T cell subpopulation increases as more T cells join that subpopulation [41], and by mutual inhibition between T-cell subpopulations. These feedback processes can be hijacked by the parasite when there is a positive relationship between parasite biomass and immunomodulation, such that the per-gram growth rate of the parasite may increase with its biomass (e.g., due to escalating manipulation; [26]).

Here we derive and analyze a minimal model for the interaction between parasite biomass and the adaptive immune response that incorporates the key feedback mechanisms above. We show that this model is capable of producing acute versus chronic thresholds. Further, we explore how changing the relative strength of different feedback mechanisms alters the possible dynamical outcomes of the system, recapitulating empirical patterns like the dose-dependence of infection outcome. We suggest that building these mechanistic immune feedbacks into models of host-parasite interactions provides testable insights into how feedback processes drive duration.

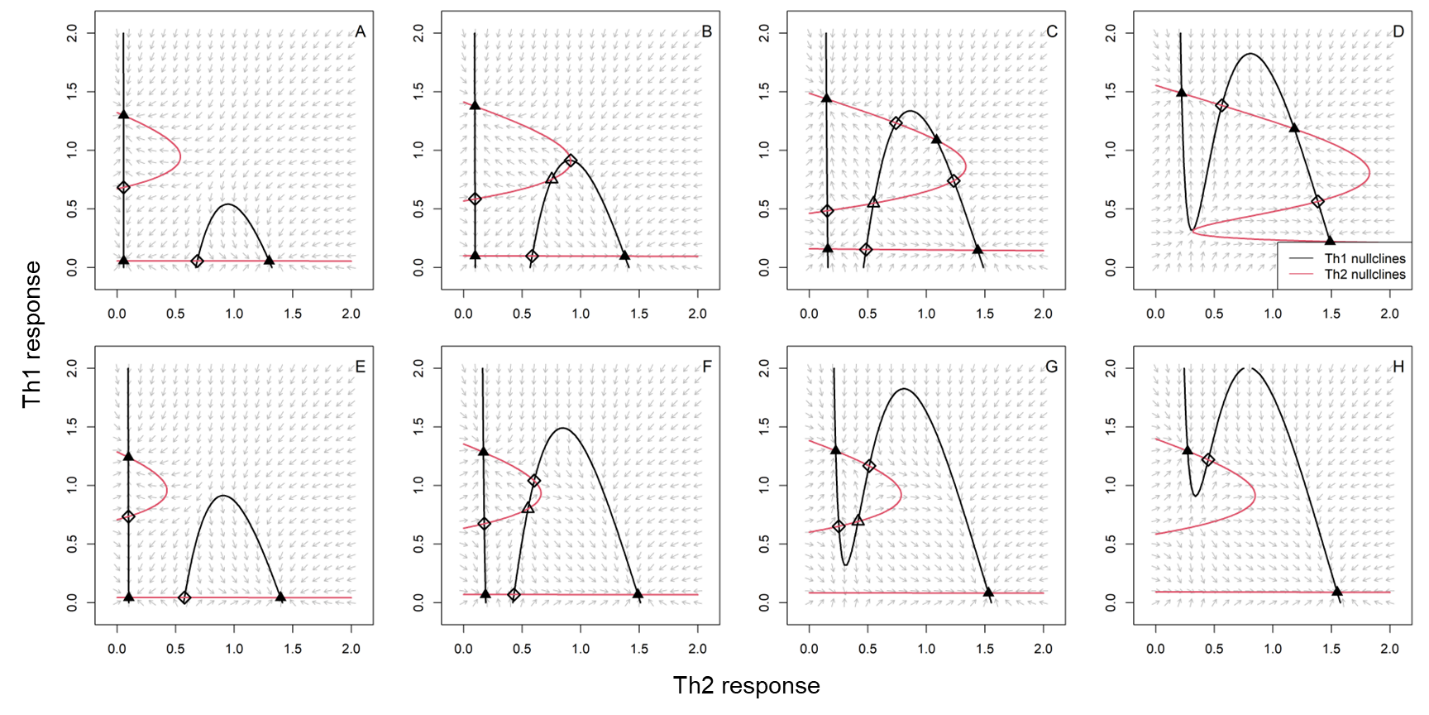
**Results**

We begin by considering the interaction between the two main “arms” of the mammalian adaptive immune response: T-helper 1 and T-helper 2 cells. Th1 and Th2 cells are the primary coordinators of the adaptive immune response, and polarization towards one cell type or the other drives parasite clearance or chronicity.

