# Mutation and the evolution of virulence

SEBASTIAN BONHOEFFER AND MARTIN A. NOWAK

*Department of Zoology, University of Oxford, South Parks Road, Oxford OX1 3PS, U.K.*

SUMMARY

The evolution of parasite virulence can be influenced by intra-host mutation of the parasite and competition among different parasite mutants in the same host. During the course of an infection, parasite variants may be generated that have some competitive advantage over the strains that initiated the infection. This intra-host competition may select for a short-term advantage. It can lead to variants without any transmission potential or variants with extremely high virulence. We develop a mathematical model which describes intra- and inter-host parasite evolution. The major conclusions are: (i) intra-host competition generates a virulence polymorphism in the parasite population; (ii) intra-host competition can shift the mean virulence beyond what would maximize the reproductive rate of the parasite in the host population; and (iii) the parasite population can evolve to intermediate levels of virulence even if there is no trade-off between transmission rate and virulence. Our model applies to highly mutating parasites with long infection periods, such as, for example, the human immunodeficiency virus (HIV).

## 1. INTRODUCTION

The evolution of highly mutating microparasites, here broadly defined as viruses, bacteria, and protozoa, takes place on two different levels: within the host population, and within an individual host. On the level of the host population, selection may tend to increase a parasite's reproductive rate in terms of the number of secondary cases caused by a single infected host (Anderson & May 1979, 1991; May & Anderson 1979). This reproductive rate depends on the parasite's transmission rate and the duration of the infectious period of the host. Because microparasites usually cease propagating after the host's death, it is often believed that parasites generally should evolve to become harmless to their natural hosts, but there is strong observational and theoretical evidence against this so called 'conventional wisdom'. The epidemics of myxoma virus in the Australian rabbit population gives an example for the evolution towards intermediate levels of virulence (Fenner & Ratcliffe 1965). Smallpox and measles are just two examples of a long list of highly virulent pathogens with a long-standing evolutionary relation with their hosts. Herre (1993) gives an example for highly virulent nematodes of fig wasps that have been coevolving with their host for several million years. Theoretical models have postulated that evolution can lead to intermediate levels of virulence if transmissibility and virulence are coupled in some way (May & Anderson 1979, 1983, 1990; Anderson & May 1981, 1982; Levin and Pimentel 1981; Bremermann & Pickering 1983; Ewald 1983; Knolle 1989; Frank 1992; Antia *et al.* 1994).

On the level of the individual infection, selection will favour parasite mutants with a competitive advantage, regardless of their ultimate 'survival' in the host population. Levin & Bull (1994) argue that, for some highly mutating parasites, virulence may be a random byproduct of intra-host evolution. They give three examples where this 'short-sighted' intra-host evolution may account for virulence: bacterial meningitis, polio and AIDS. In none of these three cases do the diseases induced by these pathogens seem to confer a transmission advantage, and there is evidence that the diseases result from an intra-host evolutionary process.

Only a few models have been put forward which combine both the intra-host and the inter-host level of microparasite evolution. Levin & Pimentel (1981) discussed a model of superinfection of two parasite strains. They showed the coexistence of an avirulent and a virulent parasite strain in the host population, assuming that the virulent strain can take over hosts infected by the avirulent type. Coexistence can arise because intra-host selection favours the virulent strain whereas inter-host selection favours the avirulent strain. Bremermann & Pickering (1983) pointed out that intra-host competition between different strains might lead to a degree of virulence larger than required for the optimal spread in the host population. Nowak & May (1994) have studied the effect of superinfection on the evolution of virulence and the diversity of the parasite population. Their basic conclusions are that superinfection leads to complex polymorphisms of many different parasite strains at virulences increased beyond what would be optimal for the parasite population (see also May & Nowak 1994).

In this paper we investigate a simple analytically solvable model, which combines intra- and inter-host parasite competition of a large number of different strains that are produced by mutation during the course of an infection. We do not deal with intra-host competition arising from superinfection (infection by a

*Proc. R. Soc. Lond.* B (1994) **258**, 133–140
*Printed in Great Britain*

new parasite strain which then dominates the infection) or coinfection (infection by several strains simultaneously). Essentially, the model is relevant for highly mutating parasites, where intra-host competition between different mutants may play a significant role in the overall evolution of the parasite. We show that rapid intra-host evolution can account for: (i) wide ranges of degrees of virulence which are simultaneously maintained in the parasite population; and (ii) an average virulence larger than required for optimal spread in the population. We analyse several examples for the functional relation between transmission rate, parasite virulence and competitive superiority within an infected host, and we derive analytical expressions for the equilibrium polymorphism of the parasite population.

## 2. THE MODEL

We consider $n$ different parasite strains and order them in a dominance ranking, such that strain $i+1$ dominates strain $i$, i.e. we assume that strain $i+1$ takes over an infection by strain $i$ if both strains are present in the same host. The generation of new dominating strains due to mutation is modelled by assuming that during an infection each strain $i$ produces mutants of type $i+1$ at a rate $q$. We call this rate the mutation rate, although more accurately it describes the rate at which a new dominating strain is produced by mutation and takes over the infection by intra-host selection. Thus $q$ reflects the rate of this combined event. In this model we assume that at any one time infections are always dominated by a single strain type. This simplification is made for reasons of mathematical tractability and represents a starting point for more general investigations where different mutants can coexist during an infection.

