UNIVERSITY OF CALIFORNIA

Los Angeles

Evolution of immune system against diverse antigens

A dissertation submitted in partial satisfaction
of the requirements for the degree
Doctor of Philosophy in Physics and Astronomy

by

Jiming Sheng

2021

ABSTRACT OF THE DISSERTATION

Evolution of immune system against diverse antigens

by

Jiming Sheng
Doctor of Philosophy in Physics and Astronomy
University of California, Los Angeles, 2021
Professor Shenshen Wang, Chair

The immune system evolves across the host's lifetime to protect against the wide array of threats in nature. While the immune system is capable of evolving and adapting to a single antigen (Ag), it becomes a challenging task to defend against diverse antigenic targets, including mutants of a same pathogen or a wide spectrum of pathogen species. First, the rapid intra-host diversification of highly-mutable pathogens such as human immunodeficiency virus (HIV) or hepatitis C virus (HCV) creates a coevolutionary arms race with the immune system. As a result, viruses persist into a chronic infection in most subjects and are only cleared in rare cases. In addition, The vaccination trials so far to elicit broadly-neutralizing antibodies (bnAbs) against highly-mutable viruses have met with failure. Second, the immune system has to allocate its finite amount of adaptive immune cells against the wide spectrum of pathogens in the environment. As memory cells accumulate from each pathogen encounter, the host's immune repertoire gradually becomes skewed: more adaptive immune cells are dedicated to the frequent pathogens while fewer are reserved for the rare pathogens. The skewed repertoire in the elderly has been correlated with immune risk phenotype and a chronic inflammatory response even in the absence of pathogen, but whether there is a mechanistic connection remains unknown.

My dissertation aims to address the following questions regarding evolution of immune

system against diverse antigenic targets: (1): What are the mechanisms and deciding factors behind the distinct coevolutionary outcomes observed in different subjects? (2): What makes a viable vaccine design strategy to guide immune system evolution towards bnAbs? (3): What are the side effects of a skewed immune repertoire, as a result of adapting to different pathogens encountered during hosts' lifetime? Also, how can human intervention alleviate these side effects?

My dissertation shows that (1): different coevolutionary outcomes are decided by the timing and efficacy of successive narrow and broad antibody (Ab) responses, which in turn are determined by the conservation level and initial diversity of Ag. (2): A viable vaccine strategy to elicit bnAbs should balance suppression of strain-specic B cells and preserving cross-reactive B cells. The corresponding optimal selection strength should increase in time as driven by the evolution of B cell cross-reactivity. (3): Mediated by adaptive-innate feedback, repeated pathogen encounters during host lifetime and resulting memory inflation may trigger a fragility, in which any encounter with a novel pathogen will cause the system to irreversibly switch from health to chronic inflammation (CI). In addition, the onset of CI strongly depends on the history of encountered pathogens; the timing of onset can be delayed drastically when the same set of infections is encountered in a specific order.

The dissertation of Jiming Sheng is approved.

Robijn Bruinsma

Giovanni Zocchi

Alexander Levine

Shenshen Wang, Committee Chair

University of California, Los Angeles

2021

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

ACKNOWLEDGMENTS

It is a genuine pleasure to express my gratitude to all those who have supported me through my Ph.D. years. To my academic advisor, Dr. Shenshen Wang, your timely advice, meticulous scrutiny, and scientific scope have helped me to reach where I am now. I am grateful for your time and patience in helping me to improve. To my committees, collaborators and others who have also offered their valuable insights and support, including but not limited to Drs. Alex Levine, Giovanni Zocchi, Robijn Bruinsma, Jean Carlson, Arvind Murugan, Andrea Graham, Micaela Martinez, Mercedes Pascual, Sarah Cobey, David Schneider, Qixin He, Eric Jones.

To my parents and all other people in the family, you have always been there for me when I needed you. To my beloved wife, Yujia Peng. I am always delighted and grateful to have you by my side. You have brightened my most difficult times, and given me support and courage without which I wouldn't be able to stand where I am today.

To Drs. Hongjing Lu and Hakwan Lau, you have helped me to gain valuable insights and support. To all my friends, including but not limited to Zheng Kang, Hongda Jiang, Steven Schulz, JD Knotts, Nick Baker, Eric Wang, William Loftus, Bernard Koch, Viyehni Carrigan, Chenlong Li. You have always cheered me up and broadened my horizon. It has been great having you along for the ride.

Chapter 4 of this dissertation is a version of Jones, Sheng, Carlson and Wang (2021). Carlson and Wang developed the study concept and directed the study. All the authors contributed to the study design. Jones and Sheng (equal contribution) conducted the study and wrote the manuscript, with editorial input from all the authors.

Portions of the works in this dissertation were supported by the CSC scholarship; by James S. McDonnell Foundation (Grant No. 220020491); by the National Science Foundation Graduate Research Fellowship under Grant No. 1650114; by the David and Lucile Packard Foundation and the Institute for Collaborative Biotechnologies through grant W911NF-09-0001 from the U.S. Army Research Office; and by funding from the Chancellor's Office of UCLA.

VITA

2010-2014	B. S. in Physics Peking University, Beijing, P.R.C.
2014-2016	M. S. in Physics University of California Los Angeles, Los Angeles, U.S.A
2017	Ph.D. candidate in Physics University of California Los Angeles, Los Angeles, U.S.A

PUBLICATIONS

Sheng, J., Wang, S. (in preparation) Dynamic control of eco-evolutionary adaptation. Manuscript in preparation

Sheng, J., Wang, S. (under review by *iScience*). Coevolutionary transitions emerging from flexible molecular recognition and eco-evolutionary feedback.

Jones*, E., Sheng*, J., Carlson, J., & Wang, S. (2021). Aging-induced fragility of the immune system. *Journal of Theoretical Biology*, 510, 110473. *:co-first authors

Sachdeva, V., Husain, K., Sheng, J., Wang, S., & Murugan, A. (2020). Tuning environmental timescales to evolve and maintain generalists. *Proceedings of the National Academy of Sciences*, 117(23), 12693-12699.

Liu, Y., Sheng, J., Wu, H., He, Q., Cheng, H. C., Shakir, M. I., ... & Duan, X. (2016). High-Current-Density Vertical-Tunneling Transistors from Graphene/Highly Doped Silicon Heterostructures. *Advanced Materials*, 28(21), 4120-4125.

Guo, J., Meng, X., Chen, J., Peng, J., Sheng, J., Li, X., Xu, L., Shi, J., Wang, E. &Jiang, Y. (2014). Real-space imaging of interfacial water with submolecular resolution. *Nature Materials,* 13(2), 184-189.

Chen, J., Guo, J., Meng, X., Peng, J., Sheng, J., Xu, L., Jiang, Y., Li, X.& Wang , E. (2014). An unconventional bilayer ice structure on a NaCl (001) film. *Nature Communications*, 5, 4056.

CHAPTER 1

Introduction

Greek philosopher Heraclitus once said: "the only constant in life is change". Indeed, the immune system evolves across the host's lifetime to protect against the wide array of threats in nature. Understanding and actively modulating the evolution of the immune system is of great importance as it not only advances our knowledge of nature but also helps in human's perpetual fight against disease and aging.

The core of immune system evolution across host life time is the evolution of adaptive immune repertoire, defined as the collection of antigen-specific receptors on the surface of B and T cells. During an antigen (Ag) encounter, a Darwinian affinity maturation (AM) process is initiated for the Ag-cognate B cells ([1]): the B cells that recognize the Ag are activated and go through somatic hypermutations with a mutation rate $\sim 10^6$ times higher than normal tissue cells([2]). The mutant B cells with highest receptor-Ag binding affinity are positively selected to expand, while the rest B cells die by apoptosis. The affinity-matured B cells create antibody (Ab) molecules to neutralize Ag, as well as become Ag-specific long-term memory cells in preparation for future encounter against the same Ag. On the contrary, the T cells do not mutate to improve receptor-Ag binding affinity after they are born from thymus. Nevertheless, like B cells, the Ag-cognate T cells are also selectively expanded during Ag encounter, facilitate Ag clearance, and generate a long-term immune memory ([3]). While the immune system is capable of evolving and dealing with a single Ag, such tasks become more challenging for the immune system when faced with diverse antigenic targets, including mutants of a same pathogen or a wide spectrum of pathogen species.

First, highly mutable pathogens with a mutation rate comparable with somatic hypermutations of B cells pose a difficult target for immune system evolution. Examples include

human immunodeficiency virus (HIV) or hepatitis C virus (HCV) that goes through rapid intra-host diversification and can reach over 10% intra-host nucleotide diversity ([4, 5]). The intra-host pathogen diversification results in coevolution, or the "red-queen dynamics", between the pathogen and adaptive immune system, where the pathogens poise as a moving goalpost to the evolving immune system and constantly escape adaptive immune recognition ([6, 7, 8]). In turn, HIV persists indefinitely in any infected individual, while HCV can be spontaneously cleared in only around 30% of those invaded ([9]).

A prospective solution to counter pathogen diversification is the broadly-neutralizing antibodies, or bnAbs ([10, 11, 12]). BnAbs target the functionally important structures of the virus that are conserved among viral mutants, as mutations of these structure would incur a severe fitness cost ([13, 14]). However, in natural infections, bnAbs hardly emerge until a late (typically more than 2 years) persistent infection has been established ([15, 16, 17]). Rare examples of early bnAbs development during acute infection and subsequent viral clearance mediated by bnAbs have been reported for HCV ([18, 14]), but not HIV. In addition, vaccination to elicit bnAbs-as human modulation of immune system evolution-has also proven challenging ([19, 20, 21]). A major obstacle in bnAbs elicitation is that strain-specific B cells targeting the variable Ag elements outcompetes cross-reactive B cells targeting conserved Ag elements during AM. The conserved pathogen structures, for example the CD4 binding site of HIV, or the E1E2 proteins of HCV, are shielded by highly mutable glycoproteins. The conserved structures are thus less sterically accessible to B cells and are less immunogenic ([22, 23]). As a result, fewer B cells recognize conserved viral structures and their affinity is also harder to improve from mutations. This in turn renders the cross-reactive B cells disadvantageous when competing against their strain-specific peers ([24]).

In answer to this challenge, vaccination with diverse Ag variants have been proposed. This is expected to provide a selective advantage to the cross-reactive B cells as they remain fit when binding different Ag variants. Unfortunately, vaccine trials against HIV or HCV so far have met with failure ([25]), and the only trial with positive protective effect against HIV reports a vaccine efficiency of only 31% ([26]).

Apart from the mutants of the same pathogen, the second major challenge to the evolving

immune system is the wide spectrum of pathogens given the limited resource of immune system. Over the host lifetime, the immune repertoire has to adapt to the environment that the host lives in, where diverse pathogen species are present. As the host ages, it experiences an increasing number of encounters with different pathogens, and pathogen-specific memory T cells are generated to protect the host from future infections of the same pathogen. However, "side effects" of memory accumulation also exist. As only a finite amount of T cells is maintained by homeostasis control ([27, 28]), the inflating number of memory cells take up the survival niche of naive cells. For example, it has been reported in a French cohort that the fraction of memory cells among T cells has increased, while the fraction of naive cells decreased, by more than twofold in people aged 75+ than aged 5-10 ([29]). As naive T cells defend the host against novel pathogens, the immune repertoire becomes more vulnerable to the unencountered pathogen species over time ([30]). In addition, an increased fraction of CD8+ T cells in the elderly caused by memory inflation has been associated with chronic inflammatory response even in the absence of pathogen stimulation ("inflamm-aging", [31]), as both are hallmarks of immune risk phenotype (IRP) that predicts impaired immune function and mortality ([32, 33, 34, 35, 36])

To resolve the challenges mentioned above regarding the evolution of immune system against diverse antigenic targets, several questions remain to be answered: (1): What are the mechanisms and deciding factors behind the distinct coevolution outcomes observed in different pathogen species and different hosts? (2): What makes a viable vaccine design strategy to guide immune system evolution towards bnAbs? (3): Is there a mechanistic connection between memory inflation and the chronic inflammation (CI) in the elderly, i.e., is the later a second side effect of memory inflation? Also, how can human intervention alleviates the side effects of memory accumulation?

To answer these questions, statistical mechanics and quantitative modeling provide a great instrument to decipher the evolution of immunity ([37, 38, 39]). Firstly, statistical-mechanical method allows modeling of evolution of immunity with simplified assumptions of biological realism, and uncovering the underlying mechanisms behind experimental observations. Both statistical mechanics and immunology seek to understand the collective

behavior of a macroscopic system, based on the knowledge of its microscopic components as well as their interactions. The collective strength of an immune response, in particular, depends on the dynamics of individual immune cells as well as the intricate immune-signaling network, which is nonlinear and includes complex feedbacks. While experiments and clinical observations yield information on the collective immune response, the cellular- or even molecular-level dynamics are difficult to access in practice. As such, similar to how statistical physicists abstracted real gas as ideal gas and successfully provided a mechanistic explanation of the gas laws, in immunology physicists can shine with mechanistic modeling ([40]) and focus on the rules governing the collective behavior of immune response. Secondly, quantitative models built by theorists can serve as "virtual experiments", and their predictions help experimentalists to invest resources on the most efficient experiment designs and most promising treatment protocols. This is crucial as real-life experiments are limited by the budget of both money and time, as well as by the availability of experimental techniques. For example, a typical HIV vaccine trial would span around 10 years and cost over 100 million dollars ([41, 42, 43]). Notable successes of quantitative modeling in immunology include the theoretical inspiration of "triple cocktail" therapy against HIV ([44]), the prediction of phasic B cell mutation rate during AM ([45]), and the explanation of reduced influenza vaccine effectiveness on second-time vaccinees than first-time vaccinees ([46]), to name just a few—many of which are predictions preceding experimental verifications.

My dissertation aims to address the unanswered questions above on immune system's evolution, from a statistical-mechanical and computational perspective. Specifically, chapter 2 investigates the intra-host coevolution of immune system and highly-mutable pathogens. Existing work on this topic often assumes constant population sizes of both immune cells and Ag ([24, 47]). Such assumption is common in population genetics ([48, 49, 50, 51, 52, 53]) and reasonably approximates the chronic stage of infections. However, it also explicitly precludes alternative outcomes other than viral persistence, such as the rare yet clinically-observed cases of spontaneous HCV clearance ([18, 14]). In addition, the assumption of constant population size neglects potential feedback between ecological dynamics and evolution in a host. Such feedback arises because B cells not only sample and follow the antigenic

environment, but also govern the absolute and relative fitness of distinct viral strains. Consequently, antibodies arising early and becoming dominant quickly modify the shared "Ag resource" that itself adapts, thereby influencing what future antibodies can evolve. Lastly, existing approaches rarely explore the possibility that physical modulation of receptor-Ag binding (e.g. locating best complementarity in real space) may strongly impact the efficacy of Ab repertoire. In this dissertation, I have established a computational framework of agent-based stochastic simulations that incorporate eco-evo dynamics of B cells and Ag, as well as the flexible Ab-Ag binding footprint ([54]). I have identified three possible co-evolution outcomes which can be compared with experimental observations: Ag clearance, persistence and rebound. The transition between different outcomes is decided by the timing and efficacy of successive narrow and broad Ab responses, which in turn are determined by the conservation level and initial diversity of Ag. In addition, clearance of structurally complex Ag relies on Ab evolution in a larger antigenic space than where selection directly acts; rebound highlights the impact of feedback between ecology and rapid evolution. Finally, immune compartmentalization can slow viral escape but may also delay clearance.

Chapter 3 studies the vaccination protocol to optimally elicit bnAbs against highly-mutable pathogens. This goal is incorporated in a broader context as the *dynamic* control of eco-evolutionary adaptation, and the competition between strain-specific and cross-reactive B cells during their germinal center (GC) evolution is studied in detail. Specifically, existing vaccination protocols involve multiple injections ("prime-boost" scheme,[55]) well separated in time (typically by months, [56, 57, 58]), and after each injection a new set of B cells, be it naive or memory, are recruited and initialize subsequent GC reactions. Correspondingly, previous modeling studies assumed a constant antigenic environment during each GC reaction ([59, 60]). However, experimental evidence has shown that bnAbs precursors start with a low breadth when newly recruited to GCs, and their breadth gradually improves during AM in GCs ([61, 62]). As such, a time-dependent antigenic environment during a continuous GC reaction can provide a dynamic selection pressure tailored for concurrent B cells, and a better outcome than constant antigenic environment can thus be expected. In addition, the continuous delivery of vaccine Ag during a single germinal reaction has been demonstrated

recently via osmotic pump technique ([63, 64]), so the clinical realization of the new dynamic vaccine protocol is expected to be feasible. In this dissertation, by quantifying B cell dynamics in GCs with ordinary differential equations, I have identified a trade-off between suppressing strain-specific B cells and preserving cross-reactive B cells at constant selection strength. Time-varying selection outperforms constant selection by fulfilling both ends of the trade-off sequentially in time. In addition, by applying optimal control theory, I demonstrate that the optimal selection strength increases in time as driven by the evolution of B cell cross-reactivity. Finally, suppressing strain-specific B cells with time-varying selection promotes both quantity and quality of cross-reactive B cells, saving the need for two distinct controls.

Chapter 4 looks into immunosenescence–the aging of immune system–with a particular focus on repertoire overspecialization and its implications. On one hand, in an earlier mathematical model of the adaptive immune response, Stromberg and Carlson demonstrated that repeated pathogen exposures could lead to an imbalanced immune repertoire that was fragile to rare pathogens, in the sense that rare pathogens proliferated significantly more than common pathogens ([30]). On the other hand, Reynolds et al. developed a model to describe the innate immune response immediately following an acute pathogen encounter ([65]). In addition, by considering both innate and adaptive immunity, Shi et al demonstrated through modeling the enhanced pathogen clearance during late stage of sepsis progression ([66]). Nevertheless, the effect of an *evolving* adaptive repertoire on innate immunity remains unexplored, and a mechanistic connection between the aging adaptive repertoire and CI is yet to be established. In the dissertation, I built an ODE model of coupled dynamics between innate and adaptive immune branches, which explicitly considers the adaptive suppression of innate immunity through anti-inflammtory cytokines such as IL-10 ([67]), or through inhibition of inflammasomes NLRP1, NLRP3 ([68]). I identify three steady states following a single pathogen encounter: health, septic death, and chronic inflammation. Mediated by adaptive-innate feedback, repeated pathogen encounters during host lifetime and resulting memory inflation may trigger a fragility in which any encounter with a novel pathogen will cause the system to irreversibly switch from health to CI. This transition is consistent with

the onset of "inflammaging", the condition observed in aged individuals who experience chronic low-grade inflammation even in the absence of pathogens. The model predicts that the onset of CI strongly depends on the history of encountered pathogens; the timing of onset can be delayed drastically when the same set of infections are encountered in a specific order. Lastly, the coupling between the innate and adaptive immune branches generates a trade-off between rapid pathogen clearance and a delayed onset of immunosenescence.

CHAPTER 2

Coevolutionary transitions emerging from flexible molecular recognition and eco-evolutionary feedback

2.1 Abstract

Highly mutable viruses evolve to evade host immunity that exerts selective pressure and adapts to viral dynamics. How this dynamic reciprocal interaction drives diverging trajectories and outcomes of evolution is not well understood, which limits our ability to mitigate viral evolution and accelerate immune control. Here we provide a framework for identifying key determinants of the mode and fate of viral-immune coevolution, by linking molecular recognition and eco-evolutionary dynamics. We find that conservation level and initial diversity of Ag jointly determine the timing and efficacy of successive narrow and broad Ab responses, which in turn control the transition between viral persistence, clearance and rebound. In particular, clearance of structurally complex Ag relies on Ab evolution in a larger antigenic space than where selection directly acts; rebound highlights the impact of feedback between ecology and rapid evolution. Finally, immune compartmentalization can slow viral escape but may also delay clearance. This work suggests that flexible molecular binding allows a plastic phenotype, which modulates evolutionary transitions of co-adapting populations. By assisting in the accumulation of potentiating variations, it is possible to shortcut long paths toward highly adapted states.

2.2 Introduction

Among the many viruses that inhabit every species on the planet, some evolve as fast as the adaptive immunity of their host. These rapidly evolving intruders share unusual characteristics [69]: They evade immune detection by putting on diverse and variable disguises (e.g. via high genetic variability) and divert immune focus by presenting multiple competing antigenic targets.

Despite strong similarity in evasion tactics employed, courses and outcomes of evolution may differ markedly among viruses. A dramatic example is that, HIV persists indefinitely in any infected individual, while HCV can be spontaneously cleared in around 30% of those invaded [9], even though the genetic diversity of HCV in a chronically infected person is comparable to or higher than that of HIV [4, 70, 71, 72, 73, 74, 5, 75]. In addition, bnAbs — capable of recognizing a vast array of mutant strains — evolve against both viruses [76, 77], presenting a potential solution to counter rapid viral evolution. Yet, HIV bnAbs often emerge years into infection and confer little protection [78, 79], whereas HCV bnAbs are found to arise earlier and contribute to viral clearance [80, 81]. Such variability in the path and fate suggests that transitions between distinct regimes may be controlled and that long routes of immune adaptation could be shortened, if key determinants of possible outcomes were identified.

Viral dynamics and outcomes cannot be understood without considering host immunity. B cells constitute the evolving branch of the adaptive immune system and produce a dynamic Ab repertoire. They undergo an accelerated evolutionary process named affinity maturation [82] in microenvironments called germinal centers [83]; therein cycles of competition, proliferation and somatic hypermutations ($\sim 10^6$ fold faster than mutations in normal tissue cells [2]) result in increasingly stronger binders to the recognized Ag. In contrast to a passive antigenic agent (e.g. vaccine) that merely sustains GC reactions, mutating viruses respond to AM and engage lymphocytes in mutual selection (Fig. 2.1A): B cells are selected for enhanced binding to Ags sampled into GCs from circulation, while viruses are selected for reduced recognition by circulating Abs (secreted B cell receptors, BCRs) output from

all GCs combined. An enduring co-adaptation relies on genetic novelty in virus and Ab populations, generated and maintained by rapid mutation at comparable rates. Driven by stochastic mutation and selection, an ensemble of coevolving trajectories form in an infected individual.

Important advances are being made to trace viral-immune coevolution during natural infection. Measurements with ever higher throughput and resolution provide a wealth of information across scales, ranging from the origin and pattern of phylogeny [84, 85, 86, 87, 88], time series of functional traits (e.g. Ab potency and breadth) [89, 54, 90, 91] and GC dynamics [92], to structure and key mutations at and near the binding interface between an Ag epitope and an Ab paratope [93, 94, 76]. Three chief observations specifically motivate our model framework. First, not every part of the B cell repertoire is equally accessible; in particular, bnAbs tend to have significantly lower germline frequencies than strain-specific Abs [95], reflecting the fact that conserved residues in the epitope (targets of bnAbs) are fewer and harder to access than variable elements [96, 97]. Second, mutational paths toward bnAbs vary in length. Although most bnAbs discovered so far accumulate a large number of genetic alterations [98], some broad Abs effective at controlling infections are generated with relatively few somatic mutations [99, 100]. This raises the question of how much AM is needed for bnAbs to emerge and what distinguish short paths from long ones. Third, clonal diversity and composition vary widely among GCs [92], suggesting that spatial segregation of B cells may serve a functional role in the face of moving targets.

Mathematical models have long shed light on the competitive dynamics of AM with a single Ag [101, 102, 103, 104, 105, 106], with quantitative insight gained by supplementing modeling with parameter inference [107]. Recent works have reported features of Ab evolution against multiple related Ags, with a particular interest in how Ag characteristics (number, dose, complexity) and temporal patterns (in series or combination) impact the chance of evolving cross-reactive Abs [108, 109, 110, 111, 112, 113]. This line of research is primarily motivated by the discovery of bnAbs against a variety of highly mutable pathogens, such as HIV, HCV, influenza, and malaria, all lacking an effective universal vaccine. In these *in silico* vaccination studies, Ags are preset and do not respond to AM.

Comparatively fewer studies have considered host-pathogen coevolution [114, 115, 116, 47]; these works often assume constant size of both populations (with recent exceptions [117, 118, 119, 120]). This assumption is typical in theoretical evolutionary biology and presumably suitable for the chronic stage of infections, and yet it precludes outcomes other than viral persistence, and neglects potential feedback between ecological dynamics and evolution *in a host*. Such feedback arises because B cells not only sample and follow the antigenic environment, but also govern the absolute and relative fitness of distinct viral strains. Consequently, Abs arising early and becoming dominant quickly modify the shared "Ag resource" that itself adapts, thereby influencing what future Abs can evolve. Furthermore, existing approaches rarely explore the possibility that physical dynamics of receptor-Ag binding (e.g. locating best complementarity) may strongly impact the efficacy of Ab repertoire.

Here, we present a computational framework of coevolution that addresses these gaps and accounts for notable observations, emphasizing how physical dynamics on the molecular level give rise to evolutionary transitions on much larger scales. We consider Ab cross-reactivity and flexible binding footprint [121, 54], variation in adaptation rate along different lineages, as well as feedback to and from population dynamics. Instead of modeling particular viruses, we predict conditions under which diverse evolutionary trends emerge, and determine ways in which features of trajectories and diversity patterns at early times inform outcomes.

Main findings are as follows. First, we identify the conservation level and initial diversity of Ag as key determinants of evolutionary outcomes. Second, we find that BCR mutations outside the direct contact region can potentiate Ab footprint shift toward sites of viral vulnerability, thus enabling clearance. Such flexibility of molecular recognition expands the search space and opens new paths of accelerated adaptation. Third, we show that functional consequences of GC compartmentalization depends on Ag variability. Our results stress the importance to consider adaptive dynamics of "Ag resource" when evaluating the viability and efficiency of immune control. We discuss how this understanding suggests new ways to enhance immune efficacy and shortcut long routes to desired outcomes.

2.3 Results

2.3.1 Coevolution model

We develop a computational model of B cell AM driven by a coevolving viral population (Fig. 2.1A). To study the effect of population subdivision on collective dynamics, we consider segregation of B cells into a variable number of GCs, independently seeded by germline clones and evolving in parallel. Ab-secreting cells output from all GCs combined accumulate in circulation and remove viruses they encounter and match. Meanwhile, circulating viruses replicate and diversify; those that, by chance, either avoid matching Abs or acquire escape mutations grow in number. A random subset of Ag is transported into each GC, presented to B cells, and fuels further rounds of AM. B cells that bind and internalize more Ag compete better for limited T cell help, while the losers apoptose. Surviving cells either recycle to mature more or differentiate and exit GCs, thereby closing the loop of reactions between B cell and Ag populations.

We simulate the stochastic processes during AM (see Methods for steps), based on rules and parameters derived from experimental studies of GC reactions [122, 123, 124, 92] (see Appendices). Our model is a coarse-grained one that leaves out migration of B cells within and between GCs [125, 126] and abstracts the molecular contexts of receptor-Ag binding. Nonetheless, this simplification allows us to focus on a few novel features that shape evolutionary dynamics and outcomes in essential ways.

2.3.1.1 Phenotypic description of BCR-Ag interaction and evolution

To describe joint dynamics and feedback of B cell and Ag populations, we extend the classic shape-space model by Perelson and Oster [127] to consider evolving phenotypic distributions. BCRs and Ags are points in a common n-dimensional Euclidean vector space named shape space. Dimensions represent groups of amino acids comprising the binding interface between an Ag epitope and a BCR paratope; contiguous dimensions correspond to spatially proximal residue groups. Coordinates describe biochemical properties relevant for binding affinity.

This geometric representation captures specificity of BCR-Ag interaction: a small distance depicts high complementarity hence strong binding, whereas a large separation reflects poor match thus weak binding.

Residues in a viral eptiope vary in accessibility to BCR: conserved elements essential for viral fitness and function are often surrounded and partially masked by highly variable residues (and covered by glycans) [96, 76]. Consequently, strain-specific B cells that target easy-to-access variable residues evolve to enhance affinity at a fast pace, whereas cross-reactive B cells directed at poorly accessible conserved residues tend to improve slowly. In this sense, spatial accessibility of binding targets dictates adaptation rate of elicited clonal lineages.

Guided by these observations, we describe the variability of adaptation rate in a phenomenological manner: B cells move fast in dimensions along which Ags can mutate and evade recognition (n_v fast/variable dimensions), but move slowly in dimensions where Ags are fixed at the origin ($n_c = n - n_v$ slow/conserved dimensions). The corresponding difference in mutation effect is encoded by a difference in the jump step size in shape space (see Appendices). As illustrated in Fig. 2.1B, germline B cells (blue ovals), starting on or outside a founder hypersphere centered at the origin (where the infecting virus resides), evolve to approach Ags (red stars) in all directions, while viruses can only move along fast/variable dimensions, attempting to escape the chase by facing it.

2.3.1.2 Binding affinity and footprint shift

Structural studies [128, 84, 121, 54, 129] indicate that not all the residues constituting a viral Ag epitope are in contact with a *particular* BCR paratope. Rather, BCRs specific for the same epitope may bind different, yet partially overlapping, portions of it, due to differences in BCR conformation or approach angle. This joint binding surface on the Ag defines the epitope. Hence, the footprint of a particular BCR spans only a fraction of the epitope; in our shape space, BCR-Ag binding takes place in a subspace. Further, once in proximity, a BCR may search across the epitope surface for best complementarity and settle therein once

located, as commonly seen in molecular recognition [130]. This translates to comparing the Ag-BCR Euclidean distance among all subspaces (comprising n_b contiguous dimensions) and taking the minimum; this operation picks out the best match from $(n - n_b + 1)$ potential binding sites.

To be specific, we define the local binding affinity A_l (in units of $k_{\mathrm{B}}T$) between a BCR $\vec{x}$ and an Ag $\vec{y}$ as follows

$$A_l(\vec{x}, \vec{y}) \equiv A_{\max} - \frac{1}{n_b}||\vec{x} - \vec{y}||_l^2. \tag{2.1}$$

Here $l \in \{1, 2, \cdots, n - n_b + 1\}$ indexes the starting dimension of a n_b-dimensional binding subspace (i.e. the l-th binding footprint); $A_{\max}$ denotes the maximum affinity at perfect match. The global affinity is identified as

$$A(\vec{x}, \vec{y}) \equiv \max_l A_l(\vec{x}, \vec{y}) = A_{\max} - \frac{1}{n_b} \min_l ||\vec{x} - \vec{y}||_l^2. \tag{2.2}$$

Hence higher affinity indicates smaller mismatch in conformation; minimization of mismatch reflects a dynamic search for the optimal binding footprint labelled by $l^*(\vec{x}, \vec{y}) \equiv \mathrm{argmax}_l A_l(\vec{x}, \vec{y})$.

This notion of binding subspace endows Abs with desired properties, such as flexibility in binding target and state-dependent effect of mutations (i.e. epistasis). Importantly, this representation captures the observed shift of Ab/BCR binding footprint on viral epitope during HIV-Ab coevolution [54, 121], which correlates the precision of targeting to conserved residues with breadth development. As shown schematically in Fig. 2.2A, as mutations induce conformational changes in BCR (purple shape) and/or the variable region of Ag (green shape), the optimal binding footprint (yellow interface) may shift toward increasingly conserved regions of the viral epitope (red shape), under conditions to be discussed below.

Footprint shift influences the distribution of mutation effect $P(\Delta A)$ in two ways: First, when a deleterious mutation ($\Delta A_{l^*} < 0$) occurs *inside* current optimal subspace l^* and lowers the affinity A_{l^*}, switch to a different subspace in which $A_l > A_{l^*} + \Delta A_{l^*}$ will buffer against large detrimental effect. Second, when a beneficial mutation ($\Delta A_l > 0$) occurs *outside* current optimal subspace ($l \neq l^*$) such that $A_l + \Delta A_l > A_{l^*}$, switch into this subspace will enhance affinity; without footprint shift, this apparently "neutral" mutation would be

wasted. Thus, as shown in Fig. 2.2B for founder B cells and the infecting virus (contrasting blue and yellow histograms), shift of binding footprint leads to fewer and smaller deleterious mutations (a shortened negative tail), along with more frequent beneficial mutations of larger sizes (an expanded positive wing). This turns out to be key to speeding up B cell adaptation and enabling viral clearance.

Based on the conservation level of the optimal binding subspace, we classify GC B cells and Ab-secreting plasma cells into B (broad) type and N (narrow) type, which, respectively, have greater than and at most $n_b - 1$ conserved dimensions when averaged over encountered Ags.

2.3.1.3 Ecological dynamics influence mutual selection

A GC B cell $\vec{x}$ is able to successfully acquire Ag in GC cycle t with the following probability

$$P_G(\vec{x}, t) = \frac{C_{Ag}(t) \sum_{\vec{y} \in \mathcal{Y}} e^{A(\vec{x},\vec{y})}}{1 + C_{Ag}(t) \sum_{\vec{y} \in \mathcal{Y}} e^{A(\vec{x},\vec{y})}}, \tag{2.3}$$

where $C_{Ag}(t)$ is the concentration of Ag presented on follicular dendritic cells (FDCs) in a GC, assumed to be proportional to the virus population size at time t in circulation; this neglects the time lag due to Ag transport and presentation. $\mathcal{Y}$ denotes the set of FDC Ags scanned by B cell $\vec{x}$ in cycle t. Competition for limited T cell help is incorporated by keeping the top 70% potent B cells in each cycle [111]. Conversely, a virus $\vec{y}$ in circulation is neutralized by Abs secreted by plasma cells it has encountered in generation t with the following probability

$$P_V(\vec{y}, t) = \frac{C_{pla}(t) \sum_{\vec{x} \in \mathcal{X}} e^{A(\vec{x},\vec{y})}}{1 + C_{pla}(t) \sum_{\vec{x} \in \mathcal{X}} e^{A(\vec{x},\vec{y})}}. \tag{2.4}$$

Here, $C_{pla}(t)$ represents the concentration of plasma cells in cycle t, among which a random subset $\mathcal{X}$ encounter virus $\vec{y}$. Therefore, through time-dependent concentrations of binding substrates (Ags and plasma cells), population dynamics directly influence two-way selection pressure, which, in turn, drives changes in population size and composition, thereby creating a feedback loop between ecology and evolution.

A central feature of our shape space is that both the virus and B cell populations evolve in a larger state space than where selection acts (i.e. $n_v, n > n_b$). In particular, while BCRs

mutate in the full space (n dimensions), mutual selection is only based on binding affinity within a subspace (n_b dimensions). Hence, B cells can enhance receptor potency without increasing selective pressure on the virus. This, therefore, allows mutations that are not immediately beneficial but may potentiate future adaptation to accumulate more efficiently than if all dimensions contribute to binding.

2.3.2 Antigenic determinants for coevolutionary outcomes

2.3.2.1 Three phases

Infections are founded by viral strains centered at the origin of our shape space (with a Gaussian distribution); germline B cells are isotropically distributed on a founder hypersphere in their respective optimal binding subspace. Following estimates based on immunological data [131], we choose the shape-space dimensionality to be $n = 8$, in which $n_b = 3$ dimensions constitute each binding subspace, corresponding to the portion of engaged residues among those comprising the epitope (see Appendices). In the studied parameter range, a total time span of 200 GC cycles is sufficient to distinguish outcomes.

Our model produces dynamics and phenotypic patterns consistent with natural infection of highly mutable viruses (Fig. 2.3). Stochastic simulations identify three distinct outcomes observed in nature — viral/Ag persistence, rebound and clearance — at different combinations of two control parameters, the number of conserved dimensions, n_c, and phenotypic diversity of Ag when they first activate immune responses, σ_A. A phase diagram delineates dynamical regimes (Fig. 2.3A): Viral persistence (blue), as commonly observed for rapidly mutating pathogens (e.g. HIV and HCV), occurs at low epitope conservation (small n_c, $n_c < n_b$) and modest initial Ag diversity (small σ_A); sufficiently conserved epitopes ($n_c \geq n_b$) combined with moderate Ag diversity at immune activation result in viral clearance (magenta), often associated with acute infections (e.g. influenza and Ebola). More unexpected is the rebound phase (grey) at large initial Ag diversity. Although viral persistence occurs in the majority of chronic infections, a recent study of HCV-Ab coevolution has identified clearance and rebound in different individuals; in both cases bnAbs emerge [80].

Figs. 2.3B and 2.3C present dynamics of virus and plasma cell populations, respectively, demonstrating defining features of three phases. Figs. 2.3G-I display characteristic phenotypic patterns developed in the shape space (at time points marked by black dots in Figs. 2.3D-F), distinguishing outcomes since early times. In the persistence phase (Fig. 2.3D), sustained oscillations in Ag abundance (red), characteristic of predator-prey cycles, are accompanied by a steady accumulation of plasma cells (blue) from GC output. The viral-immune mutual engagement in a Red-Queen state is made vivid through density profiles at several instants during oscillations (Fig. 2.3G): in a variable subspace, a deforming cloud of plasma cells (blue) is chasing after Ag clusters (red), which escape increasingly farther from the infecting strain at the origin that focuses past B cell response (Fig. 2.11).

In the rebound phase, an initially very diverse viral population rapidly falls to a low abundance (below 10% of the initial size or capacity, but above the extinction threshold at 0.1% capacity) and subsequently recovers, forming a population bottleneck; the resulting lack of Ag stimulation for GC reaction yields a plateau in plasma cell counts (Fig. 2.3E). Notably, during the Ag bottleneck, the plasma cell distribution resembles the founder hypersphere, even after Ags reemerge outside the enclosure (Fig. 2.3H bottom panel), indicating very little AM prior to viral rebound. Only after escape mutants build up in number, GC reaction and plasma cell differentiation resume. Note that, even after recovering from the bottleneck, Ags may still be cleared, if the epitope is sufficiently conserved. This stage proceeds in a similar manner as in the clearance phase (Figs. 2.3F and 2.3I). Starting from a small variance, diversifying Ags drive rounds of AM through which B cells evolve from targeting variable residues (Fig. 2.3I top panel) to recognizing a mixed region (middle panel, condensing at opposing poles of the founder sphere), and finally focusing onto conserved residues (bottom panel, high concentration at the origin of a fully conserved subspace). This shift in binding target toward viral vulnerability results in a monotonic decline of Ag population to below the extinction threshold (Fig. 2.3F, red curve).

2.3.2.2 Characterizing broad and narrow responses

To understand the emergence of diverse trajectories and patterns, we examine how n_c and σ_A jointly control the transition between distinct regimes (Fig. 2.4). While n_c governs the prevalence of germline B cells targeting conserved and variable residues (i.e. precusor frequency of broad and narrow clones), σ_A decides their relative accessibility (phenotypic distance) to Ags. As to be shown below, these parameters control the efficacy and timing of strain-specific and cross-reactive responses, which, in turn, determine evolutionary outcomes.

To characterize these components, we divide the set $\mathcal{X}$ of plasma BCRs into two subsets $\mathcal{X}_B$ and $\mathcal{X}_N$, based on the number of conserved dimensions $c(\vec{x},\vec{y})$ in the optimal binding subspace between a BCR $\vec{x}$ and a given Ag $\vec{y}$: $\mathcal{X}_B \equiv \{\vec{x} \in \mathcal{X} | c(\vec{x},\vec{y}) = n_b\}; \mathcal{X}_N \equiv \mathcal{X} \backslash \mathcal{X}_B$. Thus, for any BCR $\vec{x} \in \mathcal{X}_B$, it binds Ag $\vec{y}$ mostly strongly in a fully conserved subspace; $\mathcal{X}_B$ is empty if $n_c < n_b$, since all subspaces contain at least one variable dimension. Accordingly, we can rewrite the probability of Ag removal (Eq. 2.4) as $P_V(\vec{y},t) \equiv P_B(\vec{y},t) + P_N(\vec{y},t)$, where $P_S(\vec{y},t) = C_{pla}(t) \sum_{\vec{x}\in\mathcal{X}_S} \exp[A(\vec{x},\vec{y})] / \{1 + C_{pla}(t) \sum_{\vec{x}\in\mathcal{X}} \exp[A(\vec{x},\vec{y})]\}$, with $S = B, N$. Therefore, P_B and P_N, when averaged over circulating Ags $\{\vec{y}\}$, characterize the efficacy of broad and narrow lineages, respectively.

Fig. 2.4 shows typical trajectories of Ag population size (black) and concomitant strength of B-type (P_B; color, solid) and N-type (P_N; color, dashed) plasma cells for different pairs of n_c and σ_A (columns and rows). In cases where both types exist ($n_c \geq n_b$, right two columns) narrow clones arise first and suppress Ag population; broad clones only emerge and expand as narrow lineages decline in efficacy, signifying viral escape. Hence, the peak in P_N indicates a shift in dominance from strain-specific to cross-reactive response.

2.3.2.3 Initial Ag diversity governs clearance-to-rebound transition

A greater Ag diversity at GC onset (larger σ_A) reflects a longer lag between infection and activation of first responsive clones in the B cell repertoire. In our shape space, compared to the founder virus at the origin, a mutated Ag binds more weakly to some B cells but more strongly to others, as long as it remains enclosed by B cells. Similar behavior has

been observed in experiment where the same Ag mutations weaken affinity to certain Abs but enhance binding to others [132, 129, 133], consistent with Ags being corralled by B cells early in infection.

Less intuitively, increasing Ag diversity results in higher mean affinity of strain-specific (N-type) germline B cells without affecting cross-reactive (B-type) cells (see Appendices and Fig. 2.10), leading to faster Ag removal by specific clones at early times. This implies that Ag diversity at the response onset (judged by the start of decline in Ag population) controls the access of narrow and broad clones to Ag "resource", thereby affecting the rate of Ag consumption. Indeed, as σ_A increases, a faster fall in Ag population follows an earlier and steeper rise in P_N (Fig. 2.4 right column, bottom to top), but this also leads to weaker P_B upon viral escape, since broad clones have shorter time to evolve. Thus, through feedback between population dynamics and mutual selection, initial Ag diversity has a complex influence on B cell responses; not only does it affect the efficacy of narrow and broad clones at their peak time, it also tunes the timing of Ag escape and the ensuing expansion of broad lineages that ultimately determine evolutionary outcomes.

Such eco-evolutionary feedback manifests most dramatically as a rapid viral rebound following a deep population bottleneck, when initial Ag diversity is large (Fig. 2.4A). A modest Ag diversity, instead, results in monotonic clearance. These phases appear to capture distinct kinetics of clearance and rebound in HCV-infected individuals [80] and suggest Ag diversity at response onset as a predictor of outcomes. The key lies in the level of AM reached before Ag population falls too small to sustain GC reaction. Specifically, if Ag diversity starts modest, both narrow and broad clones begin with weak affinity and evolve toward viruses at similar antigenic distance. By the time narrow clones start to wane while broad clones are just on the rise, Ags are still relatively abundant. Consequently, broad lineages continue to gain both abundance and breadth as diversifying Ags favor their selection. This, in turn, increasingly limits viral growth, yielding a monotonic drop to extinction. Throughout this process, Ags remain inside B cell enclosure (Fig. 2.12); as the Ag cloud is consumed from the surface inward, a density gradient is generated, providing an "attractant" field that guides B cells toward the founder virus.

In contrast, at large initial diversity, "pioneer" viruses near the frontier of the Ag cloud are close to the B cell founder hypersphere; while being able to escape B cell enclosure with few mutations, these mutants risk being recognized by strain-specific clones as they transect across the frontline. As a result, Ag removal is so rapid that B cells have hardly evolved when escape mutants emerge outside the founder hypersphere (almost unchanged B cell distribution, Fig. 2.3H and Fig. 2.13; rapid rise and fall of P_N, Fig. 2.4A). Meanwhile, Ag population is already too small to sustain GC reaction and AM essentially comes to a halt (Fig. 2.4A right panel, flat P_B due to lack of AM). This thus allows escape mutants to expand unchecked, leading to a significant rebound. With a delay, this renewed supply of mutated Ag boosts AM and selects broad lineages that evolve to acquire breadth and eventually clear the virus (Fig. 2.14).

2.3.2.4 Epitope conservation determines the timing of viral rebound

When conserved residues make up a larger portion of the epitope (increasing n_c; Fig. 2.4, left to right), narrow response rises more slowly and reach a lower maximum, while broad response, uncompromised by viral escape, gains stronger dominance toward viral clearance ($n_c \geq n_b$). In the absence of fully conserved targets, infections persist ($n_c < n_b$, left column). At intermediate values of σ_A (Fig. 2.4B), dynamics and outcomes exhibit a non-monotonic trend as n_c increases: monotonic viral clearance occurs either when N-type clones are sufficiently potent early on so they remove Ags before escape mutants arise (left panel, $n_c < n_b$), or when B-type lineages have time to mature and clear Ags following their escape from narrow lineages (right panel, $n_c > n_b$). In between ($n_c \sim n_b$), however, neither narrow nor broad clones are effective at clearing Ags but they suppress each other; rapid Ag removal by specific clones leaves little time and stimuli to support breadth development, thus, viral rebound ensues.

Interestingly, increasing n_c leads to a shallower Ag population bottleneck and yet a slower rebound (Fig. 2.4A, left to right; Fig. 2.5, yellow to green). This behavior reflects a separation of timescales between the dominance of clones with different binding targets: weaker binding

to variable residues early on reduces Ag removal thus yielding a shallower bottleneck, whereas stronger binding to conserved residues later delays Ag revival; two stages are separated by Ag escape before which little AM has taken place. Note, for a given n_c, just as one would expect, a deeper bottleneck indeed lasts longer (Fig. 2.5, same color, going up and to the left). These results indicate that the conservation level of the viral epitope determines the size and duration of Ag population bottleneck, and hence the timing of viral rebound.

2.3.3 Flexible molecular recognition drives viral extinction and bnAb development

2.3.3.1 Footprint shift enables viral clearance

The clearance regime *per se* is not a surprise, if immune adaptation can outpace viral evolution via faster or larger mutations. It becomes more surprising as we rule out this possibility by choosing mutation parameters that support stable persistence. In fact, predator-prey cycles would be the only outcome even for a relatively conserved epitope, if binding footprint were fixed (Figs. 2.6A and 2.15). Ab footprint shift toward conserved viral residues has been observed during HIV-bnAb development [54, 121], but how this occurs through coevolution is unclear. Fig. 2.6B demonstrates that footprint shift retains descendants of broad germline clones with fully conserved targets (blue bar), such that they can persist to make up a substantial fraction of GC populations and clear the virus (Fig. 2.6D, blue band). In contrast, without a flexible footprint (Fig. 2.6C), broad clones are quickly outcompeted by strain-specific ones and lost once for all. Note, even in this case, despite that highly specific clones (red band) dominate early on, lineages with relatively cross-reactive ancestors (purple band) steadily grow and dominate at later times as Ags diversify. But, because truly broad clones are permanently lost, GC populations only chase after escape viral mutants, leaving an oscillatory signature in Ag population trajectories (Fig. 2.6A, black curve). Therefore, Ab footprint shift enhances Ag removal by preserving broad precusors and promoting expansion of cross-reactive descendants.

2.3.3.2 Footprint shift opens novel pathways toward broad response

One might wonder, how can slowly evolving cross-reactive clones persist, in the face of strain-specific competitors that are more numerous, potent and faster evolving? Our model suggests an intriguing possibility: Physical optimization of binding — via adjusting BCR footprint — not merely speeds up adaptation of intrinsically slow clones, but allows for a plastic phenotype along a lineage. That is, shift between binding targets of different conservation levels amounts to switch between states with different environmental sensitivity. Therefore, even if outcompeted early on, cross-reactive (B-type) clones can "regenerate" from strain-specific (N-type) ancestors, enabling sustained adaptation to diversifying Ags.

To quantify this intuition, we track the phenotypic identity of surviving B cells over the course of coevolution. Fig. 2.7B presents the time-dependent ensemble-averaged composition based on initial identity X (against founder virus) and current identity Y (against FDC Ags encountered in current GC cycle) of individual cells; hence X→Y represents current Y-type clones descending from an X-type germline precusor — one of four types of identity switch (B→B, N→B, B→N, N→N). Over a wide range of initial Ag diversity, a common pattern emerges: shortly after the response begins, B→B (blue) drops rapidly while B→N (red) and N→N (coffee) grow; subsequently, N→B (green) and B→B (blue) rise to dominance, while B→N and N→N fall minor. Accordingly, footprint shift opens multiple coexisting novel pathways toward broad clones, with N→N→B and B→N→B being more prevalent than B→B→B.

While cross-reactivity can be good for evolvability (capacity to adapt), it takes time and the right conditions to itself evolve. These unexpected pathways suggest that spontaneous (but not random) phenotype switch of individual cells — enabled by physical dynamics — allows a cell population to fulfill conflicting demands through a separation of timescales: switch to N-type early on meets the short-term need for GC survival under severe selection pressure, and switch to B-type later sustains long-term adaptation toward greater breadth and potency. Both switches are achieved by target shift that confers largest affinity gain available. This is further supported by time-dependent distributions of mutation effect in fast

(variable) and slow (conserved) dimensions (Fig. 2.16): early in response, a larger number of beneficial mutations of bigger size exist in fast dimensions, hence switch to N-type is favorable; at later times, adaptive changes of appreciable size continue to be available in slow dimensions, while mutations in fast dimensions can no longer improve affinity, thus, switch to B-type is selected. Consistently, crossing of mean affinities of B- and N-type cells occurs right at the start of N→B switch (Fig. 2.7A, vertical line). Therefore, evolution drives footprint shift and conversely, footprint shift opens new pathways and facilitates adaptation.

Interestingly, initial Ag diversity influences the relative abundance of pathways enabled by footprint shift (Figs. 2.7, 2.17 and 2.18). At small σ_A, N→B pathway dominates (Figs. 2.7B and 2.17 left, green bar), whereas large σ_A favors B→N→B pathway (Figs. 2.7B and 2.17 right, blue bar), indicating higher initial Ag diversity better preserves broad precusors. Both pathways may contribute to the experimental observation that a fraction of strain-specific precusors later gain breadth [84, 134, 135, 133]. In practice, B→N→B might be observed as N→B due to the brief presence of broad precusors. Among pathways that both start and end with broad clones, B→N→B is more abundant than B→B→B, in terms of the fraction of cells in a single metapopulation (one simulation of subdivided GCs; see statistics in Figs. 2.7C-D) and the fraction of metapopulations containing either pathway (Fig. 2.7E).

2.3.3.3 Footprint shift exploits potentiating variations

It is apparent from Fig. 2.7C that increasing initial Ag diversity σ_A (purple to red) accelerates the expansion of BNB lineages in GC populations, leading to faster viral clearance. To address how σ_A affects the pace and efficacy of the BNB pathway, we trace the mutational history along BNB lineages that survive until viral elimination, contrasting small and large σ_A (Fig. 2.8). Here, a lineage includes a germline B cell and all its descendants; Figs. 2.8A and 2.8B depict two typical lineage trees where nodes represent B cells colored by phenotypic identity (red for N-type, blue for B-type), links colored red/blue indicate affinity-altering mutations occurring in variable/conserved dimensions, and black arrows in the latest generation mark B-type clones whose offspring last until virus depletion.

With small σ_A (Fig. 2.8A), a B-type founder switches to N-type and remains N-type for extended time before switching back to B-type. When σ_A is large (Fig. 2.8B), the intermediate N-period is very brief before B-type clones take over. This contrast in N-period duration is clearly seen from the trajectories of B-type fraction in a lineage (Fig. 2.8C) and the statistics of N-period duration across lineages in many populations (Fig. 2.8D).

What drives an earlier N$\rightarrow$B switch at a higher initial Ag diversity? Recall that at larger σ_A, N-type clones have higher affinity early on and produce plasma cells more rapidly, leading to faster Ag removal (Figs. 2.4B-D, right column). Rapid reduction in viral population, in turn, makes it even harder to generate escape mutants, resulting in sooner extinction. Hence, when σ_A is large, few beneficial mutations in conserved dimensions would suffice to clear the virus. Meanwhile, only clones that improve fast enough can survive the fierce competition for rapidly diminishing Ags. Taken together, at higher σ_A, fewer but larger beneficial mutations accumulate in conserved subspaces within a shorter time (Figs. 2.8G and 2.8H), resulting in sooner dominance of B-type clones and faster viral elimination.

To illustrate how footprint shift facilitates accumulation of adaptive changes that collaborate with breadth development, we present in Figs. 2.8E and 2.8F the optimal affinity of a B cell (to encountered FDC Ags) among all possible binding subspaces (solid line) and among conserved subspaces only (dashed line), along the long-lived branch marked in panels A and B, respectively. This clearly shows that, breadth-conferring mutations accumulated in conserved subspaces are conditionally neutral: they are neutral under the current condition of selection (mutated residues being outside current contact region) but exhibit significant gain in efficacy under new conditions (strongly diversified Ags). In this way, evolving in a non-selecting environment allows "neutral" establishment of breadth-enhancing mutations (stepwise affinity changes, Figs. 2.8E and 2.8F, dashed); even occasional affinity-reducing mutations can be buffered when σ_A is small (Fig. 2.8E dashed curve, dip at an early time). This is because, modestly diverse Ags induce weak specific response thus having prolonged availability, which in turn tolerates slow adaptation of broad clones and permits a wider range of paths (Fig. 2.19, A versus B).

Although a larger σ_A results in faster Ag clearance, this gain in speed comes at a cost:

Ab quality is compromised due to limited evolution, since strong selection pressure (due to rapid Ag consumption) disfavors slow accumulation of breath-conferring mutations. As shown in Fig. 2.8I, both the affinity (solid) and prevalence (dashed) of bnAbs (affinity to conserved targets being at least $6k_{\mathrm{B}}T$) fall with increasing σ_A (hastened AM, purple to red; also see Fig. 2.20). An intermediate σ_A may balance speed and efficacy (grey, $\sigma_A = 1$); time to clearance is halved with modest reduction in bnAb affinity (relative to $\sigma_A = 0.5$).

2.3.4 Functional role of B cell compartmentalization depends on Ag variability

GCs are dynamic microstructures where B cells evolve and Abs diversify. Multiphoton imaging combined with single-cell sequencing reveals that multiple GCs form in parallel during an immune response, housing modest B cell populations varying in clonal dynamics and diversity [92]. A natural question is: What is the evolutionary advantage, if any, of segregating into many small populations, relative to assembling few large compartments?

Theoretical and experimental works have suggested that population subdivision could promote or inhibit adaptation depending on the level of epistasis [136, 137]; most of these studies consider a constant environment. What if the environment responds to system evolution? Our framework accounts for collective evolution of subdivided populations (GCs) coupled by shared adaptive environments (evolving Ags) and, therefore, lets us examine functional consequences of GC compartments on viral outcomes. It turns out that the answer depends on epitope conservation: population subdivision can slow down viral escape if Ag epitope is highly variable ($n_c < n_b$; Fig. 2.9B, red to purple), but may also delay clearance if the binding target is relatively conserved ($n_c > n_b$; Fig. 2.9A).

In both clearance and persistence regimes, moderately diverse Ags create an effective attractant gradient, drawing B cells toward the founder virus at the origin. If a fully conserved binding target exists ($n_c > n_b$), Ags will remain fixed in the corresponding subspace. Sooner clearance then solely relies on a faster and stronger convergence of B cells toward the origin; this happens to fewer larger GCs, as seen from the radial distribution of plasma cells strongly peaked near the origin (Fig. 2.9C, red histogram for a single GC) and the

shape-space snapshots (insets: left, 100 GCs; right, 1 GC). This behavior results from global competition for limited T cell help and rapid increase in selection pressure with decreasing Ag abundance. Consequently, a large effective population size reduces the drift load so that plasma cells are on average closer to the global fitness optimum (fixed at the origin). Thus, as B cells coalesce into fewer larger GCs, Ab efficacy (quantified by time- and population-averaged binding probability between plasma cells and Ags; see Appendices) rises and the time to clearance falls (Fig. 2.9E).

Conversely, if Ags are highly variable ($n_c < n_b$), BCRs effective against all variants no longer exist, and the optimal location in the shape space for Ag neutralization is no longer fixed but moving with the evolving Ag distribution. In this case, few large B cell populations, rapidly concentrated to the origin early on, are slow to redirect themselves when Ag mutants emerge outside B cell enclosure (Fig. 2.9D); the resulting plasma cells are therefore ineffective at neutralizing mutant Ags, yielding a rapid rebound following the bottleneck (Fig. 2.9B). With many modest GCs, however, B cells carrying beneficial mutations are distributed across GCs and expand to dominate locally. As such, population subdivision facilitates the generation and maintenance of clonal diversity, enhancing the "mobility" of the collective response. As a result, Abs produced are better able to contain viral mutants and slow the escape (Fig. 2.9F).

In sum, few large GCs speed up global optimization toward a fixed target, whereas many small GCs allow efficient clonal relocation in response to moving targets. Presumably, having a range of GC sizes can strike a balance in meeting these needs. In addition, GC-to-GC communication (e.g. via reactivation of migrating memory cells [126] or Ab feedback [138]) might provide a mechanism for regulating the effective population size.

2.4 Discussion

Highly mutable viruses are capable of repeatedly evading the host immune response by coevolving with it. Despite comparable pace to diversify and similar tactics to adapt, diverging courses and outcomes may result. Identifying major antigenic determinants for the path and

fate is key to predicting hence controlling the evolutionary future. This is in urgent need, as an expanding variety of rapidly adapting pathogens (notably HIV, HCV, influenza and malaria) defy universal vaccines.

Our stochastic model of coevolution describes the joint dynamics of virus and immunity in a host, accounting for strongly coupled ecological and evolutionary components. We show that distinct viral outcomes (persistence, clearance, and rebound) observed in different individuals emerge from a feedback between physical dynamics involved in molecular recognition and ecological dynamics of co-adapting populations, linked by affinity-dependent mutual selection. Without considering this feedback, clearance and rebound will not be possible.

We identify two antigenic properties that jointly control transitions between distinct regimes: conservation level of viral epitope (n_c) and Ag diversity at the onset of immune response (σ_A); these factors determine the accessibility of different parts of the B cell repertoire and, consequently, control the pace and efficacy of Ab response. By shaping the adaptive dynamics of the Ag "resource", fast arising narrow response governs when and what broad response may subsequently evolve.

Recent studies [80, 81] report bnAb-mediated clearance of HCV infection in two human subjects without treatment. In particular, one subject (designated C117, clearance subject) exhibits monotonic decline in viral load to extinction, while the other (designated C110, rebound subject) experiences a significant viral resurgence prior to eventual clearance. The contrasting viral kinetics in clearance and rebound phases as predicted by our model closely resemble those reported in the data. More importantly, our work provides a mechanistic explanation for this pronounced difference. Our model predicts that, for a sufficiently conserved epitope ($n_c \geq n_b$), as initial Ag diversity exceeds a critical value, monotonic viral clearance transitions to a distinctive rebound; a higher Ag diversity at the onset of Ab response leads to a faster initial decline in viral load. Indeed, phylogenetic and genetic distance analyses [80, 81] indicate that the rebound subject was infected with a larger diversity of viral strains than the clearance subject; moreover, a faster fall of viral load was observed in the former. In addition, our model suggests that a faster decline of Ag population in early infection should result in a retarded accumulation of Ab breadth, because depletion of

Ag stimuli halts AM, impeding the emergence and evolution of broad response. This may explain the slower development of plasma breadth in the rebound subject [81]. Finally, another observable consequence of the proposed mechanism is that the rapid fall in viral load leading to extinction (in the clearance phase) is mediated by broad antibodies, whereas the sharp drop toward Ag population bottleneck (in the rebound phase) is caused primarily by strain-specific antibodies. This is supported by data [81]: plasma samples collected from the clearance subject show similar neutralization profiles (patterns of relative potency against a panel of HCV strains) to that of bnAbs. In contrast, plasma of the rebound subject is of much lower neutralization efficacy than bnAbs prior to the viral bottleneck, indicating the dominance of specific antibodies (from which viruses escape).

We propose that flexible molecular recognition allows for plastic phenotypes: A BCR can search across a continuum of potential binding targets on the viral protein and dock at the site of best complementarity; conservation level of thus identified binding target naturally defines B-cell phenotype — broad clones primarily target conserved residues whereas narrow clones chiefly bind variable elements. Provided such flexibility in recognition, phenotype can spontaneously switch in a way that mitigates the tension in response to changing selective forces. Related concepts have been discussed in the context of tradeoff evolution in changing environments [139]. More broadly, our result suggests that physical dynamics may alter evolutionary constraints, a mechanism expected to apply to other systems where biomolecular binding mediates selection, such as the emergence of drug resistance and evolution of gene regulatory elements.

While a fixed binding target results in viral persistence in line with previous studies, our work suggests that a flexible binding footprint of BCRs enables clearance of antigenically complex pathogens. On one hand, B cells can utilize affinity-enhancing mutations outside current contact region while buffering affinity reduction within the contact, thus enhancing the rate and size of favorable changes. On the other hand, target switch allows strain-specific ancestors to generate cross-reactive descendants and hence adapt persistently to diversifying Ags.

Shift of BCR footprint opens multiple novel pathways of breadth evolution that can be

searched for using longitudinal phenotypic measurements: the BNB pathway preserves broad precusors via an intermediate period of narrow phenotypes, whereas the NB pathway expands the range of precursors to strain-specific clones that later acquire breadth-conferring mutations. During the N-period, beneficial variations accrue in conserved dimensions, which potentiate later switch to broad phenotypes. Importantly, these conditionally neutral mutations do not exacerbate selection pressure on the virus, thus prolonging the availability of Ag stimuli. In this sense, "regeneration" of B-type clones relies on BCRs evolving in a larger state space than where selection directly acts, so that lineages slowly accumulating potentiating variations stand a chance to persist.

Last but not least, our result suggests that compartmentalizing AM into multiple GCs — evolving separately but driving viral evolution collectively — may strike a balance between quickly finding an optimal solution against conserved Ag and maintaining clonal diversity to fend off pathogens with higher mutability.

Our approach employs an extended shape-space depiction of flexible molecular recognition. Different from classic models, our shape-space dimensions represent residue groups constituting the viral epitope; this representation abstracts away atomic details while retaining features that emphasize the biomolecular basis of host-pathogen interactions: First, receptors engage Ags in binding subspaces (actual contacts) but evolve in the full space (potential binding surface). Second, speed of adaptation varies among phenotypic dimensions, reflecting distinct accessibility and mutability of different parts of the epitope. Apart from inheriting desired properties of the Fisher's geometric model [140], including epistasis and drift load, our coevolution model incorporates moving fitness optima (mutating Ag targets) and hence a dynamic distribution of mutation effect for a given receptor. In addition, the "corralling" geometry with founder B cells enclosing viral Ags captures a slowing bnAb evolution amid sustained viral evolution [141], as well as viral rebound kinetics [142, 80].

This framework broadly applies to coevolutionary processes in which binding affinity constitutes a physical phenotype and evolution proceeds on similar timescales as ecological dynamics — an under-explored regime of eco-evolutionary dynamics. Our results offer a number of vaccine lessons: (1a) Create a diversity background early on such that lineages

of different origins can succeed at various stages toward highly adapted states. (1b) Maintain a search space larger than the exact target, to encourage potentiating variations that may appear neutral but can confer future advantage. In practice, these principles recommend presentation of vaccine constructs with peripheral variable residues, in addition to the conserved core of the target epitope. (2) Supply Ags at a modest yet sustained level, because strong stimulation and rapid Ag consumption hasten AM and yield predominantly low-quality B cells [143]. (3) Apply time-dependent selection pressure that first allows access to diverse ancestors and paths and later filters out inferior lineages, keeping only those capable of sustained adaptation under increasingly more severe selection; dynamic protocols can potentially shorten paths to desired outcomes [144, 145].

Experiment indicates that memory B cells can be activated by Ag, reenter GCs, and go through further AM [146]. Of particular interest is how memory reentry impacts GC dynamics and maturation outcome. Extending our model (see Appendices for details), we find that memory reentry, either by joining ongoing GCs (Fig. 2.23) or by seeding new ones (Fig. 2.24), can facilitate the development of broad antibodies and, in turn, reduce the likelihood and amplitude of viral rebound and accelerate viral clearance. In the meantime, IgM antibodies, secreted by short-lived plasma cells and not subject to GC reactions, take part in Ag removal before IgG antibodies emerge from AM. We show that an addition of IgM antibodies not only enhances Ag removal toward the viral population bottleneck, but it also aids in corralling escape mutants during rebound (Fig.2.25);

both lead to a higher chance of viral clearance in regimes with mixed outcomes. In addition to B cells, T cells and innate immune cells likely play a role in clearing viruses. Their influence is partially reflected in the initial Ag distribution: founder viruses are driven to diversify by other immune pressures before B cell response comes into play. However, not enough knowledge is yet available for modeling the full dynamics including T cell and innate responses, which may well interact with B cell dynamics in complex ways.

Our model leaves out a number of biological factors of realism, including Ag transport and recycling, cell migration within and between GCs, immune exhaustion and latent viral reservoir, whose effects deserve focused future studies. Our model assumes optimal regulation

of binding, sharp selection in T-help competition, and perfect conservation of core residues; relaxing these assumptions could improve biological precision though is unlikely to alter qualitative conclusions.

This work provides a caricature of viral-immune coevolution that highlights how physical and ecological dynamics interplay to drive non-intuitive system-level behaviors. We hope that it serves as a starting point for studying rapid eco-evolutionary dynamics, mediated by biomolecular interactions, with feedback across scales.

2.5 Methods

2.5.1 Mutation of BCR and Ag

Upon each single mutation in a BCR, one coordinate is chosen from n dimensions uniformly at random ($k \in \{1, 2, \cdots, n\}$) and changed by an amount that follows a zero-mean Gaussian distribution

$$x_k \to x_k + \delta x_k;\ \delta x_k \sim \mathcal{N}(0, \Delta_{B,k}^2). \tag{2.5}$$

For a mutation in an Ag, one coordinate randomly chosen from n_v variable dimensions ($k \in \{1, 2, \cdots, n_v\}$) is altered according to

$$y_k \to y_k + \delta y_k;\ \delta y_k \sim \mathcal{N}(0, \Delta_{Ag}^2). \tag{2.6}$$

Here the width of the Gaussian distributions characterizes the average magnitude of change in the trait value caused by a mutation. This mutation step size differs between dimensions; specifically, $\Delta_{B,k\geq n_v+1} < \Delta_{B,k\leq n_v} = \Delta_{Ag}$, i.e., on average, mutations in the conserved/slow dimensions ($k \geq n_v + 1$) have smaller effect than those in the variable dimensions ($k \leq n_v$), reflecting a lower accessibility of the conserved core than surrounding variable elements in the target epitope [96, 22].

2.5.2 Simulating coevolution

2.5.2.1 Step 0: Initialization

(i) Generation of germline B cells

We assume all the germline B cells bind the founder virus with an equal reference affinity $A_{l^*} = 0$ (i.e. barely meeting an activation threshold) in a randomly chosen optimal binding subspace labeled by l^*. According to Eq.2.1, each germline B cell generated in this way resides on a founder hypersphere of dimension n_b and radius $R_f = \sqrt{n_b A_{max}}$ in its optimal binding subspace; A_{max} denotes the maximum affinity at perfect match. Collectively, germline B cells enclose initial Ag (see (iii) for their generation), as illustrated in Fig. 2.1B. In Appendices, we consider germline B cells with a range of affinities for the founder virus, and show that the qualitative results remain while the likelihood of viral clearance increases in place of viral rebound (Fig. 2.26).

By definition, in any binding subspace other than the optimal one ($l' \neq l^*$), local binding affinity $A_{l'}$ would be no greater than zero. Thus, we draw B-cell coordinates outside each cell's respective optimal subspace (i.e. outside the binding footprint) so that $-A_{l'}$ follows an exponential distribution with mean $\lambda = 2$, mimicking declining probabilities radially outward from the founder hypersphere. In Appendices, we vary the value of λ and find that increasing λ impedes the development of both narrow and broad Ab responses and slows viral clearance (Fig. 2.27), since footprint shift becomes less likely to yield an affinity gain or buffer an affinity loss.

(ii) GC compartmentalization

Distribute 5000 germline B cells thus generated randomly and evenly among a total number of N_{GC} GCs, where they replicate without mutation until reaching the overall capacity $K_b = 10^6$.

(iii) Generation of initial Ag

Ags begin to diversify ever since the founder virus is transmitted. By the time first responsive B cells become activated, a cloud of variant Ags surrounding the infecting strain

is developed. To model different levels of Ag diversity at the beginning of GC reactions, Ag coordinates in the variable dimensions are drawn from a Gaussian distribution centered at the origin with a width σ_A. Ag population starts at the capacity $K_a = 10^5$.

2.5.2.2 Step 1: Population dynamics of GC B cells

(i) Ag presentation in GCs: Ags are transported from circulation into each GC and presented on follicular dendritic cells (FDCs), providing stimuli for B cell AM. In total $K_{a,GC} = 10^5$ Ags are distributed evenly among GCs. In each GC cycle, FDC Ags are updated with a random subset of current plasma Ags.

(ii) Ag binding and internalization: In each GC cycle, a B cell randomly encounters 100 FDC Ags and internalizes them with a probability given by Eq. 2.3. Note that different B cells may encounter different sets of FDC Ags. The time-dependent Ag concentration is given by

$$C_{Ag}(t) = C_{Ag,0}\frac{N_{Ag}(t)}{K_a}, \tag{2.7}$$

where $N_{Ag}(t)$ represents the Ag population size at cycle t and $K_a = 10^5$ the Ag carrying capacity.

(iii) Competition for limited T cell help: Rank surviving B cells in each GC by their affinity averaged over encountered FDC Ags, and keep the top fraction f_{Tfh} while removing the rest that fail to receive T cell signal and apoptose.

(iv) Memory differentiation: A fraction $p_{mem} = 5\%$ of the remaining B cells differentiate into memory cells and leave GC.

(v) Plasma cell differentiation: Among the remaining B cells whose average affinity is above a threshold $A_{pla} = 4k_{\mathrm{B}}T$, a fraction $p_{pla} = 5\%$ differentiate into plasma cells and leave GC.

(vi) B cell replication and mutation: Each remaining B cell divides into two daughter cells that independently mutate with a probability $\mu = 0.5$. If a mutation occurs, there is a chance of $p_{let} = 0.3$ that it is lethal and the daughter B cell is removed from GC. With a probability of $p_{sil} = 0.5$, the mutation is silent and the daughter cell retains the shape-space

coordinates of the parent cell. Otherwise the mutation alters affinity [147, 148, 105, 109, 111]; one of the shape-space coordinates will change according to Eq. 2.5.

(vii) GC reseeding: A GC ends when it either perishes or exceeds the initial population size K_b/N_{GC} [109, 111]. A new GC is seeded in the next cycle; reseeding includes generation and replication of germline B cells as well as sampling of FDC Ags. This step ensures a constant number of ongoing GCs.

