DIMACS Series in Discrete Mathematics
and Theoretical Computer Science
Volume 71, 2006

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

ABSTRACT. Over 100 serotypes of rhinoviruses, one of the primary causes of the common cold, co-circulate in the human population. This high diversity makes it effectively impossible to develop a vaccine, even for those at risk of complications due to asthma or cystic fibrosis. Is the high mutation rate of these viruses sufficient to explain this diversity? We use parameters estimated from the literature to study whether immune interactions between different rhinovirus serotypes also play an important role in maintaining diversity. Our mathematical models indicate that high mutation rates alone may well be responsible for the observed levels of diversity. However, careful studies of a few communities have found that some serotypes persist for many years, in conflict with the predictions of the simplest models, hinting that there might be more to the story than is yet known.

1. Introduction

Diversity is usually regarded as an ecological question and is rarely considered in the realm of virology [1]. However, the immense diversity of rhinoviruses, the most frequent cause of the common cold, makes development of treatments and a vaccine virtually impossible [15]. We use mathematical and computer simulation models to attempt to understand the maintenance of this high diversity.

There are currently 102 serotypes of the human rhinovirus (HRV) that co-circulate in the human population, divided into type A (with 76 serotypes), type B (25 serotypes), and a single type which clusters with enteroviruses [30]. HRV is a genus in the family *Picornaviridae* and shares the characteristics of other genera within this family (e.g. enterovirus, poliovirus, foot and mouth disease virus, and hepatitis A virus). These characteristics include a non-enveloped icosahedral capsid and a single-stranded positive-sense RNA genome of approximately 7.2 kb. Viruses are spherical in shape with a diameter of 25–30 nm [8]. Within *Picornaviridae*, HRV is genetically most closely related to the enteroviruses [21]. The enteroviruses can tolerate a wider range of pH and therefore multiply mainly in the alimentary tract while HRV multiplies in the nasal epithelium [31]. HRV serotypes cause as

Key words and phrases. Biodiversity, Rhinovirus, Neutral theory, Cross-immunity.
First author was supported by IGERT grant NSF DGE-0217424.

many as 40% of upper respiratory illnesses, or "common colds" [17]. They have also been associated with bacterial sinusitis and otitis media [8], and appear in over 80% of acute asthma episodes in children [14].

Once a sufficient dose, as small as 1–30 particles, has deposited itself in the nasal passages it is transported back to the adenoid area [9, 18]. Then the virus attaches to a receptor on the surface of the nasal cells. HRV-A serotypes attach to the Intercellular Adhesion Molecule-1 (ICAM-1) receptor while serotypes in HRV-B attach to the low density lipoprotein receptor (LDLR) [6]. This attachment leads to the infection process. The incubation period is 8–12 hours with the peak of infection occurring in 36–72 hours [15].

The goal of this paper is to develop a model to explain the observed diversity of HRV using characteristics from its pathology and evolution. The maintenance of diversity may be explained by the high mutation rate of HRV, the cross-reactivity between serotypes, or immunodominance from previous infections.

RNA viruses typically have high rates of mutation and can be thought of as a quasi-species [25]. It has been shown that the mutation rate for HRV is 0.67 mutations per genome, or 10^{-4} per nucleotide per replication [10]. This high rate of mutation may lead to evolution of serotypes. In 1969 it was suggested that HRV-51 had undergone sufficient variation in two to four years to produce a new serotype [32]. A recent experiment found that HRV-17 could quickly evolve to escape the antisera used to identify serotypes, while 15 other serotypes showed little or no such evolution [28].

Cross-reactivity between serotypes led to the proposal that groups of HRV share common epitopes. In 1975, Cooney et al. [7] analyzed the published literature and in combination with their own data found evidence for a strong relationship between serotypes. They found that a serotype will elicit a heterotypic antibody response from an average of 3.75 other serotypes.

A rhinovirus in the human body presents a large number of epitopes to the immune system. The immune response focuses on only a few of the many potential epitopes, a process called immunodominance [12]. The sequence in which the host encounters antigenic variants influences the specificity of the immune response. This process was termed *original antigenic sin* [3]. For example, if the primary exposure in the host was with antigen $\mathbf{A}$ which elicited an immune response $\mathbf{a}$, a secondary exposure to a related antigen $\mathbf{A}'$ could stimulate the same immune response $\mathbf{a}$. A rapid response from the memory B or T cells may keep the antigen density below a threshold required for stimulation of naive B or T cells [12]. Although there may be a better response to a particular antigen, the immune system does not respond with it because the invasion has been regulated by existing memory cells.