In our baseline model, we will ignore the dynamics of the parasite, since the positive feedbacks intrinsic to the dynamics of immunity should be sufficient to generate immunological Allee effects (Schrom et al. 2020, van den Ham and de Boer 2008, Yates et al. 2000, 2004), even in the absence of parasites. Thus we assume that there is T-cell independent activation of Th1 and Th2 production, at rates governed by the parameters and . Activated T-helper cells further increase their population sizes by producing the same cytokines that induced their own activation, termed self-activation. At the same time, the cytokines produced by activated T-helper cells prevent naïve T-helper cells from becoming T-helper cells of the opposite type, termed cross-inhibition. In reality, the activation of a naïve T-cell depends on the intracellular interaction between cytokines and the master regulator transcription factors T-bet and GATA3 (Yates et al. 2004): Th1 cytokines stimulate the expression of T-bet, and T-bet expression regulates the production of Th1 cytokines; Th2 cytokines stimulate the expression of GATA3, and GATA3 expression regulates the production of Th2 cytokines. Interestingly, T-bet and GATA3 expression are *also* self-activating and cross-inhibiting, which has been shown to produce multistability in individual immune cell phenotypes (Schrom et al. 2020, Yates et al. 2004; van den Ham and de Boer 2008). For simplicity, we here assume that both transcription factor expression in individual cells and the production of cytokines can be captured by a self-activation term that is dependent on the density of T-helper cells of the same type, according to a Hill function with an exponent of , ; and by a cross-inhibition term that is dependent on the density of T-helper cells of the opposite type, also according to a Hill function with an exponent of , . Finally, we include immune cell apoptosis due to any of a number of processes (Yates et al. 2000) with the parameter . The resulting system of two equations is:

In the Supplemental Information, we show that the possible dynamical outcomes of the model are highly sensitive to the choice of Hill exponents, so for the remainder of our analysis we will focus on the case where and , which produces switch-like behavior in immune self-activation. To analyze the possible dynamics of this system, we study the nullclines for this system: combinations of and that cause ( nullcline) or ( nullcline). Intersections of these nullclines represent the possible equilibria of the system.

Determining parameter values for this system is challenging: although there are good estimates for the parameters of underlying models of transcription factor expression and cytokine production (Schrom et al. 2020), the complex dynamics produced by such models make it infeasible to use, for example, quasi-equilibrium assumptions to estimate the parameters of this model. Instead, we will avail ourselves of the technique of nondimensionalization often favored in the analysis of these models (Ellner et al. 2021; van den Ham and de Boer 2008; Yates et al. 2000); the dimensionless parameters of the model typically have simpler interpretations. Specifically, we define dimensionless state variables and . Thus the dimensionless measure of immunity is relative to the immune abundance such that self-activation is half its maximum, and time is measured relative to the average lifespan of an immune cell. We also define the dimensionless parameters , , and . Although and do not have a meaningful biological interpretation, is a measure of cross-inhibition relative to self-activation. In particular, if , then cross-inhibition increases faster than self-activation. The dimensionless system is then:

Figure 1 shows the configuration of the nullclines as the activation rate is increased; for example, by the growth of immune-stimulating parasite biomass. Panels A-D (top row) show the nullcline configurations for the case of a balanced response (all parameters are equal); panels E-H (bottom row) show the nullclines when the system is biased towards a Th2 response (­). In a perfectly balanced system (Fig. 1, top row), at low external activation levels, there are three equilibria corresponding to a low activation state and each polarized response, with the basin of attraction being the largest for the low activation state. As external activation increases, we see the emergence many new equilibria, including a stable equilibrium representing high co-activation of both immune response “arms.” Eventually we lose the low activation equilibrium altogether, and further increases in activation would eventually result in the loss of both polarized responses as well, leaving the high coactivation equilibrium as the only possible outcome.

In the biased response case (Fig. 1, bottom row), we see that the basin of attraction for the Th2-polarized response is much larger initially, we never see the emergence of a high coactivation equilibrium, and at very high levels of activation the polarized response will be the only equilibrium. This indicates that a growing parasite population that provokes a biased immune response can eventually lead to immune polarization.

However, that ignores the reality that the immune response inhibits parasite growth. To account for this, we extend the model to include parasite growth. We assume that the parasite is inhibited by a T-helper 2 response (e.g., it is a extracellular parasite), though that assumption is not critical to our results. We further assume that the parasite can directly stimulate both Th1 and Th2 responses in a T-cell independent manner; this reflects the biological reality that the detection and activation of T-cells by parasite antigen is largely independent of the abundance of T-cells, as detection and activation are done by other cells of the immune response (Bonhoefer and Nowak 1994, Fenton et al. 2006, Alizon et al. 2008). We assume for simplicity that the parameters governing this response (the maximum activation rate and the half-saturation constant ) capture the dependence of this process on the abundances of antigen presenting cells and naïve T cells, and the rate that these cells encounter one another (Yates et al. 2000). Also for simplicity, we assume a very simple form for the parasite dynamics: the parasite grows logistically in the absence of any immune regulation, and the parasite’s mortality rate is a linear function of the abundance of Th2 cells; in reality, of course, Th2 cells are not effector cells, so the per Th2-cell cell explusion rate parameter, , is assumed to capture all of the downstream upregulation of immune effectors triggered by the activation of the Th2 response.

