Letter to the Editor

Estimating Parasitic Fitness: A Reply

J. V. Groth and J. A. Barrett

Associate professor, Department of Plant Pathology, University of Minnesota, St. Paul 55108, and research associate, Department of Genetics, University of Cambridge, Downing Street, Cambridge, England.

Scientific Journal Series Paper 10,250, Minnesota Agricultural Experiment Station, St. Paul. The authors wish to thank Richard Fleming for critical review of the manuscript.

Accepted for publication 6 May 1980.

In a recent letter to the editor, MacKenzie (9) attempted to relate two somewhat different measures of population change in plant pathogens, namely rate of disease increase, r, and relative vegetative, or as he calls it, parasitic, fitness W. As he indicated, such a relationship could prove worthwhile in bridging the gap between epidemiology and population genetics. In plant pathology the former is more developed than the latter. In recent times, however, fitness, which is one of the most fundamental variables of population genetics, has been recognized as important in plant pathology in several contexts (2,5,6,9,10).

We feel compelled to respond to this letter because the relationship between these variables can and should be further clarified and because we have reservations about when and how they should be interchanged. We also show that much of the development of this relationship is found already in the population genetics literature. There are seven points of contention or clarification.

Point 1. MacKenzie (9) does not mathematically connect Vanderplank's use of r (11) and Crow and Kimura's use of w (1); and on page 10, column 2 and page 11, column 2, the use of w as a measure of *relative fitness* and a measure of *absolute fitness* are confused. The application of concepts established in one field to another area of study often can give rise to ambiguity; some explanation of the use of the term "fitness" will facilitate understanding of this point and some of the other points in this letter.

Absolute fitness (Definition 1) is a measure of the *number* of offspring left by, and survival rate of, individuals per interval of time and is thus a measure of population growth. If $w > 1$, the population is increasing; if $w < 1$, the population is decreasing. We can define *absolute fitness* as $w = 1 + s$; when $s > 0$, s is equivalent to a rate of interest and when $-1 < s < 0$, s is equivalent to rate of depreciation (or negative interest). If $s = 0$, there is no change in the *numbers* of individuals during that time interval.

This definition is most useful in population growth models of the type used routinely by epidemiologists, since it directly equates survival and reproduction—which constitute fitness—and incremental growth. In this case s is equivalent to rate of increase r in the discrete growth model of Vanderplank (11)

$$X_t = X_0 (1 + r)^t \qquad (1)$$

In models derived from this, s and r always can be interchanged. But note that the use of r in this context is not the same as its use in exponential (or continuously compounded) growth rate equations. Consider two monomorphic populations, both starting with N_0 individuals. One population exhibits discrete growth with an absolute fitness w during each time interval. The other population is growing continuously with a rate of increase r for the same time interval. At the end of one time interval, both populations consist of N individuals. For the population wth discrete growth, $N = N_0 w$ and for the continuously growing populations $N = N_0 e^r$ (1,11), and it follows that:

$$N_0 w = N_0 e^r$$

and

$$\ln w = r$$

If $w = 1 +$ "r" or $w = 1 + s$, then

$$\ln(1 + \text{"r"}) = r$$

or

$$\ln(1 + s) = r$$

In other words, absolute fitnesses of organisms increasing in discrete steps can be converted into equivalent continuous growth rates for the same time interval (eg, days or generations, and vice versa).

The second use of w is as a measure of *relative fitness*. Here the reproductive success and survival rates of individuals or genotypes during a specified time interval are standardized by comparison with a reference individual or genotype. Conventionally, such fitnesses are defined relative to the most fit genotype or individual in the population (Definition 2a); ie, by dividing the absolute fitness of a given genotype or individual by the absolute fitness of the most fit genotype or individual; relative fitnesses in this case lie in the interval $0 < w < 1$, in which the relative fitness $w = 1 - s$ and s is the selection coefficient. This definition is most commonly used in population genetic models, but it must be remembered that this definition is purely conventional. Indeed, some workers have preferred to define fitnesses relative to the least fit individuals in the population (Definition 2b), so that $1 < w < \infty$, in which the relative fitness $w = 1 + s$ and s is the selection coefficient.

