

Linking scales to understand diversity: Immunity, host ecology, and evolutionary landscapes in the dynamics of influenza

Introduction

At root, every hypothesis is a claim about the relevance of particular scales. If the hypothesis is parsimonious, and if the phenomenon it attempts to explain is simple, most scales are irrelevant. The classic SIR model (Kermack and McKendrick 1927) is impressive because it captures the dynamics of many infectious diseases with just a few interactions. Hosts are well-mixed particles in three possible compartments whose contact rates vary as simple linear or nonlinear functions of their populations. The model omits processes on other temporal scales, such as feedback with the physical environment and evolution by the pathogen and the host. It also omits spatial scales, such as competition among pathogens within a host and differences in contact rates due to population structure.

These assumptions work well for particular systems [e.g., Grenfell et al. (1992), cf. Swinton et al. (1998)], and are the foundation of other popular models in ecology [e.g., Volterra (1926)]. By nature of their simplicity, all their potential behaviors can be predicted analytically. However, many interesting phenomena are not well explained by these assumptions. Mean-field models can poorly approximate sexually transmitted diseases (Eames and Keeling 2002) and mechanisms of disease emergence, which may depend on the effects of superspreaders (Lloyd-Smith et al. 2005). In other systems, endogenous dynamics are strongly affected by extrinsic climatic forcing (Koelle and Pascual 2004) or interactions with other populations (LoGiudice et al. 2003; Jensen et al. 2006). In fast-evolving RNA viruses, such as HIV, strong selection within individual hosts and the timing of transmission may affect the frequency of CTL escape and CXCR4 mutants in the host population (Rambaut et al. 2004).

Local interactions, biotic and abiotic extrinsic drivers, and rapid evolution are examples of spatial, organizational, or temporal scales that crucially augment or replace the null dynamical model (Pascual 2005). Often these complexities can be incorporated into standard models by adding parameters, such as modified mixing terms (Roy and Pascual 2006). If the modifications do not produce the observed behavior, they may need to be replaced by a more accurate description of the underlying process, e.g., a contact network or agent-based simulation. One of the goals in researching these systems is “to address the importance of variability at small, local scales to the dynamics of aggregated quantities measured at large, global scales. If small-scale ‘details’ matter, we need to ask how much complexity we need to incorporate into large-scale models if we seek to both

understand and predict the dynamics of global quantities” (Pascual 2005). I would add that these details do not have to be small, since we are not always studying “global” quantities. We may also ask whether patterns are shaped by extrinsic factors or dynamics—perhaps it matters that the system is open. The best hypotheses of complex systems remain parsimonious while appealing to processes occurring on other spatial, temporal, or organizational scales to describe a pattern. They are consequently challenging to evaluate, since there are so many possible processes (and thus competing hypotheses) to choose from.

I would like to use this framework to determine what does and does not regulate the diversity of a common but poorly understood disease, influenza. I define diversity as the variation in genotype and phenotype at any one time and how this variation changes over time: It encapsulates a broad set of ecological and evolutionary patterns resulting from the interaction of few or many scales. I follow the lead of recent exceptional models that have relaxed common assumptions about which scales are relevant and thereby made contributions to our understanding of flu. The following section presents biological context—that is, the mechanisms I suspect underpin broad patterns in flu diversity—for the research questions that follow.

Mechanisms and patterns in influenza

Structure and antigenicity

The genome of influenza A consists of eight RNA segments, each 850 to 2300 bases long, which code for 10 proteins (Table 1). Two of these proteins, hemagglutinin (HA) and neuraminidase (NA) are abundant on the virus’s surface, with approximately 4 HA for every NA. There are sixteen forms of HA and nine forms of NA. Combinations of HA and NA form subtypes, e.g., H3N2. Amino acid sequences of HA differ up to 20% within subtypes and 30-70% between them (Skehel and Wiley 2000).

Hemagglutinin and neuraminidase are the primary determinants of antigenicity. Antibody-binding sites of some subtypes of HA and NA have been described by X-ray crystallography and electron micrographs of monoclonal antibody escape mutants (Bizebard et al. 1995; Fleury et al. 1999; Knossow et al. 2002). These sites are grouped into antibody-binding regions, or epitopes, on the globular head of hemagglutinin (HA1). HA1 of the H3N2 subtype infecting humans has four or five epitopes, labeled A-E (Figure 1a) [epitope D may only be recognized by murine antibodies (Sato et al. 2000)]. HA1 of the H1N1 subtype has four or five recognized epitopes (Caton et al. 1982) (Gerhard et al. 1981), and the N2 NA has at least two (Gulati et al. 2002). Different subtypes of HA1 can have the same epitopes (Smirnov et al. 1999).

Antibodies neutralize influenza viruses through steric inhibition of receptor binding or membrane fusion, rather than inducing conformational change in HA. The receptor-binding site is a highly conserved pocket at the top of the HA1 (in H3N2, the site falls near epitopes A and B) and shows little variation among subtypes (Skehel and Wiley 2000). Antibodies can neutralize viruses by blocking the receptor-binding site directly (Bizebard et al. 1995) or by binding to an epitope some distance away (Fleury et al. 1999) (Figure 1b); the latter mechanism can block receptor binding or interfere with membrane fusion. Antibodies to the same epitope can compete or interact synergistically

in neutralization (Brown et al. 1990; Sanna et al. 2000). Antibody affinity is positively but loosely correlated with neutralization ability (Kostolansky et al. 2000). One study found that on average one quarter of the hemagglutinins on the viral surface had to be neutralized to prevent infection (Knossow et al. 2002). The three neutralizing antibodies they investigated all interfered with receptor binding rather than membrane fusion. Neutralization kinetics have been described as log-linear or pseudo-first-order [reviewed in Frank (2002)].

Immunity

Hosts resist infection through humoral immunity, cellular immunity, and serum inhibitors. The contribution of each kind of immunity is a very active area of research, and contributions differ among host species.

Humoral immunity is the basis of permanent strain-specific immunity and some cross-immunity within subtypes in mammals; it also may play a role in heterosubtypic immunity. The specificity of antibodies is gauged by the ability of sera to inhibit hemagglutination by viruses of a particular strain. Most antibodies in sera of infected or vaccinated humans target HA1 (Sato et al. 2004), though some individuals also mount responses to NA (Cox and Brokstad 1999). All mammals investigated demonstrate antibodies to NP (Deboer et al. 1990; Cox and Brokstad 1999). Antibody repertoire can be monoclonal (antibodies to one epitope are present in antisera) or polyclonal (antibodies to multiple epitopes are present). Children have narrower antibody repertoires than adults: 25 of 27 children under age twelve developed antibodies to a region of epitope B of H3N2 HA1, and six showed antibodies for sites A or C. In contrast, older subjects had polyclonal responses to epitopes A, B, C, and E (Sato et al. 2004).

Cellular immunity can significantly abrogate pathology and speed viral clearance, though it appears to play a lesser role in preventing infection [Liang et al. (1994) and review in Thomas et al. (2006)]. T cells only attack presenting cells, and thus they usually lag behind antibodies in appearance and proliferation during infection. Mice challenged with a virus containing the internal proteins of human H1N1 produce CD8⁺ T cells specific to all six internal proteins, though T cells specific to epitopes on PA and NP predominate and can be detected ≥ 570 days after initial infection. Secondary responses tend to be dominated by NP (Belz et al. 2000; Marshall et al. 2001). CD4⁺ T cells bolster CD8⁺ and B cell responses and appear requisite for T cell memory (Belz et al. 2002). Studies of humans generally cannot separate the effects of cellular immunity from humoral and are unhelpful in gauging protection conferred by the former. They do, however, confirm that (1) there is no evidence that cellular immunity prevents infection, except possibly when exposures are simultaneous; and (2) there is a marked attenuation of symptoms during secondary infections. One small study observed several cases of rapid reinfections with heterologous subtypes, sometimes within days of clearance of the first subtype, and the secondary infection was no more likely to be asymptomatic than the first (Frank et al. 1983); in contrast, students at high schools experiencing concurrent epidemics of H3N2 and H1N1 were less likely to suffer multiple infections than students at schools with sequential epidemics (Sonoguchi et al. 1986). Attenuation of symptoms during secondary infections has also been demonstrated during pandemics of H2N2 and H3N2 (Sonoguchi et al. 1985). However, in both cases the HA of the pandemic strain

were rapidly evolving (Matrosovich et al. 2000), and similar patterns are produced by antibody-mediated cross-immunity (Gill and Murphy 1977).

Comparatively little is known about the immunologic effects of serum inhibitors, which can select variants with altered receptor binding sites and may play a role in host range (Rogers et al. 1983; Matrosovich et al. 1998).

Evolution

Influenza viruses evolve by point mutations, reassortment of whole gene segments, and rarely by recombination (Hirst et al. 2004; Suarez et al. 2004).

Point mutations are the most frequent means of escaping immune surveillance, and they are also a means to modulate virulence, develop drug resistance, and adapt to new hosts or tissue types. Monoclonal antibody escape mutants arise in vitro every 10^4 to 10^6 viruses (Webster and Laver 1980). They avoid recognition by conformational changes or additional glycosylation sites that block antibody binding. Conformational changes tend to affect only the local structure within the surrounding epitope (Knossow et al. 1984). Hemagglutinin may be particularly tolerant of such changes: epitope A mostly consists of a loop extending from the rest of the molecule, and epitopes B and C are bulges (Wiley et al. 1981). While amino acids at certain positions, such as loops, will have a dramatically greater influence than others on antibody recognition, the location of influential positions can change over time (Nakajima et al. 2005). As for other pathogens, tertiary protein structure greatly complicates predictions of the locations of B cell epitopes (Korber et al. 2006).

Influenza viruses can also escape immune surveillance through the addition of glycosylation sites [Asn-X-Thr/Ser, where X is any amino acid except proline or potentially aspartic acid (Gallagher et al. 1992)]. Host cell carbohydrates binding to these sites form a “glycan shield” to antibodies. Glycosylation is the major mechanism of antibody escape by HIV (Wei et al. 2003) and predicts the strength of the antibody repertoire mounted by Rhesus monkeys to SIV (Reitter et al. 1998). The number of potential glycosylation sites on HA1 of H3N2 has increased from two to six or seven since the subtype emerged in humans (Abe et al. 2004). While increased glycosylation decreases receptor binding activity in vitro, it can do so without negatively affecting cell fusion. Glycosylation may thus be an especially rapid effector of antigenic change (Schulze 1997), though it is not well tolerated by all subtypes of HA (Tsuchiya et al. 2002).

The accumulation of point mutations in response to immune pressure has previously been called “antigenic drift,” and it underlies the characteristic phylogeny of HA1 in human viruses: genetic distance steadily increases from the founding strain, and no strain persists more than a few years (Figure 2a). Recently it has been shown that these strains in human H3N2 form antigenic clusters defined by cross-reactivity patterns (Figure 2b). Only one cluster appears to dominate at any time (Smith et al. 2004).

In contrast to antigenic drift, “antigenic shift” by reassortment was once thought to be rare and to always cause pandemics (Webster et al. 1992). There is growing evidence that reassortment is as common in humans (Lindstrom et al. 2004; Holmes et al.

2005) as in other hosts (Hatchette et al. 2004; Webby et al. 2004). It may lead to partial immune escape by shuffling combinations of HA, NA, and other antigenic determinants. One antigenic cluster (FU02) may have arisen through reassortment (Holmes et al. 2005), potentially without antigenically significant changes in HA1. Unlike introductions of H1N1, H2N2, and H3N2 into the human population, the emergence of the H1N2 subtype since 2001-2002 has not obviously increased incidence.

