vol. 175, no. 6 the american naturalist june 2010

E-Note

The Effect of Host Genetic Diversity on Disease Spread

Curtis M. Lively

Department of Biology, Indiana University, Bloomington, Indiana 47405

Submitted November 6, 2009; Accepted December 1, 2009; Electronically published April 13, 2010

abstract: Host genetic diversity is thought to reduce the likelihood that disease will spread in natural populations. In this study, I present an epidemiological model for the intrinsic rate of spread (R_0) for an infectious disease. The results show that the average value for R_0 ($\overline{R}_0$) is inversely related to the number of host genotypes in the population (G), assuming that each host genotype is susceptible to a different parasite genotype. Specifically, for large host populations, $\overline{R}_0$ is equal to B/G, where B is the number of infectious propagules produced by each infection that contact a different host. The results also suggest that virulent, single-strain infections, which initially spread in genetically diverse host populations, would quickly die out when the parasite depresses the frequency of susceptible hosts below $1/B$. These results are consistent with empirical studies showing that genetically diverse host populations suffer less from pathogens and parasites.

Keywords: disease spread, epidemics, genetic diversity, R_0.

Introduction

Genetic diversity for disease resistance in host populations is generally thought to reduce the risk of spread for infectious diseases. This idea is supported by empirical studies from both plant and animal hosts for a wide variety of diseases (Dwyer et al. 1997; Baer and Schmid-Hempel 1999; Zhu et al. 2000; Altermatt and Ebert 2008), but it has not received much in the way of direct theoretical analysis. The models that do exist tend to suggest that variation in host susceptibility does not affect the emergence of infectious diseases (Springbett et al. 2003; Yates et al. 2006), although it might reduce the severity of epidemics when they do occur (Springbett et al. 2003). In these models, the variation in susceptibility among hosts was caused by differences in resistance to a single strain of pathogen, but none of the host genotypes were completely resistant to infection.

In this study, I examine the effect of host genetic diversity on disease dynamics, while assuming that hosts are resistant to a majority of parasite genotypes. This resistance is based on a self-nonself recognition system, which allows hosts to detect and kill parasites that do not (at least at some level) match the characteristics of the host's cell-surface molecules. More generally, the model assumes that there is a tight genetic specificity for infection such that parasite genotypes can infect only a subset of host genotypes, as is directly shown in the study by Carius et al. (2001) for *Daphnia* and indirectly shown in the study by Dybdahl et al. (2008) for snails. The model suggests that the likelihood of disease spread is inversely proportional to the number of host genotypes for parasite resistance.

Model

Consider an annual host that is infected by an annual parasite. Successful infections produce spores that make contact with juvenile hosts in the next generation. I assume that each host resistance genotype can be infected by only one parasite genotype. Hence, there is no unconditional resistance to infection; hosts are resistant to and kill parasite genotypes that do not match the host's genotype. This model represents the matching-alleles (or matching-genotypes) model for infection (Frank 1993; Otto and Michalakis 1998; Agrawal and Lively 2002). The qualitative results presented here should be robust to the exact genetic interface for infection/resistance, as long as some degree of genetic specificity is required for infection (Agrawal and Lively 2002; Engelstadter and Bonhoeffer 2009).

The number of infected hosts having the ith genotype at time $t + 1$ can be written as

$$I_{i(t+1)} = g_{i(t+1)} N_{t+1} P_{i(t+1)}, \tag{1}$$

where $g_{i(t+1)}$ gives the frequency of the ith host genotype at $t + 1$, N_{t+1} gives the total number of hosts at time $t + 1$, and $P_{i(t+1)}$ gives the probability of infection for the ith host genotype at time $t + 1$. The probability of infection can be estimated as 1 minus the zero class in a Poisson distribution, where the zero class is $\exp(-\lambda)$ and λ is the mean number of "matching" spores that contact each host. Thus, the probability of infection for the ith host genotype at time $t + 1$ is

$$P_{i(t+1)} = 1 - e^{-\lambda} = 1 - e^{-BI_{i(t)}/N_{t+1}}, \tag{2}$$

Am. Nat. 2010. Vol. 175, pp. E149–E152. © 2010 by The University of Chicago. 0003-0147/2010/17506-51719$15.00. All rights reserved.
DOI: 10.1086/652430

where $I_{i(t)}$ is the number of infected hosts having the ith genotype at time t, and B is the number of parasite propagules produced by each infection that make contact with different hosts. If, for example, an infection generates 10,000 infectious propagules, but only 10 of these propagules come into contact with different hosts, then $B = 10$. As such, B places an upper limit on the number of secondary infections that can be generated. Assuming that a single infected individual of genotype i is introduced into the host population at time t (i.e., $I_{i(t)} = 1$), the number of secondary infections in the next time step (R_0) is

$$R_{0i} = g_{i(t+1)} N_{t+1}(1 - e^{-B/N_{t+1}}). \tag{3}$$

Hence, the spread of infection (which requires $R_0 > 1$) requires that

$$g_{i(t+1)} N_{t+1} > \frac{1}{1 - e^{-B/N_{t+1}}}. \tag{4}$$

Note that the condition for the spread of infection is both frequency dependent and density dependent. Increasing the genetic diversity in the host population (thereby reducing the frequency of each host genotype) would increase the threshold density required for the spread of infection. However, if the disease does spread and the infection reduces host fitness, then $g_{i(t+1)}$ would be expected to diminish over time, which would increase the threshold density required for persistence of the infection. Most likely, then, the disease will die out, as the condition suggested by equation is no longer met. Hence, even if diseases do initially spread in small host populations, they may not persist in genetically diverse populations.

