Polymorphism in sexual versus non-sexual disease transmission

PETER H. THRALL AND JANIS ANTONOVICS

Department of Botany, Duke University, Durham, North Carolina 27708, USA

SUMMARY

Pathogens causing sexually transmitted diseases (STDs) often consist of related strains that cause non-sexually transmitted, or 'ordinary infectious', diseases (OIDs). We use differential equation models of single populations to derive conditions under which a genetic variant with one (e.g. sexual) transmission mode can invade and successfully displace a genetic variant with a different (e.g. non-sexual) transmission mode. Invasion by an STD is easier if the equilibrium population size in the presence of an OID is smaller; conversely an OID can invade more easily if the equilibrium size of the population with the STD is larger. Invasion of an STD does not depend on the degree of sterility caused by the infection, but does depend on the added mortality caused by a resident OID. In contrast, the ability of an OID to invade a population at equilibrium with an STD decreases as the degree of sterility caused by the STD increases. When equilibrium population sizes for a population infected with an STD are above the point at which non-sexual contacts exceed sexual contacts (the sexual-social crossover point) and when equilibrium population sizes for an OID are below this point, there can be a stable genetic polymorphism for transmission mode. This is most likely when the STD is mildly sterilizing, and the OID causes low or intermediate levels of added mortality. Because we assume the strains are competitively equivalent and there are no heterogeneities associated with the transmission process, the polymorphism is maintained by density-dependent selection brought about by pathogen effects on population size.

1. INTRODUCTION

It has long been known that many common pathogens consist of multiple strains that are distinguishable either serologically or on the basis of other criteria, such as their nutritional requirements, resistance to antibiotics, or host range. The possibility that pathogen populations may be genetically variable with regard to transmission mode has been much less intensively studied, perhaps because transmission mode itself has been used as an important way to classify and distinguish diseases (Ewald 1994). Nevertheless, it is clear that transmission modes are not distinct, and that many diseases do not fall clearly into one category or another. Frequently one disease can exhibit multiple transmission modes. For example, out of 108 sexually transmitted diseases (STDs) for which reasonable estimates could be obtained of the degree of sexual transmission, 87 had alternative routes of transmission (Lockhart *et al.* 1996). This situation could result from a single genotype having several transmission routes or it may be the result of the disease being composed of several strains that are individually specialized for different routes of transmission. This issue is of particular relevance to STDs because, as has often been pointed out (Nahmias, personal communication), many of these diseases seem to show an interesting duality, where the sexually transmitted form of the pathogen has a 'relative' that is non-sexually trans-

mitted (syphilis and yaws; genital and facial herpes; pubic and head lice; genital and cutaneous warts; genital and ocular chlamydias). Strain variation with regard to transmission mode is well known. Thus the major 'biovars' of *Chlamydia trachomatis* are not only serologically distinct, but some cause primarily lymphogranuloma venereum, while others are mostly responsible for ocular trachoma. Within the trachoma class of biovars, serotype A has never been recovered from the genital tract, whereas serotype B has been isolated from both genital and ocular sites (Schachter 1990). Phylogenetic analysis of papillomaviruses that cause warts shows them to consist of two major groups characterized by mucosal (including sexual) infection and cutaneous (primarily non-sexual) infection. A third group contains types that preferentially infect cutaneous sites but are also found in mucosal areas (Zur Hausen & de Villiers 1994). Other diseases, such as syphilis and yaws, have quite different transmission modes and symptoms, yet the pathogens causing them, while often given sub-specific status, (*Treponema pallidum* ssp. *pallidum* and *T. pallidum* ssp. *pertenue*) are barely distinguishable in terms of the amino acid composition of their proteins (Norris 1993) or in the serological profiles of their antigens (Martin *et al.* 1990); the two pathogens are in all likelihood very closely related.

Even the STD/non-STD pathogen groups that probably represent quite divergent lineages (genital

and oral herpes: Gentry *et al.* 1988; McGeoch & Cook 1994; genital and cuticular warts: Bernard *et al.* 1994; head and pubic lice: Salazar 1994; and *Chlamydia trachomatis* biovars: Schachter 1990; the *Brucella* group: Nielsen & Duncan 1990) may show some immunological cross-reactivity or cross-protection. For example, while prior exposure to oral herpes (HSV-1) does not protect against genital herpes (HSV-2), epidemiological data strongly suggest that HSV-2 may provide substantial protection against HSV-1 infections (Nahmias *et al.* 1990). In mouse models, vaccination with one *Brucella* species (e.g. *B. abortus*) protects against infection with other species (e.g. *B. melitensis* and *B. ovis*) that differ in their degree of sexual transmission (Jimenez de Bagues *et al.* 1994a, b).