Ecology

Influenza is highly seasonal in temperate regions, with a four-month epidemic period in winter and few cases in the summer (Figure 3a) (Cox and Subbarao 2000). Incidence shows no clear periodicity in the tropics (Figure 3b) (Chow et al. 2006; Viboud et al. 2006a). Pandemic and interpandemic influenza A strains circulate globally over short time periods. H3N2 outbreaks have especially high spatial synchrony (Greene et al. 2006; Viboud et al. 2006b), and phylogenetic studies of strains circulating in France (Lavenu et al. 2006), Japan (Nakajima et al. 1991), and New York (Nelson et al. Submitted) demonstrate that multiple lineages seed annual epidemics in each community. Swabs from air travelers support the hypothesis that there is interhemispheric transport of strains throughout the year (Sato et al. 2000).

Incidence is usually inferred from deaths to pneumonia and influenza. Until recently, few countries tracked deaths or infections by type and subtype. Observations from WHO collaborating labs reporting to the CDC suggest that seasons dominated by H3N2 have relatively low incidence of H1N1 and influenza B and vice-versa (Figure 3c) (Thompson et al. 2003; Greene et al. 2006). Estimates of annual incidence range from 10-20% in interpandemic years and 40-50% during pandemics (Cox and Subbarao 2000), but there does not seem to be much confidence in any value in the flu community. Contact with young children is a significant risk factor for infection (Gubareva et al. 2002; Viboud et al. 2004), and vaccination of young children can dramatically reduce incidence in older contact children and adults (Monto et al. 1970; Hurwitz et al. 2000).

Many other species can be infected with influenza viruses, and multiple subtypes are endemic in swine, domestic poultry, and horses. Influenza is seasonal in many aquatic birds, which are its natural hosts. Transmission between wild birds and domestic poultry, wild birds and swine, domestic poultry and humans, and swine and humans occurs at least annually in many parts of the developed and developing world where the populations cohabitate (section 5, below).

Models

Most models of influenza have focused on explaining seasonal dynamics of flu in (temperate latitude) human populations and mechanisms of strain cycling (Pease 1987; Andreasen et al. 1997; Lin et al. 1999; Andreasen 2003; Boni et al. 2004; Dushoff et al. 2004; Lavenu et al. 2004). They assume population immunity declines as a function of antigenic drift, which itself can occur at a constant rate or as a function of epidemic size, and that cross-immunity between serotypes is fixed. Regular and irregular oscillations in incidence, dynamical resonance, and complex strain cycling can result under these

assumptions. A few studies have explored strain structure with evolution. Boni et al. (2006) predicted that strong host immunity and long epidemics would lead to the highest rates of antigenic drift. They proposed that their model could explain anecdotal reports that antigenic variants tend to arise toward the end of a season, with attack rates rising the following year. Gog et al. (2003) demonstrated that a slight increase in the duration of infectiousness can substantially increase the survival probability of a mutant strain during seasonal bottlenecks in transmission. They suggested that this mechanism may have accounted for the rapid fixation of CTL escape mutants in the 1993-1994 season.

Three models have attempted to describe broader patterns by coupling dynamics of strain evolution to population dynamics. Gog and Grenfell (2002) showed that strains evolving in one-dimensional evolutionary space form clusters if infection times are short and cross-immunity is high. They noted that the existence of clusters could yield phylogenies with short branches, but, like previous studies, they did not attempt to match their system's endogenous dynamics to empirically observed seasonality. To explain strain evolution, subtype cycling, and seasonal dynamics in tandem, Ferguson et al. (2003) constructed an agent-based model that could generate HA trees with short side branches, annual fluctuations in incidence, and subtype replacement and coexistence. Central to their model was a strain-transcendent, generalized immunity (attributed to cellular immunity) that reduced the probability of infection in a density-dependent way; they found this immunity required a half-life of six months to restrict diversity sufficiently. They also incorporated spatial structure, something akin to original antigenic sin (exposure would boost immunity to old strains even if hosts were not infected), and a 30-year host lifetime. It is unclear how robust their results are to these assumptions.

Ferguson and coauthors showed that the interaction of ecological dynamics, immune selection, and evolutionary dynamics on comparable time scales might be critical for influenza. The third and most recent model elaborates another scale of complexity, rather than invoking generalized immunity, to explain influenza evolution and seasonality. Koelle et al. (in prep) infer from the existence of antigenic clusters that large spaces of genotype space are effectively neutral, since they have the same antigenic phenotype, and that this relationship arises from the nature of the genotype-phenotype mapping and not the infection history of the strain (Figure 4a). We propose that interpandemic influenza undergoes epochal evolution. During periods of apparent evolutionary stasis, strains diffuse through genotype space. A mutant that discovers a new phenotype (cluster) encounters fewer immune hosts and proliferates. Partial cross-immunity to the old cluster causes the old cluster to rapidly go extinct. The model predicts and confirms that real influenza strains undergo a boom-and-bust cycle of genetic diversity. The model further proposes—and requires—that genotype space within clusters is *almost* neutral. Without weak positive selection, simulated strains explore genotype space too slowly to generate the characteristic increases (Figure 4b).

I am not aware of any ecological and evolutionary models of influenza in other host species, despite recent dramatic changes in host ecology and viral diversity.

Proposed research

One of the most important ends of understanding biological systems, and diseases in particular, is prediction—including the knowledge that a system might be too stochastic

or chaotic to predict reliably. Our management strategy for influenza is still largely reactive; next season's vaccine is based on the most antigenically divergent strains of the current one, and until recently, only populations most at risk of dying from infection were targeted for vaccination [2006-2007 is the first season for which the CDC has recommended vaccinating children 6-59 months old (Smith et al. 2006)]. Influenza is furthermore an exciting system because of the convergence of so many possible dynamical scales: due to their high mutation rates and short generation times, RNA viruses are especially sensitive to changes in host ecology. It is increasingly evident that the relevant ecologies are both those encountered within hosts and in host populations, manifested as adaptive immune responses (Grenfell et al. 2004). We are in a position to affect host ecology, and thus viral evolution, through vaccination, antivirals, and farming practices.

My goal is to examine linkages between viral evolution on a molecular level and immunity on molecular and population levels to identify factors involved in regulating flu diversity in humans and other species. My first three research projects investigate the roles of humoral and cellular immunity in flu evolution and strain competition. The next two projects consider the effects of other aspects of host ecologies on viral evolution. The last project explores the adaptive potential of influenza under different modes of evolution.

My specific questions are:

1. **What temporal patterns of selection are evident in H3 HA?** If the model of epochal evolution is correct, we should see episodic selection in HA corresponding to cluster transitions, and weak selection within clusters. How are selection pressures distributed across HA1? Is there evidence that positive selection is at work on other proteins, such as NA? Is the spectrum of selection potentially adaptive, by helping HA find new phenotypes?
2. **How does antibody heterogeneity affect the dynamics of partial cross-immunity, strain structure, and epitope evolution?** Diverse, epitope-specific host antibody responses mean that the effective phenotype of a strain depends on the host it is infecting. How do strains compete under such heterogeneous selection pressures? What is the effect of antibody repertoire specificity on strain and epitope evolution?
3. **What are the conditions for subtype replacement, coexistence, and interference, and are they consistent with observations and the biology?** Ferguson et al. (2003) invoked generalized immunity to explain how subtypes replace each other during pandemics and interfere during interpandemic years, though other mechanisms might generate similar patterns. Which are the most robust and biologically grounded?
4. **How does epochal evolution interact with other aspects of host ecology to modulate flu diversity in swine?** Models show that the outcome of strain competition is affected by the duration of infectiousness relative to the host lifespan. Several antigenic clusters have been known to circulate in swine concurrently. Is this diversity predicted by differences in the host's life history and epidemiology?
5. **How do host ecologies affect evolutionary opportunities for emergence?** The fitness of a pathogen's phenotype is determined by the number of hosts available to it,

which is a function of host ecology. How do the endogenous disease dynamics and interspecific contact networks of different host species affect the probability that a virus with another receptor preference—allowing a shift in host range—will emerge?

6. **How do epistasis and modularity affect the genetic potential of influenza?** The idea that populations may be able to evolve their “genetic potential,” or capacity for change, has attracted attention from theoreticians. Applications of this idea have sparked debate. Is there evidence that influenza has evolved to increase its genetic—or antigenic—potential, e.g., through codon volatility, modularity, or the propensity to reassort? Is evolution of genetic potential feasible under epistasis?

1 Evidence of episodic selection on H3N2

There are two hypotheses of how influenza undergoes continual antigenic drift while its genetic diversity remains bounded over time. Ferguson, Galvani, and Bush (2003) and Tria et al. (2005) contend that density-dependent, generalized immunity is necessary to restrict diversity. Each group invokes another factor—heterogeneity in host transmission (spatial structure) and variable strain fitness, respectively—that crucially augments the effect of generalized immunity. Koelle et al. (in prep) propose that diversity is constrained episodically by the appearance of new antigenic phenotypes that competitively displace existing strains. Selective sweeps by antigenic variants have been proposed before for H3N2 (Fitch et al. 1991) and H1N1 (Ina and Gojobori 1994). The model of Koelle and coauthors suggests that the novel phenotypes correspond to the eleven clusters of H3N2 between 1968 and 1998. Cohort studies of reinfection in humans indicate that cross-immunity is as high as 95% within clusters and 60-84% between clusters (Gill and Murphy 1977).

The model also proposes that in addition to abrupt, intense sweeps, there is weak selection within clusters. A strain begins in a neutral network, i.e., genotype set corresponding to the same phenotype or cluster, and diffuses through high-dimensional genotype space via mutations until it arrives at a node (sequence) belonging to an adjacent network. With some probability, this adjacent network is a continuous (slight) or discontinuous (major) phenotypic change, with respectively a modest ($\cong 5\%$) or substantial ($\cong 20\%$) decrease in cross-immunity. Discontinuous changes curtail diversity by precipitating cluster transitions; continuous changes increase diversity by causing accelerated diffusion from the cluster's founding strain. The effects of accelerated diffusion in genotype space may be visible in the phylogenetic trees of some clusters (Figure 4a). For example, in BE89 and BE92, multiple lineages persist from one year to the next, suggestive of neutral diffusion. The trees of WU95 and SY97, in contrast, generally have unidirectional growth, which echoes the traditional descriptions of the HA1 trees as a whole. This pattern could be due to chance, perhaps in sampling or off-season extinction, or due to positive selection on some branches.

There are several reasons why it may be useful to evaluate the strength and location of positive selection in influenza over time. The first is that there is clearly some threshold where selection ceases to augment diversity and instead reduces it. The second is that selection, as suggested above, may allow prediction of which strains will dominate from year to year within a cluster. Before the clusters were identified, Bush et al. (1999b; 1999a) found that the number of amino acid replacements of 18 codons, most located on

two epitopes of HA1, could retrospectively predict the lineages that would survive from one year to the next. They applied their technique to eleven years; interestingly, their rule does most poorly in years in which cluster transitions probably occurred. This raises the interesting possibility that not only the strength but also the targets of selection within clusters may differ from that between clusters. Further, there is evidence suggesting the possibility of strong selection on NA (Venkatramani et al. 2006). Is this selection reflected in the tree of HA? Is there any temporal correlation in selection on the two proteins? Which epitopes on each protein are targeted? Recognition of these patterns could lead to better predictive models and more sophisticated vaccines.

Selection within clusters may also be a nontrivial component of influenza's evolutionary and ecological dynamics. A strictly neutral network model presents a paradox. If the supply of susceptibles diminishes as strains diffuse through genotype space, it would be possible for the virus to go extinct before finding a new phenotype. This has not yet happened for H3N2. Are the discontinuous phenotypes frequent enough to exclude this possibility, or might positive selection within networks critically accelerate diffusion, and thus increase the probability of finding significantly new phenotypes (K. Koelle, pers. comm.)?

