21 Montague, C.T. *et al.* (1997) Congenital leptin deficiency is associated with severe early-onset obesity in humans. *Nature* 387, 903–908
22 Farooqi, I.S. *et al.* (1999) Effects of recombinant leptin therapy in a child with congenital leptin deficiency. *N. Engl. J. Med.* 341, 879–884
23 Heymsfield, S.B. *et al.* (1999) Recombinant leptin for weight loss in obese and lean adults. *J. Am. Med. Assoc.* 282, 1568–1575
24 Mantzoros, C.S. (1999) The role of leptin in human obesity and disease: a review of current evidence. *Ann. Intern. Med.* 130, 671–680
25 Ahima, R.S. *et al.* (1996) Role of leptin in the neuroendocrine response to fasting. *Nature* 382, 250–252
26 Reinhardt, H.W. and Seeliger, E. (2000) Toward an integrative concept of control of total body sodium. *News Physiol. Sci.* 15, 319–325
27 Kurokawa, K. (2001) Salt, kidney and hypertension: why and what to learn from genetic analyses? *Nephron* 89, 369–376
28 James, G.D. and Baker, P.T. (1990) Human population biology and hypertension. Evolutionary and ecological aspects of blood pressure. In *Hypertension: Pathophysiology, Diagnosis and Management* (Laragh, J.H. and Brenner, B.M., eds), pp.137–145, Raven Press
29 Carretero, O.A. and Oparil, S. (2000) Essential hypertension. Part I: definition and etiology. *Circulation* 101, 329–335
30 Carretero, O.A. and Oparil, S. (2000) Essential hypertension. Part II: treatment. *Circulation* 101, 446–453
31 Laragh, J.H. and Brenner, B.M. (1990) *Hypertension. Pathophysiology, Diagnosis and Management*, Raven Press
32 Kaplan, N.M. (2000) Diuretics as a basis of antihypertensive therapy. *Drugs* 59, 21–25
33 Burnier, M. (2001) Angiotensin II type 1 receptor blockers. *Circulation* 103, 904–912
34 Gavras, H. and Brunner, H.R. (2001) Role of angiotensin and its inhibition in hypertension, ischemic heart disease, and heart failure. *Hypertension* 37, 342–345
35 Tarazi, R.C. *et al.* (1970) Long-term thiazide therapy in essential hypertension. Evidence of persistent alteration in plasma volume and renin activity. *Circulation* 41, 709–717
36 Cody, R.J. *et al.* (1983) Renin system activity as a determinant of response to treatment in hypertension and heart failure. *Hypertension* 5, III36–III42
37 Eaton, S.B. and Konner, M. (1985) Paleolithic nutrition. A consideration of its nature and current implications. *N. Engl. J. Med.* 312, 283–289
38 Brunner, H.R. and Gavras, H. (1980) Is the renin system necessary? *Am. J. Med.* 69, 739–745
39 Lippi, G. and Guidi, G. (2000) Lipoprotein(a): from ancestral benefit to modern pathogen. *Q. J. Med.* 93, 75–84
40 Franceschi, C. *et al.* (2000) Inflamm-aging: an evolutionary perspective on immunosenescence. *Ann. New York Acad. Sci.* 908, 244–254
41 Stini, W.A. (1998) Calcium homeostasis and human evolution. *Coll. Antropol.* 22, 411–425
42 Fernandez-Real, J-M. and Ricart, W. (1999) Insulin resistance and inflammation in an evolutionary perspective: the contribution of cytokine genotype/phenotype to thriftiness. *Diabetologia* 42, 1367–1374
43 Francis, G.S. (2001) Pathophysiology of chronic heart failure. *Am. J. Med.* 110, 37S–46S
44 Bristow, M.R. (2000) β-adrenergic receptor blockade in chronic heart failure. *Circulation* 101, 558–569
45 Bond, R.A. (2001) Is paradoxical pharmacology a strategy worth pursuing? *Trends Pharmacol. Sci.* 22, 273–276
46 Schwartz, M.W. *et al.* (2000) Central nervous system control of food intake. *Nature* 404, 661–671

Ecology and evolution of the flu

David J.D. Earn, Jonathan Dushoff and Simon A. Levin

Influenza (flu) is a common infectious disease, but it is unusual in that the primary timescales for disease dynamics (epidemics) and viral evolution (new variants) are roughly the same. Recently, extraordinarily reliable phylogenetic reconstructions of flu virus evolution have been made using samples from both extant and extinct strains. In addition, because of their public health importance, flu epidemics have been monitored throughout the period over which the phylogenetic trees extend. In parallel with this empirical work, theoretical ecologists have developed mathematical and computational models that elucidate many properties of multistrain systems. In the future, to unravel and interpret the complex interactions between ecological and evolutionary forces on flu dynamics, the documented evolution of the virus must be related to the observed population dynamics of the disease. New theoretical insights are also required to simplify model structures and facilitate predictions that can be tested with accessible data.

Published online: 07 May 2002

David J.D. Earn*
Dept of Mathematics and Statistics, McMaster University, Hamilton, Ontario, Canada L8S 4K1.
*e-mail: earn@math.mcmaster.ca

Jonathan Dushoff
Simon A. Levin
Dept of Ecology and Evolutionary Biology, Princeton University, Princeton, NJ 08544-1003, USA.

We have all had the flu and we would rather not have it again. Unfortunately, no matter how many times we have battled the high fever, aches and fatigue, we might be unable to escape infection in the next 'flu season'. Annual flu epidemics are an important cause of mortality, particularly for the elderly and those with chronic illness; and severe flu pandemics – three of which have occurred in the past century – can threaten the lives of even the healthiest individuals. These issues, which represent enormous medical and public health challenges, have deep ecological and evolutionary significance that is only beginning to be appreciated and explored.

What is the flu?

