Disease Evolution
Models, Concepts, and Data Analyses

Zhilan Feng
Ulf Dieckmann
Simon Levin
Editors

American Mathematical Society
Disease Evolution
Models, Concepts, and Data Analyses
This page intentionally left blank
Disease Evolution
Models, Concepts, and Data Analyses

Zhilan Feng
Ulf Dieckmann
Simon Levin
Editors

Center for Discrete Mathematics
and Theoretical Computer Science
A consortium of Rutgers University, Princeton University,
AT&T Labs–Research, Bell Labs (Lucent Technologies),
NEC Laboratories America, and Telcordia Technologies
(with partners at Avaya Labs, HP Labs, IBM Research,
Microsoft Research, and Stevens Institute of Technology)

American Mathematical Society
Contents

Foreword vii
Preface ix
Editors’ introduction xiii

Section I: Model Infrastructure
The implications of spatial structure within populations to the evolution of parasites
Mike Boots, Masashi Kamo, and Akira Sasaki 3
Insights from Price’s equation into evolutionary epidemiology
Troy Day and Sylvain Gandon 23
Within-host pathogen dynamics: Some ecological and evolutionary consequences of transients, dispersal mode, and within-host spatial heterogeneity
Robert D. Holt and Michael Barfield 45
Evolutionary and dynamic models of infection with internal host structure
John K. Kelly 67
Basic methods for modeling the invasion and spread of contagious diseases
Wayne M. Getz and James O. Lloyd-Smith 87

Section II: Applications to Specific Diseases
Modeling the invasion and spread of contagious diseases in heterogeneous populations
Wayne M. Getz, James O. Lloyd-Smith, Paul C. Cross, Shirli Bar-David, Philip L. Johnson, Travis C. Porco, and María S. Sánchez 113
A cophylogenetic perspective on host-pathogen evolution
Michael A. Charleston and Alison P. Galvani 145
The influence of anti-viral drug therapy on the evolution of HIV-1 pathogens
Zhilan Feng and Libin Rong 161
Do rhinoviruses follow the neutral theory? The role of cross-immunity in maintaining the diversity of the common cold
William J. Koppelman and Frederick R. Adler 181
Drug resistance in acute viral infections: Rhinovirus as a case study
Alun L. Lloyd and Dominik Wodarz 193
Dynamics and control of antibiotic resistance in structured metapopulations
David L. Smith, Maciej F. Boni, and Ramanan Laxminarayan 213
This page intentionally left blank
Foreword

This volume has its genesis in the activities of the DIMACS working group on Genetics and Evolution of Pathogens. This working group held a meeting on November 24-25, 2003 at Rutgers University, and we would like to express our appreciation to Zhilan Feng for organizing and planning this successful conference. The volume represents an expansion of the efforts of this working group, and contains papers from experts in the field who were unable to attend this initial meeting. We thank the three editors Zhilan Feng, Ulf Dieckmann and Simon Levin for their efforts in the organization of the volume, and we thank Bruce Levin for his insightful Preface and also the various authors who contributed to the volume.

The meeting was part of the 2002-2007 Special Focus on Computational and Mathematical Epidemiology, and was organized by one of a number of special focus research groups called “working groups” as part of the special focus. We extend our thanks to Martin Farach-Colton, Sunetra Gupta, Donald Hoover, David Krakauer, Simon Levin, Marc Lipsitch, David Madigan, Megan Murray, S. Muthukrishnan, David Ozonoff, Fred Roberts, Burton Singer and Daniel Wartenberg for their work as special focus organizers.

The meeting brought together researchers who approach the study of epidemiology from a variety of disciplines, some applied and some theoretical. These included computer scientists, mathematicians, statisticians, and biologists together with both descriptive and analytical epidemiologists. The goal of the working group as well as of this volume is the exploration of cross-disciplinary approaches to the study of topics related to disease evolution, and how they apply to the study of specific diseases.

DIMACS gratefully acknowledges the generous support that makes these programs possible. Special thanks go to the National Science Foundation, the James S. McDonnell Foundation, the Burroughs-Wellcome Fund, the Purdue University Mathematics Department and to DIMACS partners at Rutgers, Princeton, AT&T Labs - Research, Bell Labs, NEC Laboratories America, and Telcordia Technologies, and affiliate partners Avaya Labs, HP Labs, IBM Research, and Microsoft Research.