The first chapter of my thesis will describe the strength and targets of selection in HA and NA of H3N2 over time. The null model, strict within-cluster neutrality, can be associated with three different patterns:

1. Simple diffusion with stochastic branch extinction should generate a distribution of pairwise genetic distances that changes characteristically over time. Specifically, it should increase, on average, more slowly than if fitnesses rise at other points in the network, since strains close to the founder will not be outcompeted faster than any others. Genetic distances under neutrality should also be less clustered in the graph theoretical sense, because no strain will be significantly more successful than any other.
2. There should not be an excess of nonsynonymous mutations within isolates from the same cluster. An excess of nonsynonymous mutations is the hallmark of selection.
3. *Within* clusters, branches with more amino acid substitutions or a higher ratio of nonsynonymous to synonymous replacement rates should not have a higher survival probability from year to year. Under neutrality, lineage survival within clusters should be random. Branches *linking* clusters, however, should have on average more amino acid substitutions than other branches. This expectation is based on the simple fact that the greater the number of amino acid replacements, the more likely it is that a discontinuous phenotype will be discovered.

I do not see a simple method for statistically evaluating the first prediction, though simulation may yield insight. My research will thus focus on the latter two predictions.

Using the 253 sequences antigenically typed by Smith et al. (2004), I have so far calculated d_N and d_S differences and ratios for groups of codons suggested in the literature of positive selection. Preliminary results suggest that there is weak positive

selection on at least one epitope of HA1 in every cluster and strong selection between them. Selection on epitopes A, B, and E is most often associated with cluster transitions (Figure 5). Epitope B has the greatest contrast in the strength of selection within clusters versus between them, suggesting it has greater antigenic potential than other epitopes. Epitopes C and D demonstrate fluctuating selection, e.g., there is no positive selection on epitope D after the 1970s. The strength of selection on epitopes within clusters does not appear predictive of the epitopes that are most strongly selected between clusters.

Future analysis will refine these measures. In particular, I will consider methods for detecting selection at single sites and in quasispecies (Stewart et al. 2001); methods that estimate the physicochemical effects of substitutions (Wong et al. 2006); and methods that consider selection in sites that are structurally close (Suzuki 2004). I will also broaden the analysis by searching over the entire HA1 and available NA sequences.

This work will be complemented by analysis of lineage survival, broadening the picture by Bush et al. (1999a). They report a 40% excess of amino acid substitutions on terminal branches, though the number of replacements at 18 positively selected codons was associated with lineage survival. It is telling that their rule of “fast” evolution, which predicts that the successful lineage has the greatest numbers of replacements at the 20 fastest-evolving codons (12 of which were under positive selection), performs best in the putative cluster transitions: accumulation of random changes in bulk may be more likely to lead to discontinuous transitions than selection at a few sites, perhaps because of the unpredictable effects of different residues on tertiary structure. Their work shows that it will be important to consider changes in individual codons, the locations of changes, and the number of changes for each branch.

To reevaluate their results and test my third hypothesis, I will thus build and compare trees with codon-based maximum likelihood and Bayesian substitution models, using relaxed clocks where possible (Drummond et al. 2006). More complex models become computationally tractable if trees are built of just a few clusters at a time. I will then evaluate models of positive selection using maximum likelihood and Bayesian approaches on the trees (Yang 1997; Ronquist and Huelsenbeck 2003; Kosakovsky Pond and Frost 2005; Kosakovsky Pond et al. 2005).

In my analysis of selection over time, I will also consider where the addition of glycosylation sites might have affected strain fitness by removing epitopes from antibody pressure (for example, at site D). I will also try to distinguish sites under T cell selection from those under antibody selection, using discovered (Rimmelzwaan et al. 2004; Berkhoff et al. 2005) and predicted (Korber et al. 2006) T cell epitopes.

2 Effects of antibody heterogeneity on the dynamics of partial cross-immunity, strain structure, and epitope evolution

Models of strain competition usually assume that cross-immunity between strains is invariant: all hosts infected with one strain have the same probability of being infected with another. Cross-immunity under this assumption can yield complex dynamics determined by the intensity of competition. Gupta et al. (1998) showed that for realistic ratios of infection times and host lifespans, cross-immunity could move the system

through three different dynamical regimes (Figure 6). Intense strain competition led to the dominance of one antigenically nonoverlapping set; at intermediate competition, strains cycled periodically or chaotically; and when cross-immunity was low, strain structure disappeared. Koelle et al. (in prep) also assume that the cross-immunity between strains is fixed. Strains within a cluster have almost complete cross-immunity, but their high mutation rate allows genetic and transient antigenic diversification. Slightly lower cross-immunity between clusters causes competitive exclusion by ensuring large fitness differences between serotypes. Thus, the intrasubtypic dynamics of influenza are dominated by the regime of intense competition identified by Gupta and coauthors.

It has been shown that hosts of the same species can mount different immune responses after infection with identical strains. Wang et al. (1986) found that some strains provoked more varied responses than others: antibodies to Hong Kong pandemic strains induced antibodies to sites A and B or both, whereas a strain from 1978 induced antibodies to several epitopes. In humans, differences exist between the responses of children and adults. Nakajima et al. (2000) looked at the acute phase and convalescent sera of nine people infected with H3N2 during the 1990-1991 season (during the BE89 cluster) and found that all the young children had consistently narrower responses than adults. The sera of the three and four year old children had antibodies only to site B1, and all the older children and the one adult had antibodies binding to sites A, B1, B2, C, and C/E. In a follow-up study, Sato et al. (2004) examined the sera of 35 people who had been infected with a strain of the SY97 cluster and found that almost all young children developed antibodies to B1 and many to A. Everyone else developed a unique polyclonal response, often reacting more strongly to epitopes other than B1.

Such observations have motivated proposals that variation in antibody repertoires could underlie the mechanism of antigenic drift in humans. These hypotheses differ in the role of populations with polyclonal responses. Nakajima et al. (2000) and Sato et al. (2004) posit that drift results from serial adaptation to monoclonally-responding subpopulations. Antigenic variants first escape site B1 and then sites targeted by others; despite the rarity of older children or adults with narrow responses to sites other than B1 in their study, “we [can] not exclude the possibility that on a worldwide scale, individuals of this type may be numerous” (Sato et al. 2004). This pattern, they note, fits evolution within BE89 but not SY97. Cleveland et al. (1997) recalled that major drift variants tend to have at least four amino acid substitutions in two epitopes (Wilson and Cox 1990). They predict the existence of four different “human genetic groupings” with consistent, nonoverlapping epitope biases. Viruses drift as they move from group to group, acquiring a critical amino acid change in each, and become double escape mutants. In contrast to Sato et al. (2004), they argue that polyclonal responses can select for drift mutants under particular conditions: extrapolating from responses in rabbits, they conclude that antibodies to one epitope must predominate, the titer of that antibody must be sufficiently high, and titers of other antibodies must be sufficiently low. Other models in mice and ferrets suggest that mutants can arise from polyclonal responses as long as they could escape a predominant antibody [reviewed in (Nakajima et al. 2000)].

My second project will examine the ecological and evolutionary implications of heterogeneous antibody responses. Specifically, I ask:

1. How does variable cross-immunity between strains affect the outcome of strain competition, i.e., the possible dynamical regimes proposed by Gupta et al. (1998)?
2. How do viruses evolve at the epitope level in heterogeneous populations? Do the patterns predicted above emerge?

To address the first question, I will simulate an adaptation of Gupta's model. She defines strains as having n loci, each defined by m possible alleles. In my model, each locus corresponds to an epitope, and each allele a possible phenotype of the epitope. Cross-immunity is set by γ , which indicates the reduction in transmission probability conferred by previous infection with one strain; ($0 \leq \gamma \leq 1$). Without heterogeneous immune responses, the fraction immune to a strain i , z_i , changes as

$$\frac{dz_i}{dt} = (1 - z_i)\lambda_i - \mu z_i \quad (1)$$

where λ_i is the force of infection of strain i (the per capita rate of rate of acquiring infection, which is linearly proportional to the number of infectious individuals) and μ is the birth and death rate. Gupta et al. then add a compartment w_i representing hosts immune to all strains j that share alleles with i , including i itself:

$$\frac{dw_i}{dt} = (1 - w_i) \sum_{j \sim i} \lambda_j - \mu w_i \quad (2)$$

The expression $j \sim i$ refers to all strains j sharing alleles with i . The population of individuals infectious with strain i , y_i , is then determined by

$$\frac{dy_i}{dt} = [(1 - w_i) + (1 - \gamma)(w_i - z_i)]\lambda_i - \sigma y_i \quad (3)$$

where σ is the rate of loss of infectiousness. To incorporate antibody heterogeneity into this formalism, we need to track populations immune to epitopes. Strain i is now defined as a set of epitopes, with each epitope defined by a phenotype k : $i = \{1_k, \dots, n_k\}$ where $k \in [1, m]$ and $k \in \mathbf{Z}^+$. Let p_n be the probability that an individual develops an antibody to epitope n . Assume all responses are on average monoclonal to one epitope, so that $\sum p_n = 1$. We also assume that $\gamma = 1$ if two strains share an epitope to which a host has antibody. Equation (1) does not change: all people infected with strain i will mount a specific response to one of its epitopes and will not transmit i in the future. But now not all hosts with immunity to strain j , which shares epitopes with i , will potentially have immunity to i . Only the fraction of hosts infected with j that mount antibodies to epitopes shared with i will then be immune to i . Let S_{ij} be the set of shared epitopes between strains i and j : $S_{ij} \equiv i \cap j$. The probability r_{ij} of developing antibodies to i if infected with j is

$$r_{ij} = \sum_{n \in S_{ij}} p_n \quad (4)$$

Thus equation (2) becomes

$$\frac{dw_i}{dt} = (1 - w_i) \sum_j \lambda_j r_{ij} - \mu w_i \quad (5)$$

The reduction in w_i , the number of hosts immune from infection with other strains takes the place of γ , and the infectious class simply changes as

$$\frac{dy_i}{dt} = (1 - w_i)\lambda_i - \sigma y_i \quad (6)$$

Initially I will allow only one phenotype at each epitope ($m = 1$). I will explore the effects of varying p_I over $p_I = p_n = 1/n$ to $p_I = 1$. For different values of p_I , what are the system's dynamics, and does p_I have similar thresholds to γ ? Can the bifurcation points be derived analytically? I will then explore the behavior of $m > 1$. It might be interesting to relax the assumption that responses are monoclonal, and model dynamics of two host groups analogous to young children and adults. The monoclonal assumption corresponds to $\sum p_n = 1$ and fully polyclonal to $\sum p_n = n$ (the latter corresponds to Gupta's model with $\gamma = 1$).

To address the effects of antibody heterogeneity on evolution, I will let $m > 1$ and define a cross-immunity parameter between different phenotypes of the same epitope. This cross-immunity value could be a linear or nonlinear function of the virus's genotype, represented by a bit string at each epitope. This model seems most practically evaluated by simulation. The goal of this experiment is to observe the effect of antibody bias on selection pressures on individual epitopes, measured by d_N and d_S , and virus phylogeny. I will then compare results to my findings from the first research project to see if I have produced realistic within-cluster dynamics.

3 Subtype competition: Conditions for coexistence and exclusion

Among the most striking and least well understood patterns in influenza are those of subtype replacement and coexistence (Earn et al. 2002). There is strong evidence that H3N2 (or another H3 subtype) circulated in humans before 1918 (Houswort. Wj and Spoon 1971; Enserink 2006) [and possibly H2N2 before that (Masurel and Marine 1973)]. It was then replaced by H1N1. In 1957, H2N2 replaced H1N1. In 1968, H3N2 replaced H2N2. H1N1 reentered the population in 1977 and coexisted with H3N2. H1N2 appeared sporadically in the 1980s and became widespread in 2001-2002 (Guo et al. 1992a; Xu et al. 2002). H3N2 and H1N1/influenza B appear negatively correlated in incidence from season to season (Ferguson et al. 2003). The time series indicates there are major differences in attack rates between emergence events of different subtypes (Figure 1a).

