Bluff your way in epidemic models

Mick Roberts and Hans Heesterbeek

The literature on the mathematical modelling of infectious diseases has grown enormously in recent years, both in quantity and quality. Here, we briefly point to the purposes of these modelling exercises and introduce the main ideas behind compartmental models, to act as a guide to the (bio)mathematical literature that is less directly accessible.

M.G. Roberts is in AgResearch, Wallaceville Animal Research Centre, PO Box 40063, Upper Hutt, New Zealand; J.A.P. Heesterbeek is in the Dept of Zoology, University of Oxford, South Parks Road, Oxford, UK OX1 3PS.

Epidemiology is the study of the spread, in time and space, of disease within animal and plant populations. The aim is to trace factors that are responsible for, or contribute to, the occurrence of these diseases. Here, we are only concerned with infectious diseases caused by microparasites (that is, viruses, bacteria and protozoa)[1] and the mathematical models that can describe their spread with time.

There are many reasons for using mathematical models in epidemiology[2]. The main justifications include the following[2,3]. First, models provide insight into, and understanding of, the relationships between the mechanisms operating at the level of the individual and the phenomena that result at the population level. Second, formulating mathematical models requires precision about the underlying assumptions and can reveal potentially useful working hypotheses that might otherwise go unnoticed. The analysis of mathematical models can lead to the discovery of concepts that turn out to play an important role in the epidemiology of infectious disease. The most important example of this is the basic reproduction ratio, R_0, which is, roughly speaking, the expected number of new cases from one infective individual[1,3–7]. Third, an important use of models is in clarifying which parameters have a critical influence on the predicted dynamical behaviour of the population. This may lead to the discovery of key parameters, the numerical value of which may be unknown or, alternatively, the realization that some parameters that researchers are struggling to measure are irrelevant to the dynamics of the infection. Finally, models are valuable for performing thought experiments, for example to evaluate the efficacy of control measures in cases where actual experiments are impossible because of ethical or economic constraints.

It is tempting to think that models may be useful for predicting future trends, but this is often not the case. The most complex models for specific diseases are still (highly) oversimplified, our knowledge of key parameters in the transmission process is often poor and making models more complex rapidly leads to a proliferation of parameters, hardly any of which can be 'guesstimated' with accuracy. Making accurate quantitative predictions (certainly in the long term) from complex models is therefore practically impossible. However, an approach involving simplification and mathematical analysis of a model can reveal the important underlying relationships.

In this paper, we introduce compartmental models[8] for studying the spread of infectious diseases. We give an example of a basic model that includes many of the essential ingredients and describe its dynamic behaviour. We also describe the use of this model in evaluating control strategies, and finally we explore areas in which the model can be extended to be more realistic. We end with a summary of some recent issues in the modelling of infectious diseases.

A simple model

In a compartmental model for disease transmission, the individuals in a population are divided into a number of compartments, commonly denoted by S (susceptible), I (infective) and R (removed: immune and no longer infective)[9]. The model usually consists of a set of differential equations that describe the change in the number of individuals in the compartments with time. A large number of compartmental models have been described in the literature[7–9], but we will describe one of the simplest here. For obvious reasons this model is referred to as an *SIR* model.

If we consider a constant, closed population of N hosts, S will be susceptible to a disease and I are infected and infectious. By 'closed', we mean that the addition of new susceptible hosts to the population is negligible on the time scale of disease transmission. We will also assume that contacts between susceptible and infective individuals occur randomly, and that the I infectives make βI random contacts that could result in disease transmission per unit of time (where β is a disease-specific transmission parameter that combines information about the rate of making contacts and the transmission probability). Because the contacts are random, only a fraction (S/N) of contacts are with a susceptible individual. Therefore the rate at which individuals transfer from the susceptible to the infective compartment is $\beta IS/N$ (Ref. 10). If the rate at which infectives cease to be infective is γ, and if they then take no further part in disease transmission, we obtain the set of equations in Box 1. The third class, R (removed), consists of individuals that are assumed to play no further part in the epidemic. An example is measles: after individuals have been infective for a limited period, they recover and become immune for life[1,14].