Defining $x(t)$ as the density of hosts susceptible to infection at time $t$, and $y_i(t)$ as the density of hosts infected by strain type $i$ at time $t$, we obtain the following model for the host–parasite dynamics:

$$\mathrm{d}x/\mathrm{d}t = k - ux - x\sum_{i=0}^{n}\beta_i y_i, \tag{1}$$

$$\mathrm{d}y_i/\mathrm{d}t = (\beta_i x - u - v_i - q)\,y_i + qy_{i-1}, \quad i = 0,\ldots,n. \tag{2}$$

(We define $y_{-1} = 0$.) The parameters used in this model are: $k$, for the immigration or birth rate of new susceptible hosts; $u$, for the natural death rate of hosts; $\beta_i$, for the per contact transmission rate at which hosts infected by strain $i$ infect susceptible hosts; and $v_i$, for the virulence of strain $i$ (defined as the parasite-induced host mortality). The parameter $q$ reflects the generation of new dominating mutants as described above.

The population dynamics of the uninfected hosts in the absence of parasites is governed by a simple immigration–death process (Kermack & McKendrick 1933; Anderson & May 1979, 1991). This is the simplest way to attain a stable susceptible host population. Different population dynamics for the uninfected hosts in the absence of parasites will only affect the total number of infected and uninfected hosts, but will not affect the population structure.

Note that in equation (2) strain $n$ loses hosts by generating dominant mutants. These hosts do not enter any other infectious class. To avoid this problem one could define $\mathrm{d}y_n/\mathrm{d}t = (\beta_n - u - v_n)\,y_n + qy_{n-1}$. Strains of type $n$ would then be absolutely dominant and could not be outcompeted by any other strain type. Alternatively we may require a suitable choice of parameters ($\beta_i$, $v_i$, $u$, $q$ and $n$) such that $y_{n-1}$ is vanishingly small in equilibrium. (Solving $\mathrm{d}y_n/\mathrm{d}t = 0$ for $y_n$ shows that $y_n$ vanishes for small $y_{n-1}$, and so the term $-qy_n$ becomes negligible in equation (2)). In the following sections we will always assume a suitable choice of parameters such that $y_n$ is very small.

The basic reproductive rate of a parasite strain is defined as the average number of secondary cases produced by a single infected host introduced in an entirely susceptible host population (Anderson & May 1979, 1991). For the above system of differential equations the basic reproductive rate of strain $i$ at the uninfected host equilibrium is:

$$R_i = (k/u)\,\beta_i/(q+u+v_i). \tag{3}$$

To avoid double indices we drop the conventional subscript 0 of the basic reproductive rate. The index refers to different strains. Note that the basic reproductive rate of a strain decreases with increasing mutation rate $q$. This reflects the losses resulting from intra-host competition.

The model defined by equations (1) and (2) is not the most general model for intra-host mutation, because we make the very simplifying assumption that strain $i$ only mutates into $i+1$ and not to any other strain $j$ with $j > i$. A more general model has to include a full mutation matrix $\{q_{ij}\}$ where each element $q_{ij}$ specifies the rate at which infections by strain $i$ are turned into infections by strain $j$. The model then takes the form

$$\mathrm{d}y_i/\mathrm{d}t = y_i(\beta_i x - u - v_i) + \sum_{j=1}^{i-1} q_{ij}y_j - \sum_{j=i+1}^{n} q_{ij}y_j.$$

We do not attempt to analyse this more general model in this paper, but instead concentrate on equation (2) for which we can derive an equilibrium solution. Despite the limitation of equation (2), we feel that it can reveal many interesting features of intra-host mutation and competition which will also hold for more complex models.

## 3. THE SOLUTION

There are two different equilibria depending on the basic reproductive rates $R_i$ (for any initial condition with $v_0 > 0$).

1. If $R_i < 1$ for all strains $i$, then no strain can invade or persist in the host population and the densities of infected hosts $y_i$ converge to zero for all strains $i$. (To see this, substitute the equilibrium density of infected hosts in the absence of parasites, $\hat{x} = k/u$, into equation (2).) $R_i < 1$ gives a threshold for the mutation rate, $q_t = \beta_i k/u - v_i - u$, above which the parasite is unable to persist (or invade) in the host population.

2. If $R_i > 1$ for some $i$ then the unique non-trivial equilibrium is given by

$$\left.\begin{aligned} y_i^* &= 0, && i<j,\\ y_j^* &> 0, && \\ y_i^* &= y_j^* \prod_{k=j}^{i} f_k, && i<j, \end{aligned}\right\} \tag{4}$$

where $j$ is the strain with maximum basic reproductive rate. (In case there are two degenerate maxima of the basic reproductive rate then $j$ is the strain with the higher index number in the dominance ranking.) We have defined $f_k = q/(q+u+v_k-\beta_i x^*)$ an $f_j = 1$. The asterisk indicates equilibrium frequencies. The equilibrium density of uninfected hosts can be derived from equation (2) for $i=j$; we obtain

$$x^* = (u+v_j+q)/\beta_j. \tag{5}$$

Thus the equilibrium density of susceptible hosts only depends on the parameters of the strain with maximum basic reproductive rate. This holds even if the strain with maximal basic reproductive rate might have a vanishingly small equilibrium frequency. By using equations (1) and (4), we obtain for the equilibrium abundance of hosts infected by strain $j$

$$y_j^* = \{[k\beta_j/(u+v_j+q)]-u\}\left[1\Big/\left(\sum_{i=j}^{n}\beta_i\prod_{l=j}^{i}f_l\right)\right]. \tag{6}$$