Influenza (flu) is a respiratory infection in mammals and birds. It is caused by an RNA virus in the family Orthomyxoviridae. The virus is divided into three main types (A, B and C), which are distinguished by differences in two major internal proteins [1] (Fig. 1). Influenza virus type A is the most significant epidemiologically and the most interesting from an ecological and evolutionary standpoint, because it is found in a wide variety of bird and mammal species and can undergo major shifts in immunological properties. Type B is largely confined to humans and is an important cause of morbidity. Little is known about type C, which is not an important source of morbidity. Influenza A is further divided into subtypes based on differences in the membrane proteins hemagglutinin (HA) and neuraminidase (NA), which are the most important targets for the immune system. The notation H*h*N*n* is used to refer to the subtype comprising the *h*th discovered HA protein and the *n*th discovered NA protein. There are currently two subtypes circulating in humans: H1N1 and H3N2. Subtypes are further divided into strains; each genetically distinct virus isolate is usually considered to be a separate strain.

Surprisingly little is known about the transmission of flu, and the importance of airborne transmission relative to droplet transmission remains controversial. However, it seems difficult

  PII: S0169-5347(02)02502-8

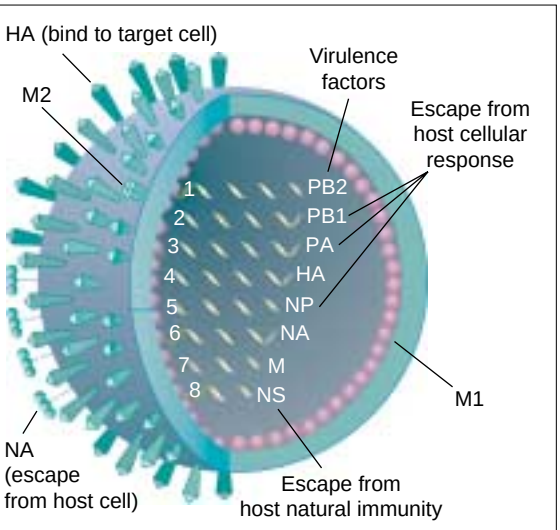

Fig. 1. Schematic representation of an influenza A virus virion (the extracellular infective form of the virus). Each of the eight strands of RNA is labelled with the protein or group of proteins for which it is known to code, and some important protein functions are indicated. The genome of type B also comprises eight strands of RNA, whereas type C has seven strands [1,57]. Types A and B can be distinguished by differences in the internal proteins NP (nucleoprotein) and M1 (one of the matrix proteins denoted together by M on strand 7). In total, the eight segments of influenza A RNA encode 11 known gene products, including the recently discovered PB1-F2 protein [58]. The surface proteins HA (hemagglutinin) and NA (neuraminidase) are the principal targets of the humoral immune response (i.e. the response involving antibodies). Subtypes of influenza A are distinguished by differences in HA and NA. [Adapted from [59]. Copyright (2001) American Association for the Advancement of Science.]

to explain the sometimes explosive spread of flu epidemics without some role for airborne transmission [2]. Schulman and Kilbourne [3] performed one of the few controlled experiments of flu transmission, using mice as a model, and concluded (using only four data points) that airborne transmission was positively correlated with low humidity and low rates of airflow. Other (unorthodox) theories about flu transmission are reviewed by Hope-Simpson [4]. After initial infection, individuals incubate the virus for roughly one to three days before becoming infectious. Infectiousness can precede clinical disease by approximately one day. The infectious period is typically three to six days, whereas the duration of the disease is typically two to seven days [5].

Most individuals recover from flu and are believed to retain lifelong immunity to strains closely related to the infecting strain. Support for this is found in the 1977 re-emergence of H1N1, which was absent in human populations from 1957 to 1977 (a laboratory accident might have been the origin of the reintroduction of this subtype [6]). H1N1 strains collected in 1957 and 1977 were nearly identical, both antigenically and genetically, and the majority of illness occurred in individuals younger than 20 years [7], presumably because most older individuals retained H1N1 immunity from the 1950s.

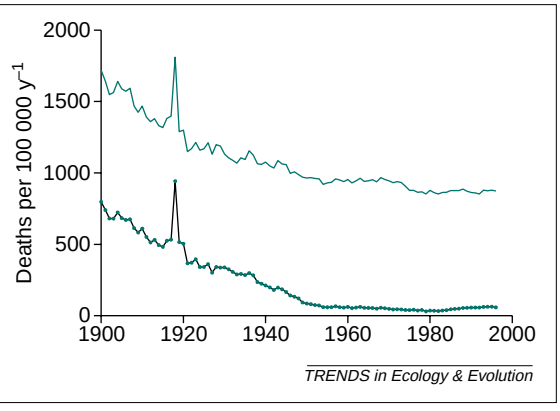

Fig. 2. Annual crude mortality rates in the USA from 1900 to 1996 (from all causes, solid line; from infectious diseases, dotted line). The great flu pandemic of 1918 is very evident, even in the curve showing deaths from all causes. Publication of US annual mortality statistics began in 1900. The area of the USA in which deaths were registered expanded over time until it included all states in 1934. Mortality before 1934 is underestimated. [Adapted from [15]. Copyright (1999) American Medical Association.]

The burden of flu

Flu is an underrated disease. Perhaps because it is a recurrent disease with which we are all familiar, and from which we usually recover naturally [8], it does not hold the terror of AIDS, tuberculosis or malaria. Yet it is a major contributor to mortality and morbidity throughout the world.

The World Health Organization (WHO) estimates that respiratory infections killed more than four million people in 1999 [9], making them the most dangerous category of infectious disease. Flu contributes to many of these deaths, but calculating how much mortality is caused directly and indirectly by flu has proven to be difficult [10]. There are several reasons for this, including: (1) flu predisposes individuals to potentially fatal secondary infection with bacterial pathogens; (2) flu or bacterial superinfections kill in conjunction with other diseases, such as chronic cardiopulmonary conditions [10]; and (3) other respiratory infections, particularly those caused by respiratory syncytial virus (RSV), are probably often mistaken for flu. Estimating the number of excess deaths (deaths above a baseline death rate) caused by flu is commonly done using correlations between seasonal patterns of flu and mortality data and is an active research area [11–13].