The condition for the spread of infection is greatly simplified in large host populations. By taking the limit for equation (3) as N goes to infinity, we get

$$R_{0i} = g_i B. \tag{5}$$

Thus, in large populations, the initial spread of infection requires that $B > 1/g_i$. If the single-strain infection reduces the fitness and, hence, the frequency of the susceptible host genotype over time, then the infection will die out when g_i becomes less than $1/B$. These expectations were verified by computer simulation (results not shown).

Suppose, however, that the pathogen is introduced by migration at a high rate. In this case, it is possible to have multiple coexisting strains of the parasite. The mean value for R_0 is then

$$\overline{R}_0 = \frac{\sum R_{0i}}{G} = \frac{N_{t+1}(1 - e^{-B/N_{t+1}})}{G}, \tag{6}$$

where G is the number of genotypes in the host population. Thus, the spread and persistence of infection depends on host population size and the number of genotypes for resistance (fig. 1). For large host populations, the result simplifies to

$$\overline{R}_0 = \frac{B}{G}. \tag{7}$$

Thus, all else being equal, the spread and persistence of infection should more easily occur in genetically homogeneous populations.

Discussion

Host heterogeneities of different kinds are known to affect the likelihood of disease spread (e.g., Hawkins et al. 1993), but there has been a paucity of theoretical work that specifically examines the effect of genetic diversity in the host population. In one direct test, results suggested that host genetic diversity would not affect R_0 but it would affect the variance in this parameter and could reduce the severity of epidemics when they occur (Springbett et al. 2003). In that study, host diversity was the quantitative genetic variation among host genotypes in susceptibility to infection by a single parasite genotype. Similar results were obtained by Nath et al. (2008), using a two-locus, two-allele model where one locus controlled susceptibility and the other locus controlled the recovery period. Finally, Yates et al. (2006) found that variation in susceptibility

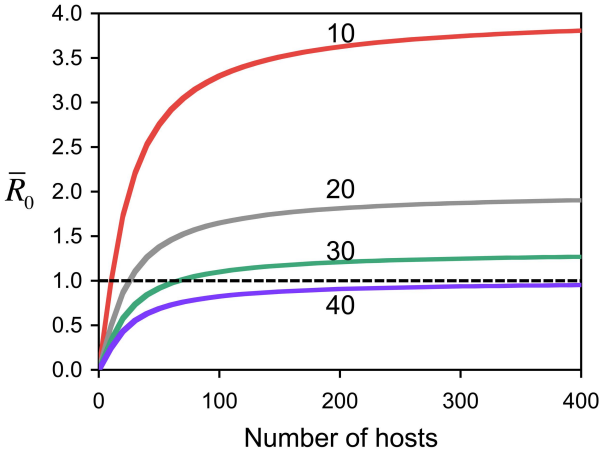

Figure 1: Effect of genetic diversity on disease spread. The numbers associated with each line gives the number of genotypes (G) in the host population. The dashed line gives $\overline{R}_0 = 1$. Disease spreads for $\overline{R}_0$ values above the dashed line, assuming $B = 40$. Note that the lines are asymptotic on B/G.

did not affect disease spread in a model where 10% of the population was 100 times more susceptible to infection than was the remaining fraction of the population.

The results of my model suggest that host genetic diversity could reduce R_0, but they also make different assumptions about the basis for the underlying genetic diversity in the host population. Specifically, this model assumes that the different host genotypes are susceptible to only one of the different parasite genotypes and are resistant to the remaining parasite genotypes. These are the basic assumptions underlying the matching-alleles model for infection, which stands as a good framework for studying self-nonself recognition systems in animals and which has some experimental support in invertebrates (Carius et al. 2001; Dybdahl et al. 2008). The results suggest that, under these simplifying assumptions, R_0 averaged over the different parasite strains is inversely proportional to the number of host genotypes (G) in the population. More specifically, $\overline{R}_0$ is equal to B/G, where B is the number of propagules produced by each infection that make contact with a different host (where a single contact is sufficient to produce infection in the matching, susceptible host genotypes). The results also suggest that virulent (fitness-reducing) single-strain infections, which would initially spread in genetically diverse host populations, would die out as they depress the frequency of the matching host genotype below $1/B$. These results are consistent with observations that genetically diverse populations have lower incidences of disease (Dwyer et al. 1997; Baer and Schmid-Hempel 1999; Zhu et al. 2000; Altermatt and Ebert 2008), but the exact form of the relationship between genetic diversity and disease risk has yet to be empirically examined.