Thus we have an analytical solution for the equilibrium of the model given by equations (1) and (2). Extensive numerical simulations suggest that this equilibrium is globally stable. (A formal proof of stability may be difficult. Certainly the equilibrium is saturated, which means that the transversal eigenvalues $\lambda_i = \partial\dot{y}_i/\partial y_i < 0$ for $i<j$.)

The equilibrium abundance of uninfected hosts as a function of the mutation rate, $q$, is

$$x^*(q) = \min\{(u+v_i+q)/\beta_i : i = 0,\ldots,n\}. \tag{7}$$

The minimum of $(u+v_i+q)/\beta_i$ corresponds to the maximum of the basic reproductive rate for a given $q$ (see equation (3)). The abundance of uninfected hosts increases continuously with increasing mutation rate (but need not be differentiable with respect to $q$).

We derive the total number of infected hosts from equations (4) and (6).

$$Y^*(q) = [k/x^*(q)-u]\left(\sum_{i=j}^{n}\prod_{l=j}^{i}f_l\right)\Big/\left(\sum_{i=j}^{n}b_i\prod_{l=j}^{i}f_l\right), \tag{8}$$

where $j$ is the strain which maximizes the basic reproductive rate (equation (3)) for a given $q$.

Our analysis has focused on equilibrium solutions only. The time to reach equilibrium depends on the choice of parameters, but qualitatively our observations on the virulence polymorphism and the increase in virulence due to intra-host competition appear to be very robust.

If the mutation rate $q$ equals zero, then equations (1) and (2) reduce to:

$$\mathrm{d}x/\mathrm{d}t = k-ux-x\sum_{i=0}^{n}\beta_i y_i, \tag{9}$$

$$\mathrm{d}y_i/\mathrm{d}t = (\beta_i x-u-v_i)\,y_i \quad i=0,\ldots,n. \tag{10}$$

If some strains have a basic reproductive rate greater than 1, the strain with maximum basic reproductive rate outcompetes all other strains, and in equilibrium the host population is infected by this strain type only. We obtain the equilibrium $x^* = (u+v_j)/\beta_j$, $y_i^* = 0$ for $i \neq j$, and $y_j^* = k/(u+v_j)-u/\beta_j$, where $j$ is the strain with maximum basic reproductive rate. (If all strains have $R_i < 1$ then $y_i^* = 0$ for all $i$.)

In summary, we observe for vanishing mutation rate ($q=0$) that the parasite strain with maximum basic reproductive rate outperforms all competitors and that the host population is infected by a single strain type at equilibrium. If there is some mutation ($q>0$), we obtain a distribution of parasite strains persisting in the host population. The strain with maximum basic reproductive rate has the smallest intra-host dominance (i.e. the lowest index number in the dominance ranking) of those strains which persist in the host population. All strains with higher index number are present with non-vanishing frequency (see equation (4)). Strains with a basic reproductive rate smaller than one arise by intra-host evolution and persist in the host population. They could not become established or persist in the absence of strains with a basic reproductive rate larger then one.

## 4. VIRULENCE AND DOMINANCE ARE EQUIVALENT

In this section we assume that the more virulent strain dominates intra-host competition. In terms of our model we have $v_i < v_{i+1}$ for all $i$. We proceed by analysing two examples showing how a strain's transmission rate, $\beta_i$, may be related to its virulence, $v_i$.

### (*a*) *All strains have the same transmission rate*

In this example, transmission rate shall not be correlated with virulence (or intra-host dominance), and we assume that $\beta_i = \beta$ for all $i$. The basic reproductive rate is then given by:

$$R_i = (k/u)\,\beta/(u+v_i+q). \tag{11}$$

It has its maximum for the smallest value of $v_i$. This is $v_0$.

Let us now consider a continuous approximation and assume $v_0 = 0$. The density of susceptible hosts at equilibrium is given by $X = (u+q)/\beta$ (see equation (5)). Substituting this into the multiplication coefficients $f_i$ (see equation (4)) we obtain $f(v) = q/v$. We can derive the virulence at which the distribution of infected hosts has its maximum by equating $f(v) = 1$, and we obtain $v_{\max} = q$. This means that the most abundant strain has a virulence such that the strain 'loses' a host at the same rate by killing a host and by generating a dominating mutant. The virulence of the most abundant strain increases linearly with mutation rate, $q$.