Definitions 2a and 2b can be interchanged; for example consider two phenotypes in a population, then:

Relative fitnesses	PHENOTYPES	
	A	*a*
Definition 2a	1	$1 - s$
Definition 2b	$1 + s'$	1

Hence, $\dfrac{1}{1-s} = \dfrac{1+s'}{1}$ and $s' = \dfrac{s}{1-s}$

It is essential to appreciate the fact that when *relative fitnesses* are used, they can only be used to follow changes in the relative proportions of different phenotypes (or genotypes) in a population. On the other hand, w as a measure of *absolute fitness* can be used to follow both the changes in numbers and the proportions of different phenotypes in a population. The two uses of $w = 1 + s$, ie:

(use 1) where s is a growth rate parameter and $0 < w < \infty$ (Definition 1);

and (use 2) where s is a measure of relative mortality or reproductive success and $0 < w < 1$ (Definition 2a) or $1 < w < \infty$ (Definition 2b) should not be confused as they are not directly interchangeable. For example, it is not obvious whether the expression on page 10, column 1 uses F as a measure of absolute or relative fitness.

Point 2. In the equation on page 10, column 2, it was not clear whether q or p as stated above and below the equation, respectively,

0031-949X/80/09084003/$03.00/0

referred to the less fit component of the mixture; this was corrected later in a published erratum (9). For clarity, the intent was that q is the frequency of the less fit component.

Point 3. The expression $q = q_0(F)^t$ is incorrect (page 10, column 1). If a population of an asexually reproducing organism consists of two phenotypes, *A* and *a*, with frequencies *p* and *q* (respectively) and absolute fitnesses F_A and F_a per time interval (respectively), then: after time *t*, the proportion of *a* in the population, will be:

$$q = \frac{q_0(F_a^t)}{p_0(F_A^t) + q_0(F_a^t)} \tag{3a}$$

To obtain relative fitnesses we divide the absolute fitnesses by the absolute fitness of a reference phenotype, say *A*. Thus:

$$q = \frac{q_0(F_a/F_A)^t}{p_0(F_A/F_A)^t + q_0(F_a/F_A)^t} = \frac{q_0(W^t)}{p_0 + q_0(W^t)} \tag{3b}$$

in which *W* is the fitness of *a* relative to *A*.

It then follows that:
for absolute fitnesses,

$$\frac{q}{p} = \frac{q_0(F_a^t)}{p_0(F_A^t) + q_0(F_a^t)} \cdot \frac{p_0(F_A^t) + q_0(F_a^t)}{p_0(F_A^t)} = \frac{q_0(F_a^t)}{p_0(F_A^t)} \tag{4a}$$

and for relative fitnesses,

$$\frac{q}{p} = \frac{q_0(W^t)}{p_0 + q_0(W^t)} \cdot \frac{p_0 + q_0(W^t)}{p_0} = \frac{q_0(W^t)}{p_0} \tag{4b}$$

Irrespective of whether *F* is relative or absolute fitness, the expression $q = q_0(F^t)$ is incorrect, since it lacks a divisor. However, the next expression (page 10, column 2) is correct for relative fitnesses (Eq. 4b) and it has been in use for several years (6).

Point 4. The two expressions $q = q_0(W^t)$ and $q = q_0e^{rt}$, page 11, column 2 are incorrect; see Eq. 3a and b and Eq. 4a and b above. The third expression (page 11, column 2) cannot be derived from the above two expressions.