Several hypotheses have been offered to explain instances of replacement or coexistence with interference. Ferguson et al. (2003) argue that short-lived, nonspecific immunity is critical for capturing the dynamics of subtype replacement during pandemics and out-of-phase oscillations when emergence is not accompanied by a pandemic. Antibody to NA of H2N2 was significantly associated with a lower probability of infection with H3N2 during its emergence in 1968 (Monto and Kendal 1973). Viboud et al. (2005) propose that different frequencies of antibody to NA was the largest factor modulating the intensity of the H3N2 pandemic in North America and Europe. Studies of heterosubtypic immunity in humans and other animals suggest no shortage of possible effectors (Table 2), e.g., pigs previously infected with H3N2 or H1N1 are partially

protected from H1N2 (Van Reeth et al. 2004). In light of the results of Gupta et al. (1998), it is clear that the outcome of competition might also be determined by chaotic dynamics. These dynamics will also be modulated by influenza's strong seasonality.

For my third project, I plan to use results from the analytic model and the simulation developed for my second project to compare hypotheses of subtype replacement and coexistence. It is clear from Table 2 that cross-reacting epitopes may be present on almost any protein, though each epitope might elicit a quite different immune response, depending on whether it is targeted by T cells or B cells. I would like to determine the probabilistic thresholds of host heterogeneity for dynamical regimes—including subtype replacement and coexistence—under different hypotheses:

1. How does the addition of a **seasonal forcing** affect previous results on antibody heterogeneity? For example, for a given amount of forcing, what is the minimum p_I for exclusion between two serotypes sharing one epitope (e.g., on nucleoprotein)? To avoid unrealistically severe bottlenecks, I would include three populations, representing the northern hemisphere, the tropics, and southern hemisphere.

The following hypotheses would be addressed with and without seasonality:

2. **Asymmetric cross-immunity** may cause antigenic drift in H3N2 (de St. Groth 1977). There are numerous observations of antibodies to later strains binding more strongly to earlier strains than antibodies to earlier strains do to later strains (1982; 1983; 1987). There are perhaps fewer opportunities for asymmetric cross-immunity among subtypes (glycosylation of HA is the candidate mechanism in H3N2), but there is evidence of asymmetry in the responses of swine to H3N2 and H1N1 (Heinen et al. 2001). How does asymmetric cross-immunity change thresholds of p_I ?
3. How would **short-lived generalized immunity**, as defined by Ferguson et al. (2003), affect dynamics of subtype replacement and coexistence? How does the time to subtype fixation vary between generalized immunity and antibody-mediated immunity?
4. Do results differ if the reduction in transmissibility acts through the **probability of infection** (similar to humoral protection) or the **duration of infection** (similar to cellular protection)?

I would then like to conduct a more in-depth review of the literature on influenza's epidemiology and immunology to evaluate hypotheses in light of both their biological plausibility and dynamical robustness. I will also identify missing information that would allow determination of the mechanisms of specific replacements.

4 Other determinants of diversity: Host ecology and influenza in swine

Until recently, swine in North America circulated only one subtype of influenza, the "classical swine" H1N1, which evolved from human H1N1 sometime before 1933. In 1997 and 1998, H3N2 appeared in the United States and became widespread within a year [reviewed in Webby et al. (2004)]. At least two lineages emerged, one a double

reassortant between human H3N2 of the SY97 cluster and classical swine H1N1, and the other a triple reassortant containing avian flu genes (Zhou et al. 2000). Since then, H3N2 in swine has acquired at least two more HA from human H3N2, and further reassortment with classical swine virus has produced at least two lineages of H1N2, which has also become widespread (Karasin et al. 2002). Reassortment between classical swine H1N1, human H3N2, and avian H1N1 had been described previously in European swine populations (Castrucci et al. 1993; Marozin et al. 2002). In the 1970s, the HK68, EN72, and VI75 clusters of human H3N2 were found circulating in Asian and Italian swine after HK68 and EN72 had disappeared from the human population (Shortridge et al. 1977; Ottis et al. 1982). It is interesting that proliferation of genetic and antigenic diversity in North American swine accompanies dramatic changes in host ecology. There are currently 100 million swine in North America; in the United States, the percentage of swine farms with ≥ 5000 swine increased from 18% in 1993 to 53% in 2002, and vaccination over the same time period became common (negligible in 1995, 44.1% of sows in 2000, and over half in 2003) (Wuethrich 2003). In poultry, vaccination has been associated with rapid antigenic drift away from vaccine strains (Lee et al. 2004).

Before the emergence of H3N2, several authors described antigenic drift in swine as slower than in humans, possibly due to their short life spans relative to the frequency of epidemics and the infrequency of vaccination [reviewed in Heinen et al. (2001)] or the high standing antigenic diversity (Olsen et al. 2000). It is also possible that smaller farm sizes would have reduced the frequency of epidemics and the spread of drift strains. The rapid spread of emerging subtypes in North America and Eurasia, and such observations as a multidrug resistant European swine influenza viruses in Hong Kong (Gregory et al. 2001), suggests that the relevant spatial scale for swine influenza ecology is closer to continents than farms. Because swine frequently exchange viruses with humans, and because they might be capable of sustaining more (or at least different) antigenic and genetic diversity than influenza in humans, it is important to have models to explain and predict these patterns.

For my fourth project, I propose to develop a null model of ecological and evolutionary dynamics of influenza in swine. This model will be functionally similar to the one developed for influenza's epochal evolution in humans (Koelle et al. in prep) and will complement similar investigations into the dynamics of avian and equine influenza. I plan four steps in the analysis:

1. Assume the genotype-phenotype map and antibody responses of swine are identical to humans', and study the effects of altered **pathology** (e.g., duration of infectiousness) and **life history** on incidence, phylogeny, and antigenic diversity. Do multiple clusters circulate? An important question that will have to be addressed is whether to force seasonality; infections U.S. swine appear to be seasonal (Olsen et al. 2000).
2. Add simple subcontinental **spatial structure**; how does drift slow and diversity increase as fragmentation increases?
3. Simulate **vaccination** by immunizing a fraction of the population each year with strains from the previous year. How are incidence and antigenic drift affected?

4. Assume swine have a more **refined (sensitive) antibody response** than humans, such that fewer mutations are neutral. Antigenic diversity should increase, but probably without commensurate increases in genetic diversity. Can this mechanism be distinguished qualitatively or quantitatively from mechanisms 1-3?

The researchers who developed antigenic maps of influenza in humans have submitted a paper on antigenic drift in swine (K. Koelle, pers. comm.). It will be interesting to see if their results agree with the trends in antigenic and genetic diversity described above. If they classify enough available sequences by antigenic type, it may be possible to perform selection analysis. If the authors use ferret sera to determine antigenic clusters, this analysis can help determine if the antigenic types are also recognized by swine sera. It might also shed light on whether the presumed clusters arose from selection or neutral drift, e.g., via spatial segregation.

5 Immunological, ecological, and evolutionary drivers of host range

Concern that a new strain of influenza virus may emerge from an animal reservoir and cause a pandemic in humans has motivated models of how the disease might spread in human populations (Ferguson et al. 2005; Longini et al. 2005). The mechanisms by which such a strain might move from animals to humans have generated much speculation. The traditional view of influenza held that the appearance of novel subtypes in humans, such as H2N2 in 1957 and H3N2 in 1968, was always preceded by reassortment between avian, swine, and possibly human strains in pig populations (Webster et al. 1992; Castrucci et al. 1993; Ludwig et al. 1995). Transmission of these reassortants to humans was considered rare and random, with humans occupying a peripheral niche in influenza's large community of hosts. In addition to humans and pigs, this community included wild waterfowl, which sustained all known subtypes of the virus, and also horses and domesticated birds. With few exceptions, influenza's dynamics in each species appeared independent of dynamics in other species. Outbreaks in turkeys in North America, which coincided with the fall migration of waterfowl, and serological surveys of asymptomatic pig farm workers were the only indicators of regular interspecific transmission (Halvorson et al. 1983; Karunakaran et al. 1983; Sivanandan et al. 1991; Campitelli et al. 1997; Olsen et al. 2002; Myers et al. 2006).

Influenza's recent activities in almost all host species indicate that ecological and evolutionary processes, and these processes in different hosts, cannot—or can no longer—be considered independent. For the first time since 1961, wild waterfowl have suffered severe morbidity and mortality from highly pathogenic H5N1, suggesting that they may not be in “evolutionary stasis” with the virus (Kida et al. 1980), cf. (Ito and Kawaoka 1998; Hatchette et al. 2004). These infections resulted from H5N1 strains adapted to poultry, which worldwide over the past nine years have experienced twice the number of outbreaks than in the previous four decades (Hirst et al. 2004; Shortridge 2005). These epidemics were often caused by the introduction of new subtypes by waterfowl throughout Asia, Europe, and North and South America (Guan et al. 2002). Many of these subtypes are still circulating in poultry and also mixing with swine (Ludwig et al. 1994; Peiris et al. 2001; Perez et al. 2003). In addition, strain diversity

within certain subtypes is increasing, perhaps as a result of the selection pressures induced by vaccination measures prompted by the outbreaks (Horimoto et al. 1995; Lee et al. 2004; Swayne 2005). This rise in viral diversity in poultry has been mirrored in swine in North America and Asia (Zhou et al. 2000; Webby et al. 2004). Perhaps the most significant development in influenza's ecology is the establishment of a new transmission route between two host species: Since 1996, influenza has been repeatedly and directly transmitted from birds to humans (Banks et al. 1998; Lin et al. 2000; Shortridge 2005). These transmissions, which have involved several avian species, at least five subtypes (involving H5, H7, and H9 HA), and hundreds to thousands of humans on different continents, demand a reexamination of almost all aspects of the virus-host interactions, and especially the ecological and evolutionary constraints on host range (Suarez 2000; Liu et al. 2003a; Enserink 2004; Palese 2004).

While no constraint holds for all subtypes all the time, the least escapable biochemical limitation in vivo and in vitro is the compatibility of virus's receptor binding site and the host cell's sialic acid receptor. In waterfowl, marine mammals, and horses, HA binds to sialic acid receptors of the $\alpha 2,3$ conformation, which are found throughout birds' gastrointestinal tracts and the respiratory epithelia of horses and marine mammals. The respiratory epithelia of humans bind in a $\alpha 2,6$ conformation. Cells in pigs and chickens have both receptor types, allowing them to be infected by viruses adapted to different ranges of hosts, and thereby permitting the generation of novel subtypes. Thus, pigs and chickens might serve as "mixing vessels" or "intermediate hosts" for influenza (Claas et al. 1994; Scholtissek et al. 1998; Matrosovich et al. 2001; Gambaryan et al. 2002a; Gambaryan et al. 2002b). Not all viruses share this constraint: strains of several subtypes have been able to infect humans while retaining their $\alpha 2,3$ preference, but they do not appear capable of human-to-human transmission (Matrosovich et al. 1999). It is also likely that hosts are commensurately ambiguous by means of their receptor availability: pigs have more $\alpha 2,6$ than $\alpha 2,3$, and swine-adapted viruses tend to prefer $\alpha 2,6$; and $\alpha 2,3$ (but not $\alpha 2,6$) receptors are found in the human eye (Olofsson et al. 2005) and lower respiratory tract (Shinya et al. 2006).

Despite many uncertain details, the consistent match between the virus's receptor preferences and the receptors available in the hosts to which they are adapted suggests a framework in which the evolution of host range can be approached. Experiments have shown that most viruses cannot replicate in host tissue of dissimilar receptor type, and viruses preferring one receptor type can usually sustain some replication in any host possessing that type, even if they are adapted to other species (Kida et al. 1994; Ito et al. 1999; Ito and Kawaoka 2000; Gambaryan et al. 2002b; Lee et al. 2005). Thus, the chemistry of receptor binding creates a tradeoff between the ability to invade cells of one type and the other. Viruses can change their receptor preference through several mutational steps, including single nucleotide substitutions at residue 226, or through reassortment with a virus having the other receptor preference on its HA (Nobusawa et al. 2000). The mutational jump from $\alpha 2,3$ to $\alpha 2,6$ occurred in domesticated swine in the 1980s following infection by poultry and can easily be obtained in vitro (Rogers et al. 1983). Reassortment appears common in most hosts (Hinshaw et al. 1980; Guo et al. 1992b; Yamnikova et al. 1993; Hatchette et al. 2004; Lindstrom et al. 2004), and can occur even when viruses are unable to replicate efficiently (Kida et al. 1994).