We can again create a dimensionless version of this model, using the same dimensionless variables and parameters as above

One of the challenges with analyzing a multistable model is identifying the most interesting combinations of initial conditions to study. In this model we can vary the initial parasite abundance and the initial abundances of both Th1 and Th2 cells. In particular, both the number and balance between Th1 and Th2 matter for the ultimate dynamical outcome. In Fig. 2, we show the dynamical outcomes as the initial parasite dose and the initial balance between Th1 and Th2 differs, but as the initial total number of Th1+Th2 cells remains fixed. Across the panels of Fig. 2, we are increasing the activation of the Th2 response by the parasite. In the Supplemental Material, we show the same figures as the initial total number of Th1+Th2 cells varies.

|  |  |  |
| --- | --- | --- |
| Parameter | Definition | Value |
|  | Activation rate of naïve T-cells |  |
|  | Maximum T-helper cell self-activation rate |  |
|  | T-helper cell density where self-activation is half its maximum value |  |
|  | T-helper cell density where cross-inhibition is half its maximum value |  |
|  | T-helper cell apoptosis rate |  |
|  | Exponents for the Hill functions governing self-activation and cross-inhibition | 2, 1 |

**Discussion**

Theoretical immunologists have long been aware that immune polarization, whether of T cell populations or individual T cell phenotypes, is driven by self-activating and cross-inhibiting feedback processes (e.g., Th1 cytokines increasing expression of the transcription factor T-bet, which regulates expression of Th1 cytokines; van den Ham and de Boer 2008, Yates et al. 2000, 2004). However, this model is the first we are aware of to connect those immune feedbacks to a model of parasite growth, showing how qualitatively distinct infection outcomes (clearance versus chronicity) can result from changing the initial immune and parasite conditions, and how changing the strength of these feedbacks can alter the potential for multistability.

These feedback mechanisms are essential components of the immune response. Theoretical ecologists often analogize immune-parasite interactions with predator-prey interactions, but it is the self-activating and cross-inhibiting aspects of the immune response that give the lie to that analogy. Parasite clearance is only possible because of these immune-driven feedback loops; in the absence of immune self-activation, when parasite abundances become low, so would immune activation, leading to coexistence of the immune system and parasite. However, these immune feedbacks can be hijacked by parasites to their own ends. For example, the mouse whipworm, *Trichuris muris*, produces an excretory/secretory product that immunomodulates the host by binding to a key Th2 cytokine (Bancroft et al. 2019); this inhibition causes the system to polarize towards an inappropriate Th1 response that is unable to clear the parasite (Cliffe et al. 2005).Thus we find, lying within the feedback mechanisms of the immune response itself, the seeds of the immune system’s failure to clear a parasitic infection.

In the model of van Leeuwen et al., the positive feedback loop is that parasite growth reduce resources that increase immunity that reduces parasite growth; in the model of Ellner et al., the positive feedback loop is that parasite growth reduces the immune response that reduces parasite growth. Thus, in both models, the acute vs. chronic thresholds are created by the parasite actively manipulating its host. Here, we show that the positive feedback (and mutual inhibition) loops inherent to the immune response can produce these thresholds as well, such that even a parasite that is not capable of direct immunomanipulation can still trigger incorrect immune polarization; for example, even a parasite that stimulates a strong Th2 response can potentially become chronic if the immune system starts out biased towards a Th1 response.

The mathematical analysis in the Supplemental Information and the numerical analysis in Fig. 1 both show that the magnitude of the activation rate strongly influences the potential for multistability, with multistability becoming less likely as activation rate increases. This makes biological sense: if the immune system is strongly stimulated, then the possible dynamical outcomes should become more limited as the immune dynamics are more strongly pushed towards polarization. The analysis in the Supplemental Information also reveals the importance of relative sensitivity of self-activation versus cross-inhibition. The figures shown in the main text keep self-activation much stronger than cross-inhibition by setting the parameter . In the Supplemental Information, we show that as that parameter decreases, and cross-inhibition becomes more sensitive to changes in T-cell abundance than self-activation, then the dynamics become simpler. In particular, if the parasite stimulates Th1 and Th2 immunity relatively equally, then the system is very likely to become trapped at a low-activation equilibrium where the immune response remains “off,” whereas if the parasite stimulates Th2 immunity more strongly than Th1, the only outcome is rapid clearance.

These results suggest possible explanations for the variety of clearance phenotypes observed in nature. For example, many intracellular pathogens have a predictable and short infection duration (at least among individuals who survive infection).

Because the dynamics of immune phenotype development is governed by positive feedback loops between cytokines and transcription factors, host-parasite interactions are governed by processes operating at the host population scale (e.g.1, transmission), at the cellular scale (e.g., immune killing), and at the molecular scale (e.g., immune cell phenotypes).