2.5.2.3 Step 2: Accumulation of plasma and memory cells

We treat both memory and plasma cells as Ab-secreting cells that accumulate in the plasma compartment and drive viral evolution, without explicitly accounting for differentiation of memory cells into plasma cells upon Ag activation. In Appendices, we study the influence of GC reentry or reseeding by activated memory cells and the effect of IgM antibodies secreted by short-lived plasma cells. It has been reported that humans can maintain circulating memory B cells for many decades after first exposure [149]. Hence, we ignore the intrinsic decay of Ab-secreting cells; instead, cells in excess of a carrying capacity $K_p = 10^6$ of the plasma compartment are randomly picked and removed.

2.5.2.4 Step 3: Population dynamics of Ag

(i) Neutralization: In every GC cycle, each Ag encounters antibodies secreted by 100 randomly chosen Ab-secreting cells. An Ag is neutralized and removed from the simulation according to the probability given by Eq. 2.4. Different Ags may encounter antibodies produced by different subsets of Ab-secreting cells, which exhibit a time-dependent concentration

$$C_{pla}(t) = C_{pla,0}\frac{N_{pla}(t)}{K_p}, \tag{2.8}$$

where $N_{pla}(t)$ denotes the population size of plasma and memory cells combined at cycle t, and K_p is the carrying capacity of these Ab-secreting cells.

(ii) Replication and mutation: Once every two GC cycles, surviving Ags replicate; during replication, they mutate at a rate of $0.01 * (2 * n_v/n)$/virion/generation, so that when $n_v =$

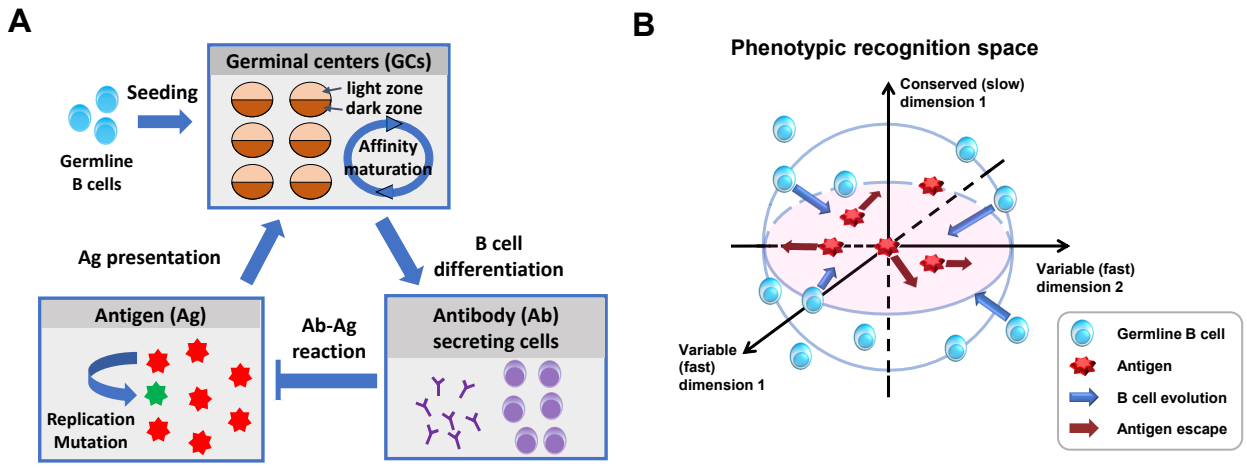

Figure 2.1: **Model: from physical space to shape space.** (A) Overview of model compartments, where key agents engage in cycles of stochastic reactions: AM — B cell proliferation and mutation in the dark zone followed by affinity-dependent selection in the light zone — in multiple GCs seeded by germline B cells, differentiation of mature B cells into Ab-secreting cells that exit GCs, replication-mutation-selection of Ags in circulation, and sampling and presentation of Ags into GCs that fuel further AM. (B) Shape-space representation of BCR-Ag interaction and coevolution. Germline B cells (blue ovals) are distributed on or outside the founder hypersphere (blue circles delineating cross-sections), enclosing the Ags (red stars). Single mutations manifest as jumps (colored arrows) with Gaussian distributed step size. B cells move slowly in conserved dimensions where Ags are fixed at the origin (vertical), while moving fast in variable dimensions along which Ags may escape (horizontal). A 3D binding subspace ($n_b = 3$) is illustrated with two variable (fast) dimensions and one conserved (slow) dimension. B cells (Ags) that get closer to (farther from) the Ags (B cells) are preferentially selected.

$n/2$, it matches the observed rate [115]. Upon mutation, the shape-space coordinates of Ags are altered according to Eq. 2.6.

(iii) Carrying capacity: Excess Ags beyond the capacity K_a are randomly chosen and removed.

2.5.2.5 Step 4: Iteration

Repeat steps 1–3 until (a) Ag population falls below an extinction threshold $K_{low} = 100$, or (b) a maximum duration of $T_{max} = 200$ GC reaction cycles is reached.

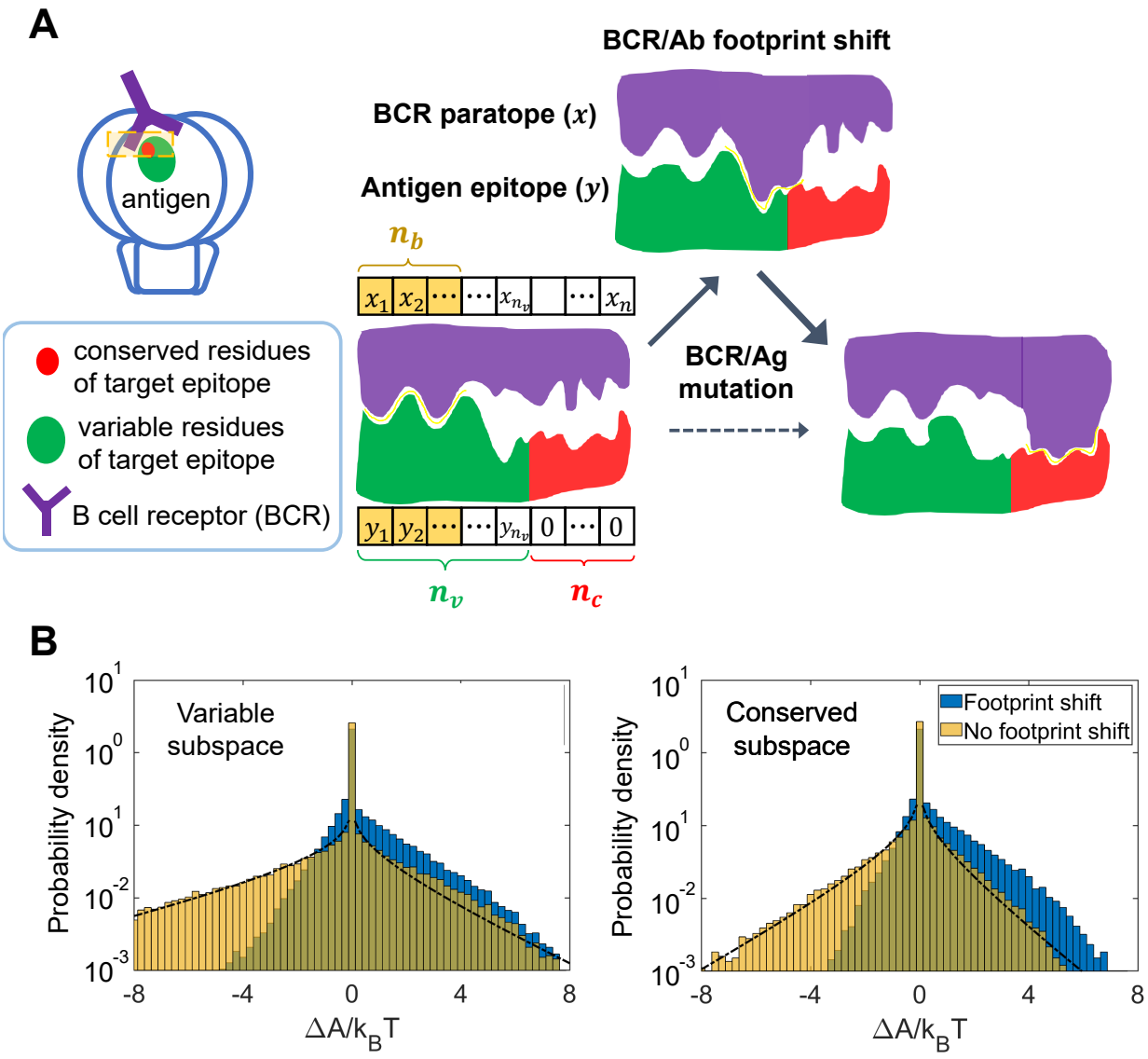

Figure 2.2: **Shift of binding footprint alters the distribution of mutation effect.** (A) Binding affinity between Ag epitope $\vec{y}$ and BCR paratope $\vec{x}$ is determined by the binding footprint—the binding site with the best complementarity (i.e. the n_b-dimensional subspace in which the shape-space distance is shortest). An initial configuration (lower left) has strongest binding in first n_b dimensions (yellow boxes); adjacent dimensions represent spatially proximal residue groups, here in total n_v (n_c) variable (conserved) ones. Upon mutation in BCR or Ag (change in purple or green shape), whether within or outside the current contact region (yellow interface), footprint may shift (top and lower right). Favorably, shift occurs from variable (green) via mixed to conserved (red) parts of the epitope (solid arrows); direct shift to conserved regions is unlikely (dashed arrow). (B) Distribution of mutation effect shown for binding of germline B cells to the founder virus, with footprint in variable (left) or conserved (right) subspace before mutation, and footprint shift being allowed (blue) or inhibited (yellow) upon mutation. The black curve indicates the theoretical distribution for fixed footprint; see Appendices for derivation. $n_c = 3$, $n_v = 5$, and $n_b = 3$.

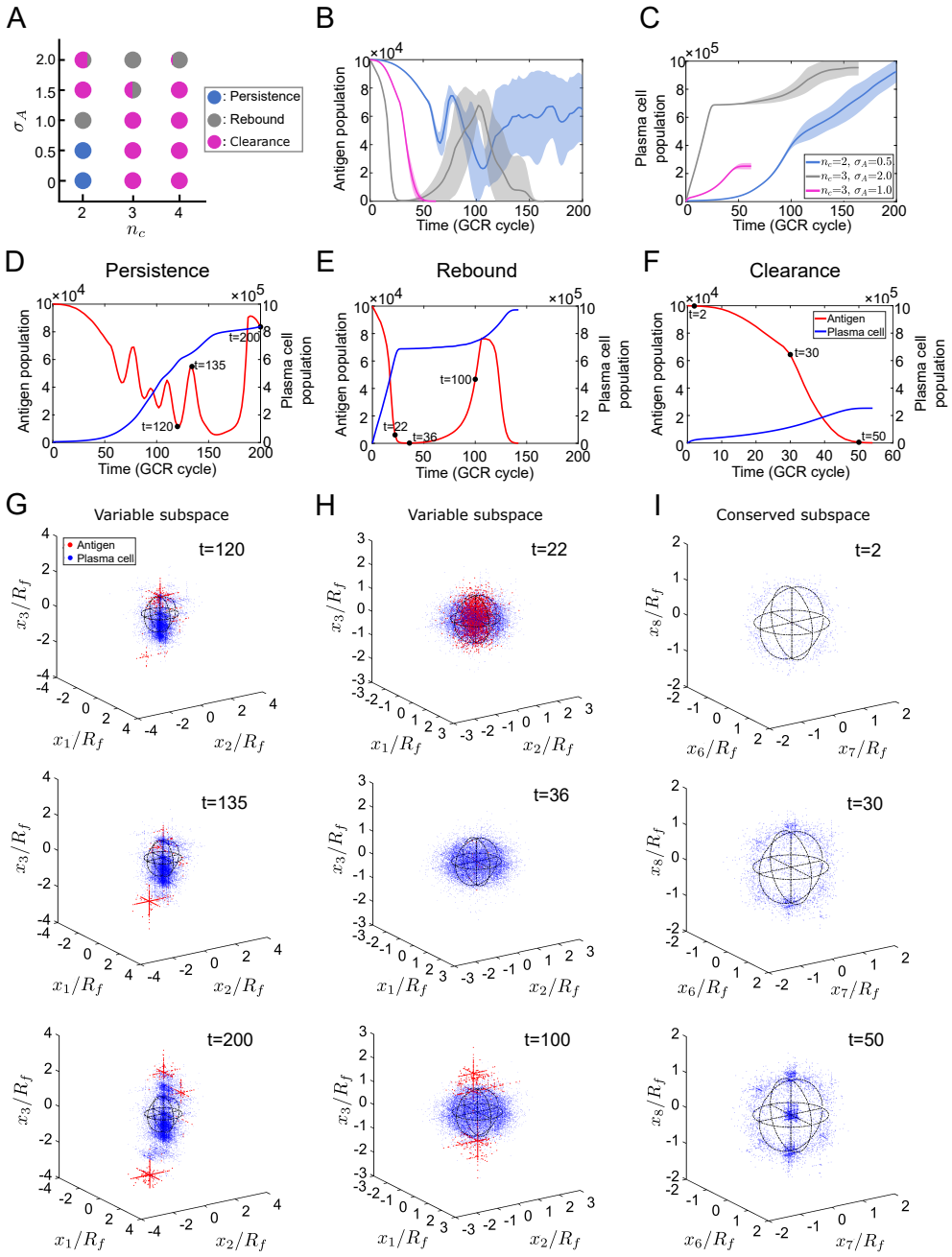

Figure 2.3: **Distinct coevolutionary outcomes and patterns emerge at different combinations of initial Ag diversity and epitope conservation.** (A) Phase diagram of three coevolution outcomes: viral clearance (magenta), rebound (grey), and persistence (blue). At high initial Ag diversity (large σ_A), both clearance and rebound could happen in repeated simulations (each hybrid symbol indicating the proportion of occurrence among 100 runs). (B, C) Population trajectories of Ags (B) and plasma cells (C) in three phases, shown with mean±SD (shade) from 10 simulations in each case. Here and after, we measure time in units of GCR cycles; one GCR cycle corresponds to 6 to 12 hours in real time. (D-I) Representative population trajectories and shape-space configurations of Ags (red) and plasma cells (blue) in each phase. (D, G) Persistence; $n_c = 2$, $\sigma_A = 0.5$. (E, H) Rebound; $n_c = 3$, $\sigma_A = 2$. (F, I) Clearance; $n_c = 3$, $\sigma_A = 1$. Snapshots in a binding subspace of $n_b = 3$ dimensions are taken at time points marked by black dots in (D–F). (G) and (H) illustrate Ag escape in a variable subspace; (I) demonstrates Ag clearance by bnAbs in a conserved subspace. For visual clarity, a random subset of 5% of plasma cells and 25% of Ags are shown. Shape-space coordinates are scaled by the radius R_f of the B cell founder hypersphere; black circles delineate the cross-sections of founder sphere with three orthogonal planes intersecting at the origin.

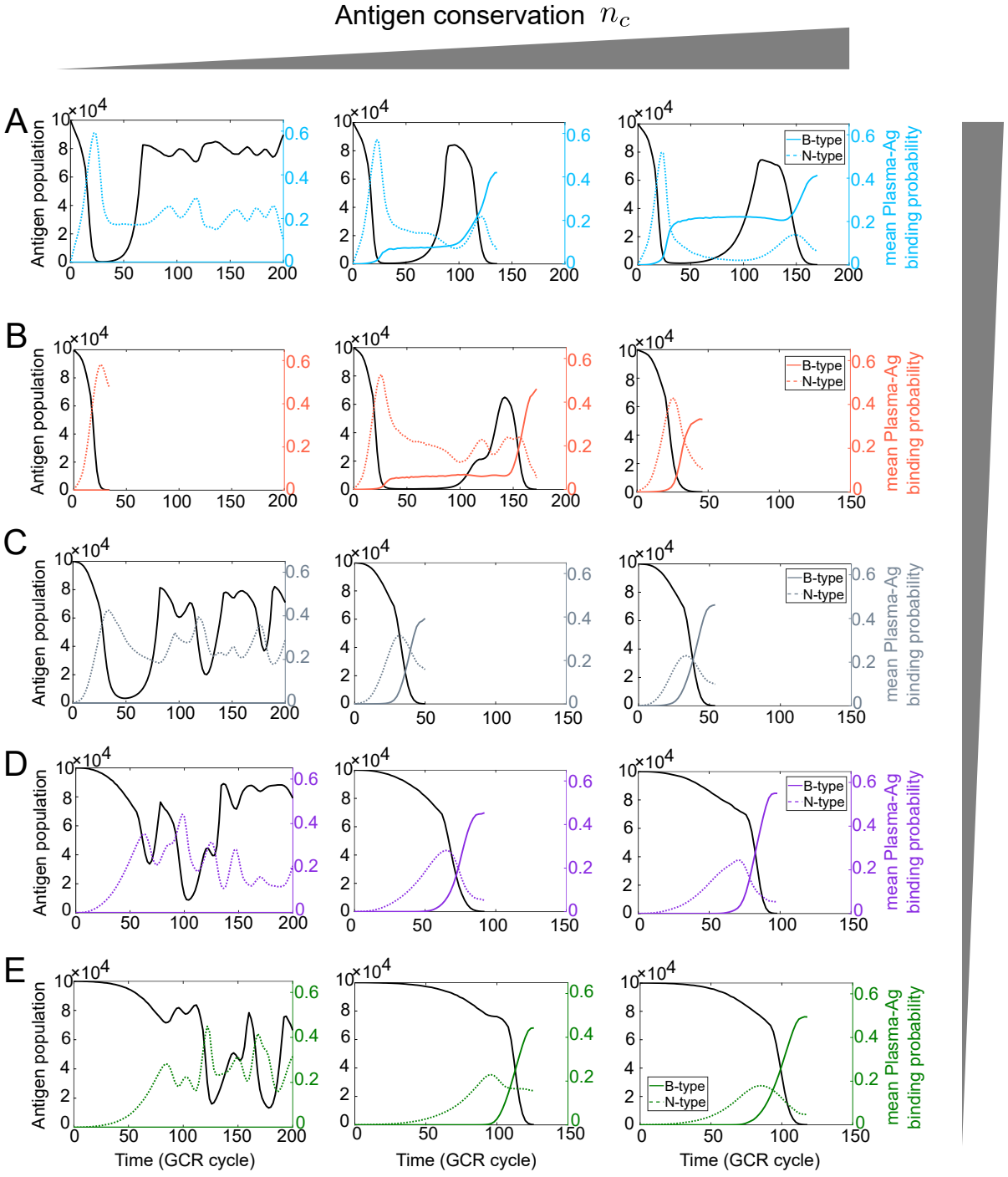

Figure 2.4: **Transitions between distinct phases are governed by the timing and efficacy of strain-specific and cross-reactive B cell responses.** Representative trajectories of Ag population (black) and average BCR-Ag binding probability (colored) for broad (B-type, colored solid; P_B) and narrow (N-type, colored dotted; P_N) B cell lineages under different combinations of epitope conservation (n_c) and initial Ag diversity (σ_A). Parameter choices correspond to the phase diagram in Fig.2.3A; left to right: $n_c = 2, 3, 4$; top to bottom: $\sigma_A = 2, 1.5, 1, 0.5, 0$. When $n_c = 2$ ($n_c < n_b$, left column), there is no contribution from B-type lineages (i.e. $P_B = 0$), since fully conserved binding subspace does not exist.

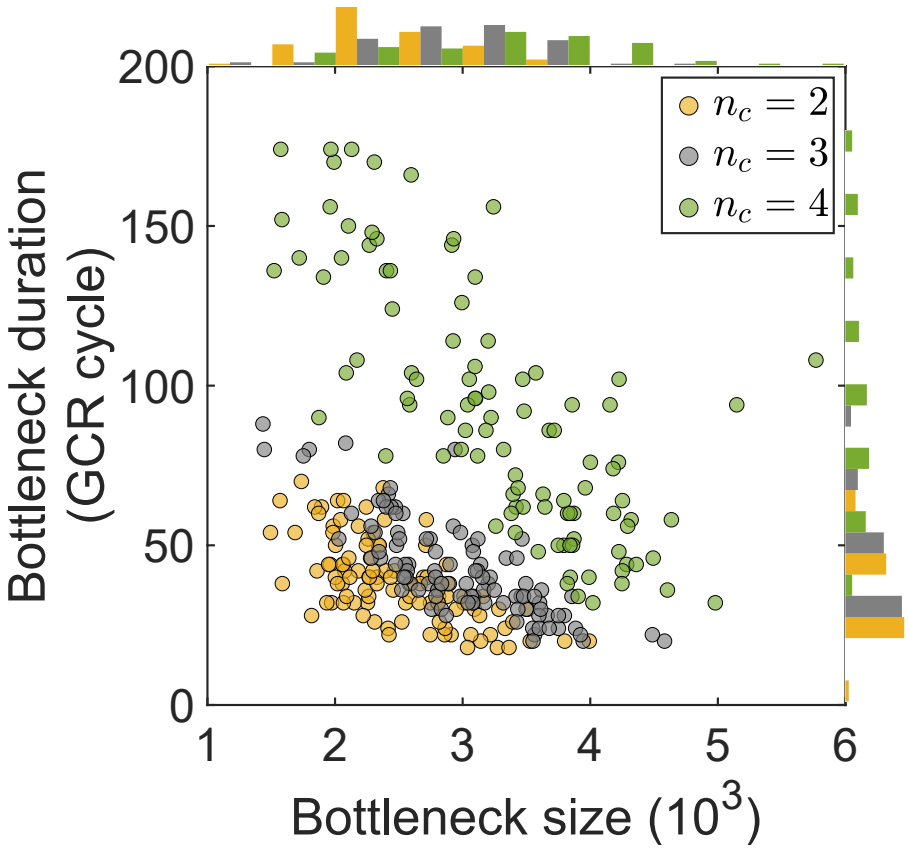

Figure 2.5: **Epitope conservation controls the size and duration of Ag population bottleneck in rebound phase.** Stronger epitope conservation (larger n_c) yields a shallower yet longer-lasting Ag population bottleneck. Bottleneck starts when Ag population first drops to below 10% capacity and ends when it recovers to above this level. Bottleneck size is defined by the time-averaged Ag count during the bottleneck. 100 simulations (scatters) are performed for each n_c value; corresponding histograms are shown along the sides. Initial Ag diversity $\sigma_A = 2$.

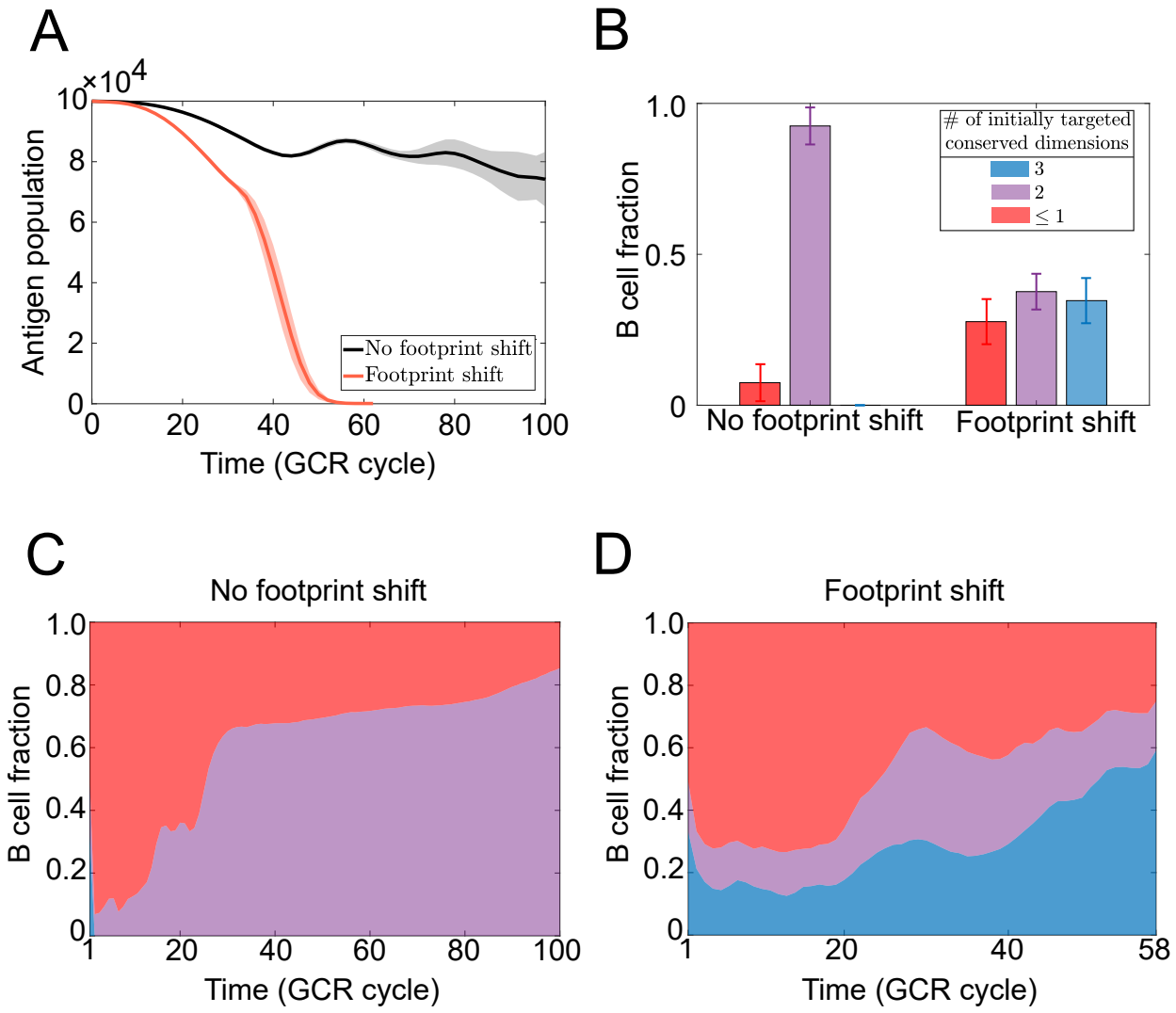

Figure 2.6: **Footprint shift enables bnAb development and viral clearance.** (A) Ag population trajectories with (red) and without (black) shift in BCR binding footprint. Ag clearance only occurs if footprint shift is allowed. (B) Evolved B cells are grouped based on binding footprint of their germline ancestors against the founder virus; color indicates the number of conserved dimensions in the optimal binding subspace. Data are collected at $t = 100$ (no footprint shift) or when Ag population drops to 10% capacity (with footprint shift). B cells that initially target a fully conserved region (blue) only survive under footprint shift. (A) and (B) are based on the same set of simulations. (C, D) Typical trajectories of GC composition (fraction of three cell groups defined in panel B), when footprint is fixed (C) or adjustable (D). $n_c = 4$, $\sigma_A = 1$.

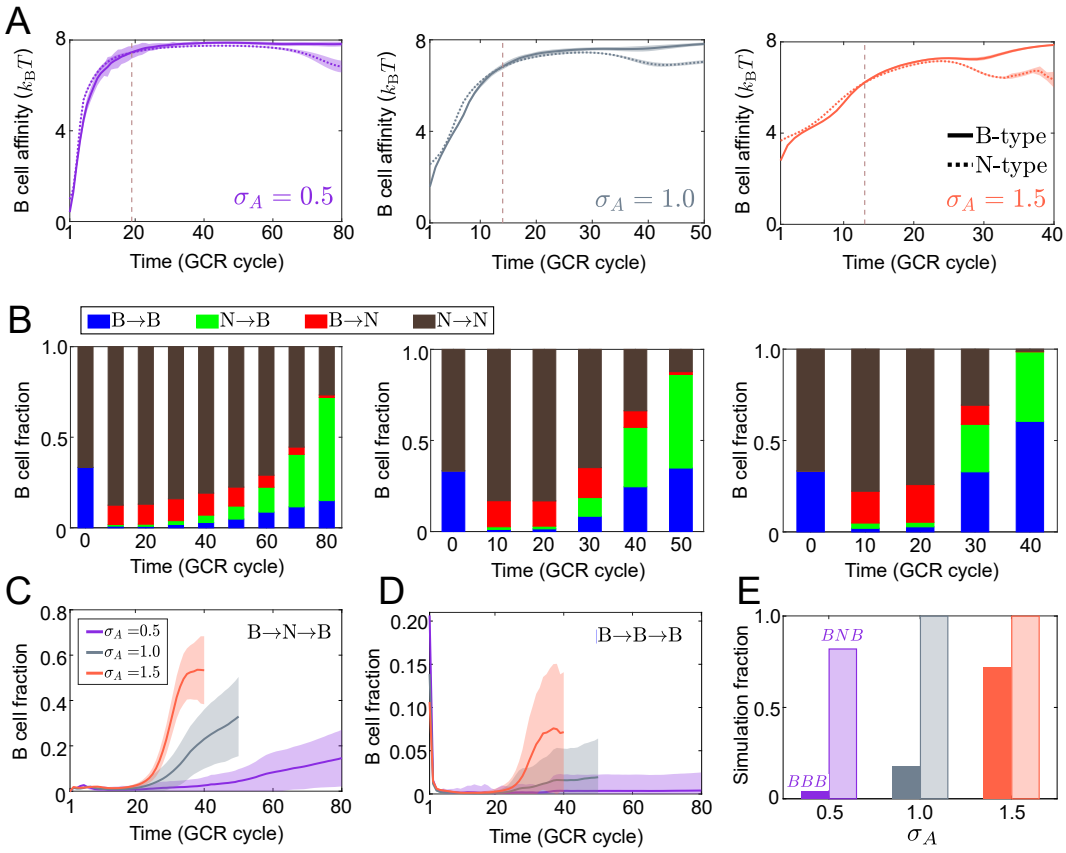

Figure 2.7: **Footprint shift opens novel pathways toward broad response.** (A) Affinity trajectories of B-type (solid) and N-type (dotted) clones under different initial Ag diversity. Vertical dashed lines mark the time since when the mean affinity of B-type exceeds that of N-type. (B) Temporal evolution of ensemble-averaged GC composition based on initial and current identity of individual cells, e.g. N→B represents current B-type clones descending from an N-type germline ancestor; see Fig. 2.17 for absolute B cell counts of each switch type. (A, B) Left to right: $\sigma_A = 0.5, 1, 1.5$. (C, D) Fraction of BNB (C) and BBB (D) cells until viral clearance; see Fig. 2.18 for prevalence of other pathways. (E) Fraction of simulations (GC ensembles) containing BBB or BNB lineages. Opaque: BBB, transparent: BNB. In all panels, data are collected from the same 50 simulations for each σ_A value; purple for $\sigma_A = 0.5$, grey for $\sigma_A = 1$, and red for $\sigma_A = 1.5$. $n_c = 4$.

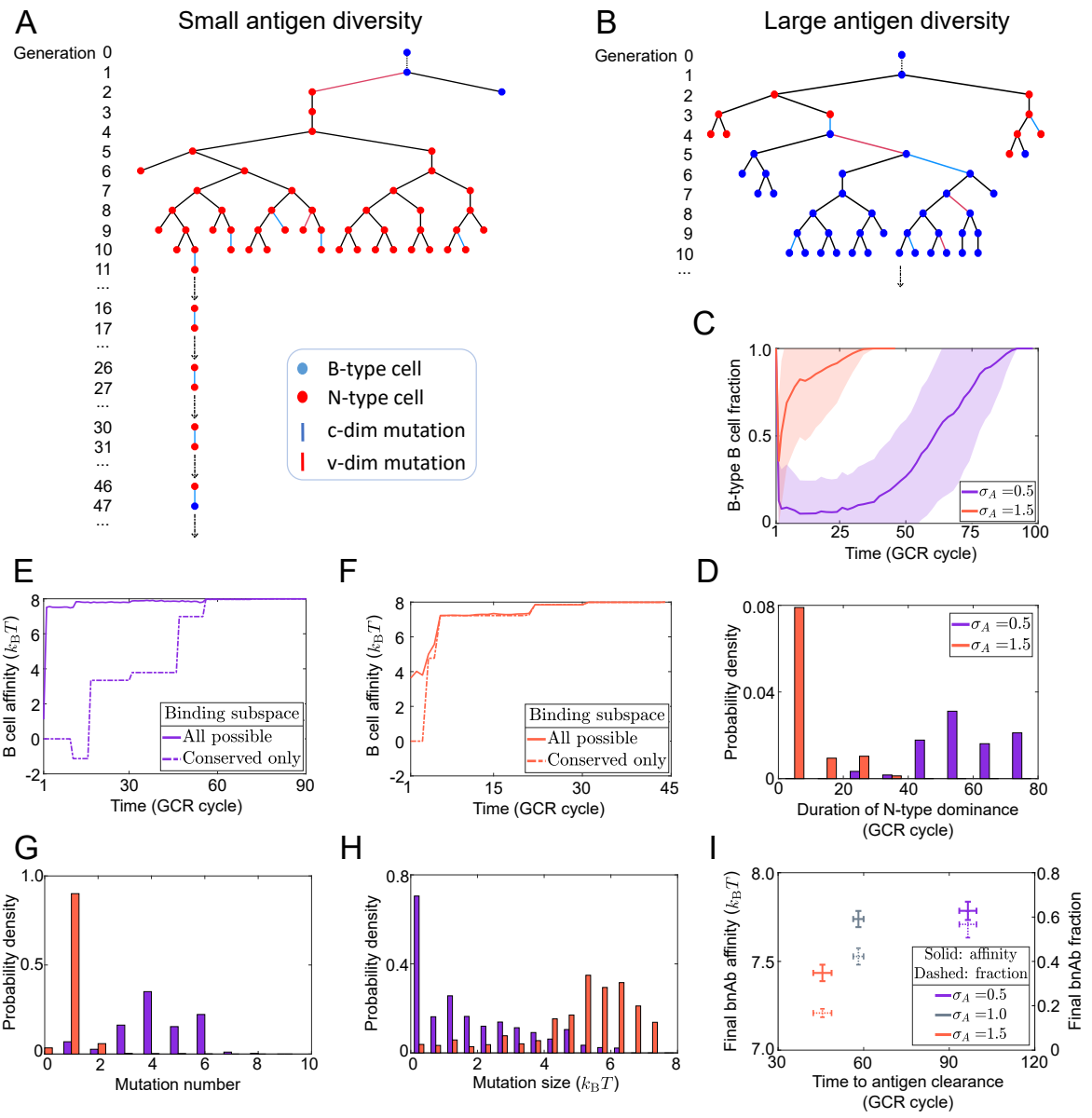

Figure 2.8: **Increasing initial Ag diversity speeds up expansion of BNB lineages at the cost of bnAb efficacy.** (A, B) Example B cell lineage trees typical of the BNB pathway under small ($\sigma_A = 0.5$) and large ($\sigma_A = 1.5$) initial Ag diversity. A BNB lineage stems from a B-type founder (blue node in generation 0) and undergoes an N-type dominated intermediate stage (takeover of red nodes) before switching back to B-type dominance (regrowth of blue nodes). Shown are long-lived lineages responsible for viral clearance, where the black arrow at the latest generation indicates the branch that lasts to the very end. Full tree structures are presented only up to generation 10. The extended branch in (A) contains the earliest reappearing B-type clone (at cycle 47) and mutations in conserved dimensions (blue links) leading to it. (C, D) Fraction of B-type cells (mean±SD, panel C) and histogram for the duration of N-period (N-type fraction being above 50%, panel D) along long-lived BNB lineages for $\sigma_A = 0.5$ (purple) and $\sigma_A = 1.5$ (red), respectively. (E, F) B cell affinity along the longest-lived branch (marked in A and B) until Ag clearance. Shown are the optimal affinity among all binding subspaces (solid) and that among fully conserved subspaces only (dashed). (G, H) Number (G) and size (H) of beneficial mutations in conserved subspaces during the N-period. (C, D) and (G, H) are all based on the same 50 simulations for each σ_A value. (I) BnAb efficacy, quantified by mean affinity to conserved targets (solid), final fraction among plasma cells (dotted) and time to clear Ag (x axis) at different initial Ag diversity. A bnAb has an optimal binding affinity to conserved targets of at least $6k_BT$. $n_c = 4$.

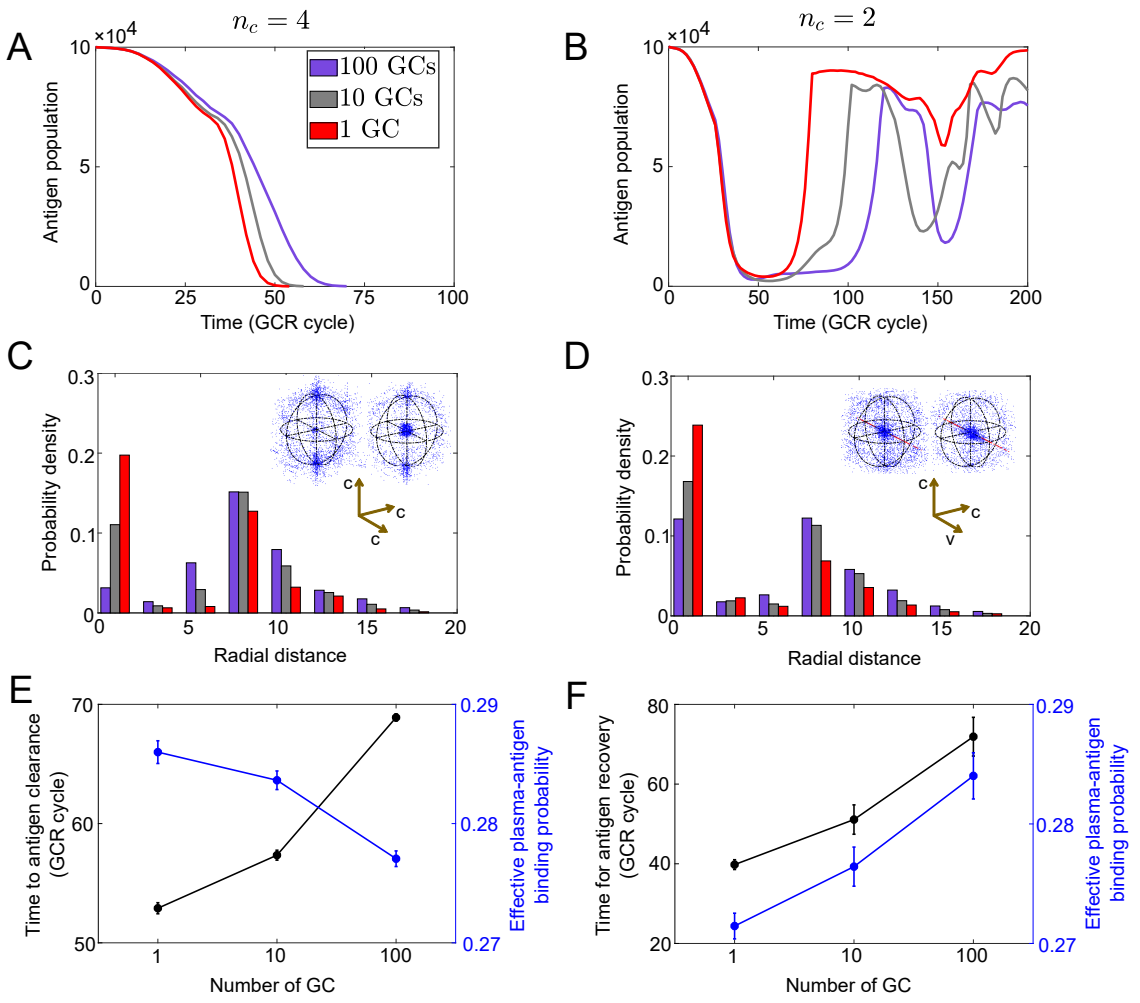

Figure 2.9: **GC compartmentalization slows viral clearance or rebound depending on epitope conservation.** Population subdivision has different functional consequences in the presence (left column, $n_c > n_b$) and absence (right column, $n_c < n_b$) of fully conserved binding targets. (A, B) A typical trajectory of Ag population for each value of GC number. (C, D) Radial distribution of plasma cells in the optimal binding subspace. Data are collected from simulations in (A) and (B), respectively, at the time when Ag population drops to 10% capacity in the single-GC case. Insets show shape-space patterns for one GC (right) and 100 GCs (left). Blue/red dots: plasma cells/Ags; black circles indicate the B cell founder hypersphere. The same color scheme applies to panels (A–D). (E, F) Time to clearance (E) or rebound (F) (black) and time-averaged effective plasma-Ag binding probability (blue) for different numbers of GCs. Time for Ag recovery (F) spans between population size falling below 10% capacity and rising above 50% capacity. 20 simulations are performed for each GC number. $\sigma_A = 1$.

APPENDIX

2.A Parameter choice

2.A.1 Dimensionality of binding subspace

There potentially is strong variability in what portion of the Ag-binding region of a BCR determines its footprint on an Ag. When choosing the dimensionality n_b of the binding subspace, we took a structural study [150] as a reference, where binding configurations of a germline Ab with diverse epitopes on multiple peptides were reported. Out of 26 Ab residues involved in binding with these peptides, on average only 11 residues determine the footprint. This led us to a choice with a similar proportion, i.e. $n_b/n \sim 11/26$, or $n_b = 3$ given $n = 8$.

2.A.2 Energy scale and mutation step size

Like in [59], we choose $A_{max} = 8k_{\mathrm{B}}T$, so that the equilibrium constant $\exp(A/k_{\mathrm{B}}T)$ for binding to the founder virus can increase by roughly 3000-fold through AM. The mutation step size Δ is chosen such that a single mutation can typically change the BCR-Ag binding affinity by 1–2 $k_{\mathrm{B}}T$.

2.A.3 Ag population size

We set Ag carrying capacity to 10^5, which is large enough to capture the variety of viral load kinetics in vivo [151] while being computationally efficient. With a proper choice of the concentration coefficients $C_{Pla,0}$ and $C_{Ag,0}$, observed viral-load variations up to three orders of magnitude were reproduced in simulations.

Table 2.1: Simulation parameters in chapter 2

Parameter	Notation	Value	Reference
Number of shape-space dimensions	n	8	[59],[152]
Number of variable dimensions	n_v	4–6	
Number of conserved dimensions	n_c	4–2	

Table 2.1 continued from previous page

Number of dimensions of BCR-Ag binding footprint	n_b	3	
Maximum B cell affinity to the founder virus ($k_{\mathrm{B}}T$)	A_{max}	8	[59]
Radius of B cell founder hypersphere	R_f	4.9	
Average affinity advantage of germline B cells' optimal footprint over other footprints ($k_{\mathrm{B}}T$)	λ	2	
B cell carrying capacity	K_b	10^6	[153]
Total number of distinct germline B cells		5000	[154]
Total number of FDC Ag	$K_{a,GC}$	10^5	
Number of FDC Ags each B cell encounters in a GC cycle		100	
Fraction of GC B cells receiving T cell help	f_{Tfh}	0.7	[59]
Probability of differentiation into memory cells	p_{mem}	0.05	[59]
Probability of differentiation into plasma cells	p_{pla}	0.05	[59]
Affinity threshold for plasma cell differentiation ($k_{\mathrm{B}}T$)	A_{pla}	5	
Probability of mutation in each daughter B cell	μ_B	0.5	[155],[156]
Fraction of lethal BCR mutation	p_{let}	0.3	[155],[156]
Fraction of affinity-altering BCR mutation	p_{aa}	0.2	[155],[156]
Fraction of silent BCR mutation	p_{sil}	0.5	[155],[156]
Average B cell mutation step size in variable dimensions	$\Delta_{B,v}$	2	[59]
Average B cell mutation step size in conserved dimensions	$\Delta_{B,c}$	1	
Initial size and capacity of Ag population	K_a	10^5	
Initial Ag diversity	σ_A	0–2	
Number of plasma cells each Ag encounters in each cycle		100	
Probability of mutation per Ag per generation	μ_{Ag}	$0.01n_v/4$	[24]
Average Ag mutation step size in variable dimensions	Δ_A	2	
Extinction threshold of Ag population		100	
Total number of GCs		$1, 10, 100$	
Concentration coefficient of FDC Ag	$C_{Ag,0}$	0.05	
Concentration coefficient of Plasma BCR	$C_{Pla,0}$	0.005	
Simulation time span (GC cycle)	T_{max}	200	
Affinity threshold of bnAbs ($k_{\mathrm{B}}T$)		6	

2.B Distribution of mutation effect without footprint shift

Below we derive the theoretical distribution of mutation effect (Fig. 2.2, black curve). We consider the change in binding affinity between germline B cells and the founder virus at the shape-space origin, due to a single affinity-altering mutation. Each germline B cell targets the founder virus in one of the $n-n_b+1$ possible binding subspaces. Cells targeting the same subspace are uniformly distributed on the surface of a n_b-dimensional hypersphere centered at the origin in that subspace, with a radius $R_f = \sqrt{n_b A_{max}}$. We set $n = 8$, $n_b = 3$, and $A_{max} = 8$, so $R_f \approx 4.9$.

Each affinity-altering mutation changes the shape-space coordinate in one randomly chosen dimension k. Mutation step size δx_k follows a centered Gaussian distribution, $\delta x_k \sim \mathcal{N}(0, \Delta_{B,k}^2)$, where $\Delta_{B,k} = 2$ for mutations in variable dimensions ($k = 1, 2, 3, ..., n_v$), and $\Delta_{B,k} = 1$ for those in conserved dimensions ($k = n_v + 1, ..., n$).

Without footprint shift, mutations outside a germline B cell's target binding subspace has no effect on affinity (i.e., $\Delta A = 0$). A mutation in the target binding subspace, which occurs with probability n_b/n, will change the cell's coordinate x_k in dimension k. The resulting change in affinity can be calculated based on Eq.2.1

$$\begin{aligned} \Delta A(x_k, \delta x_k) &= -\frac{(x_k + \delta x_k)^2 - x_k^2}{n_b} \\ &= -\frac{2x_k \cdot \delta x_k + (\delta x_k)^2}{n_b} \end{aligned} \tag{2.9}$$

One can express x_k in terms of ΔA and δx_k, i.e. $x_k(\Delta A, \delta x_k) = -\frac{n_b \Delta A}{2\delta x_k} - \frac{\delta x_k}{2}$.

The probability density of affinity change $P(\Delta A)$ can be written in general as follows:

$$P(\Delta A) = \frac{n_b}{n} \int_{-\infty}^{\infty} g(\delta x_k) h_X[x_k(\Delta A, \delta x_k)] \mathrm{d}\delta x_k \tag{2.10}$$

Here, the prefactor ensures that only mutations in the target subspace are considered, $g(\delta x_k) = \frac{1}{\sqrt{2\pi}\Delta_{B,k}} \exp(-\frac{\delta x_k^2}{2\Delta_{B,k}^2})$ is the Gaussian distribution of mutation step size, and $h_X(x_k)$ is the probability density function of the coordinates of unmutated germline B cells.

To derive $h_X(x_k)$, we consider a spherical coordinate system in the 3-dimensional binding subspace and let the polar axis be along the dimension k in which the mutation takes place.

Denoting by θ the polar angle of a germline B cell's coordinates before mutation, we have $x_k = R_f \cos\theta$. Since germline B cells targeting this subspace are uniformly distributed on the founder hypersphere, we have

$$\Pr(\theta < \Theta < \theta + \mathrm{d}\theta) = \frac{\int_0^{2\pi} \sin\theta \mathrm{d}\theta \mathrm{d}\phi}{4\pi} = \frac{\sin\theta \mathrm{d}\theta}{2} \tag{2.11}$$

which is normalized for $\theta \in [0, \pi]$. The probability density $h_X(x_k)$ for $x_k \in [-R_f, R_f]$ thus reads

$$\begin{aligned} h_X(x_k) &= \Pr(\theta < \Theta < \theta + \mathrm{d}\theta)/|dx_k| \\ &= \frac{\sin\theta}{2}(R_f \sin\theta)^{-1} \\ &= \frac{1}{2R_f}. \end{aligned} \tag{2.12}$$

That is, germline coordinates are uniformly distributed on the interval $x_k \in [-R_f, R_f]$ and vanish outside this range, since germline B cells are right on the surface of the founder hypersphere.

Substituting the expressions of $g(\delta x_k)$ and $h_X(x_k)$ into Eq.2.10, we arrive at

$$P(\Delta A) = \frac{n_b}{2R_f n\sqrt{2\pi}\Delta_{B,k}} \int_{-\infty}^{\infty} \exp(-\frac{\delta x_k^2}{2\Delta_{B,k}^2})\, I(-R_f \leq -\frac{n_b \Delta A}{2\delta x_k} - \frac{\delta x_k}{2} \leq R_f)\, \mathrm{d}\delta x_k. \tag{2.13}$$

Here $I(Y)$ is the indicator function that takes value 1 if the condition Y is satisfied and takes value 0 otherwise.

2.C Dependence of germline B cell affinity on initial Ag diversity

Below we show that the average binding affinity of germline B cells to initial Ag increases with Ag diversity, σ_A, and the number of variable dimensions, n_v. Considering Langmuir binding isotherm, we calculate the "log-mean-exponential" affinity between a founder B cell $\vec{x}$ and initial Ag $\{\vec{y}\}$, i.e., $\ln\langle\exp[A(\vec{x}, \vec{y})]\rangle_{\{\vec{y}\}}$.

From Eq.2.2, we have

$$\begin{aligned} \langle \exp\left[A(\vec{x}, \vec{y})\right]\rangle_{\{\vec{y}\}} &= \exp(A_{max})\langle \exp\left[-\sum\nolimits_{k=l^*(\vec{x},\vec{y})}^{l^*(\vec{x},\vec{y})+n_b-1} \frac{(x_k - y_k)^2}{n_b}\right]\rangle_{\{\vec{y}\}} \\ &= \exp(A_{max})\langle \prod\nolimits_{k=l^*(\vec{x},\vec{y})}^{l^*(\vec{x},\vec{y})+n_b-1} \exp\left[-\frac{(x_k - y_k)^2}{n_b}\right]\rangle_{\{\vec{y}\}}, \end{aligned} \tag{2.14}$$

where $l^*(\vec{x},\vec{y})$ denotes the starting dimension of the optimal binding subspace between a germline B cell $\vec{x}$ and an Ag $\vec{y}$. When initial Ag diversity is moderate, we can assume that the optimal subspaces (binding footprints) of germline B cells for diverse initial Ag are the same as those for the founder virus at the origin, i.e., $l^*(\vec{x},\vec{y}) = l^*(\vec{x},\vec{0})$ for all $\vec{y}$. Note this is only exact when $\sigma_A = 0$ and yields an affinity no greater than the actual value by imposing a non-optimal binding subspace for finite σ_A.

As described in the Methods section in chapter 2, the probability distribution $g(y_k)$ of the Ag coordinate in any variable dimension is Gaussian and becomes a Dirac delta function in conserved dimensions. Thus, we can write $g(y_k) = \frac{1}{\sqrt{2\pi}\sigma_k}\exp(-\frac{y_k^2}{2\sigma_k^2})$, where $\sigma_k = \sigma_A$, if $1 \le k \le n_v$; $\sigma_k \to 0$, if $n_v + 1 \le k \le n$.

Eq.2.14 now becomes

$$\begin{aligned}\langle \exp\left[A(\vec{x},\vec{y})\right]\rangle_{\{\vec{y}\}} &\approx \exp(A_{max})\langle \prod\nolimits_{k=l^*(\vec{x},\vec{0})}^{l^*(\vec{x},\vec{0})+n_b-1} \exp\left[-\frac{(x_k-y_k)^2}{n_b}\right]\rangle_{\{\vec{y}\}} \\ &= \exp(A_{max}) \prod\nolimits_{k=l^*(\vec{x},\vec{0})}^{l^*(\vec{x},\vec{0})+n_b-1} \int_{-\infty}^{\infty} \exp\left[-\frac{(x_k-y_k)^2}{n_b}\right] g(y_k)dy_k \\ &= \exp(A_{max}) \exp\left[-\sum\nolimits_{k=l^*(\vec{x},\vec{0})}^{l^*(\vec{x},\vec{0})+n_b-1} \frac{x_k^2}{2\sigma_k^2+n_b}\right] \prod\nolimits_{k=l^*(\vec{x},\vec{0})}^{l^*(\vec{x},\vec{0})+n_b-1} \frac{1}{\sqrt{1+\frac{2\sigma_k^2}{n_b}}}\end{aligned} \tag{2.15}$$

Since all germline B cells have equal affinity against the founder virus, it follows that

$$A(\vec{x},\vec{0}) = A_{max} - \sum\nolimits_{k=l^*(\vec{x},\vec{0})}^{l^*(\vec{x},\vec{0})+n_b-1} \frac{x_k^2}{n_b} \equiv 0. \tag{2.16}$$

Hence,

$$\sum\nolimits_{k=l^*(\vec{x},\vec{y})}^{l^*(\vec{x},\vec{y})+n_b-1} x_k^2 \simeq n_b A_{max} \tag{2.17}$$

for any initial Ag $\vec{y}$.

Considering target subspaces (binding footprints) being fully conserved or fully variable, i.e. $\sigma_k = 0$ or $\sigma_k = \sigma_A$ for all k (mixed subspace yielding a value bounded by these limits), the log-mean-exponential affinity from Eq.2.15 becomes

$$\begin{aligned}\ln\langle \exp[A(\vec{x},\vec{y})]\rangle_{\{\vec{y}\}} &= A_{max} - \sum\nolimits_{k=l^*(\vec{x},\vec{0})}^{l^*(\vec{x},\vec{0})+n_b-1} \left[\frac{x_k^2}{2\sigma_k^2+n_b} + \frac{1}{2}\ln(1+\frac{2\sigma_k^2}{n_b})\right] \\ &= A_{max} - \left[\frac{A_{max}}{1+\frac{2\sigma_k^2}{n_b}} + \frac{n_b}{2}\ln(1+\frac{2\sigma_k^2}{n_b})\right]\end{aligned} \tag{2.18}$$

where we have used Eq.2.17. Note there is no explicit dependence on $\vec{x}$, because both the germline B cells and initial Ag are isotropic in a fully conserved or fully variable subspace.

When $\sigma_k = \sigma_A > 0$, it follows from Eq.2.18

$$\begin{aligned}\frac{\mathrm{d}}{\mathrm{d}\sigma_A}\left[\ln\langle\exp[A(\vec{x},\vec{y})]\rangle_{\{\vec{x},\vec{y}\}}\right] &= 4\sigma_A\left[\frac{n_b A_{max}}{(2\sigma_A^2+n_b)^2} - \frac{n_b}{2}\frac{1}{2\sigma_A^2+n_b}\right] \\ &= \frac{4\sigma_A n_b}{(2\sigma_A^2+n_b)^2}\left[A_{max} - \frac{1}{2}(2\sigma_A^2+n_b)\right] > 0\end{aligned} \tag{2.19}$$

as long as $\sigma_A < \sqrt{A_{max} - n_b/2}$, which holds for the parameters used in our simulations, $A_{max} = 8$, $n_b = 3$, and $\sigma_A \leq 2$. That is, the log-mean-exponential affinity between germline B cells and initial Ag indeed increase with increasing σ_A. In addition, since $\sigma_k \equiv 0$ in conserved dimensions, germline B cells targeting variable subspaces have a higher binding affinity than those targeting conserved subspaces if $\sigma_A > 0$. Thus, increasing the number of variable dimensions, n_v, will increase the fraction of the former, and in turn, lead to higher overall affinity of the germline population.

It is worth pointing out that the result that average initial BCR-Ag affinity increases with σ_A reflects a simple geometric fact: as the the distribution widens around the shape-space origin, they are getting closer to the enclosing B cell founder hypersphere from inside (though remaining within the enclosure). Such geometric fact does not require the affinity function to be quadratic in shape-space distances. A more general definition of affinity reads

$$A_l(\vec{x},\vec{y}) \equiv A_{max} - \frac{1}{n_b}\sum\nolimits_{k=l}^{l+n_b-1}|x_k - y_k|^q\,. \tag{2.20}$$

It follows that Eq.2.15 now becomes

$$\langle\exp[A(\vec{x},\vec{y})]\rangle_{\{\vec{y}\}} \approx \exp(A_{max})\prod\nolimits_{k=l^*(\vec{x},\vec{0})}^{l^*(\vec{x},\vec{0})+n_b-1}\int_{-\infty}^{\infty}\exp\left[-\frac{|x_k-y_k|^q}{n_b}\right]g(y_k)dy_k. \tag{2.21}$$

Results for $q = 1, 2, 3$ are shown in Fig.2.10.

2.D Effective plasma-Ag binding probability

To characterize the overall efficacy of B cell response in viral suppression, we define an effective plasma-Ag binding probability, $P_{\text{eff}}(t_i, t_f)$, on the time interval $[t_i,t_f]$. Specifically,

$P_{\text{eff}}(t_i, t_f)$ determines the total fold change in Ag population size, N_{Ag}, between t_i and t_f:

$$N_{Ag}(t_f) = N_{Ag}(t_i) * (1 - P_{\text{eff}})^{t_f - t_i} * 2^{\frac{t_f - t_i]}{2}} \tag{2.22}$$

Here, $1-P_{\text{eff}}$ indicates the effective probability of Ag survival in one cycle, and $2^{\frac{t_f - t_i}{2}}$ accounts for Ag replication once every other cycle (assuming first replication occurs at t_i).

The actual dynamics is described by

$$N_{Ag}(t_f) = N_{Ag}(t_i) * \prod_{t=t_i}^{t_f - 1} [1 - \langle P_V(\vec{y}, t)\rangle_{\vec{y}}] * 2^{\frac{t_f - t_i}{2}} \tag{2.23}$$

where $P_V(\vec{y}, t)$ is the probability that Ag $\vec{y}$ will be neutralized by plasma BCR and $\langle\cdot\rangle_{\vec{y}}$ indicates an average over all Ags present at cycle t. Equating Eq. 2.22 and Eq. 2.23 leads to

$$P_{\text{eff}}(t_i, t_f) = 1 - \{\prod_{t=t_i}^{t_f - 1} [1 - \langle P_V(\vec{y}, t)\rangle_{\vec{y}}]\}^{\frac{1}{t_f - t_i}}. \tag{2.24}$$

Therefore, a larger $P_{\text{eff}}(t_i, t_f)$ indicates a higher efficiency of Ag removal in a given time span $[t_i, t_f]$.

2.E Availability of beneficial BCR mutations in slow (conserved) and fast (variable) dimensions

Fig. 2.16 presents the distribution of mutation effect in fast (variable) and slow (conserved) dimensions, for lineages with a B-type ancestor that survive until Ag clearance (like in Fig. 2.8). Mutation effect ΔA is measured by mutation-induced changes in log-mean-exponential affinity of a B cell to encountered FDC Ags, i.e.,

$$\Delta A = \log\langle\exp[A(\vec{x}', \vec{y})]\rangle - \log\langle\exp[A(\vec{x}, \vec{y})]\rangle. \tag{2.25}$$

Here $\vec{x}$ and $\vec{x}'$ are shape-space coordinates of a B cell before and after mutation, respectively; $\langle\cdot\rangle$ denotes an average over FDC Ags $\{\vec{y}\}$ that B cell $\vec{x}$ encounters prior to mutation. In the case without footprint shift (yellow, "No FS"), binding footprints between B cells and FDC Ags remain unchanged after mutation.

As seen in Fig. 2.16, at earlier times, affinity improvement is larger for mutations in fast dimensions, because of a larger mutation step size $\Delta_{B,v}$. Affinity improvement is further enhanced by footprint shift (blue compared to yellow). These factors, combined with the fact that initial BCR-Ag binding is stronger in fast dimensions when $\sigma_A > 0$, drive B→N switch for higher affinity. However, beneficial mutations in slow dimensions remain available, while those in fast dimensions are quickly depleted, partly because Ags can escape BCR recognition in fast dimensions, but not in slow dimensions. This allows B cells to continuously improve affinity through mutations in conserved dimensions and, in turn, drives N→B switch in surviving lineages. Taken together, a B→N→B pathway toward breadth develops.

2.F Memory reentry into GCs promotes bnAb development and enhances viral clearance

Experiment has suggested that memory B cells can be activated by Ag, reenter GCs, and go through further AM. In particular, memory cells can either join ongoing GC reactions, or seed new GCs as founders. Since the exact conditions of memory activation and GC reentry are not well understood, we study both scenarios. In both cases, our model predicts that memory reentry promotes the development of broad Ab response and, in turn, reduces the likelihood and amplitude of viral rebound and accelerates viral clearance.

2.F.1 Memory cells reenter ongoing GCs

At each GC reaction cycle, a memory cell $\vec{x}$ in the plasma has the following probability to be activated and join a randomly picked ongoing GC:

$$P_M(\vec{x}, t) = \frac{C_{Ag}(t) \sum_{\vec{y} \in \mathcal{Y}} e^{A(\vec{x},\vec{y})}}{1 + C_{Ag}(t) \sum_{\vec{y} \in \mathcal{Y}} e^{A(\vec{x},\vec{y})}} \cdot \alpha. \tag{2.26}$$

Here, the Langmuir isotherm determines the probability that the memory cell becomes activated by a random subset $\mathcal{Y}$ of plasma Ag it encounters during cycle t. The constant $\alpha \in [0, 1]$ denotes the fraction of activated memory cells that reenter an ongoing GC. The actual value of α is unknown, and we examine the extreme case of strong memory reentry

($\alpha = 1$) that sets an upper bound of the effect in our extended model.

With reactivated memory cells joining ongoing GCs, a notable change in the coevolutionary outcome is that viral rebound turns into clearance, if the Ag epitope is sufficiently conserved for bnAbs to develop, i.e. $n_c \geq n_b$ (Fig. 2.23A versus Fig. 2.3A; see Fig. 2.23B for clearance time and Fig. 2.23C for viral load trajectories). While the total number of Ag-removing cells in the plasma falls due to memory reentry (Fig. 2.23D), the overall quality of the plasma rises since memory cells are of lower affinity than plasma cells (Fig. 2.23E).

On the one hand, the neutralization efficacy of narrow lineages (P_N) is lower (Fig. 2.23F, dashed lines), due primarily to the reduced size of the plasma pool (despite an increase in their average binding affinity due to removal of memory cells; Fig. 2.23E). On the other hand, broad lineages that arise following viral escape gain in binding potency (P_B) compared to without memory reentry (Fig. 2.23F, solid lines, $t \gtrsim 20$). This is because, as memory cells reenter GCs and go through further AM, both the quality and abundance of broad clones increase in the plasma pool. As a result, rapid clearance without rebound becomes more likely (Figs. 2.23B-C).

2.F.2 Memory cells seed new GCs

Once a GC collapses and vacates the follicle, a new one can be seeded therein. Assume a fraction β of the founders are memory cells randomly drawn from the plasma pool and the rest are germline B cells.

The effect of memory reseeding is most pronounced and interesting when (i) Ag population goes through a bottleneck so GC reseeding will occur, and (ii) $n_c \geq n_b$ so broad Ab response can develop ($P_B > 0$). As the memory-founder fraction β increases from zero (fully naive reseeding) to one (fully memory reseeding), the proportion of trajectories exhibiting monotonic clearance without rebound increases from 20% to 100% when $n_c = 4$ and $\sigma_A = 2$; the proportion increases from 40% to 100% when $n_c = 3$ and $\sigma_A = 1.5$. No monotonic clearance is observed when $n_c = 3$ and $\sigma_A = 2$, even with memory reseeding.

As we show in Fig 2.24, a larger fraction of memory founders (larger β) leads to an

increasing likelihood that viral load remains below 10% of the initial level following the bottleneck (panel A), since broad Ab response evolves faster during the bottleneck (panel C) and more effectively suppresses or even prevents viral resurgence. This is intuitively expected: since memory cells have gone through earlier rounds of AM, on average, newly seeded GCs begin with a larger breadth than if starting anew from germline founders. Consequently, virus populations get cleared within a shorter time (panel B). Note that a higher level of naive reseeding (smaller β) results in a greater variability of clearance time (panel B).

2.G IgM antibodies aid in viral removal both prior to population bottleneck and following escape

Apart from GC B cells, another Ag-consuming agent is IgM antibodies; they are secreted by short-lived plasma cells that do not go through AM. These low-affinity IgM antibodies may remove viruses before IgG antibodies are produced by GC reactions.

To study their influence on coevolutionary dynamics and outcomes, we introduce 10^5 IgM plasma cells (10% of plasma capacity), whose shape-space coordinates were generated the same way as germline B cells, before GC reactions start. We find that the addition of non-evolving IgM antibodies leads to a higher fraction of monotonic viral clearance against rebound in regimes of mixed outcomes (Figs. 2.25A-B). This is achieved in two ways. First, with IgM antibodies, the overall efficacy of specific antibodies (P_N) increases initially (Fig. 2.25D), driving a faster decline in viral load (Fig. 2.25C). Second, the drop in P_N following viral escape is reduced (Fig. 2.25D); unlike IgG antibodies that are drawn to the founder viruses (via a self-generated Ag gradient in shape space), IgM antibodies remain in the corralling configuration and, consequently, can intercept escape mutants.

The neutralization efficacy of broad lineages (P_B) slightly decreases in the presence of IgM antibodies (Fig. 2.25D), since the latter effectively dilute broad plasma cells, making it less likely that viruses encounter broad antibodies in the plasma. Nonetheless, the overall effectiveness of narrow and broad responses combined is enhanced both before and during viral bottleneck, thereby promoting viral clearance.

2.H Alternative positioning of germline B cells

So far we have assumed that all germline B cells have the same binding affinity — in their respective optimal binding subspace — for the founder virus. In any non-optimal subspace, the initial BCR-Ag binding affinity is lower and the amplitude of deviation follows an exponential distribution with mean λ (see Methods→Simulating coevolution→Step 0). Below we show that qualitative results of coevolutionary dynamics and outcomes remain valid for alternative assumptions for germline B cell affinity (Figs. 2.26 and 2.27). In particular, we allow a range of founder affinities and vary the germline potency against non-optimal binding targets. We find that both only quantitatively modify the likelihood and speed of viral clearance.

2.H.1 Differential germline affinities

To allow a range of initial affinities of germline B cells to the founder virus, we performed simulations in which founder B cells $\{\vec{x}\}$ assume Gaussian-distributed affinities $\{\Delta A(\vec{x})\}$, with $\Delta A(\vec{x}) \sim \mathcal{N}(0,1)$ sampled independently for each cell $\vec{x}$. Accordingly, the fixed radius of the founder hypersphere, $R_f = \sqrt{n_b A_{max}}$, is now replaced by a cell-specific value $R_f(\vec{x}) = \sqrt{n_b(A_{max} - \Delta A(\vec{x}))}$.

With differential founder affinities, coevolutionary outcomes remain unchanged qualitatively (Fig .2.26A vs Fig. 2.3A), yet the proportion of monotonic viral clearance against rebound is elevated. Because the most potent clones, now of higher affinity than in the case of identical germline affinity, are preferentially expanded by selection, both narrow and broad clones reach higher efficacy in removing viruses more rapidly (Fig. 2.26D). As a result, the success rate of monotonic clearance is enhanced (Figs. 2.26A and 2.26B) and the mean time to viral extinction shortened (Figs. 2.26B and 2.26C).

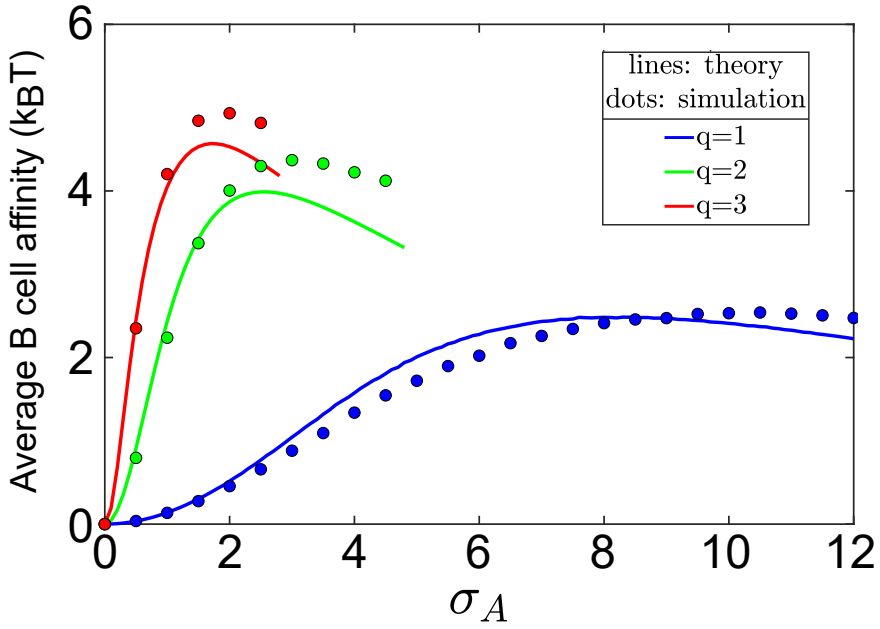

Figure 2.10: **Average binding affinity of founder B cells to initial Ag increases with Ag diversity (σ_A) when diversity is modest.** Curves, as given by Eqs. 2.18 and 2.21, terminate when σ_A reaches the radius of the founder hypersphere, $R_f = (n_b A_{max})^{1/q}$. Good agreement with simulations (symbols) is achieved, even at moderately large σ_A. Here, $n_b = 3$ and $A_{max} = 8$.

2.H.2 Varying germline quality against non-optimal binding sites

Next, we vary the affinity of germline B cells against the founder virus at non-optimal binding sites (i.e. in non-optimal binding subspaces) by varying λ. As λ increases, lower quality of binding to non-optimal targets makes it less likely that footprint shift can buffer deleterious mutations within the optimal binding site or catch beneficial mutations outside. Consequently, not only that narrow response becomes less effective (Fig. 2.27B, dashed line), but the efficacy of subsequent broad response rises more slowly (Fig. 2.27B, solid line). Therefore, bnAb development via footprint shift weakens and viral clearance by bnAbs slows (Fig. 2.27A).

