Letter to the Editor

Selection Pressures and Plant Pathogens: Robustness of the Model

Richard A. Fleming

Graduate student, Department of Zoology, The University of British Columbia, Vancouver, Canada, V6T 1W5.

Leonard (4) developed a mathematical model to examine the dynamics of a gene-for-gene relationship. He derived algebraic expressions for the frequency of the susceptibility gene in the host population and the frequency of the virulence gene in the parasite population at the single nontrivial equilibrium point. At this point, the two alleles of each population are simultaneously in equilibrium and at nonzero frequencies. By computer simulation, Leonard (4) found that as time progresses in his model the pair of gene frequencies departed from each set of nonzero initial values he used and spiralled into the nontrivial equilibrium point when stabilizing selection occurred. This suggested that his model may be globally asymptotically stable (Fig. 1A).

Subsequently Sedcole (6) offered computer simulation and analytic arguments claiming that Leonard's model is inherently unstable. He found that the pair of gene frequencies, on leaving any set of nonzero initial values, would spiral away from the nontrivial equilibrium point (Fig. 1B). Therefore, Sedcole disagreed with the suggestion that Leonard's model explains the stability of host-pathogen systems in the 'fertile crescent' area of the Middel East.

Leonard and Czochor (5) later conceded to Sedcole that the nontrivial equilibrium point of Leonard's model is analytically unstable. However, in contrast to those of Sedcole, their computer simulations indicated that Leonard's model is stable. Realizing that Sedcole's (6) analytic treatment is strictly valid only in the immediate vicinity of an equilibrium point, they resolved this apparent conflict between the analytic results and their computer simulation results by suggesting that many concentric limit cycles surround the nontrivial equilibrium point (Fig. 1C). A limit cycle is a closed trajectory such that no trajectory sufficiently near it is also closed. Thus, Leonard and Czochor (5) imply that very close to the nontrivial equilibrium point Sedcole's analysis applies and the system spirals outward to the innermost limit cycle. Farther from the equilibrium point their computer simulations show that the system can spiral inward (5), presumably until it reaches the first stable limit cycle in its path. A limit cycle is stable or unstable, respectively, if any infinitesimally small perturbation to the closed orbit decays or grows with time. For example, the inner and outer limit cycles of Fig. 1C are stable and unstable, respectively. Leonard and Czochor (5) suggest that the behavior of host-pathogen systems in the Middle East is consistent with systems that have unstable equilibrium points, but stable limit cycles.

This letter has two purposes: First, in contrast to the claims of both Sedcole (6) and Leonard and Szochor (5), it shows that Leonard's (4) model is not necessarily locally asymptotically stable. Second, and more important, it demonstrates that Leonard's model is not robust; ie, slight changes in the assumptions can drastically affect its behavior. Three alternative sets of assumptions are examined. The genetic compositions of host and pathogen populations are assumed to change: in a sequence of simultaneous and discrete steps, in a sequence of alternate and discrete steps, and continuously. Each set of assumptions produces a qualitatively different type of stability behavior (Fig. 1). In this sense Leonard's

model is inadequate as an explanation of the stability of gene-for-gene relationships in the Middle East.

Leonard's (4) model is based on Tables 1 and 2 which are adaptations of Tables 2 and 3, respectively, in his article. From Table 1 Leonard expresses the frequency of the virulence gene in the pathogen population in the $(i + 1)$st generation as

$$n_{i+1} = \frac{n_i[1 - k + (1 - q^2)\, a]}{1 - (1 - q^2)\, t + n_i\, [(1 - q^2)\,(a + k]} \tag{1}$$

Similarly, from Table 2, the change in frequency of the susceptible gene in the host population is

$$\Delta q = \frac{(1 - q)\, q^2\, [ns(a + t) + c - ts]}{1 - s + nks - (1 - q^2)\, [ns(a + t) + c - ts]} \tag{2}$$

The system is at equilibrium when $n_{i+1} = n_i$ and $\Delta q = 0$ simultaneously.
By Eq. 1 and 2

$$n_{i+1} = n_i \text{ when } n = 0,\ 1 \text{ and when } q = q^* = \sqrt{1 - k/(a+t)}$$

and
$$\tag{3}$$

$$\Delta q = 0 \text{ when } q = 0,\ 1 \text{ and when } n = n^* = (ts-c)/(sa+st).$$

Given 'stabilizing selection' (sensu Van der Plank [7]), Leonard (4) has shown that a single nontrivial equilibrium exists at (n^*, q^*). In addition there are four trivial equilibria where each population has lost an allele. These are $(n,q) = (0,0),\ (0,1),\ (1,0)$ and $(1,1)$. Leonard and Czochor (5) claim that $(n,q) = (0,q^*),\ (1,q^*),\ (n^*,0)$ and $(n^*,1)$ also are nontrivial equilibria, but either $n_{i+1} \neq n_i$ or $\Delta q \neq 0$ at each of these points (Eq. 3).