Given the common occurrence of strain variation in disease expression and transmission mode, and the likelihood that many of these strains (at least in their early divergence) are likely to have shown cross-immunity, it is particularly interesting to ask under what conditions a strain with one transmission mode could invade and displace an alternative transmission mode. In the present study, we use a general heuristic theoretical model to investigate the conditions under which a disease variant with sexual transmission can invade a pathogen showing non-sexual disease transmission, and vice versa. Moreover, we investigate if there can be stable polymorphisms for alternative transmission routes in the pathogen, and so ask whether the presence of several transmission modes could be due to the presence of polymorphic pathogens rather than due to a single pathogen genotype that is a 'generalist' with regard to several transmission modes. We assume that the two strains cannot simultaneously infect the same host; moreover, we assume a single population with no spatial or temporal variation in transmission mode, and a single host type with no genetic variation in behavioural or physiological features affecting transmission through alternative routes. Under such conditions, a stable polymorphism might not be expected; however, we show that a sexually transmitted pathogen strain can nevertheless stably persist with a pathogen strain that is transmitted through non-sexual routes.

2. THE MODEL

We begin by assuming a population of uniform individuals (i.e. with no age, genotype or environmental heterogeneity) in which two types of contact are possible: sexual, and non-sexual or 'social'. We follow previous theoretical work and assume that non-sexual transmission can be approximated by a simple density-dependent function (e.g. Anderson & May 1981), while sexual transmission is best represented by a frequency-dependent function (Getz & Pickering 1983; Thrall *et al.* 1993, 1995). The relationship between number of contacts and population density is therefore implicitly determined by linear functions (figure 1) that have a slope of 0 in the case of pure sexual transmission (the number of sexual contacts per unit time interval, c_s , is assumed to be independent of

Figure 1. Diagram showing the relationship between effective contact number and population density for the cases of frequency-dependent (sexual) and density-dependent (non-sexual) disease transmission. Effective contact number is defined as the product of the actual contact number, c , and the per contact probability of infection δ . The sexual and non-sexual disease transmission coefficients are β_s and β_1 , respectively, and k_1 is the rate at which the number of non-sexual contacts increase with increasing density. For further explanation, see text.

density), and a positive slope for non-sexual transmission (number of non-sexual contacts, c_1 , is assumed to increase proportionately with density). The force of infection (defined as the rate at which susceptible hosts become infected) is then made directly proportional to these linear functions. In the case of frequency-dependent (sexual) transmission, the force of infection is given by $\beta_s(Y/N)$ where Y and N are the density of infected hosts and total population density, respectively, and β_s is the sexual disease transmission coefficient. Expressed in terms of contacts, $\beta_s = c_s \delta_s$, where δ_s = per-contact probability of infection. In the case of density-dependent (non-sexual transmission), the force of infection is $\beta_1 Y$. Expressed in terms of contacts, this force of infection, $\beta_1 Y = c_1 \delta_1 (Y/N)$, where the subscript 'I' denotes the non-sexual route of transmission. Because the number of contacts, c_1 , is assumed to be a linear function of N , this can be rewritten as $\beta_1 Y = k_1 N \delta_1 (Y/N)$, where k_1 represents the rate at which contacts increase with increasing N . The population size at which the number of non-sexual contacts exceeds the number of sexual contacts is given by β_s / β_1 (figure 1); this point will be termed the 'social-sexual crossover point' as it will become important in the analyses below. It is important to note that $\beta_1 = k_1 \delta_1$ (not $c_1 \delta_1$) and consequently β_1 and β_s have different units (Thrall *et al.* 1995); moreover, although the β s are often referred to as 'transmission coefficients', they do not represent the actual per contact transmission rates (Antonovics *et al.* 1995).

We assume that one consequence of transmission through a sexual route is an added degree of sterility in infected hosts, but that hosts infected by non-sexually transmitted pathogens do not have reduced fertility. We further assume that transmission through the non-sexual route may result in a greater mortality rate for infected hosts, relative to infection through the sexual route. Lockhart *et al.* (1996) have shown as a general phenomenon with respect to animal diseases, that

OIDs are less likely to cause sterility than STDs, but may cause much greater mortality.