It should be noted that flu causes or contributes to death mostly among elderly people and thus leads to the loss of fewer years of life than might be implied by these high mortality estimates. However, flu poses a very real threat to people of all ages with various chronic medical conditions, and flu pandemics can cause heavy mortality in all age groups [14].

Flu ecology

Flu has probably coexisted with humans for more than 400 years [10]. Flu pandemics have had a substantial effect on human mortality. This is evident in Fig. 2, which shows infectious disease mortality in the USA from 1900 to 1996. The large spike is due to the Spanish flu pandemic of 1918–1919. (By contrast, the pandemics of 1957 and 1968 caused widespread morbidity but much less mortality.) The upturn in mortality starting in 1980 is due primarily to HIV but is also the result of respiratory infections in the increasing number of elderly people [15].

In spite of the impact of pandemics, the cumulative morbidity and mortality burden of flu is dominated by regular epidemics. By contrast with excellent time series available for some other diseases (e.g. measles and rubella), incidence data for flu are generally not available because there is no reporting requirement for flu in most countries, and the disease is easily confused with other respiratory infections.

Nevertheless, there are several sources of data that provide information about flu population dynamics (changes in the prevalence of flu infection):

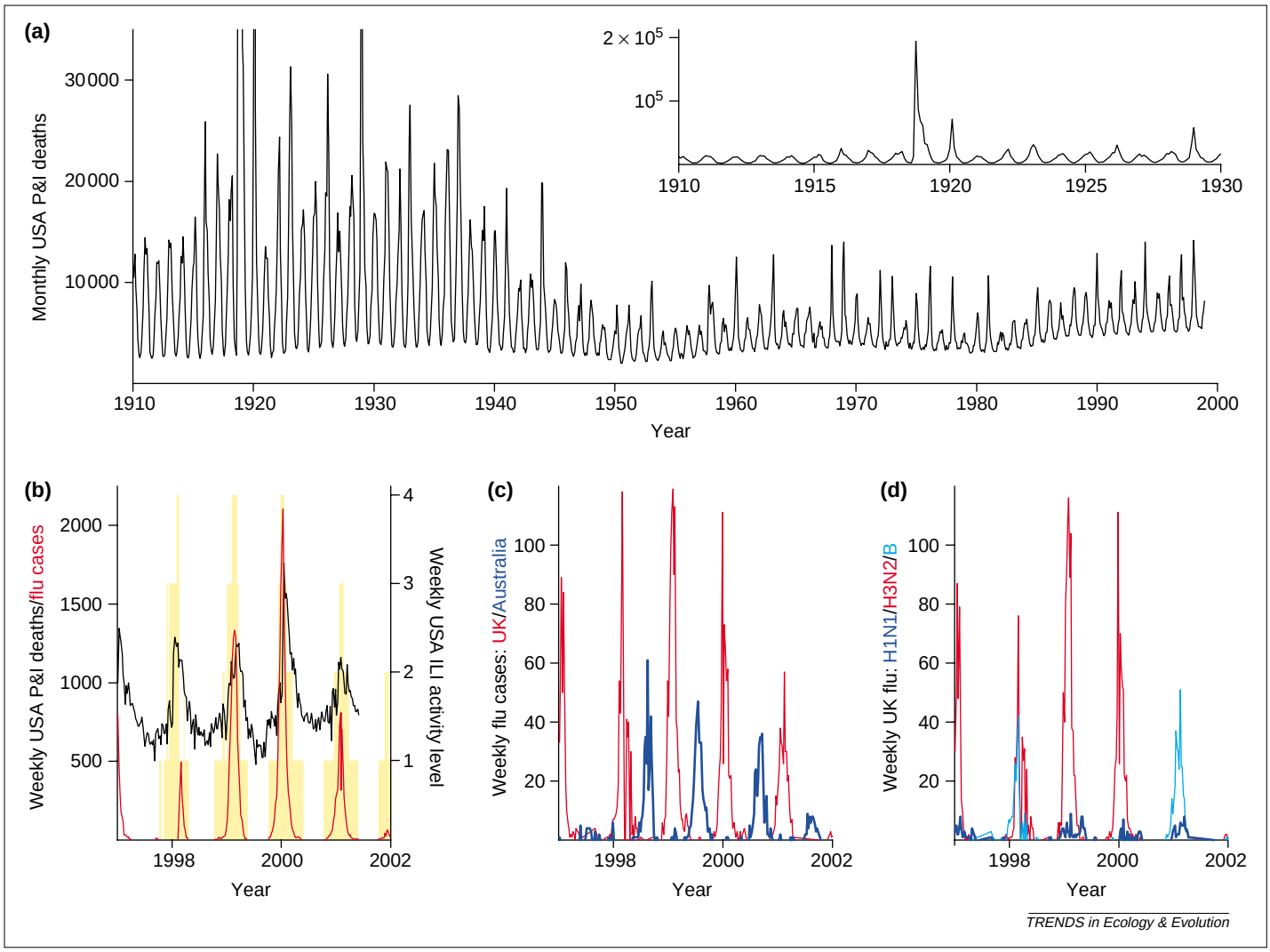

Fig. 3. Flu ecology. (a) Monthly mortality attributed to pneumonia and influenza (P&I deaths) in the USA in the 20th century. The inset plot shows the period 1910–1930 on a much larger scale, revealing the magnitude of the three peaks that extend beyond the top of the main panel: 1918–1919, 1919–1920 and 1928–1929. As in Fig. 2, mortality before 1934 is underestimated. (b) Weekly measures of flu incidence in the USA in the last five years: P&I mortality in 122 cities (black curve), laboratory-confirmed cases (red curve), and influenza-like illness (ILI activity) on a five-point scale from 0 to 4 (yellow shading). (c) Weekly laboratory-confirmed cases in the UK (red curve) and Australia (blue curve) in the last five years. (d) Weekly patterns of relative prevalence of the three major groups of strains in the UK in the last five years (H1N1, blue curve; H3N2, red curve; B, cyan curve). (d) gives a limited impression of the level of aggregation of data in (a–c). The coarsest data are those in (a), where the monthly records span periods during which different subtypes of influenza A have been present in human populations: 1918–1957 (H1N1), 1957–1968 (H2N2), 1968–1977 (H3N2), 1977-present (both H1N1 and H3N2). Data sources: [60], US CDC 122 cities data base (http://www.cdc.gov/epo/dphsi/121hist.htm), WHO FluNet database (http://oms.b3e.jussieu.fr/flunet/).