Fred S. Roberts
Director

Robert Tarjan
Co-Director for Princeton
This page intentionally left blank
Preface

Infections by microparasites (bacteria, viruses, protozoa and single celled fungi) are the primary source of human mortality in the underdeveloped world. And, despite all of the improvement in public health, hygiene, nutrition, living conditions and medical intervention over the past century, infections continue to be a major cause of morbidity and mortality in the developed world as well. Indeed, if we include people compromised by age, cancers and other diseases with immune-suppressing effects and/or treatments, coronary artery, diabetes, and other non-contagious and degenerative diseases, bacterial infections (often acquired in hospitals) may well be the major immediate cause of death even in overdeveloped countries.

Traditionally, the study of infectious diseases and their prevention and treatment has been the purview of epidemiologists, microbiologists, immunologists and clinicians – people who generally have little background in or appreciation for mathematics beyond statistics, if that. While the importance of quantitative reasoning for studies of the epidemiology of infectious diseases has been recognized for some time, this has been less so for investigations of the course of infections and their treatment within individual patients. For the most part, protocols for preventing the spread of infections in hospitals and communities and for the treatment of individual patients are based on qualitative considerations, experience and intuition, with money being the primary quantitative element in their design and implementation.

During the past two decades, studies of the epidemiology, evolution and within-host biology of infectious diseases and the development of methods for their prevention and treatment have been increasingly infiltrated by quantitative methods beyond statistics. A number of applied and not-so-applied mathematicians, mathematically trained and oriented epidemiologists, microbiologists, immunologists, ecologists, population and evolutionary biologists and even real doctors have been using mathematical and numerical models (computer simulations) to study the epidemiology, evolution and within-host dynamics of infectious diseases and to develop and evaluate protocols for their prevention and treatment. This collection is an impressive sampling of the nature and diversity of this epidemic of mathematical and numerical modelling for the studies of infectious diseases. It illustrates some of the delicious problems and opportunities for mathematicians and mathematical biologists that infectious diseases pose – problems that have the virtues of being important to human health and well-being and, at the same time, being challenging and intriguing even from the precious heights of academe.

For generality, tractability and the aesthetic appeal of closed-form mathematical analysis, traditional models of the epidemiology and evolution of infectious diseases have been deterministic and give little or no consideration to the spatial, temporal and other heterogeneities of human and other host communities and those of the microbes that infect them. In recent years, increasing numbers of modelers have been confronting these inconveniences and the unfortunate finiteness of the real world and exploring how they affect the inferences about the epidemiology and evolution of infectious diseases drawn from simpler models. Three of the chapters in this collection illustrate this trend. Mike Boots and his collaborators consider the
effects of the spatial structure of host populations on the evolution of the virulence of the microparasites that infect them. The two chapters by Wayne Getz and his cohorts examine, in a pedagogically useful as well as scholarly way, the consequences of spatial structure and stochastic processes on the spread of infectious diseases. Their chapters illustrate the utility of modelling to understanding the ascent and spread of emerging and reemerging diseases like SARS and tuberculosis, predicting their emergence and evaluating methods to control their dissemination.

In a commentary with a title that should appeal to this audience, “In Theory”, Sidney Brenner referred to molecular biology as the “great leveler” and suggested that for many it has made thinking unnecessary, a position I do not challenge. On the other side, the ease with which data and particularly those on the nucleotide sequences of DNA can be and have been gathered has also provided an opportunity for the quantitative study of evolutionary history through phylogenies. God is no longer the only one who can make a tree. Phylogenies generated from DNA sequence data – molecular phylogenies – have been the largest growth industry in the evolution business for the past decade. And, the development (if not always the application) of methods for generating, analyzing and interpreting these trees is an activity that requires serious thinking. The chapter by Charleston and Galvani is a fine example of this thinking applied to ascertaining the evolutionary relationship between interacting organisms like parasites and their hosts. “Co-phylogenetic” methods of the sort they are developing are of practical as well as academic interest. They can be used to determine the origins (original hosts) of newly emerging infectious diseases like HIV/AIDS and in that way better understand the conditions responsible for their emergence.