If we let $v_i = i\Delta v$ (i.e. the virulence classes are evenly distributed), the abundance of hosts infected by strain $i$ is (see equation (4))

$$y_i^* = (q/\Delta v)^i\,(1/i!)\,y_0^*. \tag{12}$$

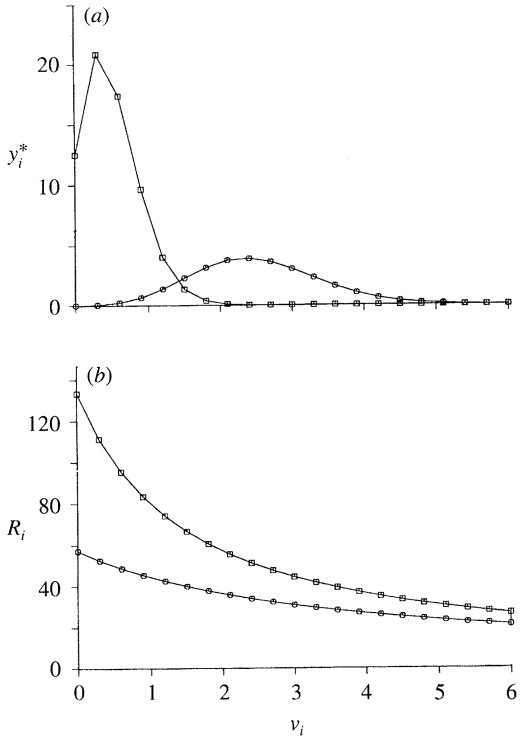


Figure 1. Example from §4*a*; transmission rate is independent of virulence. (*a*) The distribution of infected hosts $y_i^*$, and (*b*) the basic reproductive rate, against the virulence for two different mutation rates (squares denote $q = \frac{1}{2}$, circles denote $q = \frac{5}{2}$). The infected hosts $y_i^*$ are Poisson-distributed with mean $\mu = q/\Delta v$ and variance $\sigma^2 = q/\Delta v$. Interestingly we observe an evolution towards an intermediate level of virulence, although there is no trade-off between transmission rate and virulence. The parameters are: $n = 20$, $\beta = 2$, $k = 100$, $u = 1$, $\Delta v = \frac{6}{20}$.

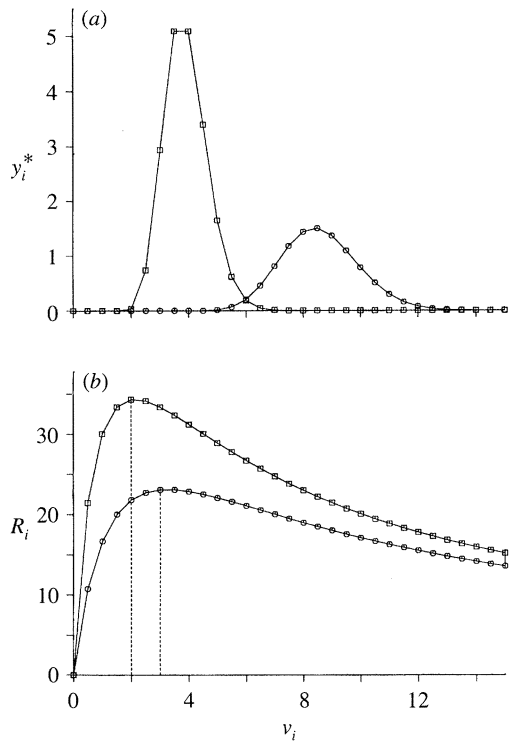


Figure 2. Example from §4*b*; transmission rate increases with virulence. (*a*) The distribution of infected hosts $y_i^*$ (equation (6)), and (*b*) the basic reproductive rate (equation (3)), against the virulence for two different mutation rates (squares denote $q = \frac{1}{2}$, and circles denote $q = \frac{5}{2}$). The distribution of infected hosts becomes broader with increasing mutation rate, $q$. The maxima of the basic reproductive rate (shown by the broken line) shift to higher virulences with increasing $q$. The parameters are: $n = 30$, $a = 1$, $c = \frac{1}{3}$, $k = 100$, $u = 1$.

From equation (6) we derive for

$$y_0^* = [k\beta/(u+q)-u]\left\{1\Big/\left[\beta\sum_{i=0}^{n}(q/\Delta v)^i\,1/i!\right]\right\} \qquad (13)$$

$$= [k/(u+q)-(u/\beta)]\,\mathrm{e}^{-q/\Delta v} \quad \text{for} \quad n\to\infty, \qquad (14)$$

and we obtain for the distribution of infected hosts (see figure 1)

$$y_i^* = [k/(u+q)-(u/\beta)]\,[(q/\Delta v)^i/i!]\,\mathrm{e}^{-q/\Delta v}. \qquad (15)$$

This is a Poisson distribution with mean $\mu = q/\Delta v$ and variance $\sigma^2 = q/\Delta v$. Thus, both mean and variance increase linearly with increasing mutation rate, $q$. The total number of infected hosts,

$$Y^*(q) = k/(u+q)-(u/\beta), \qquad (16)$$

decreases with increasing mutation rate, $q$.

Interestingly, we observe an evolution towards an intermediate level of virulence, although we do not assume an explicit trade-off between transmission rate and virulence in this example. The resulting mean level of virulence arises from the competition of parasite strains on two levels: (i) within the host population the less virulent strains are favoured because of their larger basic reproductive rate; and (ii) within a single host the more virulent strains are favoured, because of their intra-host competitive superiority.