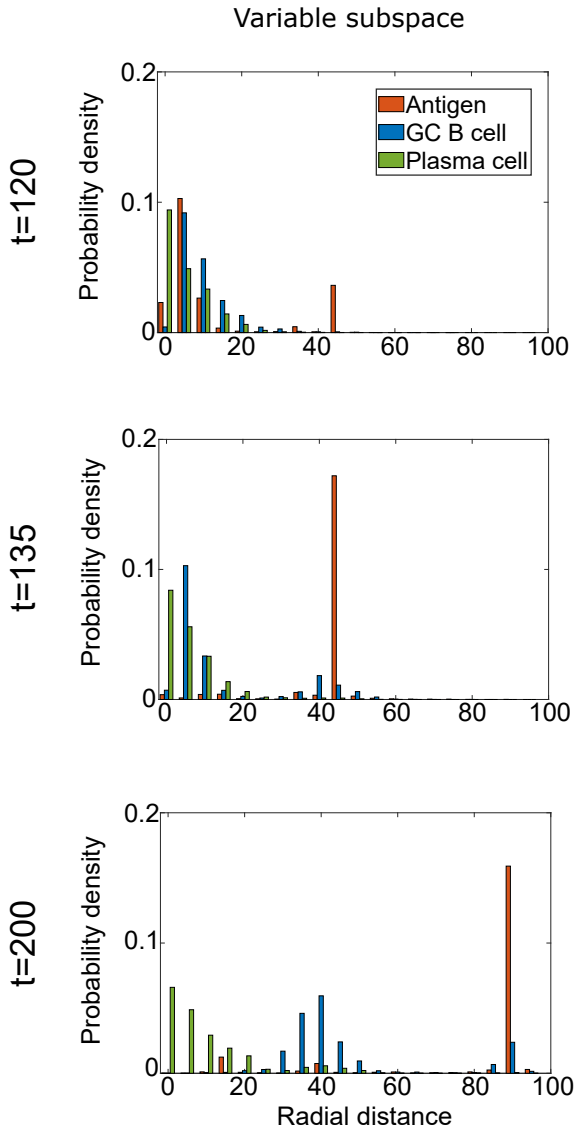

Figure 2.11: **Radial distribution of Ag, GC B cells and plasma cells in the persistence phase.** Driven by strain-specific GC B cells (blue), mutant Ag (red) emerge increasingly farther from the founder strain at the origin that focuses plasma cells (green). Data are collected from the same simulation as shown in Fig.2.3D, at time points marked therein and in the same subspace as in Fig.2.3G. Radial distance is measured by r^2/n_b, where r is the Euclidean distance between an agent and the subspace origin; $n_b = 3$ is the dimensionality of binding subspace. $n_c = 2$, $\sigma_A = 0.5$. The same legend applies to all panels.

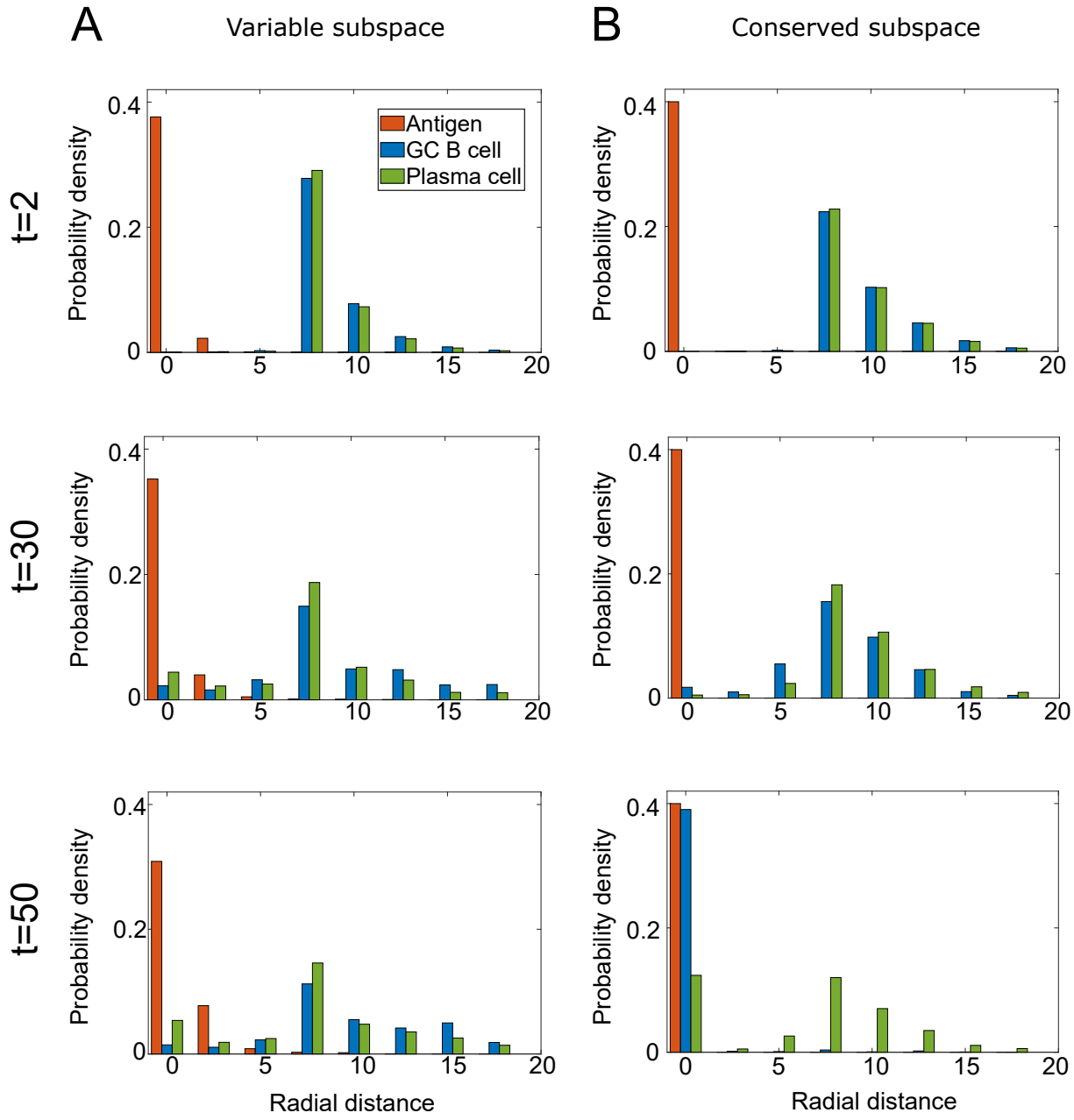

Figure 2.12: **Radial distribution of Ag, GC B cells and plasma cells in the clearance phase.** Rapid convergence of GC B cells (blue) toward the origin in the conserved subspace leads to clearance of Ag (red) soon afterward. Note that the plasma cell distribution develops a peak at the origin of the conserved subspace (B) while remaining similar in the variable subspace (A) at later times. From the same simulation as in Fig.2.3F at marked time points and in the same conserved subspace as in Fig.2.3I. (A)/(B): Radial distribution in a fully variable/conserved subspace. $n_c = 3$, $\sigma_A = 1$. The same legend applies to all panels.

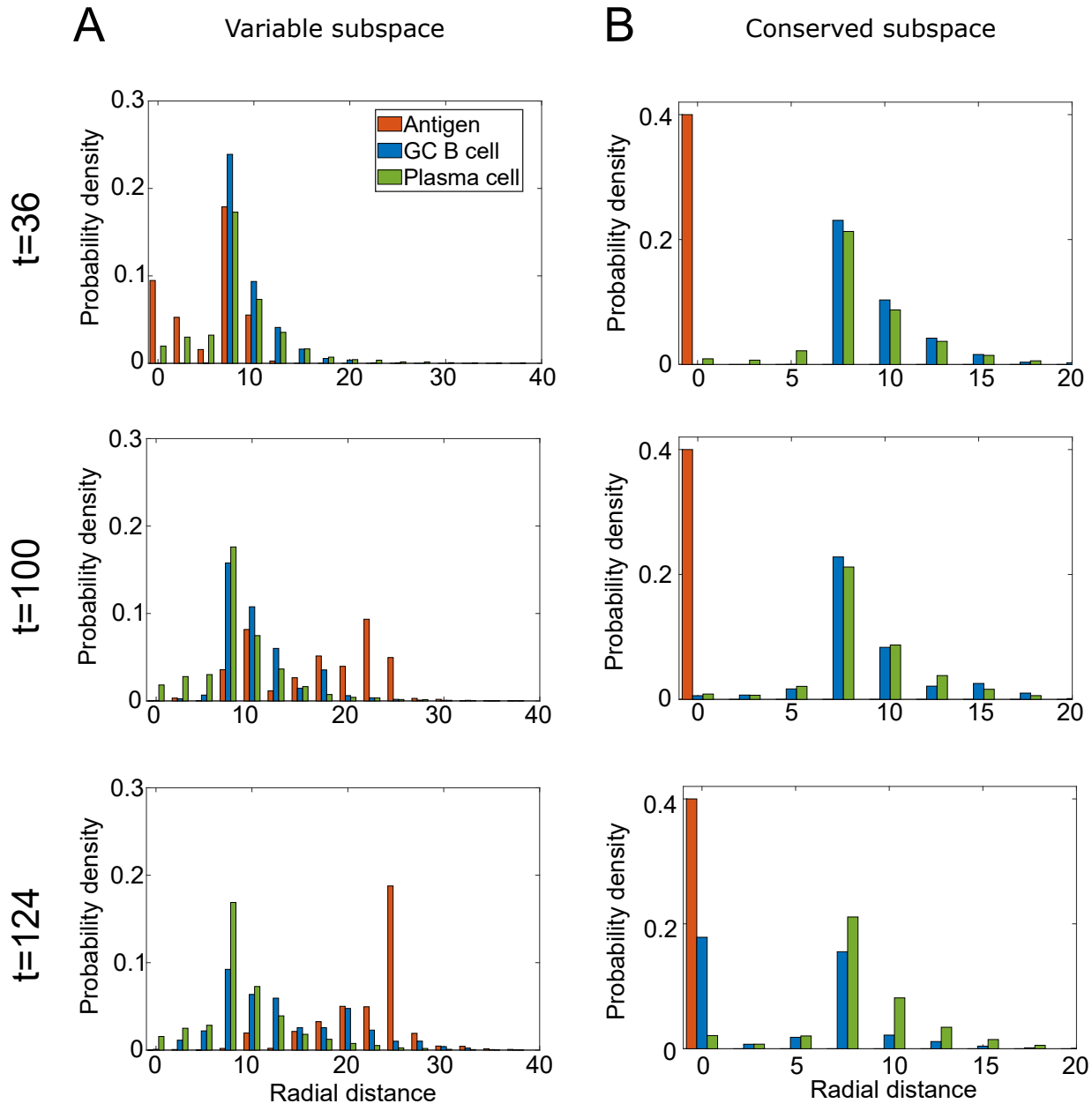

Figure 2.13: **Radial distribution of Ag, GC B cells and plasma cells in the rebound phase.** Ag population goes through the bottleneck ($t = 36$), recovers to a high level ($t = 100$), and finally declines toward extinction ($t = 124$). Ag (red) are capable of evading elimination by GC B cells (blue) in the variable subspace (panel A) but remaining at the origin in the conserved subspace (panel B); the distribution of plasma cells (green) has hardly changed in the variable subspace over this course. Same simulation as that shown in Fig.2.3E and the same variable subspace as that in Fig.2.3H. $n_c = 3$, $\sigma_A = 2$. The same legend applies to all panels.

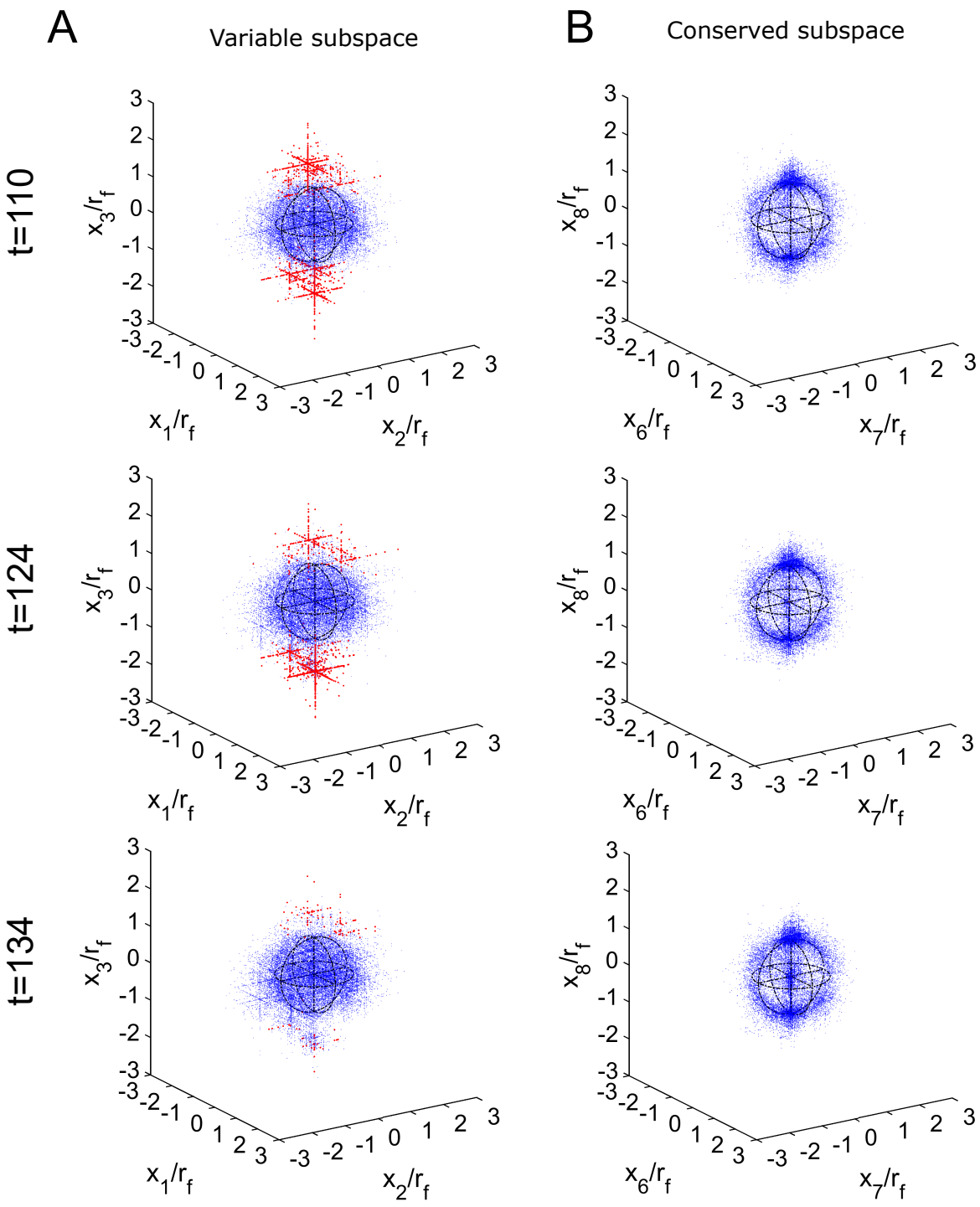

Figure 2.14: **Shape-space snapshots in the rebound phase, after recovering from population bottleneck.** Ag are eventually cleared as newly generated plasma cells converge to the origin in the conserved subspace. Data are collected from the same simulations as those in Figs.2.3E and 2.3H. Panel A/B: configurations of Ag (red) and plasma cells (blue) in a fully variable/conserved subspace. Top row: at $t = 110$, Ag population peaks following recovery from the bottleneck; middle: at $t = 124$, Ag population just falls below %50 of the peak value; bottom: Ag are close to extinction. Here $n_c = 3$, $\sigma_A = 2$.

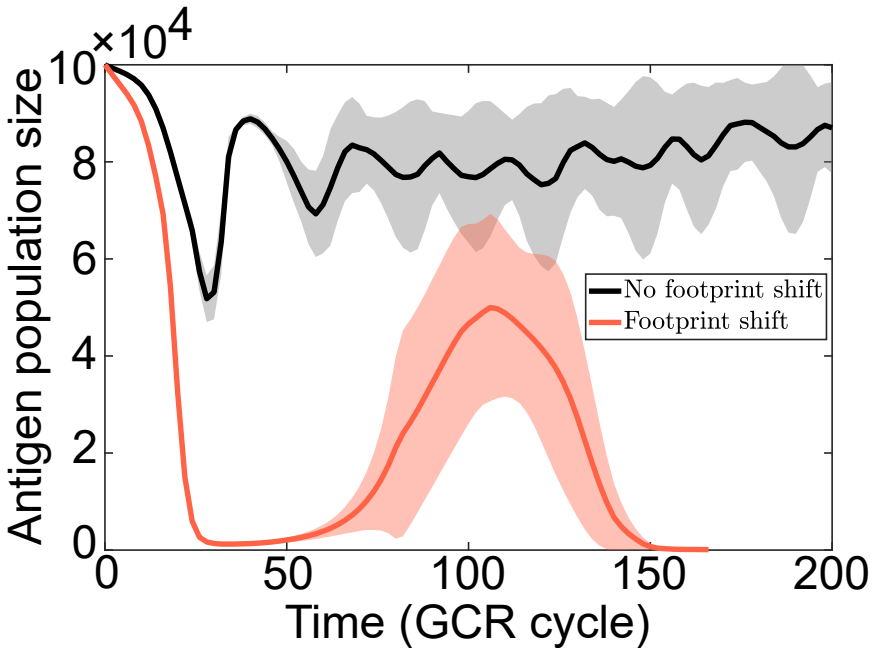

Figure 2.15: **Ag population trajectories in the rebound phase in the presence and absence of footprint shift.** Without footprint shift, Ag removal by plasma cells becomes ineffective, resulting in viral persistence instead. Each curve depicts mean±SD over 5 simulations. $n_c = 4, \sigma_A = 2$.

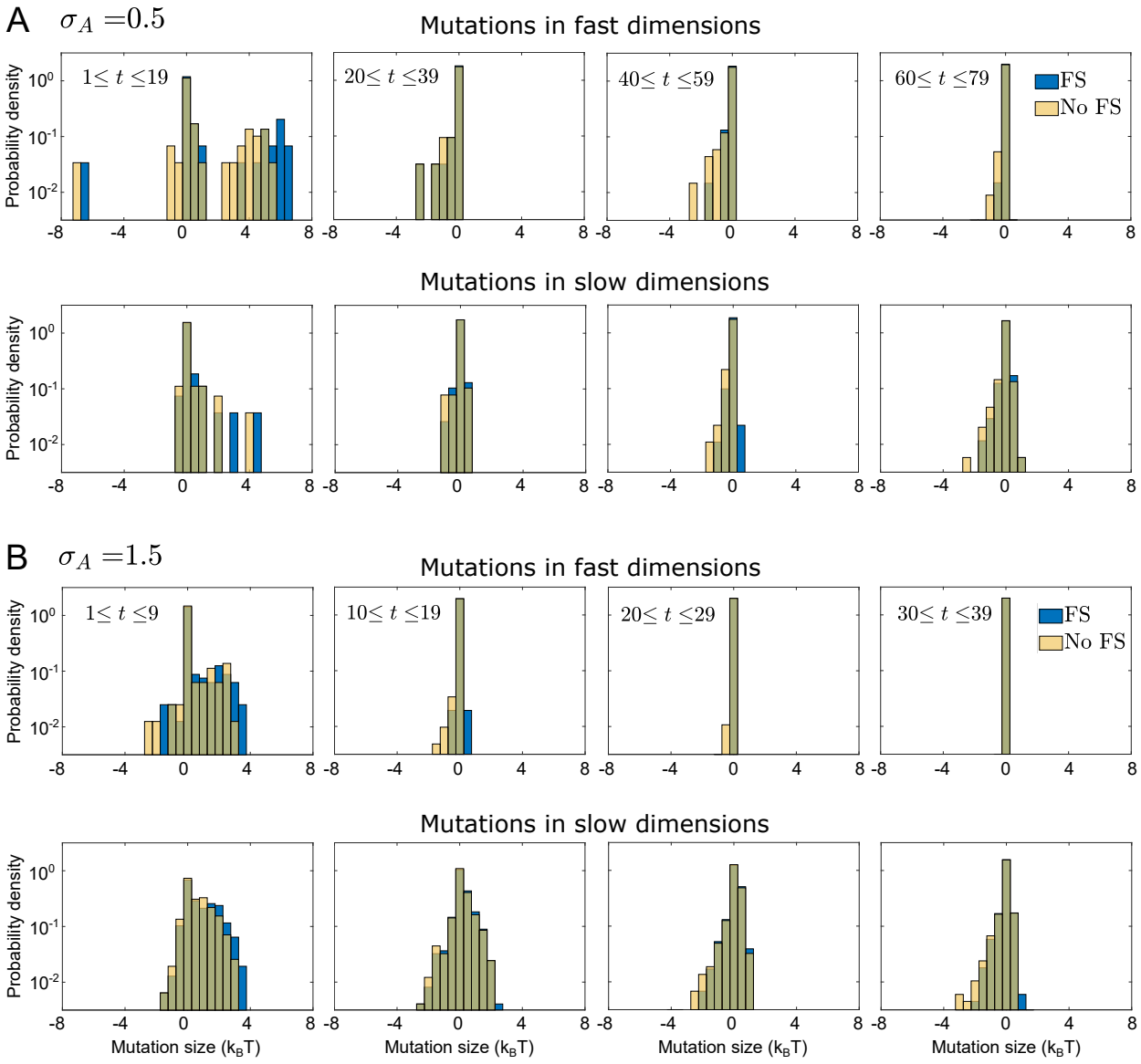

Figure 2.16: **Time-dependent distribution of mutation effect in fast (variable) and slow (conserved) dimensions.** Beneficial BCR mutations remain available in slow dimensions but are quickly depleted in fast dimensions. Mutation size is measured by the difference in log-mean-exponential affinity of a B cell to its encountered FDC Ag before and after a mutation (see Appendices text Section 5); a positive value indicates a beneficial mutation. In the case of a fixed footprint ("No FS", yellow), binding footprint is unchanged following a mutation regardless of its potential effect. Affinity-optimizing footprint shift enables the usage of mutations outside current contact and augments the beneficial effect ("FS", blue). Data are collected from long-lived lineages descending from B-type founders in repeated simulations, like in Fig.2.8. Here, $n_c = 4$, $\sigma_A = 0.5$ in (A) and $\sigma_A = 1.5$ in (B). The same color legend applies to all panels.

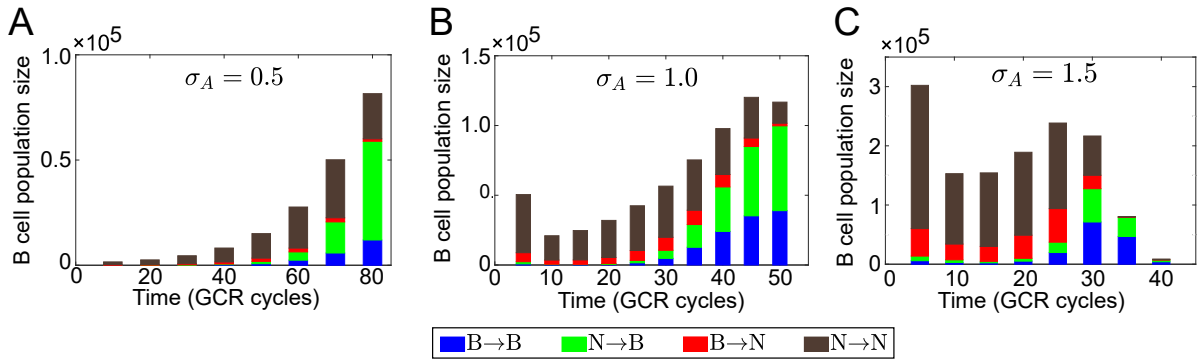

Figure 2.17: **Time evolution of absolute B cell count in each of four types of identity switch.** Switch types are defined based on the initial and current identity of individual cells, e.g. N→B represents current B-type clones descending from an N-type germline ancestor. From the same simulations in Fig.2.7B. $n_c = 4$ in all panels.

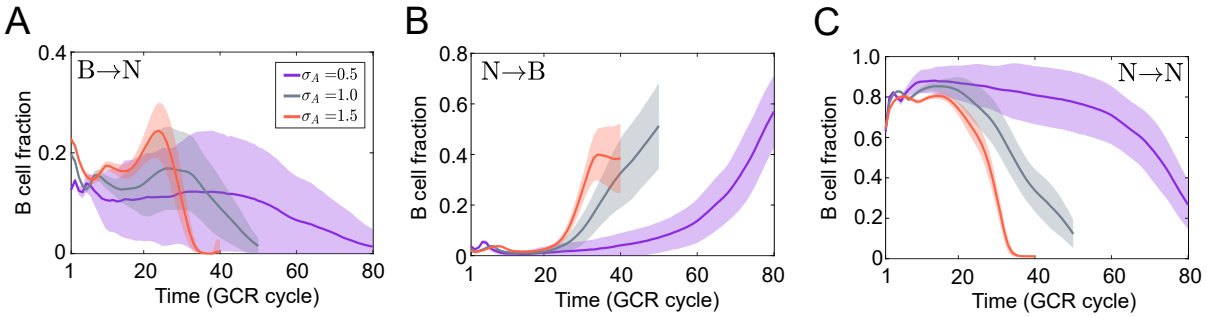

Figure 2.18: **Prevalence of B cells taking B→N, N→B and N→N pathways of identity switch.** From the same simulations as in Fig.2.7. $n_c = 4$. The same color legend applies to all panels.

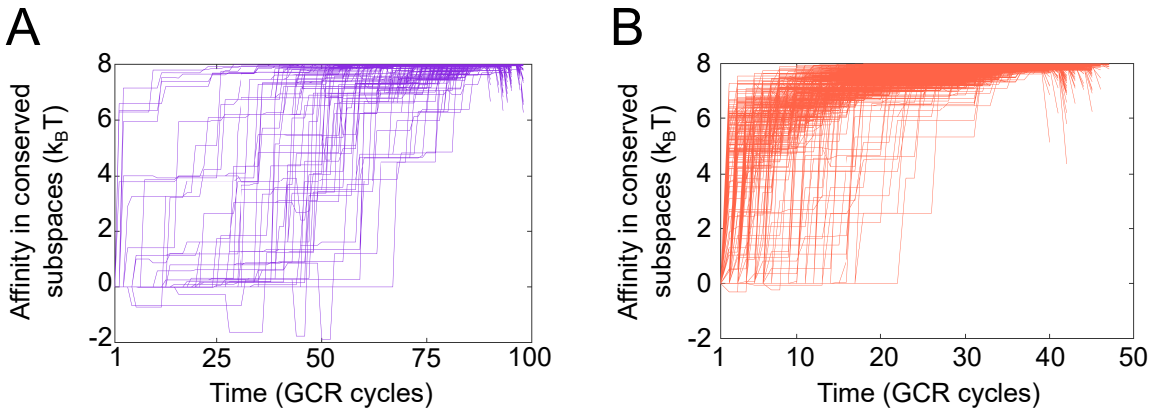

Figure 2.19: **Binding affinity to conserved targets along the B→B lineages.** (A) A modest initial Ag diversity ($\sigma_A = 0.5$) prolongs AM and permits a wide range of paths (tolerating occasional deleterious mutations). (B) A high initial Ag diversity ($\sigma_A = 1.5$) hastens maturation and yields a strong selection pressure that narrows the viable paths. Each trajectory represents a single line of B cells that both starts and ends as B-type. Data are from the same simulations as in Figs. 2.8G and 2.8H. $n_c = 4$.

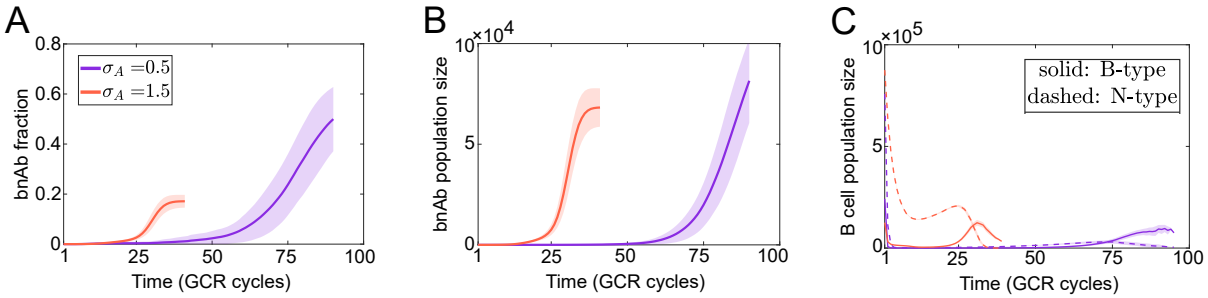

Figure 2.20: **Increasing initial Ag diversity speeds up accumulation of bnAbs and recovery of B-type cells, at the expense of hastened AM and reduced bnAb prevalence upon clearance.** Affinity threshold of bnAbs (binding fully conserved targets) is set to $6k_BT$. From the same simulations in Fig. 2.8C. $n_c = 4$. The same color legend applies to all panels.

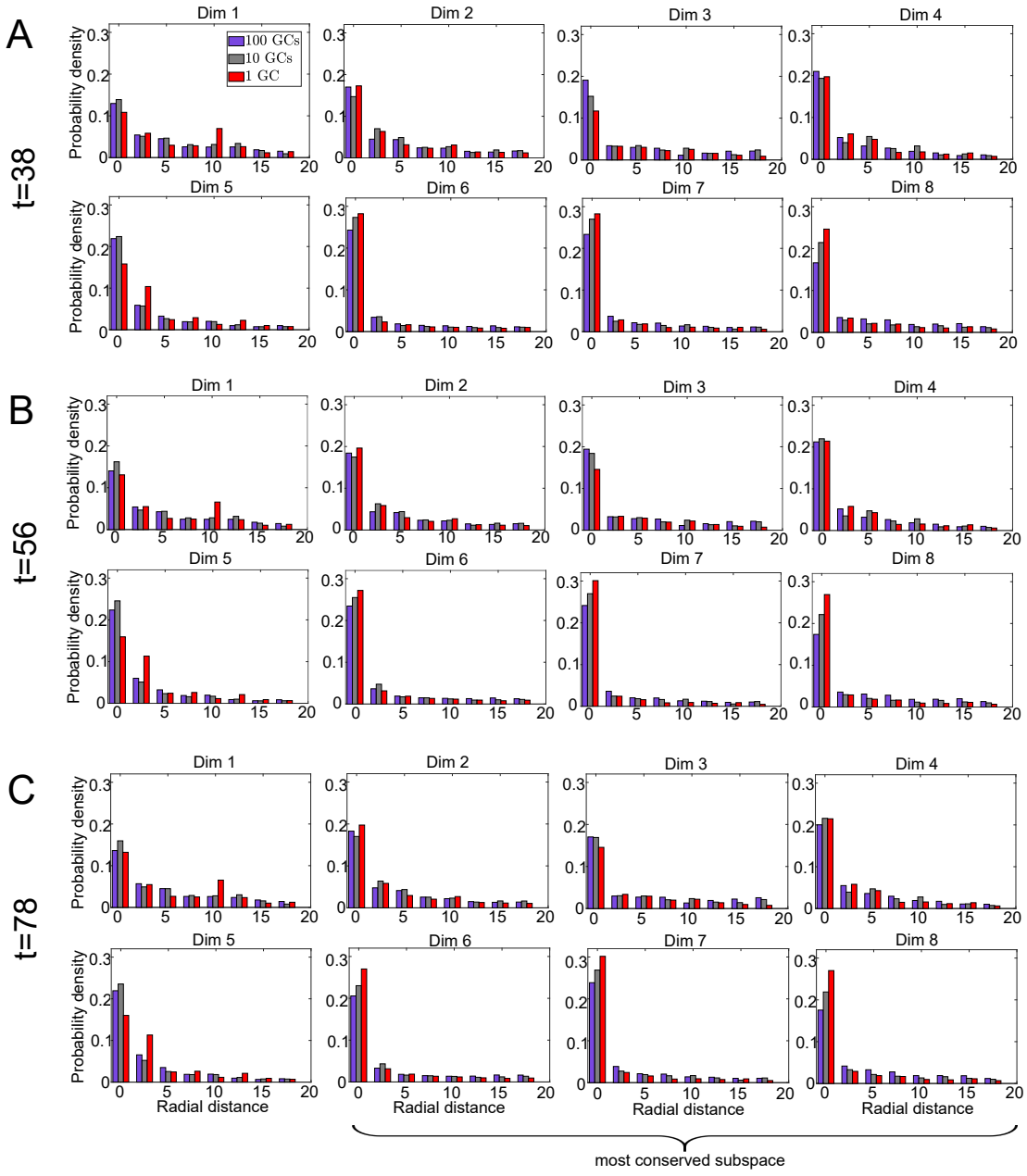

Figure 2.21: **Radial distribution of plasma cells during Ag population bottleneck at different levels of GC compartmentalization.** For fewer larger GCs (purple to red), plasma cells are more concentrated toward the origin of the most conserved subspace (dimension 6 to 8). Data are taken from the same simulations as in Fig. 2.9B; time points are at the beginning (A), the lowest point (B), and the recovery stage (C) of the Ag population bottleneck, defined as when population size falls below 10% capacity, reaches the minimum, and rises above 50% capacity, respectively, for one GC. Radial distance in dimension k is measured by x_k^2 for each plasma cell $\vec{x}$. $n_c = 2$, $\sigma_A = 1$. The same legend applies to all panels.

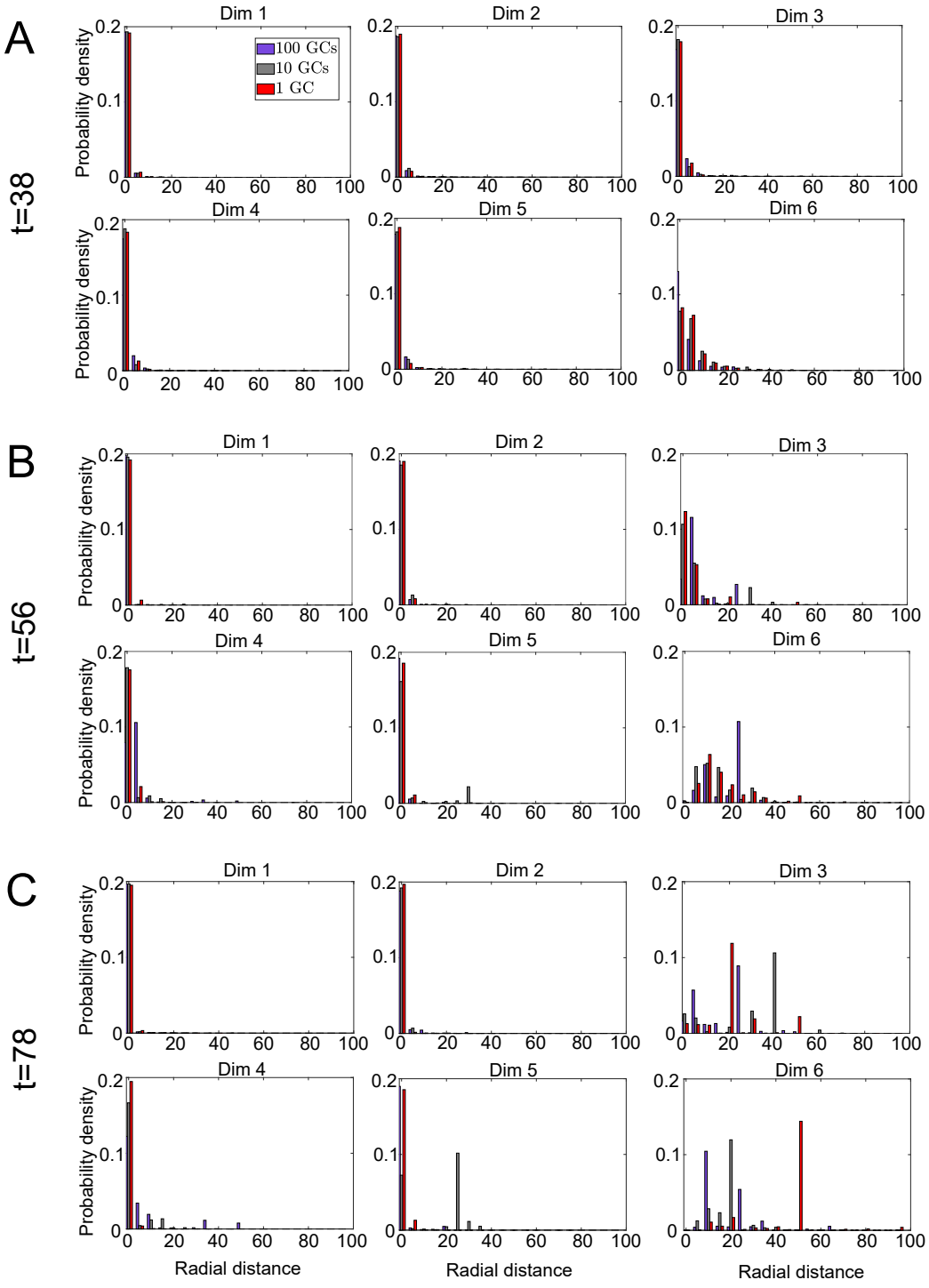

Figure 2.22: **Radial distribution of Ag during population bottleneck at different levels of GC compartmentalization.** For fewer larger GCs (purple to red), Ag are driven farther away from the origin in the variable dimension (dimension 6) of the most conserved subspace (dimension 6 to 8) during recovery from the population bottleneck. Data are taken from the same simulations as in Fig. 2.9B; time points in A to C are chosen like in Fig.2.21. Radial distance in dimension k is measured by y_k^2 for each circulating Ag $\vec{y}$. $n_c = 2$, $\sigma_A = 1$. The same legend applies to all panels.

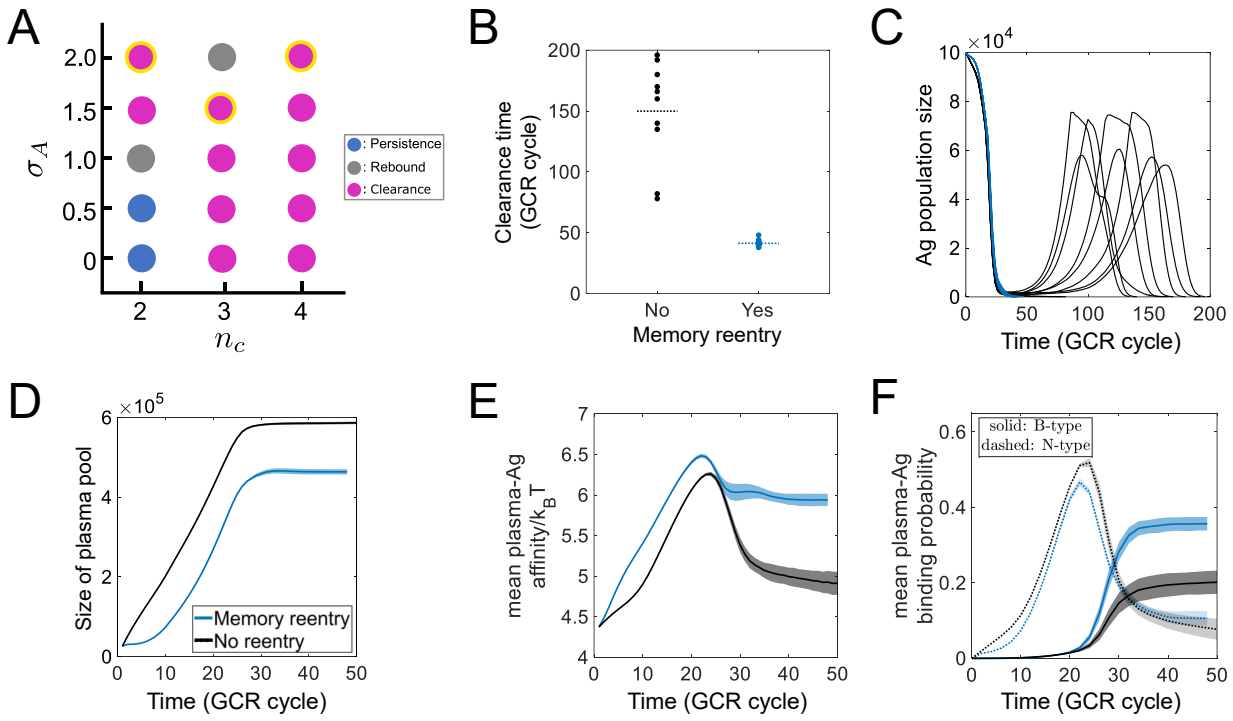

Figure 2.23: **Memory reentry into ongoing GCs promotes broad Ab response and turns viral rebound into rapid clearance.** (A) Coevolutionary outcomes with memory reentry. Yellow outlines highlight changes compared to Fig. 2.3A; memory reentry turns mixed outcomes into rapid clearance. (B) Time until viral clearance. Each symbol is an independent simulation; dashed lines indicate the mean value. (C) Ag population trajectories. Each curve is one simulation. (D, E) Overall quantity (D) and quality (E) of Ab-secreting cells (both plasma and memory cells are counted). (F) Efficacy of Ab response, measured by plasma-Ag binding probabilities P_B and P_N for broad (B-type) and narrow (N-type) antibodies, respectively. In (B–F), blue (black) indicates with (without) memory reentry. Data in (D–F) are plotted as mean ± std. Maximum memory reentry ($\alpha = 1$) in all panels. $n_c = 4$, $\sigma_A = 2$.

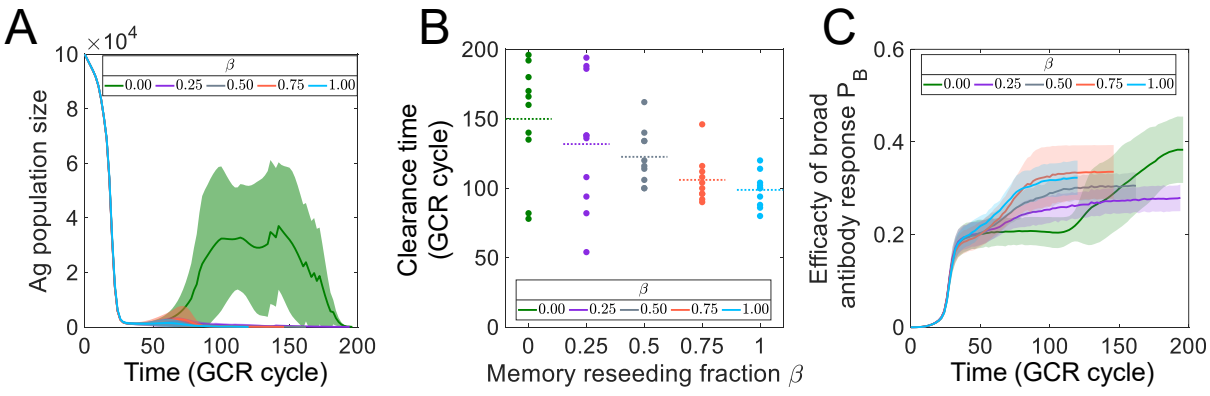

Figure 2.24: **Memory reseeding of new GCs results in a faster development of broad response and hence sooner viral clearance.** A higher fraction β of memory cells among GC founders leads to (A): reduced peak height or even prevention of Ag rebound: following the initial Ag population drop, Ag remains below 10% of capacity in 20% → 100% of simulations as β increases from 0 to 1 (not shown). In addition, higher β also leads to (B): earlier clearance of Ag, and (C) faster development of broad Ab response (measured by the plasma-Ag binding probability P_B). $n_c = 4$, $\sigma_A = 2$. Data in (A)(C) are plotted as mean ± std. Horizontal dashed lines in (B) show the mean value.

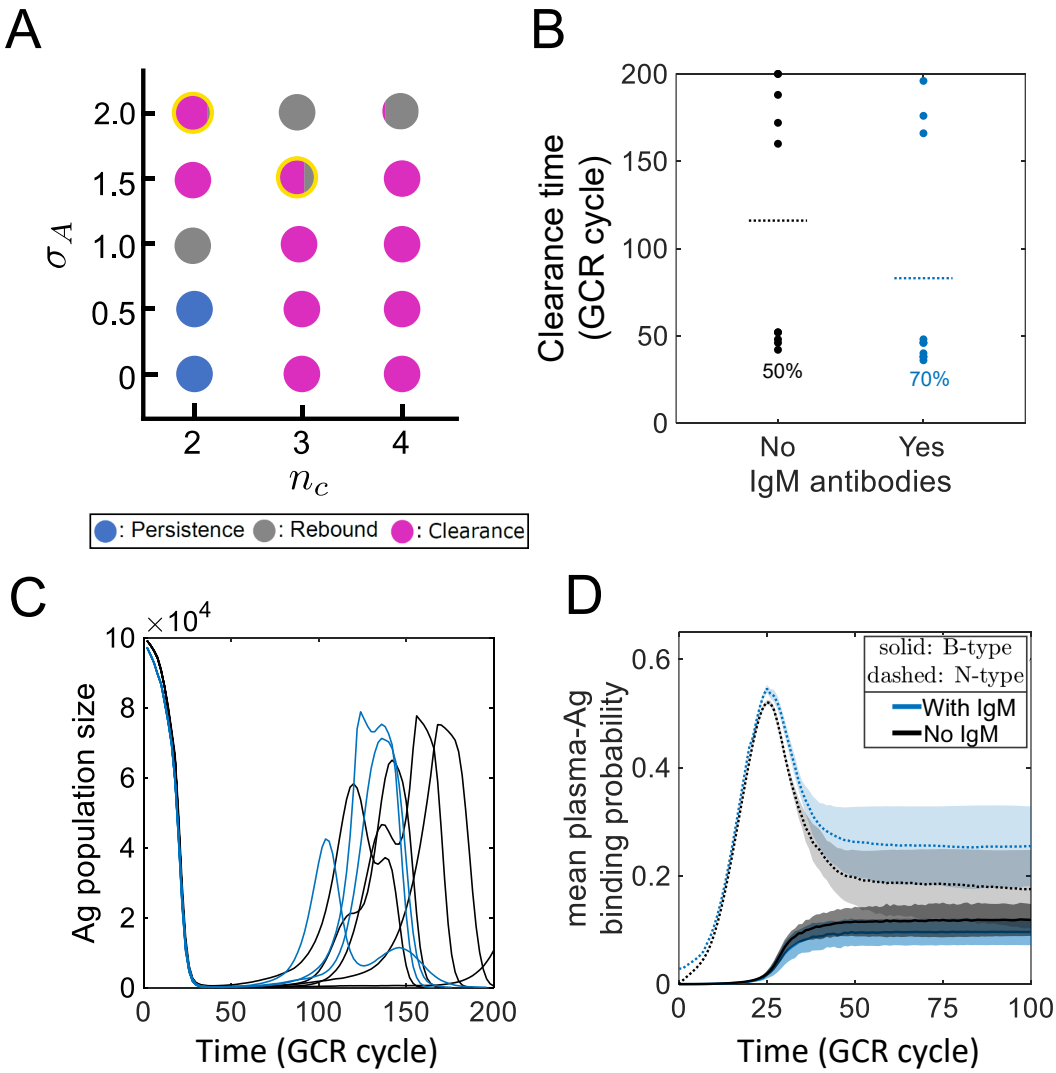

Figure 2.25: **IgM antibodies promote viral clearance by enhancing early removal and corralling escape mutants.** (A): Phase diagram of coevolution outcomes with IgM antibodies. Yellow marker outline highlights the change to coevolution outcome compared with Fig. 2.3A. (B) The time to Ag clearance is shorter with IgM antibodies. Horizontal dashed lines show the mean value averaged over an ensemble of simulations. The percentages label the fraction of simulations with clearance instead of rebound. Percentages label the fraction of simulations with clearance outcome. (C) Ag population trajectories. (D) Plasma-Ag binding probabilities from broad (P_B) and specific antibodies (P_N). Data in (D) are plotted as mean $\pm$ std. n_c=3, $\sigma_A = 1.5$ and the same color scheme is applied in (B-D).

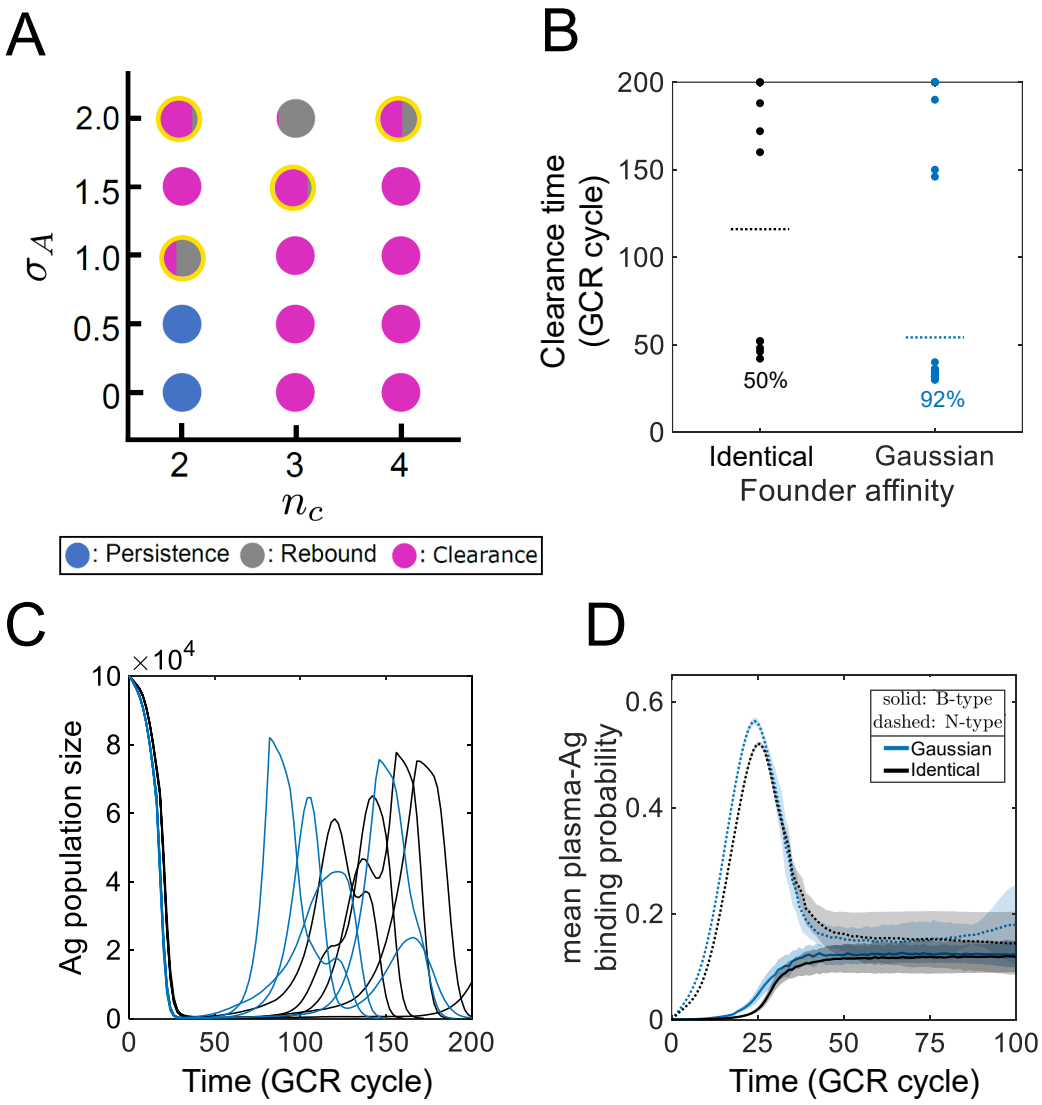

Figure 2.26: **A wider range of founder affinities speeds up the development of both broad and specific responses and, in turn, enhances viral clearance.** (A): Phase diagram of coevolution outcomes with Gaussian-distributed founder affinities. Yellow marker outlines highlight the changes to coevolution outcome compared with Fig. 2.3A. (B) The time to Ag clearance is shorter with Gaussian-distributed than identical founder affinities. Horizontal dashed lines show the mean value averaged over an ensemble of simulations. Percentages label the fraction of simulations with clearance outcome. (C) Ag population trajectories. (D) Both plasma-Ag binding probabilities from broad (P_B) and specific antibodies (P_N) grow faster with Gaussian-distributed founder affinities. Data in (D) are plotted as mean $\pm$ std. n_c=3, $\sigma_A = 1.5$ and the same color scheme is applied in (B-D).

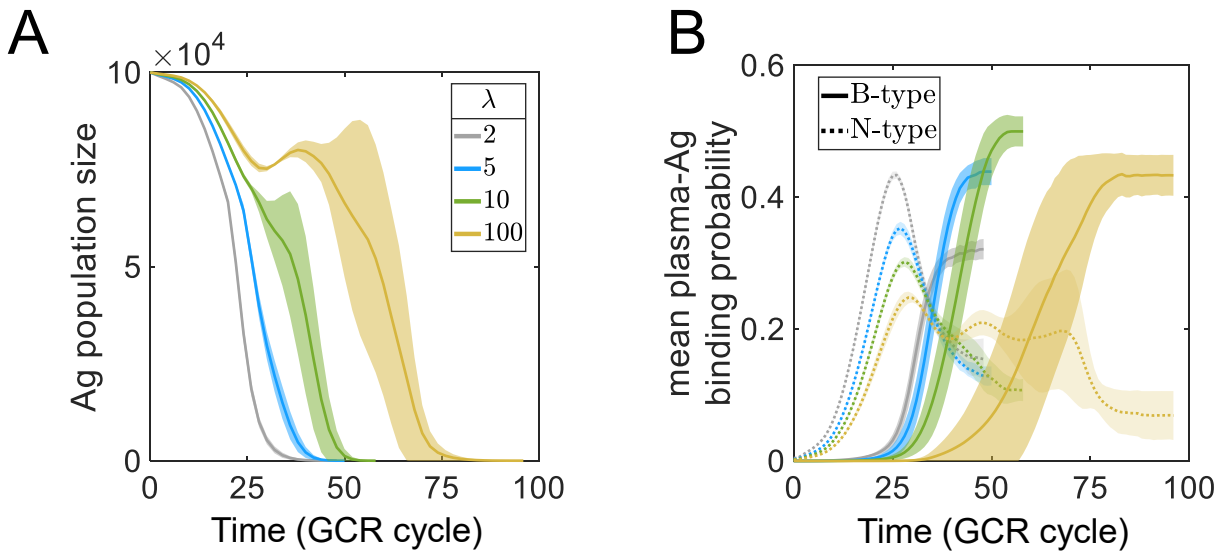

Figure 2.27: **Higher germline quality at sub-optimal binding sites against the founder virus (smaller λ) strengthens both broad and specific Ab responses and speeds viral clearance.** (A), (B) Time trajectories of Ag population size and plasma-Ag binding probabilities from broad (P_B) and specific antibodies (P_N) at different λ. Data are plotted as mean $\pm$ std and the same color scheme applies to both panels $n_c = 4$, $\sigma_A = 1.5$. In all other plots in this chapter $\lambda = 2$.

CHAPTER 3

Optimal vaccination to elicit bnAbs as dynamic control of eco-evolutionary adaptation

3.1 Abstract

Across diverse environments, generalist phenotypes remain potent, while specialist phenotypes are fitter than generalists in a specific environment but maladaptive in others. The development of generalists have great implications in disease prevention and treatment, but the optimal selection force to control the generalist-specialist competition remains unknown. Such is further complicated by the diverse interactions between a population's ecological and evolutionary dynamics, as both can act on comparable timescales. Here we explore the optimal control rules to promote generalists during eco-evolutionary adaptation. We first study the general principles to promote generalists in a minimal and biosystem-independent model. Then we establish a specific model focusing on vaccine elicitation of bnAbs against antigenically variable viruses. A trade-off between suppressing specialists and preserving generalists is identified, and we show that time-varying selection outperforms constant selection by pursuing both ends of the trade-off sequentially in time. We also demonstrate that optimally suppressing specialists can promote both generalist quantity as well as quality, thus sparing the need for two distinct controls. On the contrary, in intraspecies competition among generalist subtypes, a trade-off exists between generalists' quantity and quality, thus requiring different optimal control profiles. Overall, our findings highlight the importance of feedback between ecology and evolution, illustrate principles for guiding generalists' dynamics with time-varying selection in eco-evolutionary systems, and suggest new vaccination protocols to elicit bnAbs that are amenable to experimental testing.

3.2 Introduction

A biological system's adaptation to selection is shaped by competition among species ([157, 158]). In the face of environment heterogeneity, *generalists* remain adapted to diverse environmental conditions; *specialists*, on the contrary, are fit in a specific environment but are maladaptive in others. Prominent examples of generalists include bnAbs with high breadth against highly mutable viruses such as HIV ([159, 160, 161]), HCV ([14, 18]) or influenza([162, 163]), cancer cells/bacteria resistant to multiple drugs ([164, 165, 166]), multi-host phages ([167, 168]), etc. The evolution of generalists thus have notable implications in disease prevention and treatment, and the generalist-specialist competition has been under extensive investigation. However, the selection force that could be applied to optimally control the competition and develop generalists remains unclear.

A significant challenge in developing generalists is that they are "Jack of all trades but master-of-none": as a trade-off for generalists' robust performance across diverse environments, in any particular environment they are outperformed by specialists. Such performance-robustness trade-off has been observed in different biological systems, spanning cells ([169, 23]), viruses ([170]), bacteria ([171]), animals ([172, 173, 174, 175]), plants ([176]), and the source of the trade-off has been attributed to genetic constraints including genetic-based fitness trade-off (antagonistic pleiotropy, [177, 178, 179]) or a maintenance cost associated with generalism ([180, 181]). To counter this challenge, diverse environmental conditions, either temporally or spatially, has been applied to promote generalists. The outcomes have been mixed across different biosystems ([182, 183]), suggesting that additional factors should be considered when guiding generalist evolution.

In addition to the genetic-based trade-off, the coupling between evolution and ecology complicates competition dynamics and its control, especially when the evolutionary and ecological branches act on comparable time scales ([184, 185, 186, 187]). In such an eco-evolutionary system, diverse feedback pathways exist between eco- and evo- branches. First, density-dependent selection directly couples selection force with population abundance. Second, a higher population abundance accelerates evolution due to higher population mutation

rate ([188]) and more efficient elimination of deleterious mutations([189, 176]); Third, the range of acceptable population abundance places constraints on viable selection force when controlling a biological system. The generalist abundance should be kept above extinction threshold at all times in order to ensure survival.

However, previous ecological or evolutionary studies have mostly assumed a separation of timescales where eco- or evo- branch acts rapidly, while the other remains at (quasi) equilibrium. Population genetics focuses on different genotypes' relative abundance in a population, while the total population size remains constant ([48, 49, 50, 51, 52, 53]); ecology studies take a "gradualist" view instead ([190]), and considers evolution to be negligible when studying population abundance ([191, 192]). On the contrary, the eco-evolutionary dynamics have only been appreciated recently([193, 194, 195]). For example, it has been increasingly recognized in the last three decades that evolution can occur sufficiently fast in a declining population to allow recovery and save the population from extinction–a phenomenon known as "environmental rescue" ([196, 190, 197]). Such is the case of B cell AM in GCs, a rapid Darwinian selection process to improve B cell binding with Ag and is crucial for the development of bnAbs. B cells survive AM as their binding strength can increase by three orders of magnitude ([198]) from mutation-selection cycles and will otherwise go through apoptosis in the absence of such evolution.

Control theory has been a useful tool in the field of biology and medicine for over half a century, and has facilitated human understanding of host-pathogen interactions and improvement of disease treatment ([199, 200, 40]). For bnAbs elicitation by controlling vaccine Ag, though, optimal control studies have been absent for two reasons. First, to cover the vast immunological space of possible Ab and Ag configurations, agent-based models are predominantly used ([24, 156, 59, 60]) which are incompatible with existing optimal control framework. In particular, while ordinary differential equations have successfully described B cell evolution against a single Ag with optimal control theory applied ([45]), the same have not been achieved if multiple Ag are presented to B cells: the multiple Ag are required to provide a selective advantage to bnAbs precursors, but at the same time they also significantly expands the state space of antibodies and Ag. Second, current vaccination protocols

involve multiple injections ("prime-boost" scheme,[55]) well separated in time (typically by months, [56, 57, 58]) and a new set of B cells are recruited and replicate to form new GCs after each injection. Therefore, the control over B cell dynamics during a continuous GC reaction is limited by control's low time resolution from discrete injections. Nevertheless, the continuous delivery of vaccine Ag during a single GC reaction has been demonstrated recently via osmotic pump technique ([63, 64]). This opens up a pathway towards more time-sensitive control of B cells' eco-evolutionary dynamics.

In this chapter, we seek to elucidate general principles on optimizing eco-evolutionary adaptation by optimally controlling the selection strength, and also explore the application to vaccination aiming for bnAbs. First, we introduce a minimal and generic model featuring eco-evolutionary dynamics of generalist-specialist competition. Next, we focus on the context of B cell AM against diverse viral mutants and establish a specific model based on its biological details. The antigenic distance between vaccine virus strains (a proxy for selection strength) is treated as control. By transforming the underlying agent-based model architecture into a differential-equation description, we have successfully applied optimal control theory to bnAbs elicitation by vaccination. To the best of our knowledge, this is the first study where this is achieved.

We demonstrate that constant control is subject to a trade-off between preserving generalists and suppressing specialists. The optimal time-varying control balances such trade-off by achieving both ends sequentially in time. We also show that during interspecies generalist-specialist competition, both generalist quantity and quality can be promoted through suppresing specialists and thus by similar control profiles, owing to eco-evolutionary feedback. However, during intraspecies competition among generalist subtypes, generalist quantity and quality are subject to a trade-off and each would be optimized by different control profiles.

In all, our study highlights the importance of eco-evolutionary feedback and illustrate principles for optimally guiding the dynamics of eco-evolutionary systems. Our results also suggest potential new vaccination protocols to elicit bnAbs that are amenable to experimental testing.

3.3 Model

3.3.1 Controlling generalist-specialist competition: a generic model

The generic picture of generalist-specialist competition is illustrated in our model schematics Fig.3.1A and 3.1B. With low selection pressure in a homogeneous environment, specialists have a higher fitness than generalists as generalists' adaptation to heterogeneous environments ("generality") comes at a price of penalized fitness in a particular environment ([201, 202, 23, 182]). As a result, the generalists are outcompeted by specialists over time. Conversely, at high selection pressure in a heterogeneous environment, the absolute fitness of both species are lowered but generalists are less negatively affected and enjoy an increased relative fitness. Generalists can thus take over the specialists. Hence, by controlling the selection strength through environmental heterogeneity, one can selectively promote the generalist species.

We denote N_ℓ, Γ_ℓ, f_ℓ the abundance, mean quantitative trait and *per capita* birth rate of species $\ell = s$ (specialists) or g (generalists), respectively. The eco-evolutionary dynamics capturing the described control mechanism and generalist-specialist competition is given by the stochastic differential equations (SDE) adapted from [203]:

$$\begin{aligned}
\dot{N}_g &= \underbrace{[f_g(\Gamma_g) - d - cN_{\text{tot}}]N_g}_{\text{logistic growth}} - \underbrace{(1-\epsilon)uN_g}_{\text{ctrl. induced death}} \\
&\quad + \underbrace{\sqrt{N_g[f_g(\Gamma_g) + d + (1-\epsilon)u + cN_{\text{tot}}]}\xi_1}_{\text{demograph. noise}} \\
\dot{N}_s &= [f_s(\Gamma_s) - d - cN_{\text{tot}}]N_s - uN_s \\
&\quad + \sqrt{N_s[f_s(\Gamma_s) + d + \ \ u + cN_{\text{tot}}]}\xi_2 \\
\dot{\Gamma}_g &= \underbrace{2D_g \frac{\partial f_g}{\partial \Gamma_g} N_g}_{\text{selection}} + \underbrace{\sqrt{D_g}\xi_3}_{\text{genetic drift}} \\
\dot{\Gamma}_s &= 2D_s \frac{\partial f_s}{\partial \Gamma_s} N_s + \sqrt{D_s}\xi_4
\end{aligned} \tag{3.1}$$

Here $\{\xi_i\}$ are unit Gaussian white noises: $\langle \xi_i(t) \rangle = 0$, $\langle \xi_i(t)\xi_j(t') \rangle = \delta_{ij}\delta(t-t')$. The deterministic portion of ecological dynamics for N_ℓ depends on birth rate f_ℓ, baseline death

rate d, density-dependent death rate $cN_{\text{tot}} = c(N_s + N_g)$, and the control-induced death rate. The control-induced death rate is proportional to the control variable $0 \leq u(t) \leq u_{\max}$, which is the quantification of selection strength. Generalists' generality is measured by the constant $0 < \epsilon < 1$: larger ϵ means generalists are less susceptible to control u. The evolutionary dynamics of Γ_ℓ depends on the fitness landscape $f_\ell(\Gamma_\ell)$ ([203], also see Appendices), diffusion constants D_ℓ, and is accelerated by a larger population size N_ℓ for faster generation of mutations. To model the extinction of species ℓ, $\dot{N}_\ell$ and $\dot{\Gamma}_\ell$ are assigned zero by rule in Eq.3.1 when N_ℓ is less than extinction threshold N_c.

The effectiveness of the control function u: $[0, T] \to [0, u_{\max}]$ is measured by the ecological or evolutionary objective indices

$$\begin{aligned} J_{\text{eco}}\{u(t)\} &\equiv \int_0^T N_g(t) \frac{\mathrm{d}t}{T} \\ J_{\text{evo}}\{u(t)\} &\equiv \int_0^T f_g(t) \frac{\mathrm{d}t}{T} \end{aligned} \tag{3.2}$$

that focus on generalist quantity and quality, respectively. The integrand is also treated as zero when $N_g < N_c$ to penalize generalist extinction. Using Hamilton-Jacobi-Bellman (HJB) formalism the optimal control maximizing either objective can be solved numerically (see Methods). Also see Methods and Appendices for details on the SDE formulation and solutions for the stochastic optimal control.

3.3.2 Guiding affinity maturation (AM) towards generalists: a specific model

Guiding AM towards generalists: a specific model During AM of B cells in GCs, strain-specific B cells (specialists) bind variable Ag epitopes, and Ag mutations on these epitopes allow escape from strain-specific B cells. Cross-reactive B cells (generalists), on the contrary, target conserved Ag epitopes. These epitopes include structures that are functionally important to the virus and are shared by viral mutants ([10, 61, 204, 205, 206]). Therefore, the cross-reactive B cells remain potent against mutated viral strains. GC B cells are selected based on their binding affinity with Ag and compete for finite amount of T cell help ([1]). By controlling the composition of vaccine viral strains, we seek to frustrate the Ag binding of specialists and promote the evolution and expansion of generalists.

3.3.2.1 Vaccination protocol

The vaccination begins with a germline-targeting viral strain ("prime" strain) injected to activate naive B cells as GC founders. This is followed by diverse viral strains ("boost" strains) registered to guide B cell AM towards bnAbs ([207, 160, 208]), either through repeated injections or osmotic pumps. We consider a variable and a conserved epitope on vaccine Ag as B cell binding targets, and B cells targeting the two epitopes are the "specialists" and "generalists", respectively (Fig.3.1C). The conserved epitope contains conserved residues that are shared among viral strains and are ideal target for bnAbs. These conserved residues are shielded by variable residues on the same epitope, which act as a distraction to the breadth development of generalists. Indeed, the germline version of mature bnAbs have been reported to target variable Ag residues and gradually acquire breadth during AM ([62]). We also assume the two epitopes are structurally distinct, so B cells do not switch their target epitope during AM ([209, 62, 18]).

3.3.2.2 String representation of B cells and viral strains

Based on [152, 46, 24], BCRs and viral epitopes are represented as strings (Fig. 3.10) of length $M = 20$ and alphabet size $K = 4$. A B cell is encoded by one string $\vec{b}$, while a viral strain is encoded by two strings: ($\vec{a}^v$,$\vec{a}^c$) for its variable and conserved epitopes, respectively. Each string site represents a residue group included in BCR-virus binding. To reflect the conserved residues common to viral strains, the last M_c sites of the conserved epitope string $\vec{a}^c$ are "conserved sites" and are shared among all strains. The rest $M_v = M - M_c$ sites of $\vec{a}^c$ as well as all M sites of the variable epitope string $\vec{a}^v$ are "variable sites", and may differ between viral strains.

3.3.2.3 Controling antigenic distance between boost and prime strains

We assume that the boost strains at time t are designed and synthesized so that each variable residue group is randomly mutated relative to the prime strain with probability $\frac{\Sigma(t)}{M}$ ($0 \leq \Sigma(t) \leq M$). Similar measure of antigenic difference has been applied to influenza data

in [210]. In string representation, each variable string site from boost strains has a random different letter than the prime strain with probability $\frac{\Sigma(t)}{M}$. The Hamming distance between the variable epitope string of prime and each boost strain follows a binomial distribution with expected value $M \times \frac{\Sigma(t)}{M} = \Sigma(t)$. Therefore, we call $\Sigma(t)$ the *antigenic distance* of boost strains and treat it as the control variable in vaccine design.

3.3.2.4 Affinity class description of GC B cells

In our string representation, BCR-virus binding affinity A is modeled as a function of BCR-virus string matches (see Methods), which facilitates an affinity class description of B cell: they are binned into affinity classes based on their string matches with the prime virus strain and all B cells from the same affinity class have the same affinity against the prime strain. In Appendices we also derive the average binding affinity between a B cell and boost viral strains, based on the B cell affinity class and the antigenic distance Σ. Thus the affinity classes describe not only BCR-prime binding, but also BCR-boost binding.

A considerable advantage of the affinity class description, compared with exhaustively tracking all B cell strings, is the greatly reduced number of variables (from thousands of B cell strings to dozens of affinity classes). This allows a tractable ordinary differential equation (ODE) description of B cell dynamics when evolving against diverse Ag strains in GC, which is also readily compatible with optimal control framework (see Methods). In Appendices we also support our results from ODEs with stochastic agent-based simulations.

3.3.2.5 Dynamics of GC B cells

We denote the specialists and generalist populations with vectors $\vec{x} \in \mathbb{R}^{M+1}$ and $\vec{y} \in \mathbb{R}^{(M_v+1)(M_c+1)}$, respectively. The components $\{x_i\}$ and $\{y_i\}$ are the number of specialists and generalists in affinity class labeled by index i. For specialists, $1 \leq i \leq M+1$ relates to string matches as $m = i-1$. For generalists, $1 \leq i \leq (M_v+1)(M_c+1)$ is a linear index of all generalist matching score pairs ($0 \leq m_v \leq M_v$, $0 \leq m_c \leq M_c$). As i runs from 1 to $(M_v+1)(M_c+1)$, $(m_v, m_c) = (0,0), (1,0), ..., (M_v, 0), (0,1), (1,1), ..., (M_v, 1), ..., (M_v - 1, M_c), (M_v, M_c)$.

The GC B cell dynamics is then written as (adapted from [45]):

$$\begin{aligned}\frac{\mathrm{d}s_i}{\mathrm{d}t} \equiv f_i^s = &\underbrace{k_p\, s_i\, \theta_i^s\, h_i^s(\Sigma) \cdot (2\mathfrak{M}_{ii}^s - 1)}_{\text{replication}} \\ &+ \underbrace{k_p \sum_{j\neq i} s_j\, \theta_j^s\, h_j^s(\Sigma) \cdot 2\mathfrak{M}_{ij}^s}_{\text{mutation}} \\ &- \underbrace{k_p\, s_i\, \theta_i^s\, [1 - h_i^s(\Sigma) + \frac{|\vec{x}| + |\vec{y}|}{K_B}]}_{\text{apoptosis}}\end{aligned} \tag{3.3}$$

where $s = x,\, y$ for specialists and generalists, respectively.