If the absolute fitnesses (see Point 1 above) of two phenotypes *A* and *a*, with frequencies $(1 - q)$ and q respectively, in an asexually reproducing population are considered, then, under discrete and continuous growth models, we obtain:

Discrete

$$\frac{q}{1-q} = \frac{q_0[w_a \text{ (absolute)}]^t}{(1-q_0)\,[w_A \text{ (absolute)}]^t}$$

$$\frac{q}{1-q} = \frac{q_0 w^t}{1-q_0}$$

if the fitness of *a* relative to *A* is *w*.

Continuous

$$\frac{q}{1-q} = \frac{q_0 e^{r_a t}}{(1-q_0)\, e^{r_A t}}$$

where r_A and r_a are the exponential growth rates (11) (or natural rates of increase, Lotka [8]), of *A* and *a* respectively.
Therefore,

$$\frac{q}{1-q} = \frac{q_0 \quad e^{(r_a - r_A)\,t}}{1-q_0}$$

Let $r = r_a - r_A$ where *r* is the *relative* rate of replacement if $r < 0$, or displacement if $r > 0$, of the phenotype *a* (see above).

$$\frac{q}{1-q} = \frac{q_0 e^{rt}}{1-q_0} \tag{5}$$

It then follows that for the discrete model:

$$\ln \frac{q}{1-q} = \ln \frac{q_0}{1-q_0} - t \ln w$$

if $w = 1 - s$, then

$$\ln (1-s) = \frac{1}{t}\left[\ln \frac{q}{1-q} - \ln \frac{q_0}{1-q_0}\right]$$

but $\ln (1 + x) = x - \frac{x^2}{2} + \frac{x^3}{3} - \frac{x^4}{4} + \ldots$

$$\therefore \text{ if } s \text{ is small, } -s \cong \frac{1}{t}\left[\ln \frac{q}{1-q} - \ln \frac{q_0}{1-q_0}\right] = \frac{1}{t}\left[\ln \frac{q(1-q_0)}{q_0\,(1-q)}\right] \tag{6}$$

and for the continuous model:

$$\ln \frac{q}{1-q} - \ln \frac{q_0}{1-q_0} = rt$$

$$\therefore r = \frac{1}{t} \ln\left[\frac{q}{1-q} - \ln \frac{q_0}{1-q_0}\right] \tag{7}$$

which is another form of (v), and is given by MacKenzie on page 12, column 1.

Hence, $\ln(w) = r$, if q and q_0 are the same in both discrete and continuous models. In other words, the natural logarithm of the relative fitness of *a* is equal to the difference between the continuous growth rates of the two phenotypes.

Equation 6 was first derived by Haldane in 1924 (3,4) as an approximation for small values of *s* when the changes in phenotype frequencies per generation are small, and $\Delta q \rightarrow dq/dt$. This derivation and conclusion have been quoted frequently in population genetics literature (eg, Li [7, page 277], in which they are set as an exercise for the student and in Crow and Kimura [1] [page 193]). Equation 7 is merely the continuous generation version of this expression, so that Haldane's simplifying approximation is not required.

Point 5. MacKenzie's arithmetic (page 12) shows that Haldane's expression is a good approximation. For example, the relative fitness per day of the sensitive strain is given (page 12, column 2, paragraph 2) as:

$$w = \ln (r) = \ln (-0.160) = 0.852$$
$$s = 1.0 - 0.852 = 0.148$$

Haldane's approximation gives:

$$-s \cong \frac{1}{t} \ln \frac{q(1-q_0)}{q_0\,(1-q)} \cong -0.160$$
$$s \cong 0.160$$

So that Haldane's expression (Eq. 6) gives a value close to that given by the exact expression (Eq. 7).

Point 6. On page 12, column 2, paragraph 6, MacKenzie describes a way of estimating "relative parasitic fitness": "All one needs to do is subtract the larger *r* from the smaller *r*...." But this is the "*r*" inserted into the expression on page 11 (see Points 1 and 4 above) which he has defined as "... the rate of exponential growth." Therefore, two incompatible definitions of the parameter *r* for his third expression on page 12, column 2 have been used in his derivation; for clarity, *r* is the *relative rate of replacement* of the less fit phenotype by the more fit phenotype in his model (see Point 4 above).