A previous model by Pease [29] considered an evolutionary epidemic model in which a person would go from having immunity to a virus back to being susceptible after the virus had changed genetically, and hence immunologically. Andreasen et al. [4] developed an influenza model to consider partial cross-protection among strains. In the model, the immunity history of the host population was followed so that invading strains could be analyzed against it. This allowed for the existence of a multi-strain endemic equilibrium. Cross-immunity has been the focus of other papers as well [5, 13]. In particular, recent work has included cross-immunity, mutation, and demographic stochasticity in a study of the persistence and dynamics of a viral disease [1].

Other groups have developed a mathematical model of immunodominance in HIV-1 infections [**26, 27**]. These models focus on competition for stimulation by epitopes between immune cell lineages, concluding that the most efficient predator (immune cell lineage) reduces the prey (epitopes) down to a level where less efficient predators cannot survive. Models of immunological interactions between different dengue virus serotypes allow modelers to identify the immunological distance at which serotypes can stably coexist [**22**].

Hubbell [**20**] developed a neutral theory for species richness and relative species abundance. The theory is neutral because individual organisms do not differ in their demographic parameters, and there are no competitive interactions other than maintenance of constant population size. The theory proposes that only three parameters are needed to determine diversity: the fundamental biodiversity number (which includes metacommunity size and speciation rate), the probability of immigration, and the local community size.

Our aim is to estimate the relevant parameters for HRV, and to test what effects cross-immunity and immunodominance have on maintained diversity in comparison to the neutral case. As a test, we compare our model results with a detailed study of the small town of Tecumseh, Michigan [**24**], comparing the observed diversity and serotype frequency during the two time periods covered by the study, and checking whether the overlap in serotypes between the two periods is realistic.

2. The simulation

Rhinovirus serotypes are placed on a d-dimensional hyper-torus, with periodic boundaries of length 2. Each serotype is represented by a single hypercube with sides of length one, creating space for 2^d serotypes. The maximum distance between serotypes is $\sqrt{d}$, relative to the difference between serotypes. The difference between the most similar serotypes is roughly 10% nucleotide divergence [**28**], thus a dimension of $d = 10$ is consistent with the maximum divergence of 34 - 41% observed [**30**]. This dimension leaves room for $2^{10} = 1024$ possible serotypes, an order of magnitude more than have been observed.

Our model tracks the infectious and immune state of each individual in a population of size N. Hosts are born and die at equal and constant rates, set to give a mean lifespan of 70 years. Newborn hosts have no HRV immunity. Every time step (one day), infected hosts have a probability γ of clearing the infection. A value of $\gamma = 0.1$ is consistent with a mean infection duration of 10 days [**24**]. Upon recovery, these individuals gain immunity to the infecting serotype.

Susceptible individuals contract the illness through contact with a single pool of viruses, rather than direct contact with other individuals. Such contacts occur with a probability α. The pool of viruses is created by mixing a fraction $1 - m$ from infected individuals in the population and a fraction m of random viruses to represent immigration. We chose values of $m = 0$ and $m = 0.01$ to study the effects of immigration, assuming in the latter case that an individual re-enters the local area on about 1% of days, or four times per year. The value of α is adjusted until the disease prevalence matches a pre-determined target based on a frequency of approximately 0.5 HRV infections per person per year with a duration of 10 days [**17**]. If contact occurs, the potentially infecting strain is a mutated version of a strain chosen randomly from the pool. The mutation is in a random

direction in the d dimensional space with a magnitude of μ times a standard log-normal distribution (with mean and variance equal to 1). The value of μ is chosen so that a new serotype is generated after about 50 infections [32], meaning that most mutations do not generate a new serotype. Given a mutation rate of 10^{-4} per nucleotide per replication [10], we thus assume 20 replications per infection to achieve the 10% sequence divergence between serotypes.