The first line on the right-hand side of Eq.3.3 is B cell *replication*. k_p is a rate constant that sets the timescale of GC reactions. We set $k_p = 1$ GC reaction cycle per time unit. Since GC B cells go through 2-4 cycles per day ([1]), a typical GC reaction that lasts for 25-50 days ([59]) would span $T \approx 100$ time unit. θ_i^s is a step function of s_i: $\theta_i^s = 1$ if $s_i \geq 1$ and $\theta_i^s = 0$ otherwise. It indicates the existence of B cells of species s in class i and prevents unrealistic growth from a fraction of a cell ([45, 211]). h_i^s is the probability for B cells of species $s = x,\, y$ and class i to internalize Ag presented on the follicular dendritic cells, which serves as B cell fitness. The expression of h_i^s is given by the Langmuir isotherm:

$$h_i^s(\Sigma) = \frac{C_{\mathrm{Ag}}(t)\, \mathrm{e}^{\langle A_i^s(\Sigma)\rangle}}{1 + C_{\mathrm{Ag}}(t)\, \mathrm{e}^{\langle A_i^s(\Sigma)\rangle}} \tag{3.4}$$

where $C_{\mathrm{Ag}}(t)$ is the total concentration of all viral strains in GC and $\langle A_i^s(\Sigma)\rangle$ is the log-mean-exponential binding affinity between B cell of class i and boost strains of antigenic distance Σ. The fitness h_i^s saturates at 1 when the enumerator of Eq.3.4 is large, which will weaken the affinity discrimination of different B cell classes. Thus, at lower C_{Ag} the fitness h_i^s has a stronger dependence on affinity $\langle A_i^s(\Sigma)\rangle$ and results in a stronger selection for high-affinity B cells (Fig. 3.11), which is consistent with clinical observations ([212, 213]). The expression for $\langle A_i^s(\Sigma)\rangle$ is derived in Appendices. In particular, as Σ increases the boost strains become more distinct than the prime strain. Therefore B cells that bind strongly with the variable residues of the prime strain (large m for specialists or large m_v for generalists) are frustrated while those bind strongly with the conserved residues (large m_c for generalists) remain fit (Fig. 3.12). Thus, a larger control $\Sigma(t)$ corresponds to stronger selection pressure and grants

a stronger selective advantage to generalists targeting conserved residues. Matrix $\mathfrak{M}^s$ is the mutation matrix and its element $\mathfrak{M}^s_{ij}$ is the probability for a daughter B cell of species s to be class i if its mother cell is class j (see Appendices for the construction of matrix $\mathfrak{M}^s$). $i = j$ represents cell replication and on average $2\mathfrak{M}^s_{ii}$ daughter cells of class i are created while 1 mother cell of class i is lost.

The second line of Eq.3.3 is B cell *mutation*, where daughter B cells of class i are produced when mother B cells of class $j \neq i$ divide.

The third line of Eq.3.3 is B cell *apoptosis*, where $1 - h^s_i$ is the probability for B cells' failure to internalize Ag and receive survival signal in GC reaction. The density-dependent death term $(|\vec{x}| + |\vec{y}|)/K_B$ models B cells' competition for finite T cell help by imposing a soft capacity constraint. $|\vec{s}| \equiv \sum_i s_i \theta^s_i$ $(s = x, y)$ is the 1-d norm and constant K_B is the capacity.

3.3.2.6 Evaluation of generalist performance

To evaluate the generalist performance from different antigenic distance $\Sigma(t)$, we define an objective index Φ that accounts for both generalists' ecology and evolution during GC reaction:

$$\begin{aligned}\Phi\{\Sigma(t)\} = \gamma \underbrace{\int_0^T \frac{\sum_{m_{c,i}\geq 1} y_i}{K_B} \frac{\mathrm{d}t}{T}}_{\Phi_{\text{eco}}} \\ + (1-\gamma) \underbrace{\int_0^T \frac{\sum_{m_{c,i}\geq 1} y_i * \mathrm{e}^{A^c(m_{c,i})}}{|\vec{y}| * \mathrm{e}^{A^c(M_c)}} \frac{\mathrm{d}t}{T}}_{\Phi_{\text{evo}}}\end{aligned} \tag{3.5}$$

The first integral Φ_{eco} measures generalist quantity with at least minimal conserved binding $m_{c,i} \geq 1$. The second integral Φ_{evo} quantifies average generalist cross-reactivity where generalists are weighted by the exponential of their conserved-binding affinity $A^c(m_{c,i}) = \frac{(m_{c,i}/M)^C M_c}{M} E$ (Eq.3.10). Note that for Φ_{evo}, generalists are effectively binned into "breadth classes" as only the conserved binding score m_c is distinguished between generalists while variable binding score m_v is not. Parameter $0 \leq \gamma \leq 1$ tunes the focus of Φ between generalist quantity or quality: $\gamma = 0$ is entirely quality-/evo-oriented while $\gamma = 1$ is entirely

quantity-/eco-oriented. Integral form is chosen for Φ as B cells constantly differentiate into memory and plasma cells and exit GC during AM ([1]). We observe qualitatively similar results if terminal values are evaluated instead.

3.4 Results

3.4.1 Insights from the generic model

3.4.1.1 Dynamic control balances generalist preservation and specialist suppression

We solve for the optimal control $u^*(N_s, N_g, f_s, f_g, t)$ maximizing generalist ecological or evolutionary indices J_{eco}, J_{evo} (Eq.3.2) at different generality ϵ. Due to the linear control dynamics in Eq.3.1, the optimal control is bang-bang: u^* is either 0 or $u_{\max}$ at all time t (also see Appendices). We then performed ensembles of 1000 SDE simulations at different ϵ following the optimal control rule u^* (Fig.3.2A).

Applying positive control have two opposing direct and indirect effects on generalist growth. On one hand, control directly lowers the net growth rates of both generalists and specialists. On the other hand, via suppressing specialists, control also indirectly benefits generalists by reducing the competition they face. Correspondingly, a trade-off between preserving generalists and suppressing specialists can be identified. The optimal control to balance the trade-off is modulated by generality ϵ, as the control's direct (and undesired) effect on generalist growth is weaker at higher ϵ. As a result, different optimal control behaviors can be observed depending on the value of ϵ in Fig.3.2A.

First, at low generality (Fig.3.2A left, $\epsilon = 0.1$) the control remains off. In this case it is not worthwhile to apply control to lower specialist fitness, when the control hurts generalists nearly equally. Next, at high generality (Fig.3.2A right, $\epsilon = 0.9$) the control remains on until specialist are driven extinct. In this case the generalists are nearly immune to control, and the benefit from suppressing specialists outweighs the direct cost to generalist fitness. Finally, at intermediate generality (Fig.3.2A middle, $\epsilon = 0.5$) the control is initially on, but switches

off when both species are still present. This optimal switching time best promotes generalist survival (Fig.3.2B), and delays generalist extinction if it occurs (Fig. 3.9). Specifically, when the control is switched off late (Fig. 3.9, purple), excessively lowered generalist fitness leads to early generalist extinction. On the contrary, when the control is switched off early (Fig. 3.9, green), specialists benefit from lack of control and generalists are gradually outcompeted by specialists. A balance between generalist survival and specialist suppression is thus struck at the optimal switching time and best promotes generalist performance.

With the exception of low generality, the optimal control profile is time-varying to balance the preservation of generalists and suppression of specialists. This is also confirmed in Fig.3.2C where the ensemble average of objective index J_{eco} at optimal control outperforms constant controls.

3.4.1.2 Generalist quantity and quality are both improved by suppressing specialists

Due to the positive feedback between population growth and trait evolution (Eq.3.1), improving one process also promotes the other. Therefore, when the control optimally suppresses specialists and reduce competition faced by generalists, both generalist quantity and quality benefit. Indeed, when the optimal control aims to maximize generalist evolutionary index J_{evo} (Fig.3.2A, dashed black), the same types of control behaviors as those maximizing J_{eco} (Fig.3.2A, solid black) are observed. This in turn leads to nearly identical generalist performance between control profiles maximizing generalist quantity or quality (Fig.3.2D).

3.4.2 Promoting generalists with diverse viral strains during AM

3.4.2.1 Preserving generalists and suppressing specialists: a trade-off at constant control

Now we focus on the specific context of AM of B cells during GC reactions. Fig.3.3 shows the generalist performance when constant antigenic distance Σ is applied during GC reaction.

This serves as a reference scenario when a single vaccine recipe is registered throughout a GC reaction, when consecutive vaccine shots are temporally separated. ([59, 60]). The optimal generalist performance is achieved at an intermediate level of Σ (Fig.3.3A). At low $\Sigma \lesssim 6$, specialists outcompete generalists and dominate the GC (e.g., Fig.3.3C, $\Sigma \equiv 0$). Therefore both generalists quantity and quality are poor. Such outcome is qualitatively consistent with clinical observation from early HIV infections (less than 3 years post infection), when the Ab response is predominantly strain-specific before further virus diversification has taken place ([16, 214, 209]). As Σ increases, it suppress the specialists and provides a selective advantage to generalists, leading to a generalist response that dominates GC (e.g., Fig.3.3C, $\Sigma \equiv 9$). At even higher $\Sigma \gtrsim 12$, however, both specialists and generalists are suppressed by the strong selection pressure. This could result in an early extinction of generalists (e.g., Fig.3.3C, $\Sigma \equiv 16$) before any further generalist evolution and expansion can take place.

The distinct outcomes from different constant antigenic distance Σ reflect the trade-off we identified from the generic model to launch a successful generalist response. Specifically, first, the specialists are to be prevented from outcompeting generalists in the same GC. This calls for a high antigenic distance Σ to frustrate the specialists and provide a selective advantage to the generalists. Second, generalist founders are also susceptible to frustration from $\Sigma > 0$ and generalists' survival must be ensured so that they can eventually develop high cross-reactivity. This calls for a low antigenic distance Σ to keep generalists from early extinction. The opposing needs from both conditions thus leads to the intermediate optimal level of constant Σ.

3.4.2.2 Low-to-high control both preserves generalists and suppresses specialists sequentially

In order to achieve a better generalist performance than the trade-off at constant antigenic distance Σ would allow, we propose that a time-varying $\Sigma(t)$ function could be applied during GC reaction. In particular, a desirable vaccine scheme should start with low $\Sigma(t)$ to preserve generalists and allow their cross-reactivity m_c to evolve; this is followed by high

$\Sigma(t)$ to suppress specialists and allow further generalist evolution and expansion.

To test this idea, we design a simple $\Sigma(t)$ function that sharply switches from 0 to $\Sigma_{\max} > 0$ at $t = t_1$: $\Sigma(t < t_1) = 0$, $\Sigma(t \geq t_1) = \Sigma_{\max}$. The diamond in Fig.3.3A shows the result from optimal pair $(\Sigma^*_{\max}, t^*_1)$=(16, 20) that maximizes Φ_{evo}. As expected, a better performance than constant Σ is achieved (Fig.3.3A, Fig.3.14). Compared with constant Σ, dynamic selection pressure allows for better generalist preservation during the low-Σ period (Fig.3.3B, vertical axis). This then feedbacks onto faster evolution of generalist quality by faster accumulation of beneficial mutations. Moreover, compared with the optimal constant control $\Sigma \equiv 9$, a higher antigenic distance $\Sigma_{\max} = 16$ could be applied in the dynamic scheme, even though constant $\Sigma \equiv 16$ would drive generalists to extinction. Altogether, with the dynamics $\Sigma(t)$, generalists gain an advantage in the competition and expand while the specialists are excluded at a rate faster than constant $\Sigma(t)$ could achieve (Fig.3.3C, right two panels).

3.4.2.3 Optimal control decided by founder quality, vaccine protocol duration, and emphasis on generalist quantity vs quality

Fig.3.4 shows the optimal antigenic distance $\Sigma^*(t)$ solved using Pontryagin's maximum principle (see Methods and Appendices for details). We identify five different shapes of optimal $\Sigma^*(t)$ functions (Fig.3.4A). Example $\Sigma^*(t)$ as well as B cell quantity and quality trajectories are showcased in Figs.3.4B and 3.4C. Of note, the low-to-high switching function is naturally recovered without making any *a priori* assumptions.

The different control shapes are determined by generalist founder quality m_c, vaccination protocol duration T, and also by the objective index (Φ_{evo} or Φ_{eco}) being maximized. Among these, generalist founders' quality m_c is the most prominent factor. When founder quality $m_c < 4$, the optimized antigenic distance $\Sigma^*(t)$ exhibits the signature low-to-high switching behavior (Fig.3.4A, triangles). This can be interpreted as when GCs are seeded by naive B cells with low cross-reactivity. However, if generalist founders already have high cross-reactivity $m_c \geq 4$ (Fig.3.4A upper row, squares), the optimal solution is to turn on control

from very beginning: when the generalist founders are already cross-reactive and resistant to high antigenic distance, strong selection can be applied early to select for high-quality generalists (high m_c) without endangering generalist survival. Such can be the case if the GCs are seeded by reactivated memory cells ([215]) that have already gone through previous rounds of AM and acquired some cross-reactivity.

The objective index (Φ_{evo} or Φ_{eco}) affects the shape of optimized antigenic distance $\Sigma^*(t)$ at late time, when t is close to the vaccine protocol duration T. When maximizing the evolutionary index Φ_{evo}, $\Sigma^*(t)$ remains at high level once switched on. Such sustained duration of high $\Sigma^*(t)$ selects for more cross-reactive generalists even when specialists are suppressed and no longer poses a threat to generalist survival (Fig.3.4C, solid), as competition remains among different generalist classes. When maximizing the ecological index Φ_{eco}, however, the selection pressure from diverse boost strains is relaxed after specialist suppression (Fig.3.4C dashed, Fig. 3.16) to allow for higher generalist quantity at the cost of quality. Comparing the AM outcome when maximizing either Φ_{eco} or Φ_{ev} (Fig.3.4C) though, the difference is small as both optimal control profiles promote generalist evolution and expansion through suppressing specialists. Also, the on switch for eco-control slightly lags behind evo-control to allow more generalist growth before the strong selection pressure is applied.

Vaccine protocol duration T affects the shape of optimal control in the limit of very short protocol duration duration and low founder quality. In this regime, the optimal antigenic distance to maximize Φ_{eco} is to apply minimal $\Sigma(t)$ throughout (Fig.3.4A lower panel, circles). The low quality generalists do not have a chance to prevail over specialists in such limited timespan, even with time-varying antigenic distance. If diverse viral strains are applied, the direct negative impact on generalist quantity will only outweigh the indirect benefit from suppressing specialists, thus leading to a lower generalist quantity.

3.4.2.4 Optimal control timespan balances generalist preservation and specialist suppression

From Fig.3.4, the optimal antigenic distance $\Sigma^*(t)$ exhibits two types of switching behaviors: a low-to-high ("on") switch for both evolutionary and ecological optimal controls, and a high-to-low ("off") switch only for ecological control. Two questions then rise: what decides the optimal on and off switching times, and how would the generalist performance degrade if the switching times deviate from their optimal values?

First, we look at the optimal on switch. In Fig.3.5 we demonstrate that the optimal on switch strikes a balance between generalist preservation and specialist suppression. We focus on the case of low generalist founder quality (Fig.3.5A blue, $m_c = 0$ for generalist founders) when the evolutionary objective Φ_{evo} is most sensitive to the timing t_1 of on switch (Fig.3.5A inset). An intermediate switching time leads to peak Φ_{evo}, whose value declines sharper with an earlier non-optimal timing than a later timing.

To elucidate what makes the peak position in Fig.3.5A optimal, we track the composition of generalists in time when the switching time is early, optimal, or late (Fig.3.5B). During GC reaction, generalist mutants of increasingly higher cross-reactivity m_c gradually emerge and the progression is driven by an "avalanche" of mutation flux originating from founder B cells (Fig.3.5B, purple to red). On one hand, compared with the optimal on time, earlier switching time suppresses generalists' quantity prematurely. This slows down their evolution to higher quality (Fig.3.5B upper and middle panels, $t = 20$) and subsequent expansion after the on switch. Too early a switching can even drive generalists to extinction (Fig.3.17A), which accounts for the sharp decline of Φ_{evo} at early t_1 in Fig.3.5A. On the other hand, if the on switch is later than optimal, the generalists will temporarily have a quantity advantage over the optimal switch (Fig.3.5B middle and lower panels, $20 \leq t \leq 40$). However, without the selection force for high cross-reactivity provided by high Σ, specialists gradually exclude generalists in competition. As a result, both generalist quantity and quality are eventually surpassed by the optimal $\Sigma^*(t)$ (Fig.3.5B middle and lower panels, $t \geq 60$).

In practice, it is essential to avoid the extinction of generalists due to applying high

antigenic distance prematurely. Here we provide an estimator for the earliest viable on switch that allows generalist survival. Such an estimator also provides a lower boundary for the optimal switching time (Fig.3.5C): while the generalists barely survive the population bottleneck with the earliest viable switch (Fig. 3.17B), a later optimal switch allows for more generalist evolution and expansion before switching and, in turn, a faster recovery from population bottleneck (Fig. 3.17C).

Such an estimator is based on the earliest emergence of *self-sustainable* generalist class from GC reaction. During the avalanche of mutation flux following the onset of GC reaction, a positive growth rate is required for the survival and subsequent evolution of the leading generalist class with highest cross-reactivity m_c. Based on Eq.3.3, we can classify the growth rate f_i^y of generalist class i into *intrinsic* and *extrinsic* growth rates $f_{i,\text{int}}^y$ and $f_{i,\text{ext}}^y$ (Eq.3.13): the *intrinsic* growth rate includes the replication and apoptosis (if failed to internalize Ag during GC reaction cycle) of generalists in class i, which are independent from other B cells present in GC. In contrast, the *extrinsic* growth rate includes the mutation influx from other generalist classes $j \neq i$ into class i, as well as apoptosis induced by competition. Thus, the intrinsic growth rate is only a property of generalist class index i, while the extrinsic growth rate depends on the entire GC B cell population. As B cell population size shrinks following the onset of GC reaction, both the mutation flux as well as the competition among surviving GC B cells weakens. Thus, we define the *self-sustainable* generalist class i by neglecting the *extrinsic* growth rate and require the *intrinsic growth rate* $f_{i,\text{int}}^y(y_i, \Sigma = \Sigma_{\max}) > 0$.

In turn, the emergence of *self-sustainable* generalists from evolution acts as our estimator for earliest viable on switch (Fig.3.5C): the control should remain off ($\Sigma(t) = 0$) until the *self-sustainable* generalists have evolved, and once evolved these generalists can rely on their self-replication for survival after the on switch. Such an estimator depends only on the existence of self-sustainable BCRs, thus it can be realized by isolating the plasma antibodies and test their binding affinity to boost strains ex vivo. In particular, there is no need to monitor the entire GC B cell population which is impractical. With parameters in Fig.3.5, the self-sustainability condition $f_{i,\text{int}}^y(y_i, \Sigma = \Sigma_{\max} = 10) > 0$ is first met for generalists with $m_c \geq 3$ and an immediate on switch at $t = 0$ is viable if $m_c \geq 2$ for generalist founders

(Fig.3.5C): self-sustainability is immediately met by the founders (when founder $m_c \geq 3$) or the first mutants of improved m_c (when founder $m_c = 2$).

Second, we shift our scope to the optimal off switch for ecological objective Φ_{eco}. Like for the on switch, it strikes the balance between generalist preservation and specialist suppression , and an intermediate off time leads to the peak of objective index (Fig.3.6A).

We start with vaccine protocol duration $T = 100$ which corresponds to the typical GC duration of 25-50 days. When the high antigenic distance is switched off at optimal timing (Fig.3.6B, middle panel), the generalists already dominate the GC and have reached saturation : the later is reflected by generalists' nearly identical per-cell birth and death rates (Fig.3.6C, middle panel; see Eq.3.14 in Methods for the equations of the two rates). After the high antigenic distance is switched off, the generalist abundance further increases briefly owing to relaxed selection, until a new saturation is reached. However, if the high antigenic distance is switched off earlier than optimal, generalists' advantage over specialists is lost prematurely and specialists will take over again (Figs.3.6B and 3.6C, left panels). On the contrary, when the high antigenic distance is switched off later than optimal (Figs.3.6B and 3.6C, right panels), the duration of relaxed selection after off switch is shorter than optimal, which results in a lower value of Φ_{eco}.

When the vaccine protocol duration T is too short for generalist abundance to reach saturation, the optimal off time becomes $t_2^* \lesssim T$ to maintain suppression of specialists and maximize generalist growth (Fig.3.6D). The high antigenic distance is only switched off near the terminal time T to relax selection and achieve a final spike of generalist growth. For very long T ($T \gtrsim 400$, or 100-200 days), the high antigenic distance is only applied until specialists are driven extinct (Fig.3.6D, horizontal dashed). This avoids specialists' recovery after off-switch while also allows relaxed selection on remaining generalists.

3.4.2.5 Co-promotion of generalist quantity and quality at intermediate level of competition

In Fig.3.4C, both generalist quantity and quality are promoted by similar optimal antigenic distance $\Sigma^*(t)$, which exhibits a switch from low $\Sigma(t) = 0$ to high $\Sigma(t) = \Sigma_{\max} = 10$ at nearly identical timings. In addition, as we will show below, the co-promotion of both generalist quantity and quality requires an intermediate competition strength, where the B cell capacity K_B in GC equals or is comparable with the initial B cell abundance 10^4 at GC reaction onset (Fig.3.7A). This is also the capacity regime we believe to be most realistic: a few founder B cells expand to an abundance limited by GC capacity before GC reaction cycle begins.

At intermediate competition strength, the optimal timings to switch on high antigenic distance are nearly identical (Fig.3.7A) for both eco- and evo- objectives. When the high $\Sigma(t)$ is switched on, the generalists suffer from a reduced birth rate (Fig.3.7B) due to increased selection strength. However, as generalist cross-reactivity improves, generalist birth rate gradually recovers over time. In addition, the reduction in generalist birth rate is also quickly compensated by the reduced death rate (birth rate exceeds death rate at $t = 38$), as specialists are excluded under strong selection and ease the competition faced by generalists. Overall, the indirect gain from the reduced death rate outweighs the direct loss from the reduced birth rate with optimal control, and the net generalist growth rate is improved compared with no control. To see this directly, one can compare the area between the birth and death rate curves (which gives the integral of net growth rate) with and without control (Fig.3.7B, solid v.s. dashed). The improved generalist quantity also positively feedbacks onto accumulation of breadth-improving mutations, and the generalists with higher m_c are selected by high $\Sigma(t)$, leading to co-improved generalist quality.

On the contrary, in weak competition regime (Fig.3.7A, $K_B \gtrsim 10^4$), the optimal timing to switch on high antigenic distance for eco-objective Φ_{eco} is postponed as capacity increases, while the on timing for evolutionary control remains unchanged. The delayed on switch of ecological control is because the competition only becomes relevant when the total B cell abundance has expanded to be comparable with capacity. In particular, in the limit of

extremely large capacity $K_B \geq 10^{24}$, the optimal ecological control simply becomes $\Sigma^*(t) \equiv 0$. This is the regime when the capacity K_B has even exceeded the unrealistic B cell abundance achieved in the absence of competition. If high Σ is applied for such large K_B, the reduced generalist death rate by reducing competition (which is nearly neglible) is simply insufficient to compensate for the drop in generalist birth rate (Fig.3.7C). For evolutionary control, however, a high Σ still selects for generalists of high cross-reactivity m_c, and the same early switching as intermediate K_B remains optimal for generalist quality.

Finally, in strong competition regime ($K_B \lesssim 10^4$), the generalist population quickly shrinks and the generalists cannot accumulate enough mutations to improve cross-reactivity m_c. The generalists remain susceptible to strong selection and the elevated death rate does not compensate the slow recovery of death rate if a high Σ is applied (Fig.3.7D dashed). Thus, the optimal controls for both eco- and evo- objectives are $\Sigma^*(t) \equiv 0$ (Fig.3.7A) to elongate generalist survival.

3.4.2.6 Intraspecies competition of generalists: quantity-quality trade-off

To better elicit bnAbs through vaccination, people have explored germline-targeting immunogen designs to activate naive B cells that target conserved viral epitope ([216, 217, 218]). If a GC is seeded by these generalist founders, then the *interspecies* competition between generalists and specialists is no longer a concern and AM will instead be driven by *intraspecies* competition among different generalist B cells. We demonstrate in Fig.3.8 that different control rules emerge in this regime as a trade-off now exists between generalist quantity and quality. During competition among different generalists, diverse vaccine strains frustrate the BCR binding with variable viral residues and preferentially selects for the generalists of higher cross-reactivity m_c. Thus, under constant antigenic distance the generalist quality increases with Σ by filtering out lower-quality generalists, which reduces generalist quantity: this results in the trade-off between generalist quantity and quality (negative slope in Fig.3.8A, in contrast with the positive slope in Fig.3.3B for generalist-specialist competition).

The generalists' quantity-quality trade off also applies to dynamic antigenic distances.

While a low-to-high switching $\Sigma(t)$ function still promotes generalist quality and outperform constant Σ (Fig.3.8A; Fig.3.19A), the same dynamic $\Sigma(t)$ function no longer also boasts a high generalist quantity compared with constant Σ functions (Fig.3.8A; Fig.3.19B). Moreover, the optimized antigenic distance $\Sigma^*(t)$ takes distinct shapes (Fig.3.8B) when the objective index Φ places different weight γ on generalist quantity (Eq.3.5). Specifically, a low-to-high switching $\Sigma^*(t)$ function still optimizes generalist quality (Fig.3.8B, red). However, $\Sigma^*(t) \equiv 0$ optimizes generalist quantity (Fig.3.8B, blue) as it allows free growth of generalists though the selection force towards higher cross-reactivity is absent. As a result, in generalist-generalist competition, $\Sigma^*(t)$ functions optimized for different eco-/evo-focus yield significantly different values of Φ_{eco} and Φ_{evo} (Fig.3.8C, black). In comparison, in generalist-specialist competition both Φ_{eco} and Φ_{evo} are almost invariant from optimized $\Sigma^*(t)$ of different eco-/evo- focus (Fig.3.8C, blue; see Fig. 3.20 for corresponding optimal $\Sigma^*(t)$ functions).

3.5 Discussion

In this chapter, we elucidate the optimal control principles to guide generalist-specialist competition in an eco-evolutionary system. We treat the selection pressure from environment heterogeneity as control, and start by finding the optimal control in a minimal and generic eco-evolutionary model of generalist-specialist competition without relying on specific biological context. Next we study in detail the AM of bnAbs during GC reaction, where the competition between cross-reactive and strain-specific B cells are guided by time-varying vaccine Ag.

In both models we demonstrate a trade-off between generalists' absolute and relative fitness. In the presence of environmental heterogeneity, the generalists are less negatively affected than specialists, and thus have increased relative fitness at the cost of lowered absolute fitness. This results in a counterintuitive strategy to promote generalist quantity: applying strong control can improve generalists' absolute fitness indirectly by suppressing specialist abundance and reduce competition, even though the control also directly harms generalists'

absolute fitness. The optimal timespan to apply control balances the two opposite effects, and is decided by the generalists' generality (robustness to environment heterogeneity): control lasts longer when the generality is higher. In addition, both generalist quantity and quality can be promoted by suppressing specialists due to the positive feedback between generalist expansion and evolution. Since our generic model is independent of the underlying biological details, we believe our conclusions are robust across biological systems.

Additional insights are gained from our GC reaction model, where we incorporate the biological realism of AM process and consider the antigenic distance $\Sigma(t)$ between prime and boost viruses as control. First, we model the evolution of generality (B cell cross-reactivity) over time, and this leads to two-way interactions with optimal control scheme. On one hand, the evolution of B cell cross-reactivity results in a signature on-switch of control. In particular, zero $\Sigma(t)$ is first applied to preserve generalists when their cross-reactivity is still low; a switch to high $\Sigma(t)$ then follows to suppress specialists, after the generalists have evolved to be self-sustainable at high $\Sigma(t)$. On the other hand, the optimal control leads to improved cross-reactivity over no-control scheme. That is, the robustness-performance trade-off of generalists itself evolves over time and is modified by the application of control. Second, we explicitly model the population structure of generalists and specialists, each containing individual cells of different affinity classes. We find that a high $\Sigma(t)$ not only indirectly improves generalist quality through eco-evo feedback, but also directly selects for high-quality generalists during *intraspecies* competition. In particular, we compare the different outcomes from *interspecies* competition between generalists and specialists, and from *intraspecies* competition among generalists. In *interspecies* competition, both generalist quantity and quality are promoted by suppressing specialists and reducing competition, and thus can be achieved by similar control function; in *intraspecies* competition, however, a trade-off exists between generalist quantity and quality, as high $\Sigma(t)$ directly selects for high-quality generalists at the cost of removing low-quality ones. Thus in *intraspecies* competition, generalist quantity and quality are promoted by distinct control profiles.

In a recent work ([219]), Sachdeva et al has highlighted the importance of environmental timescale in deciding the fate of evolution outcome. Our results further support this notion

by showing that a dynamic selection pressure outperforms constant selection pressure in promoting generalist species in an eco-evolutionary system. In the specific context of AM, our results show that the optimal vaccine Ag contents should be time-varying during a lasting GC reaction and provide an increasing selection strength in time: this applies to both the strain composition of vaccine Ag as well as to their concentration (Fig. 3.22).

This scenario of time-varying vaccine Ag during a lasting GC reaction distinguishes our work from existing studies on bnAbs elicitation (e.g.,[220, 60]). In the previously-studied scenario the vaccine injections are well separated in time, and later injections would initiate new GC reactions that are seeded by previous memories. The vaccine strain composition thus remains constant during each GC reaction and GC reactions terminate when Ag has run out, either due to decay or consumption by B cells ([59, 60]). While an increasing selection strength is advocated in both scenarios, we argue that the optimal vaccine Ag should be time-varying within the timescope of a single GC reaction. It is especially so following the first vaccine injection, when GCs are seeded by naive B cells with low cross-reactivity (lower end of m_c in Fig.3.4AB) instead of memory B cells.

When implementing the optimal dynamic vaccine Ag during an ongoing GC reaction, a method of sustained Ag delivery is required to replenish the GC Ag's decay and consumption. This can be achieved with osmotic pumps ([221],[222]). Once implanted, they can provide a prolonged drug release and the rate can be controlled by pump design. In [63], it has been demonstrated that an extended Ag dosing profile can be achieved with osmotic pumps and lead to improved Ab titer than one-time Ag delivery.

The biological details of Ag presentation and retention in GCs is not yet fully understood, and this may hinder our ability to precisely control the Ag contents in GC. To begin with, the decay rate of Ag in GC is not fully characterized, though there has been some evidence on long-term Ag retention on follicular dendritic cells in GC ([223]). In addition, a delay exists between the application of vaccine Ag and when they are eventually present in GC and become available to B cells. There are experimental reports, though, that the Ag can be detected in lymph nodes within only 4 hours after injection ([224]), which is shorter than a GC reaction cycle. Besides, our results also suggest that the shape of our optimal vaccine

Ag profile $\Sigma^*(t)$ is robust: suboptimal Ag profiles with imprecise shape parameters can still perform reasonably well compared with the optimal results (Fig.3.5, Fig. 3.16) as long as the cross-reactive B cells are not driven to early extinction (Fig.3.5,Fig. 3.17A).

The design rationale of germline-activating Ag to better elicit bnAbs has been under extensive investigations. Our results underscore the importance of germline-activating Ag, as GC founders' target epitope will play an important role in deciding the optimal vaccine strategy. On one hand, when germline-activating Ag recruit both B cells targeting variable and conserved epitopes, diverse boost Ag suppress specialists and a simultaneous improvement of generalist quantity and quality is possible. On the other hand, when germline-activating Ag predominantly recruit B cells targeting conserved Ag epitope, exclusion by specialists is no longer a concern and generalist survival becomes easier. However, since stronger Ag frustration improve relative fitness of the more cross-reactive B cells at the cost of their absolute fitness, one has to face a trade-off between optimizing generalist quantity or quality.

Finally, the coarse-grained nature of our models allow extension to broader context than B cell AM. For example, to the eco-evolutionary dynamics and its intervention in bacteria, cancer cell, or even macroscopic ecosystems.

3.6 Methods

3.6.1 Optimal control in the generic model

It is convenient to first introduce some shorthand notations and rewrite Eq.3.1 in a more compact form:

$$\dot{\chi} = p(\chi, u) + \nu(\chi, u)\xi \tag{3.6}$$

Here $\chi \equiv (N_s,\, N_g,\, f_s,\, f_g) \in \mathbb{R}^{4\times 1}$ and $\xi \equiv (\xi_1,\, \xi_2,\, \xi_3,\, \xi_4) \in \mathbb{R}^{4\times 1}$ are the vectors of state variables and noises. $p(\chi, u) \equiv ((f_s - d - u - cN_{\text{tot}})N_s,\, (f_g - d - \epsilon u - cN_{\text{tot}})N_g,\, 2D_s\frac{\partial f_s}{\partial \Gamma_s}N_s,\, 2D_g\frac{\partial f_g}{\partial \Gamma_g}N_g) \in \mathbb{R}^{4\times 1}$ is the deterministic component in Eq.3.1. Diagonal matrix $\nu(\chi, u)$
$\equiv \text{Diag}\{\sqrt{N_s[f_s + d + u + cN_{\text{tot}}]},\, \sqrt{N_g[f_g + d + \epsilon u + cN_{\text{tot}}]},\, \sqrt{D_s},\, \sqrt{D_g}\} \in \mathbb{R}^{4\times 4}$ denotes the diffusive components in Eq.3.1.

The optimal control policy $u^*(N_s, N_g, f_s, f_g, t) = u^*(\chi, t)$ is given by the HJB equation, which can be derived following standard method (e.g.,see [225, 226]) and is also detailed in Appendices. Here we present the final result. Denote $R(\chi, t)$ the integrand of the objective index we seek to maximize in Eq.3.2. The optimal "cost-to-go" function V at intermediate time t and state $\chi(t)$ is defined as:

$$V(\chi(t), t) \equiv \max_{u(t\to T)} \langle \int_t^T R(\chi(q))\mathrm{d}q \rangle \tag{3.7}$$

with the maximized objective index $\max_{u(0\to T)} J = V(\chi(0), t = 0)$. The HJB equation then reads:

$$-\frac{\partial V}{\partial \mathrm{t}} = \max_u [R(\chi) + \sum_{i=1}^{4} p_i(\chi, u)\frac{\partial V}{\partial \chi_i} + \frac{1}{2}\sum_{i=1}^{4} \nu_{ii}^2(\chi, u)\frac{\partial^2 V}{\partial \chi_i^2}] \tag{3.8}$$

Once V is solved, the optimal control $u^*(\chi, t)$ can be found by maximizing the bracket in Eq.3.8.

To numerically solve the HJB equation Eq.3.8, we employ the method detailed in [227] to convert the partial differential equation into a set of ordinary differential equations. The state space of χ is discretized and $V(\chi, t)$ at each discretized χ is considered as a unknown function of time. The partial derivatives with respect to χ in Eq.3.8 are replaced by finite differences. Eq.3.8 can then be solved by integrating it backward in time, starting from terminal time T and terminal values $V(\chi, T) = 0$ for all states χ.

3.6.2 BCR-virus binding affinity in the GC reaction model

For a specialist B cell $\vec{b}^x$ and a virus of variable and conserved string pair $(\vec{a}^v, \vec{a}^c)$, the binding affinity is modeled as

$$\begin{aligned} A^x[\vec{b}^x, (\vec{a}^v, \vec{a}^c)] &= A^x[m(\vec{b}^x, \vec{a}^v)] \\ &= \frac{m(\vec{b}^x, \vec{a}^v)}{M} E \end{aligned} \tag{3.9}$$

where the superscript x denotes specialists. E is an energy scaling constant and the matching score $m(\vec{b}^x, \vec{a}^v) \equiv \sum_{\ell=1}^{M} 1(b_\ell^x = a_\ell^v)$ is the number of matching sites between two strings.

The binding affinity between a generalist B cell $\vec{b}^{\,y}$ and the virus is modeled as

$$\begin{aligned} A^y[\vec{b}^{\,y},(\vec{a}^{\,v},\vec{a}^{\,c})] &= A^v[m_v(\vec{b}^{\,y},\vec{a}^{\,c})] + A^c[m_c(\vec{b}^{\,y},\vec{a}^{\,c})] \\ &= \underbrace{\frac{m_v(\vec{b}^{\,y},\vec{a}^{\,c})}{M}E}_{\text{variable binding}} + \underbrace{\frac{\left[m_c(\vec{b}^{\,y},\vec{a}^{\,c})/M_c\right]^C M_c}{M}E}_{\text{conserved binding}} \end{aligned} \tag{3.10}$$

where the superscript y indicates generalists. Here the matching scores are defined as $m_v(\vec{b}^{\,y},\vec{a}^{\,c}) \equiv \sum_{\ell=1}^{M_v} 1(b_\ell^y = a_\ell^c)$ and $m_c(\vec{b}^{\,y},\vec{a}^{\,c}) \equiv \sum_{\ell=M_v+1}^{M} 1(b_\ell^y = a_\ell^c)$: m_v and m_c are the number of matching variable and conserved sites between the B cell $\vec{b}^{\,y}$ and conserved epitope string $\vec{a}^{\,c}$, respectively. Constant $C > 1$ penalizes the lower accessibility of conserved viral residues ([22]). Eq.3.9 and 3.10 assume the binding affinity is a function of number of string matches but is independent of their locations ([46, 228, 229, 156, 47]). This enables the affinity class description of B cells since only the matching scores have to be tracked instead of the full string contents.

3.6.3 Optimal control in the GC reaction model

We seek the optimal control function Σ^*: $[0,T] \to [0,\Sigma_{\max}]$ to maximize the objective index $\Phi\{\Sigma(t)\}$ (Eq.3.5), given initial conditions $\vec{x}(t=0)$, $\vec{y}(t=0)$ and vaccine protocol duration T. Here we also assume a constant Ag concentration C_{Ag} during vaccination. In Appendices we extend to simultaneous optimization of dynamic antigenic distance $\Sigma(t)$ and Ag concentration $C_{Ag}(t)$ (Fig. 3.22). We choose $\Sigma_{\max} = 10$ during optimization. When $\Sigma = 10$, on average $\frac{\Sigma}{M} = 50\%$ of the variable residues differ between prime and boost viral strains, which is higher than the viral diversity reported for influenza ([210]) or HIV ([230]). Also see Appendices for even larger $\Sigma_{\max} = M = 20$ (Fig. 3.21).

For more compact notations, denote $\vec{z} \equiv \begin{pmatrix} \vec{x} \\ \vec{y} \end{pmatrix} \in \mathbb{R}^{M+1+(M_v+1)(M_c+1)}$ as the population vector of all B cells and $\vec{f} \equiv \begin{pmatrix} \vec{f}^{\,x} \\ \vec{f}^{\,y} \end{pmatrix}$ as the B cell growth rate given by Eq.3.3. Also denote $\phi(\vec{z})$ the integrand of objective index Φ in Eq.3.5, i.e., $\Phi = \int_0^T \phi(\vec{z})\,\mathrm{d}t$.

From the Pontryagin's maximum principle (PMP) ([231]), we define the costate vector $\vec{\lambda}(t)$ with the same dimension as state vector $\vec{z}(t)$. The Hamiltonian H is defined as $H(\vec{z},\vec{\lambda},\Sigma,C_{Ag}) \equiv \sum_\alpha \lambda_\alpha f_\alpha(\vec{z},\Sigma,C_{Ag}) + \phi(\vec{z})$, where the subscript α runs through all B cell

classes.

The PMP requires the costate vector $\vec{\lambda}$ to obey the dynamics:

$$\begin{aligned}\frac{\mathrm{d}\lambda_\alpha}{\mathrm{d}t} &= -\frac{\partial H}{\partial z_\alpha} \\ &= -\sum_\beta \lambda_\beta \frac{\partial f_\beta(\vec{z}, \Sigma, C_{Ag})}{\partial z_\alpha} - \frac{\partial \phi(\vec{z})}{\partial z_\alpha}\end{aligned} \tag{3.11}$$

with terminal condition $\vec{\lambda}(t = T) = 0$. Also according to PMP, the optimality of control function $\Sigma^*(t)$ requires its maximization of Hamiltonian H pointwise in time, i.e.,

$$\Sigma^*(t) = \arg\max_{\Sigma'} H(\vec{z}(t), \vec{\lambda}(t), \Sigma', C_{Ag}), \forall t \in [0, T]. \tag{3.12}$$

For the control $\Sigma^*(t)$ and corresponding state and costate variables $\vec{z}^*(t)$, $\vec{\lambda}^*(t)$ to be optimal, Eq.3.3, 3.11, 3.12 all must be satisfied. The numerical solution can be found iteratively by the method of successive approximations ([232, 45, 233], Fig. 3.15). See Appendices for algorithm procedures.

3.6.4 Definition of self-sustainable generalists

From Eq.3.3, the growth rate $\frac{\mathrm{d}y_i}{\mathrm{d}t} = f_i^y$ of generalist B cell class i can be grouped into *intrinsic* and *extrinsic* terms:

$$\begin{aligned}f_i^y(\vec{x}, \vec{y}, \Sigma) &\equiv f_{i,\mathrm{int}}^y(y_i, \Sigma) + f_{i,\mathrm{ext}}^y(\vec{x}, \vec{y}, \Sigma) \\ &= \underbrace{k_p y_i \theta_i^y h_i^y(\Sigma) \cdot (2\mathfrak{M}_{ii}^y - 1) - k_p y_i \theta_i^y [1 - h_i^y(\Sigma)]}_{\text{intrinsic: replication and selection}} \\ &\quad \underbrace{+ k_p \sum_{j \neq i} y_j \theta_j \cdot h_j^y(\Sigma) \cdot 2\mathfrak{M}_{ij}^y - k_p y_i \theta_i^y \frac{|\vec{x}| + |\vec{y}|}{K_B}}_{\text{extrinsic: mutation influx and competition}}\end{aligned} \tag{3.13}$$

The intrinsic growth rate describes the replication and apoptosis (if failed to internalize Ag during GC reaction cycle) of generalists in class i, which are independent from other B cells present in GC. In contrast, the extrinsic terms describe the mutation influx from other generalist cell classes $j \neq i$ into class i as well as apoptosis induced by competition.

The self-sustainability of generalist class i at antigenic distance Σ is defined by a pos-

itive intrinsic growth rate: $f^y_{i,\text{int}}(y_i, \Sigma) > 0$. In particular, we are concerned with the self-sustainable generalists at high antigenic distance $\Sigma = 10$: $f^y_{i,\text{int}}(y_i, \Sigma = 10) > 0$

3.6.5 Net per-cell growth rate of generalists

The net per-cell growth rate of generalist species $\frac{1}{|\vec{y}|}\frac{\mathrm{d}|\vec{y}|}{\mathrm{d}t}$ can be calculated from GC B cell dynamics Eq.3.3 by summing over all generalist classes $\{y_i\}$:

$$\begin{aligned}\frac{1}{|\vec{y}|}\frac{\mathrm{d}|\vec{y}|}{\mathrm{d}t} = &\underbrace{\frac{2k_p}{|\vec{y}|}\sum_i y_i\theta_i^y h_i^y(\Sigma)\cdot(1-p_{let})}_{\text{per-cell birth rate}} \\ &\underbrace{-\,k_p(1+\frac{|\vec{x}|+|\vec{y}|}{K_B})}_{\text{per-cell death rate}}\end{aligned} \tag{3.14}$$

Here $|\vec{y}| = \sum_i y_i\theta_i^y$ is the total abundance of generalists. The constant probability of lethal mutations p_{let} (see Appendices for details) in Eq.3.14 comes from the conservation of probability $p_{let} + \sum_i \mathfrak{M}^y_{ij} = 1$ for any generalist class j. The per-cell birth rate depends on the generalists' selection by Ag strains, reflected by the probability $h_i^y(\Sigma)$ to internalize Ag during GC reaction cycles. The per-cell death rate depends on the generalist-specialist competition due to finite capacity K_B. Note that the per-cell death rate does not depend on $h_i^y(\Sigma)$, since the B cell is always lost either if it dies by apoptosis during selection by Ag (with rate $k_p(1-h_i^y)$), or if it survives and is replaced by daughter cells after division (with rate $k_p h_i^y$).

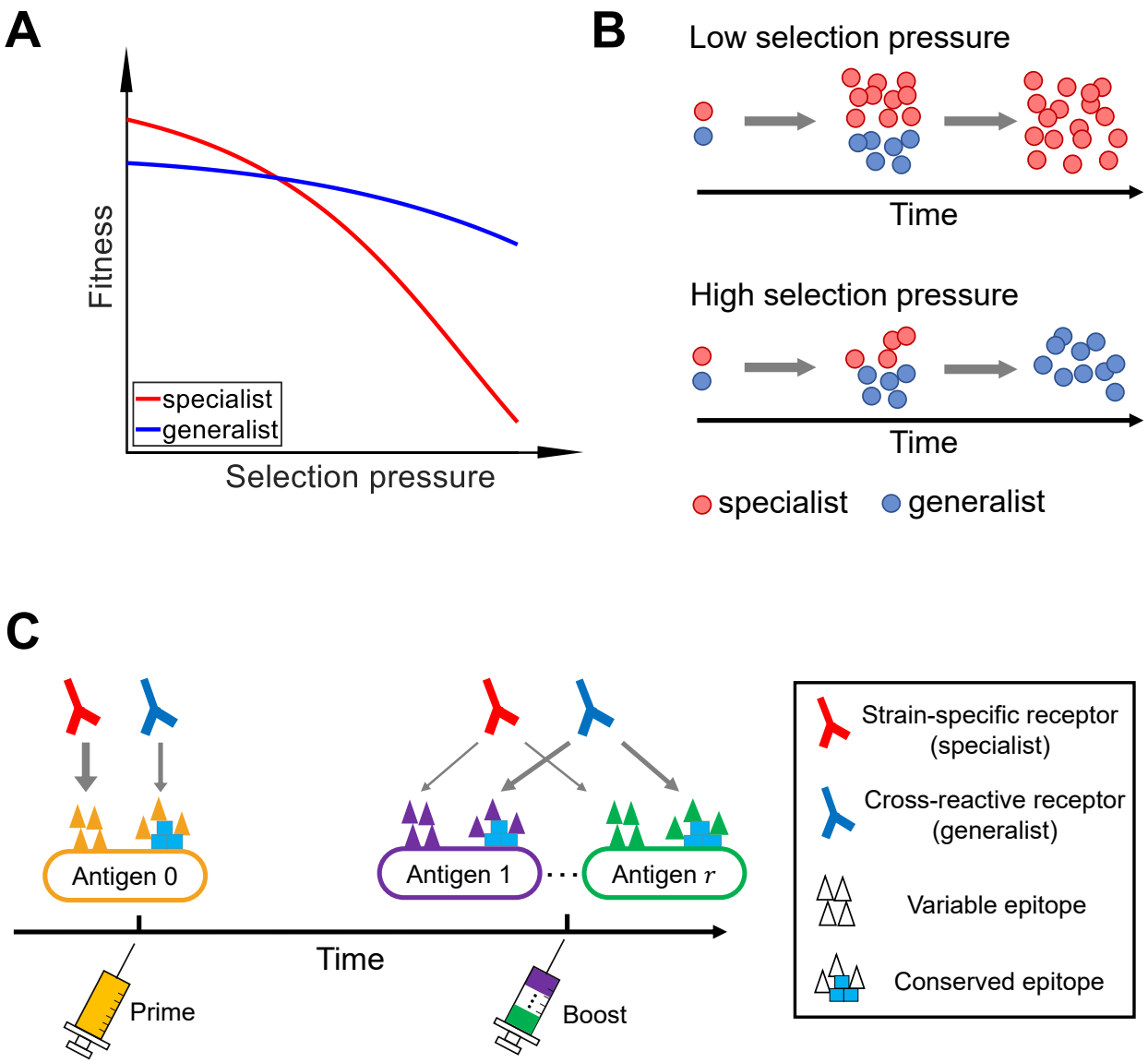

Figure 3.1: **Model schematics: controlling specialist-generalist competition through different strength of selection pressure.** (A)(B): A generic framework. (A): Applying high selection pressure from environmental heterogeneity as control lowers the absolute fitness of both species, but generalists are less negatively affected. (B) Therefore at high selection pressure, generalists can gain a relative advantage and outcompete the specialists. (C) A concrete model inspired by GC B cell evolution: specialist/generalist BCRs recognize and bind the variable/conserved epitopes of different Ag strains. The Ag strain composition is controlled by repeated injections or osmotic pumps and can change in time. Conserved Ag residues (light blue blocks) as part of conserved epitope are shared among Ag strains while variable residues (colored triangles) are strain-specific. Initially B cells are activated by a germline-activating ("prime") strain (#0, yellow). While specialist receptors bind stronger than generalists against this strain (thicker arrow), generalists have a stronger binding against antigenically different ("boost") strains (#1, 2 ...r...) due to their shared conserved residues.

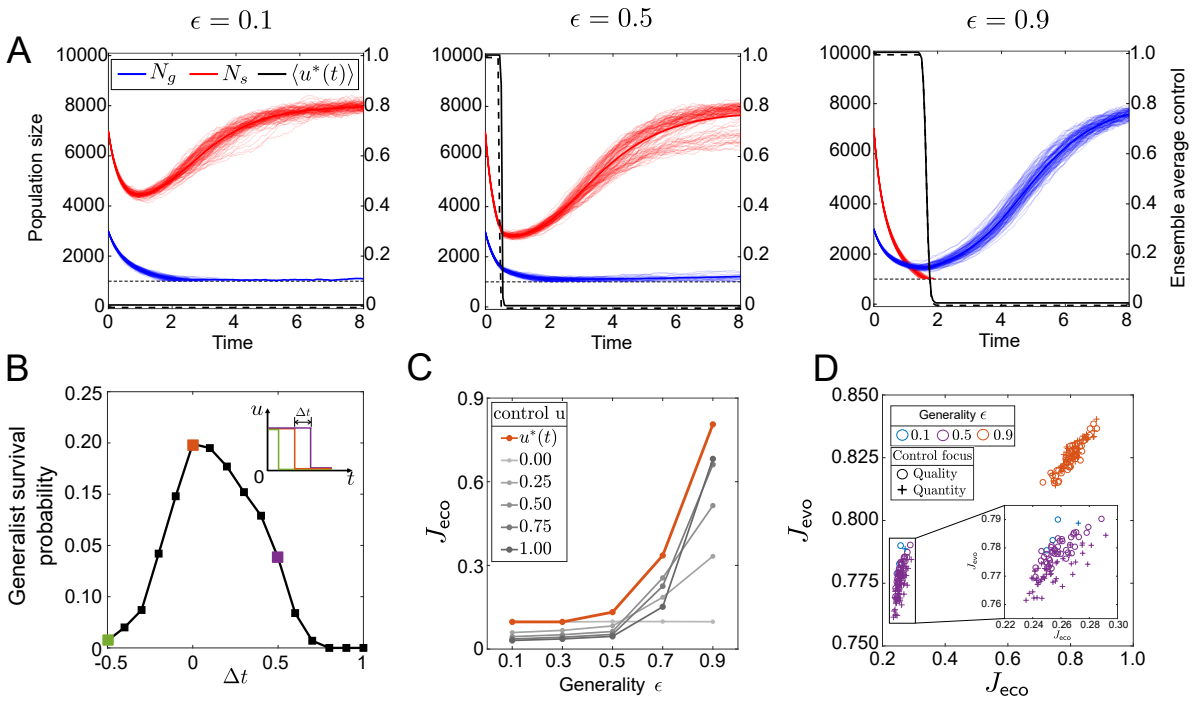

Figure 3.2: **Generalist quantity and quality are both improved by suppressing specialists and are best promoted by dynamic selection.** (A): Trajectories of population abundance $\vec{N}(t)$ and averaged optimal control profile (solid) from repeated SDE simulations optimizing ecological index J_{eco} (Eq.3.2). Thick blue/red lines show the population average. The averaged optimal control profile for evolutionary index J_{evo} (dashed black) is nearly identical to ecological control. A small vertical shift is added to evolutionary control to avoid overlapping. Initial condition is $N_s = 7000$, $N_g = 3000$, $f_s = f_g = 0.8$. Lower cutoff is $N_c = 1000$ (thin horizontal dashed line). 1000 simulations are performed for population average and plotted population trajectories are subsampled from 100 simulations for clarity. (B): Generalists survival probability is the highest when the control switches at optimal timing. Δt is the lag of control switch relative to the optimal timing (see inset; green, red, purple trajectories corresponds to early, optimal, late switches, respectively); $\Delta t > 0$ (< 0) means late (early) switch. The fraction of simulations with surviving generalists until terminal time $T = 8$ peaks at optimal off time ($\Delta t = 0$). $\epsilon = 0.5$. 1000 simulations are performed at each Δt. Also see Fig. 3.9 for the time to generalist extinction when $\Delta t = -0.5, 0, 0.5$ (green, red, purple squares, respectively). (C): Constant control (gray) is outperformed by optimized dynamic control (red). Among constant controls, maximum/intermediate/zero value is preferred at high/intermediate/low generalists' generality ϵ. 1000 simulations are performed at each data point. (D): The generalist ecological and evolutionary indices J_{eco} and J_{evo} can be simultaneously promoted by either ecological or evolutionary control. Data are from 1000 simulations in panel (A)(C) and with surviving generalists at terminal time T. The fraction of simulations with survival generalists is $< 1\%$, 19.9, and 100% when $\epsilon = 0.1$, 0.5 and 0.9, respectively (not shown). For $\epsilon = 0.5$ and 0.9, 50 simulations are subsampled and plotted for clarity. Circles/crosses: control is optimized for generalist quality J_{evo}/quantity J_{eco}.

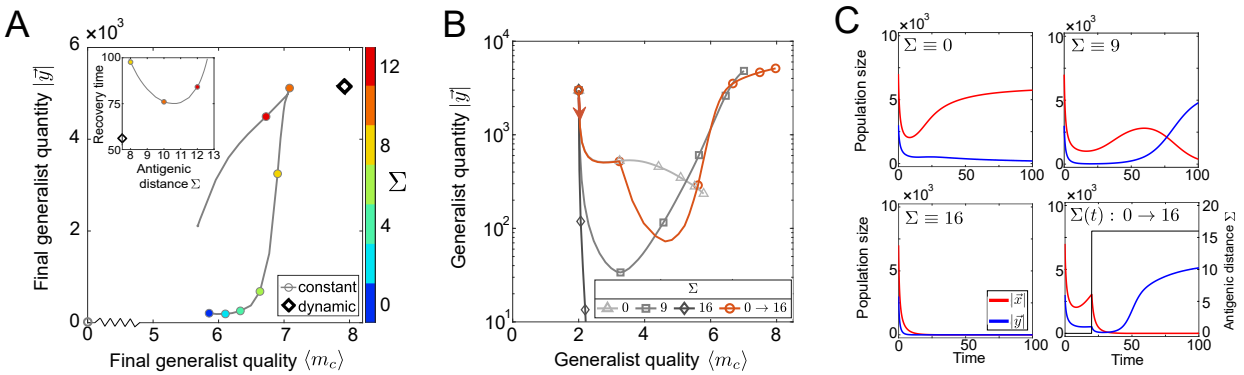

Figure 3.3: **Applying constant antigenic distance Σ is subject to a trade-off between suppressing specialists and preserving generalists, and is outperformed by time-varying antigenic distance $\Sigma(t)$.** (A) Final generalist quantity $|\vec{y}|$ and quality $\langle m_c \rangle$ at different constant antigenic distance Σ (solid line). Σ increases in counterclockwise direction (blue→red). The optimum is at intermediate level of Σ. At high constant $\Sigma > 12.7$ generalists go extinct, as represented by the gray circle at origin. Note the transition to extinction is discontinuous in final generalist quantity: when $\Sigma = 12.7$ generalists barely avoid extinction when their abundance is at the lowest, and then recover to over thousand cells at terminal time T. A better performance than constant Σ is achieved with simple dynamical $\Sigma(t)$ (black diamonds) that switches once from 0 to $\Sigma_{\max} = 16$ at $t_1 = 20$; the Σ level after switching and corresponding switching time are optimized over ($\Sigma_{\max}$, t_1) pairs. Inset: Generalists recover to initial abundance early at intermediate constant level of Σ (solid line), and even earlier with the dynamic $\Sigma(t)$ function (diamond). Also see Appendices for the corresponding objective indices Φ_{evo} and Φ_{eco} from constant and dynamic $\Sigma(t)$. (B) Phase trajectories of generalist quantity $|\vec{y}|$ and quality $\langle m_c \rangle$ over time. Gray: constant Σ at low ($\Sigma \equiv 0$), intermediate ($\Sigma \equiv 9$), or high level ($\Sigma \equiv 16$). red: dynamic $\Sigma(t)$ function that switches from 0 to $\Sigma_{\max} = 16$ at $t_1 = 20$ as in panel A. Red arrow marks the direction of time. Markers show $t = 0$, 4, 8 along the trajectory at high constant $\Sigma \equiv 16$ (gray diamonds) and $t = 0$, 20, 40, 60, 80, 100 along all others. (C) Time trajectories of generalist (blue) and specialist (red) quantities corresponding to panel B. The dynamic $\Sigma(t)$ function is shown with black solid line in lower right subplot. In all panels, GC reaction starts with 10^4 B cells including $|\vec{x}| = 7000$ specialists ($m = 8$) and $|\vec{y}| = 3000$ generalists ($m_v = 6$, $m_c = 2$). Ag concentration $C_{\text{Ag}} = 0.05$, vaccine protocol duration $T = 100$.

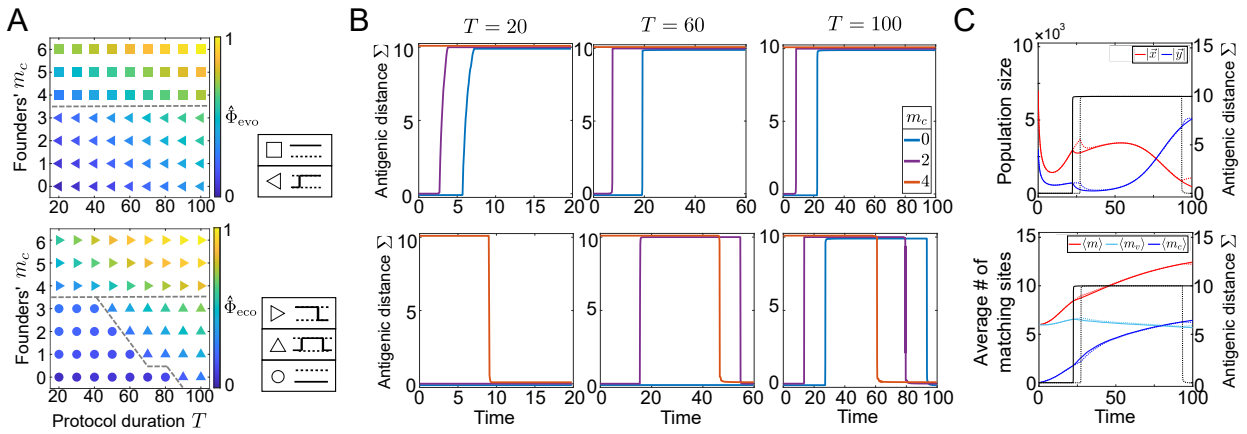

Figure 3.4: **Optimized antigenic distance $\Sigma^*(t)$: distinct shapes depending on generalist founder quality m_c and vaccine protocol duration T.** (A) Phase diagram of control shapes optimized by Pontryagin's maximum principle to maximize evolutionary or ecological objective $\Phi_{\rm evo}/\Phi_{\rm eco}$ (Eq.3.5). Different control shapes are labeled with markers (see legend for schematics; lower and upper dashed lines are guidelines for $\Sigma(t) \equiv 0$ and $\Sigma(t) \equiv \Sigma_{\rm max} = 10$, respectively). Marker color indicates the value of normalized objective indices $\hat{\Phi}_{\rm evo}/\hat{\Phi}_{\rm eco}$ by their maximum values at $m_c = 6$ and $T = 100$. GC reaction starts with 10^4 B cells including $|\vec{x}| = 7000$ specialists and $|\vec{y}| = 3000$ generalists. Ag concentration $C_{\rm Ag} = 0.1$. For generalist founders $m_v = 6$ and m_c is as labeled; for specialist founders $m = m_v + m_c$. (B) Examples of optimized antigenic distance $\Sigma^*(t)$ for different founder quality m_c and protocol duration T. A small vertical shift is added to red and blue curves to avoid overlapping. (C) Time trajectories of B cell quantity (top) and quality (bottom) with optimized $\Sigma^*(t)$. Solid and dashed: $\Sigma^*(t)$ optimizes $\Phi_{\rm evo}$ and $\Phi_{\rm eco}$, respectively. $|\vec{x}| \equiv \sum_i x_i \theta_i^x$ and $|\vec{y}| \equiv \sum_i y_i \theta_i^y$ (Eq.3.3) are the population sizes of specialists and generalists, respectively. $m_c = 0$ for generalist founders. Also see Appendices for similar results when maximal allowed $\Sigma_{\rm max} > 10$ (Fig. 3.21).

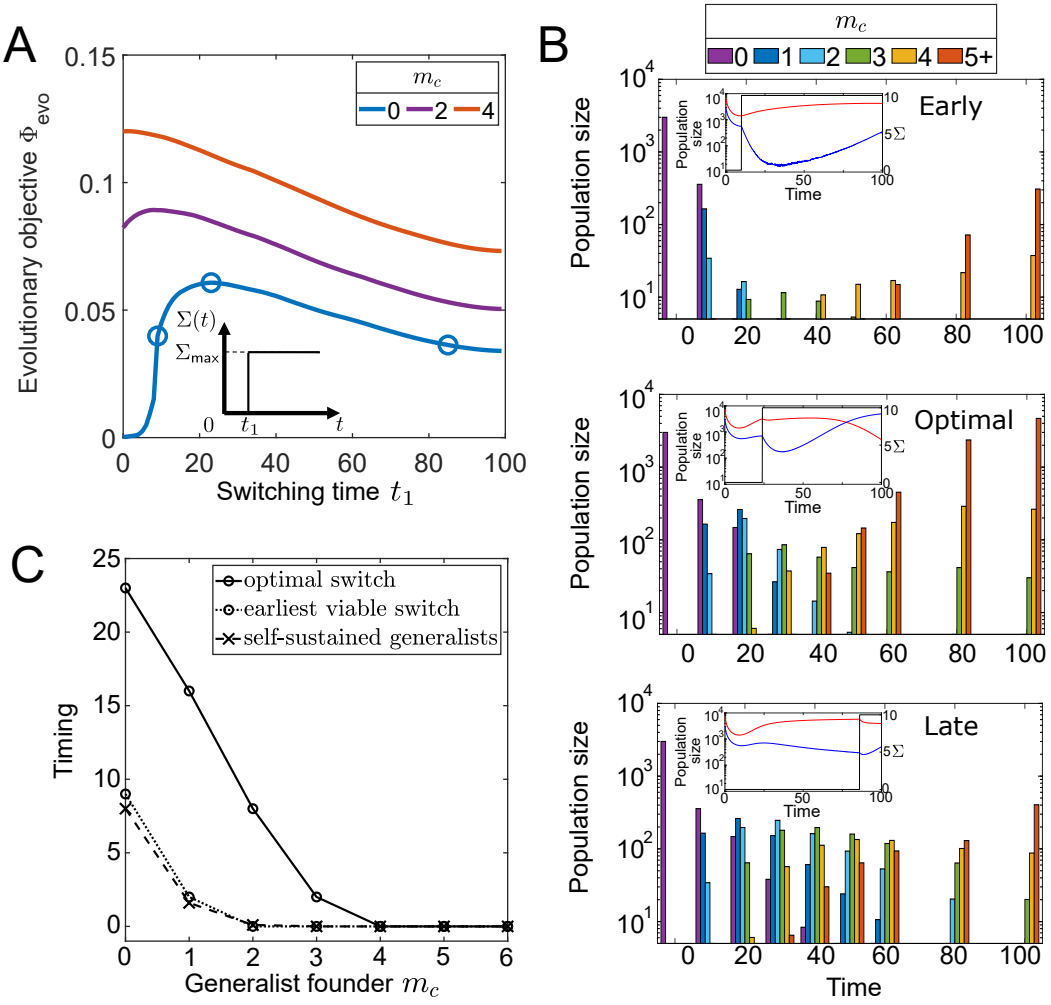

Figure 3.5: **Optimal timing to switch on high antigenic distance: a balance between early generalist evolution and subsequent specialist suppression.** (A) An intermediate time t_1 to switch on maximum antigenic distance $\Sigma_{\max}$ (see inset for schematics) maximizes the evolutionary objective Φ_{evo} for founder cross-reactivity $m_c \leq 2$. Blue circles mark the peak position ($t_1 = 24$) and when quality objective Φ_{evo} drops to 60% peak value ($t_1 = 10$ and 86). (B) The cross-reactivity of generalists improve and these high-quality generalists expand over time, which progress the fastest for optimal switching time. Generalists are binned and counted based on their cross-reactivity m_c. From top to bottom: the switching of $\Sigma(t)$ is early, optimal, and late, matching three blue circles in (A). Inset shows the corresponding antigenic distance $\Sigma(t)$ (black) and total generalist and specialist abundance $|\vec{y}|$, $|\vec{x}|$ (blue and red, respectively). (C) The emergence time of self-sustainable generalists predicts the earliest viable on-switch that avoids generalist extinction; both timings also precede the optimal on-switch. The earliest viable switch is found by scanning switching time t_1 from zero and in increment of 1, until generalists can survive the population bottleneck; the emergence time of self-sustainable generalists is found as when generalists that satisfy $f^y_{i,int}|_{\Sigma=\Sigma_{\max}} > 0$ have first evolved under $\Sigma(t) \equiv 0$. The self-sustainability condition is first met by generalists with $m_c \geq 3$. Parameters and initial B cell abundance are the same as in Fig.3.4 except that vaccine protocol duration is fixed as $T = 100$. For generalist founders $m_v = 6$; m_c is as labeled in panel (A)(C) and is zero in panel (B). For specialist founders $m = m_v + m_c$.

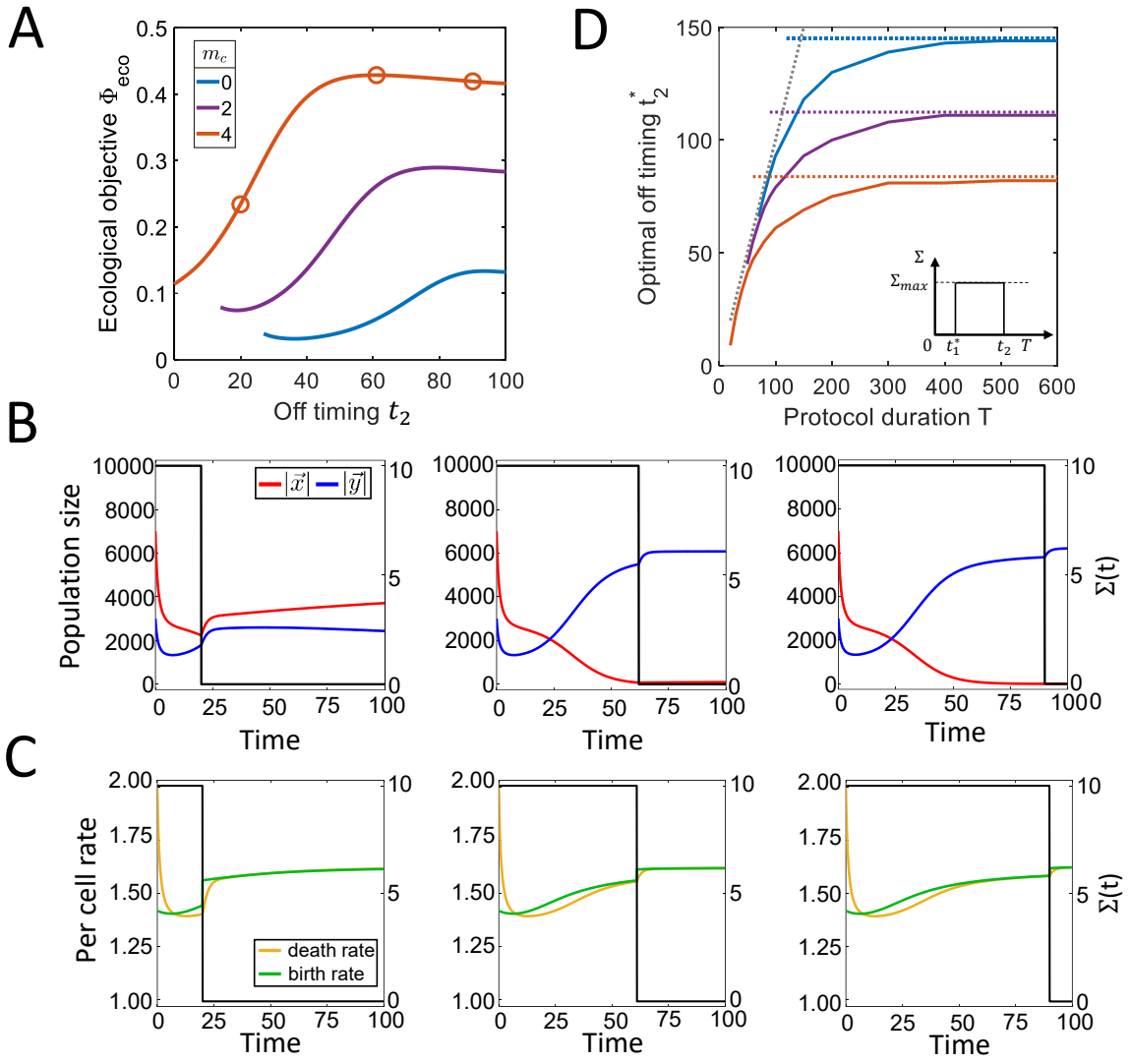

Figure 3.6: **Optimal timing to switch off high antigenic distance: a balance between specialist suppression and subsequent relaxation of generalist selection** (A): Ecological objective Φ_{eco} as a function of off-timing t_2: an intermediate value is optimal. High antigenic distance is switched on at optimal timing t_1^* that maximizes Φ_{eco} across all (t_1, t_2) pairs, and all curves start from $t_2 = t_1^*$. Also see inset of panel (D) for control schematics. $\Sigma_{\max} = 10$. Red circles mark the parameters showcased in panel (B)(C). (B) B cell population sizes and (C) generalist birth and death rates (Eq.3.14) when the control's off timing t_2 is early (left, $t_2 = 20$), optimal (middle, $t_2 = 61$), or late (right, $t_2 = 90$). Control function is shown in black (D) The optimal off timing t_2^* as a function of generalist founder quality m_c and vaccine protocol duration T (colored solid). At small T, $t_2^*(T) \lesssim T$ to maintain specialist suppression and thus maximized generalist expansion. The gray dashed line plots $t_2^* = T$ as a guideline to the eye. As T increases, $t_2^*(T)$ gradually saturates at the time point when specialists are driven extinct (horizontal dashed). Parameters and initial B cell abundance are the same as in Fig.3.5. For generalist founders $m_v = 6$; m_c is as labeled in panels (A)(D), and is $m_c = 4$ in panels (B)(C). For specialist founders $m = m_v + m_c$.

