Immunology and Evolution of Infectious Disease

STEVEN A. FRANK

Princeton University Press
Princeton and Oxford

Published by Princeton University Press,
41 William Street, Princeton, New Jersey 08540
In the United Kingdom: Princeton University Press,
3 Market Place, Woodstock, Oxfordshire OX20 1SY

Library of Congress Cataloging-in-Publication Data

Frank, Steven A., 1957–
Immunology and Evolution of Infectious Disease /
Steven A. Frank. p. cm.
Includes bibliographic references and index.
ISBN 0–691–09594–9 (cloth : alk. paper)
ISBN 0–691–09595–7 (pbk. : alk. paper)
1. Immunogenetics. 2. Host-parasite relationships—
Genetic aspects. 3. Microorganisms—Evolution.
4. Antigens. 5. Molecular evolution.
6. Parasite antigens—Variation. I. Title.
[DNLM: 1. Communicable Diseases—immunology.
2. Evolution, Molecular. 3. Genetics, Population.
4. Immunity—genetics. WC 100 F828i 2002]
QR184 .F73 2002
616.9′0479—dc21 2002018384

British Library Cataloging-in-Publication Data is available

Typeset by the author with TEX
Composed in Lucida Bright

Printed on acid-free paper. ∞

www.pupress.princeton.edu

Printed in the United States of America

10 9 8 7 6 5 4 3 2 1

10 9 8 7 6 5 4 3 2 1
(Pbk.)

Contents

PART V: STUDYING EVOLUTION

Acknowledgments

My wife, Robin Bush, read earlier drafts and helped in every way. Camille Barr provided comments on the entire manuscript. My department, led by Chair Al Bennett, gave me the freedom to read and write over nearly two years. The National Science Foundation and the National Institutes of Health funded my research. My web pages at http://stevefrank.org/ provide information and updates for this book.

1 Introduction

Multidisciplinary has become the watchword of modern biology. Surely, the argument goes, a biologist interested in the biochemical pathways by which genetic variants cause disease would also want to understand the population processes that determine the distribution of genetic variants. And how can one expect to understand the interacting parts of complex immune responses without knowing something of the historical and adaptive processes that built the immune system?

Working in the other direction, evolutionary biologists have often treated amino acid substitutions within a parasite lineage as simply statistical marks to be counted and analyzed by the latest mathematical techniques. More interesting work certainly follows when hypotheses about evolutionary change consider the different selective pressures caused by antibody memory, variation among hosts in MHC genotype, and the epidemiological contrasts between rapidly and slowly spreading infectious diseases.

Synthesis between the details of molecular biology and the lives of organisms in populations will proceed slowly. It is now hard enough to keep up in one's own field, and more difficult to follow the foreign concepts and language of other subjects. The typical approach to synthesis uses an academic discipline to focus a biological subject. I use the biological problem of parasite variation to tie together many different approaches and levels of analysis.

Why should parasite variation be the touchstone for the integration of disciplines in modern biology? On the practical side, infectious disease remains a major cause of morbidity and mortality. Consequently, great research effort has been devoted to parasites and to host immune responses that fight parasites. This has led to rapid progress in understanding the biology of parasites, including the molecular details about how parasites invade hosts and escape host immune defenses. Vaccines have followed, sometimes with spectacular success.

But many parasites escape host defense by varying their *antigenic* molecules recognized by host immunity. Put another way, rapid evolution of antigenic molecules all too often prevents control of parasite

populations. The challenge has been to link molecular understanding of parasite molecules to their evolutionary change and to the antigenic variation in populations of parasites.

On the academic side, the growth of information about antigenic variation provides a special opportunity. For example, one can find in the literature details about how single amino acid changes in parasite molecules allow escape from antibody binding, and how that escape promotes the spread of variant parasites. Evolutionary studies no longer depend on abstractions—one can pinpoint the physical basis for success or failure and the consequences for change in populations.

