

Contents

Preface — V

Chapter 1

Inference of the acting factors of evolution and basic evolutionary parameters from sequence data — 1

- 1.1 Mechanism of HIV diversity and the estimate of average selection coefficient — 2
 - 1.1.1 Sequence diversity in protease gene: data analysis — 3
 - 1.1.2 Model 1: deterministic adaptation of virus in an individual — 6
 - 1.1.3 Model 2: chain of single-clone transmission — 9
 - 1.1.4 Model 3: coinfection from independent sources — 11
 - 1.1.5 Probability of coinfection estimated for HIV — 12
 - 1.1.6 Model 4: individual variation in wild type due to MHC subtypes — 12
 - 1.1.7 Discussion — 14
 - 1.1.8 Mathematical derivations — 15
 - 1.1.8.1 Virus evolution in an individual — 15
 - 1.1.8.2 Chain of single-clone transmission — 16
 - 1.1.8.3 Coinfection from independent sources — 17
 - 1.1.8.4 Coinfection from the same source — 18
 - 1.1.8.5 Estimation of the value of q for HIV — 19
 - 1.1.8.6 Approximations — 20
- 1.2 Estimate of the effective population size — 21
 - 1.2.1 One-locus model of stochastic evolution — 23
 - 1.2.2 Three regimes of evolution — 24
 - 1.2.3 Two-locus model and linkage disequilibrium test — 27
 - 1.2.4 Estimation of the effect of recombination — 30
 - 1.2.5 Robustness to approximations — 31
 - 1.2.6 Discussion — 32
- 1.3 Estimate of recombination rate — 33
 - 1.3.1 Methods — 35
 - 1.3.1.1 Model of population — 35
 - 1.3.1.2 Linkage disequilibrium measures — 36
 - 1.3.1.3 Patient data — 36
 - 1.3.2 Simulation of HIV adaptation — 36
 - 1.3.3 Estimation of recombination rate and average selection coefficient — 39
 - 1.3.4 Discussion — 42

Chapter 2**Inference of fitness landscape from sequence data — 44**

- 2.1 Universal evolutionary footprint of epistasis — **44**
 - 2.1.1 Model of stochastic evolution with epistasis — **46**
 - 2.1.2 The footprint of epistasis for a single interacting pair in a long genome — **47**
 - 2.1.3 The long genome of isolated pairs — **50**
 - 2.1.4 Full compensation and UFE interval — **53**
 - 2.1.5 Effects of network topology — **55**
 - 2.1.6 Discussion — **59**
 - 2.1.7 Mathematical derivations — **60**
 - 2.1.7.1 Isolated pairs (Section for 2.1.3) — **61**
 - 2.1.7.2 Double arches — **64**
 - 2.1.7.3 Triple arches — **66**
 - 2.1.7.4 Long chain — **69**
 - 2.1.7.5 Large binary tree — **72**
 - 2.1.7.6 Double arches with unequal interactions — **75**
- 2.2 Detection of epistatic pairs in a single population: mission impossible — **78**
 - 2.2.1 Epistatic pairs have a distinct signature in a narrow time window — **79**
 - 2.2.2 Results are robust to the combinations of LD measures — **82**
 - 2.2.3 Parameter sensitivity analysis confirms the results — **83**
 - 2.2.4 Population divergence creates strong linkage effects — **84**
 - 2.2.5 The use of multiple populations rescues epistatic signature — **85**
 - 2.2.6 Discussion — **86**
- 2.3 Detection of epistasis by the three-way correlation method — **87**
 - 2.3.1 Simulation model to generate sequences for testing the method — **88**
 - 2.3.2 First step: averaging over populations — **88**
 - 2.3.3 Second step: three-way correlation — **90**
 - 2.3.4 Analytic test of the method — **91**
 - 2.3.4.1 Derivation for topology without loops — **91**
 - 2.3.4.2 Derivation for topology with closed squares — **93**
 - 2.3.5 Application to influenza A virus — **97**
 - 2.3.5.1 A primary mutation and compensatory sites — **99**
 - 2.3.5.2 Structural interpretation — **99**
 - 2.3.6 Discussion — **100**
- 2.4 Estimation of selection coefficients from DNA data — **102**
 - 2.4.1 Experimental distribution of selection coefficients — **104**
 - 2.4.2 Model — **105**
 - 2.4.3 Monte-Carlo simulation of genetic evolution — **105**
 - 2.4.4 Analytic derivation of universal DFE — **107**

- 2.4.4.1 Early evolution — **107**
- 2.4.4.2 Traveling wave regime — **108**
- 2.4.4.3 Quasi-equilibrium argument — **110**
- 2.4.4.4 Monte-Carlo test confirms results — **111**
- 2.4.5 Estimating selection coefficients from a sequence set — **112**
- 2.4.6 Discussion — **113**