(1) the monthly records of mortality attributed to pneumonia and influenza (P&I deaths), which have been kept in several countries; (2) weekly records of influenza-like illness (ILI) on a five-point scale, estimated by sentinel physicians around the world and compiled by the WHO since 1995; and (3) weekly laboratory-confirmed cases of flu by type (or subtype). Unfortunately, except for mortality data, these data have very limited spatiotemporal coverage.

As an example of the sort of data available, Fig. 3a shows monthly P&I deaths in the USA from 1910 to 1998, the dominant feature of which is the annual cycle. Figure 3b shows a comparison between ILI, mortality and case data for the USA in the last five years. The strong correlations among these three measures suggest that they are also well correlated with genuine flu incidence. Figure 3c compares data from the UK and Australia, showing that flu epidemics in the Northern and Southern hemispheres are about six months out of phase; this suggests a climatic influence on flu transmission, but the epidemic phase shift might result from the timing of school terms. Finally, Fig. 3d shows patterns of relative prevalence of the three major groups of strains in the UK over the past few years and indicates that epidemics of each influenza A subtype do not necessarily occur every year.

Aside from the obvious annual cycle, the monthly mortality records can be used to identify longer period incidence patterns, as are common in childhood diseases such as measles [16]. Unfortunately, such patterns can be masked by the noisy process of inferring flu incidence from P&I deaths. The situation might improve as longer time series of laboratory confirmations become available.

Flu evolution

The flu virus is unusual in that it continuously undergoes immunologically significant evolution: a flu infection

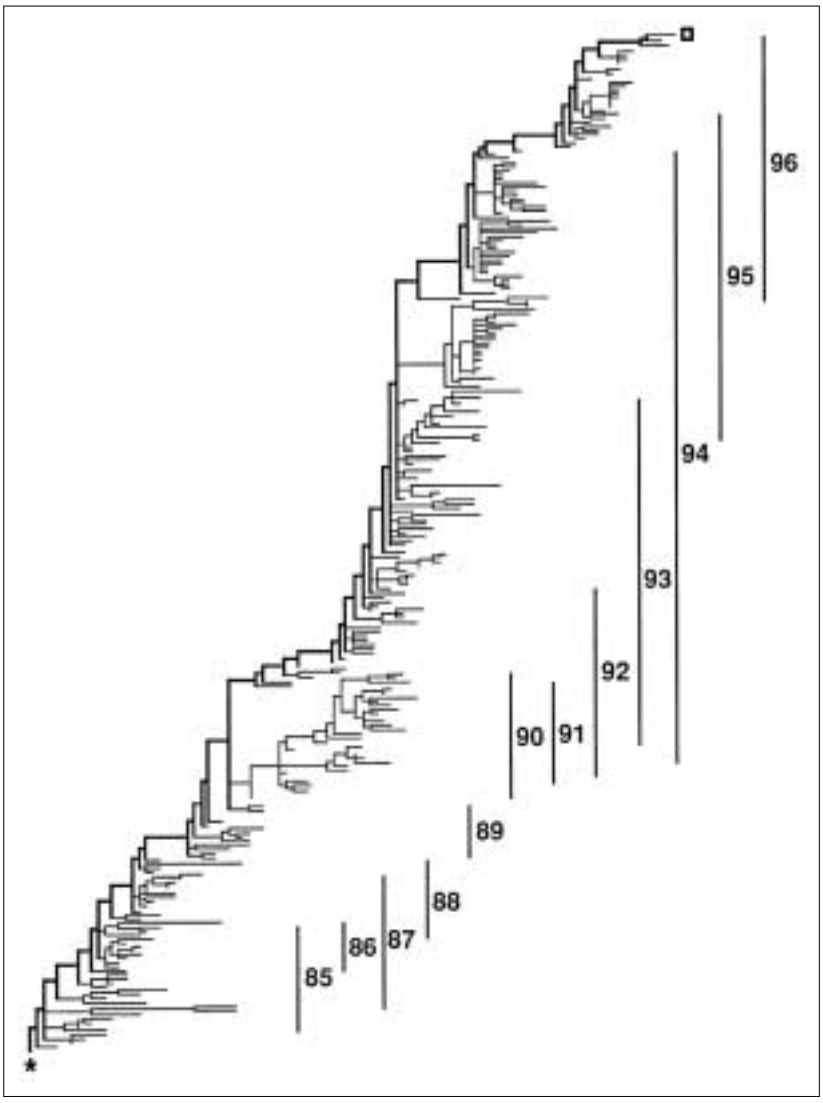

Fig. 4. Flu evolution. Maximum parsimony phylogenetic reconstruction of the evolution of the HA1 domain of the H3N2 subtype of influenza A since 1985. The thick line running from the lower left (* = root) to the upper right (open square) is called the trunk and represents the successful H3N2 lineage. The vertical lines indicate the range of isolates from each flu year (1 October to 30 September). [Reprinted from [20]. Copyright (1997) National Academy of Sciences, USA.]

brings lasting immunity to the infecting strain, but most people are susceptible to a new circulating strain of flu within a few years of infection. Both influenza A and B show what is known as drift evolution: a high rate of viable, immunologically significant mutations. This means that they can drift away from recognition by the immune system by changing the properties of antigenic sites that the immune system must recognize to suppress flu, particularly on the HA protein.

Patterns of change suggest that evolutionary modifications of antigen properties occur rapidly. For sites involved in antigen determination, nucleotide substitutions that change the amino acid are more frequent than synonymous substitutions [17,18], whereas the rest of the sites show the more common pattern of primarily synonymous variation. Thus, the evolution of the virus seems to be strongly influenced by selection for new antigenic variants to escape immune system recognition.