It is agreed that Leonard's model cycles about the phase plane (the $n \times q$ plane in which $0 < n < 1,\ 0 < q < 1$). Those involved in the controversy are arguing whether the system spirals into the nontrivial equilibrium point (n^*, q^*) as originally indicated by Leonard's (4) computer simulations (see Fig. 1A), spirals away from it as Sedcole (6) states (see Fig. 1B), or does both (depending on initial conditions) with the trajectories of the two spirals converging on a limit cycle as Leonard and Czochor (5) suspect (see Fig. 1C).

Consider Sedcole's (6) claims first. He is correct in showing mathematically that the system

$$n_{i+1} = g(n_i, q_i),\quad q_{i+1} = q_i + \Delta q = f(n_i, q_i) \tag{4}$$

is locally unstable. Furthermore, there is no error in his computer simulation which suggests global instablity (Fig. 1B).

However, as Leonard and Czochor (5) note, Sedcole's system (Eq. 4), in which $\Delta q = \Delta q\,(n_i)$, is a misinterpretation of Leonard's model which is of the form:

$$n_{i+1} = g(n_i, q_i), \quad q_{i+1} = q_i + \triangle q(n_{i+1}) = f(n_{i+1}, q_i) \qquad (5)$$

According to Eqs. 1, 2, and 5:

$$\left.\frac{\partial g}{\partial n_i}\right|_{eq} = 1$$

$$\left.\frac{\partial g}{\partial q_i}\right|_{eq} = -\frac{2q^*n^*(1-n^*)(a+t)}{1-(1-q^{*2})t} = +x_{12}, \text{ say} \qquad (6)$$

$$\left.\frac{\partial f}{\partial n_i}\right|_{eq} = \frac{(1-q^*)q^{*2}\, s\, (a+t-kt)}{(1-s-n^*ks)[1-(1-q^{*2})t]} = x_{21}, \text{ say}$$

$$\left.\frac{\partial f}{\partial q_i}\right|_{eq} = \left[\left(\left(\frac{\partial f}{\partial g}\right)_{q_i} \times \left(\frac{\partial g}{\partial q_i}\right)_{n_i} + \left(\frac{\partial f}{\partial q_i}\right)_g\right)\right]_{eq}$$

$$= \left.\frac{\partial f}{\partial n_i}\right|_{eq} \times \left.\left(\frac{\partial g}{\partial n_i}\right)^{-1}\right|_{eq} \times \left.\frac{\partial g}{\partial q_i}\right|_{eq} + \left.\frac{\partial f}{\partial q_i}\right|_{eq}$$

$$= 1 - x_{12}x_{21}$$

in which $x_{12} > 0$. By using these values of the partial derivatives at the nontrivial equilibrium it can be shown (cf Sedcole [6]) that the characteristic roots are

$$\lambda = (2 - x_{12}x_{21} \pm \sqrt{(2 - x_{12}x_{21})^2 - 4})/2.$$

When $x_{12}x_{21} < 0$ or $x_{12}x_{21} > 4, |^3\lambda| > 1$ so the nontrivial equilibrium point is locally unstable. When $0 \leqslant x_{12}x_{21} \leqslant 4, |\lambda| = 1$ so the linear analysis inconclusively describes the behavior of the nonlinear equations (Eq. 5) near the nontrivial equilibrium point.

Thus, Sedcole (6) and Leonard and Czochor (5) err in unequivocally stating that the nontrivial equilibrium point of Leonard's model is locally unstable. In fact, substitution from the suggested (4) range of parameter values into Eq. 6 indicates that generally $.0075 \leqslant x_{12}\, x_{21} \leqslant .02$. Hence, the local stability of the nontrivial equilibrium point is uncertain for all reported computer simulations (4,5,6). Thus, neither the local stability analysis, nor the *relevant* computer simulations (4,5), result in any inconsistencies with the behaviors predicted by either asymptotic stability (Fig. 1A), or concentric limit cycles (Fig. 1C), in the prescribed (4) range of parameter values.

