Host-Pathogen Interactions, Insect Outbreaks, and Natural Selection for Disease Resistance

Bret D. Elderd,[1,*] Jonathan Dushoff,[2] and Greg Dwyer[1,†]

1. Department of Ecology and Evolution, University of Chicago, Chicago, Illinois 60637;
2. Department of Biology, McMaster University, West Hamilton, Ontario L8S 4K1, Canada

Submitted August 7, 2007; Accepted July 16, 2008; Electronically published October 31, 2008

Online enhancement: appendix.

abstract: The theory of insect population dynamics has shown that heterogeneity in natural-enemy attack rates is strongly stabilizing. We tested the usefulness of this theory for outbreaking insects,

infectious cadavers, respectively, so that transmission occurs with rate parameter . After , time units, infected larvae die and become infectious cadavers at time *t*, releasing occlusion bodies onto the foliage to complete the cycle of transmission (Cory and Myers 2003). Occlusion bodies are rendered inactive by sunlight and other environmental factors at a rate .

A simple way to allow for heterogeneity is to assume that attack rates follow a probability distribution (Hassell et al. 1991). In epidemic models, the transmission parameter is effectively the natural-enemy attack rate, and so we allow for a probability distribution of values. Variability in appears to be due to variability among hosts (Dwyer et al. 1997), and so we assume in particular that there is a distribution of across hosts. This leads to the model

$$\frac{dS}{dt} = -\left[\frac{S(t)}{S(0)}\right]^{C^2} SP, \tag{3}$$

$$\frac{dP}{dt} = \left[1-\left[\frac{S(t-\tau)}{S(0)}\right]^{C^2}\right] P(t-\tau)S(t-\tau) - \gamma P, \tag{4}$$

where $\bar{\beta}$ and *C* are the average and the CV, respectively, of the distribution of , so that *C* is a measure of heterogeneity in the attack rate. We have thus substituted C^2 for CV^2, to avoid the confusion that arises from using multiple letters to refer to a single quantity. This model provides a much better description of NPV epidemics than equations (1) and (2), suggesting that heterogeneity in infection risk plays an important role in NPV transmission (Dwyer et al. 1997, 2002). To show how high heterogeneity in the attack rate produces a stable equilibrium, we extend this model to allow for multiple generations.

An important point is that the risk that insects become infected with an NPV is usually measured by exposing groups of larvae to a known dose of the virus and then observing the fraction that become infected (Watanabe 1987). Because larvae that do not consume the entire dose are discarded, such experiments effectively measure the risk of infection given exposure, which is usually thought of as susceptibility. In contrast, the transmission parameter allows for not just the risk of infection given exposure but also the risk of exposure. Variability in both types of risk can affect the overall risk that gypsy moth larvae become infected with the NPV (Dwyer et al. 2005), and so heterogeneity in provides a useful measure of variability in overall infection risk.

In previous work, Dwyer et al. (1997) attempted to measure heterogeneity in for the gypsy moth virus. The confidence intervals on the resulting estimates, however, were so large that it was impossible to determine which was more likely, a stable equilibrium or stable cycles (Dwyer et al. 2000). By refin-455.2f0).mj/F5 30.s andcome infet al0 2-1 ri- 0 /F11 11f1.05 0 TD()T7/F5 1 Tf0.5449 0 TDP

Figure 1: Dynamics of standard insect-pathogen models, which assume that infection risk is constant. *A* and *B* show the dynamics of equations (5)–(7). In *A*, host reproduction p 5.5, pathogen between-generation impact p 35, and heterogeneity in susceptibility *C* p 0.86, leading to stable cycles. The parameters in *B* are the same, except that *C* p 1.03, leading to damped oscillations and thus a stable equilibrium. *C* and *D* similarly show the effects of increasing heterogeneity for the model that incorporates stochasticity and a generalist predator by substituting equation (A21) in the online edition of the *American Naturalist* for equation (6). Parameters are the same for *C* and *D* (p 74.6, p 60, maximum predation rate *a* p 0.96, and the density at which maximum predation occurs *b* p 0.14), except that in *C*, *C* p 0.96, while in *D*, *C* p 1.12. For this model, *C* ! 1 leads to irregular cycles with a large amplitude, but *C* 1 1 produces only small-amplitude fluctuations about the equilibrium.

$$N_{t+1} = N_t[1 - i(N_t, Z_t)], \tag{6}$$

$$Z_{t+1} = fN_t i(N_t, Z_t) + Z_t. \tag{7}$$

Here f and are the survival rates of pathogen particles produced in the most recent epidemic and in previous epidemics, respectively. We allow for the possibility of differences in the two survival rates because particles produced in the current epidemic are likely to have a greater chance of infecting larvae in the following generation (Murray and Elkinton 1989, 1990). Note that, in contrast, Anderson and May's (1980) well-known insect-pathogen model unrealistically assumes overlapping generations, which is why that model shows stable cycles even in the absence of heterogeneity in attack rates.