An important quantity that arises from any epidemic model is the basic reproduction ratio (R_0)

Box 1. A simple three-compartment model

The number of susceptibles in the population is reduced as new infections take place (see the text for the definitions of terms):

$$\frac{dS}{dt} = -\beta I \frac{S}{N}$$

The number of infectives increases as new infections take place, but is reduced as hosts move into the removed compartment:

$$\frac{dI}{dt} = \beta I \frac{S}{N} - \gamma I$$

The number of removed hosts increases as the pool of susceptible hosts is used up:

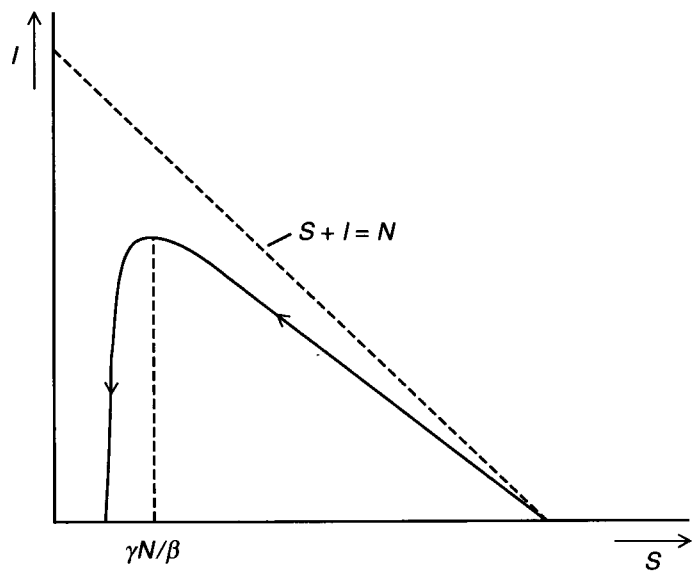
$$\frac{dR}{dt} = \gamma I$$

Note that two of these equations are sufficient to define the system, as $S + I + R = N$, which is constant for this model.

The basic reproduction ratio (R_0) is the expected number of secondary cases that would occur as the result of a typical single primary case, if the population were totally susceptible. For the simple model, $R_0 = \beta/\gamma$, as each infective individual has the potential to infect β susceptibles per unit of time and remains infective for an average of $1/\gamma$ time units. Given that an epidemic develops (i.e. when $R_0>1$), the relationship between the number of susceptibles and the number of infectives in the population typically evolves as shown in the figure below.

Recent publications have discussed the statistical estimation of R_0 (Ref. 11), the use of R_0 in evaluating control programmes[1], the definition and calculation of R_0 for models of diseases in heterogeneous populations[6] and its definition for stochastic models[12]. More general deterministic models have also been reviewed[13].

To model a vaccination programme, suppose that a given strategy leads to a proportion f of the total population being immunized at any given time, and therefore a proportion $1-f$ of the original susceptible population remains susceptible. If we define R_f as the expected number of new cases arising from a single primary case when a proportion $1-f$ of the population is susceptible, then the vaccination strategy is successful in eradicating the infection if it causes R_f to be reduced below 1. For the simple model above, $R_f = (1-f)\beta/\gamma = (1-f)R_0$ and $R_f<1$ when $f>1-1/R_0$.



The figure shows the relationship between the number of infectives (I) and the number of susceptibles (S) during the course of an epidemic, in accordance with the model above. The number of infectives rises steadily to a maximum, then decreases to zero and the infection disappears from the population. The decrease can be understood by observing that S becomes too small after a certain point for an infective individual to contact sufficient susceptibles during its infectious period to cause a new case.