There is another area of critical uncertainty in Leonard's model. He (4,5) defines the parameters (Table 2) as the rate of loss in host fitness per unit amount of pathogen fitness. However, he provides no evidence to support his choice of values ($0.05 \leqslant s \leqslant 0.02$). In fact, x_{21} is so sensitive to s (Eq. 6) that the results of the local stability analysis for the nontrivial equilibrium point are inconclusive when s is small, but show instability when s is large. Computer simulations starting at $n = 0.7$ and $q = 0.8366$ produce inward spirals towards (n^*, q^*) for $s = 0.8$. This demonstrates that Leonard's model must have at least one stable limit cycle for these *particular* parameter values ($a = 0.0, c = 0.1, k = 0.3, t = 1.0$, and $s = 0.8$). However, other sets of parameter values may induce different behaviors when s is large. Hence, both global instability (Fig. 1B), and concentric limit cycles (Fig. 1C), represent possible forms of phase plane trajectories for Eq. 5 for large s.

At this stage in the analysis of Leonard's (4) model, three behaviors: global asymptotic stability, (concentric) limit cycle(s), and global instability (Fig. 1) must be entertained as possibilities. A single stable limit cycle seems quite likely and is included under '(concentric) limit cycle(s)'. Ultimately, the numerical values of the parameters may determine which behavior the model (Eq. 5) enacts.

Sedcole's model (Eq. 4) assumes that the genetic compositions of host and pathogen populations change in a sequence of simultaneous steps. During the growing season the pathogen adapts to the initial composition of the host and the host adapts to the initial composition of the pathogen. Thus, the influence exerted by one species on the other is determined solely by its genetic composition at the *very beginning* of the growing season. Sedcole's model implicitly assumes that the genetic feedback between the two populations occurs at discrete intervals and that it is reciprocal when it does occur.

In contrast, Leonard's model (Eq. 5) assumes that the genetic compositions of host and pathogen populations change in a sequence of alternate steps. During the growing season the pathogen goes through a number of generations in which it adapts to the unchanging genetic composition of the host. Thus, during the growing season, the host affects the pathogen but the pathogen has no effect on the genetic composition of the host. At the end of the growing season, seeds are produced by the host and these initiate a new host generation for the next season. The proportion of seeds produced by susceptible plants depends upon the genetic composition of the host population, which has been constant since the beginning of the growing season, and the genetic composition of the pathogen population at the *very end* of the growing season (Eq. 5). Thus, Leonard assumes that the genetic feedback between the two populations occurs at discrete intervals and that it is nonreciprocal when it does occur.

Actually, the relative reproduction of a host genotype depends upon the relative amount of disease it has suffered during the growing season (5). This, in turn, is related to the genetic composition of the pathogen population *throughout* the growing

TABLE 1. Relative pathogen fitnesses on different hosts

| | | Relative pathogen fitness on: | |
| | | rr (susceptible) | R-(resistant) |
Pathogen genotype	Frequency	q^2	$1-q^2$
v (avirulent)	$1-n$	1	$1-t^a$
V (virulent)	n	$1-k^a$	$1-k+a^a$

[a]k = cost of virulence, t = effectiveness of resistance, and a = advantage of virulent race on hosts with corresponding gene for resistance.
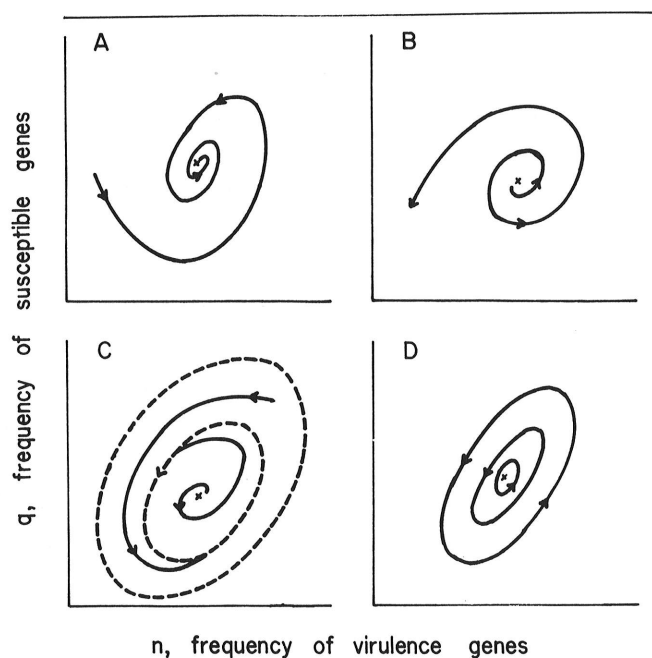

Fig. 1. Types of *qualitative* behavior which have been proposed for simple gene-for-gene relationships: **A**, asymptotically stable nontrivial equilibrium; **B**, unstable nontrivial equilibrium; **C**, stable limit cycle; and **D**, neutrally stable cycles. The position of the nontrivial equilibrium point is denoted by ×.