We assume that there is one host and two pathogen types, but that superinfection (i.e. simultaneous infection by both pathogens) does not occur. This is biologically equivalent to assuming that such double infections are rare, or that the first pathogen to invade a host individual preempts the available resources (Begon & Bowers 1995). In our models, the two pathogen types might therefore represent two strains of a single pathogen species.

We can represent the numerical dynamics of a susceptible host (X), a non-sexually transmitted pathogen (Y_1) and a sexually transmitted pathogen (Y_2) as:

$$\frac{dX}{dt} = b[X + Y_1 + (1-\epsilon)Y_2] - X\left(\beta_1 Y_1 + \beta_s \frac{Y_2}{N} + \mu\right), \quad (1)$$

$$\frac{dY_1}{dt} = Y_1(\beta_1 X - \mu'), \quad (2)$$

$$\frac{dY_2}{dt} = Y_2\left(\beta_s \frac{X}{N} - \mu\right). \quad (3)$$

In equations (1)–(3), ϵ is the ‘sterility component’ of sexual transmission (i.e. $\epsilon = 1$ means that infection by the sexually transmitted pathogen causes complete sterility), b is the host birth rate, and β is the disease transmission parameter. We represent the disease-free mortality rate as μ ; the parameter μ' is the mortality rate for hosts infected by a pathogen transmitted through non-sexual means. For all analyses, we assume that $\mu' > \mu$ (i.e. $\mu' - \mu$ is the ‘mortality component’ of non-sexual transmission), and that infection by a sexually-transmitted pathogen does not result in any added mortality. The total population size is given by N ($N = X + Y_1 + Y_2$).

We assume a density-dependent per capita reproductive rate where

$$b = \frac{b_0}{\zeta N + 1}. \quad (4)$$

The parameter b_0 is the maximum reproductive rate (i.e. the limit of N as it approaches zero) and ζ is a constant that determines the strength of density-dependence. The per capita reproductive rate given by equation (4) declines hyperbolically as population density (N) increases.

While a full stability analysis of the model system described above is complicated, we derive conditions for the ability of either an OID or an STD to invade a population of hosts at equilibrium with the alternative pathogen. In general, global extinction results if the maximum per capita birth rate (b_0) is less than the weighted mean mortality rate in the population. An approximate condition for global extinction is when $b_0 < (1/N')(\mu X + \mu' Y_1 + \mu Y_2)$ where $N' = X + Y_1 + (1-\epsilon)Y_2$ (essentially the ‘effective’ population size with respect to reproduction). Therefore, in all of the analyses that follow, we constrain b_0 to values larger than the maximum mean mortality rate (i.e. $b_0 > \max\{\mu', \mu\}$). If transmission rates for both the

normal and sexually transmitted pathogen are low ($\beta_1 < \zeta\mu\mu'/(b_0 - \mu)$, $\beta_s < \mu$), then neither can increase when rare, and in the absence of the pathogen, host density will converge to its carrying capacity, $X^* = (b_0 - \mu)/\zeta\mu$.

We begin by showing the conditions under which the two boundary equilibria, $(X^*, Y_1^*, 0)$ and $(X^*, 0, Y_2^*)$ are stable. We then derive conditions for either pathogen to invade a population at equilibrium with the other.

(i) Stability of host–OID system: invasion of STD

Provided that $\beta_1 > \zeta\mu\mu'/(b_0 - \mu)$, there will be a single internal equilibrium with both the host (X^*) and the OID (Y_1^*) present at

$$X^* = \frac{\mu'}{\beta_1}, \quad (5)$$

$$Y_1^* = \frac{\beta_1 F^* - \beta_1 - \zeta\mu'}{\zeta\beta_1}, \quad (6)$$

where F is a new variable ($F = qN + 1$; see Thrall *et al.* (1995); and Antonovics & Thrall (1995) for similar analyses); positivity of Y_1^* requires $F^* > (\beta_1 + \zeta\mu')/\beta_1$ which is always true when the invasion condition for β_1 is met. In this case, the expression for F^* is a quadratic, and we therefore do not solve for it explicitly in the following analysis (however, if second order terms and higher in μ and/or μ' are ignored, then $F^* \approx (b_0 - \mu')/\zeta\mu'$; numerical evaluation of this approximation shows close agreement with the more complicated quadratic form if the added mortality due to non-sexual transmission is small). Note that the value for X^* is simply the threshold density required for pathogen invasion.