Why does host genetic diversity matter to disease spread? In this model, parasite propagules are distributed at random among G different host genotypes. In an infinite host population, the probability of successful infection per propagule is therefore equal to the frequency of the host genotype (g_i) that is susceptible to the ith parasite genotype. Thus, R_0 is the product of g_i and B, which simply gives the number of secondary infections. (Note that the number of secondary infections is constrained to be less than or equal to B instead of the total number of susceptible hosts as in most SIR [susceptible-infectious-recovered] models [see the appendix].) Increasing genetic diversity would tend to reduce g_i and thus reduce the risk of disease spread. Alternatively, if the disease does spread and is virulent enough to reduce g_i to less than $1/B$, it will tend to die out locally. It would therefore seem, in general, that increasing host genetic diversity would reduce infection, provided that the different host genotypes are susceptible to different subsets of parasite genotypes.

Acknowledgments

I thank S. Bishop, L. Delph, D. Ebert, K. King, M. Neiman, and M. Wade for comments on the manuscript. This work was supported by the U.S. National Science Foundation (grant DEB-0640639).

APPENDIX

SIR model

Here I ask whether the main results of this article can be derived using the standard SIR (susceptible-infectious-recovered) model (Anderson and May 1991). Let I_i be the number of infected hosts having genotype i, let S_i be the number of susceptible hosts having genotype i, and let R_i be the number of recovered hosts having genotype i. The variable β is the probability of contact (c) between susceptible and infected individuals times the probability of disease transmission (a; see Otto and Day 2007), and v is the rate of host recovery from infection. The rate of change for the infected portion of the population is

$$\frac{dI_i}{dt} = \beta I_i S_i - v I_i.$$

Let $S_i = g_i N$ to get

$$\frac{dI}{dt} = \beta I_i g_i N - v I_i.$$

The number of secondary infections is then

$$R_0 = \beta g_i N - v.$$

The disease will spread if $g_i N > v/\beta$. All else being equal, increasing genetic diversity will reduce the risk of disease spread (see the main text) because it increases the threshold density for epidemics. The result, however, is different for large host populations. By taking the limit for R_0 as N goes to infinity, we find that the disease will spread for all values of g_i greater than 0. Thus, in very large host populations, host genetic diversity by itself does not affect whether the disease will spread.

Literature Cited

Agrawal, A. F., and C. M. Lively. 2002. Infection genetics: gene-for-gene versus matching-allele models, and all points in between. Evolutionary Ecology Research 4:79–90.

Altermatt, F., and D. Ebert. 2008. Genetic diversity of *Daphnia magna* populations enhances resistance to parasites. Ecology Letters 11: 918–928.

Anderson, R. M., and R. M. May. 1991. Infectious diseases of humans: dynamics and control. Oxford University Press, Oxford.

Baer, B., and P. Schmid-Hempel. 1999. Experimental variation in polyandry affects parasite loads and fitness in a bumble-bee. Nature 397:161–164.

Carius, H.-J., T. J. Little, and D. Ebert. 2001. Genetic variation in a host-parasite association: potential for coevolution and frequency-dependent selection. Evolution 55:1136–1145.

Dwyer, G., J. S. Elkinton, and J. P. Buonaccorsi. 1997. Host heterogeneity in susceptibility and disease dynamics: tests of a mathematical model. American Naturalist 150:685–707.

Dybdahl, M., J. Jokela, L. F. Delph, B. Koskella, and C. M. Lively. 2008. Hybrid fitness in a locally adapted parasite. American Naturalist 172:772–782.

Engelstädter, J., and S. Bonhoeffer. 2009. Red Queen dynamics with non-standard fitness interactions. PLoS Computational Biology 5: e1000469.

Frank, S. A. 1993. Specificity versus detectable polymorphism in host-parasite genetics. Proceedings of the Royal Society B: Biological Sciences 254:191–197.

Hawkins, B., M. Thomas, and M. E. Hochberg. 1993. Refuge theory and biological control. Science 262:1429–1432.

Nath, M., J. A. Woolliams, and S. C. Bishop. 2008. Assessment of the dynamics of microparasite infections in genetically homogeneous and heterogeneous populations using a stochastic epidemic model. Journal of Animal Science 86:1747–1757.

Otto, S. P., and T. Day. 2007. A biologist's guide to mathematical modeling in ecology and evolution. Princeton University Press, Princeton, NJ.

Otto, S. P., and Y. Michalakis. 1998. The evolution of recombination in changing environments. Trends in Ecology & Evolution 13:145–151.

Springbett, A. J., K. MacKenzie, J. A. Woolliams, and S. C. Bishop. 2003. The contribution of genetic diversity to the spread of infectious diseases in livestock populations. Genetics 165:1465–1474.

Yates, A., R. Antia, and R. R. Regoes. 2006. How do pathogen evolution and host heterogeneity interact in disease emergence? Proceedings of the Royal Society B: Biological Sciences 273:3075–3083.

Zhu, Y., H. Chen, J. Fan, Y. Wang, Y. Li, J. Chen, S. Yang, et al. 2000. Genetic diversity and disease control in rice. Nature 406:718–722.

Associate Editor: Yannis Michalakis
Editor: Ruth G. Shaw