Molecular understanding of host-parasite recognition leads to a comparative question about the forces that shape variability. Why do some viruses escape host immunity by varying so rapidly over a few years, whereas other viruses hardly change their antigens? The answer leads to the processes that shape genetic variability and evolutionary change. The causes of variability and change provide the basis for understanding why simple vaccines work well against some viruses, whereas complex vaccine strategies achieve only limited success against other viruses.

I did not start out by seeking a topic for multidisciplinary synthesis. Rather, I have long been interested in how the molecular basis of recognition between attackers and defenders sets the temporal and spatial scale of the battle. Attack and defense occur between insects and the plants they eat, between fungi and the crop plants they destroy, between viruses and the bacteria they kill, between different chromosomes competing for transmission through gametes, and between vertebrate hosts and their parasites. The battle often comes down to the rates at which attacker and defender molecules bind or evade each other. The biochemical details of binding and recognition set the rules of engagement that shape the pacing, scale, and pattern of diversity and the nature of evolutionary change.

Of the many cases of attack and defense across all of biology, the major parasites of humans and their domestic animals provide the most information ranging from the molecular to the population levels. New advances in the conceptual understanding of attack and defense will likely rise from the facts and the puzzles of this subject. I begin by putting the diverse, multidisciplinary facts into a coherent whole. From that foundation, I describe new puzzles and define the key problems for the future study of parasite variation and escape from host recognition.

I start at the most basic level, the nature of binding and recognition between host and parasite molecules. I summarize the many different ways in which parasites generate new variants in order to escape molecular recognition.

Next, I build up the individual molecular interactions into the dynamics of a single infection within a host. The parasites spread in the host, triggering immune attack against dominant antigens. The battle within the host develops through changes in population numbers—the numbers of parasites with particular antigens and the numbers of immune cells that specifically bind to particular antigens.

I then discuss how the successes and failures of different parasite antigens within each host determine the rise and fall of parasite variants over space and time. The distribution of parasite variants sets the immune memory profiles of different hosts, which in turn shape the landscape in which parasite variants succeed or fail. These coevolutionary processes determine the natural selection of antigenic variants and the course of evolution in the parasite population.

Finally, I consider different ways to study the evolution of antigenic variation. Experimental evolution of parasites under controlled conditions provides one way to study the relations between molecular recognition, the dynamics of infections within hosts, and the evolutionary changes in parasite antigens. Sampling of parasites from evolving populations provides another way to test ideas about what shapes the distribution of parasite variants.

My primary goal is to synthesize across different levels of analysis. How do the molecular details of recognition and specificity shape the changing patterns of variants in populations? How does the epidemiological spread of parasites between hosts shape the kinds and amounts of molecular variation in parasite antigens?

I compare different types of parasites because comparative biology provides insight into evolutionary process. For example, parasites that spread rapidly and widely in host populations create a higher density of immune memory in their hosts than do parasites that spread slowly and sporadically. Host species that quickly replace their populations with offspring decay their population-wide memory of antigens faster than do host species that reproduce more slowly. How do these epidemiological and demographic processes influence molecular variation of parasite antigens?

I end each chapter with a set of problems for future research. These problems emphasize the great opportunities of modern biology. At the molecular level, new technologies provide structural data on the three-dimensional shape of host antibody molecules bound to parasite antigens. At the population level, genomic sequencing methods provide detailed data on the variations in parasite antigens. One can now map the nucleotide variations of antigens and their associated amino acid substitutions with regard to the three-dimensional location of antibody binding. Thus, the spread of nucleotide variations in populations can be directly associated with the changes in molecular binding that allow escape from antibody recognition.

No other subject provides such opportunity for integrating the recent progress in structural and molecular analysis with the conceptual and methodological advances in population dynamics and evolutionary biology. My problems for future research at the end of each chapter emphasize the new kinds of questions that one can ask by integrating different levels of biological analysis.

Part I of the book gives general background. Chapter 2 summarizes the main features of vertebrate immunity. I present enough about the key cells and molecules so that one can understand how immune recognition shapes the diversity of parasites.