Successful infection occurs if the exposed individual does not have immunity from previous infections (Figure 1). Previous infections provide partial immunity out to a distance of x_i, with the probability of infection being

$$\text{Probability of infection} = \frac{\text{distance to nearest prior infection}}{x_i}.$$

The parameter x_i is varied in our simulations, but should take on values slightly larger than 1 to correspond to the observed degree of cross-reactivity. Immunity is determined by distance without explicit regard to serotype. Unsuccessful infections are included in the exposed individual's immune history when immunodominance is off, and are not included when it is on. In the latter case, we assume that the successful response was based entirely on the existing epitopes.

TABLE 1. Variables, parameters, and functions in the simulation

Parameters		
Symbol	Meaning	Value or values
N	Population size of hosts	1000 – 8000
p	Prevalence of rhinovirus infections	0.02 – 0.1
d	Dimension of genetic space	7 or 10
x_i	Cross-immunity distance	0.0 – 2.0
μ	Standard deviation of mutation distance	0.06
δ	Birth rate of hosts per day	$1/(70 \cdot 365)$
γ	Recovery rate of hosts per day	0.1
α	Contact rate of hosts with virus pool	Computed
m	Probability of immigration of infected host	0 or 0.01

3. Simulation results

How well do these parameter estimates (summarized in Table 1) predict the actual diversity of HRV? First, we compare our results with the neutral theory [20]. The fundamental biodiversity parameter θ is the product of the number $J = pN$ of infected individuals and twice the effective mutation rate ν, or

$$\theta = 2\nu pN = 2\nu J.$$

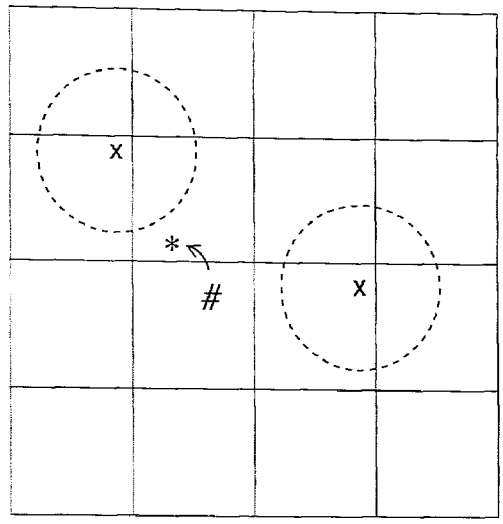

FIGURE 1. The structure of the simulation. An individual who started with the infection marked with # (displayed in a two-dimensional version of the serotype space) infects another individual by transmitting the mutated version indicated by a *. The new individual has previously had the two colds at locations marked with the x's. and has immunity (potentially only partial immunity) in the regions indicated by the dashed circles. The new infection lies outside these circles, and is thus successful. Because it has crossed one of the lines in the grid, it is designated as a new serotype.

The effective mutation rate ν includes the effects of both mutation and migration. The diversity D_ν is then

$$\begin{aligned} D_\nu &= \sum_{j=1}^{J} \frac{\theta}{\theta + j - 1} \\ &\approx \theta \int_{x=0}^{J-1} \frac{1}{\theta + x} dx \\ &= \theta \ln\left(\frac{\theta + J - 1}{\theta}\right) \\ &= 2\nu pN \ln\left(\frac{2\nu pN + pN - 1}{2\nu pN}\right) \\ &\approx 2\nu pN \ln\left(\frac{1 + 2\nu}{2\nu}\right). \end{aligned}$$