Although we tested the model using the gypsy moth NPV, in fact the model applies to many different insect-pathogen interactions. To illustrate this, we note that the basic assumptions of the model are that the pathogen is directly transmitted and fatal, that it affects only juveniles, and that it must survive in the environment between generations. Directly transmitted, fatal diseases that infect only larvae occur in a large number of insects. Pathogens with specialized stages that allow survival between host generations are also common among insects and include viruses, fungi, and protozoa (Fuxa and Tanada 1987; Miller 1997).

Among forest defoliators in particular, there are many species for which epidemics affecting the larval stages cause outbreaks to collapse (Myers 1993; Moreau and Lucarotti

Figure 2: Transmission $\bar{\beta}$ and heterogeneity in susceptibility C for each population in our transmission experiments. Populations are arranged in order of decreasing values of $\bar{\beta}$. This is a box-and-whisker plot, in which the boxes show the interquartile range, while the whiskers show the range of data values included within $1.5\times$ the interquartile range. Values of heterogeneity C were thus usually 11, but values of transmission $\bar{\beta}$ varied greatly.

and their NPVs. By using a general model, we are thus

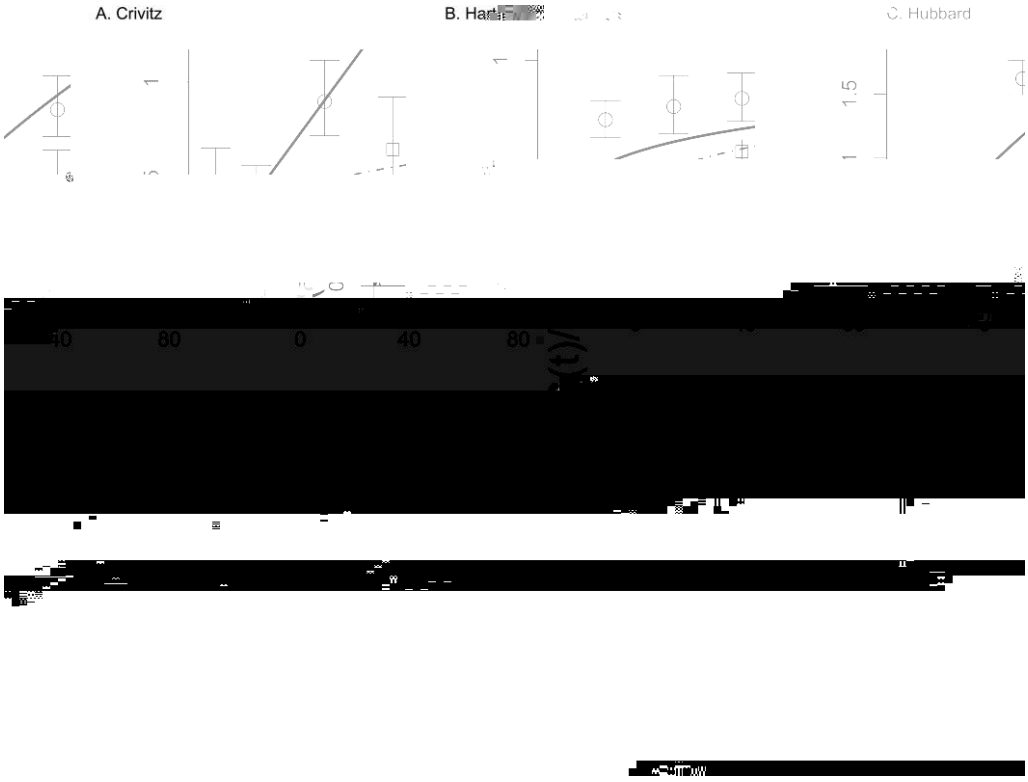

Figure 3: Experimental measurements of infection rates before and after population crashes in five gypsy moth populations. Points represent data, and lines represent the best fit of equation (8) in each population in each year. Here we plot the log-transformed fraction uninfected, $-\ln(S(T)/S(0))$, which makes it easier to detect changes in transmission parameters. Although the results vary across populations, infection risk was generally lower after crashes. Note that changes in the shape of the best-fit model are due to changes in heterogeneity C (Dwyer et al. 1997).

and Elkinton 1993; Dwyer et al. 1997, 2005; D'Amico et al. 1998). Like dose-response experiments, these experiments include a range of viral doses, but they also allow for realistic feeding behavior, and they can be used to estimate $\bar{\beta}$ and C. Indeed, estimates of $\bar{\beta}$ and C from these experiments produce model predictions that are close to the dynamics of natural epidemics (Dwyer et al. 2002), suggesting that transmission rates in our experiments are not that different from those in nature.