Chapter 3 describes various benefits that antigenic variation provides to parasites. These benefits explain why parasites vary in certain ways. For example, antigenic variation can help to escape host immunity during a single infection, extending the time a parasite can live within a particular host. Or antigenic variation may avoid the immunological memory of hosts, allowing the variant to spread in a population that previously encountered a different variant of that parasite. Different benefits favor different patterns of antigenic variation.

Part II introduces molecular processes. Chapter 4 describes the attributes of host and parasite molecules that contribute to immune recognition. The nature of recognition depends on specificity, the degree to which the immune system distinguishes between different antigens. Sometimes two different antigens bind to the same immune receptors, perhaps with different binding strength. This cross-reactivity protects

hosts against certain antigenic variants, and sets the molecular distance by which antigenic types must vary to escape recognition. Cross-reactivity may also interfere with immune recognition when immune receptors bind a variant sufficiently to prevent a new response but not strongly enough to clear the variant.

Chapter 5 summarizes the different ways in which parasites generate antigenic variants. Many parasites generate variants by the standard process of rare mutations during replication. Baseline mutation rates vary greatly, from about 10^{-5} per nucleotide per generation for the small genomes of some RNA viruses to about 10^{-11} for larger genomes. Although mutations occur rarely at any particular site during replication, large populations generate significant numbers of mutations in each generation. Some parasites focus hypermutation directly on antigenic loci. Other parasites store within each genome many genetic variants for an antigenic molecule. These parasites express only one genetic variant at a time and use specialized molecular mechanisms to switch gene expression between the variants.

Part III focuses on the dynamics of a single infection within a particular host. Chapter 6 emphasizes the host side, describing how the immune response develops strongly against only a few of the many different antigens that occur in each parasite. This immunodominance arises from interactions between the populations of immune cells with different recognition specificities and the population of parasites within the host. Immunodominance determines which parasite antigens face strong pressure from natural selection and therefore which antigens are likely to vary over space and time. To understand immunodominance, I step through the dynamic processes that regulate an immune response and determine which recognition specificities become amplified.

Chapter 7 considers the ways in which parasites escape recognition during an infection and the consequences for antigenic diversity within hosts. The chapter begins with the role of escape by mutation in persistent infections by HIV and hepatitis C virus. I then discuss how other parasites extend infection by switching gene expression between variants stored within each genome. This switching leads to interesting population dynamics within the host. The different variants rise and fall in abundance according to the rate of switching between variants, the time lag in the expansion of parasite lineages expressing a particular variant, and the time lag in the host immune response to each variant.

Part IV examines variability in hosts and parasites across entire populations. Chapter 8 considers genetic differences among hosts in immune response. Hosts differ widely in their major histocompatibility complex (MHC) alleles, which cause different hosts to recognize and focus their immune responses on different parasite antigens. This host variability can strongly affect the relative success of antigenic variants as they attempt to spread from host to host. Hosts also differ in minor ways in other genetic components of specific recognition. Finally, host polymorphisms occur in the regulation of the immune response. These quantitative differences in the timing and intensity of immune reactions provide an interesting model system for studying the genetics of regulatory control.

Chapter 9 describes differences among hosts in their molecular memory of antigens. Each host typically retains the ability to respond quickly to antigens that it encountered in prior infections. This memory protects the host against reinfection by the same antigens, but not against antigenic variants that escape recognition. Each host has a particular memory profile based on past infections. The distribution of memory profiles in the host population determines the ability of particular antigenic variants to spread between hosts. Hosts retain different kinds of immunological memory (antibody versus T cell), which affect different kinds of parasites in distinct ways.