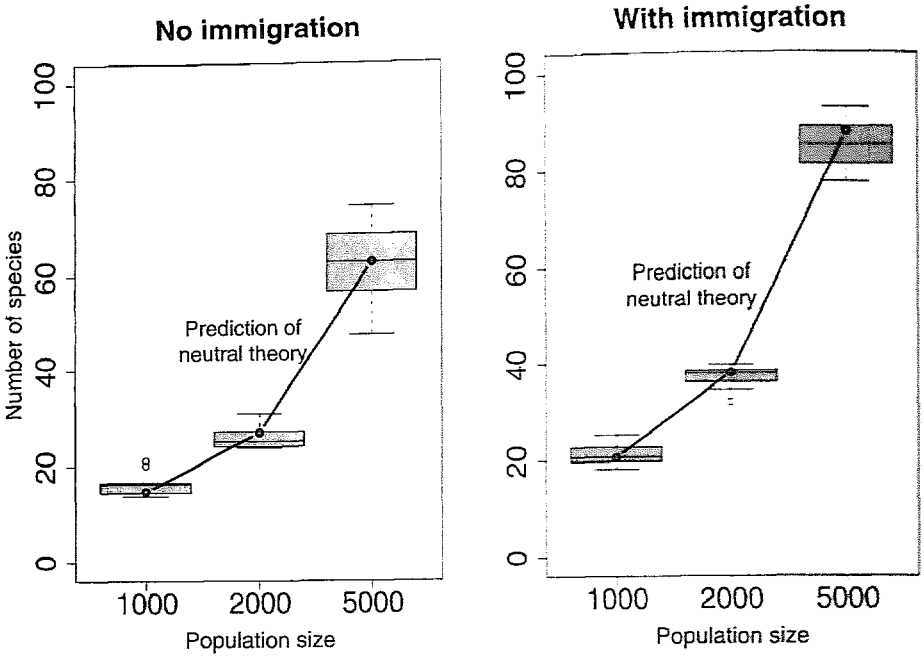

FIGURE 2. Comparison of neutral theory predictions with simulated diversity for the ranges of parameter values in Table 1 but with no cross-immunity ($x_i = 0$). To avoid disease die-out in the smaller populations, we used an unrealistically high prevalence of 10%. To match the actual number of infections carried by a person in their lifetime, we increased the birth and death rates by a factor of 7.0. **a.** With no immigration ($m = 0$). **b.** With 1% of infections from immigration ($m = 0.01$).

This model assumes that all new mutants form new serotypes, and our estimated dimension of $d = 10$ creates 1024 serotypes, a factor of 10 more than have been observed co-circulating in the human population.

We estimated a value of $\nu = 0.016$ from the simulations with $m = 0$ (fairly close to our target value of 0.02 based on requiring approximately 50 infections to create a new serotype). We used a value of $\nu = 0.016 + m$ in the cases with $m = 0.01$ because new migrants are essentially certain to be of a new serotype. In each case, the neutral theory predicts the overall diversity quite well in the absence of cross-immunity (Figure 2).

Cross-immunity generally increases diversity. For relatively large populations, the increase is maximized at an intermediate scale of cross-immunity ($x_i = 0.75$ to 1.0) roughly consistent with the scale of a serotype (Figure 3). Thus, the observed relatively limited cross-immunity between serotypes [7] may be at a level that maximizes the standing diversity of HRV. Contrary to our original hypothesis, immunodominance has no effect on the diversity maintained in the system (Figure 4).