Recent developments in molecular biology and computation have made it possible to produce remarkable phylogenetic reconstructions of flu evolution [19,20], and the sequence data are now publicly available [21]. A striking and poorly understood result from phylogenies of the H3N2 subtype (Fig. 4) is that evolution seems to follow a single track over the long term: at every branching point, it seems that one branch is destined for extinction.

Figure 4 shows that the timescale for significant drift evolution in influenza A is on the order of a year, comparable to the interepidemic interval (the vertical lines in Fig. 4 indicate the range of isolates each year). The population dynamics and evolution of this pathogen thus appear to be inextricably interlinked.

Influenza A, but not influenza B, also undergoes dramatic antigenic changes, known as shifts. Shifts are probably caused by reassortment between different strains of flu within a single host. Different subtypes resulting from shifts are categorized based on the characteristics of their HA and NA proteins, which are encoded by separate segments of the viral RNA genome.

Antigenic shifts probably result from combinations of segments from viruses circulating in the human population and segments from avian viruses. Flu is endemic and tremendously diversified in aquatic birds, which are believed to be a reservoir for new variants that are transmitted to other birds and mammals [22]. It has also been proposed that pigs might be important in facilitating such shifts, since it is known that both avian and human viruses grow well in pigs under certain conditions [23,24]. It has recently been established that avian flu viruses can infect humans without passage through an intermediate host and without acquiring gene segments from human flu viruses. In 1997, 18 people (six of whom died) were infected with avian H5N1 virus, and in 1999 two people were infected with avian H9N2 virus [25]. Neither of these avian viruses turned out to be directly transmissible among humans, but these events heighten anticipation of future shifts in human flu viruses.

There were three major antigenic shifts in the 20th century, and Webster [22] estimates that there have been 10 to 20 pandemics (presumably because of shifts) in the past 250 years. The Spanish flu pandemic in 1918 was apparently caused by the reappearance of human flu with subtype H1N1; recent and current work on archival autopsy tissue of two victims and lung tissue from a third victim who was interred in permafrost is revealing the entire genetic sequence of the 1918 virus [26–29]. Antigenic shifts have since occurred in 1957 (Asian flu), when the H2N2 subtype appeared, and in 1968 (Hong Kong flu), when H3N2 appeared. An H1N1 virus closely resembling the pre-1957 virus reappeared in 1977, and since then both H1N1 and H3N2 have co-occurred in human populations. The huge potential effects of these shifts (such as 20–40 million deaths in 1918–1919) have motivated much of the research on the virus.

Box 1. The basic SIR model

In the standard SIR model for infectious diseases, individuals are categorized according to infection status: susceptible (S), infectious (I) or recovered and immune (R). A formal model is obtained by specifying rates of flow among the three categories. The usual equations are (Eqn I–III) [a]

$$\frac{dS}{dt} = \nu N - (\beta I + \mu)S \quad \text{[Eqn I]}$$

$$\frac{dI}{dt} = \beta IS - (\gamma + \mu)I \quad \text{[Eqn II]}$$

$$\frac{dR}{dt} = \gamma I - \mu R \quad \text{[Eqn III]}$$

where the parameters are birth rate (ν), total population size (N), transmission rate (β), death rate (μ) and recovery rate (γ). The most common extension of the model is to allow a delay between inoculation and infectiousness (a latent period) by adding an 'exposed' (E) compartment. This is generally not done for flu, because other factors are considered to be much more important (Box 2).

Reference

a Anderson, R.M. and May, R.M. (1991) *Infectious Diseases of Humans: Dynamics and Control*, Oxford University Press

Interventions

A vaccine for flu was first tested in 1935–1936, only two years after the virus was isolated from humans; large-scale trials were first conducted in the early 1940s, and the vaccine has been recommended for general use since the early 1960s [30]. Because the virus evolves so quickly, a new vaccine must be designed and distributed each year. At present, in most of the world, vaccination makes use of a trivalent inactived vaccine (TIV) cultured in chicken eggs [31]. The trivalent vaccine is composed of three strains: one of influenza B, one of influenza A subtype H1N1, and one of influenza A subtype H3N2.

Annual vaccination is currently strongly recommended in the USA for people in defined risk groups (mostly elderly people, pregnant women and people with certain chronic conditions). It is additionally recommended for people in close contact with those in high-risk groups, and it is considered appropriate for anybody who wants to be protected against flu [31,32]. The USA Center for Disease Control (CDC) has set goals of vaccinating high proportions of elderly people and others in high-risk groups [32,33].

The WHO recommends the composition of the vaccine for the Northern Hemisphere flu season (November to April) the previous February and allows a similar lead time for preparation of the vaccine for the Southern Hemisphere winter [34]. The lead time is necessary to culture the vaccine but entails some risk of antigenic change in the virus in the meantime [35]. Owing to the long preparation time and varying efficacy of TIV [8], work is underway on a wide variety of other possible preparations, the most promising of which is a cold-adapted live attenuated vaccine [31].

Recent work shows that the efficacy of flu vaccines can depend critically on an individual's vaccination and infection history, because antibodies from previous exposures can inhibit the effect of current vaccines. This has potentially important implications for selection of vaccine strains [36]. Other recent findings suggest that it might be important to vaccinate individuals who are not at serious risk themselves; in particular, immunizing schoolchildren might reduce flu transmission sufficiently to protect the general population, including those at high risk [13].

There are four drugs that have been approved to fight flu: amantadine and rimantadine, which inhibit the M2 protein, preventing the activation of the viral genetic material; and zanamivir and oseltamivir, which inhibit the action of NA, preventing new virus from leaving the infected cell. The M2 inhibitors are known to induce serious resistance and have important side effects on the gastrointestinal and central nervous systems. Less is known about the NA inhibitors, which were first approved in 1999 for use in the USA. Both the M2 inhibitors and the NA inhibitors can be used prophylactically. Anti-influenza drugs are an active area of research, with a wide variety of potential targets being investigated [32,37].