The equilibrium given by equations (5) and (6) is stable whenever

$$b_0 < \frac{\mu'(\beta_1 + \zeta\mu')^2}{\beta_1^2}. \quad (7)$$

If we substitute the equilibrium expressions (X^* , Y_1^*) into equation (3) from above, then we can derive the invasion condition for a sexually transmitted pathogen:

$$\beta_s > \frac{\mu\beta_1 N_1^*}{\mu'}, \quad (8)$$

where N_1^* is the total population size at the equilibrium (X^* , Y_1^* , 0). Alternatively, the invasion criteria could be written in terms of R_0 , where R_{0s} , the STD reproductive rate, is given by $\beta_s R_{01}/\mu N_1^*$; invasion requires $R_{0s} > 1$. It is worth noting that while the invasion condition does depend on N_1^* , it is independent of the degree of sterility caused by a sexually transmitted infection.

(ii) Stability of host–STD system: invasion of OID

In this case, we begin by assuming that the OID is absent; $\beta_s > \mu$ is the necessary condition for the increase of an STD in a completely susceptible host

Figure 2. Phase diagrams showing regions of monomorphism and polymorphism of two pathogen types with sexual and non-sexual transmission modes (grey = STD wins; black = OID wins; white = coexistence). (a) Effect of varying the sexual (β_s) and non-sexual (β_1) disease transmission coefficients. Section above and to right of white lines shows region of biologically feasible equilibria. (b) Effect of increasing varying sterility and added mortality $(\mu' - \mu)/(1 - \mu)$. All equilibria are feasible. (c) Effect of varying birth rate b_0 and density dependence. Increasing birth rate and decreasing strength of density dependence has the effect of increasing the disease-free carrying capacity of the host population clockwise from this line. Section to right of diagonal white line shows region of feasible equilibria. Except where parameters are varied in the phase diagrams as indicated, canonical values are: $\beta_s = 0.5$; $\beta_1 = 0.02$; $\epsilon = 0.5$; $\mu = 0.2$; $\mu' = 0.6$; $b_0 = 1$; and $q = 0.04$.

population. Under this condition, there will be a single positive internal equilibrium at

$$X^* = \frac{\mu(F^* - 1)}{\zeta\beta_s}, \quad (9)$$

$$Y_2^* = \frac{(F^* - 1)(\beta_s - \mu)}{\zeta\beta_s}, \quad (10)$$

where, in this case, the equilibrium value of F is

$$F^* = \frac{b_0(\beta_s - \epsilon\beta_s + \epsilon\mu)}{\mu\beta_s}. \quad (11)$$

Substituting the expressions for X^* and Y_2^* into equation (2), we solve for the minimum transmission rate (β_1) for invasion of an OID into a host population at equilibrium with an STD:

$$\frac{\beta_1 > \beta_s \mu'}{N_s^* \mu}. \quad (12)$$

Note from equation (11) that F^* , and therefore N_s^* , the total population size at the equilibrium ($X^*, 0, Y_2^*$) decreases as the degree of sterility, ϵ , caused by an STD increases to its maximum value. Therefore, any sterility of individuals infected by an STD will make it more difficult for an OID to increase when rare.

(iii) *Host-STD-OID coexistence*

Combining the invasion condition for an STD to invade a host population with an OID present and vice versa, the overall condition for there to be a valid range for both pathogens to persist is

$$\frac{\mu}{\mu'} N_I < \frac{\beta_s}{\beta_1} < \frac{\mu}{\mu'} N_s^*. \quad (13)$$

This condition means that a valid range for an internal equilibrium to exist is possible only when $N_I^* < N_s^*$. The results from the above analysis indicate that this range will be broadest if the STD does not cause extensive reductions in host fecundity, and if the OID is not very virulent (i.e. does not result in greatly increased mortality). Note that if $\beta_s/\beta_1 > (\mu/\mu') N_s^*$ then the STD will persist and exclude the OID; if $\beta_s/\beta_1 < (\mu/\mu') N_I^*$ then the OID will persist and the STD will be excluded. Simulations of equations (1)–(3) for parameter values that satisfy inequality (13) confirm that stable coexistence is possible.