### (*b*) *Transmission rate increases with virulence*

We now assume that the transmission rate is also functionally related to virulence and intra-host dominance. The transmission rate shall increase with increasing virulence and saturate for large virulences. The simplest functional relation between virulence and transmission rate fulfilling these requirements is $\beta_i = av_i/(v_i+c)$.

The basic reproductive rate as a function of $v$ is

$$R(v) = (k/u)\,\{av/[(v+c)(u+v+q)]\}, \qquad (17)$$

and has its maximum at $v = \sqrt{[c(u+q)]}$. The maximum of $R(v)$ shifts to higher levels of virulence with increasing $q$, and the magnitude of the maximum decreases with increasing $q$ (see figure 2).

The maximum of the distribution of infected hosts can be derived from $f(v) = 1$; we obtain

$$v_{\max} = q/2+\sqrt{[c(u+q)]}+\sqrt{\{cq+q^2/4+q\sqrt{[c(u+q)]}\}}. \qquad (18)$$

In this example the maximum of the distribution of infected hosts depends on the mutation rate $q$, the natural death rate $u$, and the constant $c$. For large $q$ the maximum of the distribution is given by $v_{\max} = q$ as in the first example (and the virulence of the most abundant strain is such that hosts are lost by killing and intra-host dominance at the same rate). For $q = 0$,

only the strain with maximum basic reproductive rate is selected, i.e. the maximum of the distribution of infected hosts at $v_{max} = \sqrt{(cu)}$ coincides with the maximum of the basic reproductive rate $R(v)$ (see equation (17)). For small $q$ the maximum grows faster than linear (proportional to $\sqrt{q}$), and the distribution gets broader with increasing $q$ (see figure 2).

## 5. VIRULENCE AND DOMINANCE ARE NOT EQUIVALENT

In the following examples we will relax the assumption that a strain's virulence and its ability to dominate intra-host competition are equivalent. Although one might expect a correlation between these two properties, there is no *a priori* reason to believe that virulence and intra-host competitive superiority must be equivalent (Bonhoeffer & Nowak 1994*a*, *b*).

### (*a*) *Only one transmissible strain*

The simplest assumption is that all strains have the same death rate ($v_i = v$), but only one strain, say $j$, is infectious ($\beta_j = \beta$ and $\beta_i = 0$ for all $i \neq j$). For the abundance of host infected by strain $j$ we obtain $y_j^* = k/(u+v+q) - u/\beta$. The abundance of hosts infected by strains with $i > j$ is $y_i^* = f^{i-j} y_j$ where $f = f_i =$ $q/(q+v+u) < 1$. The infected hosts are distributed exponentially with mean $\mu = -1/\ln(f) + j$ and variance $\sigma^2 = 1/\ln(f)^2$.

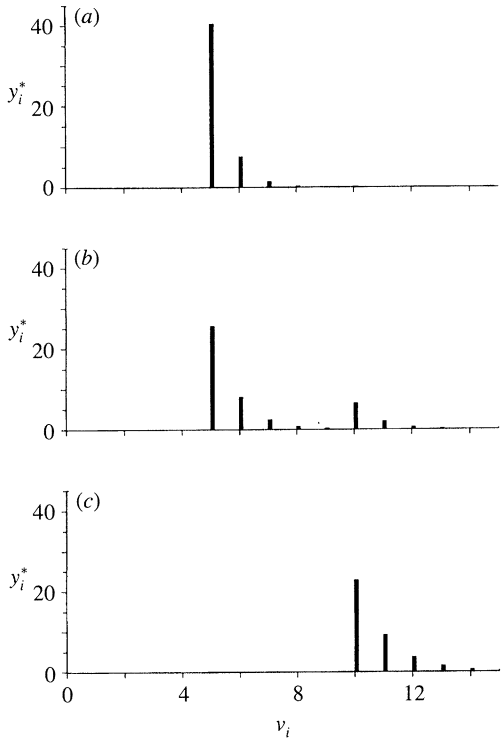


Figure 3. Example from §5*b*; two transmissible strains. These plots show the distribution of infected host $y_i^*$ (equation (6)) against the strain index $i$ for three different mutation rates (*a*) $q = 0.5$, (*b*) $q = 0.9$, (*c*) $q = 1.3$. The threshold mutation for which the distribution of infected host shifts from strain $k$ to strain $l$ is $q_c = 1$. We observe a single peak at strain 5 for $q = 0.5$, and two peaks at strains 5 and 10 for $q = 0.9$. For $q = 1.3$ the distribution shifts to strain 10. The parameters are: $\beta_i = 0$, $\beta_5 = 3$, $\beta_{10} = 4$, $v_i = 1$, $v_5 = 1$, $v_{10} = 2$, $n = 100$, $k = 100$, $u = 1$.

We can derive an expression for the total abundance of infected hosts

$$Y^* = \sum_{i=j}^{n} y_i^* = y_j^* \sum_{l=0}^{n-j} f^l \tag{19}$$

$$= [(1-f^{n-j})/(1-f)]\, y_j^*, \tag{20}$$

and in the limit of large $n$

$$\lim_{n\to\infty} Y^* = [1/(1-f)]\, y_j^* \tag{21}$$

$$= [1/(v+u)]\,\{k - [u(v+u+q)/\beta]\}. \tag{22}$$

The total abundance of infected hosts decreases linearly with increasing mutation rate $q$ and vanishes if $q > k\beta/u - v - u$ (which is equivalent to a basic reproductive rate smaller than one).