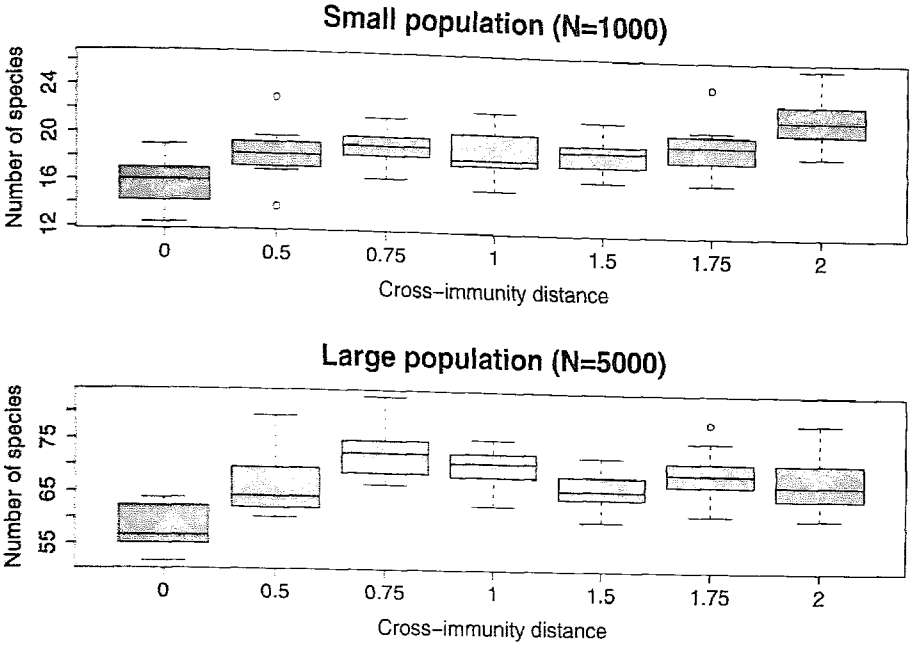

FIGURE 3. The effects of cross-immunity on diversity using the parameters in Figure 2 and $m = 0$. A cross-immunity distance of 2.0 corresponds to deriving partial cross-immunity to all cold serotypes from any infection. Bars which share the same label have serotype richnesses that are statistically indistinguishable (Tukey's honestly significant differences).

We then tested the model on a detailed data set from Tecumseh, Michigan, collected during the periods from 1966–1971 and 1976–1981 [**24**]. In this study, symptomatic individuals from the community were visited, and their infections sampled, identified and serotyped. There were 250 HRV samples taken during the first period and 194 during the latter. At that time, only 89 serotypes had been identified and fully 73 of these were identified at some point in the study. Our simulations were designed to match the size of this community (roughly 8000 people) with our other most reasonable guesses of parameter values. We randomly sampled infected individuals during two 6 year periods offset by 10 years from each other to match the design of the study.

The dominance-diversity plots match the data fairly closely, although the distributions with $m = 0.01$ have too many singletons (Figure 5). The observed diversity matches that seen in the Tecumseh study reasonably well, although the predicted diversities are too large when immigration is included at even a low level (Table 2).

However, the turnover during this time is far too large (the overlap in the last row of Table 2). In Tecumseh, most of the serotypes observed in the earlier samples reappeared in the later samples, in sharp contrast to our simulation results. The authors of the Tecumseh study argue that nearly all possible serotypes had been discovered [**24**]. There were 89 identifiable serotypes in 1987, and only 13

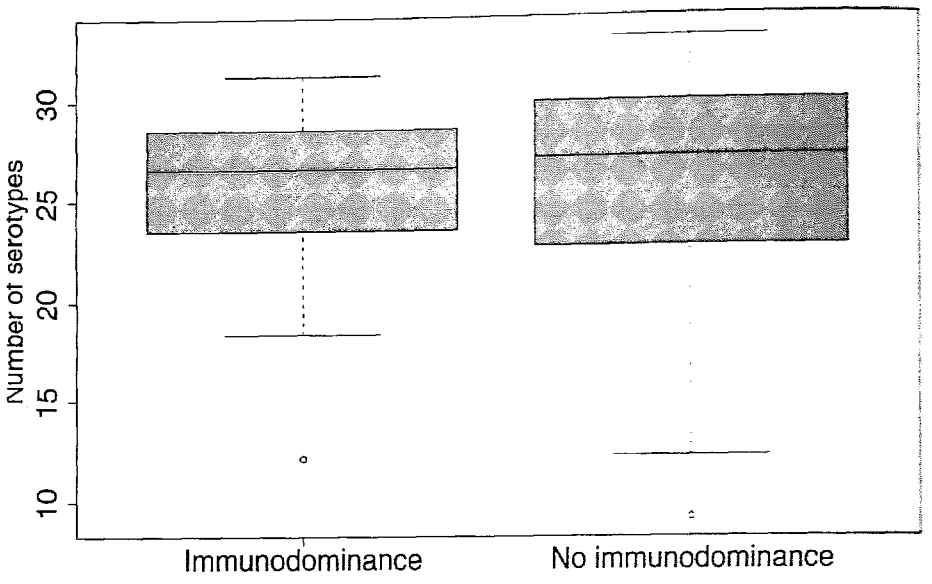

FIGURE 4. The lack of an effect of immunodominance using the parameters in Figure 2.

TABLE 2. Simulated and measured HRV serotype diversity and overlap: $d = 10$

	$x_i = 0$, $m = 0$	$x_i = 0$, $m = 0.01$	$x_i = 1$, $m = 0$	$x_i = 1$, $m = 0.01$	Measured
1966-71	45	88	73	115	62
1976-81	40	86	66	86	53
overlap	12	5	13	8	42

additional ones have been identified subsequently, at least 2 of which cross-react strongly with serotypes already identified [**23**], indicating that perhaps the numbers have saturated.