Various ecological conditions have been claimed to enable “optimal evolution” responsible for the recent spread of multiple subtypes and strains in multiple hosts (Webster and Hulse 2004). Small family farms, large commercial farms, live bird markets, chickens, pigs, and other species have been implicated in the recent changes in influenza’s host range (Bulaga et al. 2003; Liu et al. 2003b; Webster and Hulse 2004). So far, the effects of host ecology on the evolutionary opportunities available to influenza viruses have not been explored quantitatively.

The goal of my fifth research project is to develop approaches in which to evaluate these hypotheses more rigorously, and that can be applied to other pathogens whose host ranges are also partly constrained by cell recognition [reviewed in Baranowski et al. (2001)]. In particular, I would like to assess how adaptation is affected by host ecology, where hosts are defined by their receptor types, contacts, and epidemiologically relevant traits (Figure 7); and how reassortment changes this picture. I will initially address this question in two ways:

1. Analytically: Adaptive dynamics can be more applicable than R_0 maximization in studying systems with fluctuating selection pressures (Dieckmann et al. 2002). Here, selection pressure ($\alpha_{2,3}$ or $\alpha_{2,6}$ receptor availability and host immunity) changes as a function of the endogenous disease dynamics and any exogenous factors (e.g., seasonality or changes in contact rates between species). I would like to model the invasion fitness of $\alpha_{2,3}$ - or $\alpha_{2,6}$ -adapted mutants against the resident strain as a function of the availability of different host species. The most tractable approach might involve working with the equilibrium prevalence of the disease in each host species and assuming constant rates of contact between them.
2. Stochastic simulation: Fluctuations in disease prevalence resulting from endogenous dynamics, the effects of small population sizes, and the conditions corresponding to emergence are probably best modeled by stochastic simulation. I have constructed an event-driven, agent-based model in which to compare the effects of these factors against predictions.

I would like to tie this research to other theoretical models of pathogen evolution on networks. Among these models’ findings are that networks can support higher pathogen diversity (Buckee et al. 2004) and select intermediate levels of traits (Rauch et al. 2003; van Ballegooijen and Boerlijst 2004) relative to their mean-field counterparts.

6 Genotype-phenotype maps and the evolution of evolvability

Finally, I would like to analyze a group of provocative arguments about the evolution of genetic control over phenotypic variation in the context of influenza.

One of the ways in which populations may adapt to changing environments is through “genetic potential”—that is, “a heightened sensitivity to the effects of mutation that facilitates rapid evolution to novel states” (Meyers et al. 2005). Using a toy model, Meyers and coauthors showed that populations accumulate genetic potential fastest when mutation rates are high and the exogenously defined fitnesses of two possible phenotypes

alternate every generation. This idea has also been described by “codon volatility,” which measures the proportion of one-point neighbors of a codon that code for another amino acid (under the Hamming metric) or the change in the stereochemical properties of the amino acid (under the Miyata metric) (Plotkin and Dushoff 2003). Plotkin and Dushoff calculated that codons in HA1 epitopes are significantly more biased than codons in NA and NP. They later claimed that past selection can be inferred from measuring the codon volatility of a single sequence, which they demonstrated with *Plasmodium* (Plotkin et al. 2004). Their idea was countered with many examples, including genomes of HIV and other *Plasmodium*, in which volatility did not reflect selection (Friedman and Hughes 2005; Pillai et al. 2005). It was suggested that correlations between volatility and selection instead result from translational biases (Stoletzki et al. 2005). Zhang (2005) showed in simulations that strong directional selection could not increase codon volatility. We expect that selection on HA1 of influenza is frequency dependent, but that with high phenotypic diversity it more closely resembles directional selection: any mutation enabling antibody escape is advantageous, and mutations accumulate over time (Nakajima et al. 2005). For the first part of this project, I would like to revisit Plotkin and Dushoff’s calculation of volatility to determine if the volatile codons might reflect a translational bias, and if there is evidence of frequency dependent selection on HA1 codons over short time scales.

Next, I would like to explore genetic potential more broadly. It would be interesting to see if volatile codons are also the most accessible; the two properties are not identical, and random mutations should land sequences in the networks corresponding to more accessible amino acids. The theory of genetic potential has also largely ignored the consequences of higher-level interactions: Is it possible to evolve to a state of “heightened sensitivity to the effects of mutation that facilitates rapid evolution to novel [phenotypes]” when amino acids interact in complex, “epistatic” ways? Meyers and coauthors (2005) assume that traits under selection are controlled by independent coding regions. Kauffman (1995) argued that there exists a place in the NK adaptive landscape corresponding to phase transition between excessive order and chaos, where mutations to near neighbors usually fall in the same basin of attraction; this behavior emerges where each node interacts with only two of its neighbors ($K = 2$). In the NK model of Koelle et al. (in prep), even the lowest amount of epistasis ($K = 1$) among the amino acids comprising an epitope frequently resulted in phenotype shifts after one or two mutations. In other words, under very modest assumptions of interaction, the neutral network had too many edges for sequences to accumulate “genetic potential”—there was too much potential everywhere. This is worth simulating more formally: what is the range of genetic potential under the NK model? *Is there an appreciable range?*

Finally, I would like to search the literature for other models of interacting amino acids that could accommodate more genetic potential. Do such landscapes have more holes? Are they based on hydrophobicity or size? If such models do not exist or have very constrained assumptions, then the concept of genetic potential either has limited application or it is not based on epistasis. Have lower K values—in other words, increasing modularity—been selected? I would like to test the compatibility of these hypotheses on genotype-phenotype landscapes and the evolution of evolvability through simulations and, if possible, mathematical analysis. Site-directed mutation experiments in different strains of H3N2 over time, including comparisons of the fitness effects of

mutations *in vitro* and *in vivo* (Nakajima et al. 2003), will be useful in illuminating this problem.

Collaborators

In addition to working closely with my advisor, Mercedes Pascual, I will continue collaborating with Katia Koelle at Penn State University. I will take the lead on all projects described here.

References

- (1982) Current trends antigenic analysis of recent influenza isolates. *Morbidity and Mortality Weekly Report* 31(9): 110-111.
- (1983) Current trends antigenic analysis of recent influenza virus. *Morbidity and Mortality Weekly Report* 32(15): 195-196.
- (1987) Current trends update on influenza activity worldwide and World Health Organization and United States recommendations for influenza vaccine composition for the 1987-1988 season. *Morbidity and Mortality Weekly Report* 36(13): 193-195.
- Abe Y, Takashita E, Sugawara K, Matsuzaki Y, Muraki Y et al. (2004) Effect of the addition of oligosaccharides on the biological activities and antigenicity of influenza A/H3N2 virus hemagglutinin. *J Virol* 78(18): 9605-9611.
- Andreasen V (2003) Dynamics of annual influenza A epidemics with immuno-selection. *Journal of Mathematical Biology* 46(6): 504-536.
- Andreasen V, Lin J, Levin SA (1997) The dynamics of cocirculating influenza strains conferring partial cross-immunity. *Journal of Mathematical Biology* 35(7): 825-842.
- Banks J, Speidel E, Alexander DJ (1998) Characterisation of an avian influenza A virus isolated from a human - is an intermediate host necessary for the emergence of pandemic influenza viruses? *Archives of Virology* 143(4): 781-787.
- Baranowski E, Ruiz-Jarabo CM, Domingo E (2001) Evolution of cell recognition by viruses. *Science* 292(5519): 1102-1105.
- Belz GT, Xie WD, Altman JD, Doherty PC (2000) A previously unrecognized H-2D(b)-restricted peptide prominent in the primary influenza A virus-specific CD8(+) T-cell response is much less apparent following secondary challenge. *J Virol* 74(8): 3486-3493.
- Belz GT, Wodarz D, Diaz G, Nowak MA, Doherty PC (2002) Compromised influenza virus-specific CD8(+)-T-Cell memory in CD4(+)-T-cell-deficient mice. *J Virol* 76(23): 12388-12393.
- Berkhoff EGM, de Wit E, Geelhoed-Mieras MM, Boon ACM, Symons J et al. (2005) Functional constraints of influenza A virus epitopes limit escape from cytotoxic T lymphocytes. *J Virol* 79(17): 11239-11246.
- Bizebard T, Gigant B, Rigolet P, Rasmussen B, Diat O et al. (1995) Structure of Influenza-Virus Hemagglutinin Complexed with a Neutralizing Antibody. *Nature* 376(6535): 92-94.
- Boni MF, Gog JR, Andreasen V, Christiansen FB (2004) Influenza drift and epidemic size: the race between generating and escaping immunity. *Theoretical Population Biology* 65(2): 179-191.
- Boni MF, Gog JR, Andreasen V, Feldman MW (2006) Epidemic dynamics and antigenic evolution in a single season of influenza A. *Proceedings of the Royal Society B-Biological Sciences* 273(1592): 1307-1316.
- Boon aCM, de Mutsert G, van Baarle D, Smith DJ, Lapedes AS et al. (2004) Recognition of homo- and heterosubtypic variants of influenza A viruses by human CD8(+) T lymphocytes. *Journal of Immunology* 172(4): 2453-2460.
- Brown LE, Murray JM, White DO, Jackson DC (1990) An Analysis of the Properties of Monoclonal-Antibodies Directed to Epitopes on Influenza-Virus Hemagglutinin. *Archives of Virology* 114(1-2): 1-26.
- Buckee CO, Koelle K, Mustard MJ, Gupta S (2004) The effects of host contact network structure on pathogen diversity and strain structure. *Proceedings of the National Academy of Sciences of the United States of America* 101(29): 10839-10844.