Chapter 3

Evolutionary role of a trait — 116

- 3.1 Evolution of HIV toward AIDS — **116**
 - 3.1.1 Model of virus dynamics — **119**
 - 3.1.1.1 Model 1: no immune response — **119**
 - 3.1.1.2 Model 2: CD8 T-cell immune response — **120**
 - 3.1.2 The rate of HIV adaptation to a host — **121**
 - 3.1.3 Predicting the speed of progression to AIDS — **122**
 - 3.1.3.1 Negative correlation between time to AIDS and virus load — **123**
 - 3.1.3.2 Parameter estimates for AIDS prognosis — **124**
 - 3.1.4 Alternative models of progression to AIDS: impaired homeostasis — **126**
 - 3.1.5 Discussion — **128**
- 3.2 An evolutionary role for HIV latency — **130**
 - 3.2.1 Mathematical models of lentiviral transmission — **133**
 - 3.2.2 Effect of latency depends on both mucosal and systemic infection — **136**
 - 3.2.2.1 Latency increases the probability of systemic infection — **136**
 - 3.2.2.2 Latency decreases the inoculum — **137**
 - 3.2.3 Probability of latency near 50% optimizes transmission — **139**
 - 3.2.3.1 Result is robust to the presence of nonlatent routes — **140**
 - 3.2.3.2 Experiments in cell culture and primates confirm predictions — **141**
 - 3.2.4 Why latency in patients is rare: including the immune response — **141**
 - 3.2.4.1 Models incorporating the immune response fit patient data — **144**
 - 3.2.5 Prediction: depletion of CD8⁺ T cells in macaques will increase the latent reservoir — **146**
 - 3.2.6 Discussion — **147**
 - 3.2.6.1 Latency and other mechanisms of initial viral survival — **148**
 - 3.2.6.2 Potential therapy: suppressing latent reactivation in mucosa — **148**

- 3.2.6.3 Implications for alternate antiviral therapy approaches — **148**
- 3.2.7 Derivations for stochastic dynamics in mucosa ($R_0^{\text{muc}} > 1$) — **148**
- 3.2.8 Two-compartment model ($R_0^{\text{muc}} < 1$ and $R_0^{\text{LT}} > 1$) — **153**
- 3.2.9 Parameters sensitivity analysis — **155**
- 3.2.10 Robustness to model variations — **156**
- 3.3 Recombination and the optimal mutation rate of polio virus — **160**
 - 3.3.1 Experiments on strains with altered recombination and mutation rates — **163**
 - 3.3.1.1 Mutation D79H decreases recombination rate 10-fold — **163**
 - 3.3.1.2 Mutation D79H does not alter mutation rate — **165**
 - 3.3.1.3 Mutation D79H slows down viral adaptation in culture — **166**
 - 3.3.1.4 Mutation D79H does not affect fitness in culture — **166**
 - 3.3.1.5 Double mutants impair viral adaptation in mice — **167**
 - 3.3.2 Mathematical modeling and fitting data — **169**
 - 3.3.2.1 Adaptation rate is maximal near wild-type mutation rate — **169**
 - 3.3.2.2 Recombination and mutation affect adaptation independently — **173**
 - 3.3.2.3 Mathematical modeling fits the mice survival data — **175**
 - 3.3.2.4 Sensitivity to parameters — **175**
 - 3.3.3 Optimal mutation rate replaces the concept of “error catastrophe” — **176**
 - 3.3.4 Steady-state derivation — **178**
 - 3.3.5 The point of no adaption in a short-term evolution — **179**

Chapter 4

Evolutionary escape from an opposing species (Red Queen effect) — **181**

- 4.1 Evolution of antibody epitopes of influenza virus — **181**
 - 4.1.1 Model of influenza transmission in a population — **183**
 - 4.1.1.1 Strain-structured epidemiological model — **183**
 - 4.1.1.2 Including mutation and random genetic drift — **184**
 - 4.1.2 Two-component traveling wave — **186**
 - 4.1.2.1 Density of recovered individuals — **186**
 - 4.1.2.2 Moving fitness landscape — **187**
 - 4.1.3 Connecting to the evolution theory — **189**
 - 4.1.3.1 Antigenic diversity and the speed of evolution — **189**
 - 4.1.3.2 Time to the most recent common ancestor — **192**
 - 4.1.4 Comparison with data on influenza A — **193**
 - 4.1.5 Robustness to additional dimensions and old memory — **194**
 - 4.1.6 Discussion — **195**