TABLE 2. Relative host fitnesses interacting with different pathogens

| Host genotype | Frequency | Relative host fitnesses when infected by: | |
| | | v (avirulent) | V (virulent) |
		$1-n$	n
rr (susceptible)	q^2	$1-s$	$1-s(1-k)$
R (resistant)	$1-q^2$	$1-c-s(1-t)$	$1-c-s(1-k+a)$

[a]c = cost of resistance, s multiplied by pathogen fitness = disease severity.

season (7). However, the models presented thus far implicitly assume that feedback occurs only at discrete intervals. Hence, the relative reproduction of host genotypes in these models is related, not to the relative amount of disease suffered and the genetic composition of the pathogen *throughout* the growing season, but rather, to the genetic composition of the pathogen at a *single instant* of the growing season.

An alternative set of assumptions leads to a third model. Suppose that any change in the genetic composition of one species immediately influences the genetic composition of the other. Then, if the populations are large, the discreteness introduced by individual "births" and "deaths" is lost in the large total during the growing season and both selection and feedback become approximately continuous processes. A model with continuous and reciprocal genetic feedback results. In this case the relative reproduction of host genotypes depends on the genetic composition of the pathogen *throughout* the growing season. However, this model has a weakness not present in either of the discrete models: it ignores the discontinuities introduced by seasonal and life-history phenomena in the life cycles of the two species.

The continuous value corresponding to the discrete fitness, W, is

$$m = \log_e W = \ln W$$

converting the fitnesses of Tables 1 and 2 with this expression, a continuous version of the discrete models can be adapted from Crow and Kimura (1):

$$\frac{dn}{dt} = n(1-n)\,[A(1-q^2) + Bq^2]$$

in which $A = \ln \dfrac{1-k+a}{1-t} > 0$, and $B = \ln(1-k) < 0$;

and

$$\frac{dq}{dt} = q^2(1-q)\,[Dn + H(1-n)] \qquad (7)$$

in which $D = \ln \dfrac{1-x(1-k)}{1-c-s(1-k+a)} > 0$, and $H = \ln \dfrac{1-s}{1-c-s(1-t)} < 0$

In common with the models of Leonard and Sedcole, Eq. 6 implicitly assumes that each population has either non-overlapping generations or a stable age distribution. Age distributions are approximately stable when the rate of change in total population size is slow relative to the lifespan.

The nontrivial equilibrium of Eq. 6 is:

$$(n^*, q^*) = [H/(H-D), \sqrt{A/(A-/B)}] \qquad (8)$$

The values of the partial derivatives at (n^*, q^*) are:

$$\left.\frac{\partial}{\partial n}\frac{dn}{dt}\right|_{eq} = \left.\frac{\partial}{\partial q}\frac{dq}{dt}\right|_{eq} = 0 ,$$

$$\left.\frac{\partial}{\partial n}\frac{dq}{dt}\right|_{eq} = (D-H)q^{*2}(1-q^*) > 0 ,$$

and $\left.\dfrac{\partial}{\partial q}\dfrac{dn}{dt}\right|_{eq} = (B-A)n^*(1-n^*)2q^* < 0.$

Kaplan (3) shows that because

$$\left.\frac{\partial}{\partial n}\frac{dn}{dt}\right|_{eq} + \left.\frac{\partial}{\partial q}\frac{dq}{dt}\right|_{eq} = 0$$

and $\left.\dfrac{\partial}{\partial n}\dfrac{dn}{dt}\right|_{eq} \times \left.\dfrac{\partial}{\partial q}\dfrac{dq}{dt}\right|_{eq} - \left.\dfrac{\partial}{\partial q}\dfrac{dn}{dt}\right|_{eq} \times \left.\dfrac{\partial}{\partial n}\dfrac{dq}{dt}\right|_{eq} > 0$

the nontrivial equilibrium is locally neutrally stable.