(Ref. 1). This is defined as the expected number of secondary cases that would occur as the result of a single, typical, primary case if the population were totally susceptible[1,6]. If $R_0<1$ then the infection cannot establish itself in the population as each case gives rise to less than one subsequent case. If, however, $R_0>1$ then an epidemic will occur. This does not imply that all individuals in the population will become infective with time. If, subsequently, infective individuals permanently enter a removed compartment, then the proportion that is susceptible decreases, and eventually S/N becomes too small for infectives to make contact with a sufficient number of susceptibles to keep the epidemic going. The fraction of the population that is still susceptible at the end of the epidemic is a decreasing function of R_0 (Ref. 4).

Contact structure and demographics

The model described above and in Box 1 is oversimplified, in particular because of the assumption that the host population size is constant, and that the population is closed. In reality, populations change in size (N) because of births and deaths. The number of contacts (C) made by an individual will tend to increase as the population size increases (Fig. 1), and this effect must be incorporated into the model[16].

In modifying the model shown in Box 1, we still assume that contacts between susceptible and infective individuals occur randomly, but we allow the contact rate to be a function of the population size [contact rate=$C(N)$]. The definition of 'contact' depends on the mode of transmission of the infection that is studied. The transmission parameter β is now

explicitly separated into a product, $pC(N)$, where p is the transmission probability per contact. For each contact, the probability that the 'contactee' is susceptible is S/N, as before, and hence the rate at which individuals transfer from the susceptible to the infective compartment is $pC(N)IS/N$.

Some rudimentary demographics of the host population are now added to the basic model, and a general contact-rate function $C(N)$ is included. We will disregard recovery for the moment, so that the only way to leave the infective compartment I is by dying, either from the infection or from other natural causes. For this model, we consider a population consisting only of susceptible and infective individuals, which has a birth rate that exceeds its death rate. Hence, in the absence of infection, the population increases exponentially but, in the presence of infection, there are four different possibilities for the long-term behaviour of the system[7,17] (Fig. 2).

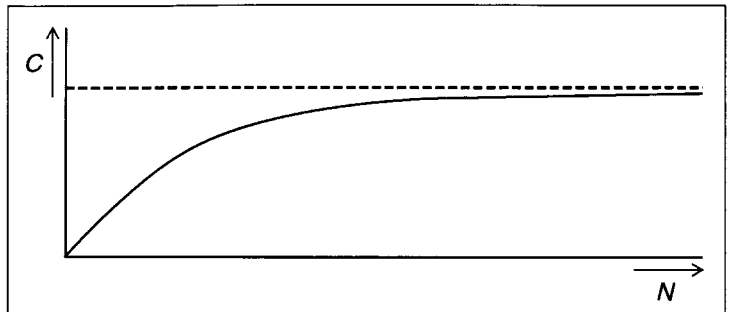


Fig. 1. The general shape of the contact rate function $C(N)$. In reality, the contact rate is a function of the local population density[10,15], but this can be approximately translated into the relationship shown if the population enlarges the area it occupies when the density becomes too high, but not if the density is less than a 'comfortable' level.

Density-dependent demographics

Host populations do not grow in an unconstrained manner. In reality, the birth and/or death rate will depend on the host population density, with birth rate decreasing and death rate increasing with increasing population density. There is a direct relationship between population density and size[7], and we use $B(N)$ and $D(N)$ for the per capita birth and death rates respectively (Box 2). In the absence of infection, the population size tends over time to the carrying capacity of the environment (K), that is, the population size at which birth and death rates are equal. Often logistic population growth is assumed, but this is essentially phenomenological and has no mechanistic biological foundation[18]. Nonlinear effects in population growth are in response to changes in the local population density, and an examination of the community structure is necessary to ensure that any model is correctly defined.

The long-term dynamics are summarized in Box 2. The major difference from the model described in Fig. 2 is that population growth is now controlled in the absence of infection. If $R_0<1$, the infection cannot persist. If $R_0>1$, the infection persists in the population and becomes endemic. Results of this type are often referred to as 'threshold theorems', because $R_0=1$ defines a threshold above which an infection can become established in a population.

As R_0 depends on K, an alternative way of expressing the threshold theorem is to define the critical community size as the carrying capacity for which $R_0=1$. If the population size is below this critical level, the parasite cannot sustain itself. For example measles rarely persists in communities of less than 250 000 people[19].