- Bulaga LL, Garber L, Senne DA, Myers TJ, Good R et al. (2003) Epidemiologic and surveillance studies on avian influenza in live-bird markets in New York and New Jersey, 2001. *Avian Diseases* 47: 996-1001.
- Bush RM, Fitch WM, Bender CA, Cox NJ (1999a) Positive selection on the H3 hemagglutinin gene of human influenza virus A. *Molecular Biology and Evolution* 16(11): 1457-1465.
- Bush RM, Bender CA, Subbarao K, Cox NJ, Fitch WM (1999b) Predicting the evolution of human influenza A. *Science* 286(5446): 1921-1925.
- Campitelli L, Donatelli I, Foni E, Castrucci MR, Fabiani C et al. (1997) Continued evolution of H1N1 and H3N2 influenza viruses in pigs in Italy. *Virology* 232(2): 310-318.
- Castrucci MR, Donatelli I, Sidoli L, Barigazzi G, Kawaoka Y et al. (1993) Genetic Reassortment between Avian and Human Influenza-a Viruses in Italian Pigs. *Virology* 193(1): 503-506.
- Caton AJ, Brownlee GG, Yewdell JW, Gerhard W (1982) The antigenic structure of the influenza virus A/PR/8/34 hemagglutinin (H1 subtype). *Cell* 31: 417-427.
- Chow a, Ma S, Ling AE, Chew SK (2006) Influenza-associated deaths in tropical Singapore. *Emerging Infectious Diseases* 12(1): 114-121.
- Claas ECJ, Kawaoka Y, Dejong JC, Masurel N, Webster RG (1994) Infection of Children with Avian-Human Reassortant Influenza-Virus from Pigs in Europe. *Virology* 204(1): 453-457.
- Cleveland SM, Taylor HP, Dimmock NJ (1997) Selection of neutralizing antibody escape mutants with type A influenza virus HA-specific polyclonal antisera: Possible significance for antigenic drift. *Epidemiology and Infection* 118(2): 149-154.
- Cox NJ, Subbarao K (2000) Global epidemiology of influenza: Past and present. *Annual Review of Medicine* 51: 407-421.
- Cox RJ, Brokstad KA (1999) The postvaccination antibody response to influenza virus proteins. *Apmis* 107(3): 289-296.
- Danke NA, Kwok WW (2003) HLA class II-restricted CD4(+) T cell responses directed against influenza viral antigens postinfluenza vaccination. *Journal of Immunology* 171(6): 3163-3169.
- de St. Groth F. Antigenic, adaptive, and adsorptive variants of the influenza A hemagglutinin; 1977 March 21-23; Baden near Vienna.
- Deboer GF, Back W, Osterhaus aDME (1990) An Elisa for Detection of Antibodies against Influenza-a Nucleoprotein in Humans and Various Animal Species. *Archives of Virology* 115(1-2): 47-61.
- Dieckmann U, Metz JAJ, Sabelis MW, Sigmund K (2002) Adaptive Dynamics of Infectious Diseases: In Pursuit of Virulence Management; Dieckmann U, Metz JAJ, editors. Cambridge, U.K.: Cambridge University Press.
- Drummond AJ, Ho SYW, Phillips MJ, Rambaut A (2006) Relaxed phylogenetics and dating with confidence. *Plos Biology* 4(5): 699-710.
- Dushoff J, Plotkin JB, Levin SA, Earn DJD (2004) Dynamical resonance can account for seasonality of influenza epidemics. *Proceedings of the National Academy of Sciences of the United States of America* 101(48): 16915-16916.
- Eames KTD, Keeling MJ (2002) Modeling dynamic and network heterogeneities in the spread of sexually transmitted diseases. *Proceedings of the National Academy of Sciences of the United States of America* 99(20): 13330-13335.
- Earn DJD, Dushoff J, Levin SA (2002) Ecology and evolution of the flu. *Trends in Ecology & Evolution* 17(7): 334-340.
- Enserink M (2004) Infectious diseases - Bird flu infected 1000, Dutch researchers say. *Science* 306(5296): 590-590.
- Enserink M (2006) What came before 1918? Archaeovirologist offers a first glimpse. *Science* 312: 1725.
- Ferguson NM, Galvani AP, Bush RM (2003) Ecological and immunological determinants of influenza evolution. *Nature* 422(6930): 428-433.
- Ferguson NM, Cummings DAT, Cauchemez S, Fraser C, Riley S et al. (2005) Strategies for containing an emerging influenza pandemic in Southeast Asia. *Nature* 437(7056): 209-214.
- Fitch WM, Leiter JME, Li X, Palese P (1991) Positive Darwinian evolution in human influenza A viruses. *Proceedings of the National Academy of Sciences of the United States of America* 88: 4270-4274.
- Fleury D, Barrere B, Bizebard T, Daniels RS, Skehel JJ et al. (1999) A complex of influenza hemagglutinin with a neutralizing antibody that binds outside the virus receptor binding site. *Nature Structural Biology* 6(6): 530-534.

- Frank AL, Taber LH, Wells JM (1983) Individuals infected with two subtypes of influenza A virus in the same season. *The Journal of Infectious Diseases* 147(1): 120-124.
- Frank SA (2002) *Immunology and Evolution of Infectious Disease*. Princeton: Princeton University Press.
- Friedman R, Hughes AL (2005) Codon volatility as an indicator of positive selection: Data from eukaryotic genome comparisons. *Molecular Biology and Evolution* 22(3): 542-546.
- Gallagher PJ, Henneberry JM, Sambrook JF, Gething MJH (1992) Glycosylation Requirements for Intracellular-Transport and Function of the Hemagglutinin of Influenza-Virus. *J Virol* 66(12): 7136-7145.
- Gambaryan A, Webster R, Matrosovich M (2002a) Differences between influenza virus receptors on target cells of duck and chicken. *Archives of Virology* 147(6): 1197-1208.
- Gambaryan AS, Yamnikova SS, Lvov DK, Robertson JS, Webster RG et al. (2002b) Differences in receptor specificity between the influenza viruses of duck, chicken, and human. *Molecular Biology* 36(3): 429-435.
- Gerhard W, Yewdell J, Frankel ME, Webster R (1981) Antigenic Structure of Influenza-Virus Hemagglutinin Defined by Hybridoma Antibodies. *Nature* 290(5808): 713-717.
- Gill PW, Murphy AM (1977) Naturally acquired immunity to influenza type A: a further prospective study. *Medical Journal of Australia* 2: 761-765.
- Gog JR, Grenfell BT (2002) Dynamics and selection of many-strain pathogens. *Proceedings of the National Academy of Sciences of the United States of America* 99(26): 17209-17214.
- Gog JR, Rimmelzwaan GF, Osterhaus ADME, Grenfell BT (2003) Population dynamics of rapid fixation in cytotoxic T lymphocyte escape mutants of influenza A. *Proceedings of the National Academy of Sciences of the United States of America* 100(19): 11143-11147.
- Greene SK, Ionides EL, Wilson ML (2006) Patterns of influenza-associated mortality among US elderly by geographic region and virus subtype, 1968-1998. *American Journal of Epidemiology* 163(4): 316-326.
- Gregory V, Lim W, Cameron K, Bennett M, Marozin S et al. (2001) Infection of a child in Hong Kong by an influenza A H3N2 virus closely related to viruses circulating in European pigs. *Journal of General Virology* 82: 1397-1406.
- Grenfell BT, Lonergan ME, Harwood J (1992) Quantitative Investigations of the Epidemiology of Phocine Distemper Virus (Pdv) in European Common Seal Populations. *Science of the Total Environment* 115(1-2): 15-29.
- Grenfell BT, Pybus OG, Gog JR, Wood JLN, Daly JM et al. (2004) Unifying the epidemiological and evolutionary dynamics of pathogens. *Science* 303(5656): 327-332.
- Guan Y, Peiris M, Kong KF, Dyrting KC, Ellis TM et al. (2002) H5N1 influenza viruses isolated from Geese in southeastern China: Evidence for genetic reassortment and interspecies transmission to ducks. *Virology* 292(1): 16-23.
- Gubareva LV, Novikov DV, Hayden FG (2002) Assessment of hemagglutinin sequence heterogeneity during influenza virus transmission in families. *The Journal of Infectious Diseases* 186: 1575-1581.
- Gulati U, Hwang CC, Venkatramani L, Gulati S, Stray SJ et al. (2002) Antibody epitopes on the neuraminidase of a recent H3N2 influenza virus (A/memphis/31/98). *J Virol* 76(23): 12274-12280.
- Guo YJ, Xu XY, Cox NJ (1992a) Human Influenza-a (H1N2) Viruses Isolated from China. *Journal of General Virology* 73: 383-388.
- Guo YJ, Wang M, Kawaoka Y, Gorman O, Ito T et al. (1992b) Characterization of a New Avian-Like Influenza-a Virus from Horses in China. *Virology* 188(1): 245-255.
- Gupta S, Ferguson NM, Anderson RM (1998) Chaos, persistence, and evolution of strain structure in antigenically diverse infectious agents. *Science* 280: 912-915.
- Halvorson D, Karunakaran D, Senne D, Kelleher C, Bailey C et al. (1983) Epizootiology of Avian Influenza - Simultaneous Monitoring of Sentinel Ducks and Turkeys in Minnesota. *Avian Diseases* 27(1): 77-85.
- Hatchette TF, Walker D, Johnson C, Baker A, Pryor SP et al. (2004) Influenza A viruses in feral Canadian ducks: extensive reassortment in nature. *Journal of General Virology* 85: 2327-2337.
- Heinen PP, de Boer-Luijze EA, Bianchi ATJ (2001) Respiratory and systemic humoral and cellular immune responses of pigs to a heterosubtypic influenza A virus infection. *Journal of General Virology* 82: 2697-2707.

- Hinshaw VS, Bean WJ, Webster RG, Sriram G (1980) Genetic Reassortment of Influenza a-Viruses in the Intestinal-Tract of Ducks. *Virology* 102(2): 412-419.
- Hirst M, Astell CR, Griffith M, Coughlin SM, Moksa M et al. (2004) Novel avian influenza H7N3 strain outbreak, British Columbia. *Emerging Infectious Diseases* 10(12): 2192-2195.
- Holmes EC, Ghedin E, Miller N, Taylor J, Bao Y et al. (2005) Whole-genome analysis of human influenza A virus reveals multiple persistent lineages and reassortment among recent H3N2 viruses. *Public Library of Science* 3(9): 1-10.
- Horimoto T, Rivera E, Pearson J, Senne D, Krauss S et al. (1995) Origin and Molecular-Changes Associated with Emergence of a Highly Pathogenic H5N2 Influenza-Virus in Mexico. *Virology* 213(1): 223-230.
- Houswort. Wj, Spoon MM (1971) Age Distribution of Excess Mortality during a2 Hong Kong Influenza Epidemics Compared with Earlier a2 Outbreaks. *American Journal of Epidemiology* 94(4): 348-&.
- Hurwitz ES, Haber M, Chang A, Shope T, Teo S et al. (2000) Effectiveness of influenza vaccination of day care children in reducing influenza-related morbidity among household contacts. *Jama-Journal of the American Medical Association* 284(13): 1677-1682.
- Ina Y, Gojobori T (1994) Statistical analysis of nucleotide sequences of the hemagglutinin gene of human influenza A viruses. *Proceedings of the National Academy of Sciences of the United States of America* 91: 8388-8392.
- Ito T, Kawaoka Y (1998) Avian Influenza. In: Nicholson KG, Webster RG, Hay AJ, editors. *Textbook of Influenza*. Malden, MA: Blackwell Science. pp. 126-136.
- Ito T, Kawaoka Y (2000) Host-range barrier of influenza A viruses. *Veterinary Microbiology* 74(1-2): 71-75.
- Ito T, Kawaoka Y, Nomura A, Otsuki K (1999) Receptor specificity of influenza A viruses from sea mammals correlates with lung sialyloligosaccharides in these animals. *Journal of Veterinary Medical Science* 61(8): 955-958.
- Jameson J, Cruz J, Ennis FA (1998) Human cytotoxic T-lymphocyte repertoire to influenza A viruses. *J Virol* 72(11): 8682-8689.
- Jensen MA, Faruque SM, Mekalanos JJ, Levin BR (2006) Modeling the role of bacteriophage in the control of cholera outbreaks. *Proceedings of the National Academy of Sciences of the United States of America* 103(12): 4652-4657.
- Karasin aI, Landgraf J, Swenson S, Erickson G, Goyal S et al. (2002) Genetic characterization of H1N2 influenza A viruses isolated from pigs throughout the United States. *Journal of Clinical Microbiology* 40(3): 1073-1079.
- Karunakaran D, Hinshaw V, Poss P, Newman J, Halvorson D (1983) Influenza-a Outbreaks in Minnesota Turkeys Due to Subtype-H10N7 and Possible Transmission by Waterfowl. *Avian Diseases* 27(2): 357-366.
- Kauffman S (1995) *At Home in the Universe: The Search for the Laws of Self-Organization and Complexity*. New York: Oxford University Press.
- Kermack WO, McKendrick AG (1927) A contribution to the mathematical theory of epidemics. *Proceedings of the Royal Society of London Series A* 115(700-721).
- Kida H, Yanagawa R, Matsuoka Y (1980) Duck Influenza Lacking Evidence of Disease Signs and Immune-Response. *Infection and Immunity* 30(2): 547-553.
- Kida H, Ito T, Yasuda J, Shimizu Y, Itakura C et al. (1994) Potential for Transmission of Avian Influenza-Viruses to Pigs. *Journal of General Virology* 75: 2183-2188.
- Kim WI, Wu WH, Janke B, Yoon KJ (2006) Characterization of the humoral immune response of experimentally infected and vaccinated pigs to swine influenza viral proteins. *Archives of Virology* 151(1): 23-36.
- Knossow M, Daniels RS, Douglas aR, Skehel JJ, Wiley DC (1984) 3-Dimensional Structure of an Antigenic Mutant of the Influenza-Virus Hemagglutinin. *Nature* 311(5987): 678-680.
- Knossow M, Gaudier M, Douglas A, Barrere B, Bizebard T et al. (2002) Mechanism of neutralization of influenza virus infectivity by antibodies. *Virology* 302(2): 294-298.
- Koelle K, Pascual M (2004) Disentangling extrinsic from intrinsic factors in disease dynamics: A nonlinear time series approach with an application to cholera. *American Naturalist* 163(6): 901-913.
- Koelle K, Cobey S, Grenfell B, Pascual M (in prep) Epochal evolution shapes the phylodynamics of interpandemic influenza.