Because N_I^* and N_s^* are themselves rather complex functions of β_1 and β_s (as well as of μ and μ'), we explored this condition numerically for several particular cases (figure 2). At low and moderate transmission rates for the OID, coexistence was most likely when the amount of disease-induced mortality was intermediate (figure 2a). This was because N_I^* was lowest at these intermediate values; when disease-induced mortality was low ($\mu' \approx \mu$), the disease had little impact on equilibrium population sizes, whereas when disease-induced mortality was high, diseased individuals only lived for a short time, thus reducing the overall opportunity for disease transmission. In-

creasing the sterility effect (ϵ) of the STD reduced the opportunity for STD/OID coexistence because it reduced N_s^* . Coexistence was also favoured when β_s was greater than β_i (but note that these β s are in different units, and are therefore not strictly comparable) and when the β values were high rather than low (figure 2*b*). Changing the per capita rate of reproduction (b_0) and the strength of density dependence (q) also had predictable effects on coexistence through their effects on carrying capacity. STDs were favoured at low population sizes, and OIDs were favoured when population sizes were large (figure 2*c*); transmission mode polymorphisms were stable only at intermediate population sizes.

3. DISCUSSION

A major generalization from the comparative biology of diseases is that disease characteristics will often be determined by transmission opportunity (Ewald 1994). In our model, we use the classical representations of sexual transmission and non-sexual transmission: in STDs the probability of an individual becoming diseased is proportional to the frequency of infectives, whereas in non-sexually transmitted diseases (OIDs) transmission is a mass action process where the rate at which individuals become diseased is proportional to the density of infectives (Getz & Pickering 1983). In keeping with much of the epidemiological literature, we have distinguished between these two modes of transmission using simple 'frequency-' and 'density-dependent' functions. Recently, the validity of the frequency versus density dichotomy has been called into question, both on theoretical and empirical grounds (de Jong *et al.* 1994). Elsewhere, we have also emphasized (Antonovics *et al.* 1995) that the distinction between frequency and density dependence is artificial, and we treat both types of transmission as special cases of monotonically increasing asymptotic functions (Type II functional responses; Hassell 1978) that relate per individual contact rates and population density. Therefore, while in this paper we use 'pure' density and frequency dependence to caricature the distinction between sexual and non-sexual transmission, we recognize that such functions will not always accurately describe the dependency of transmission on density for a particular disease, nor will it always encapsulate the STD/OID distinction. More generally, the model we present reflects the situation where two strains of a pathogen differ in the contact rate/density function, such that one strain is transmitted more effectively at a lower density whereas the other is transmitted more effectively at a higher density. This may be the result of one being sexually and the other being non-sexually transmitted, but it could also pertain to two strains that show non-sexual transmission, but which occur in circumstances where individuals have different types of social interactions. We have recently investigated conditions for pathogen invasion when the pathogens differ in the form of their functional responses, and the qualitative features found in the present study also emerge (Thrall & Antonovics 1997). Clearly, the formulations used in the present

paper represent extreme situations and, as such, are unlikely to provide accurate descriptions of real world transmission processes. Nevertheless, the model presented here does allow us to investigate in a simple way the situation where transmission dynamics vary differentially with density (i.e. contact rates asymptote at different densities), as is likely to be the case for sexually versus non-sexually transmitted pathogens. We therefore discuss our results in this latter context.

We show graphically (figure 1) that these assumptions lead directly to the expectation that transmission opportunity by the sexual route is greater at lower densities, and transmission opportunity via the non-sexual route is greater at higher densities. More formally, we show that the invasion conditions for an STD into an OID (and vice versa) are functions of the equilibrium host population sizes (figure 2*c*); invasion of an STD is easier (i.e. β_s can be smaller) if the equilibrium population size in the presence of an OID is smaller, and conversely an OID can more easily invade (i.e. β_i can be smaller) an STD if the equilibrium size of the population with the STD is larger. Our results therefore reflect the general expectation that sexual transmission would be favoured in low density populations, whereas non-sexual transmission would be favoured at high densities.

In the present study we assumed that the two pathogens represent related strains, and that only one strain could infect a host at any particular time (either because the two strains compete directly for resources or they exhibit cross-immunity). If we assume that pathogens with alternative transmission modes are neutral with respect to each other (in that they can coexist without mutual interference in the host and their effects on host mortality or fecundity are additive), then it is straightforward to show that an STD can always invade an OID as there is no threshold host-density for an STD to spread; conversely, an OID can only invade an STD if the equilibrium population size in the presence of the STD is above the threshold value for OID invasion.