A prominent approach to drawing inferences about the nature and direction of evolution of parasites and their virulence has been to study their ecology (population dynamics and demography) within individual hosts or communities of hosts. In this perspective, the fitness of the microparasites is proportional to their reproductive number, $R_0$ – the number of secondary infections in a largely (or better yet, wholly) uninfected population of cells or tissues in an infected host or among individuals in a community of hosts. While this ecological approach to evolutionary inferences is explicit about nature and functional form of the selection pressures responsible for evolution, it does not consider the genetic basis of the variability upon which that selection is operating. In different ways and with different foci, two chapters consider ways to meld the ecological approach to the study of the evolution of microparasites and their virulence with those of population and quantitative genetics in situations where the nature of inheritance is explicit but where the ecological basis of selection is not. In their chapter, Troy Day and Sylvain Gandon consider how to apply classy population genetic approaches, like the Price equation, to studies of the evolution of microparasites and their virulence in communities of hosts. In his chapter, John Kelly uses a combination of ecological and population genetic methods to explore the contribution of tissue heterogeneity to the evolution of viruses in within infected hosts.

The contribution of the heterogeneity of the within-host habitat to the population and evolutionary dynamics of microparasites is also the focus of three other chapters in this collection. While modelers as well as experimentalists have the convenience of separately studying microparasite ecology and evolution within infected host and in communities of hosts, in the real world microbes have no choice but
to deal with both of these elements of their ecology and evolution. Although their models focus primarily on the within-host population and evolutionary dynamics of viruses and the contribution of within-host heterogeneity to that evolution, in their chapter Robert Holt and Michael Barfield consider how the within-host biology of microparasites contributes to their evolution in communities of hosts. In the chapter by Zhilan Feng and Libin Rong, the within-host heterogeneity of concern is comprised of the selective environments imposed by treatment with multiple drugs that act at different stages in the microparasite replication cycle. Using an age-structured model, they consider the treatment of HIV/AIDS with reverse-transcriptase – and protease – inhibitors, and how this treatment contributes to the evolution of resistance and rates of viral replication. Antimicrobial chemotherapy, heterogeneity and resistance are also the subjects of the chapter by David Smith and his collaborators. In their case, the drugs are antibacterial (antibiotics) rather than antiviral; the heterogeneity is both spatial and in the extent to which the drugs are employed; and the focus is the epidemiology and evolution of resistance in communities of hosts rather than in individual treated patients.

Investigators studying the evolution of infectious disease are almost invariably adaptationists; they assume that selection in the host, parasite or both populations is responsible for the virulence of the parasite and for maintaining genetic diversity in the parasite population. In their article on the serological diversity of the rhinoviruses responsible for the common cold, William Koppelman and Frederick Adler consider the neutral, null hypothesis alternative – that the 100 or so serotypes of Rhinoviruses responsible for colds are consequences of a high mutation rate and genetic drift rather than immune-mediated selection. Rhinoviruses and the cross-immunity they engender are also stage center in the chapter by Alun L. Lloyd and Dominik Wodarz, but the focus of their investigation of these ubiquitous and annoying, albeit rarely lethal, viruses is chemotherapy and the contribution of the host immune response to the evolution of resistance to the antiviral drugs employed.

This collection can be and I believe should be seen as a testimony to the work of Roy Anderson and Robert May. While they are not the discoverers of infectious diseases (at least I don’t think they are) or even the first to use mathematical models to investigate them, their research more than that of any other investigators has been responsible for the renaissance (epidemic) in the use of models for studying infectious diseases and their control. The research reported in almost all of the chapters in this volume have antecedents in Anderson and May’s work. While there is no formal dedication to them in the front matter of this volume, that dedication is where it really counts. The contributions of either Robert May and/or Roy Anderson are acknowledged at least once in every chapter and now, appropriately, in this Preface.

Enjoy,
Bruce R. Levin
Atlanta, September 2005
This page intentionally left blank
Editors' introduction

The goal of this volume is to show how to use mathematical tools to understand the evolution of infectious diseases. Inspiration for this project comes from work of the DIMACS Working Group on Genetics and Evolution of Pathogens, which is organized under the auspices of DIMACS' Special Focus on Computational and Mathematical Epidemiology.

This volume is divided into two sections: Model Infrastructure and Applications to Specific Diseases. Section I discusses the impact on disease evolution of various factors, including spatial structure, transient dynamics, coupling of within-host and between-host dynamics, heterogeneity in host populations, and drug resistance. Section II is concerned with investigations associated with specific infectious diseases such as rhinovirus, HIV/AIDS, tuberculosis, and malaria.