To test this hypothesis, we reran our simulation in seven dimensions, creating only 128 possible serotypes, much closer to the number of identified serotypes. Not surprisingly, we found a much larger proportion of overlap between the two sampling periods, more in line with the observed results (Table 3).

We used logistic regression [**19**] to test whether abundance in the earlier time period predicts presence in the later period. We find a highly significant effect ($p < 0.001$) with the actual data (after including the 16 absent serotypes with abundance zero), but no significant relationship in the samples from the simulations with any parameter values.

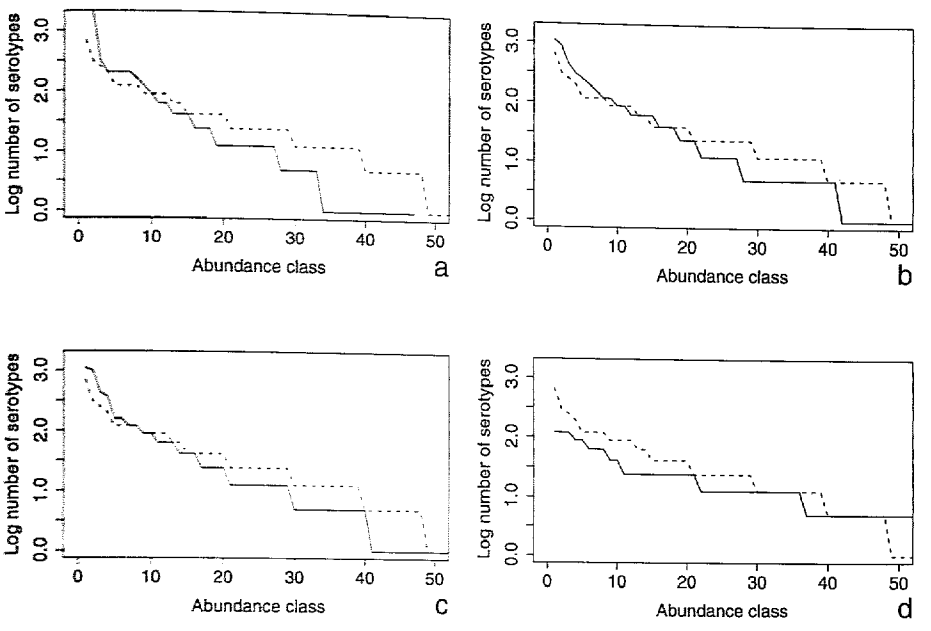

FIGURE 5. The predicted dominance-diversity curves for Tecumseh by running the simulation with $N = 8000$ and a more realistic prevalence of 10 days of infection per year (solid lines) compared with the actual data (dashed lines). The abundance rank of each serotype is based on the total number of individuals observed with that serotype. The Tecumseh study serotyped 250 individuals from 1966-1971. and we sampled the same number of individuals over the course of 6 years from the simulations. Parameter values are **a.** $x_i = 0$ and $m = 0$, **b.** $x_i = 0$ and $m = 0.01$, **c.** $x_i = 1$ and $m = 0$, and **d.** $x_i = 1$ and $m = 0.01$.

TABLE 3. Simulated and measured HRV serotype diversity and overlap: $d = 7$

	$x_i = 0$, $m = 0$	$x_i = 0$, $m = 0.01$	$x_i = 1$, $m = 0$	$x_i = 1$, $m = 0.01$	Measured
1966-71	29	69	56	70	62
1976-81	50	57	55	70	53
overlap	10	32	26	38	42

4. Discussion

We have found that a simple model with realistic parameters does a fairly good job of estimating the standing diversity of HRV. In general, the neutral theory well

approximates the results of simulations, although our effective mutation rate ν differed slightly from that estimated directly from the underlying process. Inclusion of competition through cross-immunity increases diversity by as much as 25%. For larger population sizes (5000 individuals in this study), maximum diversity is maintained when cross-immunity acts at an intermediate scale which roughly matches scale seen in nature. Immunodominance, which describes the immune response to failed infections, has no detectable effect on diversity.