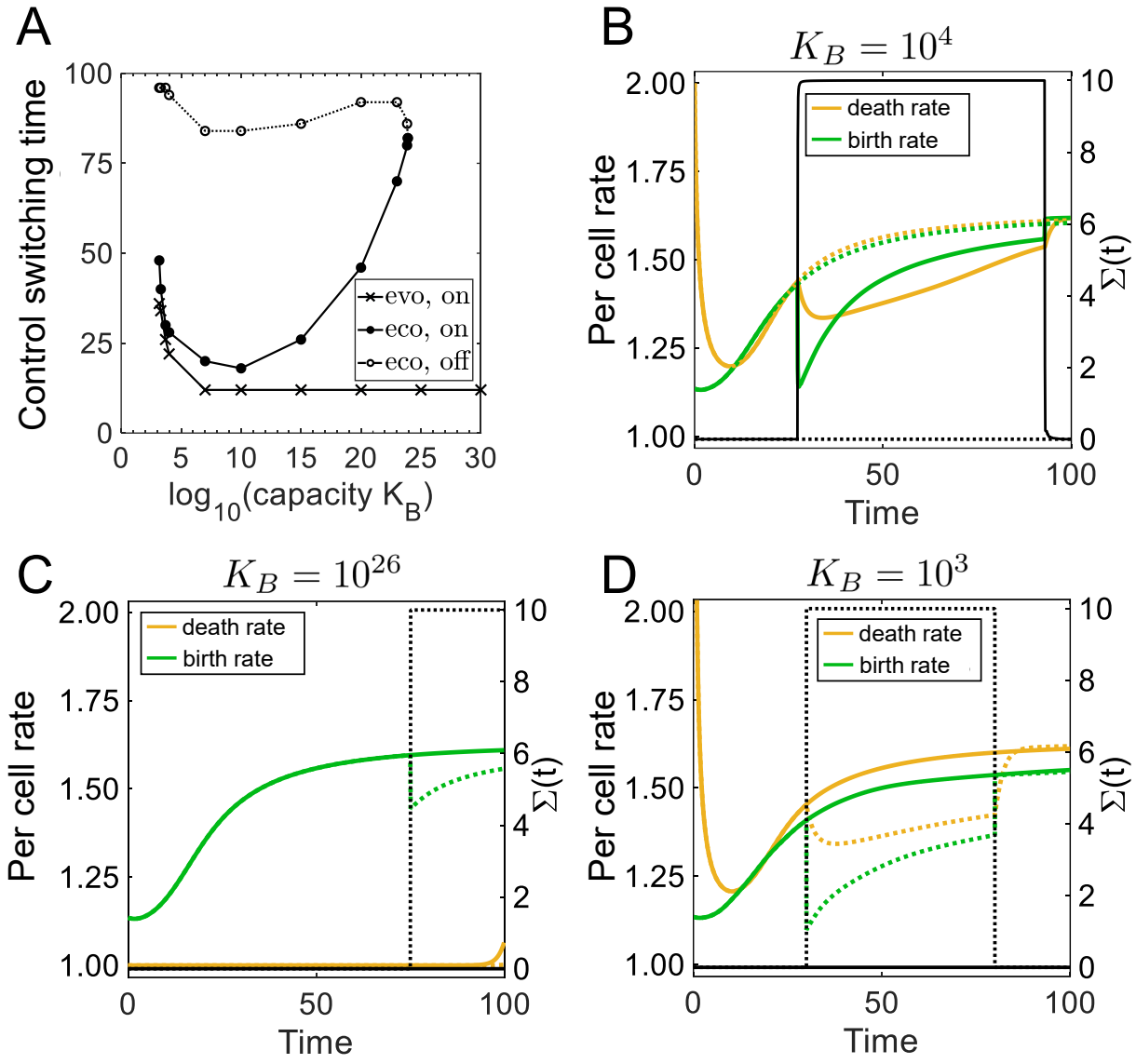

Figure 3.7: **Switching on high antigenic distance Σ simultaneously promotes both generalist quantity and quality for realistic range of GC B cell capacity K_B.** (A): Optimal timing to switch on/off high antigenic distance $\Sigma(t) = \Sigma_{\max}$ to promote generalist quantity (ecological objective Φ_{eco}; see Eq.3.5) or quality (evolutionary objective Φ_{evo}). Data in (A) are acquired by scanning through all pairs of on and off timing between $t = 0$ and $t = 100$ (the vaccine protocol duration); the resolution is $\Delta t = 2$. The optimal control for $K_B \leq 10^3$ is $\Sigma(t) \equiv 0$ for both eco- and evo- objectives. (B)-(D) optimal (solid) and example non-optimal (dashed) ecological controls (black) and corresponding generalist birth and death rates (colored, Eq.3.14), for (B) realistic (10^4), (C) large (10^{26}), and (D) small (10^3) capacities K_B. The control on and off timing pairs (t_1, t_2)=(28, 93) in panel (B), (75, N/A) in panel (C), and (30, 80) in panel (D). Parameters and initial B cell abundance are the same as in Fig.3.5 except that capacity K_B is as labeled. For generalist founders $m_v = 6$, $m_c = 0$. For specialist founders $m = 6$.

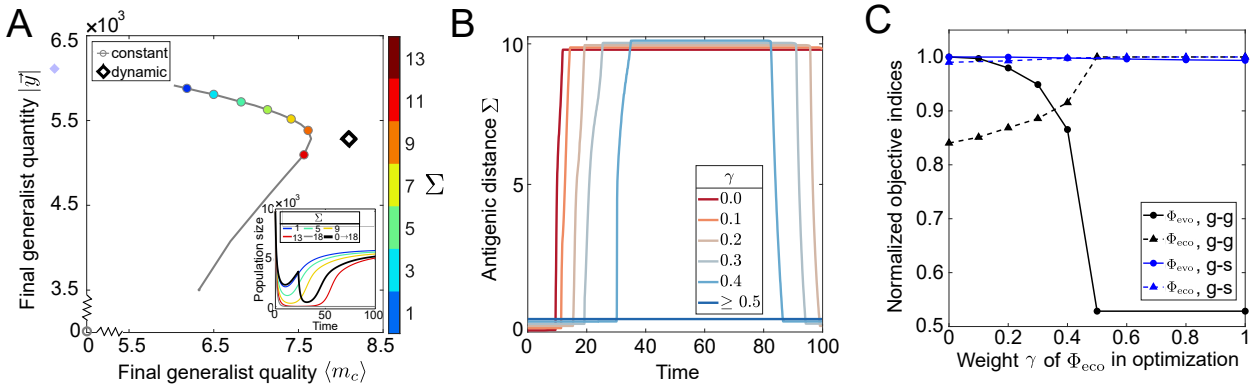

Figure 3.8: **Intraspecies competition of generalists targeting the same conserved epitope: trade-off between generalist quantity and quality** (A) Final generalist quantity $|\vec{y}|$ and quality $\langle m_c \rangle$ for different constant antigenic distance Σ (solid line), when the GC is founded only by generalists targeting the conserved epitope. Σ increases in clockwise direction (blue→red). At high constant $\Sigma > 13.9$ generalists go extinct, as represented by the gray circle at origin. The transition to extinction is discontinuous in final generalist quantity for the same reason as in Fig.3.3A. Diamonds show the outcome of a simple dynamic $\Sigma(t)$ that switches once from 0 to $\Sigma_{max} = 18$ at $t_1 = 23$; the (Σ_{max}, t_1) pairs are optimized to maximize Φ_{evo}. Generalist quantity and quality are negatively correlated as Σ increases from zero, indicated by the negative slope before curves turn towards extinction. The dynamic $\Sigma(t)$ function can thus no longer promote both generalist quantity and quality compared with constant Σ as in Fig.3.3A. Parameters are the same as Fig.3.3 except that all 10^4 founders are generalists. Inset: Generalist population trajectories at different constant Σ (colored) or dynamic $\Sigma(t)$ (black). Also see Fig. 3.19 for the values of objective indices. (B) Optimized antigenic distance $\Sigma^*(t)$ from Pontryagin's maximum principle for objective indices $\Phi = \gamma\Phi_{eco} + (1-\gamma)\Phi_{evo}$ (Eq.3.5) at different γ: more focus on generalist quantity is placed as γ increases from 0 (red, quality only) to 1 (blue, quantity only). A small vertical shift is added to all curves to avoid overlapping. Same parameters as Fig.3.4B and 3.4C is used except all founders are generalists. See Appendices for optimized $\Sigma^*(t)$ at different γ for generalist-specialist competition (Fig. 3.20). (C) Normalized objective indices Φ_{evo} (solid) and Φ_{eco} (dashed) when the optimal antigenic distance $\Sigma^*(t)$ maximizes $\Phi = \gamma\Phi_{eco} + (1-\gamma)\Phi_{evo}$ for different γ. Data are rescaled by the maximum value for each curve so all data ≤ 1. Blue: generalist-specialist competition where B cells target two distinct epitopes; black: intraspecies competition among generalists where all B cells target the conserved epitope.

APPENDIX

3.A Model formulation and control optimization

3.A.1 Generic model of eco-evolutionary dynamics of generalist-specialist competition

In the generic model (Eq.3.1) we consider two species $\ell = g$ (generalist), s (specialist) competing for finite resources. Their eco-evolutionary dynamics of population abundance $\{N_\ell\}$ and mean trait $\{\Gamma_\ell\}$ is adapted from [203].

The ecological dynamics of N_ℓ depends on the birth rate f_ℓ, baseline death rate d, density-dependent death rate $cN_{\text{tot}} = c(N_s + N_g)$ and the death rate u or ϵu from varied selection strength. Compared with original formulation in [203], we include two coexisting species coupled by finite capacity instead of one; our demographic noise is also proportional to the sum of birth and death rates ([234, 235, 236, 237]), as in our formulation we do not constrain the sum to be constant.

For the evolutionary dynamics of trait Γ_ℓ, the diffusion constant for species ℓ is decided by $D_\ell = U_\ell \epsilon_\ell^2$ where U_ℓ is the total mutation rate and ϵ_ℓ^2 the mean squared effect of mutations at genomic loci encoding the trait Γ_ℓ ([203]). Additionally, ecology feedbacks onto evolution through population mutation rate, reflected by N_ℓ in the drift term. Finally, we define a sigmoidal fitness landscape $f_\ell(\Gamma_\ell) = f_M \frac{\exp(\Gamma_\ell)}{1+\exp(\Gamma_\ell)}$ where constant f_M is the maximum birth rate. Such choice is motivated by the Langmuir isotherm in BCR-Ag binding and Γ_ℓ can be interpreted as a proxy for binding affinity.

To emphasize the effect of ecological feedback on evolution, we set the initial birth rates of generalists and specialists to be identical, and the initial abundance of generalists to be lower than specialists. We also set the diffusion constants $D_g = D_s$. As such, according to Eq.3.1, the difference in two species' birth rates $f_g(t)$ and $f_s(t)$ will be solely due to ecological feedback.

3.A.2 Derivation of Hamilton-Jacobi-Bellman equation for the generic model

By definition of Eq. 3.7, the optimal "cost-to-go" function $V(\chi(t), t)$ obeys a recurrence relation for $t < t' < T$ (written in compact notation introduced in Eq. 3.6):

$$\begin{aligned} V(\chi(t), t) &= \max_{u(t\to T)} \langle \int_t^{t'} R(\chi(q))\mathrm{d}q + \int_{t'}^{T} R(\chi(q))\mathrm{d}q \rangle \\ &= \max_{u(t\to t')} \langle \int_t^{t'} R(\chi(q))\mathrm{d}q + \max_{u(t'\to T)} \langle \int_{t'}^{T} R(\chi(q))\mathrm{d}q \rangle \rangle \\ &= \max_{u(t\to t')} \langle V(\chi(t'), t') + \int_t^{t'} R(\chi(q))\mathrm{d}q \rangle \end{aligned} \tag{3.15}$$

By choosing $t' = t + \mathrm{d}t$ we can Taylor-expand $V(\chi(t'), t') = V(\chi(t + \mathrm{d}t), t + \mathrm{d}t)$. From Eq. 3.6 and 3.7

$$V(\chi(t + \mathrm{d}t)), t + \mathrm{d}t)) = V(\chi(t), t) + \frac{\partial V}{\partial \mathrm{t}}\mathrm{d}t + (\sum_{i=1}^{4} \frac{\partial V}{\partial \chi_i} p_i)\mathrm{d}t + \frac{1}{2}(\sum_{i=1}^{4} \frac{\partial^2 V}{\partial \chi_i^2} \nu_{ii}^2)\mathrm{d}t \tag{3.16}$$

where we have used the fact that $\langle \mathrm{d}\chi_i \rangle = p_i \mathrm{d}t$ and $\langle \mathrm{d}\chi_i \mathrm{d}\chi_j \rangle = \nu_{ij}^2 \mathrm{d}t$.

Plugging Eq.3.16 into Eq.3.15 and divide by $\mathrm{d}t$, we have the HJB equation Eq.3.8. The terminal values are $V(\chi, T) = 0$ for all χ by definition of Eq.3.7.

3.A.3 Construction of the mutation matrix $\mathfrak{M}$ in AM model

We start with different types of B cell mutations and their and effect in our model. During AM, a B cell mutation can be lethal with probability p_{let}, affinity-altering with probability p_{aa}, or silent with probability $p_{\mathrm{s}} = 1 - p_{\mathrm{let}} - p_{\mathrm{aa}}$ ([155, 156, 59]). For a lethal mutation, the daughter B cell goes through apoptosis and is removed from GC. For an affinity-altering mutation, the daughter B cell becomes a different affinity class than the mother cell. Specifically, we assume an affinity-altering mutation affects one randomly chosen residue group (i.e., effectively no mutation hotspot, [155, 156, 24, 47]) from the B cell complementarity-determining region. As a result, a random site of the B cell string switches to a different letter, picked randomly from the remaining $(K - 1)$ letters in the alphabet. For a silent mutation or in the absence of mutations, the B cell affinity class remains the same between daughter and mother cells.

Next, we define $\hat{\mathfrak{M}}^x \in \mathbb{R}^{(M+1)\times(M+1)}$ and $\hat{\mathfrak{M}}^y \in \mathbb{R}^{(M_v+1)(M_c+1)\times(M_v+1)(M_c+1)}$ as the *mutation matrices of affinity-altering mutations* for specialist ($s = x$) and generalist ($s = y$) B cells, respectively. The matrix element $\hat{\mathfrak{M}}^s_{ij}$ is the probability for a daughter B cell to be class i conditioned on an affinity-altering mutation, when the mother cell is of species s and class j. Denote μ the mutation rate per daughter B cell per cell division during AM. The mutation matrix $\mathfrak{M}^s$ in Eq. 3.3 can be written in terms of $\hat{\mathfrak{M}}^x$ as

$$\mathfrak{M}^s = p_{aa}\mu\,\hat{\mathfrak{M}}^s + (1 - p_{\mathrm{aa}}\mu - p_{\mathrm{let}}\mu)I^s \tag{3.17}$$

where I^s is the identity matrix of same size as $\hat{\mathfrak{M}}^s$.

Finally, we find the matrix elements of $\hat{\mathfrak{M}}^s_{ij}$. For a specialist B cell of matching score $0 \le m \le M$ against prime virus, an affinity-altering mutation leads to one of the three outcomes: (1): The mutation site matches the prime virus before mutation, then the match will be destroyed after the B cell mutation. This occurs with probability $\Pr(m \to m - 1|m, aa) = \frac{m}{M}$ where aa denotes affinity-altering mutations; (2): The mutation site does not match the prime virus before mutation, but a match is created after B cell mutation . This occurs with probability $\Pr(m \to m + 1|m, aa) = \frac{M-m}{M}\frac{1}{K-1}$; (3): The mutation site does not match the prime virus either before or after the mutation. This occurs with probability $\Pr(m \to m|m, aa) = \frac{M-m}{M}(1 - \frac{1}{K-1})$. The specialists' matrix of affinity-altering mutations $\hat{\mathfrak{M}}^x$ can then be written as $\hat{\mathfrak{M}}^x_{ij} = \Pr(j-1 \to i-1|j-1, aa)$. $\hat{\mathfrak{M}}^x$ is a tridiagonal matrix by construction as $\hat{\mathfrak{M}}^x_{ij} = 0$ unless $j = i$ or $i \pm 1$.

For generalists B cells, the affinity-altering mutation matrix $\hat{\mathfrak{M}}^y$ can be built in the same manner. Specifically, an affinity-altering mutation changes the matching score m_v or m_c (but not both) by 0 or ± 1. This leads to one of five possible outcomes with the following

probabilities

$$\begin{aligned}
\Pr(m_v \to m_v - 1, m_c \to m_c | m_v, m_c, aa) &= \frac{m_v}{M} \\
\Pr(m_v \to m_v + 1, m_c \to m_c | m_v, m_c, aa) &= \frac{M_v - m_v}{M} \frac{1}{K-1} \\
\Pr(m_v \to m_v, m_c \to m_c - 1 | m_v, m_c, aa) &= \frac{m_c}{M} \\
\Pr(m_v \to m_v, m_c \to m_c + 1 | m_v, m_c, aa) &= \frac{M_c - m_c}{M} \frac{1}{K-1} \\
\Pr(m_v \to m_v, m_c \to m_c | m_v, m_c, aa) &= \frac{M - m_v - m_c}{M} (1 - \frac{1}{K-1})
\end{aligned} \tag{3.18}$$

If the affinity classes i of the daughter and j of mother generalist B cell belong to these five scenarios, the matrix element $\hat{\mathfrak{M}}^y_{ij}$ is given by the corresponding probability in Eq.3.18. Otherwise the transition from class j to i cannot be achieved by a single mutation and $\hat{\mathfrak{M}}^y_{ij} = 0$.

3.A.4 Deriving the average binding affinity between B cell and boost viruses within affinity class formulation

Here we derive the log-mean-exponential binding affinity $\langle A^s_i(\Sigma) \rangle$ (Eq. 3.4) between B cell of affinity class i and boost virus strains of antigenic distance Σ. Species $s = x,\ y$ for specialists and generalists, respectively. Denote the prime virus as variable and conserved epitope string pair $(\vec{a}^{0,v}, \vec{a}^{0,c})$, and N_a boost viruses as $\{(\vec{a}^{r,v}, \vec{a}^{r,c})\}$, $r = 1, 2, 3, ...N_a$.

First, consider a specialist B cell $\vec{b}^{\,s}$ whose matching score against the prime virus is $m = m(\vec{b}^{\,s}, \vec{a}^{0,v})$. At the m BCR-prime matching sites, the rth boost virus $\vec{a}^{r,v}$ differs with the prime virus $\vec{a}^{0,v}$ by Σ^r_1 sites. Per definition of antigenic distance Σ, Σ^r_1 obeys a binomial distribution with trial number m and probability of success $\frac{\Sigma}{M}$: $\Sigma^r_1 \sim \mathrm{B}(m, \frac{\Sigma}{M})$. At any site ℓ among these Σ^r_1 sites, B cell $\vec{b}^{\,s}$ will have a mismatch with the rth boost virus since $b^s_\ell = a^{0,v}_\ell \neq a^{r,v}_\ell$. Similarly, at the $M - m$ sites where the B cell and prime virus mismatch, the boost and prime viruses' strings differ by Σ^r_2 sites with $\Sigma^r_2 \sim \mathrm{B}(M - m, \frac{\Sigma}{M})$. At any site ℓ among these Σ^r_2 sites, with probability $1/(K-1)$ the B cell matches the rth boost virus, i.e., $b^s_\ell = a^{r,v}_\ell \neq a^{0,v}_\ell$. That is, viral mutation from prime to boost viruses can create new string matches with B cells. Denote Δ^r the number of BCR-boost matches created in this

way, then Δ^r follows a binomial distribution with number of trials $M-m$ and probability of success $\frac{\Sigma}{M}\frac{1}{K-1}$: $\Delta^r \sim \mathrm{B}(M-m, \frac{\Sigma}{M}\frac{1}{K-1})$. In all, we have

$$m^r \equiv m(\vec{b}^{\,s}, \vec{a}^{r,v}) = m - \Sigma_1^r + \Delta^r \tag{3.19}$$

Using the law of total probability, the probability mass function of m^r is calculated as

$$\Pr(m^r = m'|m,\Sigma) = \sum_{\Delta'} \Pr(\Delta^r = \Delta') \cdot \Pr[\Sigma_1^r = \Delta' - (m'-m)] \tag{3.20}$$

and the binomial distributions of Σ_1^r and Δ^r can be plugged into Eq.3.20. From Eq.3.20 and the definition of affinity Eq. 3.9, we now can express the average binding affinity between a specialist B cell and boost viruses in our affinity class description, without explicit dependence on the full specialist string $\vec{b}^s$:

$$\begin{aligned} e^{\langle A_i^x(\Sigma)\rangle} &= \langle e^{A^x[\vec{b}^{\,s},(\vec{a}^{r,v},\vec{a}^{r,c})]}\rangle_r = \langle e^{(m^r/M)E}\rangle_{m^r} \\ &= \sum_{r=1}^{N_a} c^r e^{(m^r/M)E} \\ &\approx \sum_{m'=0}^{M} \Pr(m^r = m'|m,\Sigma)\, e^{(m'/M)E} \end{aligned} \tag{3.21}$$

In Eq.3.21, specialist B cell class i is characterized by its string matches m with the prime virus' variable epitope string; $\{c^r\}$ are the mass fractions of all boost viruses with $\sum_{r=1}^{N_a} c^r = 1$. We assume equal fraction of all strains: $c^r = 1/N_a$. In the last line of Eq.3.21 we replace the average over all N_a boost viruses with an ensemble average over all possible realizations of boost virus at given Σ. This approximation is thus more accurate when more boost strains are included in the vaccination. Stochastic agent-based simulations with $N_a = 5$ (Fig.3.23) already confirms the existence of optimal intermediate on and off switching times as predicted from affinity class formulations and the approximation for BCR-boost affinity above.

Next, for a generalist $\vec{b}^g$, the same calculation can be repeated for the M_v variable sites on the conserved epitope strings $\vec{a}^{0,c}$ and $\{\vec{a}^{r,c}\}$, $r = 1, 2, ...N_a$. Note that a generalist always has the same string matches m_c against all prime and boost viruses at the M_c conserved sites, since the conserved sites are shared among all prime and boost viruses. Denote (m_v, m_c)

the generalist $\vec{b}^g$'s matching score with the prime virus. Similar to Eq.3.19 we have

$$m_v^r \equiv m_v(\vec{b}^{\,g}, \vec{a}^{r,v}) = m_v - \Sigma_{v,1}^r + \Delta_v^r \tag{3.22}$$

with $\Sigma_{v,1}^r \sim \mathrm{B}(m_v, \frac{\Sigma}{M})$ and $\Delta_v^r \sim \mathrm{B}(M_v - m_v, \frac{\Sigma}{M}\frac{1}{K-1})$. Using the law of total probability,

$$\Pr(m_v^r = m_v'|m_v, \Sigma) = \sum_{\Delta_v'} \Pr(\Delta_v^r = \Delta_v') \cdot \Pr[\Sigma_{v,1}^r = \Delta_v' - (m_v' - m_v)] \tag{3.23}$$

From Eq. 3.10,

$$\begin{aligned} e^{\langle A_i^y(\Sigma)\rangle} &= \left\langle e^{A^y[\vec{b}^g,(\vec{a}^{r,v},\vec{a}^{r,c})]} \right\rangle = \left\langle e^{(m_v^r/M)E} e^{(m_c/M_c)^C(M_c/M)E} \right\rangle \\ &= e^{(m_c/M_c)^C(M_c/M)E} \sum_{i=1}^{N_a} c^r e^{(m_v^r/M)E} \\ &\approx e^{(m_c/M_c)^C(M_c/M)E} \sum_{m_v'=0}^{M_v} \Pr(m_v^r = m_v'|m_v, \Sigma)\, e^{(m_v'/M)E} \end{aligned} \tag{3.24}$$

which can be calculated numerically using Eq.3.23. Again, the average binding affinity between the B cell and boost viruses depends on the B cell's affinity class (characterized by m_v and m_c for generalists), but not explicitly on the full string contents.

3.B Opposite selection force on variable residue binding when antigenic distance Σ crosses a critical value $M(1-1/K)$

In Fig.3.12 we plot probabilities $h_i^s(\Sigma)$ (Eq. 3.4) for B cells of affinity class i to internalize viruses of antigenic distance Σ. Species $s = x$, y for specialists and generalists, respectively. An intriguing observation from Fig.3.12 is that when the boost viruses are similar with the prime virus ($\Sigma < \Sigma_c \equiv 15$), B cells with stronger variable binding with the prime virus also bind stronger with the boost viruses: fitness $h_i^s(\Sigma)$ of B cells increases with m or m_v; however, when the boost strains are distinct enough with the prime virus ($\Sigma > \Sigma_c = 15$), the selection force on the variable binding with prime virus is reversed and $h_i^s(\Sigma)$ increases with lowered m or m_v. The later scenario resembles the "original antigenic sin" effect, in that immune response mounted against a previously-encountered Ag is ineffective against mutated Ag.

We now prove that the critical antigenic distance $\Sigma_c = M(1-1/K)$, where M and K are string length and alphabetical size, respectively. When $M = 20$ and $K = 4$, $M(1-1/K) = 15$ matches exactly with Fig.3.12. Here only the calculation for specialist classes of different m are demonstrated, as generalist classes of different m_v can be treated in the same manner.

We start by rewriting Eq.3.19 as

$$\begin{aligned} m^r &= M - (M - m) - \Sigma_1^r + \Delta^r \\ &= M - [\Sigma_1^r + (M - m - \Delta^r)] \end{aligned} \tag{3.25}$$

Recall in Eq.3.19 that $\Sigma_1^r \sim \mathrm{B}(m, \frac{\Sigma}{M})$ is the number of string sites where the B cell matches the prime virus but mismatches the rth boost virus; $\Delta^r \sim \mathrm{B}(M - m, \frac{\Sigma}{M}\frac{1}{K-1})$ is the number of string sites where the B cell mismatches with the prime virus but matches with the rth boost virus. Since $M - m$ is the total number of string mismatches between the B cell and the prime virus, the last term $M - m - \Delta^r$ in Eq.3.25 is then the number of string sites where the B cell mismatches both the prime and the rth boost virus. As such, Eq.3.25 can be interpreted as: number of BCR-boost string matches (m^r) = total string length (M) − number of string sites where BCR mismatches the boost virus but matches the prime virus (Σ_1^r) − number of string sites where BCR mismatches both the boost and prime virus ($M - m - \Delta^r$).

Since the binomially-distributed Δ^r is the number of successful trials out of $(M - m)$ repeated trials with success probability $\frac{\Sigma}{M}\frac{1}{K-1}$, $(M-m-\Delta^r)$ gives the number of failed trials and is also binomially distributed: $M-m-\Delta^r \sim \mathrm{B}(M-m, 1-\frac{\Sigma}{M}\frac{1}{K-1})$. In our string model, a successful/failed trial corresponds to a BCR-boost string match/mismatch. Therefore, the sum $\Sigma_1^r + (M-m-\Delta^r)$ in the bracket of Eq.3.25 can be viewed mathematically as the total number of successes from two independent sets of binomial experiments, with M trials in total: the first set includes m trials and the success probability is $\frac{\Sigma}{M}$; the second set includes $M - m$ trails and the success probability is $1 - \frac{\Sigma}{M}\frac{1}{K-1}$.

When $\Sigma = M(1 - 1/K)$, $\frac{\Sigma}{M} = 1 - \frac{\Sigma}{M}\frac{1}{K-1} = 1 - \frac{1}{K}$. That is, the two sets of binomial experiments share the same success probability, and the total number of successes obeyes a simple binomial distribution: $\Sigma_1^r+(M-m-\Delta) \sim B(m+M-m, 1-\frac{1}{K}) = B(M, 1-\frac{1}{K})$. This

combined with Eq.3.25 prove that the probability distribution of m^r $\Pr(m^r = m'|m, \Sigma = M(1-1/K))$ is independent of m. In turn, specialists' average BCR-boost binding affinity $\langle A_i^x(\Sigma)\rangle$ (Eq.3.21), as well as the probability $h_i^x(\Sigma)$ to internalize boost Ag (Eq. 3.4), are also independent of m.

Finally, when $\Sigma > M(1-1/K)$, $\frac{\Sigma}{M} > 1-\frac{\Sigma}{M}\frac{1}{K-1}$. That is, the success probability is higher for $\Sigma_1^r \sim \mathrm{B}(m, \frac{\Sigma}{M})$ than for $M-m-\Delta^r \sim \mathrm{B}(M-m, 1-\frac{\Sigma}{M}\frac{1}{K-1})$. When m increases, the binomial distribution for Σ_1^r will include more trials and the distribution of $\Sigma_1^r+(M-m-\Delta^r)$ will be more skewed towards higher values. Thus, from Eq.3.25 the distribution of m^r will be more skewed towards lower values. As a result, from Eq.3.21 the average BCR-boost binding affinity $\langle A_i^x(\Sigma)\rangle$ decreases with m when $\Sigma > M(1-1/K)$. Similarly, $\langle A_i^x(\Sigma)\rangle$ increases with m when $\Sigma < M(1-1/K)$, matching exactly with Fig.3.12.

3.C Equilibrium point of GC B cell dynamics shows the asymptotic effect of Ag concentration and antigenic distance

The equilibrium point $\vec{z}_{\mathrm{eqm}}(C_{Ag}, \Sigma)$ of GC B cells is defined as the solution to $\mathrm{d}\vec{z}/\mathrm{d}t = 0$, where $\vec{z} \equiv \binom{\vec{x}}{\vec{y}} \in \mathbb{R}^{M+1+(M_v+1)(M_c+1)}$ is the B cell population vector including both specialists $\vec{x}$ and generalists $\vec{y}$ (see Methods). $\vec{z}_{\mathrm{eqm}}$ is the terminal state that the GC B cells will be evolving towards in the asymptotic regime of $t \to \infty$, given constant vaccine parameters including the concentration C_{Ag} of Ag in GC and the antigenic distance Σ of boost viruses.

To solve for $\vec{z}_{\mathrm{eqm}}$, we first rewrite the growth rates of specialists and generalists (Eq. 3.3) together as

$$\begin{aligned}\frac{\mathrm{d}z_\alpha}{\mathrm{d}t} =& k_p\, z_\alpha\, \theta_\alpha\, h_\alpha(C_{Ag}, \Sigma)\cdot(2\mathfrak{M}_{\alpha\alpha}-1)\\ &+ k_p \sum_{\beta\neq\alpha} z_\beta\, \theta_\beta\, h_\beta(C_{Ag}, \Sigma)\cdot 2\mathfrak{M}_{\alpha\beta}\\ &- k_p\, z_\alpha\, \theta_\alpha\, [1-h_\alpha(C_{Ag}, \Sigma)+\frac{|\vec{z}|}{K_B}]\end{aligned} \tag{3.26}$$

where $\vec{h} \equiv \begin{pmatrix}\vec{h}^x\\ \vec{h}^y\end{pmatrix}$ and $\mathfrak{M} \equiv \begin{pmatrix}\mathfrak{M}^x & 0\\ 0 & \mathfrak{M}^y\end{pmatrix}$ are the fitness function and mutation matrix for all GC B cells, respectively. Both C_{Ag} and Σ enter Eq.3.26 through Ag internalization probability

$\vec{h}$ (Eq. 3.4).

By substituting Eq.3.26's left-hand side with zero and by separating the linear and the quadratic terms in $\vec{z}$ on the right-hand side, the equation $\mathrm{d}\vec{z}/\mathrm{d}t = 0$ can be rewritten in the matrix form

$$L\vec{z} = (\frac{k_p}{K_B}\,|\vec{z}|)\vec{z} \tag{3.27}$$

where L is a constant matrix: $L_{\alpha\beta} = k_p[2h_\beta(\Sigma)\mathfrak{M}_{\alpha\beta} - \delta_{\alpha\beta}]$. Note that we have removed the non-linearity from step functions θ_α in Eq.3.26 by replacing them with constant 1, since B cell classes below existence threshold ($z_\alpha < 1$) will have a negligible effect in deciding the equilibrium point.

At equilibrium, total B cell number $|\vec{z}_{\mathrm{eqm}}|$ is a constant as $\mathrm{d}\vec{z}_{\mathrm{eqm}}/\mathrm{d}t = 0$. From Eq.3.27, for any nonzero equilibrium point $\vec{z}_{\mathrm{eqm}}$, $\frac{k_p}{K_B}\,|\vec{z}_{\mathrm{eqm}}|$ should be an eigenvalue of matrix L, and $\vec{z}_{\mathrm{eqm}}$ is L's corresponding eigenvector. The nonzero equilibrium point $\vec{z}_{\mathrm{eqm}}$ can thus be solved numerically by the following steps:

(1): Find all the eigenvalues $\{\Lambda\}$ of matrix L. Keep only the positive ones since $0 < |\vec{z}| = \Lambda\frac{K_B}{k_p}$

(2): Find corresponding eigenvectors $\{\vec{v}\}$ and keep only those whose all components are of the same sign or zero. This is because $\vec{v}$ is parallel to $\vec{z}_{\mathrm{eqm}}$ and all of $\vec{z}_{\mathrm{eqm}}$'s components are positive or zero.

(3): The remaining eigenvalue and eigenvector pairs $\{\Lambda, \vec{v}\}$ lead to nonzero equilibrium points $\vec{z}_{\mathrm{eqm}} = \frac{\vec{v}}{|\vec{v}|}\frac{K_B}{k_p}\Lambda$.

Following these steps, at most two solutions are found for varied Ag concentration C_{Ag} and antigenic distance Σ: one with only specialists and one only generalists—no coexistence of the two species is identified at equilibrium.

Fig.3.13 plots the generalist population size $|\vec{y}|$ and average conserved binding score $\langle m_c\rangle$ at equilibrium. The generalist cross-reactivity $\langle m_c\rangle$ improves at lower Ag concentration C_{Ag} as the selection is stronger, and this comes at the cost of lowered generalist quantity. Interestingly, at fixed C_{Ag}, the optimal generalist cross-reactivity at equilibrium peaks at $\Sigma = \Sigma_c = M(1 - 1/K) = 15$: from Fig.3.12, generalists can reach higher fitness h^y by

evolving towards higher m_v when $\Sigma < \Sigma_c = 15$, or towards lower m_v when $\Sigma > \Sigma_c$ (see section 3.B). As a result, generalists of higher m_c do not impose higher fitness than those of lower m_c as they can have different m_v values, and the selection for high cross-reactivity m_c is thus not the most efficient. At $\Sigma = \Sigma_c$, however, there is no fitness improvement of generalists from the evolution of variable binding m_v, and the selection for high cross-reactivity m_c is the most efficient.

3.D Optimal antigenic distance at large $\Sigma_{\max} > M(1 - 1/K)$

In this chapter we optimized antigenic distance $\Sigma(t)$ when $0 \leq \Sigma(t) \leq \Sigma_{\max} = 10$. This corresponds to less than $\Sigma_{\max}/M = 50\%$ antigenic difference between the prime and boost viruses, which is higher than the viral diversity reported for influenza ([210]) or HIV ([230]). Here we demonstrate that when $\Sigma_{\max} > M(1-1/K) = 15$, the optimal control still exhibits the on switch from 0 to $\Sigma_{\max}$, and the high $\Sigma(t)$ period lasts to suppress variable binding and select for generalists of high cross-reactivity. Unlike when $\Sigma_{\max} < M(1 - 1/K) = 15$ though, we show that additional oscillations in $\Sigma^*(t)$ further improves Φ_{evo} when $\Sigma_{\max} > M(1 - 1/K) = 15$ (Fig.3.21). Nevertheless, the additional oscillations only brings marginal improvement over the simple switching $\Sigma(t)$ function from zero to $\Sigma_{\max}$, and therefore does not warrant applications when considering the significantly increased complexity to the vaccine protocol.

We employ a steepest ascent method (see Algorithm 2 for details) to find the optimal antigenic distance that maximizes generalist evolutionary index Φ_{evo} for different $\Sigma_{\max}$ and T. Such method gives *local* optima $\Sigma^\dagger(t)$ in the space of $\Sigma(t)$ functions on a discretized grid of antigenic distance Σ and time t. The switching antigenic functions from PMP are reproduced at small $\Sigma_{\max} < 15$ (Fig.3.21B, lower panel). However, at large $\Sigma_{\max} > 15$ and long protocol duration T (Fig.3.21A, diamonds), our iteration algorithm based on PMP (Algorithm 1 in Appendices) does not yield convergence while using the steepest ascent method we are able to identify chattering behavior in the local optima $\Sigma^\dagger(t)$ (Fig.3.21B, upper panel). We speculate the chatterings ([238, 239, 240]) are what lead to numerical

divergence in Algorithm 1.

As diverse boost viruses suppress variable binding and promote the evolution of generalist cross-reactivity, it appears counter-intuitive at first that oscillations between $\Sigma(t) = \Sigma_{\max} > 15$ and $\Sigma(t) = 0$ would improve generalist quality. However, such improvement is indeed confirmed when we compare the evolutionary index Φ_{evo} with or without oscillations, and with different oscillation frequencies (Fig.3.21C). It also turns out that the intervals with $\Sigma(t) = \Sigma_{\max}$ should still dominate during chattering and last longer than the intervals with $\Sigma(t) = 0$, so that a higher $\Phi_{\max}$ than no chattering can be achieved. That is, the optimal antigenic distance $\Sigma^{\dagger}(t)$ still maintains the overall switching shape as when $\Sigma_{\max} < 15$, and is decorated with short intervals of $\Sigma(t) = 0$ after the on switch.

To understand the benefit of short intervals with $\Sigma(t) = 0$ after on switch, we track the average number of generalist string matches $\langle m_v \rangle$ and $\langle m_c \rangle$ over time (Fig.3.21D). Before high $\Sigma(t)$ is first switched on ($t < 50$), $\Sigma(t) = 0$ and m_v increases over time for stronger variable binding. After switching, $\Sigma(t) = 20 > M(1 - 1/K) = 15$ and the selection force on m_v becomes the opposite: the generalist fitness is now higher with lower m_v (see section 3.B). With the chattering between $\Sigma(t) = 0 < 15$ and $\Sigma(t) = 20 > 15$, the evolution of m_v towards higher fitness is frustrated by the opposite selection forces. As a result, generalists with strong conserved binding (higher m_c) are more preferred by selection which improves generalist evolution. Finally, since such chatterings are caused by the opposite selection forces on variable binding when $\Sigma(t)$ crosses $M(1 - 1/K) = 15$, the chatterings will only emerge when maximum allowed $\Sigma_{\max} > 15$.

3.E Simultaneous optimization of antigenic distance and Ag concentration

When designing a vaccine protocol, both the boost viral strain composition (Σ) and Ag concentration in GC (C_{Ag}, which we assume to be proportional to the vaccine dosage ([59])) can be dynamic. Extension of our model to simultaneous optimization of both parameters is straightforward using PMP. To do this, in addition to the dynamics of state (Eq.3.26) and

costate vectors (Eq. 3.11), simultaneous optimization of Hamiltonian H pointwise in time by both optimal $\Sigma^*(t)$ and $C^*_{Ag}(t)$ is required, i.e., Eq.3.12 is replaced by

$$(\Sigma^*(t), C^*_{Ag}(t)) = \underset{(\Sigma', C'_{Ag})}{\arg\max}\, H(\vec{z}(t), \vec{\lambda}(t), \Sigma', C'_{Ag}), \forall t \in [0, T]. \tag{3.28}$$

The method of successive approximation for numerically finding the optimal control (Algorithm 1) still holds, while both $\Sigma(t)$ and $C_{Ag}(t)$ are updated in each iteration.

3.F Stochastic agent-based simulation of GC reaction

GC B cell dynamics written as ordinary differential equations (Eq. 3.3 or Eq.3.26) can be interpreted as rate equations of cell division and apoptosis processes. Thus, Eq.3.26 can be realized with stochastic agent-based simulations using Gillespie algorithm ([241]), and by tracking full strings of B cells and viruses over time. This can also be complemented by Tau-leaping ([242]) method to improve simulation efficiency.

To apply Gillespie algorithm to Eq.3.26, one has to find the B cell fitness h in a full string representation. For a B cell $\vec{b}$ and N_a viral strains $\mathfrak{a} \equiv \{(\vec{a}^{r,v}, \vec{a}^{r,c})\}$ (r = 1, 2, ...N_a), the BCR-virus binding affinities are found by counting the string matches for each BCR-virus pair and plug into Eq. 3.9 or 3.10 for specialists or generalists, respectively. The fitness $h[\vec{b}, \mathfrak{a}(t)]$ can then be calculated as

$$h[\vec{b}, \mathfrak{a}(t)] = \frac{[C_{Ag}(t)/N_a] \sum_{r=1}^{N_a} \mathrm{e}^{A[\vec{b},(\vec{a}^{r,v},\vec{a}^{r,c})]}}{1 + [C_{Ag}(t)/N_a] \sum_{r=1}^{N_a} \mathrm{e}^{A[\vec{b},(\vec{a}^{r,v},\vec{a}^{r,c})]}} \tag{3.29}$$

According to Eq.3.26, the cell division process for a B cell $\vec{b}$:

$$\vec{b} \to \vec{b}\,' + \vec{b}\,'' \tag{3.30}$$

proceeds at rate $k_p h[\vec{b}, \mathfrak{a}(t)]$. Affinity-altering or lethal mutations can occur in each daughter cell with probability $p_{\mathrm{aa}}\mu$ or $p_{\mathrm{let}}\mu$, respectively (Eq.3.17, also see section 3.A.3). In the event of an affinity-altering mutation, a random string site of daughter B cell is mutated to a random new letter; in the event of a lethal mutation, the daughter B cell is immediately removed from simulation: $\vec{b}\,'$ (or $\vec{b}\,''$) $\to \varnothing$.

The apoptosis process for a B cell $\vec{b}$:

$$\vec{b} \to \varnothing \tag{3.31}$$

occurs at rate $k_p\{1 - h[\vec{b}, \mathfrak{a}(t)] + \frac{N_B(t)}{K_B}\}$. $N_B(t)$ is the total number of B cells at time t.

Algorithm 1: Successive approximation algorithm for finding the optimal antigenic distance from Pontryagin's maximum principle

Step 1: Initialization

Trial control $\Sigma^0(t)$;

Iteration counter $k = 0$;

Iteration flag *Next*=true;

Step 2: Iterations

while *Next* **do**

  $k = k + 1$;

  Find $\vec{z}^k$ by integrating $\frac{\mathrm{d}\vec{z}^k}{\mathrm{d}t} = \vec{f}(\vec{z}^k, \Sigma^k)$ forward in time with $\vec{z}^k(t = 0) = \vec{z}_0$ fixed;

  Find $\vec{\lambda}^k$ by integrating $\frac{\mathrm{d}\lambda_\alpha^k}{\mathrm{d}t} = -\sum_\beta \lambda_\beta \frac{\partial f_\beta(\vec{z},\Sigma)}{\partial z_\alpha} - \frac{\partial \phi(\vec{z})}{\partial z_\alpha}$ backward in time with $\vec{\lambda}^k(t = T) = 0$;

  Find $\hat{\Sigma}^k(t) = \arg\max_{\Sigma'} H(\vec{z}^k(t), \vec{\lambda}^k(t), \Sigma')$;

  Find iteration error $\delta \equiv \int_0^T |\Sigma^k - \hat{\Sigma}^k| \frac{\mathrm{d}t}{T}$

  if $\delta <$ *Tolerance* ***or*** $k >$ *#MaxIterations* **then**

    Next=false;

  else

    Update the control function: $\Sigma^{k+1}(t) = \rho\Sigma^k(t) + (1 - \rho)\hat{\Sigma}^k(t)$ with damping constant ρ

  end

end

Algorithm 2: Steepest ascent algorithm for finding the local optima of antigenic distance function

Step 1: Discretization of control

Antigenic distance values: $[0, \Sigma_{\max}] \to 0,\ \delta\Sigma,\ 2\delta\Sigma, \ldots \Sigma_{\max}$;

Time: $[0, T] \to [0, \delta t],\ [\delta t, 2\delta t],\ \ldots [T - \delta t, T]$; antigenic distance is constant on each interval

Step 2: Iterations

Number of local optima found $Num\_sol = 0$;

while $Num\_sol < 50$ **do**

  Randomly initialize trial antigenic distance $\Sigma^0(t)$ with discretized Σ and t in Step 1;

  Iteration counter $k = 0$;

  Iteration flag *Next*=true;

  while *Next* **do**

    $k = k + 1$;

    Solve B cell trajectories $\vec{z}^k(t)$ for $\Sigma^k(t)$ from Eq.3.26;

    Find generalist evolutionary index Φ^k_{evo} from Eq. 3.5;

    Enumerate all $n = \frac{T}{\delta t}\frac{\Sigma_{\max}}{\delta\Sigma}$ "nearest neighbors" $\{\widetilde{\Sigma}^{k,i}(t)\}$ $(i = 1, 2, \ldots n)$ of the current solution $\Sigma^k(t)$: they differ with $\Sigma^k(t)$ on exactly one discretized time interval;

    Solve B cell trajectories $\{\widetilde{\vec{z}}^{k,i}(t)\}$ for all nearest neighbors $\{\widetilde{\Sigma}^{k,i}(t)\}$;

    Find generalist evolutionary indices $\{\widetilde{\Phi}^{k,i}_{\text{evo}}\}$ for all nearest neighbors $\{\widetilde{\Sigma}^{k,i}(t)\}$;

    if $\Phi^k_{evo} > \max_i\{\widetilde{\Phi}^{k,i}_{evo}\}$ **then**

      Next=false;

      Num_sol=*Num_sol*+1;

      Record current $\Sigma^k(t)$ as a local optimum;

    else

      Update the antigenic distance function: $\Sigma^{k+1}(t) = \widetilde{\Sigma}^{k,i^*}(t)$ where $i^* = \arg\max_i\{\widetilde{\Phi}^{k,i}_{\text{evo}}\}$

    end

  end

end

Table 3.1: **Simulation parameters in chapter 3**

(parameters from generic/AM model are marked white/gray)

Parameter	Value	Description	Reference or descriptions
d	1	Baseline death rate	
c	$1e-4$	Reciprocal of population capacity	
f_M	2	Maximum birth rate	
D_g	$4e-5$	Specialist trait diffusion constant	
D_s	$4e-5$	Specialist trait diffusion constant	
u_M	1	Maximum control strength	
N_c	1000	Lower population threshold	
M	20	B cell and virus string length	[152, 46]
M_v	10	Number of variable sites on conserved epitope string	
M_c	10	Number of conserved sites on conserved epitope string	
K	4	String alphabet size	[152, 46, 24]
k_p	1	GC reaction rate	1 GC cycle (6-12h, [1]) per time unit
K_B	10000	GC B cell capacity	[153]
E	10	Affinity scaling constant	Exponential binding affinity can improve by thousand-fold ([198, 243]) from founder to full string match
C	1.1	Constant in Eq. 3.10 that penalizes the lower target accessibility of generalists	
μ	0.5	Mutation probability per daughter cell per division	[155, 156]
p_{let}	0.3	Fraction of lethal BCR mutations	[155, 156]
p_{aa}	0.2	Fraction of affinity-altering BCR mutations	[155, 156]

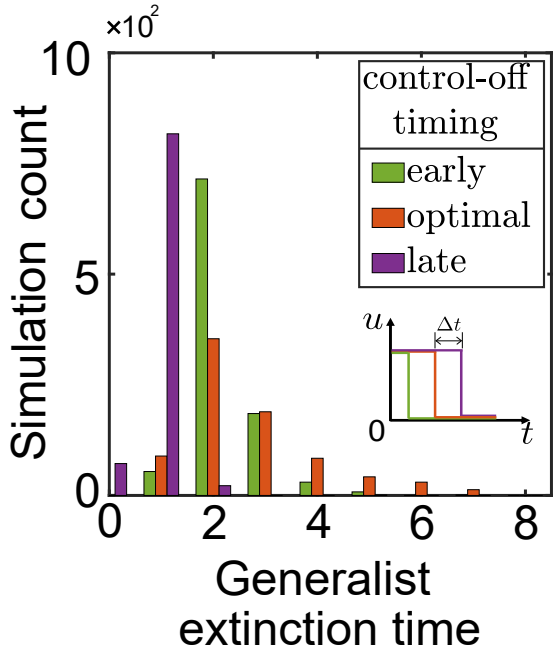

Figure 3.9: **Generalists survive the longest at optimal control switch, related to Fig. 3.2B.** Bar colors match the colored squares in Fig. 3.2B. Red: switch is optimal ($\Delta t = 0$); green (purple): switched is earlier(later) than optimal by $|\Delta t| = 0.5$. Data are collected from the same 1000 simulations as Fig. 3.2B for each condition.

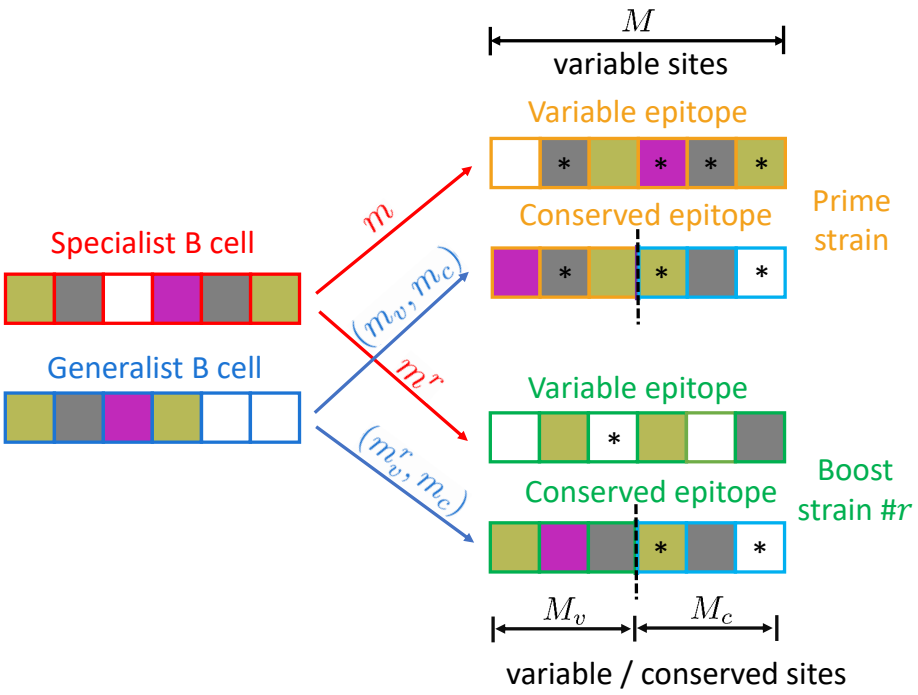

Figure 3.10: **Illustration of string representation:** B cells and virus strains are encoded by strings of length M. The last M_c sites (light blue boundary) on the conserved epitope strings are conserved among viral strains while the first $M_v = M - M_c$ sites are variable and strain-specific. A B cell's matching score is found by comparing with its target (variable or conserved) epitope string at each site and count the matches; for generalists the matches from variable and conserved sites m_v and m_c are counted separately. In the example illustration $M = 6$, $M_v = M_c = 3$. For the specialist B cell $m = 4$ against the prime strain and $m^r = 1$ against a boost strain; for generalist B cell $m_v = 1$ against the prime strain and $m_v^r = 0$ against the boost strain; $m_c = 2$ against both strains. String contents are coded by filling colors and matching sites with B cells are marked by asterisks.

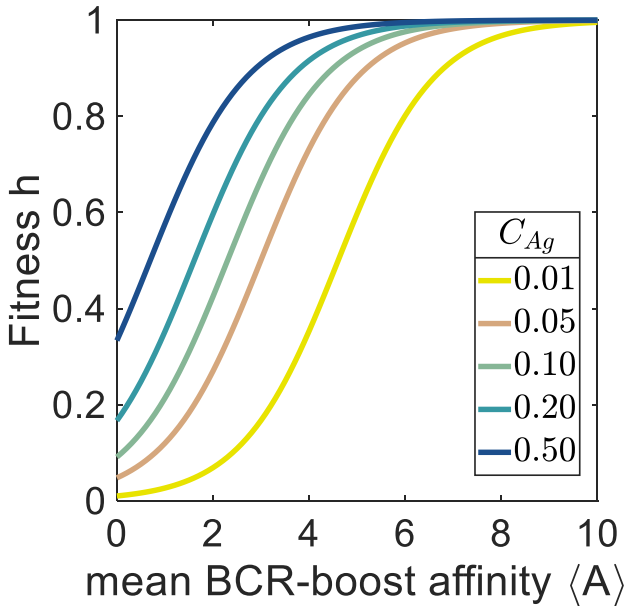

Figure 3.11: **B cell fitness h as a function of mean binding affinity $\langle A \rangle$ between the B cell and boost Ag.** At lower Ag concentration C_{Ag}, fitness h is more sensitive to binding affinity and the selection for high-affinity B cell is thus stronger. Fitness is calculated as $h = \frac{C_{Ag}\exp(\langle A \rangle)}{1+C_{Ag}\exp(\langle A \rangle)}$ for both generalists and specialists according to Eq. 3.4

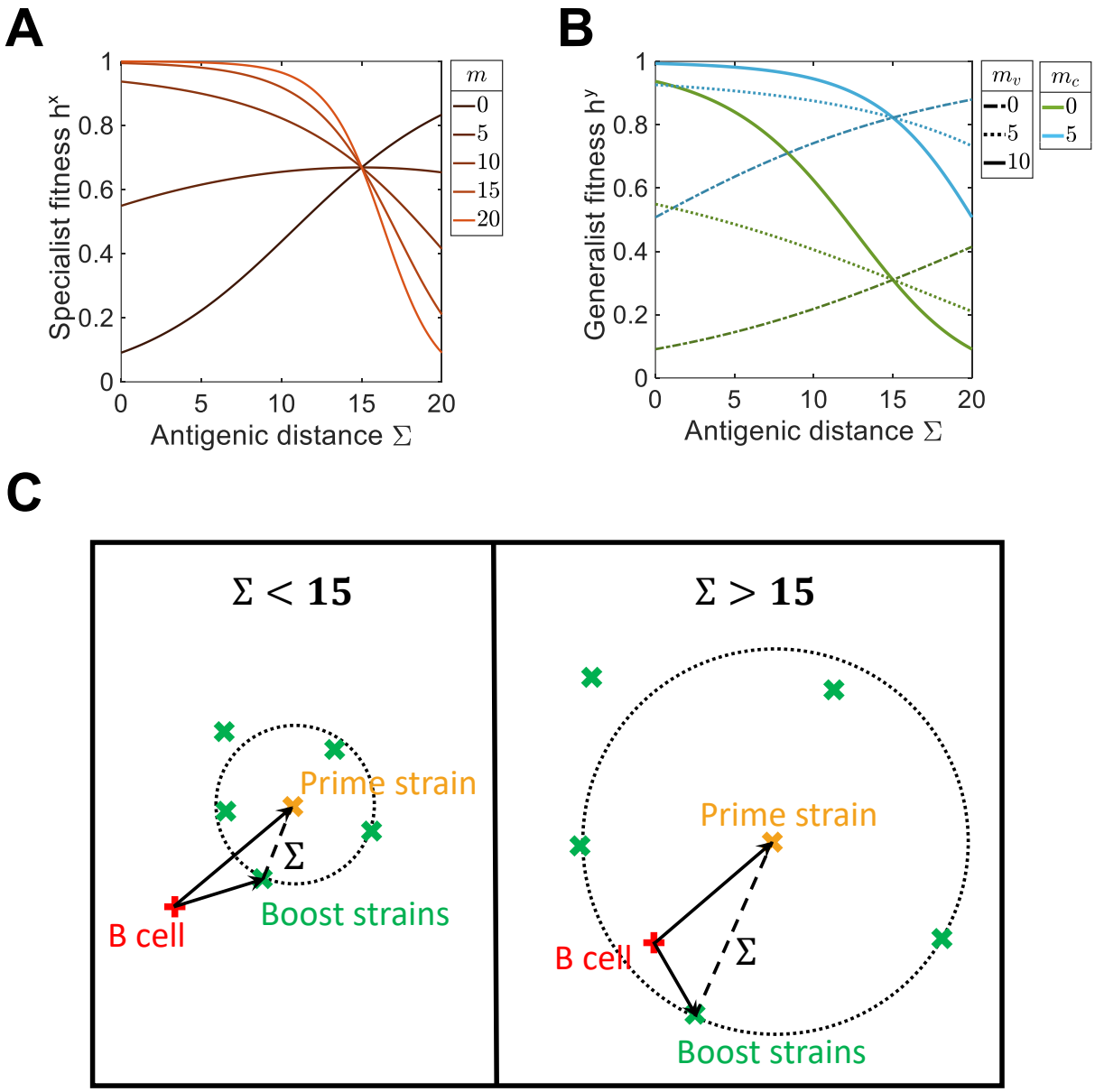

Figure 3.12: **Fitness h of different B cell classes**. (A)/(B): Fitness h as in Eq.3.4 for different specialist/generalist classes. At small antigenic distance $\Sigma < 15 = M(1 - \frac{1}{K})$, B cells that bind stronger with the prime strain also binds stronger with the boost strains, so specialists/generalists with higher m/m_v have higher fitness. At large $\Sigma > 15$, boost strains are so distinct from the prime Ag that B cells binding stronger with the prime Ag now binds weaker with the boost Ag; thus specialists/generalists with lower m/m_v have higher fitness. Regardless of Σ, stronger conserved binding (larger m_c for generalists) always promotes fitness. Total string length $M = 20$, alphabet size $K = 4$, Ag concentration $C_{Ag} = 0.1$. (C): Geometric illustration of the opposite selection force on variable binding m/m_v in panels (A)(B) when Σ crosses critical value 15. Closer distance corresponds to smaller Hamming distance in string space. Boost strains are distributed around a circle of radius Σ centered at the prime strain. At small Σ, B cell is on the exterior of the circle and evolving closer to its center (prime) also means closer to its boundary (boost). At large Σ, B cell is on the interior of the circle and evolving closer to the center (prime) now brings the B cells further away from the boundary (boost).

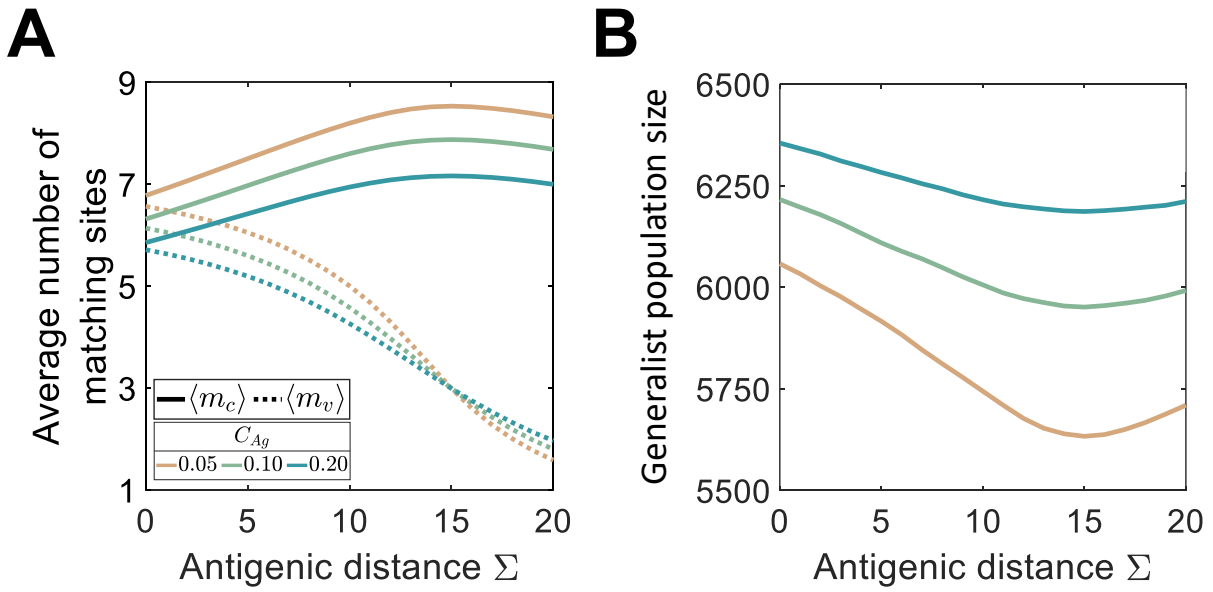

Figure 3.13: **Generalist quality and quantity at equilibrium state with constant Ag concentration C_{Ag} and antigenic distance** Σ. (A)/(B): Generalist string matches and population size at equilibrium, for different constant Σ and Ag concentration C_{Ag}. Stronger selection at lower C_{Ag} selects for higher generalist quality $\langle m_c \rangle$ at the cost of quantity. Note that from Fig.3.12, generalists are fitter with higher/lower m_v when $\Sigma < 15$ or > 15. Thus lower C_{Ag} selects for higher/lower $\langle m_v \rangle$ when $\Sigma < 15$ or > 15.

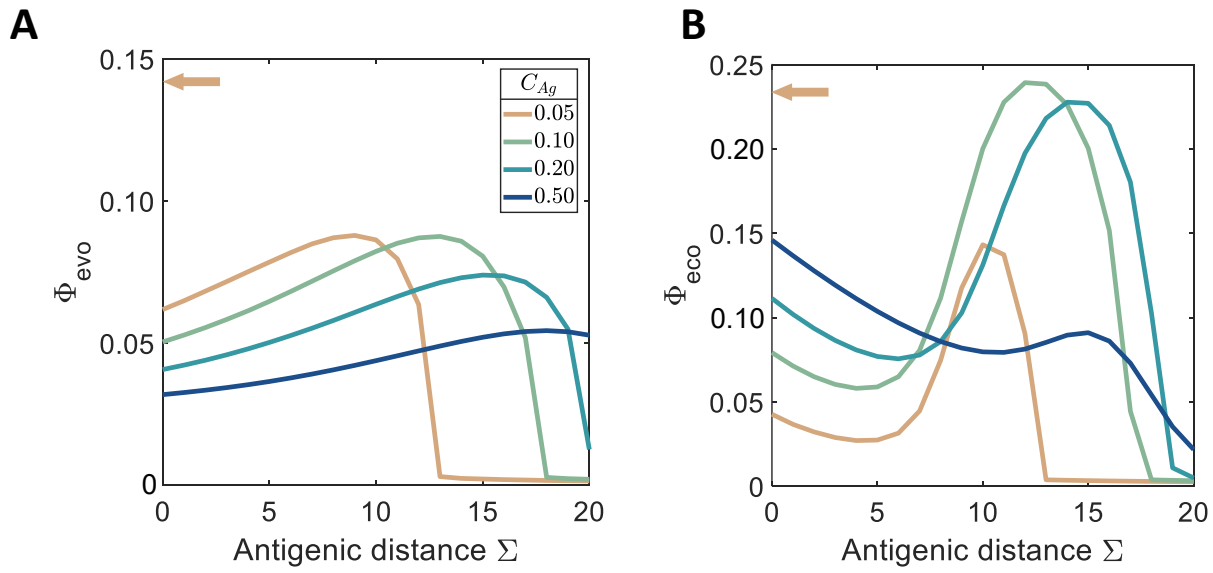

Figure 3.14: **Intermediate constant antigenic distance leads to peak objective indices, and is outperformed by dynamic antigenic distance function; related to Fig. 3.3. Evolutionary and ecological objective indices Φ_{evo} and Φ_{eco} (Eq. 3.5) as a function of constant antigenic distance Σ and Ag concentration C_{Ag}. Arrows on vertical axis show the results from the dynamic switching $\Sigma(t)$ function that maximizes Φ_{evo} when $C_{\text{Ag}} = 0.05$ (which correspond to the gray diamond as in Fig. 3.3A).**

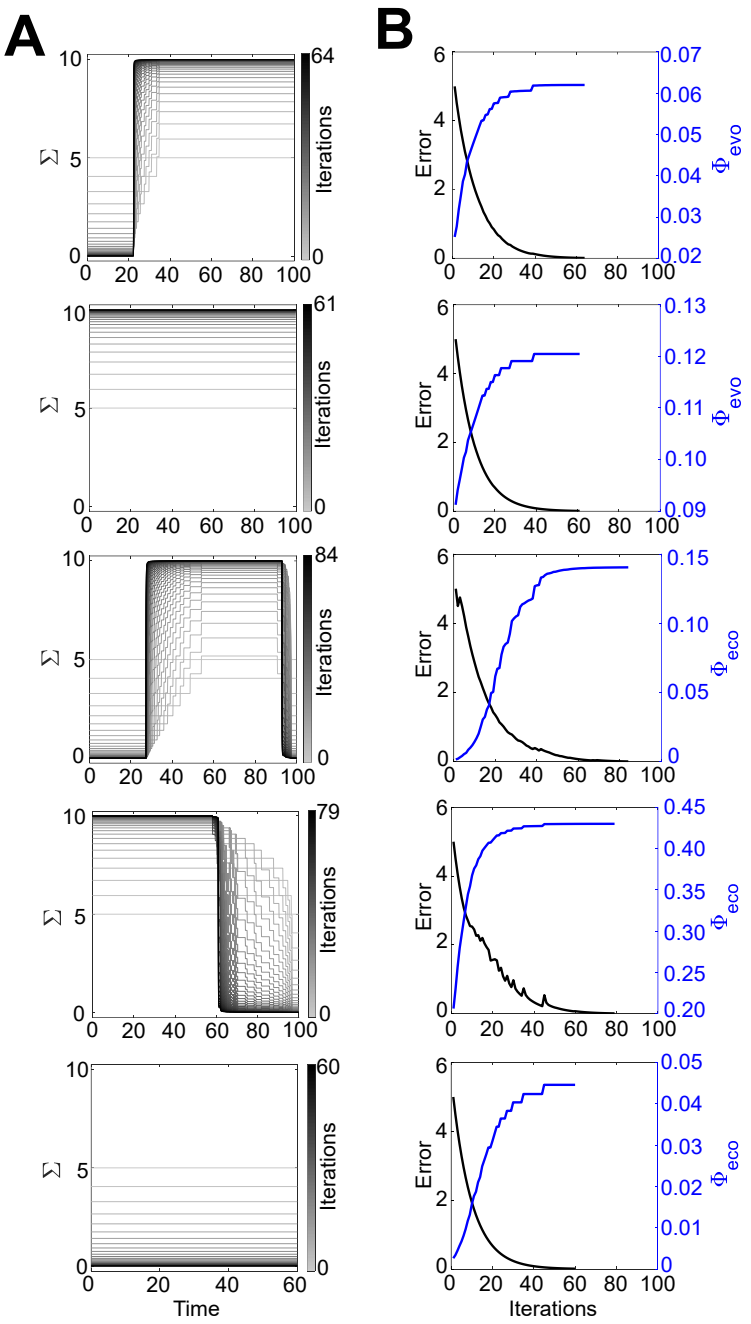

Figure 3.15: **Evolution of control function and objective index during PMP optimization, related to Fig. 3.4.** (A): Control function $\Sigma(t)$ over the iterations. (B): Values of iteration error and objective index over the iterations. Error is as defined in Algorithm 1. Five rows correspond to five control shapes as in Fig. 3.4A. From top to bottom the parameters and optimized objective indices from Fig. 3.4 are: $m_c = 0, T = 100$, Φ_{evo}; $m_c = 4, T = 100$, Φ_{evo}; $m_c = 0, T = 100$, Φ_{eco}; $m_c = 4, T = 100$, Φ_{eco}; $m_c = 0, T = 60$, Φ_{eco}.

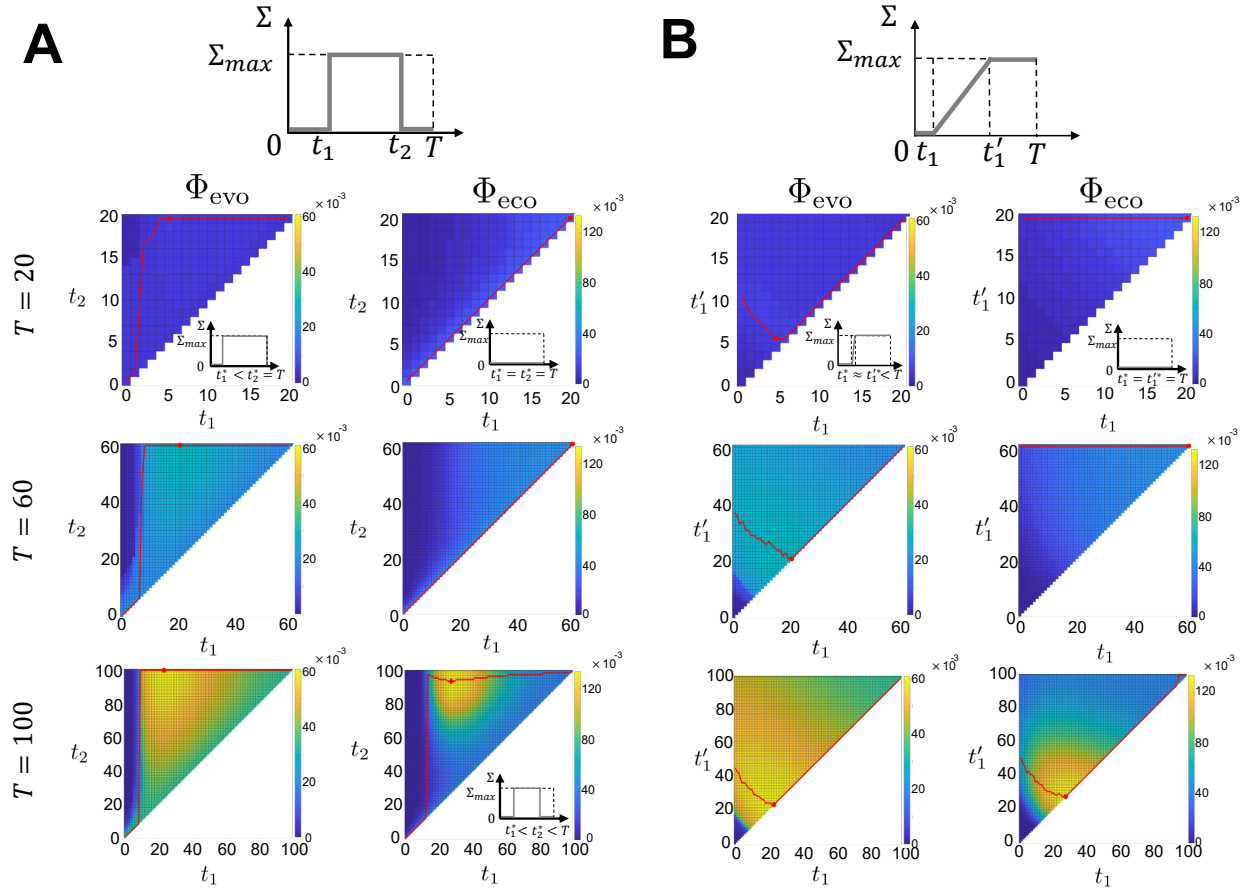

Figure 3.16: **Objective indices' dependence on control switching time and steepness at different vaccine protocol duration T, related to Fig. 3.4.** (A) Objective indices Φ_{evo} or Φ_{eco} as a function of timings $t_1 \leq t_2$ to turn on/off control Σ_{max} (see the top schematics). Red lines show the optimal t_2^* for each t_1 and red dots mark the global optimal over (t_1,t_2) pairs. A peak of finite size surrounds the global optimum and tolerate imprecision when switching on/off high Σ. (B) Objective indices Φ_{evo} or Φ_{eco} as a function of the rising edge position $t_1 \leq t_1'$ when transitioning from 0 to Σ_{max} is linear instead of bang-bang (see top schematics). Red line shows the optimal $t_1'^*$ for each t_1 and red dots mark the global optimal over (t_1, t_1') pairs. Like in (A), a peak of finite size surrounds the global optimum and tolerates imprecision when switching on high control. Schematic insets of both panels showcase distinct shapes of optimal $\Sigma(t)$ functions which are consistent with PMP results in Figs.4A and 4B. $C_{Ag} = 0.1$, $\Sigma_{\max} = 10$, and initial conditions are the same as Fig.fig:OptCtrlC.