In this example we observe a stable population of parasite strains with exponentially distributed levels of virulence persisting in the host population, although only a single strain type is infectious. Strains with vanishing transmission and basic reproductive rate can nevertheless persist in the host population.

### (*b*) *Two transmissible strains*

We extend the previous example and consider two infectious strains ($\beta_i = 0$ for all $i$ except $k$, $l$ for which $\beta_k, \beta_l > 0$). All strains shall have the same disease-induced death rate ($v_i = v$) except strains $k$ and $l$ which have $v_k$ and $v_l$. We assume that $k < l$.

The basic reproductive rate for strain $k$ is $R_k = \beta_k/(v_k+u+q)$, and for strain $l$ $R_l = \beta_l/(v_l+u+q)$. The maximum of $R$ depends on $q$. Let $R_k > R_l$ for small $q$, but $R_l > R_k$ for large $q$. This yields the conditions $\beta_l > \beta_k$ and $\beta_k/(v_k+u) > \beta_l/(v_l+u)$. The maximum of the basic reproductive rate and the distribution of infected hosts switches from strain $k$ to strain $l$ at

$$q_c = [\beta_k(v_l+u) - \beta_l(v_k+u)]/(\beta_l - \beta_k). \tag{23}$$

We derive for the total abundance of infected hosts (for large $n$):

$$Y^* = \begin{cases} [k\beta_k/(v_k+u+q) - u] \\ \quad \{(1-f^{l-k-1})/(1-f) + f^{l-k-1} f[1 + 1/(1-f)]\}/ \\ \qquad (\beta_k + \beta_l f_l f^{l-k-1}), & q < q_c, \\ 1/(v+u)\,(k(q+v+u)/(q+v_l+u) \\ \qquad - [u(q+v+u)/v_l\}, & q > q_c, \end{cases} \tag{24}$$

where $f = q/(q+v+u)$ and $f_l = q\beta_k/[\beta_k(q+v_l+u) - \beta_l(q+v_k+u)]$. If $l-k-1$ is sufficiently large, the above expression simplifies considerably and we obtain

$$Y^* = \begin{cases} 1/(v+u)\,[k(q+v+u)/(q+v_k+u) \\ \qquad - u(q+v+u)/\beta_k], & q < q_c \\ 1/(v+u)\,[k(q+v+u)/(q+v_l+u) \\ \qquad - u(q+v+u)/\beta_l], & q > q_c. \end{cases} \tag{25}$$

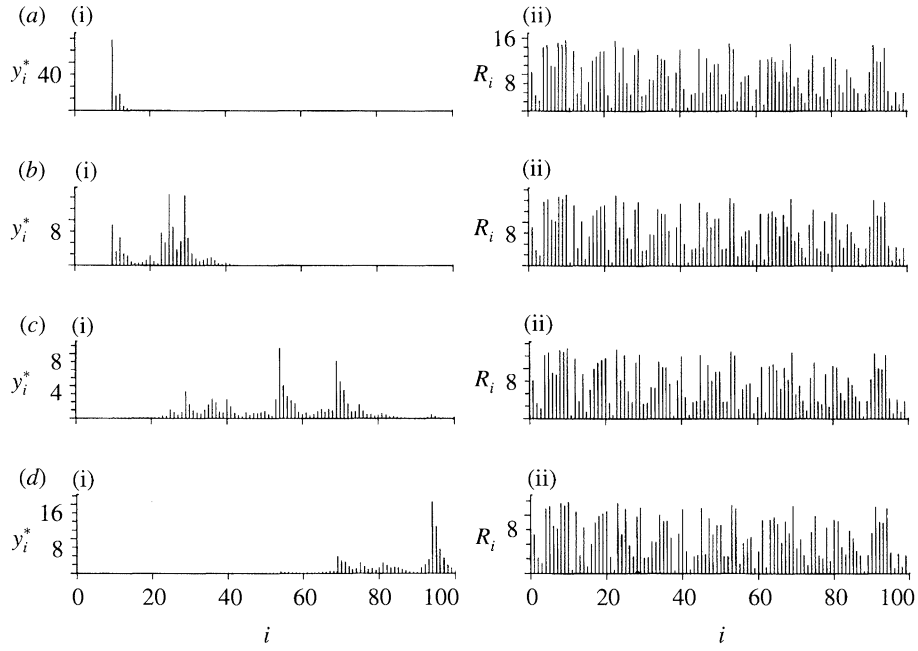


Figure 4. Example from §5*c*; no correlation between virulence and transmission rate. (i) The distribution of infected hosts $y_i^*$ and (ii) the basic reproductive rate $R_i$, for four different mutation rates, (*a*) $q = \frac{1}{4}$, (*b*) $q = \frac{2}{4}$, (*c*) $q = \frac{3}{4}$, and (*d*) $q = \frac{4}{4}$. The transmission rate $\beta_i$ and the virulence $v_i$ of each strain are chosen from a uniform random distribution. We observe a highly complex multipeaked parasite distribution. Parameters: $u = 1$, $n = 100$, $k = 100$.