Models for disease control and eradication

Although the primary reason for constructing a model is to understand the epidemiology of a disease, the possibility of intervening to control or eradicate the disease is often a consideration. Once the correct model is in place, it is often easy to modify it to model the effects of control procedures.

For example, the idea of a critical population size below which an infection cannot persist has led to disease eradication policies based on regular culling (killing) of wild animals. Such policies are used where the animal acts as a reservoir for a disease of humans (for example rabies in foxes[20]) or of domestic animals (for example tuberculosis in badgers[21] or possums[22]). It is easy to adapt the model in Box 2 to incorporate culling of reservoir hosts: if animals are killed at a rate of δ per year, then the overall death rate is increased to $D(N) + \delta$. A simple calculation can

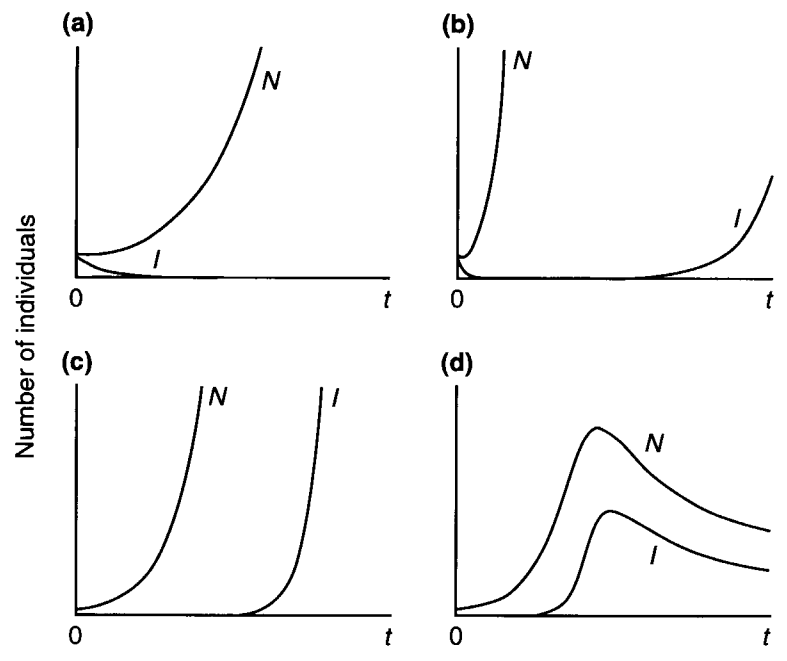


Fig. 2. Solutions of a model similar to that of Box 2, but with birth and death rates constant and birth rate exceeding death rate. The host population size (N) and the infective population size (I) are shown. The four possible types of behaviour are, for different parameter values, **(a)** the disease becomes extinct and the host population continues to grow, **(b)** both the total and infective populations grow, but the proportion of individuals that are infective tends to zero (that is, N grows faster than I), **(c)** both the total and infective populations grow, but the proportion of individuals that are infective tends to a positive limit, and **(d)** both the total and infective populations tend to a (finite) positive limit.

Box 2. A model with density-dependent demographics

The total population size is increased by new births, and reduced by deaths from natural causes and by deaths from infection (see the text for the definitions of terms):

$$\frac{dN}{dt} = [B(N) - D(N)]N - \alpha I$$

$B(N)$, the birth rate, is a decreasing, or at least a nonincreasing, function of N, the population size. $D(N)$, the death rate, is a nondecreasing function of N. The value α is the increase in mortality caused by infection.