Chapter 10 reviews the genetic structure of parasite populations. The genetic structure of nonantigenic loci provides information about the spatial distribution of genetic variability, the mixing of parasite lineages by transmission between hosts, and the mixing of genomes by sexual processes. The genetic structure of antigenic loci can additionally be affected by the distribution of host immunological memory, because parasites must avoid the antigen sets stored in immunological memory. Host selection on antigenic sets could potentially structure the parasite population into distinct antigenic strains. Finally, each host forms a separate island that divides the parasite population from other islands (hosts). This island structuring of parasite populations can limit the exchange of parasite genes by sexual processes, causing a highly inbred structure. Island structuring also means that each host receives a small and stochastically variable sample of the parasite population. Stochastic fluctuations may play an important role in the spatial distribution of antigenic variation.

Part V considers different methods to study the evolutionary processes that shape antigenic variation. Chapter 11 contrasts two different ways to classify parasite variants sampled from populations. Immunological assays compare the binding of parasite isolates to different immune molecules. The reactions of each isolate with each immune specificity form a matrix from which one can classify antigenic variants according to the degree to which they share recognition by immunity. Alternatively, one can classify isolates phylogenetically, that is, by time since divergence from a common ancestor. Concordant immunological and phylogenetic classifications frequently arise because immunological distance often increases with time since a common ancestor, reflecting the natural tendency for similarity by common descent. Discordant patterns of immunological and phylogenetic classifications indicate some evolutionary pressure on antigens that distorts immunological similarity. I show how various concordant and discordant relations point to particular hypotheses about the natural selection of antigenic properties in influenza and HIV.

Chapter 12 introduces experimental evolution, a controlled method to test hypotheses about the natural selection of antigenic diversity. This chapter focuses on foot-and-mouth disease virus. This well-studied virus illustrates how one can measure multiple selective forces on particular amino acids. Selective forces on amino acids in viral surface molecules include altered binding to host-cell receptors and changed binding to host antibodies. The selective forces imposed by antibodies and by attachment to host-cell receptors can be varied in experimental evolution studies to test their effects on amino acid change in the parasite. The amino acid substitutions can also be mapped onto three-dimensional structural models of the virus to analyze how particular changes alter binding properties.

Chapter 13 continues with experimental evolution of influenza A viruses. Experimental evolution has shown how altering the host species favors specific amino acid changes in the influenza surface protein that binds to host cells. Experimental manipulation of host-cell receptors and antibody pressure can be combined with structural data to understand selection on the viral surface amino acids. These mechanistic analyses of selection can be combined with observations on evolutionary change in natural populations to gain a better understanding of how selection shapes the observed patterns of antigenic variation.

Chapter 14 discusses experimental evolution of antigenic escape from host T cells. The host T cells can potentially bind to any short peptide of an intracellular parasite, whereas antibodies typically bind only to the surface molecules of parasites. T cell binding to parasite peptides depends on a sequence of steps by which hosts cut up parasite proteins and present the resulting peptides on the surfaces of host cells. Parasite escape from T cell recognition can occur at any of the processing steps, including the digestion of proteins, the transport of peptides, the binding of peptides by the highly specific host MHC molecules, and the binding of peptide-MHC complexes to receptors on the T cells. One or two amino acid substitutions in a parasite protein can often abrogate binding to MHC molecules or to the T cell receptors. Experimental evolution has helped us to understand escape from T cells because many of the steps can be controlled, such as the MHC alleles carried by the host and the specificities of the T cell receptors. Parasite proteins may be shaped by opposing pressures on physiological performance and escape from recognition.

Chapter 15 turns to samples of nucleotide sequences from natural populations. A phylogenetic classification of sequences provides a historical reconstruction of evolutionary relatedness and descent. Against the backdrop of ancestry, one can measure how natural selection has changed particular attributes of parasite antigens. For example, one can study whether selection caused particular amino acids to change rapidly or slowly. The rates of change for particular amino acids can be compared with the three-dimensional structural location of the amino acid site, the effects on immunological recognition, and the consequences for binding to host cells. The changes in natural populations can also be compared with patterns of change in experimental evolution, in which one controls particular selective forces. Past evolutionary change in population samples may be used to predict which amino acid variants in antigens are likely to spread in the future.

The last chapter recaps some interesting problems for future research that highlight the potential to study parasites across multiple levels of analysis.