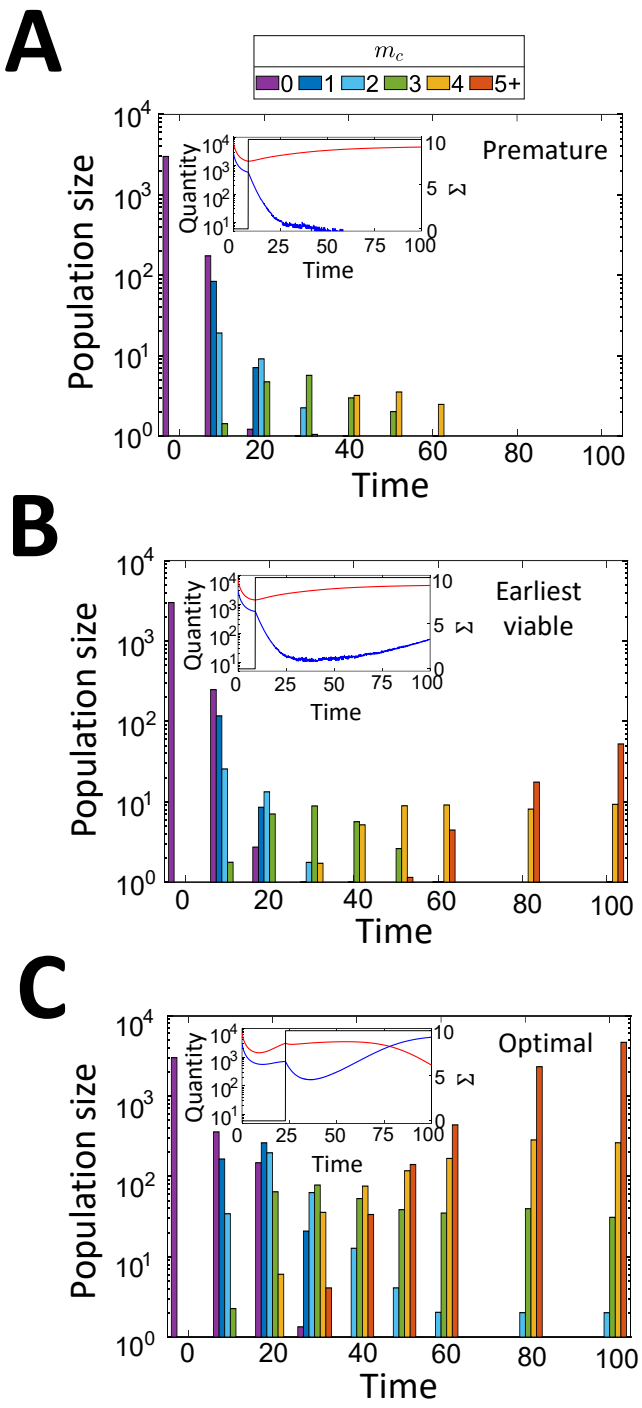

Figure 3.17: **The earliest viable switch for generalist survival precedes the optimal switching time, related to Fig. fig:SWTiming.** Generalists' composition and population size at different switching times. Generalists are binned and counted based on their m_c values at each plotted time point. (A) the switching time $t_1 = 8$ is premature when generalists remain at low abundance and quality, and eventually go extinct; (B)the switching time $t_1 = 9$ is the earliest viable switching time when generalists go through a deep bottleneck and barely recover; (C) the switching time $t_1 = 24$ is optimal when the generalists are allowed enough time to evolve before $\Sigma_{\max}$ is turned on, so that they go through a shallower bottleneck and recover faster. Inset shows the corresponding antigenic distance function $\Sigma(t)$ (black) and total generalist and specialist population size $|\vec{y}|$, $|\vec{x}|$ (blue and red, respectively). $m_c = 0$ for generalist founders.

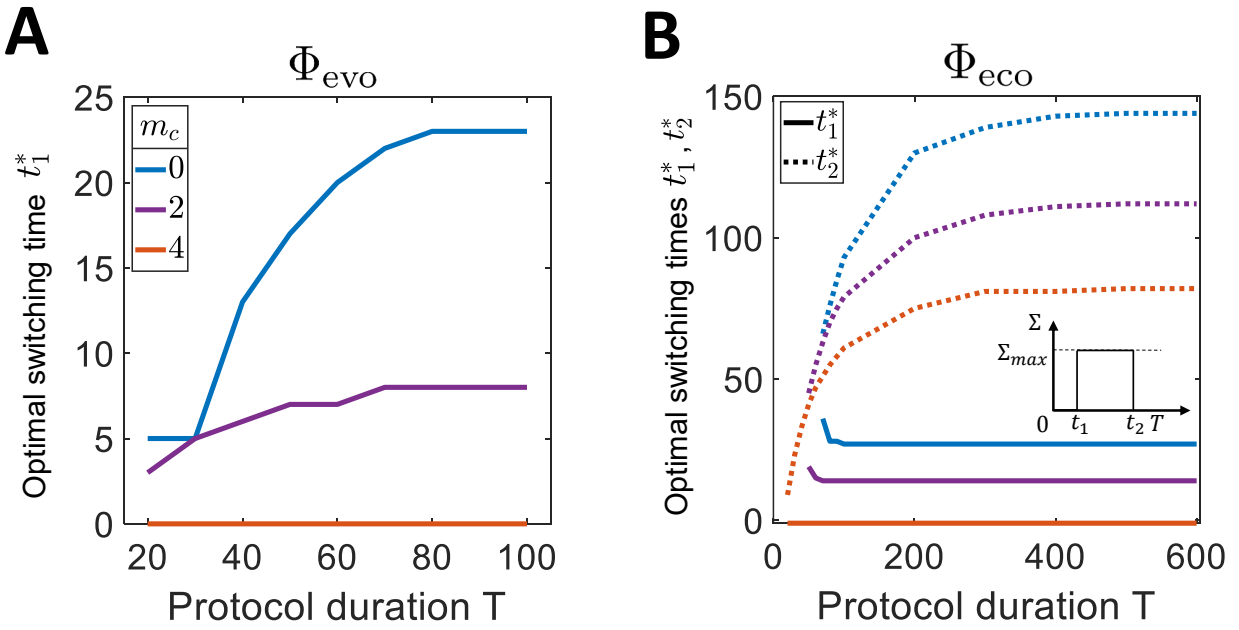

Figure 3.18: **Dependence of optimal control switching times t_1^* and t_2^* on vaccination protocol duration T.** (A) control maximizes generalist quality (Φ_{evo}); optimal off-timing is $t_2^*(T) = T$ (not plotted). (B) control maximizes generalist quantity (Φ_{eco}). The control schematic is shown in inset of panel B. $C_{Ag} = 0.1$ and generalist founders' $m_v = 6$; m_c are as labeled. Results are scanned among all (t_1^*,t_2^*) pairs with time resolution $\Delta t = 1$.

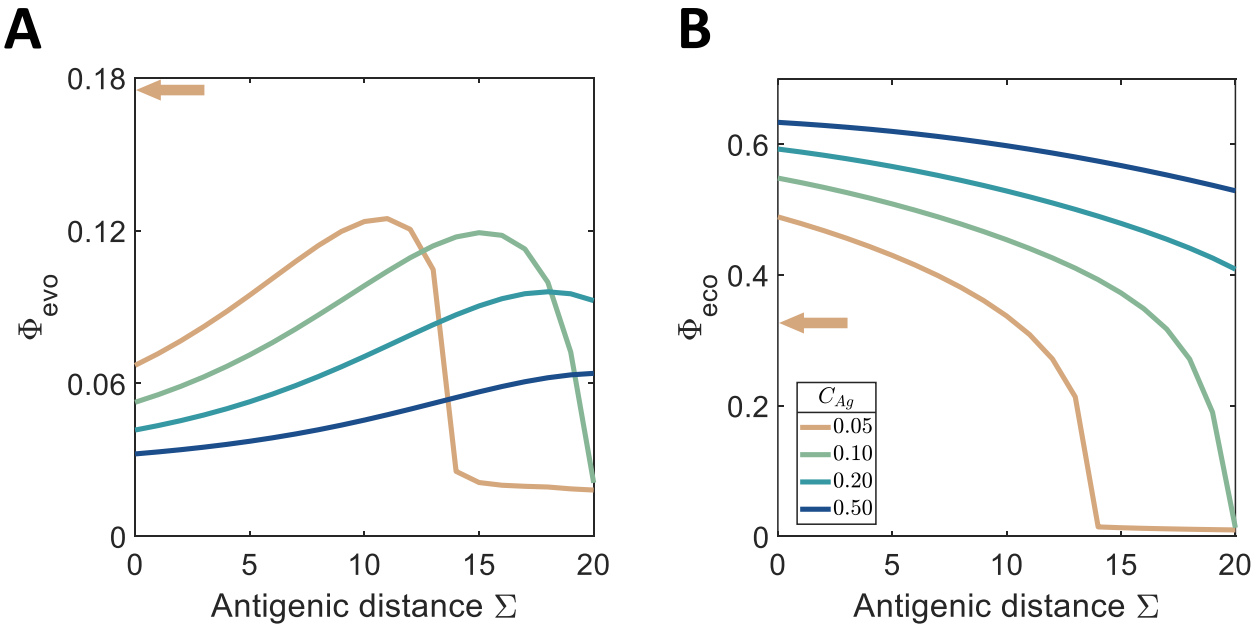

Figure 3.19: **Negatively correlated generalist quality and quantity when GC is only founded by generalists, related to Fig. 3.8**. Evolutionary and ecological objective indices Φ_{evo} and Φ_{eco} (Eq. 3.5) as a function of constant antigenic distance Σ and Ag concentration C_{Ag} (solid). Parameters are the same as in Fig. 3.8A. While generalists' quality (Φ_{evo}) is promoted by an intermediate level of Σ, their quantity (Φ_{eco}) is always higher at lower Σ. Arrows on vertical axis show the results from the dynamic switching $\Sigma(t)$ function that maximizes Φ_{evo} when $C_{Ag} = 0.05$ (which correspond to the gray diamond as in Fig. 3.6A).

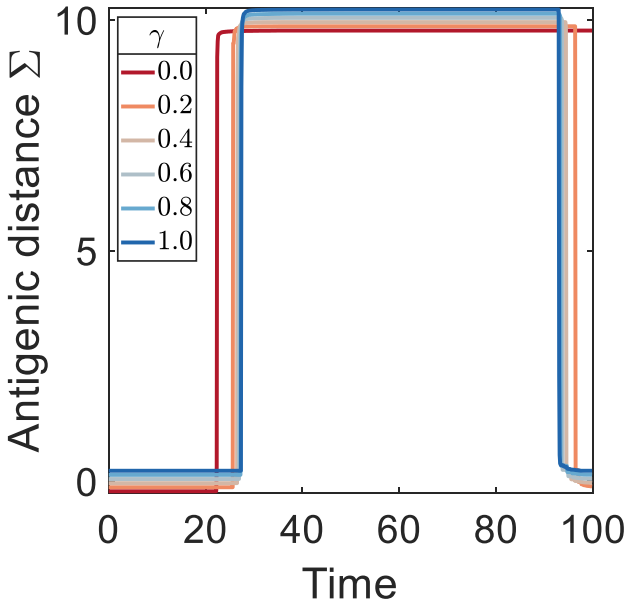

Figure 3.20: **Optimized antigenic distance functions for different generalist ecological/evolutionary focus in generalist-specialist competition, related to Fig. 3.8.** Optimized antigenic distance functions $\Sigma^*(t)$ from Pontryagin's maximum principle maximizing $\Phi = \gamma\Phi_{\text{eco}} + (1-\gamma)\Phi_{\text{evo}}$ (Eq. 3.5) for different weight γ: more focus is placed on generalist quantity (Φ_{eco}) as γ increases from 0 (red, quality only) to 1 (blue, quantity only). A small vertical shift is added to all curves to avoid overlapping. These $\Sigma^*(t)$ functions are used to generate the data in blue in Fig. 3.8C.

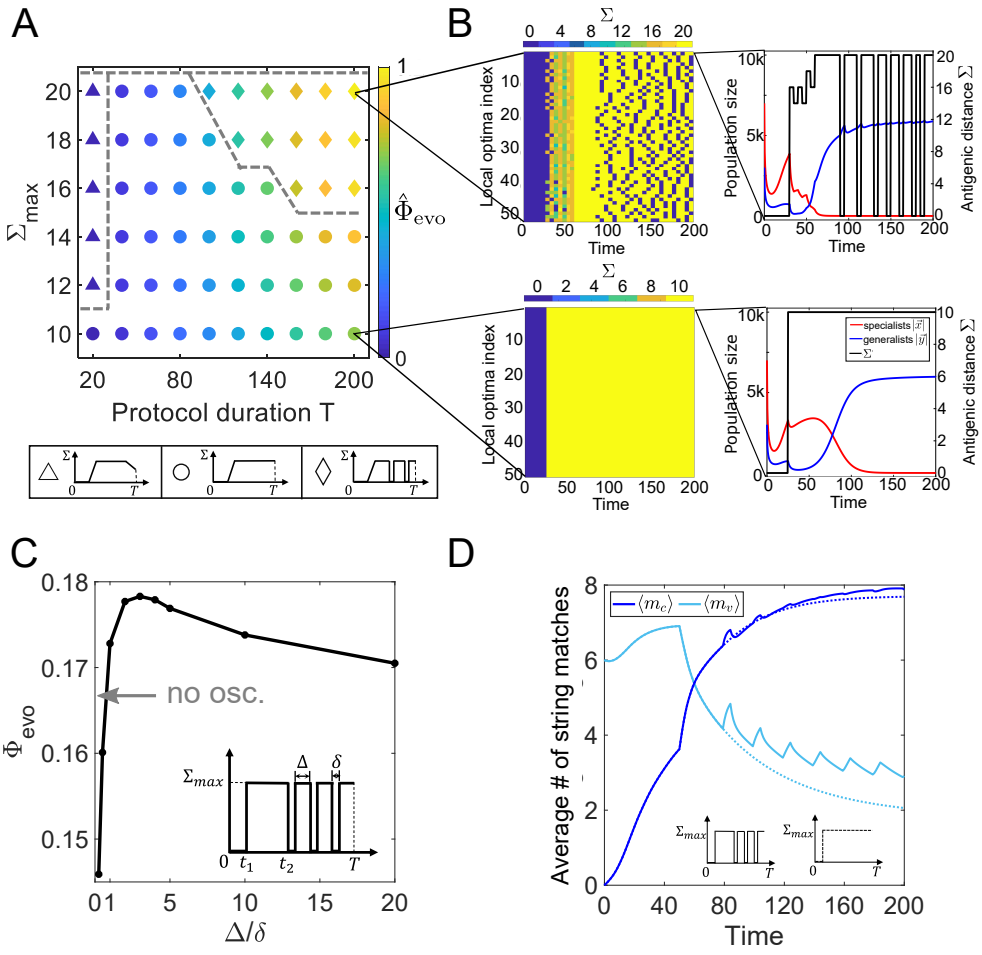

Figure 3.21: **Oscillations between zero and large antigenic distance $\Sigma_{\max}$ frustrate evolution of variable binding m_v and accelerate evolution of conserved binding m_c.** (A) Phase diagram of control shapes optimized by steepest ascent method (see section 3.D) to maximize evolutionary objective Φ_{evo} (Eq. 3.5). Different control shapes are labeled with markers: at extremely short $T = 20$ and with large $\Sigma_{\max} \geq 12$ (triangle), control first rises from zero to positive to select for high cross-reactivity m_c, and drops towards the end to slow down generalist extinction; at long T and with large $\Sigma_{\max} \geq 16$ (diamonds), control first rises from zero to $\Sigma_{\max}$, followed by oscillations between $\Sigma(t) = \Sigma_{\max}$ and 0 ("chattering"); at intermediate T the control rises from zero to $\Sigma_{\max}$. Marker color indicates the value of objective index averaged over 50 local optima solutions, and is normalized by the maximum value when $\Sigma_{\max} = 20$ and $T = 200$. GC reaction starts with 10^4 B cells including $|\vec{x}| = 7000$ specialists and $|\vec{y}| = 3000$ generalists. For generalist founders $m_v = 6$ and $m_c = 0$; for specialist founders $m = 6$. $C_{\text{Ag}} = 0.1$. (B) Example local optima of $\Sigma(t)$ (left) and population trajectories (right) when $T = 200$ and $\Sigma_{\max} = 20$ (top) or 10 (bottom). The 50 local optima are ranked in descending order of Φ_{evo}, with the highest and lowest differ by less than 0.5%. The population trajectories correspond to the best local optima of $\Sigma(t)$. (C) Oscillations improve evolutionary objective Φ_{evo} when the high-Σ period is longer than zero-Σ period ($\Delta/\delta \geq 1$). Antigenic distance functions $\Sigma(t)$ are shown as inset. The gray arrow on vertical axis marks the Φ_{evo} without oscillations in $\Sigma(t)$, i.e.,$\delta \to 0$. $\Sigma_{max} = 20$, $T = 200$, $\delta = 5$. $t_1 = 50$ and $t_2 = 80$ are from the local optima in panel (B). (D) Oscillations between zero and large antigenic distance $\Sigma_{\max}$ frustrate evolution of generalists' variable binding m_v and thus accelerate evolution of conserved binding m_c. Generalists evolve towards lower (higher) m_v for higher fitness when $\Sigma(t) > 15$ (< 15) (Fig.3.12B). The oscillations between $\Sigma(t) = 20$ and 0 thus lead to opposite selection forces for m_v and frustrate its evolution towards lower values. As a result generalists with high m_c are more preferred in selection. Solid/dashed: $\Sigma(t)$ functions with/without oscillations (inset). Parameters are the same as panel (C) with $\Delta/\delta = 3$.

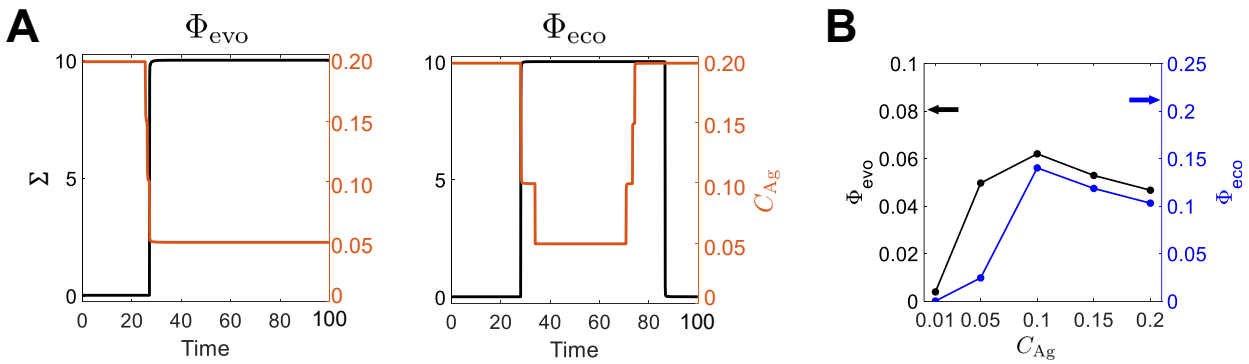

Figure 3.22: **Reinforced selection pressure by simultaneously controlling antigenic distance Σ and Ag concentration C_{Ag}.** (A): Optimal control profiles when both Σ and C_{Ag} are treated as control variables. The selection pressure shows the similar trend as when C_{Ag} is fixed: left panel: when maximizing generalist quality (Φ_{evo}), selection pressure starts as minimum and later rises to maximum (by both maximizing $\Sigma(t)$ and minimizing $C_{\mathrm{Ag}}(t)$); right panel: when maximizing generalist quantity (Φ_{eco}), selection pressure starts as minimal, then rises to maximum, and finally drops to minimal again. (B): The generalist performance with optimized $C_{\mathrm{Ag}}(t)$ (arrows next to respective axis) exceeds constant C_{Ag} (solid line). GC reaction starts with 10^4 B cells including $|\vec{x}| = 7000$ specialists and $|\vec{y}| = 3000$ generalists. For generalist founders $m_v = 6$ and $m_c = 0$; for specialist founders $m = 6$. $0.05 \leq C_{\mathrm{Ag}}(t) \leq 0.2$ during optimization.

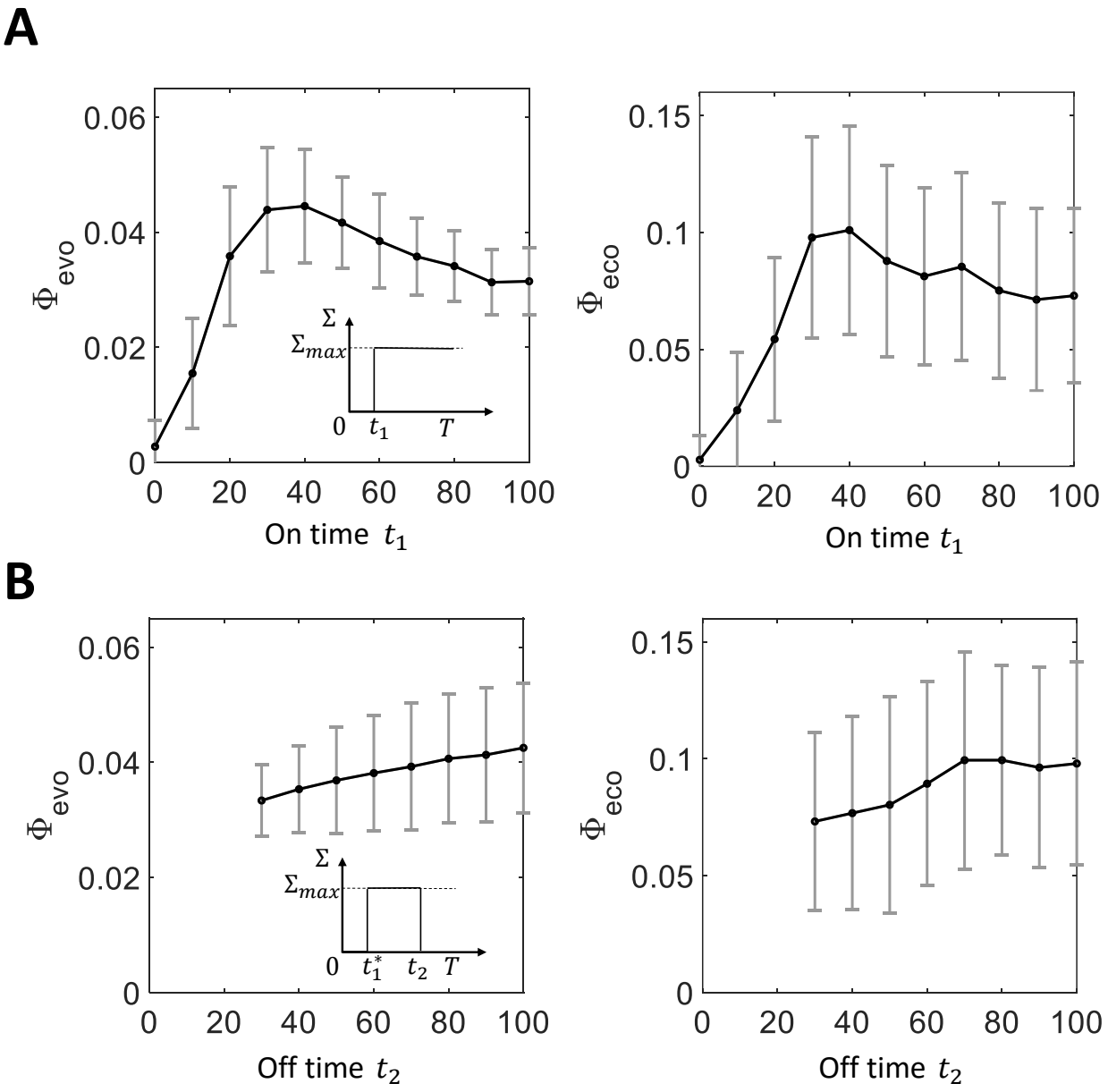

Figure 3.23: **Agent-based simulations confirm the optimal intermediate control switching time.** The objective indices Φ_{evo} (left) and Φ_{eco} (right) from stochastic agent-based simulations (see Appendices for details) as a function of control switching times. (A) Control is switched on at different time t_1 (see inset in left panel). Both Φ_{evo} and Φ_{eco} peak at an intermediate timing of 30. (B) Control is switched off at different time t_2 (see inset in left panel). Φ_{evo} improves with later off time t_2 but Φ_{eco} peaks at an intermediate timing of 70. The on time is fixed as $t_1^* = 28$ from deterministic model results (Fig. 3.4C, Fig.3.16). GC reaction starts with 10^4 B cells including $|\vec{x}| = 7000$ specialists ($m = 6$) and $|\vec{y}| = 3000$ generalists ($m_v = 6$, $m_c = 0$). Number of Ag strains $N_a = 5$, Ag concentration $C_{\text{Ag}} = 0.1$, vaccine protocol duration $T = 100$. $\Sigma_{\text{max}} = 10$. 500 simulations are performed at each data point. Errorbars show 1 std.

CHAPTER 4

Aging-induced fragility of the immune system

4.1 Abstract

The adaptive and innate branches of the vertebrate immune system work in close collaboration to protect organisms from harmful pathogens. As an organism ages its immune system undergoes immunosenescence, characterized by declined performance or malfunction in either immune branch, which can lead to disease and death. In this study we develop a mathematical model of the immune system that incorporates both the innate and adaptive immune compartments, named the integrated immune branch (IIB) model, and investigate how immune behavior changes in response to a sequence of pathogen encounters. We find that repeated pathogen exposures induce a fragility, in which exposure to novel pathogens may cause the immune response to transition to a chronic inflammatory state. The chronic inflammatory state of the IIB model is qualitatively consistent with "inflammaging," a clinically-observed condition in which aged individuals experience chronic low-grade inflammation even in the absence of pathogens. Thus, the IIB model quantitatively demonstrates how immunosenescence can manifest itself in the innate compartment as inflammaging. In particular, the onset of a persistent inflammatory response strongly depends on the history of encountered pathogens; the timing of its onset differs drastically when the same set of infections occurs in a different order. Lastly, the coupling between the two immune branches generates a trade-off between rapid pathogen clearance and a delayed onset of immunosenescence. By considering complex feedbacks between immune compartments, our work suggests potential mechanisms for immunosenescence and provides a theoretical framework to account for clinical observations.

4.2 Introduction

Infectious diseases as diverse as bacterial pneumonia, influenza, tuberculosis, herpes zoster, and most recently COVID-19 have an increased morbidity and mortality among the elderly [244, 245, 246, 247, 248, 249]. Especially in conjunction with global demographics that broadly reflect increases in age across the world's populations (both due to prolonged life expectancy and declining birth rates), the prevalence of disease among the elderly underscores the need for a better understanding of how physiology changes with age [250]. In particular, it is acutely important to identify the causes of immunosenescence, the readily observed yet mechanistically vague deterioration of immune function in aged individuals.

The vertebrate immune system targets and clears pathogens through the collaborative efforts of innate and adaptive immune responses: the innate immune system reacts quickly and non-specifically to pathogenic threats, while the adaptive immune system acts more slowly and generates a pathogen-specific response through clonal expansion of cognate T and B lymphocytes. To orchestrate this division of responsibility, extensive bidirectional interactions exist between the innate and adaptive immune compartments [251, 252, 253, 254, 255, 256]. For example, dendritic cells in the innate compartment mediate the presentation of Ag to the adaptive compartment [257]. Conversely, T cells in the adaptive arm reduce the production of inflammatory cytokines and thus limit tissue damage caused by the innate immune response [68, 258, 259]. For example, experiments with nude mice (a mutant mouse strain with low T cell levels) showed that death can ensue without this adaptive suppression of inflammation [258, 259].

Immunosenescence manifests itself in both the innate and the adaptive immune branches. In the adaptive branch aging is partially driven by thymus involution, which reduces the output of new naive T cells [260, 261, 262]. Furthermore, a lifetime of persistent pathogen exposures (e.g. chronic infections like cytomegalovirus) leads to oligoclonal expansion of memory T cells specific to those pathogens. These physiological mechanisms lead to an "imbalanced" repertoire of immune cells that is predominately populated by memory cells specific to frequently encountered pathogens, which limit the ability of the adaptive branch

to respond to novel pathogens. Indeed, the increased fraction of CD8+ T cells caused by memory inflation has been associated with the immune risk phenotype (IRP), which has been found to predict impaired immune function and mortality [263, 264, 32].

In the innate compartment, aging is associated with the development of a chronic low-grade inflammatory response even in the absence of pathogen stimulation, called "inflammaging" [31]. The elderly often experience CI and possess elevated levels of pro-inflammatory cytokines [32, 265, 266, 267], which have been found to be strong predictors of mortality (for example, interleukin 6 has been associated with the IRP) [32]. Prior theories suggest that inflammaging is facilitated by long-lasting pathogen encounters, cell debris and stress, and the reduced efficiency of the adaptive immune response [31, 268, 269]. Still, the mechanisms underlying the onset of inflammaging— and in particular its connection to aging in the adaptive immune system— require further study.

In an earlier mathematical model of the adaptive immune response, Stromberg and Carlson found that repeated pathogen exposures could lead to an imbalanced immune repertoire that was vulnerable to rare pathogens [270]. Around the same time, Reynolds *et al.* developed a model of the innate immune response immediately following a pathogen encounter. Based on these earlier models, in this chapter we construct a mathematical model of the coupled innate-adaptive immune system called the integrated immune branch model, and demonstrate how immunosenescence can develop and trigger a chronic inflammatory response. Here, as in the earlier adaptive immune model, the onset of immunosenescence arises purely from an imbalanced immune repertoire, and occurs without any further assumptions regarding the degradation of cellular function with age. The IIB model explicitly incorporates the inflammatory response of the innate branch, T cell dynamics of the adaptive branch, and innate-adaptive crosstalk. It recapitulates several clinically-observed signatures of immunosenescence: the ratio of naive to memory cells decreases over time [271], repeated exposure to chronic infections (e.g. human cytomegalovirus) induces immune fragility [272], and this fragility is characterized by CI ("inflammaging") [271, 32].

With the IIB model, we first characterize the dynamics and steady states of the immune system in response to a single infection event. Then, the system is exposed to a series of

pathogens that form an infection history. This sequence of infection events causes overspecialization in the adaptive compartment and triggers CI. In particular, the order in which infections are encountered strongly influences health outcomes, and can hasten or delay the onset of CI.

Further, by tuning resource allocation in the adaptive compartment toward pathogen clearance versus suppression of inflammation, crosstalks between immune compartments may be directly manipulated. This manipulation reveals a trade-off between a delayed onset of CI and rapid pathogen clearance. Our model provides a mechanistic explanation of how accumulated pathogen exposures can cause immune fragility that leads to inflammaging and immunosenescence, and may serve as a foundation for quantitative studies of immune crosstalk and aging.

4.3 Mathematical model

Extensive mathematical and computational modeling efforts have been made to better understand both the innate and adaptive branches of vertebrate immune system [40, 274, 275]. Additionally, a rich literature exists regarding the inflammatory innate response [276, 277, 278, 279, 280] and the adaptive immune repertoire [281, 282, 283, 284]. In particular, Reynolds *et al.* studied the positive feedback between activated phagocytes and collateral tissue damage [273], and Stromberg and Carlson modeled the accumulated loss of memory cell diversity over the course of a lifetime of infections [270]. In this study, we modify and synthesize the models of Reynolds *et al.* and Stromberg and Carlson to develop an integrated immune branch model, in which the innate and adaptive immune branches work collaboratively to clear pathogens.

A schematic of the IIB model is depicted in Fig. 4.1, and a thorough accounting of the individual components of the innate (pathogen P_i, phagocytes N^*, tissue damage D, and anti-inflammatory cytokines C_A) and adaptive (pathogen P_i, naive cells N_i, memory cells M_i, and effector cells E_i) models is provided in Table 4.1. The coupled model is described in full in Table 4.2, and parameter descriptions and values are provided in Table 4.3. For

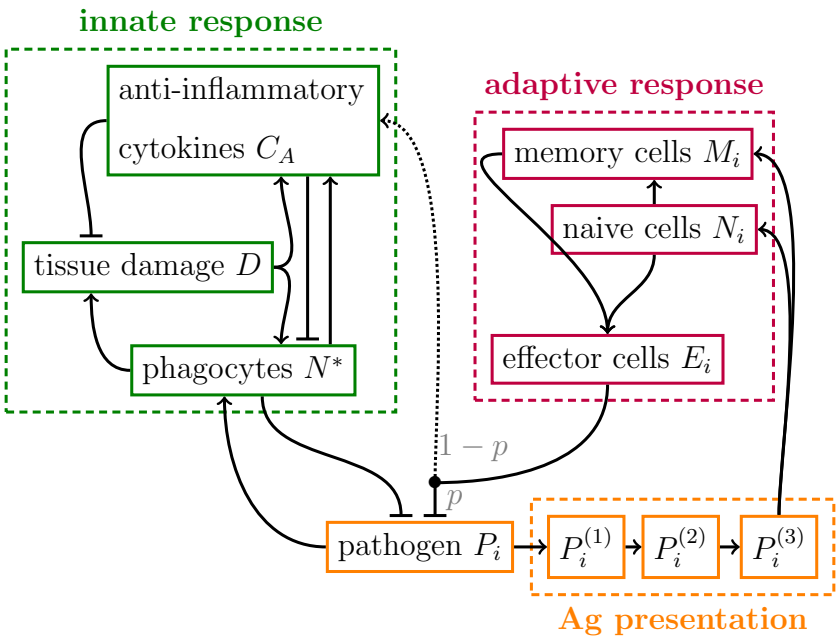

Figure 4.1: **Schematic of the IIB model.** An introduced pathogen prompts innate and adaptive immune responses that seek to eliminate the pathogen. The innate response (green) is adapted from the model of Reynolds *et al.* [273], in which the presence of a pathogen activates phagocytes that induce inflammation and the subsequent production of anti-inflammatory cells. The adaptive immune response (magenta) is adapted from the model of Stromberg and Carlson [270], in which presented pathogens (orange) activate naive and memory T cells specific to that pathogen, causing them to divide into effector cells that target the pathogen. In this model, the delay in the adaptive response due to Ag presentation is hard-coded via the compartments $P_i^{(1)}$, $P_i^{(2)}$, and $P_i^{(3)}$ using the linear chain technique. The state variables of this model are described in Table 4.1, the model itself is given explicitly in Table 4.2, and the parameters of this model are provided in Table 4.3.

more complete descriptions of the separate innate and adaptive immune models, we refer the reader to the original publications [273, 270].

4.3.1 Innate immune response

In the model formulated by Reynolds *et al.* [273] (depicted in the green box of Fig. 4.1), once a pathogen P_i is introduced, phagocytes N^* (which include, e.g., neutrophils and macrophages) are recruited to the site of the infection by activation or migration. These phagocytes kill the pathogen P_i by phagocytosis, degranulation, or by creating neutrophil extracellular traps.

Notation	Immune component	Description
P_i	Pathogen	• Harmful exogenous stimulants (e.g., bacteria or viruses) that activate an immune response • Pathogen of shape i (shape space formulation)
N^*	Activated phagocytes	• Phagocytes (which include neutrophils and macrophages) that are activated by any pathogen P_i or by pro-inflammatory cytokines • Responsible for removing pathogens P_i, but cause collateral tissue damage D
D	Tissue damage	• Caused by activated phagocytes N^* • Causes release of pro-inflammatory cytokines that recruit additional phagocytes N^*
C_A	Anti-inflammatory cytokines	• Small protein molecules that reduce the efficiency and recruitment of activated phagocytes N^* • Production encouraged by activated phagocytes N^* and tissue damage D, and also by effector cells E (innate-adaptive crosstalk)
N_i	Naive cells	• Mature T cells with receptor specificity represented by shape i that are agnostic to previous pathogen encounters • Divide and differentiate into memory M_i and effector E_i cells when activated by pathogens P_i • Subject to homeostasis control mechanisms
M_i	Memory cells	• Long-lived cells differentiated from naive cells N_i with the same pathogen specificity • Divide and differentiate into memory M_i and effector E_i cells when activated by pathogens P_i • Subject to homeostasis control mechanisms
E_i	Effector cells	• Short-lived cells differentiated from naive N_i and memory M_i cells • Remove pathogen P_i and produces anti-inflammatory cytokines C_A

Table 4.1: **Major biological components in the IIB model.** Model equations are provided in full in Table 4.2.

At the same time, these phagocytes release inflammatory cytokines that induce inflammation and tissue damage D in the host. Damaged tissue in turn releases additional inflammatory cytokines, further promoting phagocyte activation. Following this initial inflammatory response, a wave of anti-inflammatory cytokines C_A is released to downregulate phagocyte recruitment and to reduce inflammation and tissue damage.

The steady-state behavior of this model was investigated by Reynolds *et al.* [273], who found that this formulation of the innate immune response led to three possible steady states:

health, in which the pathogen is cleared and the phagocytes and tissue damage vanish; *septic death*, in which the innate response is unable to clear the pathogen; and *chronic inflammation*, in which the innate response clears the pathogen at the expense of inducing a constant inflammatory response that fails to dissipate, even after pathogen clearance. In the model's original formulation this CI steady state was named *aseptic death*, but in this study we allow for the possible interpretation of this steady state as "inflammaging," which is not necessarily fatal, rather than death [285, 286]. As we will show, the IIB model retains these three steady states while introducing interplay between the non-specific innate response and the specific adaptive response.

4.3.2 Adaptive immune response

The IIB model also incorporates the adaptive immune response (shown in the magenta box of Fig. 4.1), which is based on the shape-space adaptive immune model of Stromberg and Carlson [270]. By assigning discrete "shapes" to pathogen epitopes and adaptive immune cells, this formulation allows pathogen-specific immune memory and immune responses to be developed [287]. Thus, an introduced pathogen P_i of shape i induces an adaptive response consisting of naive cells N_i, memory cells M_i, and effector cells E_i that all specifically target the introduced pathogen. There are S_{max} available shapes. In the original model of Stromberg and Carlson, pathogens P_i of shape i could interact with adaptive cells of some different shape j with a lowered binding affinity, but in this study for computational efficiency we require that adaptive responses be activated by a pathogen of identical shape.

The process of Ag presentation delays the activation of the adaptive immune response. This phenomenon is encoded in the IIB model with the linear chain technique (shown in the orange box of Fig. 4.1), in which the populations $P_i^{(1)}$, $P_i^{(2)}$, and $P_i^{(3)}$ are intermediate states during Ag presentation [288]. Eventually, the presented Ag $P_i^{(3)}$ induces naive cells N_i to divide into memory cells M_i and effector cells E_i; and memory cells M_i to divide into additional memory cells M_i and effector cells E_i. Once created, these pathogen-specific effector cells E_i work to clear the pathogen P_i.

Equation	Interpretation
$\frac{\mathrm{d}P_i}{\mathrm{d}t} = k_{pg}P_i\left(1-\frac{\lvert\vec{P}\rvert}{P_\infty}\right) - \frac{k_{pm}s_m}{\mu_m + k_{mp}\lvert\vec{P}\rvert}P_i - k_{pn}f(N^*)P_i - p\gamma P_i E_i - \Delta * P_i$	Pathogen P_i of shape i changes according to: • logistic growth, carrying capacity P_∞ ($\lvert\cdot\rvert$ denotes the 1-norm) • inhibition by a non-local immune response • clearance by innate phagocytes N^*, effect is mediated by anti-inflammatory cytokines C_A via $f(\cdot)$ • clearance by adaptive effector cells E_i • sequestration by dendritic cells for Ag presentation
$\frac{\mathrm{d}P_i^{(1)}}{\mathrm{d}t} = \Delta\left[P_i - P_i^{(1)}\right]$ $\frac{\mathrm{d}P_i^{(2)}}{\mathrm{d}t} = \Delta\left[P_i^{(1)} - P_i^{(2)}\right]$ $\frac{\mathrm{d}P_i^{(3)}}{\mathrm{d}t} = \Delta\left[P_i^{(2)} - P_i^{(3)}\right] - \beta\gamma P_i^{(3)}(N_i + M_i)$	• Ag presentation occurs with hard-coded delay (linear chain technique) of $3/\Delta$ units of time on average required for a pathogen P_i to transition to compartment $P_i^{(3)}$ • compartments $P_i^{(1,\,2,\,3)}$ correspond to intermediate states during Ag presentation • once Ag arrives in compartment $P_i^{(3)}$ it activates naive cells N_i and memory cells M_i
$\frac{\mathrm{d}N_i}{\mathrm{d}t} = -\alpha\gamma N_i P_i^{(3)} + \theta_N - \delta_N N_i \frac{\lvert\vec{M}\rvert + \lvert\vec{N}\rvert}{R_0}$	Naive cells N_i of shape i change according to: • division into effector cells E_i (with rate f) and memory cells M_i (with rate $1-f$) • constant production at rate θ_N • return to homeostatic equilibrium (timescale $1/\delta_N$)
$\frac{\mathrm{d}M_i}{\mathrm{d}t} = (2-2f)\alpha\gamma N_i P_i^{(3)} + (1-2f)\alpha\gamma M_i P_i^{(3)} - \delta_M M_i \frac{\lvert\vec{M}\rvert + \lvert\vec{N}\rvert}{R_0}$	Memory cells M_i of shape i change according to: • division into effector cells E_i (with rate f) and memory cells M_i (with rate $1-f$) • growth from naive and memory cell division • decay at rate δ_M
$\frac{\mathrm{d}E_i}{\mathrm{d}t} = 2f\alpha(M_i + N_i)\gamma P_i^{(3)} - \delta_E E_i$	Effector cells E_i of shape i change according to: • production by naive and memory cells proportional to Ag presentation rate α • decay at rate δ_E
$\frac{\mathrm{d}N^*}{\mathrm{d}t} = \frac{s_{nr}R}{\mu_{nr} + R} - \mu_n N^*$	Innate phagocytes N^* change according to: • activation by the presence of other phagocytes, pathogen, or tissue damage (encapsulated by R) • decay at rate μ_n
$\frac{\mathrm{d}D}{\mathrm{d}t} = k_{dn}f_s(f(N^*,\ C_A)) - \mu_d D$	Tissue damage D changes according to: • induced by activated phagocytes N^*, but ameliorated by the presence of anti-inflammatory cytokines C_A via $f(\cdot)$ • decay at rate μ_d
$\frac{\mathrm{d}C_A}{\mathrm{d}t} = s_c + k_{cn}\frac{f(N^* + k_{cnd}D,\ C_A)}{1 + f(N^* + k_{cnd}D,\ C_A)} - \mu_c C_A + (1-p)\,k_{ce}\frac{\lvert\vec{E}\rvert}{\lvert\vec{E}\rvert + E_{1/2}}$	Anti-inflammatory cytokines C_A change according to: • production at constant rate s_c • production related to phagocyte and tissue damage levels • decay at rate μ_c • stimulation by effector cells $\lvert\vec{E}\rvert$
$R = f(k_{nn}N^* + k_{np}P + k_{nd}D)$ $f(x,\ C_A) = \frac{x}{1 + \left(\frac{C_A}{C_\infty}\right)^2}$ $f_s(y) = \frac{y^6}{x_{dn}^6 + y^6}$	• R is an aggregation of signals that trigger the innate immune response • $f(x,\ C_A)$ mediates the value of x according to the level of anti-inflammation cytokines C_A • $f_s(y)$ was phenomenologically fit by Reynolds *et al.* in their original formulation [273]

Table 4.2: **Full equations of the IIB model that govern the dynamics of the immunological state variables described in Table 4.1**. A full list of parameter values and descriptions is given in Table 4.3

.

4.3.3 Integrated immune branch (IIB) model

The IIB model is described in full in Table 4.2, with descriptions and values of the parameters given in Table 4.3. Next, we emphasize the modifications that synthesized the two separate

Parameter	Value	Description and dimension	Source	Parameter	Value	Description and dimension	Source
k_{pg}	**0.6**	pathogen logistic growth rate; $[T^{-1}]$	[273]	P_∞	**20**	pathogen logistic carrying capacity; $[P]$	[273]
k_{pm}	**0.6**	pathogen clearance rate by nonspecific response; $[T^{-1}]$	[273]	s_m	**0.005**	source rate of nonspecific response; $[T^{-1}]$	[273]
μ_m	**0.002**	decay rate of nonspecific response; $[T^{-1}]$	[273]	k_{mp}	**0.01**	rate of nonspecific exhaustion per pathogen; $[P^{-1}T^{-1}]$	[273]
k_{pn}	**1.8**	rate of pathogen clearance by innate response; $[(N^*)^{-1}T^{-1}]$	[273]	γ	0.02	binding rate between pathogens and adaptive cells of the same type; $[C^{-1}T^{-1}]$	[270]
p	0.9	proportion of effector cell resources allocated to pathogen clearance; [nondim.]		Δ	0.1	rate of Ag presentation; $[T^{-1}]$	
β	0.01	efficacy of $P_j^{(3)}$ depletion by Ag presentation; [nondim.]		α	0.1	efficacy of adaptive cell activation by Ag presentation; $[CP^{-1}]$	[270]
θ_N	5	naive cell creation rate; $[CT^{-1}]$		R_0	7200	total naive and memory cell logistic carrying capacity; $[C]$	
δ_N	0.025	naive cell homeostasis rate; $[T^{-1}]$		δ_M	4e-5	memory cell decay rate; $[T^{-1}]$	
f	0.4	proportion of memory and naive cells that divide into effector cells; [nondim.]	[270]	δ_E	0.05	effector cell decay rate; $[T^{-1}]$	[270]
s_{nr}	**0.08**	maximum phagocyte recruitment rate; $[N^*T^{-1}]$	[273]	μ_{nr}	**0.12**	phagocyte recruitment half-saturation constant; $[T^{-1}]$	[273]
μ_n	**0.05**	phagocyte decay rate; $[T^{-1}]$	[273]	k_{dn}	**0.35**	rate of tissue damage due to phagocytes; $[DT^{-1}]$	[273]
μ_d	**0.02**	tissue damage decay rate; $[T^{-1}]$	[273]	s_c	**0.0125**	source rate of anti-inflammatory cytokines; $[C_AT^{-1}]$	[273]
k_{cn}	**0.04**	maximum activation of anti-inflammatory cytokines by phagocytes and tissue damage $[C_AT^{-1}]$	[273]	k_{cnd}	**48**	conversion rate between tissue damage and phagocyte abundance; $[N^*D^{-1}]$	[273]
μ_c	**0.1**	anti-inflammatory cytokine decay rate; $[T^{-1}]$	[273]	k_{nn}	**0.01**	conversion rate between phagocyte abundance and aggregate innate response R; $[(N^*)^{-1}T^{-1}]$	[273]
k_{np}	**0.1**	conversion rate between pathogen abundance and aggregate innate response R; $[P^{-1}T^{-1}]$	[273]	k_{nd}	**0.02**	conversion rate between tissue damage and aggregate innate response R; $[D^{-1}T^{-1}]$	[273]
k_{ce}	0.4	maximum anti-inflammatory cytokine production rate by effector cells; $[C_AT^{-1}]$		$E_{1/2}$	10	half-saturation constant for cytokine production by effector cells; $[C]$	
C_∞	**0.28**	scaling factor for anti-inflammatory cytokine abundance; $[C_A]$	[273]	x_{dn}	**0.06**	phenomenologically-inferred half-saturation constant; $[N^*]$	[273]
S_{max}	**36**	number of pathogen shapes in shape space; [nondim.]	[270]				

Table 4.3: **Typical parameters of the immune model in Table 4.2.** The parameters listed in this table are used to generate Fig. 4.5, while the other figures are created with slightly modified parameters as detailed in the Appendices. Most innate parameters were originally described in the Reynolds *et al.* model [273], while most adaptive parameters were originally described in the Stromberg and Carlson model [270]. Parameter values that are the same as those used in the original models are bold-faced. Dimensions are given in square brackets, with $[T]$ denoting time, other symbols denoting the concentrations of their corresponding immune variables, and $[C]$ denoting concentrations of adaptive immune cells (i.e. naive cells N_i, memory cells M_i, or effector cells E_i).

innate and adaptive immune models.

In the IIB model, the innate and adaptive components are linked in two ways. First, the

two compartments are implicitly linked through the pathogen population: a higher pathogen load P_i not only activates more phagocytes N^* in the innate compartment, but also leads to a higher rate of Ag presentation and subsequent activation of naive N_i, memory M_i, and effector E_i cells. Thus, depletion of pathogen by either response (by phagocytes N^* or by effector cells E_i) affects both compartments.

Second, the two compartments are explicitly linked since effector cells E_i can create anti-inflammatory cytokines C_A that weaken the innate response. The suppression of inflammatory responses by effector cells has been observed experimentally [68, 258, 259]. In the IIB model, effector cells are allocated either to clear pathogens or to promote the production of anti-inflammatory cytokines, in proportion to p and $1 - p$, respectively (dashed line in Fig. 4.1). The pathogen removal efficiency p may be varied from 0 to 1, so that in the limit that effector cells are solely responsible for pathogen clearance ($p = 1$) no additional anti-inflammatory cytokines are produced. Therefore, the synthesized model is capable of quantitatively comparing the two adaptive immune functions of pathogen clearance and inflammation attenuation.

The IIB model introduces homeostatic constraints that regulate the capacity of naive and memory cells. The rates of both homeostatic responses are dependent on the sum of the bulk naive cell population $|\vec{N}| \equiv \sum_i N_i$ and the bulk memory cell population $|\vec{M}| \equiv \sum_i M_i$. In the process of clearing a pathogen, memory cells accumulate. Afterwards, to satisfy the homeostatic constraints, naive cells must become less abundant than they were before the infection. In the Appendices, we derive analytic approximations to the immune dynamics that are generated by these homeostatic relaxations; when the timescale of pathogen clearance is much shorter than the timescale of homeostatic relaxation, these expressions can be used as part of a dynamic programming approach to significantly speed up numerical simulations.

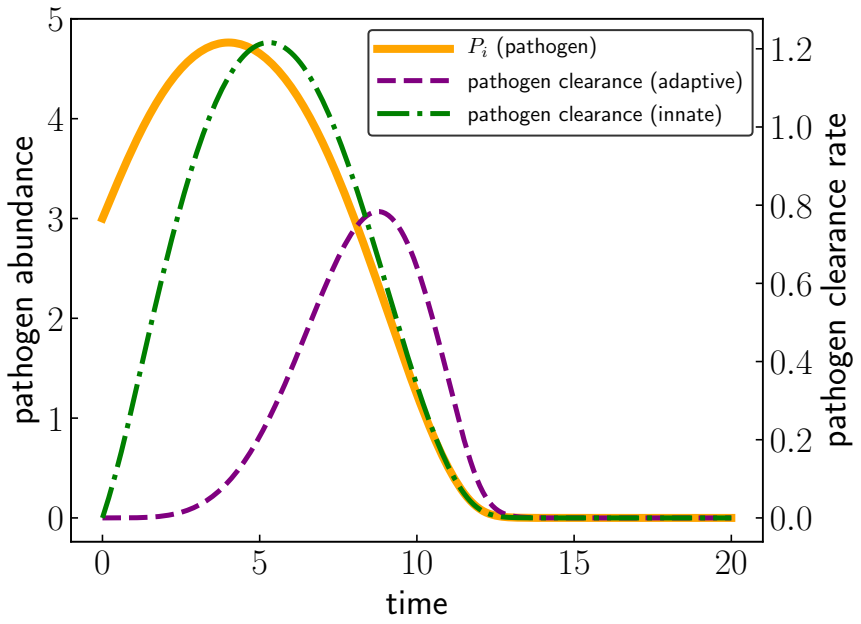

Figure 4.2: **Pathogen abundance (orange, solid) is regulated by the innate and adaptive immune responses in the IIB model.** The clearance rates of the innate response (green dash-dotted, rate given by $k_{pn} f(N^*) P_i$) and the adaptive response (purple dashed, rate given by $p\gamma P_i E_i$), as described in the $\frac{\mathrm{d}P_i}{\mathrm{d}t}$ equation of Table 4.2, are plotted. The innate response is activated immediately, while the adaptive response is delayed due to the Ag presentation process (encoded with the linear chain technique $P_i \to P_i^{(1)} \to P_i^{(2)} \to P_i^{(3)}$). Ultimately the combined immune responses manage to clear the pathogen. The parameters used to generate this figure are given in Table 4.3 and in Table 4.4 of the Appendices.

4.4 Results

4.4.1 Clearance of a single infection by the coupled immune response

In the IIB model, once a pathogen is introduced it initiates a cascade of immune responses in the innate and adaptive compartments, and these responses combat the logistic growth of the pathogen and attempt to drive it to extinction. Fig. 4.2 depicts a representative pathogen encounter and clearance, and plots the pathogen abundance (orange) as well as the pathogen clearance rates due to the adaptive effector cell (purple) and innate phagocyte (green) responses given by the quantities $p\gamma P_i E_i$ and $k_{pn} f(N^*) P_i$, respectively, as given in Table 4.2. Note that the adaptive response is specific to the pathogen shape, while the innate response is nonspecific. For clarity only the lumped contributions of the innate and adaptive compartments to pathogen clearance are plotted in Fig. 4.2. The populations of

every immunological variable are plotted for several scenarios in Fig. 4.3.

In Fig. 4.2 the logistic growth of the pathogen drives the initial pathogen spike, which is mildly tempered by a non-local immune response as described by Reynolds *et al.* [273]. Phagocytes are immediately recruited and attack the pathogen, leading to the increase in innate pathogen clearance (green). Simultaneously, the collateral tissue damage inflicted by phagocytes causes the production of anti-inflammatory cytokines, which suppress further phagocyte recruitment and tissue damage. Anti-inflammatory cytokines in conjunction with a decreasing pathogen population cause the decrease in innate pathogen clearance. Once the pathogen is presented to the adaptive immune branch— a delay that is hard-coded in the IIB model with the auxiliary immunological variables $P_i^{(1)}$, $P_i^{(2)}$, and $P_i^{(3)}$— the naive and memory cells specific to the presented Ag divide into effector cells. These effector cells subsequently contribute to the increase in adaptive pathogen clearance (purple). Ultimately, with the given parameter values (provided in the Appendices) the innate and adaptive responses overpower the pathogen and drive it to extinction. In the process, memory cells specific to this pathogen shape proliferate and provide future protection in case the same pathogen is faced again in the future, since the higher initial abundance of pathogen-specific memory cells will result in a more immediate adaptive response.

4.4.2 Steady state analysis of the IIB model

The IIB model described in Table 4.2 exhibits steady states when the time derivatives of all the populations vanish. This coupled model inherits many of the steady state characteristics of the constituent innate and adaptive immune models. In particular, this model exhibits steady states of (a) *health*, characterized by vanishing pathogen and immune response; (b) *chronic inflammation*, in which the pathogen clears but the innate immune response is sustained in a positive feedback loop; and (c) *septic death*, characterized by the chronic presence of pathogen and activated immune responses. Realizations of these three steady states are displayed in Fig. 4.3, and the parameter sets used to generate each outcome are provided in Table 4.3 and in Table 4.4 of the Appendices. The steady states attained

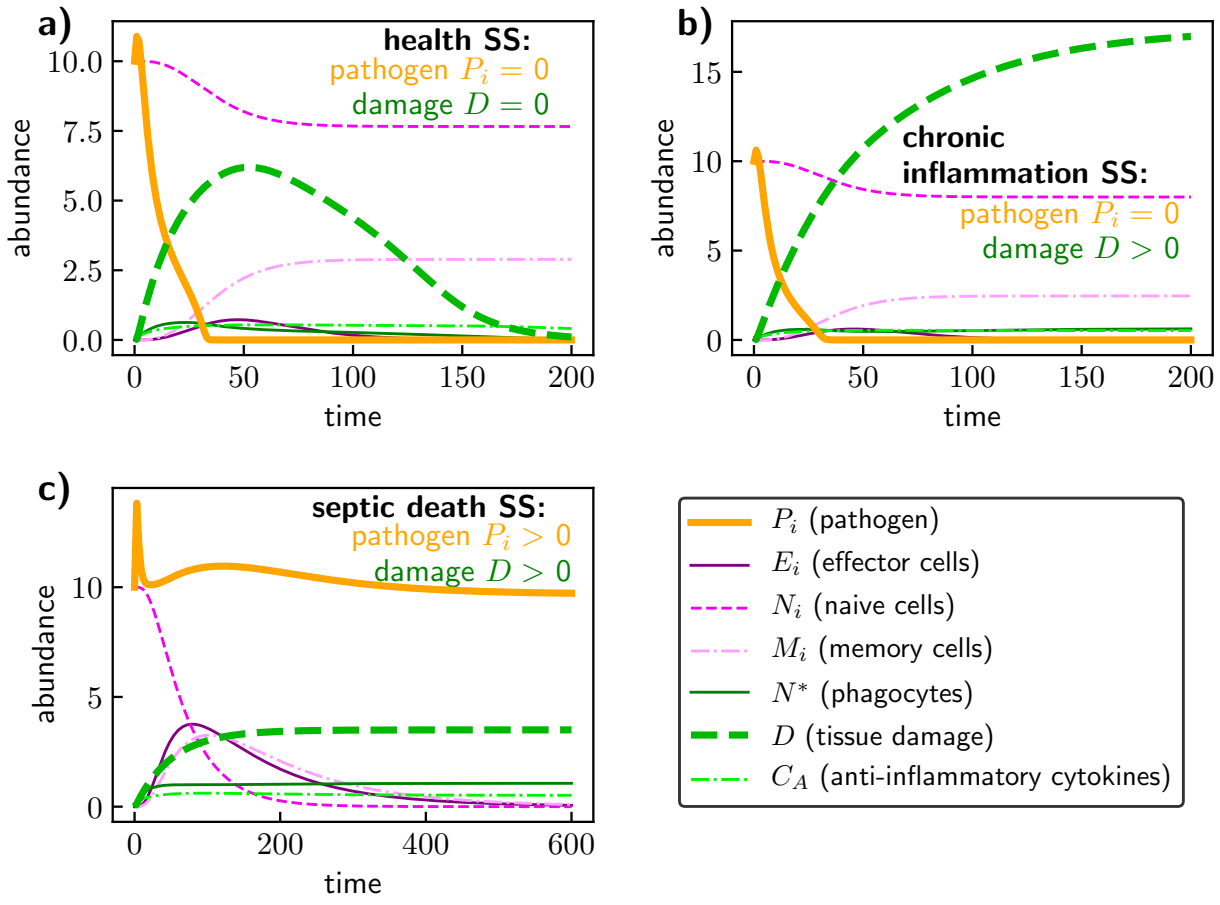

Figure 4.3: **The IIB immune model exhibits (a) health, (b) chronic inflammation, and (c) septic death steady states.** Components of the adaptive response are plotted in purple, while components of the innate response are plotted in green. (a) An inoculated pathogen activates phagocytes, which in turn induce tissue damage. Following Ag presentation, naive cells divide into memory cells and effector cells. The phagocytes and effector cells jointly suppress the pathogen, which goes extinct, and the tissue damage gradually decays resulting in the health steady state. (b) The innate and adaptive immune responses clear the pathogen, but in the process the innate response enters a positive feedback loop between phagocyte recruitment and tissue damage leading to persistent tissue damage and phagocyte activation, called the chronic inflammation steady state. (c) The innate and adaptive immune responses do not clear the pathogen, leading to the septic death steady state characterized by the presence of pathogen and tissue damage. The CI steady state was obtained with smaller innate clearance rate k_{pm} and smaller tissue damage decay rate μ_d than were used to obtain the health steady state. The septic death steady state was obtained with a larger proportion of cognate cells that divide into effector cells f and a larger pathogen growth rate k_{pg} than were used to obtain the health steady state. Explicit values of the parameters used for each panel are given in Table 4.3 and in Table 4.4 of the Appendices.

by the IIB model are sensitive to parameter values, and in Fig. 4.3 different steady states were attained by varying the innate cell clearance rate k_{pm}, the tissue damage decay rate μ_d, the proportion f of cells that divide into effector cells, and the pathogen growth rate k_{pg}. In Fig. 4.4, phase diagrams demonstrate how different parameter regimes and initial conditions lead to different steady states. To observe the presence of these steady states

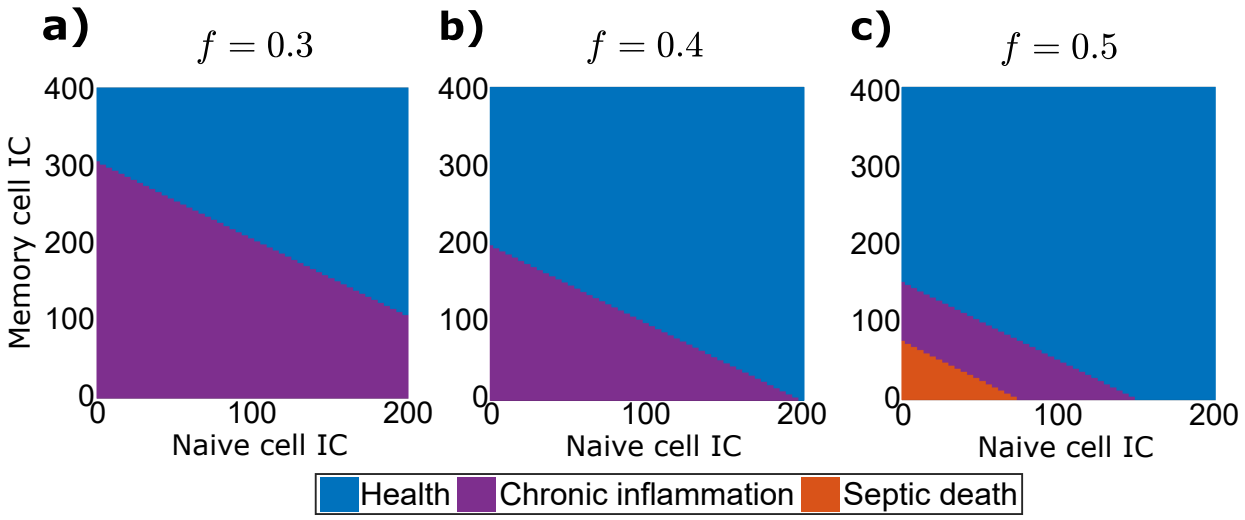

Figure 4.4: **Phase diagram of immunological steady states as a function of naive cell and memory cell initial conditions.** In IIB model, the immune system can reach health, CI, or septic death steady state following a pathogen encounter, depending on the level of cognate naive and memory cells when the pathogen is introduced. The proportion f of cognate cells that divide into effector cells significantly influences the steady-state phase diagram; in particular, septic death (red zone) can only occur for $f \geq 0.5$. Results are calculated in the absence of homeostatic response ($\theta_N = \delta_N = \delta_M = 0$ in Table 4.2) during single infection events and initial pathogen level is $P_i = 1$. Other parameters for generating the phase diagrams are as stated in Table 4.2. Except in current figure, $f = 0.4$ is always used in this study.

mathematically, note that a quantity P_i may be factored out of the pathogen dynamics $\frac{\mathrm{d}P_i}{\mathrm{d}t}$, from which it is clear that at steady state the pathogen population $\bar{P}_i$ must either be 0 or some nonzero quantity that satisfies $\frac{1}{P_i}\frac{\mathrm{d}P_i}{\mathrm{d}t} = 0$. The three steady states immediately follow from the implications of choosing $\bar{P}_i$ to be zero or non-zero.

(a) Health steady state— If $\bar{P}_i = 0$ at equilibrium, then the intermediate pathogen states must vanish as well ($\bar{P}_i^{(1)} = \bar{P}_i^{(2)} = \bar{P}_i^{(3)} = 0$). In the absence of presented Ag, the memory cells M_i decay with timescale $1/\delta_M$. When this timescale is slow relative to the homeostatic dynamics of the naive cells (whose dynamics are of timescale $1/\delta_N$), the naive cells N_i tend towards their homeostatic equilibrium $\bar{N}_i$ as described in Eq. 4.14 in the Appendices. Lastly, all effector cells E_i decay with timescale $1/\delta_E$, which is assumed to be fast compared to the naive and memory cell dynamics. Thus, in the absence of pathogen, the adaptive immune response turns off and becomes dormant. In a steady state with $\bar{P}_i = 0$ the innate immune response can either be "inactive" or it can be "active," which lead to the steady states of *health* or *chronic inflammation*, respectively. When the innate immune response

is inactive, the activated phagocytes N^* and tissue damage D are both zero. This implies that the aggregate innate response R is zero as well. Lastly, in the absence of the innate response and effector cells, anti-inflammatory cytokines equilibrate to a constant level. It is straightforward to check that setting $\bar{P}_i = \bar{N}^* = \bar{D} = 0$ and $\bar{C}_A = s_c/\mu_c$ leads to a steady state in this model.

This health steady state is demonstrated numerically in Fig. 4.3a. In this figure, an initial pathogen response activates innate (green) and adaptive (purple) responses that eventually vanish. The naive and memory cells change over the course of the adaptive immune response, and then remain at constant steady state values.

(b) CI steady state— When $\bar{P}_i = 0$ but the innate immune response is active, the model reaches the CI steady state. This steady state is inherited from the Reynolds *et al.* innate immune model [273], and occurs due to a positive feedback loop between tissue damage D and phagocytes N^* (as can be schematically understood based on Fig. 4.1). In particular, there exist equilibrium quantities $\bar{N}^*$ and $\bar{D}$ that precisely balance the activation of phagocytes and accumulation of tissue damage with their respective decays $\mu_n \bar{N}^*$ and $\mu_d \bar{D}$. This CI steady state is shown in Fig. 4.3b: the pathogen is cleared and effector cells dissipate, but the innate response is perpetually sustained.

(c) Septic death steady state— Lastly, the steady state in which the pathogen population is sustained is called septic death. For a steady state with nonzero pathogen $\bar{P}_i$, the values of the intermediate pathogen states $\bar{P}_i^{(1)}$, $\bar{P}_i^{(2)}$, and $\bar{P}_i^{(3)}$ will be nonzero as well. Subsequently, the presented pathogen $\bar{P}_i^{(3)}$ sustains the activation of naive, memory, and effector cells. In the innate compartment, the nonzero pathogen presence implies a nonzero aggregate innate response $\bar{R}$, which implies a nonzero equilibrium population of phagocytes $\bar{N}^*$, which in turn implies a nonzero equilibrium population of tissue damage $\bar{D}$. Therefore, the septic death steady state is characterized by activity in both the innate and the adaptive immune compartments. Naive and memory cells of the adaptive compartment continue to predominantly divide into effector cells; eventually all adaptive cells are exhausted and vanish while failing to clear the pathogen. This steady state is depicted in Fig. 4.3c.

The rest of this paper uses a parameter regime that does not exhibit the septic death steady state: in particular, simulations of this paper set f (the proportion of cognate cells that divide into effector cells) equal to 0.4, which causes memory cells to accumulate until they are able to produce enough effector cells to suppress the pathogen. When f is larger than 0.5 memory cells deplete over the course of an immune response, which can lead to septic death. A phase diagram of steady state behaviors at different values of f for different naive and memory cell initial conditions is plotted in Fig. 4.4. In particular, septic death is only reachable when $f \geq 0.5$: if $f < 0.5$ the number of memory cells will strictly increase over time, eventually leading to a sufficiently strong adaptive immune response capable of clearing any pathogen (and thus prohibiting septic death). In what follows, we focus on the transition from health to CI, a process phenomenologically similar to inflammaging.

4.4.3 The onset of CI results from a fragility induced by a lifetime of infections

Next we consider the immunological consequences that result from encountering a sequence of infection events. Infection sequences are composed of discrete infection events, and each infection event consists of the time course following the encounter with a particular pathogen shape until a steady state of the system (i.e. health, CI, or septic death) is reached. For each infection event an immune response is generated by simulating the IIB model, given in Table 4.2. When pathogen encounters are evenly spaced in time, the number of infection events acts as a measure of age. This infection sequence encodes a lifetime of infection events, is different for different individuals, and serves as a vehicle with which to explore the variable immunological outcomes experienced by different individuals over their lifetimes.