The number of infectives increases because of new infectious contacts and decreases from deaths in the I compartment and from deaths due to disease. The number of susceptibles is $S = N - I$:

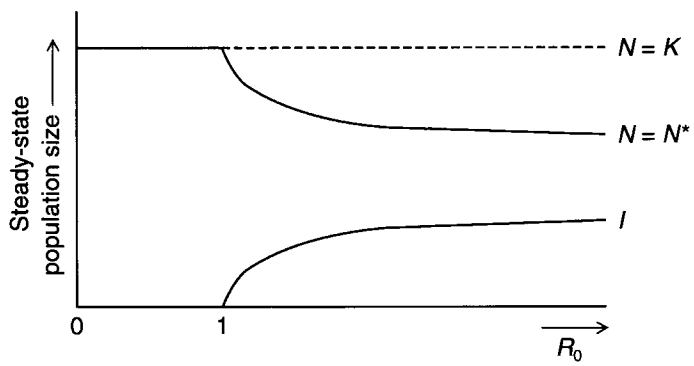
$$\frac{dI}{dt} = p\frac{C(N)}{N}SI - [\alpha + D(N)]I$$

The model has three steady-state solutions (solutions where $dN/dt = dI/dt = 0$): the trivial solution ($N = I = 0$), which is unstable, the disease-free steady state [$N = K$, $I = 0$, where $B(K) = D(K)$] and the endemic steady state ($N = N^*, I = I^*$), which only exists if the basic reproduction ratio

$$R_0 = \frac{pC(K)}{\alpha + D(K)}$$

is greater than 1. If $R_0 < 1$ the infection cannot persist, and all solutions tend to the disease-free steady state ($N = K$, $I = 0$). If $R_0 > 1$ the infection persists in the population, and solutions tend to the endemic steady state (N^*, I^*), where $N^* < K$. These steady states are illustrated below.

R_0 does not depend on the population size N in the above model, nor in other models that incorporate true mass-action kinetics for the infection process[10]. This conclusion was supported by a recent experiment, where no significant difference was found between the values of R_0 for pseudorabies (Aujeszky's disease) in groups of 10 and 40 pigs kept at the same population density[10].



The figure shows the possible steady-state values of the population size (N) for different values of the basic reproduction ratio (R_0). The solution $N = 0$ is always unstable. The disease-free solution $N = K$ is stable for $R_0 < 1$ and unstable for $R_0 > 1$ (shown as a dashed line). The endemic steady state $N = N^*$ is stable when it exists, that is when $R_0 > 1$. Also shown is the size of the infective population, I^*.

then determine the level of δ required to eradicate the infectious agent.

If the reason for wishing to control a disease is to conserve the host population, then a culling programme is self-defeating. Another option may be to vaccinate animals, either by the mass distribution of baits containing oral vaccine (for example against rabies in foxes[20] or raccoons[23]) or by capturing, vaccinating and releasing animals (for example against phocine distemper virus in seals[24]). For human populations, culling is not an option and vaccination programmes have been used against many diseases[1].

If, for a given vaccination strategy, a fraction f of the susceptible population is immune due to vaccination at any given time, one can use R_0 to calculate how large f must be to eradicate the disease (Box 1). For example R_0 has been estimated to be 10–20 for measles in western Europe[1] and hence 90–95% of the population must be immunized at any one time to eradicate the disease. This is at the limit of the vaccination level that can be achieved in western Europe, and therefore knowledge of the R_0 of an infection leads to an immediate assessment of the probability of success of a vaccination programme.

Other possible control policies include mass chemotherapy of populations and forms of biological control. The latter involves the deliberate encouragement of a pathogen to reduce the population size of a pest (for example myxomatosis in rabbits[25]). These strategies can be modelled in much the same way as we have discussed above.

Further complications

Considering the transmission dynamics of particular diseases may reveal further complications, even within the framework of an *SI* model. For example, a disease may reduce the fecundity of the host population, leading to a modification of the birth rate in the equation for the overall population size[17]. For other diseases, mechanisms for vertical disease transmission (parent to offspring) could be included in the equations, which removes the restriction that all new-born hosts are susceptible[26].