- 4.1.7 Analytic derivation for the 1D model — **197**
 - 4.1.7.1 Traveling wave solution — **197**
 - 4.1.7.2 Effective selection coefficient — **198**
 - 4.1.7.3 Wave speed — **198**
 - 4.1.7.4 Time to the most recent ancestor — **200**
 - 4.1.7.5 Comparison to a previous 1D model of influenza evolution — **200**
- 4.1.8 Multidimensional antigenic space — **201**
 - 4.1.8.1 Two dimensions with one antigenic coordinate — **201**
 - 4.1.8.2 Two antigenic coordinates — **202**
 - 4.1.8.3 Many dimensions are equivalent to tree topology — **202**
- 4.1.9 Approximations — **203**
- 4.2 Evolution of CD8 T-cell epitopes of HIV — **204**
 - 4.2.1 Model of HIV dynamics in the presence of multiple epitopes — **206**
 - 4.2.1.1 Simplified model to study the order of escape mutations — **207**
 - 4.2.2 Simulations of the dynamics of antigenic escape — **210**
 - 4.2.2.1 Phases of HIV infection — **210**
 - 4.2.2.2 The determinants of the escape rate of a mutant strain — **210**
 - 4.2.2.3 The trajectory of escape mutations in the cost–benefit plane — **211**
 - 4.2.3 Correlation between escape cost and benefit in Pol gene is explained — **213**
 - 4.2.4 Three patterns of antigenic escape in an epitope with two sites — **214**
 - 4.2.5 Approximations — **217**
 - 4.2.6 Derivation of the steady state, escape rate, and clone contraction — **220**
 - 4.2.6.1 Contraction of CTL clones after escape — **221**
 - 4.2.7 Leapfrog pattern of escape — **222**
 - 4.2.7.1 Necessary conditions — **222**
 - 4.2.7.2 The phase region of leapfrog pattern — **223**
 - 4.2.8 Relationship between Δr and HLA binding loss from three experiments — **224**
 - 4.2.8.1 Finding Δr from virus dynamics in the presence of CTL — **224**
 - 4.2.8.2 HLA-binding loss ΔB from competition binding assay — **224**
 - 4.2.8.3 ΔB and Δr from the measurement of CTL activity — **224**
- 4.3 Stability of HIV in the presence of defective interference particles in a cell and a host — **226**
 - 4.3.1 Missing lentiviral DIP and evolutionary stability — **227**

- 4.3.2 DIP interference with HIV by competition for genomic RNA leads to divergent evolution — **229**
- 4.3.3 An alternative mechanism of DIP interference: protein stealing — **232**
 - 4.3.3.1 The single-cell model with capsid stealing — **233**
 - 4.3.3.2 The individual-host model — **235**
 - 4.3.3.3 Parameter values — **237**
 - 4.3.3.4 Dynamically stable suppression of HIV at the host level — **237**
 - 4.3.3.5 HIV suppression is due to high multiplicity of DIP infection — **239**
- 4.3.4 Testing evolutionary stability: the effective selection coefficient — **240**
- 4.3.5 Discussion — **242**
- 4.3.6 Derivation of HIV and DIP loads — **245**
 - 4.3.6.1 Single cell — **245**
 - 4.3.6.2 Small waste parameter, $K \ll 1$ — **247**
 - 4.3.6.3 Individual host — **247**
 - 4.3.6.4 Dynamic stability of DIP in a host — **249**
- 4.3.7 Selection coefficient of HIV in terms of intracellular parameters — **250**
 - 4.3.7.1 Change in waste parameter — **252**
 - 4.3.7.2 Change in capsid-to-genome production ratio — **253**
- 4.3.8 Estimate of parameters κ and η in infected individuals — **253**
- 4.4 Stability of HIV in the presence of defective interference particles in a host population: evolutionary conflicts and the “tragedy of the commons” — **255**
 - 4.4.1 Three-scale model of HIV and DIP dynamics — **257**
 - 4.4.2 Conditions of HIV and DIP coexistence and HIV suppression in high-risk populations — **261**
 - 4.4.3 HIV escape mutants that are resistant to DIP face conflicting selection pressures — **262**
 - 4.4.4 Evolutionary conflicts prevent the establishment of DIP-resistant HIV mutants — **264**
 - 4.4.5 Discussion — **266**
 - 4.4.5.1 Cheaters and the “tragedy of the commons” — **267**
 - 4.4.5.2 Model assumptions and limitations — **268**
 - 4.4.5.3 Frequency-dependent selection on the population scale — **269**
 - 4.4.5.4 DIP as a resistance-proof therapy? — **269**
 - 4.4.6 Derivation of DIP and HIV prevalence in a population of hosts — **270**
 - 4.4.6.1 Link to the individual-host and single-cell scales — **270**
 - 4.4.6.2 Steady state — **271**
 - 4.4.6.3 Condition for DIP spread and stability in a population — **273**

- 4.4.7 Derivation of the evolutionary stability of HIV in the presence of DIP — **274**
- 4.4.8 Robustness to model variations — **276**
 - 4.4.8.1 T-cell division and homeostasis — **276**
 - 4.4.8.2 DIP preinfects individuals — **276**
 - 4.4.8.3 Sensitivity to κ — **278**
 - 4.4.8.4 Timing of HIV transmission — **278**
 - 4.4.8.5 Other approximations — **278**

References — 279

Index — 303