- Korber B, LaBute M, Yusim K (2006) Immunoinformatics comes of age. *Plos Computational Biology* 2(6): 0484-0492.
- Kosakovsky Pond SL, Frost SDW (2005) Datamonkey: rapid detection of selective pressure on individual sites of codon alignments. *Bioinformatics* 21(10): 2531-2533.
- Kosakovsky Pond SL, Frost SDW, Muse SV (2005) HyPhy: hypothesis testing using phylogenies. *Bioinformatics* 21(5): 676-679.
- Kostolansky F, Vareckova E, Betakova T, Mucha V, Russ G et al. (2000) The strong positive correlation between effective affinity and infectivity neutralization of highly cross-reactive monoclonal antibody IIB4, which recognizes antigenic site B on influenza A virus haemagglutinin. *Journal of General Virology* 81: 1727-1735.
- Kumar S, Tamura K, Nei M (2004) MEGA3: Integrated software for molecular evolutionary genetics analysis and sequence alignment. *Briefings in Bioinformatics* 5(2): 150-163.
- Lavenu a, Valleron AJ, Carrat F (2004) Exploring cross-protection between influenza strains by an epidemiological model. *Virus Research* 103(1-2): 101-105.
- Lavenu a, Leruez-Ville M, Chaix ML, Boelle PY, Rogez S et al. (2006) Detailed analysis of the genetic evolution of influenza virus during the course of an epidemic. *Epidemiology and Infection* 134(3): 514-520.
- Lee CW, Senne DA, Suarez DL (2004) Effect of vaccine use in the evolution of Mexican lineage H5N2 avian influenza virus. *J Virol* 78(15): 8372-8381.
- Lee CW, Suarez DL, Tumpey TM, Sung HW, Kwon YK et al. (2005) Characterization of highly pathogenic H5N1 avian influenza A viruses isolated from South Korea. *J Virol* 79(6): 3692-3702.
- Liang SH, Mozdzanowska K, Palladino G, Gerhard W (1994) Heterosubtypic Immunity to Influenza Type-a Virus in Mice - Effector Mechanisms and Their Longevity. *Journal of Immunology* 152(4): 1653-1661.
- Lin J, Andreasen V, Levin SA (1999) Dynamics of influenza A drift: the linear three-strain model. *Mathematical Biosciences* 162(1-2): 33-51.
- Lin YP, Shaw M, Gregory V, Cameron K, Lim W et al. (2000) Avian-to-human transmission of H9N2 subtype influenza A viruses: Relationship between H9N2 and H5N1 human isolates. *Proceedings of the National Academy of Sciences of the United States of America* 97(17): 9654-9658.
- Lindstrom SE, Cox NJ, Klimov A (2004) Genetic analysis of human H2N2 and early H3N2 influenza viruses, 1957-1972: evidence for genetic divergence and multiple reassortment events. *Virology* 328: 101-119.
- Linnemann T, Jung G, Walden P (2000) Detection and quantification of CD4(+) T cells with specificity for a new major histocompatibility complex class II-restricted influenza A virus matrix protein epitope in peripheral blood of influenza patients. *J Virol* 74(18): 8740-8743.
- Liu M, Guan Y, Peiris M, He S, Webby RJ et al. (2003a) The quest of influenza A viruses for new hosts. *Avian Diseases* 47: 849-856.
- Liu M, He SQ, Walker D, Zhou NN, Perez DR et al. (2003b) The influenza virus gene pool in a poultry market in South Central China. *Virology* 305(2): 267-275.
- Lloyd-Smith JO, Schreiber SJ, Kopp PE, Getz WM (2005) Superspreading and the effect of individual variation on disease emergence. *Nature* 438(7066): 355-359.
- LoGiudice K, Ostfeld RS, Schmidt KA, Keesing F (2003) The ecology of infectious disease: Effects of host diversity and community composition on Lyme disease risk. *Proceedings of the National Academy of Sciences of the United States of America* 100(2): 567-571.
- Longini IM, Nizam A, Xu SF, Ungchusak K, Hanshaoworakul W et al. (2005) Containing pandemic influenza at the source. *Science* 309(5737): 1083-1087.
- Ludwig S, Hausteina, Kaleta EF, Scholtissek C (1994) Recent Influenza-a (H1N1) Infections of Rigs and Turkeys in Northern Europe. *Virology* 202(1): 281-286.
- Ludwig S, Stitz L, Planz O, Van H, Fitch WM et al. (1995) European Swine Virus as a Possible Source for the Next Influenza Pandemic. *Virology* 212(2): 555-561.
- Marozin S, Gregory V, Cameron K, Bennett M, Valette M et al. (2002) Antigenic and genetic diversity among swine influenza A H1N1 and H1N2 viruses in Europe. *Journal of General Virology* 83: 735-745.
- Marshall DR, Turner SJ, Belz GT, Wingo S, Andreansky S et al. (2001) Measuring the diaspora for virus-specific CD8(+) T cells. *Proceedings of the National Academy of Sciences of the United States of America* 98(11): 6313-6318.

- Masurel N, Marine WM (1973) Recycling of Asian and Hong-Kong Influenza a Virus Hemagglutinins in Man. *American Journal of Epidemiology* 97(1): 44-49.
- Matrosovich M, Gao P, Kawaoka Y (1998) Molecular mechanisms of serum resistance of human influenza H3N2 virus and their involvement in virus adaptation in a new host. *J Virol* 72(8): 6373-6380.
- Matrosovich M, Zhou N, Kawaoka Y, Webster R (1999) The surface glycoproteins of H5 influenza viruses isolated from humans, chickens, and wild aquatic birds have distinguishable properties. *J Virol* 73(2): 1146-1155.
- Matrosovich M, Tuzikov A, Bovin N, Gambaryan A, Klimov A et al. (2000) Early alterations of the receptor-binding properties of H1, H2, and H3 avian influenza virus hemagglutinins after their introduction into mammals. *J Virol* 74(18): 8502-8512.
- Matrosovich MN, Krauss S, Webster RG (2001) H9N2 influenza a viruses from poultry in Asia have human virus-like receptor specificity. *Virology* 281(2): 156-162.
- Meyers LA, Anceel FD, Lachmann M (2005) Evolution of genetic potential. *Plos Computational Biology* 1(3): 236-243.
- Monto aS, Kendal aP (1973) Effect of Neuraminidase Antibody on Hong-Kong Influenza. *Lancet* 1(7804): 623-625.
- Monto aS, Davenport.Fm, Napier Ja, Francis T (1970) Modification of an Outbreak of Influenza in Tecumseh, Michigan by Vaccination of Schoolchildren. *Journal of Infectious Diseases* 122(1-2): 16-&.
- Myers KP, Olsen CW, Setterquist SF, Capuano AW, Donham KJ et al. (2006) Are swine workers in the United States at increased risk of infection with zoonotic influenza virus? *Clinical Infectious Diseases* 42(1): 14-20.
- Nakajima K, Nobusawa E, Tonegawa K, Nakajima S (2003) Restriction of amino acid change in influenza A virus H3HA: Comparison of amino acid changes observed in nature and in vitro. *J Virol* 77(18): 10088-10098.
- Nakajima K, Nobusawa E, Nagy A, Nakajima S (2005) Accumulation of amino acid substitutions promotes irreversible structural changes in the hemagglutinin of human influenza AH3 virus during evolution. *J Virol* 79(10): 6472-6477.
- Nakajima S, Nobusawa E, Nakajima K (2000) Variation in response among individuals to antigenic sites on the HA protein of human influenza virus may be responsible for the emergence of drift strains in the human population. *Virology* 274: 220-231.
- Nakajima S, Nakamura K, Nishikawa F, Nakajima K (1991) Genetic-Relationship between the Ha Genes of Type-a Influenza-Viruses Isolated in Off-Seasons and Later Epidemic Seasons. *Epidemiology and Infection* 106(2): 383-395.
- Nelson MI, Simonsen L, Viboud C, Miller MA, Taylor J et al. (Submitted) Stochastic processes are key determinants of the short-term evolution of influenza A virus.
- Nobusawa E, Ishihara H, Morishita T, Sato K, Nakajima K (2000) Change in receptor-binding specificity of recent human influenza A viruses (H3N2): A single amino acid change in hemagglutinin altered its recognition of sialylogosaccharides. *Virology* 278: 587-596.
- Olofsson S, Kumlin U, Dimock K, Arnberg N (2005) Avian influenza and sialic acid receptors: more than meets the eye? *Lancet Infectious Diseases* 5(3): 184-188.
- Olsen CW, Carey S, Hinshaw L, Karasin AI (2000) Virologic and serologic surveillance for human, swine and avian influenza virus infections among pigs in the north-central United States. *Archives of Virology* 145(7): 1399-1419.
- Olsen CW, Brammer L, Easterday BC, Arden N, Belay E et al. (2002) Serologic evidence of H1 swine influenza virus infection in swine farm residents and employees. *Emerging Infectious Diseases* 8(8): 814-819.
- Ottis K, Sidoli L, Bachmann Pa, Webster RG, Kaplan MM (1982) Human Influenza-a Viruses in Pigs - Isolation of a H3N2-Strain Antigenically Related to a/England/42/72 and Evidence for Continuous Circulation of Human Viruses in the Pig-Population. *Archives of Virology* 73(2): 103-108.
- Ozaki H, Sugiura T, Sugita S, Imagawa H, Kida H (2001) Detection of antibodies to the nonstructural protein (NS1) of influenza A virus allows distinction between vaccinated and infected horses. *Veterinary Microbiology* 82(2): 111-119.
- Palese P (2004) Influenza: old and new threats. *Nature Medicine* 10(12): S82-S87.
- Pascual M (2005) Computational ecology: From the complex to the simple and back. *Public Library of Science* 1(2): 0101-0105.