If the two strains cannot simultaneously infect the same host, and there is no host heterogeneity or spatial and temporal variation in transmission mode, then we might naively predict that the strain with the higher transmission rate would spread to fixation. However, from figure 1 and the inequality in equation (13), it can be seen that coexistence of an STD and an OID becomes possible whenever the equilibrium population size with the STD alone is above the social-sexual crossover (β_s/β_i) and the equilibrium population size with the OID is below this point (assuming there is no added mortality, or redefining the crossover point to discount the effects of added mortality). STD-OID coexistence is therefore made possible by a unique form of density-dependent selection, where the divergent densities are generated by the disease dynamics itself: selection favours the STD at low population densities produced by the OID, but favours the OID at high densities produced by the STD. Given the model assumptions of equal resource use and no other heterogeneities, it is clear that this coexistence could only be demonstrated by explicit inclusion of numerical

dynamics into the model. A number of other studies have emphasized the novel results that emerge from inclusion of explicit numerical dynamics in genetic models (May & Anderson 1983; Antonovics 1994; Antonovics & Thrall 1995; Lipsitch & Nowak 1995).

The equilibrium population sizes in the presence of an STD or OID are affected by many features of the host–pathogen relationship. Populations infected by an STD are largest when the sterilizing effects of the STD on the host population are small. Populations infected by an OID are smallest at intermediate levels of virulence; a small amount of disease-induced mortality reduces population size, but as this mortality effect becomes large, pathogen longevity and effective transmission decreases. This leads to the prediction that coexistence is most likely if STDs are mildly sterilizing, and OIDs are also of low or intermediate virulence.

The various ways in which the sterilizing effects of an STD influence STD–OID coexistence are particularly interesting. First, the sterilizing effect of the STD has no effect on its likelihood of invading an OID. However, because the equilibrium population size of a population with an STD is always lower when there is increased sterility, it becomes harder for an OID to displace an STD if the latter is sterilizing. Moreover, we have shown for STDs (Thrall *et al.* 1993) that host–pathogen coexistence is more likely if the disease sterilizes the host (assuming that density dependence acts more strongly on the healthy class). Therefore, sterility may not favour the spread of a sexual transmission mode, but it may stabilize such a relationship and ‘protect’ it from invasion by an alternative transmission mode.

Apart from these population dynamical consequences of sterility, there are likely to be other reasons to expect an association between sterility and sexual transmission. Sterility may be a consequence of disease lesions associated with the sexual organs, and it may be advantageous for the transmission of an STD by promoting continued sexual activity (especially in females, by preventing fertilization or inducing abortion). A comparison of STDs and OIDs in mammals (Lockhart *et al.* 1996) showed that STDs induced a greater amount of sterility, but this difference was not significant. A path analysis revealed no significant causative effect (and small path coefficients) of sexual transmission on either male or female sterility. However, within the class of diseases that were predominantly sexually transmitted, greater sexual transmission was associated with increased male and female sterility. Path analysis based on two causal models could not clearly differentiate whether greater sexual transmission caused increased sterility or vice versa. These empirical data support the theoretical expectation that the sterilizing effects of an STD may not affect its invasion into an OID, but that once the STD is established the degree to which it is sexually transmitted may be strongly influenced by sterility, perhaps for a variety of reasons.

Our results further suggest that increased birth rates, by increasing equilibrium host population size, would favour non-sexual transmission; this leads to the prediction that in small mammals (where per capita

birth rates and population densities are higher than in large mammals) incidence of OIDs relative to STDs should be correspondingly greater than in species with lower per capita birth rates. Results from our recent comparative study indicate that this may indeed be the case. For example, in an overall survey of the occurrence of STDs in wild and domesticated animals, we found far fewer reports of STDs in rats and mice than in larger mammals such as primates, or artiodactyls (Lockhart *et al.* 1996). A similar prediction was made by De Leo & Dobson (1996), who used allometric relationships of body size to longevity and population size in mammals to predict expected rates of pathogen increase for STDs and OIDs. STDs were predicted to be predominantly found in large mammals with small population sizes and long lifespans.