To reduce the uncertainty in our estimates of $\bar{\beta}$ and C, we followed Dwyer et al. (2005) in instituting several improvements over previous experiments (Dwyer et al. 1997). The most important improvement is that we used larvae that had all reached the fourth instar, or stage, within 24 h, thereby avoiding the changes in susceptibility that occur within the first few days of the instar (Grove and Hoover 2007). Given this change, we again allowed the initially uninfected larvae to feed for a week in the field, and then we reared them in the lab to see which of them had become infected while in the field. To estimate $\bar{\beta}$ and C, we fitted the epidemic model to the resulting data. Because in our experiments branches were enclosed in mesh bags that prevent the breakdown of the virus (G. Dwyer, unpublished data), we can set $dP/dt = 0$ in equations (3) and (4), which allows us to solve equation (3) as follows (Dwyer et al. 1997):

$$\frac{S(T)}{S(0)} = (1 - \bar{\beta}C^2P(0)T)^{-1/C^2}. \tag{8}$$

Here T is the length of time for which an experiment ran, so that $S(T)$ and $S(0)$ are the densities of uninfected larvae at the end and the beginning of the experiment, respectively, and $P(0)$ is the initial density of virus, in the form of infectious cadavers; $S(T)/S(0)$ is thus the fraction uninfected at the end of the experiment, and $\bar{\beta}$ and C can then be estimated by fitting equation (8) to the data, using maximum likelihood (Pawitan 2001) and nonlinear fitting routines (Venables et al. 2005; also see appendix).

To assess variability over space and time, we used test larvae reared from egg masses collected over 4 years across a wide area of the northeastern and midwestern United

Figure 4: Consequences of the changes in infection rates shown in figure 3 for infection risk in a full epidemic, $i(N_t, Z_t)$, as calculated from equation (5). Because changes in heterogeneity C modulate the effects of density, we considered a range of initial host densities N_t, but for simplicity, we assumed a constant initial infection rate of 5% (Woods and Elkinton 1987). In general, infection risk was reduced in most populations at most densities, suggesting that infection risk is reduced after population crashes.

States (table A1 in the online edition of the *American Naturalist*). Egg masses were collected from areas of a few square kilometers or less within larger forests of tens to hundreds of square kilometers. Each collection site, and thus each population, was at least 50 km from all other

Figure 5:

Figure 6: Dynamics of host-pathogen models in which natural selection drives fluctuations in infection risk. *A* shows host and pathogen densities for the host-pathogen-only model, equations (9)–(12), while *B* shows densities for the corresponding host-pathogen-predator model, which includes stochasticity and a generalist predator by substituting equation (A22) in the online edition of the *American Naturalist* for equation (9). *C* and *D* show the corresponding fluctuations in the average transmission rate for each model to illustrate that average transmission rises and falls in synchrony with host density. In *A* and *C*, baseline reproduction $r = 0.2$, rate of increase of reproduction with increasing susceptibility $= 9$, pathogen long-term survival $= 0.2$, heterogeneity in susceptibility $C = 2$, and pathogen between-generation impact $= 14$. In *B* and *D*, the parameters are the same as in *A* and *C*, with the additions that the maximum fraction of prey consumed $a = 0.967$ and the density at which predation is maximized $b = 0.14$. Note that *B* and *D* show irregular fluctuations, as in insect outbreaks in nature (Dwyer et al. 2004).

susceptibility and reduced digestive efficiency and thus to a cost of resistance. Second, gypsy moth larvae in particular are known to avoid anything on a leaf surface that has the consistency of a virus-infected cadaver, including not just cadavers but also molasses (Capinera et al. 1976). Ongoing work in G. Dwyer's lab has suggested that this trait may cause larvae to reject leaf tissue simply because of leaf scars (L. Eakin and G. Dwyer, unpublished data). Increases in the sensitivity of the trait might therefore reduce both infection risk and feeding efficiency, thereby leading to reduced egg mass size and thus a cost of resistance.