Using parameters based on the detailed Tecumseh study [**24**], we find a good quantitative fit to the number of serotypes present. However, we find that the turnover between serotypes is far too fast unless we restrict the number of possible serotypes. If only approximately 100 serotypes are viable, it would be extremely interesting to understand the constraints operating on the sites which determine serotype [**16**]. Even in this case, however, the simulations differ from the actual measurements in showing no relationship between abundance in the first census and presence in the later census. If the observed relationship were due to a much lower mutation rate than we assume, the observed diversity could not be maintained. Alternatively, the relationship could be due to heterogeneity among the serotypes, either in detectability or in transmission.

Our model deals with disease transmission in only the simplest way. We neglect age-dependence of prevalence and duration of infection [**17**], seasonality [**14**], and heterogeneity in contact rates. As noted above, we also neglect differences among serotypes in their overall parameters and in their unstudied effects on people of different ages or immune functions. We do not consider the fact that immunity may well last no more than a few years [**8**], and neglect the possible importance of recombination [**2**].

More detailed study of these models needs to address the observed phylogenetic relationships among coexisting serotypes [**30**]. Most simply, the models could use the two types A and B as roughly independent replicates of the experiment. The simulation could be made more explicitly genetic, rather than taking place in an abstract serotype space, and the results could then be presented explicitly as phylogenetic trees. Comparison with related studies of influenza, which shows a very different pattern of replacement of strains through antigenic drift and antigenic shift, could be particularly useful [**11**].

Given the wealth of genetic and epidemiological data available on this widespread and important class of viral infections, we believe that the common cold can provide new insights into the applicability of ecological theories, and that ecological theory can provide new insights into the transmission and evolution of the common cold.

Acknowledgments

Thanks to Lissy Coley and Jim Keener for useful comments on the Master's Thesis from which this chapter was developed, and to two anonymous reviewers for constructive comments.

References

[1] L. J. Abu-Raddad and N. M. Ferguson, The impact of cross-immunity, mutation and stochastic extinction on pathogen diversity, *Proc. Roy. Soc. Lond. B* **271** (2004), 2431–2438.

[2] V. I. Agol. Recombination and other genomic rearrangements in picornaviruses, *Seminars in Virology* **8** (1997), 77–84.

[3] D. E. Anderson, M. P. Carlos, L. Nguyen, and J. V. Torres, Overcoming original (antigenic) sin. *Clinical Immunology* **101** (2001), 152–157.

[4] V. Andreasen, J. Lin, and S. A. Levin, The dynamics of cocirculating influenza strains conferring partial cross-immunity. *J. Math. Biol.* **35** (1997), 825–842.

[5] R. Antia, M. A. Nowak, and R. M. Anderson, Antigenic variation and the within-host dynamics of parasites. *Proc. Natl. Acad. Sci. USA* **93** (1996), 985–989.

[6] S. Blomqvist, C. Savolainen, L. Raman, M. Roivainen, and T. Hovi, Human rhinovirus 87 and enterovirus 68 represent a unique serotype with rhinovirus and enterovirus features, *J. Clinical Microbiology* **40** (2002), 4218–4223.

[7] M. K. Cooney, J. A. Wise, G. E. Kenny, and J. P. Fox, Broad antigenic relationships among rhinovirus serotypes revealed by cross-immunization of rabbits with different serotypes, *J. Immunology* **114** (1975), 635–639.

[8] R. B. Couch, Rhinoviruses, in (B. N. Fields, D. M. Knipe, and P. M. Howley, eds.), *Fields Virology*, Lippincott-Raven Publishers, Philadelphia, 3rd edition, 1996, 713–734.

[9] R. G. Douglas Jr, Pathogenesis of rhinovirus common colds in human volunteers, *Annals of Otology, Rhinology, and Laryngology* **79** (1970), 563–571.

[10] J. W. Drake and J. J. Holland, Mutation rates among RNA viruses, *Proc. Natl. Acad. Sci. USA* **96** (1999), 13910–13913.

[11] N. M. Ferguson, A. P. Galvani, and R. M. Bush, Ecological and immunological determinants of influenza evolution, *Nature* **422** (2003), 428–433.

[12] S. A. Frank, *Immunology and Evolution of Infectious Disease*, Princeton University Press, Princeton, 2002, chapter 6.

[13] J. R. Gog and B. T. Grenfell, Dynamics and selection of many-strain pathogens, *Proc. Natl. Acad. Sci. USA* **99** (2002), 17209–17214.