Our model illustrates that a disease that shows both a sexual and a non-sexual transmission mode may well be a mixture of two (or more) specialized strains coexisting in the same population. Moreover, it seems likely that evidence for such transmission polymorphisms in nature will be found, given the recent marked advances in DNA-based strain identification and strain specific epidemiological tracing (Small & Moss 1993; Versalovic *et al.* 1993; Delwart *et al.* 1995; O’Rourke *et al.* 1995). Many pathogen pairs that show both sexual and non-sexual transmission modes (genital and oral herpes; genital and cuticular warts; pubic and head lice; lymphogranuloma venereum and trachoma biovars) appear to consist of evolutionarily divergent lineages that now occupy different regions of the body. However, it is quite likely that a mutant to a new transmission mode would still show immunological cross-reactivity with the original strain; such cross-reactivity would then be lost as the ‘new’ disease diverges evolutionarily. Quite divergent strains or species of pathogens with different transmission modes still retain immunological cross-reactivity (Nahmias *et al.* 1990; Martin *et al.* 1993; Jimenez de Bagues *et al.* 1994*a, b*). Changes in transmission mode may well have involved a transitional stage of polymorphism, followed by subsequent specialization and niche divergence of the individual morphs. Our models have only considered the extreme case where strains are transmitted exclusively by either sexual or non-sexual means. Individual pathogen genotypes might also show different degrees of sexual transmission; this more complicated scenario is the subject of another study (Thrall & Antonovics 1997).

The present study shows that simple heuristic models incorporating different transmission modes can lead to a number of predictions regarding the maintenance of genetic polymorphism and the characteristics favouring the displacement of pathogens with one transmission mode by those with another. We have not been able to test these predictions rigorously, but have shown them to be consistent with data on diseases in natural and human populations. The rigorous testing of these ideas is likely to come with advances in the techniques for studying inter-strain variation and strain-specific epidemiology. We hope the present theoretical study will stimulate empirical work in these directions.

We thank Andy Dobson and Marc Lipsitch for much useful discussion and comments on many of the ideas contained in this manuscript. This paper was substantially improved by the comments of two anonymous referees. The support of National Science Foundation grant DEB-9408207 is gratefully acknowledged.