In short, it seems reasonable to allow for a cost of resistance. Also, for simplicity, we assume that heterogeneity C is constant and that offspring have the same phenotype as their parents, so that we ignore sexual reproduction. These assumptions lead to the following model (see appendix):

$$N_{t+1} = \tag{9}$$

$$N_t[1 - i(N_t, Z_t, \bar{\ }_t)]\{r - \bar{\ }_t[1 - i(N_t, Z_t, \bar{\ }_t)]^{C^2}\},$$

$$Z_{t+1} = fN_t i(N_t, Z_t, \bar{\ }_t) + \ Z_t, \tag{10}$$

$$\bar{\ }_{t+1} = \tag{11}$$

$$\frac{r\bar{\ }_t[1 - i(N_t, Z_t, \bar{\ }_t)]^{C^2} + \ (C^2 + 1)\bar{\ }_t^2[1 - i(N_t, Z_t, \bar{\ }_t)]^{2C^2}}{}$$

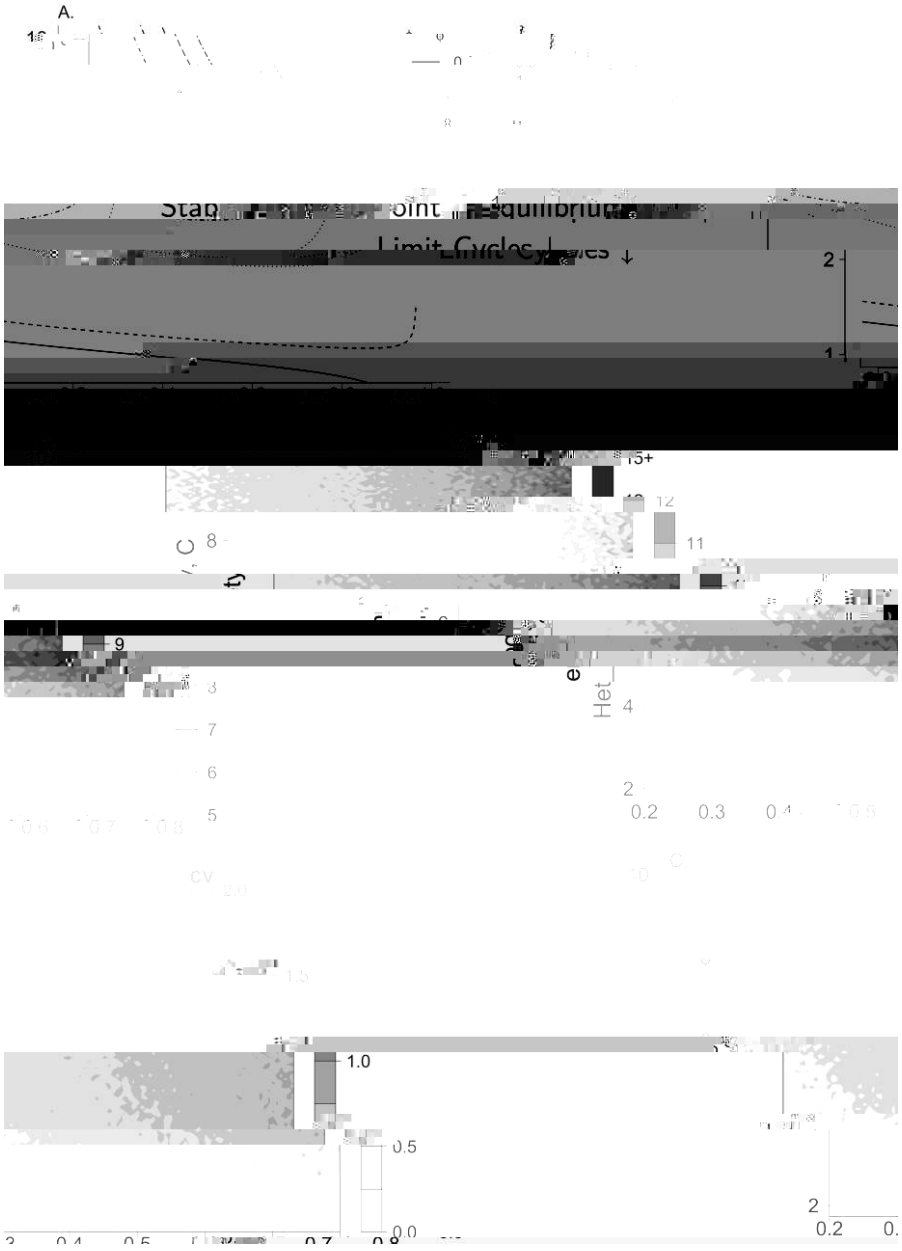

Figure 7: Range of parameter values for which cycles occur for the models in which infection risk is affected by natural selection. *A*, Host-pathogen model, equations (9)–(12). Each line represents the boundary between cycles and stability for different values of pathogen between-generation impact , such that limit cycles occur for values of heterogeneity *C* below each line. There is thus a large region of parameter space for which cycles occur, even for $C \leq 1$. *B*, Time between outbreaks, averaged over 100 realizations, for the host-pathogen-predator model with natural selection, equations (10)–(12) and (A22) in the online edition of the *American Naturalist*, with $p = 5$. *C*, Coefficient of variation (CV) of time between outbreaks for the host-pathogen-predator model with natural selection and $p = 5$. For gypsy moth outbreaks in nature, the average time between outbreaks is 6–10 years (Johnson et al. 2005), with CV values between 0.2 and 0.7 (Dwyer et al. 2004). The host-pathogen-predator model thus produces realistic cycles for a wide range of parameter values.

of individuals with transmission parameter . . The key difference from equations (5)–(7) is thus that average transmission is now a dynamic variable, $\bar{.}_t$.

For this model, we require first that $r ! 1$ because, otherwise, completely resistant hosts increase without bound (appendix). Long-period, large-amplitude cycles then occur even for $C 1 1$ (fig. 6). As figure 6 shows, in this model repeated epidemics cause average transmission $\bar{.}_t$ to drop sharply after outbreaks. The fecundity cost of resistance then causes transmission to slowly rise between outbreaks, until rising virus levels again lead to strong selection for resistance. Cycles therefore occur even when C is high because changes in average transmission reduce the stabilizing effect of highly resistant individuals.

This model thus reconciles our experimental data with the occurrence of cycles. More quantitatively, our estimates of heterogeneity C produce cycles in the evolutionary model for a wide range of values of the other parameters (fig. 7*A*). Moreover, if we allow for both pr[illegible]on a[illegible] stochasticity by substituting equation (A22) in the online edition of the *American Naturalist* for equation (9), the average (fig. 7*B*) and the CV (fig. 7*C*) of the time between outbreaks are close to data from real populations (Dwyer et al. 2004; Johnson et al. 2005).