We thank Bruce Levin for his excellent Preface. We also express our gratitude to members of DIMACS' staff who kindly helped with the support of the workshop and the preparation of this volume. The leadership of Fred Roberts in developing the multi-year epidemiology program has been an inspiration to many researchers, and this volume owes its existence to his efforts. We also thank all the Purdue Mathematics Department for providing technical support. Finally, we thank the authors for their outstanding contributions.

The workshop and the preparation of this volume were partially supported by an NSF grant to DIMACS, and by NSF and James S. McDonnell Foundation grants to ZF.

Zhilan Feng (Purdue University)
Ulf Dieckmann (International Institute for Applied Systems Analysis, Austria)
Simon Levin (Princeton University)
Titles in This Series

71 Zhilan Feng, Ulf Dieckmann, and Simon Levin, Editors, Disease Evolution: Models, Concepts, and Data Analyses
70 James Abello and Graham Cormode, Editors, Discrete Methods in Epidemiology
69 Siemion Fajtlowicz, Patrick W. Fowler, Pierre Hansen, Melvin F. Janowitz, and Fred S. Roberts, Editors, Graphs and Discovery
68 A. Ashikhmin and A. Barg, Editors, Algebraic Coding Theory and Information Theory
67 Ravi Janardan, Michiel Smid, and Debasis Dutta, Editors, Geometric and Algorithmic Aspects of Computer-Aided Design and Manufacturing
66 Piyush Gupta, Gerhard Kramer, and Adriaan J. van Wijngaarden, Editors, Advances in Network Information Theory
65 Santosh S. Vempala, The Random Projection Method
64 Melvyn B. Nathanson, Editor, Unusual Applications of Number Theory
63 J. Nešetřil and P. Winkler, Editors, Graphs, Morphisms and Statistical Physics
62 Gerard J. Foschini and Sergio Verdú, Editors, Multiantenna Channels: Capacity, Coding and Signal Processing
61 M. F. Janowitz, F.-J. Lapointe, F. R. McMorris, B. Mirkin, and F. S. Roberts, Editors, Bioconsensus
60 Saugata Basu and Laureano Gonzalez-Vega, Editors, Algorithmic and Quantitative Real Algebraic Geometry
59 Michael H. Goldwasser, David S. Johnson, and Catherine C. McGeoch, Editors, Data Structures, Near Neighbor Searches, and Methodology: Fifth and Sixth DIMACS Implementation Challenges
58 Simon Thomas, Editor, Set Theory: The Hajnal Conference
57 Eugene C. Freuder and Richard J. Wallace, Editors, Constraint Programming and Large Scale Discrete Optimization
56 Alexander Barg and Simon Litsyn, Editors, Codes and Association Schemes
55 Ding-Zhu Du, Panos M. Pardalos, and Jie Wang, Editors, Discrete Mathematical Problems with Medical Applications
54 Erik Winfree and David K. Gifford, Editors, DNA Based Computers V
53 Nathaniel Dean, D. Frank Hsu, and R. Ravi, Editors, Robust Communication Networks: Interconnection and Survivability
52 Sanguthevar Rajasekaran, Panos Pardalos, and D. Frank Hsu, Editors, Mobile Networks and Computing
51 Pierre Hansen, Patrick Fowler, and Maolin Zheng, Editors, Discrete Mathematical Chemistry
50 James M. Abello and Jeffrey Scott Vitter, Editors, External Memory Algorithms
49 Ronald L. Graham, Jan Kratochvíl, Jaroslav Nešetřil, and Fred S. Roberts, Editors, Contemporary Trends in Discrete Mathematics
48 Harvey Rubin and David Harlan Wood, Editors, DNA Based Computers III
47 Martin Farach-Colton, Fred S. Roberts, Martin Vingron, and Michael Waterman, Editors, Mathematical Support for Molecular Biology
46 Peng-Jun Wan, Ding-Zhu Du, and Panos M. Pardalos, Editors, Multichannel Optical Networks: Theory and Practice
45 Marios Mavronicolas, Michael Merritt, and Nir Shavit, Editors, Networks in Distributed Computing
44 Laura F. Landweber and Eric B. Baum, Editors, DNA Based Computers II
43 Panos Pardalos, Sanguthevar Rajasekaran, and José Rolim, Editors, Randomization Methods in Algorithm Design
42 Ding-Zhu Du and Frank K. Hwang, Editors, Advances in Switching Networks
TITLES IN THIS SERIES