Modelling flu

Mathematical modelling has contributed significantly to our understanding of the epidemiology of infectious diseases [38] and is beginning to have a substantial impact on research at the immunological and molecular levels [39]. Epidemiological models allow us to investigate hypotheses about the mechanisms responsible for epidemics and to reject hypotheses that yield predictions inconsistent with documented epidemic patterns. Early analyses provided fundamental insights about thresholds for disease spread [40] and the shape of epidemic curves [41], whereas more recent work has elucidated effects of heterogeneity in age, genetics, social and spatial structure, and strategies for eradication and control [38].

For viral and bacterial respiratory infections, the dominant epidemiological modelling framework has been the box model [38], in which the host population is divided into compartments according to infection status (Box 1). For a variety of childhood diseases, the standard susceptible–infectious–recovered (SIR) model [38] and simple variations [42] are adequate to explain many features of real epidemic time series. For flu, it seems impossible to avoid a much greater degree of model complexity.

The primary obstacle to simple compartmental modelling of flu is antigenic drift. Strains are usually divided into what we will call variants, which are assumed to be immunologically distinct (modellers often call these variants 'strains', which can lead to confusion). Susceptibility to a given variant will depend on a person's history of infection with all of the variants in a given subtype. To capture complete immunological histories for a model with n variants of one subtype, it would be necessary to divide the population into more than 2^n compartments [43] (Box 2). Even for modest numbers of variants, the problem becomes analytically and computationally

Box 2. Multi-variant SIR models

To extend the susceptible–infectious–recovered (SIR) model (Box 1) to multiple variants, it is necessary to keep track of large amounts of information. Andreasen and co-workers [a,b] used set notation to keep track of infection histories. To get a flavour for this formalism, we consider the equation for the dynamics of just one class, the individuals S_J who have been infected with (or vaccinated against) a set of variants J, but who are susceptible to all other variants.

One must account for all of the ways in which individuals can enter or leave the class S_J. In particular, for each variant k in the set J, one must keep track of all classes $I^k_{J\setminus k}$ (comprising individuals who have been infected with all variants in J except k, and currently are infected with k). It is these individuals whose recovery (at rate γ, which might depend on the variant) will provide the new members of S_J. Similarly, individuals infected with variant $l \notin J$ can infect susceptible individuals and remove them from class S_J, at a rate modified by the factor σ^l_J (reflecting immunological cross-reaction between l and the variants in J).

If the overall force of infection [c] of variant l is written Λ^l, then the dynamics of the class S_J follow (Eqn I)

$$\frac{dS_J}{dt} = \sum_{k\in J}\gamma I^k_{J\setminus k} - \mu S_J - \sum_{l\notin J}\sigma^l_J \Lambda^l S_J \qquad \text{[Eqn I]}$$

where μ is the natural mortality rate, as in Box 1. Various assumptions can be made about how cross-reactivity works (through Λ^l, through σ^l_J, or both [a,d]).

Similar equations must also be constructed for the infectivity classes. In spite of the complexity of this system, methods have been developed that yield the various equilibria of the resulting high-dimensional system recursively and reveal the potential for periodic and complicated dynamics [a,b]. Gupta and co-workers [d] have elucidated the importance of such phenomena, especially concerning the organization of the strains into 'quasi-species' – collections of variants that interact strongly with each other dynamically and (especially) evolutionarily and much more weakly with other quasi-species (the notion of quasi-species is discussed in [e,f]).

A key problem with the multivariant formalism is that it is not clear how to measure cross-reactivity, that is, how to determine the values of the parameters σ^l_J. Immunological assays (e.g. hemagglutination inhibition experiments) could provide some insight into the strength of cross-reactivity in relation to genetic distances [g]. However, cross-immunity is in effect an epidemiological relation and must be inferred ultimately from population studies, which are sparser and less transparent sources of data.

References

a Andreasen, V. *et al.* (1997) The dynamics of cocirculating influenza strains conferring partial cross-immunity. *J. Math. Biol.* 35, 825–842

b Lin, J. *et al.* (1999) Dynamics of influenza A drift: the linear three-strain model. *Math. Biosci.* 162, 33–51

c Anderson, R.M. and May, R.M. (1991) *Infectious Diseases of Humans: Dynamics and Control*, Oxford University Press

d Gupta, S. *et al.* (1998) Chaos, persistence, and evolution of strain structure in antigenically diverse infectious agents. *Science* 280, 912–915

e Jenkins, G.M. *et al.* (2001) Nonquasispecies evidence for the evolution of RNA viruses. *Mol. Biol. Evol.* 18, 987–994

f Domingo, E. *et al.* (2002) *Quasispecies and RNA Virus Evolution: Principles and Consequences*, R.G. Landes Co.

g Smith, D.J. *et al.* (1999) Variable efficacy of repeated annual influenza vaccination. *Proc. Natl. Acad. Sci. U. S. A.* 96, 14001–14006

Acknowledgements
We thank Viggo Andreasen, Sigal Balshine, Freddy Christiansen, Juan Lin, Mark Loeb, John Oxford, Peter Palese, Joshua Plotkin and Martin Weigert for valuable discussions and comments, and Greg Armstrong for sending annual mortality data. We were supported by NIH grant 1-R01-GM60729 and the Natural Sciences and Engineering Research Council of Canada.

intractable. Furthermore, the implicit assumption of deterministic modelling – that there are a large number of individuals in each compartment – becomes ridiculous in cases where the number of compartments exceeds the total number of individuals.

A variety of approaches have been used to circumvent this difficulty. Andreasen *et al.* [44], following previous work by Pease [45], have used models where variants are constrained to evolve along a straight-line path (with further simplifying assumptions) to estimate drift rates. Gupta and others [46,47] have used a simple allele structure, and the assumption that cross-immunity affects only transmission (not infection), to reduce dimensionality and analyse coexistence and dynamics in multivariant models. Andreasen *et al.* [43] and Lin *et al.* [48] have also made use of symmetries in the equations to analyse unsimplified systems with a handful of variants.