REFERENCES

- 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.
- Antonovics, J. 1994 The interplay of numerical and gene-frequency dynamics in host–pathogen systems. In *Ecological genetics* (ed. L. Real), pp. 146–170. Princeton University Press.
- Antonovics, J., Iwasa, Y. & Hassell, M. P. 1995 A generalized model of parasitoid, venereal, and vector-based transmission processes. *Am. Nat.* **145**, 661–675.
- Antonovics, J. & Thrall, P. H. 1995 The cost of resistance and the maintenance of genetic polymorphism in host–pathogen systems. *Proc. R. Soc. Lond. B* **257**, 105–110.
- Begon, M. & Bowers, R. G. 1995 Beyond host–pathogen dynamics. In *Ecology of infectious diseases in natural populations* (ed. B. T. Grenfell & A. P. Dobson), pp. 478–509. Cambridge University Press.
- Bernard, H. U., Chan, S. Y. & Delius, H. 1994 Evolution of papillomaviruses. *Curr. Top. Microbiol. Immunol.* **86**, 34–53.
- de Jong, M. C. M., Diekmann, O. & Heesterbeek, H. 1994 How does transmission of infection depend on population size? In *Epidemic models: their structure and relation to data* (ed. D. Mollison), pp. 84–94. Cambridge University Press.
- De Leo, G. A. & Dobson, A. P. 1996 Allometry and simple epidemic models for microparasites. *Nature, Lond.* **379**, 720–722.
- Delwart, E. L., Busch, M. P., Kalish, M. L., Mosley, J. W. & Mullins, J. I. 1995 Rapid molecular immunology of human immunodeficiency virus transmission. *AIDS Res. Hum. Retroviruses* **11**, 1081–1093.
- Ewald, P. W. 1994 *Evolution of infectious disease*. Oxford University Press.
- Gentry, G. A., Lowe, M., Alford, G. & Nevins, R. 1988 Sequence analyses of herpesviral enzymes suggest an ancient origin for human sexual behavior. *Proc. Natn. Acad. Sci. USA* **85**, 2658–2661.
- Getz, W. M. & Pickering, J. 1983 Epidemic models: thresholds and population regulation. *Am. Nat.* **121**, 892–898.
- Hassell, M. P. 1978 *The dynamics of arthropod predator–prey systems*. Princeton University Press.
- Jimenez de Bagues, M. P., Elzer, P. H., Blasco, J. M., Marin, C. M., Gamazo, C. & Winter, A. J. 1994a Protective immunity to *Brucella ovis* in BALB/c mice following recovery from primary infection or immunization with subcellular vaccines. *Infect. Immun.* **62**, 632–638.
- Jimenez de Bagues, M. P., Elzer, P. H., Jones, S. M., Blasco, J. M., Enright, F. M., Schurig, G. G. & Winter, A. J. 1994b Vaccination with *Brucella abortus* rough mutant RB51 protects BALB/c mice against virulent strains of *Brucella abortus*, *Brucella melitensis*, and *Brucella ovis*. *Infect. Immun.* **62**, 4990–4996.
- Lipsitch, M. & Nowak, M. A. 1995 The evolution of virulence in sexually transmitted HIV/AIDS. *J. Theor. Biol.* **174**, 427–440.
- Lockhart, A. B., Thrall, P. H. & Antonovics, J. 1996 Sexually transmitted diseases in animals: ecological and evolutionary implications. *Biol. Rev. Cambridge Phil. Soc.* **71**, 415–471.
- Martin, P. M., Cockayne, A., Georges, A. J., & Penn C. W. 1990 Immune response to *Treponema pertense* and *Treponema pallidum* Nichols in patients with yaws. *Res. Microbiol.* **141**, 181–186.
- McGeoch, D. J. & Cook, S. 1994 Molecular phylogeny of the Alphaherpesvirinae subfamily and a proposed evolutionary time scale. *J. Mol. Biol.* **238**, 9–22.
- 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.
- Nahmias, A. J., Lee, F. K., & Beckman-Nahmias, S. 1990 Sero-epidemiological and -sociological patterns of herpes simplex virus infection in the world. *Scand. J. Infect. Dis. Suppl.* **69**, 19–36.
- Nielsen, K. & Duncan, J. R. 1990 *Animal brucellosis*. Boca Raton, Florida: CRC Press.
- Norris, S. J. 1993 Polypeptides of *Treponema pallidum*: progress toward understanding their structural, functional, and immunologic roles. *Microbiol. Rev.* **57**, 750–779.
- O'Rourke, M., Ison, C. A., Renton, A. M. & Spratt, B. G. 1995 Opa-typing: a high resolution tool for studying the epidemiology of gonorrhoea. *Mol. Microbiol.* **17**, 865–875.
- Salazar, A. P. R. 1994 Filogenia de los piojos (Insecta: Anoplura) de los monos del Viejo Mundo (Catarrhini). *Rev. Biol. Trop.* **42**, 633–638.
- Schachter, J. 1990 Biology of *Chlamydia trachomatis*. In *Sexually transmitted diseases*, 2nd edn (ed. K. K. Holmes, P. A. Mardh, P. F. Sparling, P. J. Wiesner, W. Cates, S. M. Lemon & W. E. Stamm), pp. 167–180. New York: McGraw-Hill.
- Small, P. M. & Moss, A. 1993 Molecular epidemiology and the new tuberculosis. *Infect. Agents Dis.* **2**, 132–138.
- Smith, G. & Dobson, A. P. 1992 Sexually transmitted diseases in animals. *Parasitol. Today* **8**, 159–166.
- Thrall, P. H., Antonovics, J. & Hall, D. W. 1993 Host and pathogen coexistence in vector-borne and venereal diseases characterized by frequency-dependent disease transmission. *Am. Nat.* **142**, 543–552.
- Thrall, P. H., Biere, A. & Uyenoyama, M. K. 1995 Frequency-dependent disease transmission and the dynamics of the *Silene–Ustilago* host–pathogen system. *Am. Nat.* **145**, 43–62.
- Thrall, P. H. & Antonovics, J. 1997 Allocation to sexual vs non-sexual disease transmission. *Am. Nat.* (In the press.)
- Versalovic, J., Woods, C. R., Georgioui, P. R., Hamill, R. J. & Lupski, J. R. 1993 DNA-based identification and epidemiological typing of bacterial pathogens. *Arch. Pathol. Lab. Med.* **117**, 1088–1098.
- Zur Hausen, H. & de Villiers, E. M. 1994 Human papillomaviruses. *A. Rev. Microbiol.* **48**, 427–447.

Received 18 November 1996; accepted 13 December 1996