- Pease CM (1987) An Evolutionary Epidemiologic Mechanism, with Applications to Type-a Influenza. *Theoretical Population Biology* 31(3): 422-452.
- Peiris JSM, Guan Y, Markwell D, Ghose P, Webster RG et al. (2001) Cocirculation of avian H9N2 and contemporary "human" H3N2 influenza A viruses in pigs in southeastern China: Potential for genetic reassortment? *J Virol* 75(20): 9679-9686.
- Perez DR, Lim W, Seiler JP, Yi G, Peiris M et al. (2003) Role of quail in the interspecies transmission of H9 influenza A viruses: Molecular changes on HA that correspond to adaptation from ducks to chickens. *J Virol* 77(5): 3148-3156.
- Pillai SK, Pond SLK, Woelk CH, Richman DD, Smith DM (2005) Codon volatility does not reflect selective pressure on the HIV-1 genome. *Virology* 336(2): 137-143.
- Plotkin JB, Dushoff J (2003) Codon bias and frequency-dependent selection on the hemagglutinin epitopes of influenza A virus. *Proceedings of the National Academy of Sciences of the United States of America* 100(12): 7152-7157.
- Plotkin JB, Dushoff J, Fraser HB (2004) Detecting selection using a single genome sequence of *M. tuberculosis* and *P. falciparum*. *Nature* 428(6986): 942-945.
- Rambaut a, Posada D, Crandall KA, Holmes EC (2004) The causes and consequences of HIV evolution. *Nature Reviews Genetics* 5(1): 52-61.
- Rauch EM, Sayama H, Bar-Yam Y (2003) Dynamics and genealogy of strains in spatially extended host-pathogen models. *Journal of Theoretical Biology* 221(4): 655-664.
- Reitter JN, Means RE, Desrosiers RC (1998) A role for carbohydrates in immune evasion in AIDS. *Nature Medicine* 4(6): 679-684.
- Rimmelzwaan GF, Boon ACM, Voeten JTM, Berkhoff EGM, Fouchier RAM et al. (2004) Sequence variation in the influenza A virus nucleoprotein associated with escape from cytotoxic T lymphocytes. *Virus Research* 103(1-2): 97-100.
- Rogers GN, Paulson JC, Daniels RS, Skehel JJ, Wilson Ia et al. (1983) Single Amino-Acid Substitutions in Influenza Hemagglutinin Change Receptor-Binding Specificity. *Nature* 304(5921): 76-78.
- Ronquist F, Huelsenbeck JP (2003) MrBayes 3: Bayesian phylogenetic inference under mixed models. *Bioinformatics* 19(12): 1572-1574.
- Roy M, Pascual M (2006) On representing network heterogeneities in the incidence rate of simple epidemic models. *Ecological Complexity* 3(1): 80-90.
- Sanna PP, Ramiro-Ibanez F, De Logu A (2000) Synergistic interactions of antibodies in rate of virus neutralization. *Virology* 270(2): 386-396.
- Sato K, Morishita T, Nobusawa E, Tonegawa K, Sakae K et al. (2004) Amino-acid change on the antigenic region B1 of H3 haemagglutinin may be a trigger for the emergence of drift strain of influenza A virus. *Epidemiology and Infection* 132(3): 399-406.
- Sato K, Morishita T, Nobusawa E, Suzuki Y, Miyazaki Y et al. (2000) Surveillance of influenza viruses isolated from travellers at Nagoya International Airport. *Epidemiology and Infection* 124(3): 507-514.
- Scholtissek C, Hinshaw VS, Olsen CW (1998) Influenza in Pigs and their Role as the Intermediate Host. In: Nicholson KG, Webster RG, Hay AJ, editors. *Textbook of Influenza*. Malden, MA: Blackwell Science. pp. 137-145.
- Schulze IT (1997) Effects of glycosylation on the properties and functions of influenza virus hemagglutinin. *Journal of Infectious Diseases* 176: S24-S28.
- Shinya K, Ebina M, Yamada S, Ono M, Kasai N et al. (2006) Influenza virus receptors in the human airway. *Nature* 440(7083): 435-436.
- Shortridge KF (2005) Avian influenza viruses in Hong Kong: zoonotic considerations. In: Schrijver RS, Koch G, editors. *Avian Influenza: Prevention and Control*: Springer. pp. 9-18.
- Shortridge KF, Webster RG, Butterfield WK, Campbell CH (1977) Persistence of Hong-Kong Influenza-Virus Variants in Pigs. *Science* 196(4297): 1454-1455.
- Sivanandan V, Halvorson Da, Laudert E, Senne Da, Kumar MC (1991) Isolation of H13N2 Influenza-a Virus from Turkeys and Surface-Water. *Avian Diseases* 35(4): 974-977.
- Skehel JJ, Wiley DC (2000) Receptor binding and membrane fusion in virus entry: The influenza hemagglutinin. *Annual Review of Biochemistry* 69: 531-569.
- Smirnov Ya, Lipatov AS, Gitelman AK, Okuno Y, Van Beek R et al. (1999) An epitope shared by the hemagglutinins of H1, H2, H5, and H6 subtypes of influenza A virus. *Acta Virologica* 43(4): 237-244.

- Smith DJ, Lapedes AS, de Jong JC, Bestebroer TM, Rimmelzwaan GF et al. (2004) Mapping the antigenic and genetic evolution of influenza virus. *Science* 305(5682): 371-376.
- Smith NM, Bresee JS, Shay DK, Uyeki TM, Cox NJ et al. (2006) Prevent and Control of Influenza: Recommendations of the Advisory Committee on Immunization Practices (ACIP). *MMWR Weekly* 55 (Early Release): 1-41.
- Sonoguchi T, Naito H, Hara M, Takeuchi Y, Fukumi H (1985) Cross-Subtype Protection in Humans during Sequential, Overlapping, and or Concurrent Epidemics Caused by H3N2 and H1N1 Influenza-Viruses. *Journal of Infectious Diseases* 151(1): 81-88.
- Sonoguchi T, Sakoh M, Kunita N, Satsuta K, Noriki H et al. (1986) Reinfection with influenza A (H2N2, H3N2, and H1N1) viruses in soldiers and students in Japan. *The Journal of Infectious Diseases* 153(1): 33-40.
- Stewart JJ, Watts P, Litwin S (2001) An algorithm for mapping positively selected members of quasispecies-type viruses. *BMC Bioinformatics* 2(1).
- Stoletzki N, Welch J, Hermisson J, Eyre-Walker A (2005) A dissection of volatility in yeast. *Molecular Biology and Evolution* 22(10): 2022-2026.
- Suarez DL (2000) Evolution of avian influenza viruses. *Veterinary Microbiology* 74(1-2): 15-27.
- Suarez DL, Senne DA, Banks J, Brown IH, Essen SC et al. (2004) Recombination resulting in virulence shift in avian influenza outbreak, Chile. *Emerging Infectious Diseases* 10(4): 693-699.
- Suzuki Y (2004) Three-dimensional window analysis for detecting positive selection at structural regions of proteins. *Molecular Biology and Evolution* 21(12): 2352-2359.
- Swayne DE (2005) Avian influenza, Poultry vaccines (02) [2]. ProMED-mail: International Society for Infectious Diseases.
- Swinton J, Harwood J, Grenfell BT, Gilligan CA (1998) Persistence thresholds for phocine distemper virus infection in harbour seal *Phoca vitulina* metapopulations. *Journal of Animal Ecology* 67(1): 54-68.
- Thomas PG, Keating R, Hulse-Post DJ, Doherty PC (2006) Cell-mediated protection in influenza infection. *Emerging Infectious Diseases* 12(1): 48-54.
- Thompson WW, Shay DK, Weintraub E, Brammer L, Cox N et al. (2003) Mortality associated with influenza and respiratory syncytial virus in the United States. *Jama-Journal of the American Medical Association* 289(2): 179-186.
- Tria F, Lassig M, Peliti L, Franz S (2005) A minimal stochastic model for influenza evolution. *Journal of Statistical Mechanics-Theory and Experiment*: -.
- Tsuchiya E, Sugawara K, Hongo S, Matsuzaki Y, Muraki Y et al. (2002) Effect of addition of new oligosaccharide chains to the globular head of influenza A/H2N2 virus haemagglutinin on the intracellular transport and biological activities of the molecule. *Journal of General Virology* 83: 1137-1146.
- van Ballegooijen WM, Boerlijst MC (2004) Emergent trade-offs and selection for outbreak frequency in spatial epidemics. *Proceedings of the National Academy of Sciences of the United States of America* 101(52): 18246-18250.
- Van Reeth K, Brown I, Essen S, Pensaert M (2004) Genetic relationships, serological cross-reaction and cross-protection between H1N2 and other influenza - A virus subtypes endemic in European pigs. *Virus Research* 103(1-2): 115-124.
- Venkatramani L, Bochkareva E, Lee JT, Laver UGWG, Bochkarev A et al. (2006) An epidemiologically significant epitope of a 1998 human influenza virus neuraminidase forms a highly hydrated interface in the NA-Antibody complex. *Journal of Molecular Biology* 356(3): 651-663.
- Viboud C, Alonso WJ, Simonsen L (2006a) Influenza in tropical regions. *Plos Medicine* 3(4): 468-471.
- Viboud C, Bjornstad ON, Smith DL, Simonsen L, Miller MA et al. (2006b) Synchrony, waves, and spatial hierarchies in the spread of influenza. *Science* 312(5772): 447-451.
- Viboud C, Boelle PY, Cauchemez S, Lavenue A, Valleron AJ et al. (2004) Risk factors of influenza transmission in households. *British Journal of General Practice* 54(506): 684-689.
- Viboud C, Grais RF, Lafont BAP, Miller MA, Simonsen L et al. (2005) Multinational impact of the 1968 Hong Kong influenza pandemic: Evidence for a smoldering pandemic. *Journal of Infectious Diseases* 192(2): 233-248.
- Volterra V (1926) Fluctuations in the abundance of a species considered mathematically. *Nature* 118(558-560).
- Wang M-L, Skehel JJ, Wiley DC (1986) Comparative analysis of the specificities of anti-influenza hemagglutinin antibodies in human sera. *Journal of Virological* 57: 124-128.

- Webby RJ, Rossow K, Erickson G, Sims Y, Webster R (2004) Multiple lineages of antigenically and genetically diverse influenza A virus co-circulate in the United States swine population. *Virus Research* 103(1-2): 67-73.
- Webster RG, Laver WG (1980) Determination of the Number of Nonoverlapping Antigenic Areas on Hong-Kong (H3N2) Influenza-Virus Hemagglutinin with Monoclonal-Antibodies and the Selection of Variants with Potential Epidemiological Significance. *Virology* 104(1): 139-148.
- Webster RG, Hulse DJ (2004) Microbial adaptation and change: avian influenza. *Rev Sci Tech Oie* 23(2): 453-465.
- Webster RG, Bean WJ, Gorman OT, Chambers TM, Kawaoka Y (1992) Evolution and Ecology of Influenza-A Viruses. *Microbiological Reviews* 56(1): 152-179.
- Wei XP, Decker JM, Wang SY, Hui HX, Kappes JC et al. (2003) Antibody neutralization and escape by HIV-1. *Nature* 422(6929): 307-312.
- Wiley DC, Wilson Ia, Skehel JJ (1981) Structural Identification of the Antibody-Binding Sites of Hong-Kong Influenza Hemagglutinin and Their Involvement in Antigenic Variation. *Nature* 289(5796): 373-378.
- Wilson Ia, Cox NJ (1990) Structural Basis of Immune Recognition of Influenza-Virus Hemagglutinin. *Annual Review of Immunology* 8: 737-&.
- Wong WSW, Sainudiin R, Nielsen R (2006) Identification of physicochemical selective pressure on protein encoding nucleotide sequences. *Bmc Bioinformatics* 7: -.
- Wuethrich B (2003) Infectious disease - Chasing the fickle swine flu. *Science* 299(5612): 1502-+.
- Xu XY, Smith CB, Mungall BA, Lindstrom SE, Hall HE et al. (2002) Intercontinental circulation of human influenza A(H1N2) reassortant viruses during the 2001-2002 influenza season. *Journal of Infectious Diseases* 186(10): 1490-1493.
- Yamnikova SS, Mandler J, Bekhochir ZH, Dachtzeren P, Ludwig S et al. (1993) A Reassortant H1N1 Influenza-a Virus Caused Fatal Epizootics among Camels in Mongolia. *Virology* 197(2): 558-563.
- Yang Z (1997) PAML: a program package for phylogenetic analysis by maximum likelihood. *Computer Applications in BioSciences* 13: 555-556.
- Zhang JZ (2005) On the evolution of codon volatility. *Genetics* 169(1): 495-501.
- Zhou NN, Senne DA, Landgraf JS, Swenson SL, Erickson G et al. (2000) Emergence of H3N2 reassortant influenza A viruses in North American pigs. *Veterinary Microbiology* 74(1-2): 47-58.