Point 7. In deriving a relationship between fitness and population growth, and then determining fitness from the r difference, MacKenzie introduces to the field of plant pathology a useful idea. However, he implies that fitness can be determined as accurately using *r* differences as it can using more direct methods; eg, mixture or replacement studies (6). Mixtures are subject to most of the same errors of estimation as are the determinations of *r*, but

there are some experimental limitations associated with the use of r to estimate fitness that can at least be partly avoided in mixture studies. Two special problems in the use of r to estimate fitnesses need to be mentioned.

(Problem 1) Estimation of fitnesses from r values can only be as good as the estimates of r. Separate r's are, and probably will continue to be, measured in separate experiments. If the r values are not measured in the same micro-environments, as they cannot be in separate experiments, large experimental errors are likely. Mixture studies avoid this problem. They allow the pathogen phenotypes to reproduce, and sampling to be done, not only on the same plants, but in the case of foliar pathogens, on the same leaves.

(Problem 2) In the early stages of an epidemic, the accuracy with which disease level can be measured is low, as compared with later stages of the epidemic. Yet early measurements are often included of necessity in order to be able to estimate r over a long period of increase before decreased availability of host tissue becomes a limiting factor. Such a practice can be expected to lower confidence in the estimate of r, and hence in the estimate of w. Direct measurement of w by using mixtures should not be subject to this source of error, since the proportions of the different pathogen genotypes, and not the absolute amount of disease, are all that matter. Fitness may well be density dependent, especially at higher disease levels. If these levels are avoided, as can be done through manipulation of the host stand (thus allowing disease to infect healthy plants periodically), fitness changes due to crowding should be minimal. Likewise the confidence with which disease is measured at the end of an epidemic is lowered by the fact that removal of susceptible host tissue is accounted for in a way that is itself an approximation, and that can be expected to add an increasingly important source of error as the effect of removals on r becomes more prominent.

In conclusion, the concepts that MacKenzie (9) presented are sound, but the development of these concepts was not clear, with the results that some readers may not have been able to fully appreciate their meaning and value. We hope that fitness theory and fitness measurement will, in the future, become an important subject of research in plant pathology.

LITERATURE CITED

1. CROW, J. F. and M. KIMURA. 1970. An introduction to population genetics theory. Harper and Row, New York, NY. 591 pp.
2. GROTH, J. V. 1976. Multilines and "super races": A simple model. Phytopathology 66:937-939.
3. HALDANE, J. B. S. 1924. A mathematical theory of natural and artificial selection. Trans. Camb. Phil. Soc. 23:19-41.
4. HALDANE, J. B. S. 1932. The Causes of Evolution. Longmans, Green, and Co., London, England. 235 pp.
5. KIYOSAWA, S. 1972. Theoretical comparison between mixture and rotation cultivations of disease resistant varieties. Ann. Phytopathol. Soc., Jpn. 38:52-59.
6. LEONARD, K. J. 1977. Virulence, temperature optima, and competitive abilities of isolines of rate T and O of *Bipolaris maydis*. Phytopathology 67:1273-1279.
7. LI, C. C. 1955. Population genetics. University of Chicago Press, Chicago, IL. 366 pp.
8. LOTKA, A. J. 1924. Elements of Physical Biology. The Williams and Wilkins Co., Baltimore, MD. (Reprinted 1956 as: "Elements of Mathematical Biology." Dover Publications, New York, NY, 465 pp.
9. MacKENZIE, D. R. 1978. Estimating parasitic fitness. Phytopathology 68:9-13, 1108 (erratum).
10. PERSON, C., J. V. GROTH, and O. M. MYLYK. 1976. Genetic change in host-parasite populations. Annu. Rev. Phytopathol. 14:177-188.
11. VANDERPLANK, J. E. 1963. Plant Disease: Epidemics and Control. Academic Press, New York, NY. 349 pp.