The model is also consistent with the declines in infection risk that we observed after epidemics (fig. 3). As our AIC analyses make clear, however (table A3), the data suggest that both $\bar{.}$ and C change over time, even though the model allows for changes only in $\bar{.}$. Because allowing C to change over generations leads to an extremely complicated model, in comparing the model to the data we instead reanalyze the data, assuming that C is constant but that $\bar{.}$ changes over time. Table A5 in the online edition of the *American Naturalist* then shows that, in four cases out of five, average transmission declined over time, as predicted by the model. Although there were clearly sources of variability that acted on the data that are not accounted for by the model, the model with natural selection is nevertheless best able to reconcile our experimental data with the occurrence of outbreaks.

Testing Whether Infection Risk Is Heritable

Methods. A crucial untested assumption of the evolutionary models is that infection risk is heritable. In 2006 and 2007, we therefore carried out additional experiments to test for family effects on infection risk and thus to test for heritable variation. To do this, we reared larvae from individual egg masses in family groups, using full-sibling groups in two experiments and half-sibling groups in a third experiment. The distinction here is that in our experiments from 2000 to 2003, larvae from a given population were reared from 25–50 egg masses mixed together, whereas in our 2006 and 2007 experiments, larvae from a given egg mass were reared separately from larvae from other egg masses. Mixing egg masses provided more larvae and thus more replicates per population, which in turn allowed more accurate estimates of heterogeneity C. In contrast, rearing larvae in family groups from individual egg masses did not permit us to estimate parameters as accurately, but it allowed us to test directly for effects of family on virus transmission.

Figure 8: Results of transmission experiments with full- or half-sibling larvae. *A* shows data for full-sibling larvae from 2006. Each panel shows infection rates for a different family, with points representing the data and lines representing the best-fit version of equation (8). *B* shows data for full-sibling larvae in 2007, and *C* shows data for half-sibling larvae in 2007. In 2007, we used only one virus density, and so each point in *B* and *C* represents data for a different group. On the horizontal axis in *C*, at each tick mark, the number before the colon is the sire number, while the number after the colon is the sire-specific dam number; 12:2 thus refers to larvae produced by the second dam that was mated to the twelfth sire. All dams were mated to only one sire. For logistic reasons, the number of groups varied among experiments, but in all cases each treatment was replicated eight times.