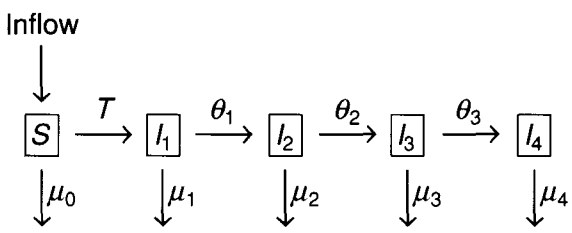
Complications posed by a long latent period, acquired immunity and heterogeneity among individuals are of greater biological significance. A latent period is modelled by introducing an equation for an additional compartment (E) containing individuals that have been infected, but are not yet infective. The model is then referred to as an *SEI* model. A period of latency reduces the value of R_0, but only if an infected individual can die during this period, before it reaches the infective compartment. This modification may have more effect on the predicted dynamics of the disease than simply to reduce R_0. For example in a model of the dynamics of rabies in European foxes, the host population size tends either to a steady state or to a stable oscillation with time, depending on the values of the parameters. This shows that the delay in transmission caused by a long latent period can destabilize the steady state of a system and cause the population size and the number of infective hosts to exhibit

Box 3. Compartmental models with complications: HIV

The epidemiology of HIV infections illustrates some of the many complications that can arise when trying to make realistic compartmental models for the dynamics of a disease. Broadly speaking, there are three categories of complication. First, the infectivity of an individual is not constant, but is a function of the time that has elapsed since infection. To describe this in a compartmental model[1,2,30] a number of infective compartments, for example four, can be introduced, I_1–I_4, representing the initial stage of infection (high infectivity but short duration), the intermediate stage (low infectivity and long duration), the pre-AIDS stage and the full-blown AIDS stage respectively (see the figure).

The second type of complication involves variation in susceptibility and the differences between individuals that influence sexual behaviour. Some of the many important characteristics are gender, age, sexual activity (that is, the propensity to make new sexual contacts), sexual orientation and personal habits (for example injecting drugs)[1,2]. Further complications arise if viral variation is taken into account[31]. To model these heterogeneities compartmentally each compartment is split into subcompartments and then different parameters are assigned for the various transmission rates between compartments.

The third category of complication involves the contact pattern in the community. Everybody does not have contact with everybody else, but differences in mixing between risk groups are important factors that influence the spread of HIV (Refs 2,32–34). These differences complicate the transmission term T by affecting the transmission rates between subcompartments (see the figure below). Individuals also tend to interact with the same partner for a significant time, rather than randomly (within a given mixing framework) and instantaneously. The incorporation of this effect leads to pair formation models for sexually transmitted diseases[35,36].



The figure shows the sequence of compartments in a model for HIV infection. Susceptibles (S) become infected by contact with individuals from each of four compartments: I_1, those newly infective; I_2, those in the intermediate (low-infectivity) stage; I_3, those in the infective pre-AIDS stage; and I_4, those with full-blown AIDS. Progression from compartment I_j to compartment I_{j+1} is at rate θ_j. The transmission term T is defined by:

$$T = \sum_{i=1}^{i=4} \frac{\beta_i I_i S}{N}$$

where the β_i terms are the transmission rates. The natural death rate of susceptibles is μ_0, and the increase in death rate in compartment j is μ_j.

periodic cycles[27]. Similarly, diseases such as measles are well known for their cyclical nature. Apparently irregular behaviour can result from random fluctuations in epidemiological parameters[28] or may be an intrinsic property of the model[29].

Heterogeneity in the susceptibility, infectiousness or social behaviour of individuals may be accounted for by increasing the number of compartments in the model. For many diseases, age has important effects on these factors, and it is common to model sexually transmitted diseases with separate S, E, I and R compartments for males and females. Models for the spread of HIV include many of the complications that have been discussed (Box 3).

With models becoming increasingly complicated, we reiterate Black and Singer[37]: 'A model is essential to any rational plan for controlling infectious disease, but more refined models require more and better biological data. ... Communication between [microbiology and mathematics] needs to be improved.' We hope that this review facilitates one side of the exchange.