More concretely, the IIB model is simulated for n_{tot} infections that are $\Delta T = 1000$ time units apart. The time ΔT is chosen to be sufficiently large so that the system reaches a steady state between infection events (i.e., infection events are well separated in time). The shape space in the IIB model is discrete, consisting of $S_{max} = 36$ available pathogen shapes. For each infection event, the probability p_i that a pathogen of shape i is encountered is given

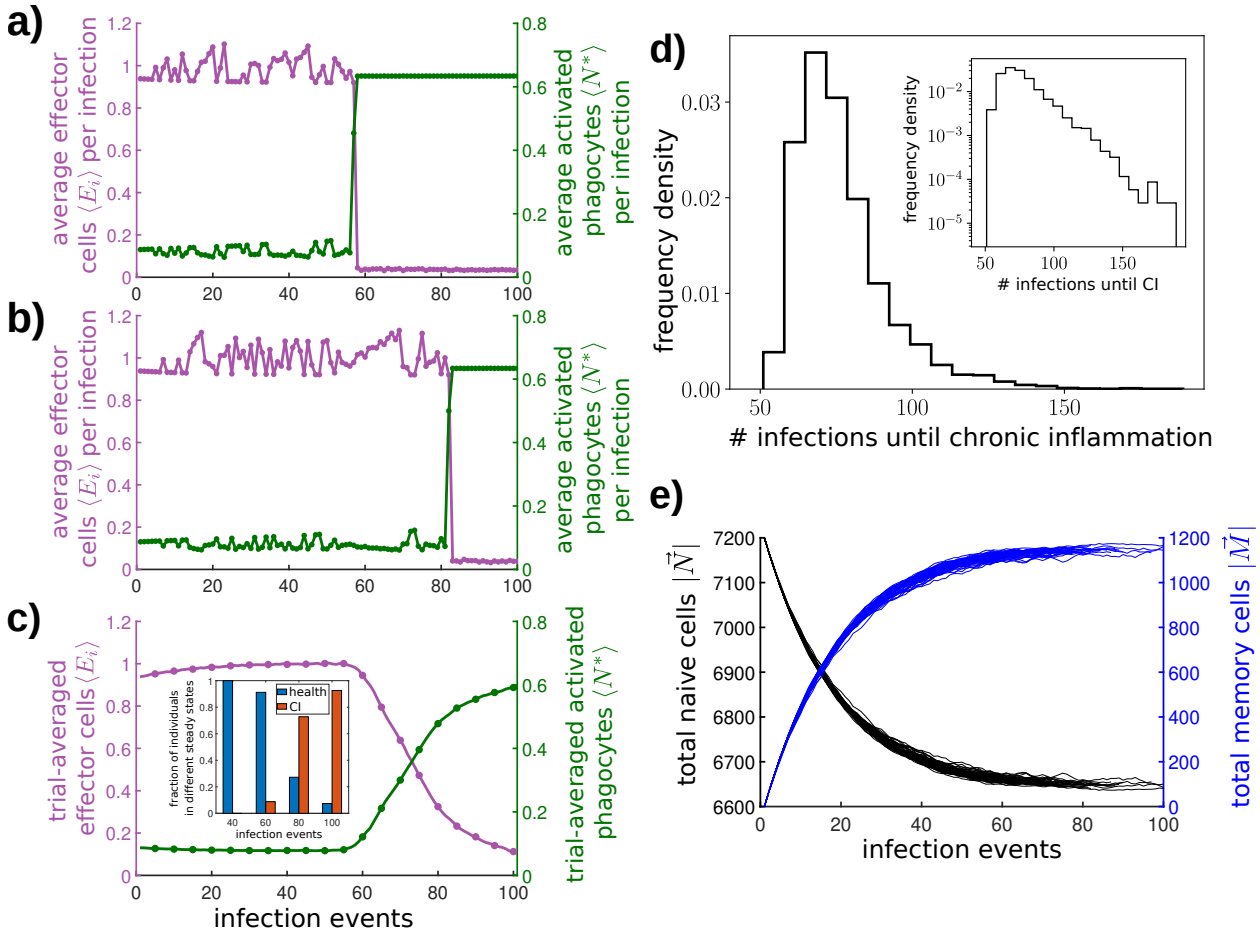

Figure 4.5: **Timing of the transition to CI is highly variable and depends on previous pathogen encounters.** (a, b) Activity of the innate (green) and adaptive (purple) immune responses over the course of 100 infection events for two different infection sequences drawn from the same statistical distribution. These regularly-spaced infection events are used to measure the age of the immune system. As a proxy for these responses, the average number of effector cells $\langle E \rangle \equiv \frac{1}{\Delta T} \int_{t_\ell}^{t_{\ell+1}} |\vec{E}(t)| \, dt$ and phagocytes $\langle N^* \rangle \equiv \frac{1}{\Delta T} \int_{t_\ell}^{t_{\ell+1}} N^*(t) \, dt$ for each infection event ℓ are plotted. The sharp transitions indicate the onset of the CI state, and occur at the 57th and 82nd infection events, respectively. (c) An ensemble average of the innate and adaptive immune responses over 1000 infection sequences (each drawn from the same pathogen distribution) smooths the variability in transition timing, though the distribution of immune responses is bimodal (c, inset). CI: chronic inflammation (d) The distribution of transition times to CI is concentrated at earlier times (on average after 76 infections). (e) The accumulation of memory cells (blue) and depletion of naive cells (black) drives immune fragility and vulnerability to new pathogen shapes (50 infection sequences shown). These figures are generated with the parameters given in Table 4.3, and with randomly generated pathogen sequences as described in Eq. 4.1.

by

$$p_i = \zeta e^{-i/\xi}, \quad i = 1,\ 2, \ldots,\ S_{max}, \tag{4.1}$$

where $\xi = 20/3$ and ζ satisfies $\sum_{i=1}^{S_{max}} p_i = 1$. This distribution allows for some pathogen shapes to be more common than others, and is similar to the one originally used by Stromberg

and Carlson [270]. A lifetime of infections is explicitly encoded in the infection sequence $\{S_\ell\}$, where each S_ℓ is the pathogen shape encountered for the ℓth infection event. Then, for each infection event $\ell = 1,\ 2, \ldots,\ n_{tot}$, a single unit of pathogen P_{S_ℓ} is added to the system at time $t_\ell = 1000(\ell - 1)$. In the simulations used to create Figs. 4.5, 4.6, and 4.7, the system is initialized with zero memory cells $M_i(0) = 0$ and a uniform distribution of naive cells $N_i(0) = 200$ across all possible pathogen shapes $i = 1,\ 2, \ldots,\ S_{max}$.

There are four important timescales in the IIB model: the time τ_{infec} required for pathogen clearance, the interval ΔT between infection events, and the timescales of naive and memory cell homeostasis control. Pathogen clearance is the fastest process, during which the homeostasis control still has little effect, and its timescale is on the order of days [273]. The timescales of naive and memory cell homeostasis are characterized by the reciprocal of their decay rates, given by $1/\delta_N$ and $1/\delta_M$, respectively. Experimental data suggest their orders as months [289] and decades or longer [290], respectively. Due to the longevity of immune memory, the interval between infections ΔT was chosen to be much shorter than the timescale of memory decay. Additionally, as in Stromberg and Carlson naive cells are assumed to regenerate and equilibrate quickly relative to the rate at which infection events occur [270]. Thus, between infection events the homeostatic naive cell population depends on a slowly-decaying memory population. More explicitly, the four timescales in IIB model are chosen to satisfy $\tau_{\text{infec}} < 1/\delta_N < \Delta T \ll 1/\delta_M$.

As the immune system ages (i.e., over the course of an infection sequence), early infection events (e.g. before the 50th infection event) are successfully cleared by the immune system, and the system returns to the health steady state. However, for later infection events the system fails to recover and instead transitions to the CI steady state, where it remains thereafter. This age-driven, history-dependent transition to CI is qualitatively similar to "inflammaging."

Depending on the details of the infection sequence, the timing of the onset of CI is highly variable. Two instances of the transition to CI, with different sequences of pathogen encounters generated from the same statistical distribution of pathogen frequencies, are displayed in Fig. 4.5a and Fig. 4.5b. The strengths of the adaptive (purple) and innate

(green) responses during each infection event are plotted in Fig. 4.5a and Fig. 4.5b, quantified by the average number of effector cells $\langle E \rangle$ and average number of phagocytes $\langle N^* \rangle$ over the course of each infection, respectively. For the ℓth infection,

$$\begin{aligned} \langle E \rangle &\equiv \frac{1}{\Delta T} \int_{t_\ell}^{t_{\ell+1}} |\vec{E}(t)| \, \mathrm{d}t, \quad \text{and} \\ \langle N^* \rangle &\equiv \frac{1}{\Delta T} \int_{t_\ell}^{t_{\ell+1}} N^*(t) \, \mathrm{d}t, \end{aligned} \tag{4.2}$$

where $|\vec{E}(t)| \equiv \sum_{i=1}^{S_{max}} E_i(t)$, and t_ℓ is the starting time of the ℓth infection event. Once the system attains the CI steady state the heightened inflammatory response will rapidly clear pathogens in any future infection, thus limiting the activity of the adaptive response. Accordingly, this leads to the sharp transition behavior of the two trajectories observed in Fig. 4.5a and 4.5b. Therefore, the onset of CI causes the innate response to strengthen while the adaptive response weakens.

When averaged over an ensemble of infection sequences, variability in the timing of the sharp transition from health to CI smooths into the crossing displayed in Fig. 4.5c (though the actual distribution of effector cells and phagocytes across infection sequences is bimodal, as seen in the inset of Fig. 4.5c). This crossing behavior is consistent with a longitudinal study of Swedish people, in which middle-aged people exhibited steady lymphocyte and neutrophil counts over the three-year span of the study, while older people (>85 years of age) exhibited significantly increased neutrophil counts and significantly decreased lymphocyte counts over the same span [291]. In addition, people with an immune risk phenotype (IRP) (which often precedes immune decline and death) commonly possess a weakened adaptive immune repertoire and experience inflammaging [32]. Taking $\langle E \rangle$ and $\langle N^* \rangle$ as proxies for adaptive and innate immune function, the age-dependent transition in the IIB model resembles the shift in immune function experienced by people with an IRP.

The number of infection events that are encountered before CI is reached is plotted as a distribution across 10,000 randomly-generated infection sequences in Fig. 4.5d. The distribution in Fig. 4.5d decays approximately exponentially for large numbers of infection events. Different choices of pathogen shape distributions result in different distributions of transition times that are qualitatively similar but in general not exponentially distributed,

as demonstrated in Fig. 4.8. Thus, the sequence of infection events is a significant driver of variability in the timing of the transition to CI.

To identify the relationship between an infection sequence and the transition from health to CI, we examine the effect of accumulated pathogen exposures on the bulk immune state variables $|\vec{M}| \equiv \sum_{i=1}^{S_{max}} M_i$ and $|\vec{N}| \equiv \sum_{i=1}^{S_{max}} N_i$; see Fig. 4.5e. Over the course of an infection sequence, memory cells are generated in response to new pathogen encounters at a faster rate than their decay. When fewer novel pathogens are encountered, memory accumulation slows, since previously encountered pathogens are cleared more quickly and adaptive immune cells are stimulated for a shorter amount of time. As the memory cell repertoire grows, the naive cell population shrinks according to the homeostatic constraints. Eventually, the immune system develops a fragility to novel pathogens due to the depleted naive cell population that results from aging; once this fragility is developed, CI will be triggered when any novel pathogen is encountered.

4.4.4 The transition to CI is influenced by previous pathogen encounters

The disparate immunological outcomes of different individuals demonstrated in Fig. 4.5 are necessarily determined by the difference in their infection sequences, since the model equations in Table 4.2 are otherwise deterministic. Until the transition to CI, the system always returns to the health steady state. In the health steady state most immune variables assume values that are independent of previous pathogen encounters; only the naive and memory cell populations occupy values that are potentially different after each infection event. Therefore different infection sequences lead to differences in memory and naive cell populations, which in turn are directly responsible for the transition to CI.

In Fig. 4.6a, the number of cognate T cells (the sum of pathogen-specific naive and memory cells $N_{S_\ell} + M_{S_\ell}$) at the beginning of each infection event ℓ are plotted for 1000 infection sequences, each consisting of 100 infection events sampled from the pathogen shape distribution Eq. (4.1). Due to the accumulation of immune memory, the number of cognate T cells specific to a pathogen shape will be greater if that pathogen has been previously

encountered. To demonstrate this, the colored points in Fig. 4.6a encode the number of times that a pathogen shape S_ℓ was encountered in the $\ell - 1$ previous infection events. A shorter time interval between infections of the same pathogen shape leads to less memory decay and more cognate T cells specific to that pathogen, as shown in the black trajectories in Fig. 4.6a. The average time before a pathogen shape is re-encountered is randomly distributed, which causes the variability of each color band (computed from 1000 simulated infection sequences). The number of cognate cells specific to novel pathogens is plotted in blue: in this case no pathogen-specific memory cells exist, and so the cognate cell and naive cell counts are the same. Therefore, the decline in the dark blue dots demonstrates the gradual decay of naive cell counts over the course of an infection sequence (i.e., as the immune system ages).

The infection events that trigger the transition to CI are indicated by the red circles in Fig. 4.6a. These transitions always occur (i) in response to a novel pathogen shape (red circles are laid on top of the blue band), and (ii) when the number of cognate naive cells falls below a threshold (approximately 180 in Fig. 4.6a). When both conditions are met, the adaptive immune response is low in magnitude and unable to produce sufficient anti-inflammatory cytokines to suppress the innate immune branch. This weakened adaptive immune response, itself a function of the infection sequence, is the principal driver of CI in the IIB model. Accordingly, the IIB model exhibits a "robust yet fragile" behavior [270, 292]: it is robust to frequently-encountered pathogens, yet fragile to novel pathogens.

4.4.5 Manipulating immune system fragility via synthetic infection sequences

To probe the variability in the timing of the transition to CI as the immune system ages, we examine three synthetic infection sequences that are reorderings of an "authentic" infection sequence sampled from Eq. (4.1). These sequences, showcased schematically in Fig. 4.6b and detailed in the text below, deliberately structure the order of pathogen encounters to induce different levels of fragility towards novel pathogen shapes. The synthetic and authentic infection sequences affect the rate of memory cell accumulation and naive cell loss, and

in turn, alter the timing of the onset of CI. An example authentic sequence, along with its corresponding synthetic sequences, is illustrated in Fig. 4.6c. In Fig. 4.6d naive cell statistics of 50 authentic infection sequences are compared with the statistics generated by their synthetic counterparts.

In the *clustered* synthetic ordering (Fig. 4.6b-d, orange), the infection sequence is ordered so that the most common pathogens are encountered first and the rarest pathogens are encountered last. In this case, pathogens are immediately reencountered so memory cells do not significantly decay between infections, and the accumulated immune memory causes an accelerated immune response that generates fewer memory cells. Therefore, this reordering is a *lower bound* for the naive cell loss rate. Indeed, the clustered sequence leads to the slowest loss of naive cells among the authentic and synthetic sequences in Fig. 4.6d. The clustered sequence provides reliable protection that delays the transition to CI until infection 89, compared with the authentic sequence that transitions after infection 66, as shown in Fig. 4.6c.

In the *cyclic* synthetic ordering (Fig. 4.6b-d, yellow) infections are ordered so that all available pathogen types are encountered as early as possible: in this ordering each pathogen type is encountered once before any pathogen is encountered for the second time, then each pathogen type is encountered twice before any pathogen is encountered for the third time, and so on. Constantly encountering new pathogen types drives the accumulation of memory cells and in turn naive cell loss at an accelerated rate. Thus the cyclic sequence yields an *upper bound* for the naive cell loss rate, as in Fig. 4.6d. At the same time, since this synthetic sequence is structured to front-load every pathogen type that can be encountered early in the infection history, the generated memory cells eventually provide full protection against each pathogen type, and the CI state never occurs, as in Fig. 4.6c. Thus, broad exposure to pathogens early in an individual's infection life history can provide adaptive-mediated protection from CI in the IIB model.

Lastly, the *incomplete cyclic* synthetic ordering (Fig. 4.6b-d, purple) is similar in construction to the cyclic ordering, except that one rare pathogen is intentionally omitted from the initial pathogen cycles. Then, this pathogen is presented at a later time to trigger CI.

The incomplete cyclic ordering induces a fragile immune response: naive cells deplete nearly as quickly as for the cyclic ordering, and incomplete immune memory coverage causes vulnerability to novel pathogens. Thus, the onset of CI is accelerated in the incomplete cyclic ordering, with the onset occurring during the 48th infection in Fig. 4.6c.

The clustered and cyclic synthetic immune histories demonstrate how some pathogen sequences can delay the onset of CI, either by prolonging the abundance of naive cells or by quickly acquiring full immune memory coverage across all pathogen shapes. In contrast, the incomplete cyclic sequence demonstrates how pathogen sequences can induce immune fragility, by quickly depleting naive cells while remaining vulnerable to novel pathogens; alternatively, the incomplete cyclic sequence shows how the introduction of a new pathogen species, either through mutation or migration to a new environment, can break existing memory coverage and lead to immune fragility.

4.4.6 The adaptive immune response is subject to a trade-off between pathogen clearance and suppressing inflammation

In the original adaptive immune model by Stromberg and Carlson [270], the sole function of effector cells was to clear pathogens. However, the diverse repertoire of effector T cells— including helper T cells, cytotoxic T cells, and regulatory T cells— can additionally exhibit anti-inflammatory functions [256, 259, 293]. Incorporating these features in the IIB model leads to a trade-off between pathogen clearance and inflammation suppression that can be explored quantitatively.

Specifically, in the IIB model a proportion p of effector cells are allocated to pathogen clearance (a responsibility of cytotoxic T cells with the aid of helper T cells [294]), while a proportion $1 - p$ of effector cells are allocated to the production of anti-inflammatory cytokines (a responsibility of regulatory T cells [293]). These dual functions are presented schematically in Fig. 4.1 and explicitly in Table 4.2.

For smaller values of p, effector cells are increasingly used to combat inflammation, which delays the onset of CI as demonstrated in Fig. 4.7a. On the other hand, at smaller values of

p more resources are allocated to combat inflammation, so the diminished innate response slows the rate of pathogen clearance. This is quantified in Fig. 4.7b, which plots an ensemble average (across 1000 infection sequences) of the cumulative pathogen load L_ℓ for each infection event: for the ℓth infection,

$$L_\ell \equiv \int_{t_\ell}^{t_{\ell+1}} P_{S_\ell}(t)\ \mathrm{d}t, \tag{4.3}$$

where S_ℓ is the pathogen shape and t_ℓ the starting time of the ℓth infection event. This measure was originally introduced in Stromberg and Carlson [270] where it was called the "loss function," and we similarly use it here as a proxy for the damage the pathogen inflicts during each infection event. In Fig. 4.7b, smaller values of p lead to greater cumulative pathogen load for each infection event and vice versa.

Note that for p as large as 0.9 there is a further drop in L_ℓ for $\ell \gtrsim 60$ due to the onset of the CI steady state in a larger proportion of individuals. Once the CI state is reached, the non-zero activated phagocyte population N^* rapidly responds to infection events, efficiently clears pathogens, and returns to the CI steady state. In this sense, the onset of CI in aging immune systems acts as a protective mechanism that shields the immune response from future pathogen encounters and minimizes the damage that pathogen inflicts. Interpreted biologically, it is advantageous for organisms to minimize cumulative pathogen load while also avoiding the early onset of CI. Evolutionarily, these two opposing selection forces should lead to an intermediate optimal p in which effector cells allocate resources both to pathogen clearance and innate suppression.

4.5 Discussion

4.5.1 Immunosenescence as an emergent immune response

Though immunosenescence affects every aging individual, the mechanisms through which it develops are not yet fully understood. In this study we demonstrate through quantitative modeling how physiological markers of immunosenescence can arise from the accumulated effect of pathogen encounters. In particular, clonal expansion and homeostatic maintenance

lead to an increase in memory cells and a decrease in naive cells, which are qualitatively consistent with their clinically-observed abundances [268]. Accumulated memory cells protect the immune system against previously-encountered pathogens, but the shrinkage in the naive cell pool renders the immune system vulnerable to novel pathogens and is the key indicator of immune system fragility. While a similar mechanism was demonstrated in the model by Stromberg and Carlson [270], in their study the overspecialized immune repertoire led to increased cumulative pathogen load. In the IIB model, the acquired immune fragility is characterized by a transition to a chronic inflammatory state, and the timing of this transition is highly variable and depends on the infection history.

In addition to this mechanism of imbalanced immunological space, several other immune functions vary with age and could play a role in the development of immunosenescence. For example, clinical studies observed that the average cytotoxicity of natural killer cells decreases with age [295, 296], cell signaling between immune cells can become impaired with age [297], and thymus involution leads to decreased T cell production with age [298, 299]. The current formulation of the IIB model exhibits an inflammaging-like behavior without taking these additional factors into account. However, in future work physiological parameters of the model could be used as a proxy for these observed behaviors: for example, the decreased cytotoxicity of natural killer cells (which are innate) could be incorporated by decreasing k_{pn} with age, the impaired cell signaling in T cells (which are adaptive) could be achieved by decreasing γ with age, or the reduced thymus output could be modeled by decreasing θ_N with age. The calculated immune outcomes that result from these modifications could shed light on the relative contributions to immunosenescence from memory-induced fragility in the adaptive response, an impaired innate response, and an impaired adaptive response.

4.5.2 CI as inflammaging and the collaboration between innate and adaptive responses

The CI steady state has two physiological interpretations. First, the runaway tissue damage caused by the sustained inflammatory response may cause death in the host, as was implied

by Reynolds *et al.* when they called this steady state "aseptic death." Second, if the sustained inflammatory response is relatively minor, the CI state can be interpreted as "inflammaging," a chronic low-grade inflammation that is common among the elderly [31]. In this study we choose this second interpretation and construe the transition to the CI state as inflammaging. Accordingly, the mechanisms of the IIB model that induce this transition might inform the biological mechanisms that they emulate.

For example, recent work has suggested that the development of inflammaging might be a result of immune system remodeling: as immunosenescence lessens the efficacy of the adaptive immune response, the body relies on inflammaging for protection against pathogens via the innate immune response [268]. Similarly, the adaptive response in the IIB model is subject to a trade-off between clearing pathogens and suppressing inflammation. In part based on recent work demonstrating that the adaptive response can act to suppress a hazardous innate response [259], recent theories suggest that this suppression might have been the evolutionary driver that promoted the development of an adaptive immune response [258]. Correspondingly, in the IIB model when pathogens are introduced to a system in the CI state, they are cleared almost immediately since the inflammatory response is already primed.

Evolutionarily, the innate immune response preceded the creation of the adaptive response [300]. This is consistent with the taxonomic complexity of organisms, in which invertebrates possess only an innate response while vertebrates possess the additional capacity for pathogen-specific immune memory [301]. Additionally, adaptive immune components are dependent on innate cells— for example, the activation of an adaptive response through Ag presentation relies on dendritic cells. The evolutionary drivers of the adaptive immune response could be explored with immune models that quantify the added benefit of possessing an adaptive immune system.

4.5.3 Age-dependent strength of immune response

The efficiency of the human immune system changes in a non-monotonic manner as one ages: it is weak in infancy and dependent on maternal antibodies; then it grows stronger as the innate and adaptive responses mature and as immune memory is accumulated; and finally it plummets in the elderly [302]. As people age, effector T cell levels drop, CI builds [32, 303], and immune outcomes among the elderly become extremely variable [303].

In this work we present a potential mechanism for these clinically-observed aging trends, driven by overspecialization of the adaptive immune repertoire. The accumulation of memory cells initially strengthens the immune response against previously-encountered pathogens. Eventually, memory cells become overspecialized and restrict the growth of naive cells, rendering aged individuals vulnerable to rare pathogen types. In the IIB model the onset of the CI state is variable, and dependent on the history of previous pathogen encounters. The age-dependent immune system efficiency observed in the IIB model is consistent with the previously mentioned clinically-observed immune behaviors.

4.5.4 Imprinting and vaccines

The shape space formulation of the adaptive immune response produces results that are qualitatively similar to the clinically-observed behaviors of immune imprinting [304] and the decreased efficacy of vaccines in the elderly [296]. Immune imprinting occurs when individuals exhibit sustained memory to the pathogens they were exposed to early in their life. In the IIB model naive cells are more abundant at the beginning of an infection sequence, and as memory cells accumulate over time, homeostatic pressures drive down the population of naive cells. Thus, during the first several infection events the larger naive cell pool will induce a stronger adaptive response and therefore generate a stronger memory for encountered pathogens. On the contrary, near the end of an infection sequence the diminished naive pool will induce a weaker adaptive response to a novel pathogen, and generate a weaker immune memory. If we interpret vaccination as an exposure to a novel pathogen, then the clinically-observed characteristics of immune imprinting and vaccination in the elderly are

qualitatively captured by the IIB model.

4.6 Conclusion

The progression towards immunosenescence is a dynamical process influenced by a lifetime of pathogen encounters, physiological alterations, genetic factors, and general lifestyle choices. This blend of factors makes it difficult to isolate and identify the most relevant causative agents of immunosenescence. Therefore, mathematical models hold great utility in their ability to probe the mechanisms of immunosenescence.

In this study we developed the IIB model, which incorporates the structure of the innate and adaptive immune branches, and exhibits behaviors that are qualitatively consistent with clinically-observed phenomena. We found that repeated pathogen encounters cause an overspecialization of memory cells and depletion of naive cells as the immune system ages. Over time these effects render the immune system fragile to novel pathogens, the encounter of which will trigger an irreversible transition of the system to a CI state. By describing immune dynamics with a mathematical model, we demonstrated how the feedback between innate and adaptive immune responses could give rise to diverse immune courses and outcomes. Going forward, experimental studies combined with knowledge-based quantitative models will continue to illuminate the impact of aging on immune efficacy.

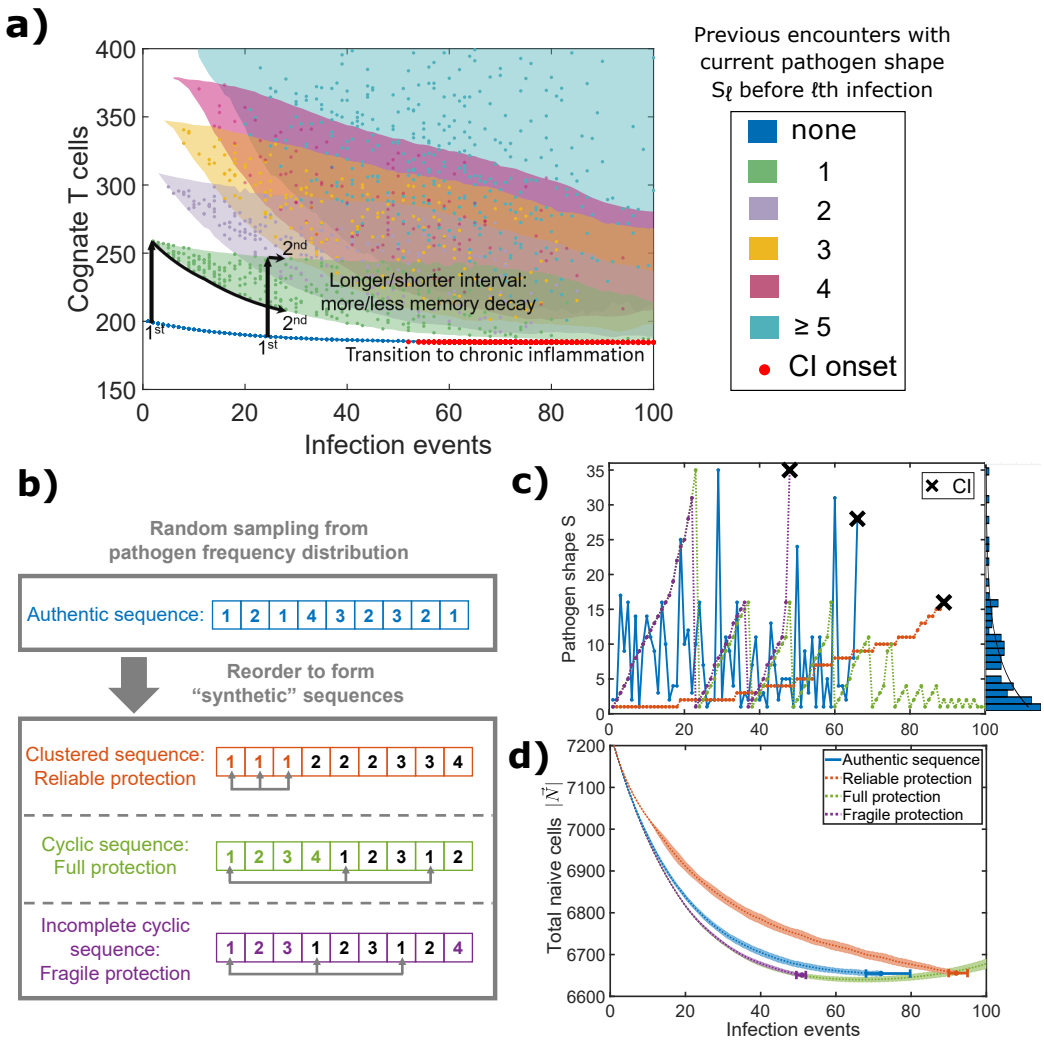

Figure 4.6: **Aging-induced transition to CI is driven by depletion of naive cells and lack of protection from memory cells.** The number of encountered infections is used as a proxy for the age of the immune system. (a) The number of cognate T cells specific to a novel pathogen shape (equal to the sum of naive and memory cells) is the key indicator for whether an infection event will trigger the CI steady state. Here, cognate T cell counts specific to an encountered pathogen P_ℓ are plotted for each infection event ℓ across 20 infection sequences sampled from Eq. (4.1). The color of each point indicates the number of times that the encountered pathogen P_ℓ has been previously encountered. The colored bands are generated from 1000 infection sequences, and envelope the observed cognate cell counts. The large red circles in the lower-right corner mark the infections events that trigger CI across all 1000 infection sequences, which occur when a novel pathogen is encountered after naive cells have been depleted below some threshold. A shorter time interval between pathogen encounters of the same shape results in less memory cell decay and hence more cognate T cells, and this effect causes the shape of the colored bands. (b) We consider three synthetic reorderings of each "authentic" randomly generated pathogen sequence: the clustered sequence orders pathogens according to their prevalence; the cyclic sequence orders them to ensure immediate exposure to all pathogen types; and the incomplete cyclic sequence induces fragility by quickly depleting naive cells and then introducing a novel pathogen. (c) The authentic sequence and three synthetic sequences transition to CI at different times (black crosses). The pathogen shape distribution for this infection history (right histogram) is drawn from the theoretical shape distribution (black line overlaid) given by Eq. (4.1). (d) The naive cell pool is depleted at different rates depending on how infection events are ordered. Naive cell counts and their variation across the 50 authentic sequences considered in panel (a) are shown for the three synthetic sequences. Error bars for the timing of CI are 50% confidence intervals.

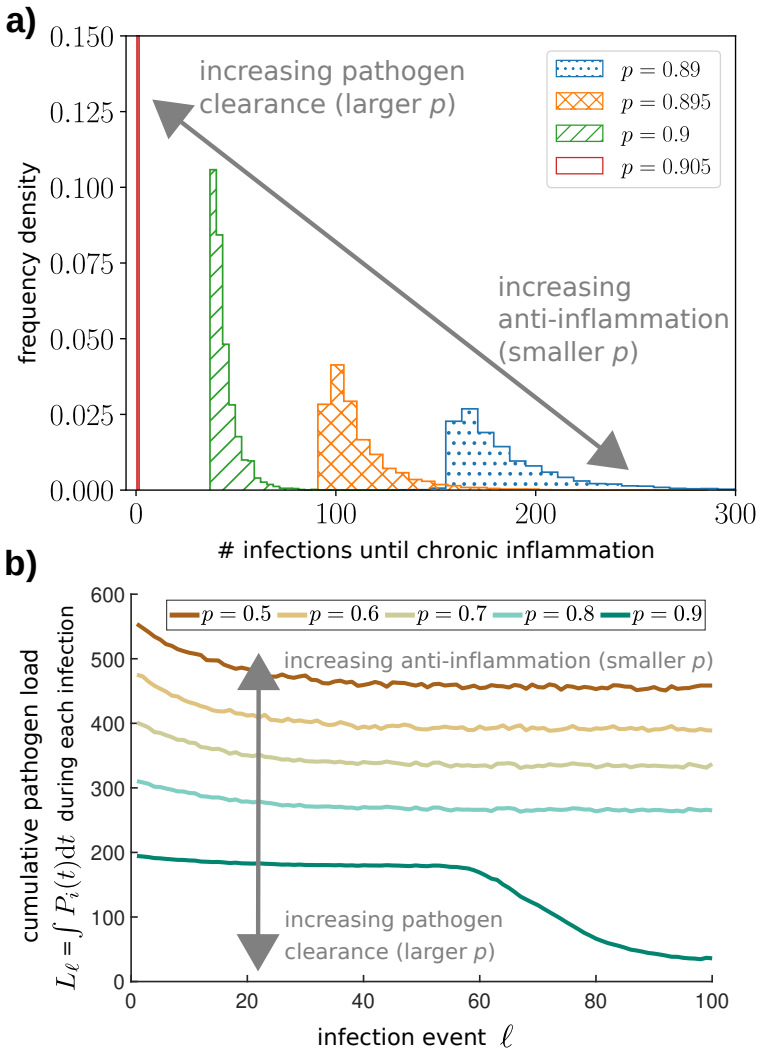

Figure 4.7: **Effector cells are subject to a trade-off between clearing pathogens and suppressing inflammation as the immune system ages.** (a) The onset of CI (histograms as in Fig. 4.6d) is delayed for lower values of p, i.e. when the anti-inflammatory role of effector cells is increased. (b) The cumulative pathogen load L_ℓ over the course of each infection event (averaged over 1000 infection sequences) is larger for smaller values of p. The drop in L_ℓ after the 60th infection event for $p = 0.9$ is caused by the onset of CI, which compensates for the overspecialized adaptive immune repertoire. To generate the statistics in panel (a), the homeostatic parameters δ_N, δ_M, R_0, and θ_N were modified to ensure that the timescales of infection clearance and homeostatic response were separated enough for us to use the adaptive programming method, as described in Table 4.4 of the Appendices. The simulations in panel (b) are generated with the parameters given in Table 4.3.

APPENDIX

4.A Immune response timescales in the IIB model

There are three principle timescales in the IIB model: a short timescale associated with pathogen clearance by the immune response; a medium timescale associated with the homeostasis of naive cells; and a long timescale that describes the gradual attrition of memory cells. When these three timescales are sufficiently separated, and when pathogen encounters are successfully cleared by the immune system, the full immune model in Table 4.2 may be considerably simplified by separating the innate and adaptive immune responses from the effects of the homeostatic response on naive and memory cells. These analytic approximations underpin the dynamic programming approach used to simulate Fig. 4.7a, and explicitly inform how the populations of naive and memory cells change during the homeostatic response.

4.A.1 Pathogen clearance (fast timescale)

First, when the immune response to a pathogen is much faster than any homeostatic contributions, then the immune response to a pathogen infection of type P_i at time t' is exclusively a function of the naive and memory cell populations specific to that pathogen type at that time, $N_i(t')$ and $M_i(t')$. In particular, this response is not a function of the bulk naive or memory cell populations $|\vec{N}(t')|$ or $|\vec{M}(t')|$, where $|\cdot|$ denotes the 1-norm. When a pathogen is successfully cleared the immune system will return to a health steady state. In the absence of a homeostatic response, this steady state has unique equilibrium values for every immunological variable except for the naive and memory cells $\vec{N}$ and $\vec{M}$— at steady state, the naive and memory cell populations are free to be any fixed value.

4.A.2 Carrying capacity R_0

In the IIB model the parameter R_0 corresponds to the carrying capacity of naive and memory cells, which is why it appears in the homeostasis terms. In this study, simulations typically are initialized with $N_i(0) = R_0/S_{max}$, and $M_i(0) = 0$, for all pathogen shapes $i \in 1, \ldots, S_{max}$. We require this initial condition to be steady state of the system, which requires the homeostatic response to vanish, so that

$$\frac{\mathrm{d}N_i}{\mathrm{d}t} = \theta_N - \delta_N N_i \frac{|\vec{M}| + |\vec{N}|}{R_0} = \theta_N - \delta_N \frac{R_0}{S_{max}} \frac{R_0}{R_0} = 0. \tag{4.4}$$

Thus, in order for this initial condition to be a steady state of the system, we require that the parameter θ_N satisfies $\theta_N = \delta_N R_0/S_{max}$.

4.A.3 Naive cell homeostasis (medium timescale)

Once the pathogen is cleared at some time t'', say, the naive and memory cell populations $\vec{N}(t'')$ and $\vec{M}(t'')$ are driven by homeostasis to their equilibrium values, while all other immunological variables remain fixed at their health state equilibrium values. Once the pathogen has been cleared, and using $\theta_N = \delta_N R_0/S_{max}$, the dynamics of the naive and memory cells are given by

$$\begin{aligned} \frac{\mathrm{d}N_i}{\mathrm{d}t} &= \delta_N \left[\frac{R_0}{S_{max}} - N_i \frac{|\vec{M}| + |\vec{N}|}{R_0} \right], \quad \text{and} \\ \frac{\mathrm{d}M_i}{\mathrm{d}t} &= -\delta_M M_i \frac{|\vec{M}| + |\vec{N}|}{R_0}, \end{aligned} \tag{4.5}$$

where as before $|\cdot|$ is the 1-norm. Explicitly, the assumption that naive cell homeostasis is much faster than memory cell attrition is a statement that $\delta_N >> \delta_M$. Thus, on the timescale of naive cells, the bulk population of memory cells $|\vec{M}|$ and the populations of each shape of memory cell M_i are approximately constant.

The dynamics of the bulk naive cells $|\vec{N}|$ are given by

$$\frac{\mathrm{d}|\vec{N}|}{\mathrm{d}t} = \sum_{i=1}^{S_{max}} \frac{\mathrm{d}N_i}{\mathrm{d}t} = \delta_N \left[R_0 - |\vec{N}| \frac{|\vec{M}| + |\vec{N}|}{R_0} \right]. \tag{4.6}$$

We note that in the absence of any memory cells $|\vec{M}| = \vec{0}$, $|\vec{N}| = R_0$, which is how we initially defined the carrying capacity R_0. When memory cells are present at a presumed constant abundance $|\vec{M}|$, the steady state bulk population of naive cells $|\vec{N}|^*$ is

$$|\vec{N}|^* = \frac{-|\vec{M}| + \sqrt{|\vec{M}|^2 + 4R_0^2}}{2}. \tag{4.7}$$

The bulk dynamics $|\vec{N}(t)|$ in Eq. 4.6 starting from initial condition $|\vec{N}(0)|$ may be solved analytically, with the solution

$$|\vec{N}(t)| = \frac{|\vec{M}|}{2}\left[-1 + \sqrt{1 + \frac{4R_0^2}{|\vec{M}|^2}}\tanh\left((t+K)\delta_N\sqrt{\frac{|\vec{M}|^2}{4R_0^2}+1}\right)\right], \tag{4.8}$$

where

$$K = \frac{1}{\delta_N}\sqrt{\frac{1}{1+|\vec{M}|^2/(4R_0^2)}}\tanh^{-1}\left[\frac{|\vec{N}(0)| + |\vec{M}|/2}{R_0}\sqrt{\frac{1}{1+|\vec{M}|^2/(4R_0^2)}}\right]. \tag{4.9}$$

If the number of pathogen shapes S is large, then any given infection will induce a relatively small change in $|\vec{N}|$ (of order $|\vec{N}|/S$). In this limit where an infection acts as a perturbation away from the bulk naive cell equilibrium value, so that $|\vec{N}(t'')|$ is close to $|\vec{N}|^*$, these dynamics may be approximated as

$$|\vec{N}(t)| = |\vec{N}|^* - \left(|\vec{N}|^* - |\vec{N}(0)|\right)\exp\left(-\frac{\delta_N}{R_0}(2|\vec{N}|^* + |\vec{M}|)\ t\right), \tag{4.10}$$

where we use the approximation $1 - \tanh(x) \approx 2e^{-2x}$ for large x. Thus, the bulk naive cell population approaches its equilibrium with timescale $\frac{R_0}{\delta_N(2|\vec{N}|^* + |\vec{M}|)}$.

The individual naive cell types each obey dynamics

$$\frac{\mathrm{d}N_i}{\mathrm{d}t} = \delta_N\left[\frac{R_0}{S} - N_i\frac{|\vec{M}| + |\vec{N}|}{R_0}\right]. \tag{4.11}$$

At steady state, assuming $|\vec{M}|$ is constant and $|\vec{N}|$ has reached equilibrium at $|\vec{N}|^* = \frac{-|\vec{M}| + \sqrt{|\vec{M}|^2 + 4R_0^2}}{2}$, the individual naive cells approach their equilibrium state N_i^* at

$$N_i^* = \frac{2R_0^2}{S\left(|\vec{M}| + \sqrt{|\vec{M}|^2 + 4R_0^2}\right)} = \frac{|\vec{N}|^*}{S}. \tag{4.12}$$

If the bulk dynamics are assumed to be at equilibrium $|\vec{N}(t)| = |\vec{N}|^*$, then the individual dynamics $N_i(t)$ in Eq. 4.11 starting from initial condition $N_i(0)$ may be solved analytically with the solution

$$N_i(t) = N_i^* \ - (N_i^* - N_i(0)) \exp\left(-\frac{\delta_N}{R_0}(|\vec{M}| + |\vec{N}|^*)\ t\right). \tag{4.13}$$

Thus, in this approximation the individual naive cells approach their equilibrium with timescale $\frac{R_0}{\delta_N(|\vec{M}|+|\vec{N}|^*)}$.

By further assuming that $|\vec{N}(t)|$ is approximated by Eq. (4.10) and that the deviation of bulk immune cells from their equilibrium $|\vec{N}|^* - |\vec{N}(0)|$ is small, the individual naive cells may be analytically computed to first order in $|\vec{N}|^* - |\vec{N}(0)|$ as

$$\begin{aligned} N_i(t) = N_i^* \ &- (N_i^* - N_i(0)) \exp\left(-\frac{\delta_N}{R_0}(|\vec{M}| + |\vec{N}|^*)\ t\right) \\ &+ \frac{|\vec{N}|^* - |\vec{N}(0)|}{S}\left[\exp\left(-\frac{\delta_N(|\vec{M}| + |\vec{N}|^*)}{R_0}\ t\right) - \exp\left(-\frac{\delta_N(|\vec{M}| + 2|\vec{N}|^*)}{R_0}\ t\right)\right]. \end{aligned} \tag{4.14}$$

4.A.4 Memory cell homeostasis (long timescale)

On long timescales, the bulk and individual naive cell populations will reach their equilibrium values at $|\vec{N}|^* = \frac{-|\vec{M}|+\sqrt{|\vec{M}|^2+4R_0^2}}{2}$ and $N_i^* = \frac{|\vec{N}|^*}{S}$. At this point, the dynamics of the bulk memory cell populations will be given by

$$\frac{\mathrm{d}|\vec{M}|}{\mathrm{d}t} = -\frac{\delta_M}{2R_0}|\vec{M}|\left(|\vec{M}| + \sqrt{|\vec{M}|^2 + 4R_0^2}\right). \tag{4.15}$$

When $|\vec{M}|^2 << 4R_0^2$, these dynamics reduce to

$$\frac{\mathrm{d}|\vec{M}|}{\mathrm{d}t} = -\frac{\delta_M}{2R_0}|\vec{M}|^2 - \delta_M|\vec{M}|, \tag{4.16}$$

which lead to bulk memory cell populations

$$|\vec{M}(t)| = \frac{2R_0|\vec{M}(0)|}{(2R_0 + |\vec{M}(0)|)e^{\delta_M t} - |\vec{M}(0)|}, \tag{4.17}$$

indicating that the bulk memory cells vanish with time constant $1/\delta_M$. In the limit where $t << 1/\delta_M$, these dynamics further reduce to

$$|\vec{M}(t)| = |\vec{M}(0)|\left[1 - \left(1 + \frac{|\vec{M}(0)|}{2R_0}\right)\delta_M t\right] \tag{4.18}$$

Individual memory cells have dynamics given by

$$\frac{\mathrm{d}M_i}{\mathrm{d}t} = -\delta_M M_i \frac{|\vec{M}| + |\vec{N}|}{R_0}, \tag{4.19}$$

and after making the same assumptions that were used to approximate $|\vec{M}(t)|$, the dynamics of individual memory cells may be approximated as

$$\frac{\mathrm{d}M_i}{\mathrm{d}t} = M_i(0)\left[1 - \left(1 + \frac{|\vec{M}(0)|}{2R_0}\right)\delta_M t\right], \tag{4.20}$$

indicating that individual memory cells decay at the same rate as bulk memory cells.

4.B Adaptive programming method for efficient simulation of immunological trajectories

To efficiently simulate the infection trajectories required to generate the histograms in Fig. 4.7a, an adaptive programming method was implemented that leveraged the analytic approximations in the previous section to the homeostatic dynamics of the IIB model. In the absence of homeostatic effects (i.e., when $\delta_M = \delta_N = 0$), the immunological trajectory of an initially healthy state exposed to a fixed quantity of encountered pathogen P_i will depend only on the number of pathogen-specific memory and naive cells, M_i and N_i, respectively. If the system returns to health, only the steady-state values of the pathogen-specific memory and naive cells will have changed. Additionally, in the absence of homeostatic effects, there is no crosstalk between memory and naive cells associated with different pathogen shapes. Thus, in the dynamic programming method, we create a look-up table gridded by naive and memory cell initial conditions (see Fig. 4.9 for a heat map representation of this look-up table). Then, an instantiation of the IIB model is run for each combination of naive and memory cell initial conditions, and the final naive and memory cell populations (and whether the system enters the CI state) are recorded. In the look-up table used to create Fig. 4.7a, naive and memory cell initial conditions were resolved to 7 cells, spanning from 0 cells to 1333 cells.

When the timescales of pathogen clearance and homeostasis are well separated, this adaptive programming approach first approximates the time course of an infection event

by looking up the predicted final naive and memory cell populations. Then, the analytic homeostatic approximations detailed in the last section— in particular the approximation of homeostatic naive cell dynamics $N_i(t)$ given by Eq. 4.14, and the approximation of homeostatic memory cell dynamics $M_i(t)$ given by Eq. 4.20— are used to update every naive and memory cell population. This process is repeated for each subsequent pathogen encounter. Notably, once the look-up table is created, both of these steps (in terms of computational complexity) are constant in time.

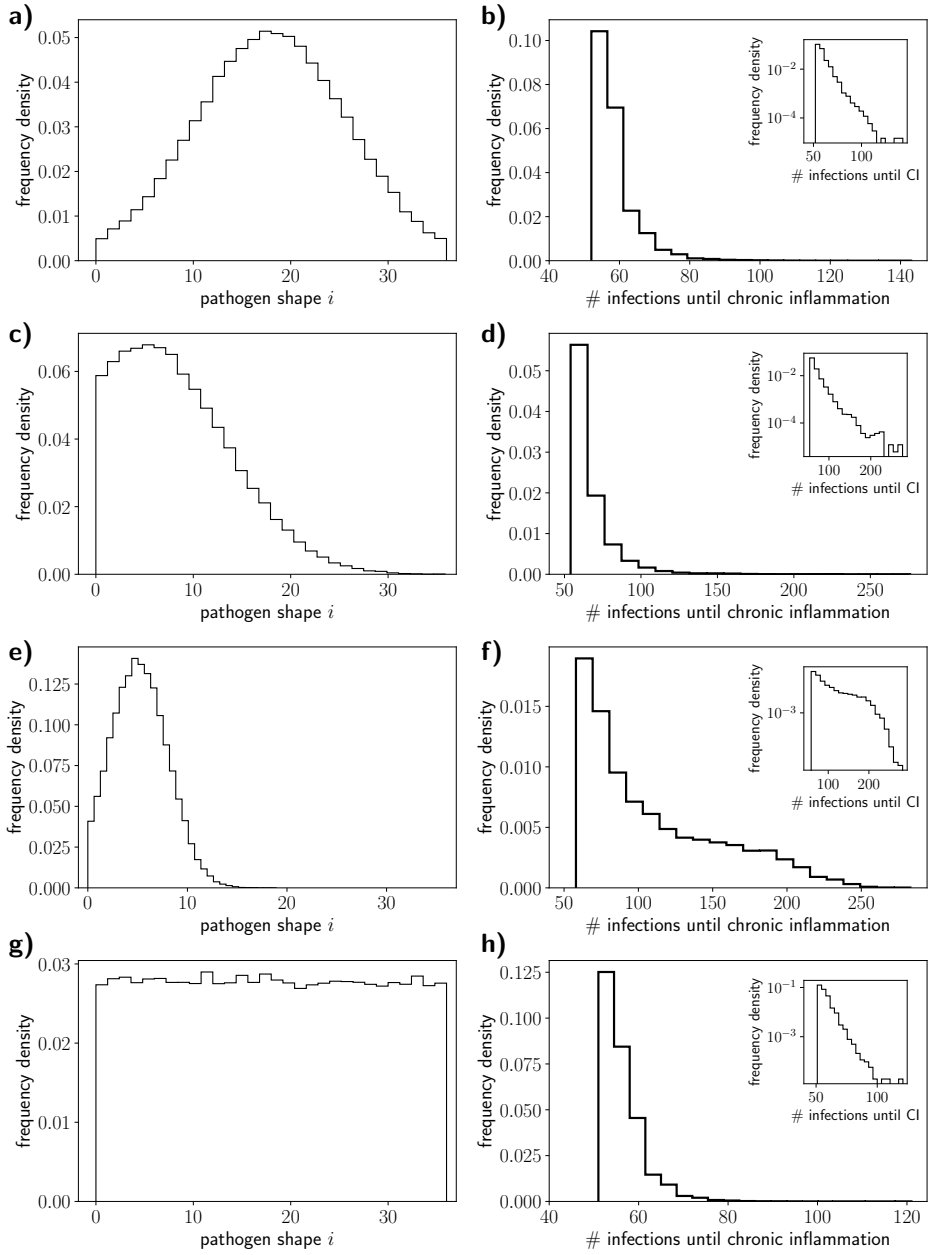

Figure 4.8: **The onset of CI for different pathogen shape distributions.** In this chapter uses an exponential pathogen shape distribution, given in Eq. 4.1, that produces a distribution of transition times to CI displayed in Fig. 4.5d. Here, additional pathogen shape distributions beyond are explored (a, c, e, g) that result in different distributions of transition times to CI (b, d, f, h). The four pathogen shape distributions used are: (a, b) a truncated normal with a mean of 18 and standard deviation of 8; (c, d) a truncated normal with a mean of 5 and standard deviation of 8; (e, f) a truncated normal with a mean of 5 and standard deviation of 3; and (g, h) a uniform distribution. Notably, the qualitative shape of the distribution of transition times to CI are similar regardless of pathogen shape distribution, although they are in general not exponential. Simulations were efficiently run using an adaptive programming approach, and each histogram contains the results from 15,000 simulations.

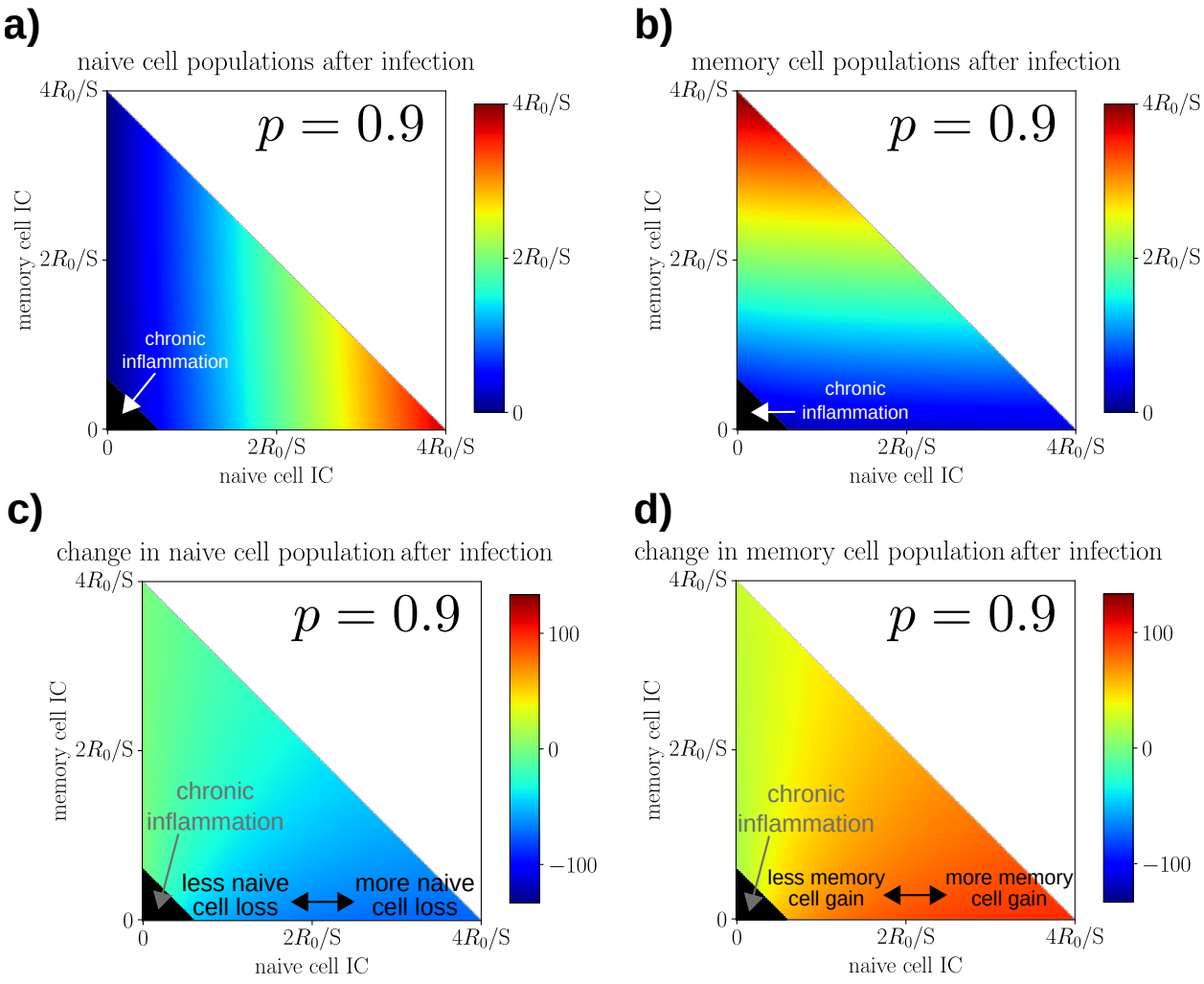

Figure 4.9: **A heatmap representation of the pathogen encounter lookup table generated by the adaptive programming method.** For an infection event in which 1 unit of pathogen is encountered, for some pathogen-specific naive and memory cell initial conditions, this lookup table stores the number of (a) naive cells and (b) memory cells that will exist following pathogen clearance. Panels (c) and (d) display the change in naive and memory cell populations following an infection event. When the sum of the initial populations of pathogen-specific naive and memory cells is too small, the system transitions into the CI state (black triangle). For larger values of p, there are more naive and memory cell initial conditions that lead to CI. In these simulations, $S = S_{max}$, the total number of pathogen shapes.

.

Figure #	Parameters values that differ from Table 4.2
Figure 2	$S_{max} = 1$, $k_{pm} = 1$, $\mu_d = 0.05$, $\gamma = 1$; $\delta_N = \delta_M = \theta_N = 0$ (no homeostatic response); initial conditions $P_1(0) = 3$, $N_1(0) = 10$, $M_1(0) = 0$
Figure 3a	$S_{max} = 1$, $k_{pm} = 1$, $\mu_d = 0.05$; $\delta_N = \delta_M = \theta_N = 0$ (no homeostatic response); initial conditions $P_1(0) = 10$, $N_1(0) = 10$, $M_1(0) = 0$
Figure 3b	$S_{max} = 1$, $f = 0.7$, $k_{pg} = 1.1$, $k_{dn} = 0.07$; $\delta_N = \delta_M = \theta_N = 0$ (no homeostatic response); initial conditions $P_1(0) = 10$, $N_1(0) = 10$, $M_1(0) = 0$
Figure 3c	$S_{max} = 1$, $k_{pg} = 0.5$; $\delta_N = \delta_M = \theta_N = 0$ (no homeostatic response); initial conditions $P_1(0) = 10$, $N_1(0) = 10$, $M_1(0) = 0$
Figure 7a	$\delta_N = 5e{-}4$, $\delta_M = 1e{-}7$, $R_0 = 8000$, $\theta_N = 0.11$, p as in figure legend; initial conditions $N_i(0) = 222$, $M_i(0) = 0$

Table 4.4: **Parameters used to generate each figure in chapter 4** Most simulations in this chapter were run using parameters as described in Table 4.2. In Figs.4.2,4.3,4.7, slightly modified parameters were used, as described in this table. Figs.4.2 and 4.3 demonstrate the dynamics of a single infection event, and do not include a homeostatic response. Fig. 4.7a was generated using the adaptive programming method, and the homeostatic parameters δ_N, δ_M, R_0, and θ_N were selected to ensure a separation of timescales between the infection event, naive cell homeostasis, and memory cell homeostasis.

CHAPTER 5

Conclusions and Discussions

Taken together, my dissertation has contributed towards understanding the immune system's evolution against diverse Ag. The studies presented cover both evolutions in natural and controlled (through vaccination) scenarios, and both the timespan of single infections and host lifetime. In particular, I have investigated (1) the determinants of coevolutionary outcomes between immune system and highly-mutable pathogens, (2) viable vaccine strategies to evolve bnAbs against these highly-mutable pathogens, and (3) consequences and corresponding remedies of an aged immune repertoire, as a result of adapting to diverse pathogen species over host lifetime.

Chapter 2 and 3 focus on the evolution of immune system against diverse variants of a pathogen. Chapter 2 shows that in natural infections, conservation level and initial diversity of Ag jointly determine the timing and efficacy of successive narrow and broad Ab responses, which in turn accounts for the coevolutionary transition between viral persistence, clearance and rebound. In particular, clearance of structurally complex Ag relies on Ab evolution in a larger antigenic space than where selection directly acts; rebound highlights the impact of feedback between ecology and rapid evolution. In addition, immune compartmentalization can slow viral escape but may also delay clearance. This work illuminates that physical dynamics can optimize molecular binding, and, by creating a plastic phenotype, modulate evolutionary transitions of adapting populations. By facilitating the accumulation of potentiating variations, it is possible to shortcut long paths toward highly adapted states. Chapter 3 shows that during vaccination, a trade-off exists between suppressing strain-specific antibodies ("specialists") and preserving broadly-neutralizing antibodies ("generalists") within a continuous GC reaction. Time-varying selection can pursue both ends of the tradeoff se-

quentially in time, thus outperforms constant selection. In particular, the optimal selection force exhibits a signature low-to-high switch driven by the evolution of generalist breadth. In addition, in generalist-specialist competition, optimally suppressing specialists can promote both generalist quantity as well as quality, thus sparing the need for two distinct controls. On the contrary, in intraspecies competition among generalist subclasses, a trade-off exists between generalists' quantity and quality, thus requiring different optimal control profiles. Overall, this work highlights the importance of feedback between ecology and evolution during bnAbs development, and illustrates optimal elicitation of bnAbs by applying time-varying selection force. Vaccination protocols amenable to experimental testing are also proposed. Chapter 4 studies the aging of immune repertoire as it adapts to diverse pathogen species over host lifetime. Repeated pathogen encounters specialize the immune repertoire, which becomes robust against frequent pathogens but fragile against rare pathogens. Mediated by the collaboration between adaptive and innate immune branches, repertoire overspecialization eventually triggers a fragility where any encounter with a novel pathogen will cause the system to irreversibly switch from health to CI. This transition is consistent with the onset of "inflammaging", a condition observed in aged individuals who experience chronic low-grade inflammation even in the absence of pathogens. This work also predicts that the onset of CI strongly depends on the history of encountered pathogens. In particular, the timing of onset can be delayed drastically when the same set of infections is encountered in a clustered or cyclic order, so as to prolong the abundance of naive cells or quickly acquire full immune memory coverage across all pathogen species. Lastly, the coupling between the innate and adaptive immune branches generates a trade-off between rapid pathogen clearance and a delayed onset of immunosenescence. Together, this study suggests the evolution of immune repertoire as a potential mechanism for immunosenescence and provides a theoretical framework at the system level and on the scale of an organism's lifetime to account for clinical observations.

In conclusion, the research in the thesis examines the evolution of immune system against diverse pathogens, illustrates its mechanisms, limitations, and corresponding remedies through human intervention. Additional implications and potential future directions

are discussed below.

5.1 Additional Ag escape mechanisms and its effect on immune-Ag coevolution

In chapter 2, the red-queen dynamics between BCRs and highly-mutable pathogens is studied in detail. In reality, highly-mutable pathogens, particularly HIV, also employ additional immune evasion mechanisms besides rapid mutation. The effect of these additional mechanisms on immune-Ag coevolution is worth investigation for a more comprehensive understanding of immune system evolution, and in principle this can be achieved by expanding the coevolution model in chapter 2 to account for the additional immune evasion mechanisms.

To begin with, HIV targets and destroys CD4+ T cells, which are a crucial component of immune system and act as helpers to promote B cell Ab production. As a result, immune function of those infected with HIV declines over time as CD4+ T cell count shrinks, and eventually people are diagnosed with AIDS when their CD4+ T cell count drops below a critical level of 200 cells/mm^3. ([305]). If the viral damage on the immune function is explicitly modeled, a narrow yet fast Ab response is expected to be more beneficial in controlling the viral damage; this then better allows subsequent development of slow yet broad Ab response and clearance of Ag.

In addition, HIV can create latent reservoirs in CD4+ T cells where viral genomes stay dormant and no viral proteins are produced ([306, 307]). Such latent state can persist for months or even years ([308, 309, 310, 311]). As a result, viral load can rebound even after the actively-replicating viruses have been cleared. With the viral reservoir considered: (1) the broad Ab response will be harder to develop as Ag escape mutants can stay latent without providing a selection advantage to broad antibodies; (2) the resulting narrow Ab response is expected to be less efficient when faced simultaneously with current Ag and Ag variants of earlier time revived from a latent reservoir. Together the viral clearance is expected to be more challenging.

5.2 Degradation of immune cell efficiency with age and their effect on repertoire evolution

Chapter 4 demonstrated an increasingly skewed immune repertoire with age as the repertoire adapts to diverse pathogen species over host lifetime. This eventually culminated in immunosenescence, reflected by a transition from health to CI state even in the absence of pathogens. In additional to the accumulating history of pathogen encounters, several other immune functions vary with age and could play a role in the development of imbalanced immunological space as well as the emergence of immunosenescence. For example, clinical studies observed that cell signaling between immune cells can become impaired with age ([312]), and thymus involution leads to decreased T cell production with age ([299, 298]). The current formulation of the IIB model in chapter 4 exhibits a skewed repertoire and an inflammaging-like behavior without taking these additional factors into account. However, in future work physiological parameters of the model could be used as a proxy for these observed behaviors: for example, the impaired cell signaling in T cells could be achieved by decreasing γ with age, or the reduced thymus output could be modeled by decreasing θ_N with age. The calculated immune outcomes that result from these modifications could shed light on their effect on repertoire evolution (for example, a reduced production of naive cells is expected to accelerate the repertoire overspecialization), and also separate the relative contributions to immunosenescence from memory-induced fragility in the adaptive response, as opposed to from an impaired immune cell efficiency.

5.3 Evolution of immune system structure against diverse pathogens.

In foregoing chapters, the longest timescale studied is the host lifetime, during which the adaptive immune repertoire evolves. On an even longer timescale spanning multiple host generations, the structure of immune system itself also evolves in adaptation to the host species' diverse pathogenic environment. Therefore, insights on the evolutionary origin of immune system structures can be gained by studying their functional benefits ([313, 314]),

as the immune system structure has constantly been optimized by selection over deep-time evolution.

By applying this methodology, the evolutionary origin of B cell compartmentalization can be speculated from chapter 2. It was shown in chapter 2 that spatial separation of B cells into parallel GCs have a two-fold effect on coevolutionary outcome: it slows down viral escape if Ag epitope is highly variable, or delays clearance if the Ag epitope is relatively conserved. When both types of Ag are present in the host species' environment, a tradeoff then emerges on the optimal number of GCs during pathogen encounter: a few large GCs speed up clearance of pathogens with relatively conserved epitope, while many small GCs allow efficient suppression of pathogens with highly variable epitope. The number of parallel GCs in reality should achieve a balance so that the host remain fit against both types of pathogens. Similarly, chapter 4 provides a potential explanation to the functional division between cytotoxic and helper T cells vs regulatory T cells. While diverting more T cells into the former group facilitates faster pathogen clearance, maintaining more regulatory T cells help to rein in the excess innate immune response and delay the inflammaging transition. Over deep-time scale, these two opposing selection forces should lead to a balanced distribution where T cells allocate resources both to pathogen clearance and innate suppression.

On a further note, the emergence of adaptive immune system itself is an outcome of evolution, which occurred approximately 500 million years ago in jawed fish ([315, 316]). Evolutionarily the innate immune response preceded the creation of the adaptive response ([317]) and the adaptive immune components are dependent on innate cells- for example, the activation of an adaptive response through Ag presentation relies on dendritic cells. The evolutionary drivers of the adaptive immune response could thus be explored with immune models that quantify the added benefit of possessing an adaptive immune system. The comparison could be made by, for example, expanding on the IIB model in chapter 4 and comparing the quantitative cost and gain associated with maintaining adaptive immune cells against diverse pathogens.

REFERENCES

[1] Gabriel D Victora and Michel C Nussenzweig. Germinal centers. *Annual review of immunology*, 30:429–457, 2012.

[2] Javier M Di Noia and Michael S Neuberger. Molecular mechanisms of antibody somatic hypermutation. *Annu. Rev. Biochem.*, 76:1–22, 2007.

[3] Lauren M Sompayrac. *How the immune system works*. John Wiley & Sons, 2019.

[4] Terry McNearney, Zuzana Hornickova, Richard Markham, Anahid Birdwell, Max Arens, Alfred Saah, and Lee Ratner. Relationship of human immunodeficiency virus type 1 sequence heterogeneity to stage of disease. *Proceedings of the National Academy of Sciences*, 89(21):10247–10251, 1992.

[5] Joseph C Forbi, David S Campo, Michael A Purdy, Zoya E Dimitrova, Pavel Skums, Guo-liang Xia, Lili T Punkova, Lilia M Ganova-Raeva, Gilberto Vaughan, Yousr Ben-Ayed, et al. Intra-host diversity and evolution of hepatitis C virus endemic to côte d'ivoire. *Journal of medical virology*, 86(5):765–771, 2014.

[6] Jean Claude Ameisen, Jean-Daniel Lelièvre, and Olivier Pleskoff. HIV/host interactions: new lessons from the red queen's country. *Aids*, 16:S25–S31, 2002.

[7] Hua-Xin Liao, Rebecca Lynch, Tongqing Zhou, Feng Gao, S Munir Alam, Scott D Boyd, Andrew Z Fire, Krishna M Roskin, Chaim A Schramm, Zhenhai Zhang, et al. Co-evolution of a broadly neutralizing HIV-1 antibody and founder virus. *Nature*, 496(7446):469–476, 2013.

[8] Jin Zhong, Pablo Gastaminza, Josan Chung, Zania Stamataki, Masanori Isogawa, Guofeng Cheng, Jane A McKeating, and Francis V Chisari. Persistent hepatitis C virus infection in vitro: coevolution of virus and host. *Journal of virology*, 80(22):11082–11093, 2006.

[9] David L Thomas, Chloe L Thio, Maureen P Martin, Ying Qi, Dongliang Ge, Colm O'hUigin, Judith Kidd, Kenneth Kidd, Salim I Khakoo, Graeme Alexander, et al. Genetic variation in il28b and spontaneous clearance of hepatitis C virus. *Nature*, 461(7265):798–801, 2009.

[10] Devin Sok, Brian Moldt, and Dennis R Burton. Snapshot: broadly neutralizing antibodies. *Cell*, 155(3):728–728, 2013.

[11] Dennis R Burton, Pascal Poignard, Robyn L Stanfield, and Ian A Wilson. Broadly neutralizing antibodies present new prospects to counter highly antigenically diverse viruses. *Science*, 337(6091):183–186, 2012.

[12] Valerie J Kinchen, Andrea L Cox, and Justin R Bailey. Can broadly neutralizing monoclonal antibodies lead to a hepatitis C virus vaccine? *Trends in microbiology*, 26(10):854–864, 2018.

[13] Rebecca M Lynch, Patrick Wong, Lillian Tran, Sijy O'Dell, Martha C Nason, Yuxing Li, Xueling Wu, and John R Mascola. HIV-1 fitness cost associated with escape from the VRC01 class of CD4 binding site neutralizing antibodies. *Journal of virology*, 89(8):4201–4213, 2015.

[14] Valerie J Kinchen, Muhammad N Zahid, Andrew I Flyak, Mary G Soliman, Gerald H Learn, Shuyi Wang, Edgar Davidson, Benjamin J Doranz, Stuart C Ray, Andrea L Cox, et al. Broadly neutralizing antibody mediated clearance of human hepatitis C virus infection. *Cell host & microbe*, 24(5):717–730, 2018.

[15] Anne Piantadosi, Dana Panteleeff, Catherine A Blish, Jared M Baeten, Walter Jaoko, R Scott McClelland, and Julie Overbaugh. Breadth of neutralizing antibody response to human immunodeficiency virus type 1 is affected by factors early in infection but does not influence disease progression. *Journal of virology*, 83(19):10269–10274, 2009.

[16] Elin S Gray, Maphuti C Madiga, Tandile Hermanus, Penny L Moore, Constantinos Kurt Wibmer, Nancy L Tumba, Lise Werner, Koleka Mlisana, Sengeziwe Sibeko, Carolyn Williamson, et al. The neutralization breadth of HIV-1 develops incrementally over four years and is associated with CD4+ T cell decline and high viral load during acute infection. *Journal of virology*, 85(10):4828–4840, 2011.

[17] Alexander P Underwood, Melanie R Walker, Nicholas A Brasher, Auda A Eltahla, Lisa Maher, Fabio Luciani, Andrew R Lloyd, and Rowena A Bull. Understanding the determinants of bnab induction in acute HCV infection. *Viruses*, 10(11):659, 2018.

[18] Justin R Bailey, Andrew I Flyak, Valerie J Cohen, Hui Li, Lisa N Wasilewski, Anna E Snider, Shuyi Wang, Gerald H Learn, Nurgun Kose, Leah Loerinc, et al. Broadly neutralizing antibodies with few somatic mutations and hepatitis C virus clearance. *JCI insight*, 2(9), 2017.

[19] Peter D Kwong, John R Mascola, and Gary J Nabel. Broadly neutralizing antibodies and the search for an HIV-1 vaccine: the end of the beginning. *Nature Reviews Immunology*, 13(9):693–701, 2013.

[20] Barton F Haynes and Dennis R Burton. Developing an HIV vaccine. *Science*, 355(6330):1129–1130, 2017.

[21] Tiza Ng'uni, Caroline Chasara, and Zaza M Ndhlovu. Major scientific hurdles in HIV vaccine development: Historical perspective and future directions. *Frontiers in Immunology*, 11:2761, 2020.

[22] Richard Wyatt and Joseph Sodroski. The HIV-1 envelope glycoproteins: fusogens, antigens, and immunogens. *Science*, 280(5371):1884–1888, 1998.

[23] Leopold Kong and Quentin J Sattentau. Antigenicity and immunogenicity in HIV-1 antibody-based vaccine design. *Journal of AIDS & clinical research*, page 003, 2012.