$Y^*$ may be discontinuous and can even increase with $q$ locally.

This example illustrates two points (see also figure 3): (i) there is a critical mutation rate at which the distribution of infected host shifts from one peak to another; and (ii) the total number of infected hosts may be related to the mutation rate in a complicated way.

### (*c*) *Arbitrary values for transmission rate and virulence, no correlation*

As a final example we investigate the situation where all parasite strains have different transmission rates and virulences. In §5*a*, *b* we assumed a direct correlation between virulence and dominance (and transmission rate, respectively). Now we consider the other end of the spectrum by assuming as a null model that all three properties (transmission rate, virulence and dominance) are independent of each other. Thus we keep a dominance ranking but assign arbitrary values to the transmission rate $\beta_i$ and virulence $v_i$ of each strain. The analysis of the previous examples allows us to understand the observed complex distribution of infected hosts and its dependence on the mutation rate. In figure 4 we show the distribution of infected hosts and the basic reproductive rate for randomly distributed transmission rates $\beta_i$ and virulences $v_i$. The transmission rates and virulences were chosen from a uniform random distribution. In figure 4 we observe a highly complex multipeaked distribution of infected hosts. As the mutation rate increases, the distribution becomes broader and the total number of infected hosts decreases. This example demonstrates that intra-host competition (arising from mutation only) can account for a highly complex parasite population structure.

## 6. CONCLUSION

We have analysed a combined model of intra- and inter-host parasite evolution and investigated the effect of mutation on the evolution of virulence. In the present model, competition between mutants during an infection only arises from (intra-host) mutation of the parasite. We explicitly excluded the possibility of superinfection (May & Nowak 1994; Nowak & May 1994) or co-infection (M. A. Nowak & R. M. May, unpublished results).

Our model is relevant for those parasites that have: (i) a significant production rate of new dominating mutants during an infection (this may correlate with a high nucleotide misincorporation rate); and (ii) a large replication rate and a long duration of the infection (this results in a large number of parasite generations and will allow selectively advantageous mutants to grow to high frequencies during an infection). Therefore, the model provides some insight into the evolution of virulence for persistent bacterial or viral infections such as tuberculosis, measles or retroviral infections. An interesting example is, of course, the human immunodeficiency virus (HIV). HIV replication is error prone (Preston *et al.* 1988; Roberts *et al.* 1988). Sequential virus samples taken at several time points during an individual infection demonstrate that different mutants prevail at different time points in the infection (Holmes *et al.* 1992). HIV replication occurs throughout the course of infection (Embretson *et al.* 1993; Pantaleo *et al.* 1993). The intra-host evolution of HIV may even be the mechanism of disease progression (Nowak *et al.* 1990, 1991). Ewald (1993) argues that HIV evolves to become more virulent if the transmission opportunities increase. Although our model does not specifically describe the dynamics of sexually transmitted diseases (Lipsitch & Nowak 1994), it never-

theless suggests that the virulence of HIV may be determined by intra-host evolution. Hence, changing the epidemiological transmission pattern may have little effect on the virulence of HIV.

Intra- and inter-host evolution of microparasites may work in opposing directions. Intra-host selection will favour dominating mutants regardless of their effect on host mortality, whereas inter-host selection may work to decrease virulence. It is the balance between inter- and intra-host selection advantages which allows for the stable polymorphism in the parasite population.

We summarize the most important observations.

1. The strain with the largest basic reproductive rate is the lowest in the dominance ranking of those strains which persist in the host population. It can occur with very low abundance in the parasite population.

2. Strains with a basic reproductive rate smaller than 1 can persist in the host population. Even strains without any potential for transmission can be held in equilibrium by the mutation–extinction balance.

3. Intra-host parasite competition shifts the distribution of parasite strains to larger virulences if intra-host dominance and virulence are correlated. The resulting level of virulence is determined by the trade-offs between transmission rate, virulence and intra-host dominance.

4. We can observe an evolution towards an intermediate level of virulence even if there is no trade-off between virulence and transmission rate.

5. A high mutation rate can be disadvantageous for the parasite, because it may enhance the effect of 'short-sighted' intra-host competition and may prevent the parasite from adapting the optimal strategy (in respect to virulence and transmissibility) for infecting a large number of new hosts. Indeed, the total number of infected hosts can decline with increasing mutation or take-over rate, $q$. Essentially a high mutation rate increases the relative importance of intra-host over inter-host competition, and the latter determines the spread of the parasite in the host population.

6. The equilibrium density of susceptible hosts depends only on the parameters of the strain with maximum basic reproductive rate.

7. Mutation can lead to a stable polymorphism of parasites in the host population.

8. Increasing mutation broadens the distribution of parasite virulence in the host population.

9. There is a threshold for the rate of intra-host evolution above which the parasite is unable to persist in the host population.

We thank Robert May and Karl Sigmund for interesting discussion. This work was supported by the Boehringer Ingelheim Fonds (to S.B.), and a Wellcome Trust Senior Research Fellowship and the E. P. Abraham Junior Research Fellowship at Keble College (to M.A.N.).