Acknowledgements

This work was conceived and completed during the programme on epidemic models at the Isaac Newton Institute for Mathematical Sciences, University of Cambridge. We thank the organizers of the programme and the Institute for the opportunity to participate. M.G.R. also received financial assistance from a Trimble Agricultural Research Fellowship and from MSD–AgVet Ltd and an extended period of leave from the New Zealand Pastoral Agriculture Research Institute Ltd. J.A.P.H. was financially supported by the Netherlands Organization for Scientific Research, NWO, the Royal Society and SERC Grant GR G59981. Helpful comments were received from three anonymous referees. Gary Smith made comments on an earlier draft. He is absolutely the last person that the authors wish to thank.

References

1 Anderson, R.M. and May, R.M. (1991) *Infectious Diseases of Humans: Dynamics and Control*, Oxford University Press
2 Hethcote, H.W. and Van Ark, J.W. (1992) *Modelling HIV Transmission and AIDS in the United States, Lecture Notes in Biomathematics* 95, Springer–Verlag
3 Diekmann, O. (1991) in *Ordinary and Partial Differential Equations* (Vol. III) (Sleeman, B.D. and Jarvis, R.J., eds), pp. 67–79, Longman
4 Kermack, W.O. and McKendrick, A.G. (1927) *Proc. R. Soc. Edinb. Sect. A* 115, 700–721 [reprinted (1991) *Bull. Math. Biol.* 53, 33–55]
5 Hethcote, H.W. and Yorke, J.A. (1984) *Gonorrhea Transmission Dynamics and Control, Lecture Notes in Biomathematics* 56, Springer–Verlag
6 Diekmann, O., Heesterbeek, J.A.P. and Metz, J.A.J. (1990) *J. Math. Biol.* 28, 365–382
7 Heesterbeek, J.A.P. and Roberts, M.G. in *Ecology of Infectious Diseases in Natural Populations* (Grenfell, B.T. and Dobson, A.P., eds), Cambridge University Press (in press)
8 Jacquez, J.A. (1985) *Compartmental Analysis in Biology and Medicine* (2nd edn), University of Michigan Press
9 Hethcote, H.W. in *Frontiers of Mathematical Biology, Lecture Notes in Biomathematics* 100 (Levin, S.A., ed.), Springer–Verlag (in press)
10 De Jong, M.C.M., Diekmann, O. and Heesterbeek, J.A.P. in *Epidemic Models: their Structure and Relation to Data* (Mollison, D., ed.), Cambridge University Press (in press)

11 Dietz, K. (1993) *Stat. Methods Med. Res.* 2, 23–41
12 Nåsell, I. in *Epidemic Models: their Structure and Relation to Data* (Mollison, D., ed.), Cambridge University Press (in press)
13 Diekmann, O., Heesterbeek, J.A.P. and Metz, J.A.J. in *Epidemic Models: their Structure and Relation to Data* (Mollison, D., ed.), Cambridge University Press (in press)
14 Grenfell, B.T. and Bolker, B.M. in *Ecological Epidemiology of Parasitic and Infectious Diseases* (Scott, M.E. and Smith, G., eds), Academic Press (in press)
15 Heesterbeek, J.A.P. and Metz, J.A.J. (1993) *J. Math. Biol.* 31, 529–539
16 Thieme, H.R. (1992) *Math. Biosci.* 111, 99–130
17 Diekmann, O. and Kretzschmar, M. (1991) *J. Math. Biol.* 29, 539–570
18 Kingsland, S.E. (1985) *Modeling Nature: Episodes in the History of Population Ecology*, University of Chicago Press
19 Bartlett, M.S. (1960) *J. R. Stat. Soc.* 123, 37–44
20 Anderson, R.M. *et al.* (1981) *Nature* 289, 765–771
21 Anderson, R.M. and Trewhella, W. (1985) *Philos. Trans. R. Soc. London Ser. B Biol. Sci.* 310, 327–381
22 Roberts, M.G. (1992) *IMA J. Math. Appl. Med. Biol.* 9, 19–28
23 Coyne, M.J., Smith, G. and McAllister, F.E. (1989) *Am. J. Vet. Res.* 50, 2148–2154
24 Grenfell, B.T., Longergan, M.E. and Harwood, J. (1992) *Sci. Total Environ.* 115, 15–28
25 Anderson, R.M. and May, R.M. (1982) *Parasitology* 85, 411–426
26 Busenberg, S.N. and Cooke, K. (1993) *Vertically Transmitted Diseases: Models and Dynamics*, Springer-Verlag
27 Anderson, R.M. (1982) in *The Population Dynamics of Infectious Diseases: Theory and Applications* (Anderson, R.M., ed.), pp. 242–261, Chapman & Hall
28 Rand, D.A. and Wilson, H.B. (1991) *Proc. R. Soc. London Ser. B Biol. Sci.* 246, 179–184
29 Schaffer, W.M. (1985) *IMA J. Math. Appl. Med. Biol.* 2, 221–252
30 Blythe, S.P. and Anderson, R.M. (1988) *IMA J. Math. Appl. Med. Biol.* 5, 181–200
31 McLean, A.R. (1993) *Trends Microbiol.* 1, 9–13
32 Morris, M. in *Infectious Human Diseases* (Isham, V. and Medley, G., eds), Cambridge University Press (in press)
33 Gupta, S., Anderson, R.M. and May, R.M. (1989) *AIDS* 3, 807–817
34 Jacquez, J.A., Simon, C.P. and Koopman, J.S. in *Epidemic Models: their Structure and Relation to Data* (Mollison, D., ed.), Cambridge University Press (in press)
35 Dietz, K. and Hadeler, K.P. (1988) *J. Math. Biol.* 26, 1–25
36 Dietz, K. (1989) in *Sesquicentennial Invited Paper Sessions* (Gail, M.H. and Johnson, N.L., eds), pp. 609–621, American Statistical Association
37 Black, F.L. and Singer, B. (1987) *Annu. Rev. Microbiol.* 41, 677–701