[24] Shishi Luo and Alan S Perelson. Competitive exclusion by autologous antibodies can prevent broad HIV-1 antibodies from arising. *Proceedings of the National Academy of Sciences*, 112(37):11654–11659, 2015.

[25] Justin Pollara, David Easterhoff, and Genevieve G Fouda. Lessons learned from human HIV vaccine trials. *Current opinion in HIV and AIDS*, 12(3):216, 2017.

[26] Supachai Rerks-Ngarm, Punnee Pitisuttithum, Sorachai Nitayaphan, Jaranit Kaewkungwal, Joseph Chiu, Robert Paris, Nakorn Premsri, Chawetsan Namwat, Mark de Souza, Elizabeth Adams, et al. Vaccination with ALVAC and AIDSVAX to prevent HIV-1 infection in thailand. *New England Journal of Medicine*, 361(23):2209–2220, 2009.

[27] Stephen C Jameson. Maintaining the norm: T-cell homeostasis. *Nature Reviews Immunology*, 2(8):547–556, 2002.

[28] Afonso RM Almeida, Benedita Rocha, Antonio A Freitas, and Corine Tanchot. Homeostasis of T cell numbers: from thymus production to peripheral compartmentalization and the indexation of regulatory T cells. In *Seminars in immunology*, volume 17, pages 239–249. Elsevier, 2005.

[29] Pasquine Saule, Jacques Trauet, Virginie Dutriez, Véronique Lekeux, Jean-Paul Dessaint, and Myriam Labalette. Accumulation of memory T cells from childhood to old age: central and effector memory cells in CD4+ versus effector memory and terminally differentiated memory cells in CD8+ compartment. *Mechanisms of ageing and development*, 127(3):274–281, 2006.

[30] Sean P Stromberg and Jean Carlson. Robustness and fragility in immunosenescence. *PLoS computational biology*, 2(11):e160, 2006.

[31] Claudio Franceschi, Massimiliano Bonafè, Silvana Valensin, Fabiola Olivieri, Maria De Luca, Enzo Ottaviani, and Giovanna De Benedictis. Inflamm-aging: an evolutionary perspective on immunosenescence. *Annals of the new York Academy of Sciences*, 908(1):244–254, 2000.

[32] Anders Wikby, Bengt-Olof Nilsson, Rosalyn Forsey, Julie Thompson, Jan Strindhall, Sture Löfgren, Jan Ernerudh, Graham Pawelec, Frederick Ferguson, and Boo Johansson. The immune risk phenotype is associated with il-6 in the terminal decline stage: findings from the swedish nona immune longitudinal study of very late life functioning. *Mechanisms of ageing and development*, 127(8):695–704, 2006.

[33] Frederick G Ferguson, Anders Wikby, Pamela Maxson, Jadwiga Olsson, and Boo Johansson. Immune parameters in a longitudinal study of a very old population of swedish people: a comparison between survivors and nonsurvivors. *The Journals of Gerontology Series A: Biological Sciences and Medical Sciences*, 50(6):B378–B382, 1995.

[34] Anders Wikby, Pamela Maxson, Jadwiga Olsson, Boo Johansson, and Frederick G Ferguson. Changes in CD8 and CD4 lymphocyte subsets, T cell proliferation responses and non-survival in the very old: the swedish longitudinal octo-immune study. *Mechanisms of ageing and development*, 102(2-3):187–198, 1998.

[35] Hillary Klonoff-Cohen, Elizabeth L Barrett-Connor, and Sharon L Edelstein. Albumin levels as a predictor of mortality in the healthy elderly. *Journal of clinical epidemiology*, 45(3):207–212, 1992.

[36] Matty P Weijenberg, Edith JM Feskens, John HM Souverijn, and Daan Kromhout. Serum albumin, coronary heart disease risk, and mortality in an elderly cohort. *Epidemiology*, pages 87–92, 1997.

[37] Alan S Perelson and Gerard Weisbuch. Immunology for physicists. *Reviews of modern physics*, 69(4):1219, 1997.

[38] Arup K Chakraborty and Andrej Košmrlj. Statistical mechanical concepts in immunology. *Annual review of physical chemistry*, 61:283–303, 2010.

[39] Mario Castro. A physicist's approach to immunology. In *Mathematical Models and Immune Cell Biology*, pages 339–350. Springer, 2011.

[40] Yoram Vodovotz, Ashley Xia, Elizabeth L Read, Josep Bassaganya-Riera, David A Hafler, Eduardo Sontag, Jin Wang, John S Tsang, Judy D Day, Steven H Kleinstein, et al. Solving immunology? *Trends in immunology*, 38(2):116–127, 2017.

[41] Kristen Jill Kresge. Special report: Thai trial results. `https://www.vaxreport.org/articles/special-report-thai-trial-results`. Accessed: 2021-04-28.

[42] Jon Cohen. Another HIV vaccine strategy fails in large-scale study. `https://www.sciencemag.org/news/2020/02/another-hiv-vaccine-strategy-fails-large-scale-study`. Accessed: 2021-04-28.

[43] NADIA WHITEHEAD. A $100 million HIV vaccine project failed. but all hope is not lost. `https://www.npr.org/sections/goatsandsoda/2020/03/02/809729417/a-100-million-hiv-vaccine-project-failed-but-all-hope-is-not-lost`. Accessed: 2021-04-28.

[44] David D Ho, Avidan U Neumann, Alan S Perelson, Wen Chen, John M Leonard, and Martin Markowitz. Rapid turnover of plasma virions and CD4 lymphocytes in HIV-1 infection. *Nature*, 373(6510):123–126, 1995.

[45] Thomas B Kepler and Alan S Perelson. Somatic hypermutation in B cells: an optimal control treatment. *Journal of theoretical biology*, 164(1):37–64, 1993.

[46] Derek J Smith, Stephanie Forrest, David H Ackley, and Alan S Perelson. Variable efficacy of repeated annual influenza vaccination. *Proceedings of the National Academy of Sciences*, 96(24):14001–14006, 1999.

[47] Armita Nourmohammad, Jakub Otwinowski, and Joshua B Plotkin. Host-pathogen coevolution and the emergence of broadly neutralizing antibodies in chronic infections. *PLoS genetics*, 12(7), 2016.

[48] Ronald A Fisher. Xxi.-on the dominance ratio. *Proceedings of the royal society of Edinburgh*, 42:321–341, 1923.

[49] Sewall Wright. Evolution in mendelian populations. *Genetics*, 16(2):97, 1931.

[50] Patrick Alfred Pierce Moran. Random processes in genetics. In *Mathematical proceedings of the cambridge philosophical society*, volume 54, pages 60–71. Cambridge University Press, 1958.

[51] Thomas Nagylaki. *Introduction to theoretical population genetics*. Springer Berlin Heidelberg, 1992.

[52] John H Gillespie. *Population genetics: a concise guide*. JHU Press, 1998.

[53] Warren J Ewens. *Mathematical population genetics 1: theoretical introduction*, volume 27. Springer Science & Business Media, 2012.

[54] Mattia Bonsignori, Tongqing Zhou, Zizhang Sheng, Lei Chen, Feng Gao, M Gordon Joyce, Gabriel Ozorowski, Gwo-Yu Chuang, Chaim A Schramm, Kevin Wiehe, et al. Maturation pathway from germline to broad HIV-1 neutralizer of a CD4-mimic antibody. *Cell*, 165(2):449–463, 2016.

[55] Kimia Kardani, Azam Bolhassani, and Sepideh Shahbazi. Prime-boost vaccine strategy against viral infections: mechanisms and benefits. *Vaccine*, 34(4):413–423, 2016.

[56] RV144 vaccine trial. `https://www.avac.org/trial/rv-144`. Accessed: 2021-01-12.

[57] HVTN 505 vaccine trial. `https://www.avac.org/hvtn-505`. Accessed: 2021-01-12.

[58] HVTN 702 vaccine trial. `https://www.avac.org/trial/hvtn-702`. Accessed: 2021-01-12.

[59] J Scott Shaffer, Penny L Moore, Mehran Kardar, and Arup K Chakraborty. Optimal immunization cocktails can promote induction of broadly neutralizing abs against highly mutable pathogens. *Proceedings of the National Academy of Sciences*, 113(45):E7039–E7048, 2016.

[60] Kayla G Sprenger, Joy E Louveau, Pranav M Murugan, and Arup K Chakraborty. Optimizing immunization protocols to elicit broadly neutralizing antibodies. *Proceedings of the National Academy of Sciences*, 117(33):20077–20087, 2020.

[61] Tongqing Zhou, Rebecca M Lynch, Lei Chen, Priyamvada Acharya, Xueling Wu, Nicole A Doria-Rose, M Gordon Joyce, Daniel Lingwood, Cinque Soto, Robert T Bailer, et al. Structural repertoire of HIV-1-neutralizing antibodies targeting the CD4 supersite in 14 donors. *Cell*, 161(6):1280–1292, 2015.

[62] Mattia Bonsignori, Tongqing Zhou, Zizhang Sheng, Lei Chen, Feng Gao, M Gordon Joyce, Gabriel Ozorowski, Gwo-Yu Chuang, Chaim A Schramm, Kevin Wiehe, et al. Maturation pathway from germline to broad HIV-1 neutralizer of a CD4-mimic antibody. *Cell*, 165(2):449–463, 2016.

[63] Hok Hei Tam, Mariane B Melo, Myungsun Kang, Jeisa M Pelet, Vera M Ruda, Maria H Foley, Joyce K Hu, Sudha Kumari, Jordan Crampton, Alexis D Baldeon, et al. Sustained antigen availability during germinal center initiation enhances antibody responses to vaccination. *Proceedings of the National Academy of Sciences*, 113(43):E6639–E6648, 2016.

[64] Kimberly M Cirelli, Diane G Carnathan, Bartek Nogal, Jacob T Martin, Oscar L Rodriguez, Amit A Upadhyay, Chiamaka A Enemuo, Etse H Gebru, Yury Choe, Federico Viviano, et al. Slow delivery immunization enhances HIV neutralizing antibody and germinal center responses via modulation of immunodominance. *Cell*, 177(5):1153–1171, 2019.

[65] Angela Reynolds, Jonathan Rubin, Gilles Clermont, Judy Day, Yoram Vodovotz, and G Bard Ermentrout. A reduced mathematical model of the acute inflammatory response: I. derivation of model and analysis of anti-inflammation. *Journal of theoretical biology*, 242(1):220–236, 2006.

[66] Zhenzhen Shi, Chih-Hang J Wu, David Ben-Arieh, and Steven Q Simpson. Mathematical model of innate and adaptive immunity of sepsis: a modeling and simulation study of infectious disease. *BioMed research international*, 2015, 2015.

[67] Kevin N Couper, Daniel G Blount, and Eleanor M Riley. Il-10: the master regulator of immunity to infection. *The Journal of Immunology*, 180(9):5771–5777, 2008.

[68] Greta Guarda, Catherine Dostert, Francesco Staehli, Katrin Cabalzar, Rosa Castillo, Aubry Tardivel, Pascal Schneider, and Jürg Tschopp. T cells dampen innate immune responses through inhibition of nlrp1 and nlrp3 inflammasomes. *Nature*, 460(7252):269–273, 2009.

[69] Charles D Murin, Ian A Wilson, and Andrew B Ward. Antibody responses to viral infections: a structural perspective across three different enveloped viruses. *Nature microbiology*, 4(5):734–747, 2019.

[70] Maria Martell, Juan I Esteban, Josep Quer, Joan Genesca, Amy Weiner, R Esteban, Jaume Guardia, and Jordi Gomez. Hepatitis C virus (HCV) circulates as a population of different but closely related genomes: quasispecies nature of HCV genome distribution. *Journal of virology*, 66(5):3225–3229, 1992.

[71] P Simmonds, EC Holmes, T-A Cha, S-W Chan, F McOmish, B Irvine, E Beall, PL Yap, J Kolberg, and MS Urdea. Classification of hepatitis C virus into six major genotypes and a series of subtypes by phylogenetic analysis of the NS-5 region. *Journal of general virology*, 74(11):2391–2399, 1993.

[72] Oliver G Pybus, Michael A Charleston, Sunetra Gupta, Andrew Rambaut, Edward C Holmes, and Paul H Harvey. The epidemic behavior of the hepatitis C virus. *Science*, 292(5525):2323–2325, 2001.

[73] Yuan-Ding Chen, Ming-Ying Liu, Wen-Lin Yu, Jia-Qi Li, Qin Dai, Zhen-Quan Zhou, and Sergio G Tisminetzky. Mix-infections with different genotypes of HCV and with HCV plus other hepatitis viruses in patients with hepatitis C in China. *World Journal of Gastroenterology: WJG*, 9(5):984, 2003.

[74] Peter Simmonds. Genetic diversity and evolution of hepatitis C virus–15 years on. *Journal of General Virology*, 85(11):3173–3188, 2004.

[75] Ali Teimoori, Saeedeh Ebrahimi, Narges Keshtkar, Soheila Khaghani, Shokrollah Salmanzadeh, and Shokouh Ghafari. Prevalence and genetic diversity of HCV among HIV-1 infected individuals living in Ahvaz, Iran. *BMC infectious diseases*, 19(1):1–7, 2019.

[76] Dennis R Burton and Lars Hangartner. Broadly neutralizing antibodies to HIV and their role in vaccine design. *Annual review of immunology*, 34:635–659, 2016.

[77] Andrew I Flyak, Stormy Ruiz, Michelle D Colbert, Tiffany Luong, James E Crowe Jr, Justin R Bailey, and Pamela J Bjorkman. HCV broadly neutralizing antibodies use a CDRH3 disulfide motif to recognize an E2 glycoprotein site that can be targeted for vaccine design. *Cell host & microbe*, 24(5):703–716, 2018.

[78] Anne Piantadosi, Dana Panteleeff, Catherine A Blish, Jared M Baeten, Walter Jaoko, R Scott McClelland, and Julie Overbaugh. Breadth of neutralizing antibody response to human immunodeficiency virus type 1 is affected by factors early in infection but does not influence disease progression. *Journal of virology*, 83(19):10269–10274, 2009.

[79] Elin S Gray, Maphuti C Madiga, Tandile Hermanus, Penny L Moore, Constantinos Kurt Wibmer, Nancy L Tumba, Lise Werner, Koleka Mlisana, Sengeziwe Sibeko, Carolyn Williamson, et al. The neutralization breadth of HIV-1 develops incrementally over four years and is associated with CD4+ T cell decline and high viral load during acute infection. *Journal of virology*, 85(10):4828–4840, 2011.

[80] Justin R Bailey, Andrew I Flyak, Valerie J Cohen, Hui Li, Lisa N Wasilewski, Anna E Snider, Shuyi Wang, Gerald H Learn, Nurgun Kose, Leah Loerinc, et al. Broadly neutralizing antibodies with few somatic mutations and hepatitis C virus clearance. *JCI insight*, 2(9), 2017.

[81] Valerie J Kinchen, Muhammad N Zahid, Andrew I Flyak, Mary G Soliman, Gerald H Learn, Shuyi Wang, Edgar Davidson, Benjamin J Doranz, Stuart C Ray, Andrea L Cox, et al. Broadly neutralizing antibody mediated clearance of human hepatitis C virus infection. *Cell host & microbe*, 24(5):717–730, 2018.

[82] Herman N Eisen and Gregory W Siskind. Variations in affinities of antibodies during the immune response. *Biochemistry*, 3(7):996–1008, 1964.

[83] Gabriel D Victora and Michel C Nussenzweig. Germinal centers. *Annual review of immunology*, 30:429–457, 2012.

[84] Hua-Xin Liao, Rebecca Lynch, Tongqing Zhou, Feng Gao, S Munir Alam, Scott D Boyd, Andrew Z Fire, Krishna M Roskin, Chaim A Schramm, Zhenhai Zhang, et al. Co-evolution of a broadly neutralizing HIV-1 antibody and founder virus. *Nature*, 496(7446):469, 2013.

[85] Constantinos Kurt Wibmer, Jinal N Bhiman, Elin S Gray, Nancy Tumba, Salim S Abdool Karim, Carolyn Williamson, Lynn Morris, and Penny L Moore. Viral escape from HIV-1 neutralizing antibodies drives increased plasma neutralization breadth through sequential recognition of multiple epitopes and immunotypes. *PLoS Pathog*, 9(10):e1003738, 2013.

[86] Jinal N Bhiman, Colin Anthony, Nicole A Doria-Rose, Owen Karimanzira, Chaim A Schramm, Thandeka Khoza, Dale Kitchin, Gordon Botha, Jason Gorman, Nigel J Garrett, et al. Viral variants that initiate and drive maturation of V1V2-directed HIV-1 broadly neutralizing antibodies. *Nature medicine*, 21(11):1332, 2015.

[87] Fabio Zanini, Johanna Brodin, Lina Thebo, Christa Lanz, Göran Bratt, Jan Albert, and Richard A Neher. Population genomics of intrapatient HIV-1 evolution. *Elife*, 4:e11282, 2015.

[88] James Hadfield, Colin Megill, Sidney M Bell, John Huddleston, Barney Potter, Charlton Callender, Pavel Sagulenko, Trevor Bedford, and Richard A Neher. Nextstrain: real-time tracking of pathogen evolution. *Bioinformatics*, 34(23):4121–4123, 2018.

[89] Feng Gao, Mattia Bonsignori, Hua-Xin Liao, Amit Kumar, Shi-Mao Xia, Xiaozhi Lu, Fangping Cai, Kwan-Ki Hwang, Hongshuo Song, Tongqing Zhou, et al. Cooperation of B cell lineages in induction of HIV-1-broadly neutralizing antibodies. *Cell*, 158(3):481–491, 2014.

[90] Amelia Escolano, Jon M Steichen, Pia Dosenovic, Daniel W Kulp, Jovana Golijanin, Devin Sok, Natalia T Freund, Alexander D Gitlin, Thiago Oliveira, Tatsuya Araki, et al. Sequential immunization elicits broadly neutralizing anti-HIV-1 antibodies in ig knockin mice. *Cell*, 166(6):1445–1458, 2016.

[91] Colin Anthony, Talita York, Valerie Bekker, David Matten, Philippe Selhorst, Roux-Cil Ferreria, Nigel J Garrett, Salim S Abdool Karim, Lynn Morris, Natasha T Wood, et al. Cooperation between strain-specific and broadly neutralizing responses limited viral escape and prolonged the exposure of the broadly neutralizing epitope. *Journal of virology*, 91(18), 2017.

[92] Jeroen MJ Tas, Luka Mesin, Giulia Pasqual, Sasha Targ, Johanne T Jacobsen, Yasuko M Mano, Casie S Chen, Jean-Claude Weill, Claude-Agnès Reynaud, Edward P Browne, et al. Visualizing antibody affinity maturation in germinal centers. *Science*, 351(6277):1048–1054, 2016.

[93] Anthony P West Jr, Louise Scharf, Johannes F Scheid, Florian Klein, Pamela J Bjorkman, and Michel C Nussenzweig. Structural insights on the role of antibodies in HIV-1 vaccine and therapy. *Cell*, 156(4):633–648, 2014.

[94] Haoqing Wang, Alexander A Cohen, Rachel P Galimidi, Harry B Gristick, Grant J Jensen, and Pamela J Bjorkman. Cryo-EM structure of a CD4-bound open HIV-1 envelope trimer reveals structural rearrangements of the gp120 V1V2 loop. *Proceedings of the National Academy of Sciences*, 113(46):E7151–E7158, 2016.

[95] Robert K Abbott, Jeong Hyun Lee, Sergey Menis, Patrick Skog, Meghan Rossi, Takayuki Ota, Daniel W Kulp, Deepika Bhullar, Oleksandr Kalyuzhniy, Colin Havenar-Daughton, et al. Precursor frequency and affinity determine B cell competitive fitness in germinal centers, tested with germline-targeting HIV vaccine immunogens. *Immunity*, 48(1):133–146, 2018.

[96] Jean-Philippe Julien, Albert Cupo, Devin Sok, Robyn L Stanfield, Dmitry Lyumkis, Marc C Deller, Per-Johan Klasse, Dennis R Burton, Rogier W Sanders, John P Moore, et al. Crystal structure of a soluble cleaved HIV-1 envelope trimer. *Science*, 342(6165):1477–1483, 2013.

[97] Raymond HY Louie, Kevin J Kaczorowski, John P Barton, Arup K Chakraborty, and Matthew R McKay. Fitness landscape of the human immunodeficiency virus envelope protein that is targeted by antibodies. *Proceedings of the National Academy of Sciences*, 115(4):E564–E573, 2018.

[98] Florian Klein, Ron Diskin, Johannes F Scheid, Christian Gaebler, Hugo Mouquet, Ivelin S Georgiev, Marie Pancera, Tongqing Zhou, Reha-Baris Incesu, Brooks Zhongzheng Fu, et al. Somatic mutations of the immunoglobulin framework are generally required for broad and potent HIV-1 neutralization. *Cell*, 153(1):126–138, 2013.

[99] Leslie Goo, Vrasha Chohan, Ruth Nduati, and Julie Overbaugh. Early development of broadly neutralizing antibodies in HIV-1–infected infants. *Nature medicine*, 20(6):655–658, 2014.

[100] Cassandra A Simonich, Katherine L Williams, Hans P Verkerke, James A Williams, Ruth Nduati, Kelly K Lee, and Julie Overbaugh. HIV-1 neutralizing antibodies with limited hypermutation from an infant. *Cell*, 166(1):77–87, 2016.

[101] Thomas B Kepler and Alan S Perelson. Cyclic re-entry of germinal center B cells and the efficiency of affinity maturation. *Immunology today*, 14(8):412–415, 1993.

[102] Mihaela Oprea and Alan S Perelson. Somatic mutation leads to efficient affinity maturation when centrocytes recycle back to centroblasts. *The Journal of Immunology*, 158(11):5155–5162, 1997.

[103] Michael Meyer-Hermann, Andreas Deutsch, and Michal Or-Guil. Recycling probability and dynamical properties of germinal center reactions. *Journal of Theoretical Biology*, 210(3):265–285, 2001.

[104] Marc Thilo Figge, Alexandre Garin, Matthias Gunzer, Marie Kosco-Vilbois, Kai-Michael Toellner, and Michael Meyer-Hermann. Deriving a germinal center lymphocyte migration model from two-photon data. *The Journal of experimental medicine*, 205(13):3019–3029, 2008.

[105] Jingshan Zhang and Eugene I Shakhnovich. Optimality of mutation and selection in germinal centers. *PLoS computational biology*, 6(6):e1000800, 2010.

[106] Assaf Amitai, Luka Mesin, Gabriel D Victora, Mehran Kardar, and Arup K Chakraborty. A population dynamics model for clonal diversity in a germinal center. *Frontiers in microbiology*, 8:1693, 2017.

[107] Marco Molari, Klaus Eyer, Jean Baudry, Simona Cocco, and Remi Monasson. Quantitative modeling of the effect of antigen dosage on b-cell affinity distributions in maturating germinal centers. *bioRxiv*, 2020.

[108] Sidhartha Chaudhury, Jaques Reifman, and Anders Wallqvist. Simulation of B cell affinity maturation explains enhanced antibody cross-reactivity induced by the polyvalent malaria vaccine AMA1. *The Journal of Immunology*, 193(5):2073–2086, 2014.

[109] Shenshen Wang, Jordi Mata-Fink, Barry Kriegsman, Melissa Hanson, Darrell J Irvine, Herman N Eisen, Dennis R Burton, K Dane Wittrup, Mehran Kardar, and Arup K Chakraborty. Manipulating the selection forces during affinity maturation to generate cross-reactive HIV antibodies. *Cell*, 160(4):785–797, 2015.

[110] Lauren M Childs, Edward B Baskerville, and Sarah Cobey. Trade-offs in antibody repertoires to complex antigens. *Philosophical Transactions of the Royal Society B: Biological Sciences*, 370(1676):20140245, 2015.

[111] J Scott Shaffer, Penny L Moore, Mehran Kardar, and Arup K Chakraborty. Optimal immunization cocktails can promote induction of broadly neutralizing abs against highly mutable pathogens. *Proceedings of the National Academy of Sciences*, 113(45):E7039–E7048, 2016.

[112] Shenshen Wang. Optimal sequential immunization can focus antibody responses against diversity loss and distraction. *PLoS computational biology*, 13(1):e1005336, 2017.

[113] Kayla G Sprenger, Joy E Louveau, Pranav M Murugan, and Arup K Chakraborty. Optimizing immunization protocols to elicit broadly neutralizing antibodies. *Proceedings of the National Academy of Sciences*, 117(33):20077–20087, 2020.

[114] Christel Kamp and Stefan Bornholdt. Coevolution of quasispecies: B-cell mutation rates maximize viral error catastrophes. *Physical Review Letters*, 88(6):068104, 2002.

[115] Shishi Luo and Alan S Perelson. Competitive exclusion by autologous antibodies can prevent broad HIV-1 antibodies from arising. *Proceedings of the National Academy of Sciences*, 112(37):11654–11659, 2015.

[116] Sarah Cobey, Patrick Wilson, and Frederick A Matsen IV. The evolution within us, 2015.

[117] Serena Bradde, Marija Vucelja, Tiberiu Teşileanu, and Vijay Balasubramanian. Dynamics of adaptive immunity against phage in bacterial populations. *PLoS computational biology*, 13(4):e1005486, 2017.

[118] Madeleine Bonsma-Fisher, Dominique Soutière, and Sidhartha Goyal. How adaptive immunity constrains the composition and fate of large bacterial populations. *Proceedings of the National Academy of Sciences*, 115(32):E7462–E7468, 2018.

[119] Jacopo Marchi, Michael Lässig, Thierry Mora, and Aleksandra M Walczak. Multilineage evolution in viral populations driven by host immune systems. *Pathogens*, 8(3):115, 2019.

[120] Hongda Jiang and Shenshen Wang. Trait-space patterning and the role of feedback in antigen-immunity coevolution. *Physical Review Research*, 1(3):033164, 2019.

[121] Tongqing Zhou, Rebecca M Lynch, Lei Chen, Priyamvada Acharya, Xueling Wu, Nicole A Doria-Rose, M Gordon Joyce, Daniel Lingwood, Cinque Soto, Robert T Bailer, et al. Structural repertoire of HIV-1-neutralizing antibodies targeting the CD4 supersite in 14 donors. *Cell*, 161(6):1280–1292, 2015.

[122] Claudia Berek and Cesar Milstein. Mutation drift and repertoire shift in the maturation of the immune response. *Immunological reviews*, 96(1):23–41, 1987.

[123] Christopher DC Allen, Takaharu Okada, H Lucy Tang, and Jason G Cyster. Imaging of germinal center selection events during affinity maturation. *Science*, 315(5811):528–531, 2007.

[124] Gabriel D Victora, Tanja A Schwickert, David R Fooksman, Alice O Kamphorst, Michael Meyer-Hermann, Michael L Dustin, and Michel C Nussenzweig. Germinal center dynamics revealed by multiphoton microscopy with a photoactivatable fluorescent reporter. *Cell*, 143(4):592–605, 2010.

[125] João Pedro Pereira, Lisa M Kelly, and Jason G Cyster. Finding the right niche: B-cell migration in the early phases of t-dependent antibody responses. *International immunology*, 22(6):413–419, 2010.

[126] Richard J Bende, Febe van Maldegem, Martijn Triesscheijn, Thera AM Wormhoudt, Richard Guijt, and Carel JM van Noesel. Germinal centers in human lymph nodes contain reactivated memory B cells. *Journal of Experimental Medicine*, 204(11):2655–2665, 2007.

[127] Alan S Perelson and George F Oster. Theoretical studies of clonal selection: minimal antibody repertoire size and reliability of self-non-self discrimination. *Journal of theoretical biology*, 81(4):645–670, 1979.

[128] Dhruv K Sethi, Anupriya Agarwal, Venkatasamy Manivel, Kanury VS Rao, and Dinakar M Salunke. Differential epitope positioning within the germline antibody paratope enhances promiscuity in the primary immune response. *Immunity*, 24(4):429–438, 2006.

[129] Jinghe Huang, Byong H Kang, Elise Ishida, Tongqing Zhou, Trevor Griesman, Zizhang Sheng, Fan Wu, Nicole A Doria-Rose, Baoshan Zhang, Krisha McKee, et al. Identification of a CD4-binding-site antibody to HIV that evolved near-pan neutralization breadth. *Immunity*, 45(5):1108–1121, 2016.

[130] Riccardo Baron and J Andrew McCammon. Molecular recognition and ligand association. *Annual review of physical chemistry*, 64:151–175, 2013.

[131] Derek J Smith, Stephanie Forrest, Ron R Hightower, and Alan S Perelson. Deriving shape space parameters from immunological data. *Journal of theoretical biology*, 189(2):141–150, 1997.

[132] Penny L Moore, Elin S Gray, C Kurt Wibmer, Jinal N Bhiman, Molati Nonyane, Daniel J Sheward, Tandile Hermanus, Shringkhala Bajimaya, Nancy L Tumba, Melissa-Rose Abrahams, et al. Evolution of an HIV glycan–dependent broadly neutralizing antibody epitope through immune escape. *Nature medicine*, 18(11):1688–1692, 2012.

[133] Mattia Bonsignori, Edward F Kreider, Daniela Fera, R Ryan Meyerhoff, Todd Bradley, Kevin Wiehe, S Munir Alam, Baptiste Aussedat, William E Walkowicz, Kwan-Ki Hwang, et al. Staged induction of HIV-1 glycan–dependent broadly neutralizing antibodies. *Science translational medicine*, 9(381):eaai7514, 2017.

[134] Nicole A Doria-Rose, Chaim A Schramm, Jason Gorman, Penny L Moore, Jinal N Bhiman, Brandon J DeKosky, Michael J Ernandes, Ivelin S Georgiev, Helen J Kim, Marie Pancera, et al. Developmental pathway for potent V1V2-directed HIV-neutralizing antibodies. *Nature*, 509(7498):55–62, 2014.

[135] Daniel T MacLeod, Nancy M Choi, Bryan Briney, Fernando Garces, Lorena S Ver, Elise Landais, Ben Murrell, Terri Wrin, William Kilembe, Chi-Hui Liang, et al. Early antibody lineage diversification and independent limb maturation lead to broad HIV-1 neutralization targeting the env high-mannose patch. *Immunity*, 44(5):1215–1226, 2016.

[136] Sergey Kryazhimskiy, Daniel P Rice, and Michael M Desai. Population subdivision and adaptation in asexual populations of saccharomyces cerevisiae. *Evolution: International Journal of Organic Evolution*, 66(6):1931–1941, 2012.

[137] Anne-Florence Bitbol and David J Schwab. Quantifying the role of population subdivision in evolution on rugged fitness landscapes. *PLoS Comput Biol*, 10(8):e1003778, 2014.

[138] Yang Zhang, Michael Meyer-Hermann, Laura A George, Marc Thilo Figge, Mahmood Khan, Margaret Goodall, Stephen P Young, Adam Reynolds, Francesco Falciani, Ari Waisman, et al. Germinal center B cells govern their own fate via antibody feedback. *Journal of Experimental Medicine*, 210(3):457–464, 2013.

[139] Mikhail Tikhonov, Shamit Kachru, and Daniel S Fisher. A model for the interplay between plastic tradeoffs and evolution in changing environments. *Proceedings of the National Academy of Sciences*, 117(16):8934–8940, 2020.

[140] O Tenaillon. The utility of fisher's geometric model in evolutionary genetics. *Annual review of ecology, evolution, and systematics*, 45:179–201, 2014.

[141] Zizhang Sheng, Chaim A Schramm, Mark Connors, Lynn Morris, John R Mascola, Peter D Kwong, and Lawrence Shapiro. Effects of darwinian selection and mutability on rate of broadly neutralizing antibody evolution during HIV-1 infection. *PLoS computational biology*, 12(5):e1004940, 2016.

[142] Penny L Moore, Nthabeleng Ranchobe, Bronwen E Lambson, Elin S Gray, Eleanor Cave, Melissa-Rose Abrahams, Gama Bandawe, Koleka Mlisana, Salim S Abdool Karim, Carolyn Williamson, et al. Limited neutralizing antibody specificities drive neutralization escape in early HIV-1 subtype C infection. *PLoS pathogens*, 5(9):e1000598, 2009.

[143] Kimberly M Cirelli, Diane G Carnathan, Bartek Nogal, Jacob T Martin, Oscar L Rodriguez, Amit A Upadhyay, Chiamaka A Enemuo, Etse H Gebru, Yury Choe, Federico Viviano, et al. Slow delivery immunization enhances HIV neutralizing antibody and germinal center responses via modulation of immunodominance. *Cell*, 177(5):1153–1171, 2019.

[144] Shenshen Wang and Lei Dai. Evolving generalists in switching rugged landscapes. *PLoS computational biology*, 15(10):e1007320, 2019.

[145] Vedant Sachdeva, Kabir Husain, Jiming Sheng, Shenshen Wang, and Arvind Murugan. Tuning environmental timescales to evolve and maintain generalists. *Proceedings of the National Academy of Sciences*, 117(23):12693–12699, 2020.

[146] Tanja A Schwickert, Randall L Lindquist, Guy Shakhar, Geulah Livshits, Dimitris Skokos, Marie H Kosco-Vilbois, Michael L Dustin, and Michel C Nussenzweig. In vivo imaging of germinal centres reveals a dynamic open structure. *Nature*, 446(7131):83, 2007.

[147] MJ Shlomchik, P Watts, MG Weigert, and S Litwin. Clone: a monte-carlo computer simulation of B cell clonal expansion, somatic mutation, and antigen-driven selection. In *Somatic Diversification of Immune Responses*, pages 173–197. Springer, 1998.

[148] Michele Shannon and Ramit Mehr. Reconciling repertoire shift with affinity maturation: the role of deleterious mutations. *The Journal of Immunology*, 162(7):3950–3956, 1999.

[149] Xiaocong Yu, Tshidi Tsibane, Patricia A McGraw, Frances S House, Christopher J Keefer, Mark D Hicar, Terrence M Tumpey, Claudia Pappas, Lucy A Perrone, Osvaldo Martinez, et al. Neutralizing antibodies derived from the B cells of 1918 influenza pandemic survivors. *Nature*, 455(7212):532, 2008.

[150] Dhruv K Sethi, Anupriya Agarwal, Venkatasamy Manivel, Kanury VS Rao, and Dinakar M Salunke. Differential epitope positioning within the germline antibody paratope enhances promiscuity in the primary immune response. *Immunity*, 24(4):429–438, 2006.

[151] John Coffin and Ronald Swanstrom. HIV pathogenesis: dynamics and genetics of viral populations and infected cells. *Cold Spring Harbor perspectives in medicine*, 3(1):a012526, 2013.

[152] Derek J Smith, Stephanie Forrest, Ron R Hightower, and Alan S Perelson. Deriving shape space parameters from immunological data. *Journal of theoretical biology*, 189(2):141–150, 1997.

[153] R Küppers, M Zhao, ML Hansmann, and K Rajewsky. Tracing B cell development in human germinal centres by molecular analysis of single cells picked from histological sections. *The EMBO journal*, 12(13):4955–4967, 1993.

[154] Jose Faro and Michal Or-Guil. How oligoclonal are germinal centers? a new method for estimating clonal diversity from immunohistological sections. In *BMC bioinformatics*, volume 14, page S8. BioMed Central, 2013.

[155] MJ Shlomchik, P Watts, MG Weigert, and S Litwin. Clone: a monte-carlo computer simulation of B cell clonal expansion, somatic mutation, and antigen-driven selection. In *Somatic Diversification of Immune Responses*, pages 173–197. Springer, 1998.

[156] Shenshen Wang, Jordi Mata-Fink, Barry Kriegsman, Melissa Hanson, Darrell J Irvine, Herman N Eisen, Dennis R Burton, K Dane Wittrup, Mehran Kardar, and Arup K Chakraborty. Manipulating the selection forces during affinity maturation to generate cross-reactive HIV antibodies. *Cell*, 160(4):785–797, 2015.

[157] Minna Pekkonen, Tarmo Ketola, and Jouni T Laakso. Resource availability and competition shape the evolution of survival and growth ability in a bacterial community. *PLoS One*, 8(9):e76471, 2013.

[158] Andrei Kucharavy, Boris Rubinstein, Jin Zhu, and Rong Li. Robustness and evolvability of heterogeneous cell populations. *Molecular biology of the cell*, 29(11):1400–1409, 2018.

[159] Rajesh Kumar, Huma Qureshi, Suprit Deshpande, and Jayanta Bhattacharya. Broadly neutralizing antibodies in HIV-1 treatment and prevention. *Therapeutic advances in vaccines and immunotherapy*, 6(4):61–68, 2018.

[160] Maria Trovato, Luciana D'Apice, Antonella Prisco, and Piergiuseppe De Berardinis. HIV vaccination: a roadmap among advancements and concerns. *International journal of molecular sciences*, 19(4):1241, 2018.

[161] Yubin Liu, Wei Cao, Ming Sun, and Taisheng Li. Broadly neutralizing antibodies for HIV-1: efficacies, challenges and opportunities. *Emerging Microbes & Infections*, 9(1):194–206, 2020.

[162] Lifen Han, Cong Chen, Xianlin Han, Shujin Lin, Xiulan Ao, Xiao Han, Jianmin Wang, and Hanhui Ye. Structural insights for anti-influenza vaccine design. *Computational and structural biotechnology journal*, 17:475–483, 2019.

[163] Masayuki Kuraoka, Yu Adachi, and Yoshimasa Takahashi. Hide and seek: interplay between influenza viruses and B cells. *International Immunology*, 2020.

[164] Hiroshi Nikaido. Multidrug resistance in bacteria. *Annual review of biochemistry*, 78:119–146, 2009.

[165] Behzad Mansoori, Ali Mohammadi, Sadaf Davudian, Solmaz Shirjang, and Behzad Baradaran. The different mechanisms of cancer drug resistance: a brief review. *Advanced pharmaceutical bulletin*, 7(3):339, 2017.

[166] Xuan Wang, Haiyun Zhang, and Xiaozhuo Chen. Drug resistance and combating drug resistance in cancer. *Cancer Drug Resist*, 2:141–160, 2019.

[167] Alex R Hall, Pauline D Scanlan, and Angus Buckling. Bacteria-phage coevolution and the emergence of generalist pathogens. *The American Naturalist*, 177(1):44–53, 2011.

[168] Rodrigo Bacigalupe, María Ángeles Tormo-Mas, José R Penadés, and J Ross Fitzgerald. A multihost bacterial pathogen overcomes continuous population bottlenecks to adapt to new host species. *Science advances*, 5(11):eaax0063, 2019.

[169] Peter D Kwong, John R Mascola, and Gary J Nabel. Rational design of vaccines to elicit broadly neutralizing antibodies to HIV-1. *Cold Spring Harbor perspectives in medicine*, 1(1):a007278, 2011.

[170] Veljko M Nikolin, Klaus Osterrieder, Veronika Von Messling, Heribert Hofer, Danielle Anderson, Edward Dubovi, Edgar Brunner, and Marion L East. Antagonistic pleiotropy and fitness trade-offs reveal specialist and generalist traits in strains of canine distemper virus. *PloS one*, 7(12):e50955, 2012.

[171] Daniel Dykhuizen and Maxine Davies. An experimental model: bacterial specialists and generalists competing in chemostats. *Ecology*, 61(5):1213–1227, 1980.

[172] David Berger, Richard J Walters, and Wolf U Blanckenhorn. Experimental evolution for generalists and specialists reveals multivariate genetic constraints on thermal reaction norms. *Journal of Evolutionary Biology*, 27(9):1975–1989, 2014.

[173] Elisabeth A Bernays. Neural limitations in phytophagous insects: implications for diet breadth and evolution of host affiliation. *Annual review of entomology*, 46(1):703–727, 2001.

[174] Karen Strickler. Specialization and foraging efficiency of solitary bees. *Ecology*, 60(5):998–1009, 1979.

[175] Frank Seebacher, Varlérie Ducret, Alexander G Little, and Bart Adriaenssens. Generalist–specialist trade-off during thermal acclimation. *Royal Society open science*, 2(1):140251, 2015.

[176] Timothy Griffith and Sonia E Sultan. Field-based insights to the evolution of specialization: plasticity and fitness across habitats in a specialist/generalist species pair. *Ecology and Evolution*, 2(4):778–791, 2012.

[177] Sara Via. The genetic structure of host plant adaptation in a spatial patchwork: demographic variability among reciprocally transplanted pea aphid clones. *Evolution*, 45(4):827–852, 1991.

[178] David Sloan Wilson and Jin Yoshimura. On the coexistence of specialists and generalists. *The American Naturalist*, 144(4):692–707, 1994.

[179] Marina C Caillaud and Sara Via. Specialized feeding behavior influences both ecological specialization and assortative mating in sympatric host races of pea aphids. *The American Naturalist*, 156(6):606–621, 2000.

[180] Thomas J DeWitt, Andrew Sih, and David Sloan Wilson. Costs and limits of phenotypic plasticity. *Trends in ecology & evolution*, 13(2):77–81, 1998.

[181] Samuel M Scheiner and David Berrigan. The genetics of phenotypic plasticity. viii. the cost of plasticity in daphnia pulex. *Evolution*, 52(2):368–378, 1998.

[182] Lisa M Bono, Catharine L Gensel, David W Pfennig, and Christina L Burch. Evolutionary rescue and the coexistence of generalist and specialist competitors: an experimental test. *Proceedings of the Royal Society B: Biological Sciences*, 282(1821):20151932, 2015.

[183] Dennis R Burton. Advancing an HIV vaccine; advancing vaccinology. *Nature Reviews Immunology*, 19(2):77, 2019.

[184] Frank John Fenner. The florey lecture, 1983-biological control, as exemplified by smallpox eradication and myxomatosis. *Proceedings of the Royal Society of London. Series B. Biological Sciences*, 218(1212):259–285, 1983.

[185] Brendan JM Bohannan and Richard E Lenski. Linking genetic change to community evolution: insights from studies of bacteria and bacteriophage. *Ecology letters*, 3(4):362–377, 2000.

[186] Takehito Yoshida, Laura E Jones, Stephen P Ellner, Gregor F Fussmann, and Nelson G Hairston. Rapid evolution drives ecological dynamics in a predator–prey system. *Nature*, 424(6946):303–306, 2003.

[187] Peter R Grant and B Rosemary Grant. Evolution of character displacement in darwin's finches. *science*, 313(5784):224–226, 2006.

[188] Ali R Vahdati, Kathleen Sprouffske, and Andreas Wagner. Effect of population size and mutation rate on the evolution of rna sequences on an adaptive landscape determined by rna folding. *International journal of biological sciences*, 13(9):1138, 2017.

[189] Tadeusz J Kawecki, NH Barton, and James D Fry. Mutational collapse of fitness in marginal habitats and the evolution of ecological specialisation. *Journal of Evolutionary Biology*, 10(3):407–429, 1997.

[190] Graham Bell. Evolutionary rescue and the limits of adaptation. *Philosophical Transactions of the Royal Society B: Biological Sciences*, 368(1610):20120080, 2013.

[191] Christian Guill and Barbara Drossel. Emergence of complexity in evolving niche-model food webs. *Journal of theoretical biology*, 251(1):108–120, 2008.

[192] Korinna T Allhoff and Barbara Drossel. When do evolutionary food web models generate complex networks? *Journal of theoretical biology*, 334:122–129, 2013.

[193] Andrew P Hendry and Michael T Kinnison. Perspective: the pace of modern life: measuring rates of contemporary microevolution. *Evolution*, 53(6):1637–1653, 1999.

[194] Thomas W Schoener. The newest synthesis: understanding the interplay of evolutionary and ecological dynamics. *science*, 331(6016):426–429, 2011.

[195] Lynn Govaert, Emanuel A Fronhofer, Sébastien Lion, Christophe Eizaguirre, Dries Bonte, Martijn Egas, Andrew P Hendry, Ayana De Brito Martins, Carlos J Melián, Joost AM Raeymaekers, et al. Ecoevolutionary feedbackstheoretical models and perspectives. *Functional Ecology*, 33(1):13–30, 2019.

[196] Richard Gomulkiewicz and Robert D Holt. When does evolution by natural selection prevent extinction? *Evolution*, pages 201–207, 1995.

[197] Andrew Gonzalez, Ophélie Ronce, Regis Ferriere, and Michael E Hochberg. Evolutionary rescue: an emerging focus at the intersection between ecology and evolution, 2013.

[198] Priscilla L Yang and Peter G Schultz. Mutational analysis of the affinity maturation of antibody 48g7. *Journal of molecular biology*, 294(5):1191–1201, 1999.

[199] Fred S Grodins, John S Gray, Karl R Schroeder, Arthur L Norins, and Richard W Jones. Respiratory responses to CO2 inhalation. a theoretical study of a nonlinear biological regulator. *Journal of applied physiology*, 7(3):283–308, 1954.

[200] Alan S Perelson. Applications of optimal control theory to immunology. In *Recent Developments in Variable Structure Systems, Economics and Biology*, pages 272–287. Springer, 1978.

[201] Daniel H Janzen. Herbivores and the number of tree species in tropical forests. *The American Naturalist*, 104(940):501–528, 1970.

[202] Robert H MacArthur. *Geographical ecology: patterns in the distribution of species*. Princeton University Press, 1984.

[203] Michael Lässig and Ville Mustonen. Eco-evolutionary control of pathogens. *Proceedings of the National Academy of Sciences*, 117(33):19694–19704, 2020.

[204] Jinghe Huang, Byong H Kang, Elise Ishida, Tongqing Zhou, Trevor Griesman, Zizhang Sheng, Fan Wu, Nicole A Doria-Rose, Baoshan Zhang, Krisha McKee, et al. Identification of a CD4-binding-site antibody to HIV that evolved near-pan neutralization breadth. *Immunity*, 45(5):1108–1121, 2016.

[205] Donald D Raymond, Goran Bajic, Jack Ferdman, Pirada Suphaphiphat, Ethan C Settembre, M Anthony Moody, Aaron G Schmidt, and Stephen C Harrison. Conserved epitope on influenza-virus hemagglutinin head defined by a vaccine-induced antibody. *Proceedings of the National Academy of Sciences*, 115(1):168–173, 2018.

[206] Muriel Lavie, Stéphane Sarrazin, Roland Montserret, Véronique Descamps, Thomas F Baumert, Gilles Duverlie, Karin Séron, François Penin, and Jean Dubuisson. Identification of conserved residues in hepatitis C virus envelope glycoprotein E2 that modulate virus dependence on CD81 and SRB1 entry factors. *Journal of virology*, 88(18):10584–10597, 2014.

[207] Scott A Brown, Sherri L Surman, Robert Sealy, Bart G Jones, Karen S Slobod, Kristen Branum, Timothy D Lockey, Nanna Howlett, Pamela Freiden, Patricia Flynn, et al. Heterologous prime-boost HIV-1 vaccination regimens in pre-clinical and clinical trials. *Viruses*, 2(2):435–467, 2010.

[208] Jean-Louis Excler and Jerome H Kim. Novel prime-boost vaccine strategies against HIV-1. *Expert review of vaccines*, 18(8):765–779, 2019.

[209] Jinal N Bhiman, Colin Anthony, Nicole A Doria-Rose, Owen Karimanzira, Chaim A Schramm, Thandeka Khoza, Dale Kitchin, Gordon Botha, Jason Gorman, Nigel J Garrett, et al. Viral variants that initiate and drive maturation of V1V2-directed HIV-1 broadly neutralizing antibodies. *Nature medicine*, 21(11):1332, 2015.

[210] Vishal Gupta, David J Earl, and Michael W Deem. Quantifying influenza vaccine efficacy and antigenic distance. *Vaccine*, 24(18):3881–3888, 2006.

[211] Thomas B Kepler and Alan S Perelson. Modeling and optimization of populations subject to time-dependent mutation. *Proceedings of the National Academy of Sciences*, 92(18):8219–8223, 1995.

[212] Herman N Eisen and Gregory W Siskind. Variations in affinities of antibodies during the immune response. *Biochemistry*, 3(7):996–1008, 1964.

[213] Myungsun Kang, Timothy J Eisen, Ellen A Eisen, Arup K Chakraborty, and Herman N Eisen. Affinity inequality among serum antibodies that originate in lymphoid germinal centers. *PLoS One*, 10(10):e0139222, 2015.

[214] Iliyana Mikell, D Noah Sather, Spyros A Kalams, Marcus Altfeld, Galit Alter, and Leonidas Stamatatos. Characteristics of the earliest cross-neutralizing antibody response to HIV-1. *PLoS Pathog*, 7(1):e1001251, 2011.

[215] Ismail Dogan, Barbara Bertocci, Valérie Vilmont, Frédéric Delbos, Jérome Mégret, Sébastien Storck, Claude-Agnès Reynaud, and Jean-Claude Weill. Multiple layers of B cell memory with different effector functions. *Nature immunology*, 10(12):1292, 2009.

[216] Joseph Jardine, Jean-Philippe Julien, Sergey Menis, Takayuki Ota, Oleksandr Kalyuzhniy, Andrew McGuire, Devin Sok, Po-Ssu Huang, Skye MacPherson, Meaghan Jones, et al. Rational HIV immunogen design to target specific germline B cell receptors. *Science*, 340(6133):711–716, 2013.

[217] Andrew T McGuire, Matthew D Gray, Pia Dosenovic, Alexander D Gitlin, Natalia T Freund, John Petersen, Colin Correnti, William Johnsen, Robert Kegel, Andrew B Stuart, et al. Specifically modified env immunogens activate b-cell precursors of broadly neutralizing HIV-1 antibodies in transgenic mice. *Nature communications*, 7(1):1–10, 2016.

[218] Joseph G Jardine, Daniel W Kulp, Colin Havenar-Daughton, Anita Sarkar, Bryan Briney, Devin Sok, Fabian Sesterhenn, June Ereño-Orbea, Oleksandr Kalyuzhniy, Isaiah Deresa, et al. HIV-1 broadly neutralizing antibody precursor B cells revealed by germline-targeting immunogen. *Science*, 351(6280):1458–1463, 2016.

[219] Vedant Sachdeva, Kabir Husain, Jiming Sheng, Shenshen Wang, and Arvind Murugan. Tuning environmental timescales to evolve and maintain generalists. *Proceedings of the National Academy of Sciences*, 117(23):12693–12699, 2020.

[220] Marco Molari, Klaus Eyer, Jean Baudry, Simona Cocco, and Rémi Monasson. Quantitative modeling of the effect of antigen dosage on b-cell affinity distributions in maturating germinal centers. *Elife*, 9:e55678, 2020.

[221] Rajesh A Keraliya, Chirag Patel, Pranav Patel, Vipul Keraliya, Tejal G Soni, Rajnikant C Patel, and MM Patel. Osmotic drug delivery system as a part of modified release dosage form. *ISRN pharmaceutics*, 2012, 2012.

[222] Amit Sharma, Damit Kumar, and Neelam Painuly. A review on osmotically controlled drug delivery systems. *Asian Journal of Pharmaceutical Research and Development*, 6(4):101–109, 2018.

[223] Balthasar A Heesters, Priyadarshini Chatterjee, Young-A Kim, Santiago F Gonzalez, Michael P Kuligowski, Tomas Kirchhausen, and Michael C Carroll. Endocytosis and recycling of immune complexes by follicular dendritic cells enhances B cell antigen binding and activation. *Immunity*, 38(6):1164–1175, 2013.

[224] James J Moon, Heikyung Suh, Adrienne V Li, Christian F Ockenhouse, Anjali Yadava, and Darrell J Irvine. Enhancing humoral responses to a malaria antigen with nanoparticle vaccines that expand tfh cells and promote germinal center induction. *Proceedings of the National Academy of Sciences*, 109(4):1080–1085, 2012.

[225] Hilbert J Kappen. Path integrals and symmetry breaking for optimal control theory. *Journal of statistical mechanics: theory and experiment*, 2005(11):P11011, 2005.

[226] Christodoulos A Floudas and Panos M Pardalos. *Encyclopedia of optimization*. Springer Science & Business Media, 2008.

[227] SE Fadugba, OH Edogbanya, and SC Zelibe. Crank nicolson method for solving parabolic partial differential equations. *IJA2M*, 1(3):8–23, 2013.

[228] Vincent Detours and Alan S Perelson. Explaining high alloreactivity as a quantitative consequence of affinity-driven thymocyte selection. *Proceedings of the National Academy of Sciences*, 96(9):5153–5158, 1999.

[229] Vincent Detours and Alan S Perelson. The paradox of alloreactivity and self mhc restriction: quantitative analysis and statistics. *Proceedings of the National Academy of Sciences*, 97(15):8479–8483, 2000.

[230] Brian Gaschen, Jesse Taylor, Karina Yusim, Brian Foley, Feng Gao, Dorothy Lang, Vladimir Novitsky, Barton Haynes, Beatrice H Hahn, Tanmoy Bhattacharya, et al. Diversity considerations in HIV-1 vaccine selection. *Science*, 296(5577):2354–2360, 2002.

[231] Lev Pontryagin, Vladmir Boltyanskii, Revaz Gamkrelidze, and Evgenii Mishechenko. *Mathematical Theory of Optimal Processes*. John Wiley & Sons, 1962.

[232] FL Chernousko and AA Lyubushin. Method of successive approximations for solution of optimal control problems. *Optimal Control Applications and Methods*, 3(2):101–114, 1982.

[233] Qianxiao Li, Long Chen, Cheng Tai, and E Weinan. Maximum principle based algorithms for deep learning. *The Journal of Machine Learning Research*, 18(1):5998–6026, 2017.

[234] Anna Melbinger and Massimo Vergassola. The impact of environmental fluctuations on evolutionary fitness functions. *Scientific reports*, 5:15211, 2015.

[235] Jonathan Desponds, Thierry Mora, and Aleksandra M Walczak. Fluctuating fitness shapes the clone-size distribution of immune repertoires. *Proceedings of the National Academy of Sciences*, 113(2):274–279, 2016.

[236] Shenshen Wang. Optimal sequential immunization can focus antibody responses against diversity loss and distraction. *PLoS computational biology*, 13(1):e1005336, 2017.

[237] Carl Boettiger. From noise to knowledge: how randomness generates novel phenomena and reveals information. *Ecology letters*, 21(8):1255–1267, 2018.

[238] Irmgard Fluegge-Lotz and Hubert Halkin. Pontryagin's maximum principle and optimal control. Technical report, STANFORD UNIV CA, 1961.

[239] Sergey V Drakunov and Vadim I Utkin. Sliding mode control in dynamic systems. *International Journal of Control*, 55(4):1029–1037, 1992.

[240] Vadim I Utkin. Sliding mode control design principles and applications to electric drives. *IEEE transactions on industrial electronics*, 40(1):23–36, 1993.

[241] Daniel T Gillespie. Exact stochastic simulation of coupled chemical reactions. *The journal of physical chemistry*, 81(25):2340–2361, 1977.

[242] Daniel T Gillespie. Approximate accelerated stochastic simulation of chemically reacting systems. *The Journal of chemical physics*, 115(4):1716–1733, 2001.

[243] Arjun K Mishra and Roy A Mariuzza. Insights into the structural basis of antibody affinity maturation from next-generation sequencing. *Frontiers in immunology*, 9:117, 2018.

[244] Thomas T Yoshikawa. Important infections in elderly persons. *Western Journal of Medicine*, 135(6):441, 1981.

[245] Theodore C Eickhoff, Ida L Sherman, and Robert E Serfling. Observations on excess mortality associated with epidemic influenza. *Jama*, 176(9):776–782, 1961.

[246] Kenneth E Powell and Laurence S Farer. The rising age of the tuberculosis patient: a sign of success and failure. *The Journal of Infectious Diseases*, 142(6):946–948, 1980.

[247] Laurence H Miller and Philip A Brunell. Zoster, reinfection or activation of latent virus?: Observations on the antibody response. *The American Journal of Medicine*, 49(4):480–483, 1970.

[248] Thomas T Yoshikawa. Perspective: aging and infectious diseases: past, present, and future. *Journal of Infectious Diseases*, 176(4):1053–1057, 1997.

[249] Kai Liu, Ying Chen, Ruzheng Lin, and Kunyuan Han. Clinical features of COVID-19 in elderly patients: A comparison with young and middle-aged patients. *Journal of Infection*, 80(6):e14 – e18, 2020.

[250] United Nations Department of Economic and Social Affairs, Population Division. World population prospects 2019: Ten key findings., 2019.

[251] Godfrey S Getz. Bridging the innate and adaptive immune systems, 2005.

[252] Anil Shanker, Menaka C. Thounaojam, Manoj K. Mishra, and Mikhail M. Dikov. Innate-adaptive immune crosstalk 2016. *Journal of Immunology Research*, 2017, 2017.

[253] Caetano Reis e Sousa. Activation of dendritic cells: translating innate into adaptive immunity. *Current Opinion in Immunology*, 16(1):21–25, 2004.

[254] Martin Pelletier, Laura Maggi, Alessandra Micheletti, Elena Lazzeri, Nicola Tamassia, Claudio Costantini, Lorenzo Cosmi, Claudio Lunardi, Francesco Annunziato, Sergio Romagnani, et al. Evidence for a cross-talk between human neutrophils and Th17 cells. *Blood*, 115(2):335–343, 2010.

[255] Tara M Strutt, K Kai McKinstry, and Susan L Swain. Control of innate immunity by memory CD4 T cells. In *Crossroads between Innate and Adaptive Immunity III*, pages 57–68. Springer, 2011.

[256] Anil Shanker. Adaptive control of innate immunity. *Immunology Letters*, 131(2):107–112, 2010.

[257] Clotilde Thery and Sebastian Amigorena. The cell biology of antigen presentation in dendritic cells. *Current Opinion in Immunology*, 13(1):45 – 51, 2001.

[258] Noah W Palm and Ruslan Medzhitov. Not so fast: adaptive suppression of innate immunity. *Nature Medicine*, 13(10):1142, 2007.

[259] Kwang Dong Kim, Jie Zhao, Sogyong Auh, Xuanming Yang, Peishuang Du, Hong Tang, and Yang-Xin Fu. Adaptive immune cells temper initial innate responses. *Nature Medicine*, 13(10):1248, 2007.

[260] Danielle Aw and Donald B Palmer. The origin and implication of thymic involution. *Aging and Disease*, 2(5):437, 2011.

[261] Jörg J Goronzy and Cornelia M Weyand. T cell development and receptor diversity during aging. *Current Opinion in Immunology*, 17(5):468–475, 2005.

[262] Daryl P Shanley, Danielle Aw, Nancy R Manley, and Donald B Palmer. An evolutionary perspective on the mechanisms of immunosenescence. *Trends in Immunology*, 30(7):374–381, 2009.

[263] Aisha Souquette, Justin Frere, Megan Smithey, Delphine Sauce, and Paul G Thomas. A constant companion: immune recognition and response to cytomegalovirus with aging and implications for immune fitness. *Geroscience*, 39(3):293–303, 2017.

[264] Marcia A Blackman and David L Woodland. The narrowing of the CD8 T cell repertoire in old age. *Current Opinion in Immunology*, 23(4):537–542, 2011.

[265] Helle Bruunsgaard, Peter Skinhøj, Jesper Qvist, and Bente Klarlund Pedersen. Elderly humans show prolonged in vivo inflammatory activity during pneumococcal infections. *The Journal of Infectious Diseases*, 180(2):551–554, 1999.

[266] H Bruunsgaard, S Ladelund, Agnes Nadelmann Pedersen, M Schroll, T Jørgensen, and BK Pedersen. Predicting death from tumour necrosis factor-alpha and interleukin-6 in 80-year-old people. *Clinical & Experimental Immunology*, 132(1):24–31, 2003.

[267] Stefano Volpato, Jack M Guralnik, Luigi Ferrucci, Jennifer Balfour, Paulo Chaves, Linda P Fried, and Tamara B Harris. Cardiovascular disease, interleukin-6, and risk of mortality in older women: the women's health and aging study. *Circulation*, 103(7):947–953, 2001.

[268] Tamas Fulop, Anis Larbi, Gilles Dupuis, Aurélie Le Page, Eric H Frost, Alan A Cohen, Jacek M Witkowski, and Claudio Franceschi. Immunosenescence and inflamm-aging as two sides of the same coin: friends or foes? *Frontiers in Immunology*, 8:1960, 2018.

[269] Claudio Franceschi, Paolo Garagnani, Paolo Parini, Cristina Giuliani, and Aurelia Santoro. Inflammaging: a new immune–metabolic viewpoint for age-related diseases. *Nature Reviews Endocrinology*, 14(10):576, 2018.

[270] Sean P Stromberg and Jean Carlson. Robustness and fragility in immunosenescence. *PLOS Computational Biology*, 2(11):1–7, 11 2006.

[271] Eric Boren and M.Eric Gershwin. Inflamm-aging: autoimmunity, and the immune-risk phenotype. *Autoimmunity Reviews*, 3(5):401–406, 2004.

[272] Anna Aiello, Farzin Farzaneh, Giuseppina Candore, Calogero Caruso, Sergio Davinelli, Caterina Maria Gambino, Mattia Emanuela Ligotti, Nahid Zareian, and Giulia Accardi. Immunosenescence and its hallmarks: How to oppose aging strategically? A review of potential options for therapeutic intervention. *Frontiers in Immunology*, 10:2247, 2019.

[273] Angela Reynolds, Jonathan Rubin, Gilles Clermont, Judy Day, Yoram Vodovotz, and G. Bard Ermentrout. A reduced mathematical model of the acute inflammatory response: I. Derivation of model and analysis of anti-inflammation. *Journal of Theoretical Biology*, 242(1):220–236, 2006.

[274] Raluca Eftimie, Joseph J Gillard, and Doreen A Cantrell. Mathematical models for immunology: current state of the art and future research directions. *Bulletin of Mathematical Biology*, 78(10):2091–2134, 2016.

[275] Arup K Chakraborty. A perspective on the role of computational models in immunology. *Annual Review of Immunology*, 35:403–439, 2017.

[276] Joanne L Dunster, Helen M Byrne, and John R King. The resolution of inflammation: a mathematical model of neutrophil and macrophage interactions. *Bulletin of Mathematical Biology*, 76(8):1953–1980, 2014.

[277] Sridevi Nagaraja, Anders Wallqvist, Jaques Reifman, and Alexander Y Mitrophanov. Computational approach to characterize causative factors and molecular indicators of chronic wound inflammation. *The Journal of Immunology*, 192(4):1824–1834, 2014.

[278] Yoram Vodovotz, Carson C Chow, John Bartels, Claudio Lagoa, Jose M Prince, Ryan M Levy, Rukmini Kumar, Judy Day, Jonathan Rubin, Greg Constantine, et al. In silico models of acute inflammatión in animals. *Shock*, 26(3):235–244, 2006.

[279] Yoram Vodovotz, Gregory Constantine, Jonathan Rubin, Marie Csete, Eberhard O Voit, and Gary An. Mechanistic simulations of inflammation: current state and future prospects. *Mathematical Biosciences*, 217(1):1–10, 2009.

[280] Qian Yang, Francois Berthiaume, and Ioannis P Androulakis. A quantitative model of thermal injury-induced acute inflammation. *Mathematical Biosciences*, 229(2):135–148, 2011.

[281] Rob J De Boer and Alan S Perelson. T cell repertoires and competitive exclusion. *Journal of Theoretical Biology*, 169(4):375–390, 1994.

[282] Emily R Stirk, Carmen Molina-París, and Hugo A van den Berg. Stochastic niche structure and diversity maintenance in the T cell repertoire. *Journal of Theoretical Biology*, 255(2):237–249, 2008.

[283] Courtney L Davis and Frederick R Adler. Mathematical models of memory CD8+ T-cell repertoire dynamics in response to viral infections. *Bulletin of Mathematical Biology*, 75(3):491–522, 2013.

[284] Kimberly J Schlesinger, Sean P Stromberg, and Jean M Carlson. Coevolutionary immune system dynamics driving pathogen speciation. *PloS One*, 9(7):e102821, 2014.

[285] Grace Y Chen and Gabriel Nuñez. Sterile inflammation: sensing and reacting to damage. *Nature Reviews Immunology*, 10(12):826, 2010.

[286] Kenneth L Rock, Eicke Latz, Fernando Ontiveros, and Hajime Kono. The sterile inflammatory response. *Annual Review of Immunology*, 28:321–342, 2009.

[287] Alan S. Perelson and George F. Oster. Theoretical studies of clonal selection: Minimal antibody repertoire size and reliability of self-non-self discrimination. *Journal of Theoretical Biology*, 81(4):645–670, 1979.

[288] Norman MacDonald and Time Lags. Lecture notes in biomathematics. *Springer*, 17:1059–1062, 1978.

[289] Rob J De Boer and Alan S Perelson. Quantifying T lymphocyte turnover. *Journal of theoretical biology*, 327:45–87, 2013.

[290] Daniel K Choo, Kaja Murali-Krishna, Rustom Anita, and Rafi Ahmed. Homeostatic turnover of virus-specific memory CD8 T cells occurs stochastically and is independent of CD4 T cell help. *The Journal of Immunology*, 185(6):3436–3444, 2010.

[291] Anders Wikby, Boo Johansson, Frederick Ferguson, and Jadwiga Olsson. Age-related changes in immune parameters in a very old population of swedish people: a longitudinal study. *Experimental Gerontology*, 29(5):531–541, 1994.

[292] Jean M Carlson and John Doyle. Highly optimized tolerance: Robustness and design in complex systems. *Physical Review Letters*, 84(11):2529, 2000.

[293] Kevin J Maloy, Laurence Salaun, Rachel Cahill, Gordon Dougan, Nigel J Saunders, and Fiona Powrie. CD4+ CD25+ TR cells suppress innate immune pathology through cytokine-dependent mechanisms. *Journal of Experimental Medicine*, 197(1):111–119, 2003.

[294] Lauren M Sompayrac. *How the immune system works*. Wiley-Blackwell, 1999.

[295] Danielle Aw, Alberto B. Silva, and Donald B. Palmer. Immunosenescence: emerging challenges for an ageing population. *Immunology*, 120(4):435–446, Apr 2007.

[296] B. Weinberger and B. Grubeck-Loebenstein. Vaccines for the elderly. *Clinical Microbiology and Infection*, 18:100–108, 2012.

[297] Tamas Fulop, Aurélie Le Page, Carl Fortin, Jacek M Witkowski, Gilles Dupuis, and Anis Larbi. Cellular signaling in the aging immune system. *Current Opinion in Immunology*, 29:105–111, 2014.

[298] Donald B Palmer. The effect of age on thymic function. *Frontiers in immunology*, 4:316, 2013.

[299] Heather E Lynch, Gabrielle L Goldberg, Ann Chidgey, Marcel RM Van den Brink, Richard Boyd, and Gregory D Sempowski. Thymic involution and immune reconstitution. *Trends in immunology*, 30(7):366–373, 2009.

[300] Christopher J. Bayne. Origins and evolutionary relationships between the innate and adaptive arms of immune systems. *Integrative and Comparative Biology*, 43(2):293–299, 04 2003.

[301] Ludmila Müller, Tamas Fülöp, and Graham Pawelec. Immunosenescence in vertebrates and invertebrates. *Immunity & Ageing*, 10(1):12–12, Apr 2013.

[302] A Katharina Simon, Georg A Hollander, and Andrew McMichael. Evolution of the immune system in humans from infancy to old age. *Proceedings of the Royal Society B: Biological Sciences*, 282(1821):20143085, 2015.

[303] Chan C Whiting, Janet Siebert, Aaron M Newman, Hong-wu Du, Ash A Alizadeh, Jorg Goronzy, Cornelia M Weyand, Eswar Krishnan, C Garrison Fathman, and Holden T Maecker. Large-scale and comprehensive immune profiling and functional analysis of normal human aging. *PLoS One*, 10(7):e0133627, 2015.

[304] Katelyn M. Gostic, Monique Ambrose, Michael Worobey, and James O. Lloyd-Smith. Potent protection against H5N1 and H7N9 influenza via childhood hemagglutinin imprinting. *Science*, 354(6313):722–726, 2016.

[305] Centers for Disease Control and Prevention. AIDS diagnosis. `https://www.cdc.gov/hiv/basics/whatishiv.html`. Accessed: 2021-04-23.

[306] Roger J Pomerantz. Reservoirs of human immunodeficiency virus type 1: the main obstacles to viral eradication. *Clinical infectious diseases*, 34(1):91–97, 2002.

[307] Welkin E Johnson and Ronald C Desrosiers. Viral persistence: HIV's strategies of immune system evasion. *Annual review of medicine*, 53(1):499–518, 2002.

[308] Tae-Wook Chun, Lieven Stuyver, Stephanie B Mizell, Linda A Ehler, Jo Ann M Mican, Michael Baseler, Alun L Lloyd, Martin A Nowak, and Anthony S Fauci. Presence of an inducible HIV-1 latent reservoir during highly active antiretroviral therapy. *Proceedings of the National Academy of Sciences*, 94(24):13193–13197, 1997.

[309] Diana Finzi, Monika Hermankova, Theodore Pierson, Lucy M Carruth, Christopher Buck, Richard E Chaisson, Thomas C Quinn, Karen Chadwick, Joseph Margolick, Ronald Brookmeyer, et al. Identification of a reservoir for HIV-1 in patients on highly active antiretroviral therapy. *Science*, 278(5341):1295–1300, 1997.

[310] Joseph K Wong, Marjan Hezareh, Huldrych F Günthard, Diane V Havlir, Caroline C Ignacio, Celsa A Spina, and Douglas D Richman. Recovery of replication-competent HIV despite prolonged suppression of plasma viremia. *Science*, 278(5341):1291–1295, 1997.

[311] Theodore Pierson, Justin McArthur, and Robert F Siliciano. Reservoirs for HIV-1: mechanisms for viral persistence in the presence of antiviral immune responses and antiretroviral therapy. *Annual review of immunology*, 18(1):665–708, 2000.

[312] Tamas Fulop, Aurélie Le Page, Carl Fortin, Jacek M Witkowski, Gilles Dupuis, and Anis Larbi. Cellular signaling in the aging immune system. *Current opinion in immunology*, 29:105–111, 2014.

[313] Andreas Mayer, Thierry Mora, Olivier Rivoire, and Aleksandra M Walczak. Diversity of immune strategies explained by adaptation to pathogen statistics. *Proceedings of the National Academy of Sciences*, 113(31):8630–8635, 2016.

[314] Oskar H Schnaack and Armita Nourmohammad. Optimal evolutionary decision-making to store immune memory. *arXiv preprint arXiv:2007.01363*, 2020.

[315] SF Schluter, RM Bernstein, H Bernstein, and JJ Marchalonis. 'big bang'emergence of the combinatorial immune system. *Developmental and comparative immunology*, 23(2):107–111, 1999.

[316] Martin F Flajnik and Masanori Kasahara. Origin and evolution of the adaptive immune system: genetic events and selective pressures. *Nature Reviews Genetics*, 11(1):47–59, 2010.

[317] Christopher J Bayne. Origins and evolutionary relationships between the innate and adaptive arms of immune systems. *Integrative and comparative biology*, 43(2):293–299, 2003.