## REFERENCES

Anderson, R. M. & May, R. M. 1979 Population biology of infectious diseases I. *Nature, Lond.* **280**, 361–367.

Anderson, R. M. & May, R. M. 1981 The population dynamics of microparasites and their invertebrate hosts. *Phil. Trans. R. Soc. Lond.* B **291**, 451–524.

Anderson, R. M. & May, R. M. 1982 Coevolution of hosts and parasites. *Parasitology* **85**, 411–426.

Anderson, R. M. & May, R. M. 1991 *Infectious diseases of humans*. Oxford University Press.

Antia, R., Levin, B. & May, R. M. 1994 Within-host population dynamics and the evolution and maintenance of microparasite virulence. *Am. Nat.* **144**, 457–472.

Bonhoeffer, S. & Nowak, M. A. 1994*a* Can live attenuated viruses work as a post-exposure treatment? *Immunol. Today.* (In the press.)

Bonhoeffer, S. & Nowak, M. A. 1994*b* Intra-host and inter-host selection: viral evolution of immune function impairment. *Proc. natn. Acad. Sci. U.S.A.* **91**, 8062–8066.

Bremermann, H. J. & Pickering, J. 1983 A game-theoretical model of parasite virulence. *J. theor. Biol.* **100**, 411–426.

Embretson, J., Zupancic, M., Ribas, J., Burke, A., Racz, P., Tenner-Racz, K. & Haase, A. 1993 Massive covert infection of helper T lymphocytes and macrophages by HIV during the incubation period of AIDS. *Nature, Lond.* **362**, 359–362.

Ewald, P. W. 1983 Host–parasite relations, vectors, and the evolution of disease severity. *A. Rev. Ecol. Syst* **14**, 465–485.

Ewald, P. W. 1993 Evolution of virulence. *Sci. Am.* **268** (4), 86–93.

Fenner, F. & Ratcliffe, F. 1965 *Myxomatosis*. Cambridge University Press.

Frank, S. A. 1992 A kin selection model for the evolution of virulence. *Proc. R. Soc. Lond.* B **250**, 195–197.

Herre, E. A. 1993 Population structure and the evolution of virulence in nematode parasites of fig wasps. *Science, Wash.* **259**, 1442–1445.

Holmes, E., Zhang, L., Simmonds, P., Ludlam, C. & Brown, A. L. 1992 Convergent and divergent sequence evolution in the surface envelope glycoprotein of human immunodeficiency virus type 1 within a single infected patient. *Proc. natn. Acad. Sci. U.S.A.* **89**, 4835–4839.

Kermack, W. O. & McKendrick, A. G. 1933 Contributions to the mathematical theory of epidemics. *Proc. R. Soc. Lond.* A **141**, 94–122.

Knolle, H. 1989 Host density and the evolution of parasite virulence. *J. theor. Biol.* **136**, 199–207.

Levin, B. R. & Bull, J. J. 1994 Short-sighted evolution and the virulence of pathogenic microorganisms. *Trends Microbiol.* **2**, 76–81.

Levin, S. & Pimentel, D. 1981 Selection of intermediate rates of increase in parasite–host systems. *Am. Nat.* **117**, 308–315.

Lipsitch, M. & Nowak, M. A. 1994 The evolution of virulence in sexually transmitted HIV/AIDS. *J. theor. Biol.* (In the press.)

May, R. M. & Anderson, R. M. 1979 Population biology of infectious diseases II. *Nature, Lond.* **280**, 455.

May, R. M. & Anderson, R. M. 1983 Epidemiology and genetics in the coevolution of parasites and hosts. *Proc. R. Soc. Lond.* B **219**, 281–313.

May, R. M. & Anderson, R. M. 1990 Parasite–host coevolution. *Parasitology* **100**, 89–101.

May, R. M. & Nowak, M. A. 1994 Superinfection, metapopulation dynamics, and the evolution of diversity. *J. theor. Biol.* **170**, 95–114.

Nowak, M. A. & May, R. M. 1994 Superinfection and the evolution of parasite virulence. *Proc. R. Soc. Lond.* B **255**, 81–89.

Nowak, M. A., May, R. M. & Anderson, R. M. 1990 The evolutionary dynamics of HIV-1 quasispecies and the

development of immunodeficiency disease. *AIDS* **4**, 1095–1103.

Nowak, M. A., Anderson, R. M., McLean, A. R., Wolfs, T. F., Goudsmit, J. & May, R. M. 1991 Antigenic diversity thresholds and the development of AIDS. *Science, Wash.* **254**, 963–969.

Pantaleo, G., Graziosi, C., Demarest, J., Butini, L., Montroni, M., Fox, C., Orenstein, J., Kotler, D. & Fauci, A. 1993 HIV infection is active and progressive in the lymphoid tissue during the clinically latent stage of disease. *Nature, Lond.* **362**, 355–358.

Preston, B. D., Poiesz, B. J. & Loeb, L. A. 1988 Fidelity of HIV-1 reverse transcriptase. *Science, Wash.* **242**, 1168–1171.

Roberts, J. D., Bebenek, K. & Kunkel, T. A. 1988 The accuracy of reverse transcriptase from HIV-1. *Science, Wash.* **242**, 1171–1173.

*Received 15 June 1994, accepted 1 August 1994*