It will be not be easy to produce more realistic models of flu drift that take into account the appearance of new variants, the complex geometry of the interactions between envelope proteins and the immune system [49], and the mechanisms of cross-immunity [36]. One possible avenue is through individual-based models [50], which have the capability to incorporate many relevant features, but at the cost of analytic intractability. Another approach would be the application of the technique of moment-closure, which has recently been applied to several different ecological questions [51–53] to simplify models; this method can be used to develop approximate dynamical equations for the mean, variance and higher moments of observable quantities.

Few attempts have been made to model the spatial spread of flu. Rvachev and Longini [54] modelled spatial spread on a global scale, whereas Bonabeau *et al.* [55] have recently modelled spatial epidemic patterns in France. Much more can be done, but it might be necessary to wait for better spatial data to accumulate before proceeding.

Box 3. Outstanding questions

In spite of considerable progress in understanding the cellular and molecular biology of flu viruses and recent advances in phylogenetics, some of the most basic ecological and evolutionary questions about the flu remain:

- What makes the disease so strongly seasonal?
- How important is inter-hemispheric transmission versus over-summer carry-over?
- How do strain structure and patterns of cross-reactivity self-organize over evolutionary time (e.g. into quasi-species), and how does this organization feed back to affect dynamics?
- Why do drifting subtypes persist and yet not diverge?
- What factors most strongly influence when new pandemics will occur, and what causes some (e.g. that of 1918) to be so lethal?
- What are the mechanisms of interaction between subtypes: why have there been two apparent exclusions (H1N1 → H2N2 and H2N2 → H3N2) and one case of coexistence (H3N2 and H1N1)?
- What are the implications of these issues for vaccination and antiviral drug strategies, and what are the reciprocal effects of intervention programmes upon the organization of the viral population?

Conclusions

Many questions about the flu remain (Box 3), in part because of a lack of data and the overwhelming complexity of models. Data collection and accessibility are improving, but more precise data are needed to allow us to uncover some of the most pressing mysteries of the flu. As a start, it would help enormously if confirmatory testing of cases of influenza-like illness

and pneumonia were to become standard practice and if laboratory-confirmed cases were always subtyped. Fast (36-h turnaround) flu testing using the polymerase chain reaction (PCR) is now possible [56], so the potential exists for tremendous improvements in the quality of surveillance data. Ideally, we would want information about the population dynamics of variants within each subtype and greater understanding of the cross-reactivity among strains. Such data would allow us to test models of drift evolution and relate them to flu phylogenies (which tell us little about the abundances of the various strains each year). Unfortunately, molecular sequencing is currently extremely costly and time consuming, so it is likely to be a long time before such analyses can be done routinely. In the meantime, new ideas are desperately needed to overcome severe modelling challenges and data limitations.