41 David Aldous and James Propp, Editors, Microsurveys in Discrete Probability
40 Panos M. Pardalos and Dingzhu Du, Editors, Network Design: Connectivity and Facilities Location
39 Paul W. Beame and Samuel R Buss, Editors, Proof Complexity and Feasible Arithmetics
38 Rebecca N. Wright and Peter G. Neumann, Editors, Network Threats
37 Boris Mirkin, F. R. McMorris, Fred S. Roberts, and Andrey Rzhetsky, Editors, Mathematical Hierarchies and Biology
36 Joseph G. Rosenstein, Deborah S. Franzblau, and Fred S. Roberts, Editors, Discrete Mathematics in the Schools
35 Dingzhu Du, Jun Gu, and Panos M. Pardalos, Editors, Satisfiability Problem: Theory and Applications
34 Nathaniel Dean, Editor, African Americans in Mathematics
33 Ravi B. Boppana and James F. Lynch, Editors, Logic and Random Structures
32 Jean-Charles Grégoire, Gerard J. Holzmann, and Doron A. Peled, Editors, The SPIN Verification System
31 Neil Immerman and Phokion G. Kolaitis, Editors, Descriptive Complexity and Finite Models
30 Sandeep N. Bhatt, Editor, Parallel Algorithms: Third DIMACS Implementation Challenge
29 Doron A. Peled, Vaughan R. Pratt, and Gerard J. Holzmann, Editors, Partial Order Methods in Verification
28 Larry Finkelstein and William M. Kantor, Editors, Groups and Computation II
27 Richard J. Lipton and Eric B. Baum, Editors, DNA Based Computers
26 David S. Johnson and Michael A. Trick, Editors, Cliques, Coloring, and Satisfiability: Second DIMACS Implementation Challenge
25 Gilbert Baumslag, David Epstein, Robert Gilman, Hamish Short, and Charles Sims, Editors, Geometric and Computational Perspectives on Infinite Groups
23 Panos M. Pardalos, David I. Shalloway, and Guoliang Xue, Editors, Global Minimization of Nonconvex Energy Functions: Molecular Conformation and Protein Folding
22 Panos M. Pardalos, Mauricio G. C. Resende, and K. G. Ramakrishnan, Editors, Parallel Processing of Discrete Optimization Problems
21 D. Frank Hsu, Arnold L. Rosenberg, and Dominique Sotteau, Editors, Interconnection Networks and Mapping and Scheduling Parallel Computations
20 William Cook, László Lovász, and Paul Seymour, Editors, Combinatorial Optimization
19 Ingemar J. Cox, Pierre Hansen, and Bela Julesz, Editors, Partitioning Data Sets
17 Eric Sven Ristad, Editor, Language Computations
16 Panos M. Pardalos and Henry Wolkowicz, Editors, Quadratic Assignment and Related Problems

For a complete list of titles in this series, visit the AMS Bookstore at www.ams.org/bookstore/.
Infectious diseases are continuing to threaten humankind. While some diseases have been controlled, new diseases are constantly appearing. Others are now reappearing in forms that are resistant to drug treatments. A capacity for continual re-adaptation furnishes pathogens with the power to escape our control efforts through evolution. This makes it imperative to understand the complex selection pressures that are shaping and reshaping diseases. Modern models of evolutionary epidemiology provide powerful tools for creating, expressing, and testing such understanding.

Bringing together international leaders in the field, this volume offers a panoramic tour of topical developments in understanding the mechanisms of disease evolution. The volume’s first part elucidates the general concepts underlying models of disease evolution. Methodological challenges addressed include those posed by spatial structure, stochastic dynamics, disease phases and classes, single- and multi-drug resistance, the heterogeneity of host populations and tissues, and the intricate coupling of disease evolution with between-host and within-host dynamics. The book’s second part shows how these methods are utilized for investigating the dynamics and evolution of specific diseases, including HIV/AIDS, tuberculosis, SARS, malaria, and human rhinovirus infections.

This volume is particularly suited for introducing young scientists and established researchers with backgrounds in mathematics, computer science, or biology to the current techniques and challenges of mathematical evolutionary epidemiology.