The ebb and flow of a fungal genome

Katherine F. Dobinson and John E. Hamer

Transposable DNA elements have only recently been described in a few species of filamentous fungi, but may be more abundant than previously believed. Several different elements have been isolated from the rice blast pathogen *Magnaporthe grisea*. The distribution and amplification of these elements suggest a potential role in the evolution of the fungal genome.

K.F. Dobinson is at Agriculture Canada, London Research Centre, 1391 Sandford St, London, Ontario, Canada N5V 4T3; J.E. Hamer is at the Dept of Biological Sciences, Hansen Life Sciences Building, Purdue University, West Lafayette, IN 47907, USA.

Filamentous fungi have proved valuable in elucidating the genetic basis of biochemical reactions and in many other fundamental genetic studies. However, most of our knowledge of transposable elements has come from other organisms, because the laboratory strains of the most well-studied fungi, *Aspergillus nidulans* and *Neurospora crassa*, appear to be devoid of active transposons. In fact, a mechanism exists in *N. crassa* to destroy specific types of repeated DNA sequences[1]. Such a mechanism may not be common to all fungi, since an assortment of mobile elements have recently been discovered in a number of species. In these diverse and poorly understood fungi, transposable elements have aided our understanding of population structure and epidemiology. Because of their rapid life cycles and large populations, fungi may prove useful in determining the effects of mobile elements on speciation and genome evolution.

Transposons recently identified in filamentous fungi

All but one of the transposable elements that have so far been unearthed from the nuclear genomes of filamentous fungi (see Table 1) are either retrotransposons, which contain long terminal repeats (LTRs) and are structurally similar to retroviruses, or retroelements, which are similar to the long interspersed nuclear elements (LINEs) first described in mammals[2] (Fig. 1a and b). For at least some of these elements there is evidence for active transposition. LINE-like elements exist both in *N. crassa* (a single strain from Africa, known as Adiopodoume) and in the rice blast pathogen *Magnaporthe grisea*. Both elements are transcribed and the *Neurospora* element has been shown to transpose through a cytoplasmic intermediate[3,4].

LTR retroelements have been isolated from *M. grisea* and from the tomato pathogen *Cladosporium fulvum*[5,6]. The structural integrity of cloned elements suggests that at least some strains harbour