The existence of a Liapunov function constitutes mathematical proof that the results of local stability analysis extend beyond the immediate vicinity of (n^*, q^*).

A general expression for Liapunov functions of continuous gene-for-gene relationships can be adapted from Goh (2):

$$V(n,q) = \int_{n^*}^{n} \frac{f_n(u)}{g_n(u)}\,du + \int_{q^*}^{q} \frac{f_q(u)}{g_q(u)}\,du \qquad (9)$$

Here $f_x(u)$ and $g_x(u)$ are continuous functions of gene frequency, $u \in (0,1)$, such that $f_x(u) \gtreqless 0$ when $u \gtreqless x^*$, respectively, and $g_x(u) > 0$. In addition, $f_x(u)$ and $g_x(u)$ allow $V(n,q) \to \infty$ as n or $q \to 0^+$ or 1^-.

To support the neutrally stable conclusion of the local analysis it must be shown that

$$\dot{V}(n,q) = \frac{f(n)}{g(n)}\frac{dn}{dt} + \frac{f_q(q)}{g_q(g)}\frac{dq}{dt}$$

vanishes throughout the admissible phase plane. Hence, assume $\dot{V}(n,q) = 0$ and let $g_n(u) = u(1-u)$ and $g_q(u) = u^2(1-u)$ so that, after substituting Eq. 7 into this expression for $\dot{V}(n,q)$, $f_n(u) = (D-H)u + H$ and $f_q(u) = (A-B)u^2 - A$. Substituting these relationships into Eq. 8 before integrating,

$$V(n,q) = D\ln\left[\frac{1-n^*}{1-n}\right] - H\ln\frac{n^*}{n} + A\left[\ln\frac{q^*}{q} + \frac{1}{q} - \frac{1}{q^*}\right] - B\ln\left[\frac{1-q^*}{1-q}\right]$$

This proves mathematically that this continuous version of Leonard's model has behavior different from any previously suggested: the maintenance of arbitrarily large elliptical orbits in acceptable phase space. The size of the orbit is determined solely by the initial conditions (Fig. 1D).

In summary, it has been shown that three different models of the same gene-for-gene relationship each produce a qualitatively different type of stability behavior. This variation in behavior is not the result of mathematical error; rather, it is due to differences in assumptions made in constructing the models. In particular, the assumptions concerning the reciprocity and continuity of feedback between host and parasite genetic compositions are responsible.

In other words, the conclusions of the models are not robust; they are extremely sensitive to assumptions about the genetic feedback between the populations. But note that the term 'robust' is not a comment on the logical methods. The same sure mathematical rigor can produce both robust and nonrobust conclusions. Robustness is distinct from mathematics: it concerns the sensitivity of the conclusions to the assumptions used to build the model. In population genetics, only those familiar with the observational phenomena and the related mathematical approaches can determine the robustness of any particular model.

In conclusion, it has been shown that the stability of the gene-for-

gene models is critically dependent on the assumptions under which they were conceived. Since the assumptions specific to each model appear to have both favorable and unfavorable aspects, the qualitative behavior of the gene-for-gene relationship is in doubt. Hence, none of these models provides a satisfactory explanation of the stability of gene frequencies in host-pathogen systems in the Middle East. Stabilizing selection alone is not enough to ensure the stability of simple gene-for-gene relationships; other factors operating in conjunction with stabilizing selection are involved.

LITERATURE CITED

1. CROW, J. F., and M. KIMURA. 1970. An Introduction to Population Genetics Theory. Harper and Row, New York, NY.
2. GOH, B. S. 1978. Stability of some multispecies population models. To appear in: T. A. Burton, ed., Proc. Regional Conf. on Modeling and Differential Equations in Biology., Marcel Dekker, Amsterdam, The Netherlands.
3. KAPLAN, W. 1958. Ordinary Differential equations. Addison-Wesley, Reading, MA.
4. LEONARD, K. J. 1977 Selection pressures and plant pathogens. Ann. N. Y. Acad. Sci. 287:207-222.
5. LEONARD, K. J., and R. J. CZOCHOR. 1978. In response to 'Selection pressures and plant pathogens: stability of equilibria', Phytopathology 68:971-973.
6. SEDCOLE, J. R. 1978. Selection pressures and plant pathogens: stability of equilibria. Phytopathology 68:967-970.
7. VAN DER PLANK, J. E. 1968. Disease Resistance in Plants. Academic Press, New York, NY.