References

1 Webster, R.G. *et al.* (1992) Evolution and ecology of influenza A viruses. *Microbiol. Rev.* 56, 152–179
2 Goldmann, D.A. (2001) Epidemiology and prevention of pediatric viral respiratory infections in health-care institutions. *Emerg. Infect. Dis.* 7, 249–253
3 Schulman, J.L. and Kilbourne, E.D. (1962) Airborne transmission of influenza virus infection in mice. *Nature* 195, 1129–1130
4 Hope-Simpson, R.E. (1992) *The Transmission of Epidemic Influenza*, Plenum Press
5 Fine, P. (1982) Applications of mathematical models to the epidemiology of influenza: a critique. In *Influenza Models: Prospects for Development and Use* (Selby, P., ed.), pp. 15–85, MTP Press
6 Oxford, J.S. (2000) Influenza A pandemics of the 20th century with special reference to 1918: virology, pathology and epidemiology. *Rev. Med. Vir.* 10, 119–133
7 Cox, N.J. and Fukuda, K. (1998) Influenza. *Infect. Dis. Clin. North Am.* 12, 27–38
8 Cox, N.J. and Subbarao, K. (1999) Influenza. *Lancet* 354, 1277–1282
9 World Health Organization (2000) *World Health Report 2000*, World Health Organization
10 Cox, N.J. and Subbarao, K. (2000) Global epidemiology of influenza: past and present. *Annu. Rev. Med.* 51, 407–421
11 Simonsen, L. *et al.* (1997) The impact of influenza epidemics on mortality: introducing a severity index. *Am. J. Public Health* 87, 1944–1950
12 Simonsen, L. (1999) The global impact of influenza on morbidity and mortality. *Vaccine* 17, S3–S10
13 Reichert, T.A. *et al.* (2001) The Japanese experience with vaccinating schoolchildren against influenza. *N. Engl. J. Med.* 344, 889–896
14 Simonsen, L. *et al.* (1998) Pandemic versus epidemic influenza mortality: a pattern of changing age distribution. *J. Infect. Dis.* 178, 53–60
15 Armstrong, G.L. *et al.* (1999) Trends in infectious disease mortality in the United States during the 20th century. *J. Am. Med. Assoc.* 281, 61–66
16 Cliff, A. *et al.* (1993) *Measles: An Historical Geography of a Major Human Viral Disease*, Blackwell
17 Ina, Y. and Gojobori, T. (1994) Statistical-analysis of nucleotide-sequences of the hemagglutinin gene of human influenza-A viruses. *Proc. Natl. Acad. Sci. U. S. A.* 91, 8388–8392
18 Bush, R.M. *et al.* (1999) Positive selection on the H3 hemagglutinin gene of human influenza virus A. *Mol. Biol. Evol.* 16, 1457–1465
19 Fitch, W.M. *et al.* (2000) Predicting the evolution of human influenza A. *J. Hered.* 91, 183–185
20 Fitch, W.M. *et al.* (1997) Long term trends in the evolution of H(3) HA1 human influenza type A. *Proc. Natl. Acad. Sci. U. S. A.* 94, 7712–7718
21 Macken, C. *et al.* (2001) The value of a database in surveillance and vaccine selection. In *Options for the Control of Influenza IV*, (Osterhaus, A.D.M.E. *et al.*, eds), Elsevier (http://www.flu.lanl.gov/)
22 Webster, R.G. (1998) Influenza: an emerging disease. *Emerg. Infect. Dis.* 4, 436–441
23 Brown, I.H. (2000) The epidemiology and evolution of influenza viruses in pigs. *Vet. Microbiol.* 74, 29–46
24 Claas, E.C.J. (2000) Pandemic influenza is a zoonosis, as it requires introduction of avian-like gene segments in the human population. *Vet. Microbiol.* 74, 133–139
25 Subbarao, K. and Shaw, M.W. (2000) Molecular aspects of avian influenza (H5N1) viruses isolated from humans. *Rev. Med. Virol.* 10, 337–348
26 Taubenberger, J.K. *et al.* (1997) Initial genetic characterization of the 1918 'Spanish' influenza virus. *Science* 275, 1793–1796
27 Reid, A.H. *et al.* (1999) Origin and evolution of the 1918 'Spanish' influenza virus hemagglutinin gene. *Proc. Natl. Acad. Sci. U. S. A.* 96, 1651–1656
28 Taubenberger, J.K. *et al.* (2000) The 1918 influenza virus: a killer comes into view. *Virology* 274, 241–245
29 Basler, C.F. *et al.* (2001) Sequence of the 1918 pandemic influenza virus nonstructural gene (NS) segment and characterization of recombinant viruses bearing the 1918 NS genes. *Proc. Natl. Acad. Sci. U. S. A.* 98, 2746–2751
30 Dowdle, W.R. (1981) Influenza immunoprophylaxis after 30 years' experience. In *Genetic Variation Among Influenza Viruses* (Nayak, D.P., ed), pp. 525–534, Academic Press
31 Subbarao, K. (1999) Influenza vaccines: present and future. *Adv. Virus Res.* 54, 349–373
32 Advisory Committee on Immunization Practices (2001) Prevention and control of influenza. *Morb. Mortal. Wkly Rep.* 50, 1–46
33 US Department of Health and Human Services Public Health Service (1991) *Healthy People 2000: National Health Promotion and Disease Prevention Objectives*, Technical Report 91-50212, US Department of Health and Human Services
34 World Health Organization (2001) Recommended composition of influenza virus vaccines for use in the 2001–2002 season. *Wkly. Epidemiol. Rec.* 76, 58–61
35 de Jong, J.C. *et al.* (2000) Mismatch between the 1997/1998 influenza vaccine and the major epidemic A(H3N2) virus strain as the cause of an inadequate vaccine-induced antibody response to this strain in the elderly. *J. Med. Virol.* 61, 94–99
36 Smith, D.J. *et al.* (1999) Variable efficacy of repeated annual influenza vaccination. *Proc. Natl. Acad. Sci. U. S. A.* 96, 14001–14006
37 Luscher-Mattli, M. (2000) Influenza chemotherapy: a review of the present state of art and of new drugs in development. *Arch. Virol.* 145, 2233–2248
38 Anderson, R.M. and May, R.M. (1991) *Infectious Diseases of Humans: Dynamics and Control*, Oxford University Press
39 Nowak, M.A. and May, R. (2000) *Virus Dynamics: The Mathematical Foundations of Immunology and Virology*, Oxford University Press
40 Ross, R. (1910) *The Prevention of Malaria*, E.P. Dutton and Co.
41 Kermack, W.O. and McKendrick, A.G. (1927) A contribution to the mathematical theory of epidemics. *Proc. R. Soc. Lond. Ser. A.* 115, 700–721
42 Earn, D.J.D. *et al.* (2000) A simple model for complex dynamical transitions in epidemics. *Science* 287, 667–670
43 Andreasen, V. *et al.* (1997) The dynamics of cocirculating influenza strains conferring partial cross-immunity. *J. Math. Biol.* 35, 825–842
44 Andreasen, V. *et al.* (1996) A model of influenza A drift evolution. *Z. Angew. Math. Mech.* 76, 421–424
45 Pease, C.M. (1987) An evolutionary epidemiologic mechanism, with applications to Type-A influenza. *Theor. Popul. Biol.* 31, 422–452
46 Gupta, S. *et al.* (1996) The maintenance of strain structure in populations of recombining infectious agents. *Nat. Med.* 2, 437–442
47 Gupta, S. *et al.* (1998) Chaos, persistence, and evolution of strain structure in antigenically diverse infectious agents. *Science* 280, 912–915
48 Lin, J. *et al.* (1999) Dynamics of influenza A drift: the linear three-strain model. *Math. Biosci.* 162, 33–51
49 Lapedes, A. and Farber, R. (2001) The geometry of shape space: application to influenza. *J. Theor. Biol.* 212, 57–69
50 Epstein, J.M. and Axtell, R. (1996) *Growing Artificial Societies: Social Science from the Bottom Up*, The Brookings Institution Press
51 Turelli, M. and Barton, N.H. (1994) Genetic and statistical analyses of strong selection on polygenic traits – what, me normal? *Genetics* 138, 913–941
52 Bolker, B.M. and Pacala, S.W. (1999) Spatial moment equations for plant competition: understanding spatial strategies and the advantages of short dispersal. *Am. Nat.* 153, 575–602
53 Dwyer, G. *et al.* (2000) Pathogen-driven outbreaks in forest defoliators revisited: building models from experimental data. *Am. Nat.* 156, 105–120
54 Rvachev, L.A. and Longini, I.M. (1985) A mathematical model for the global spread of influenza. *Math. Biosci.* 75, 3–23
55 Bonabeau, E. *et al.* (1998) The geographical spread of influenza. *Proc. R. Soc. Lond. Ser. B* 265, 2421–2425
56 Carman, W.F. *et al.* (2000) Rapid virological surveillance of community influenza infection in general practice. *Br. Med. J.* 321, 736–737
57 Cox, N.J. and Bender, C.A. (1995) The molecular epidemiology of influenza viruses. *Semin. Virol.* 6, 359–370
58 Chen, W.S. *et al.* (2001) A novel influenza A virus mitochondrial protein that induces cell death. *Nat. Med.* 7, 1306–1312
59 Webster, R.G. (2001) Virology — a molecular whodunit. *Science* 293, 1773–1774
60 United States Bureau of the Census (1910–1992) *Vital Statistics of the United States (before 1937: Mortality Statistics)*, Government Printing Office