[14] S. B. Greenberg, Respiratory consequences of rhinovirus infection, *Archives of Internal Medicine* **163** (2003), 278–284.

[15] J. M. Gwaltney Jr, J. O. Hendley, G. Simon, and W. S. Jordan Jr, Rhinovirus infections in an industrial population. II. Characteristics of illness and antibody response, *J. American Medical Association* **202** (1967), 494–500.

[16] D. T. Haydon and M. E. Woolhouse, Immune avoidance strategies in RNA viruses: fitness continuums arising from trade-offs between immunogenicity and antigenic variability, *J. Theor. Biol.* **193** (1998), 601–612.

[17] T. Heikkinen and A. Jarvinen, The common cold, *Lancet* **361** (2003), 51–59.

[18] J. O. Hendley, W. P. Edmondson Jr, and J. M. Gwaltney Jr, Relation between naturally acquired immunity and infectivity of two rhinoviruses in volunteers, *J. Infect. Dis.* **125** (1972), 243–248.

[19] D. W. Hosmer and S. Lemeshow, *Applied Logistic Regression*, Wiley-Interscience, New York, 2000.

[20] S. P. Hubbell, *The Unified Neutral Theory of Biodiversity and Biogeography*, Princeton University Press, Princeton, 2001.

[21] A. L. Hughes, Phylogeny of the Picornaviridae and differential evolutionary divergence of picornavirus proteins, *Infection, Genetics and Evolution* **4** (2004), 143–152.

[22] I. Kawaguchi, A. Sasaki, and M. Boots, Why are dengue virus serotypes so distantly related? Enhancement and limiting serotype similarity between dengue virus strains, *Proc. R. Soc. Lond. B* **270** (2003), 2241–2247.

[23] R. M. Ledford, N. R. Patel, T. M. Demenczuk, A. Watanyar, T. Herbertz, M. S. Collett, and D. C. Pevear, VP1 sequencing of all human rhinovirus serotypes: Insights into genus phylogeny and susceptibility to antiviral capsid-binding compounds, *J. Virol.* (2004), 3663–3674.

[24] A. S. Monto, E. R. Bryan, and S. Ohmit, Rhinovirus infections in Tecumseh, Michigan: frequency of illness and number of serotypes, *J. Infect. Dis.* **156** (1987), 43–49.

[25] M. A. Nowak, What is a quasispecies? *Trends in Ecology and Evolution* **7** (1992), 118–121.

[26] M. A. Nowak, R. M. May, R. E. Phillips, S. Rowland-Jones, D. G. Lalloo, S. McAdam, P. Klenerman, B. Köppe, K. Sigmund, C. R. M. Bangham, and A. J. McMichael, Antigenic oscillations and shifting immunodominance in HIV-1 infections, *Nature* **375** (1995), 606–611.

[27] M. A. Nowak, R. M. May, and K. Sigmund, Immune responses against multiple epitopes, *J. Theor. Biol.* **175** (1995), 325–353.

[28] L. J. Patterson and V. V. Hamparian, Hyper-antigenic variation occurs with human rhinovirus type 17, *J. Virol.* **71** (1997), 1370–1374.

[29] C. M. Pease, An evolutionary epidemiological mechanism, with applications to type A influenza, *Theor. Pop. Biol.* **31** (1987), 422–452.

[30] C. Savolainen, S. Blomqvist, M. N. Mulders, and T. Hovi, Genetic clustering of all 102 human rhinovirus prototype strains: serotype 87 is close to human enterovirus 70, *J. General Virology* **83** (2002), 333–340.

[31] C. Savolainen, P. Laine, M. N. Mulders, and T. Hovi, Sequence analysis of human rhinoviruses in the RNA-dependent RNA polymerase coding region reveals large within-species variation, *J. General Virology* **85** (2004), 2271–2277.

[32] E. J. Stott and M. Walker, Antigenic variation among strains of rhinovirus type 51, *Nature* **224** (1969), 1311–1312.

Department of Mathematics, University of Utah, 155 South 1400 East, Salt Lake City, UT 84112-0900

E-mail address: william@math.utah.edu

Department of Mathematics and Department of Biology, University of Utah, 155 South 1400 East, Salt Lake City, UT 84112-0900

E-mail address: adler@math.utah.edu