

Estimation of Relative Parasitic Fitness of Pathogens in Heterogeneous Host Populations

Ping Sun and Shimai Zeng

Associate professor and professor, Department of Plant Protection, Beijing Agricultural University, Beijing 100094, P. R. China. This research was supported in part by the National Science Foundation of China (grant 39230260). We thank the reviewers for their valuable suggestions. Accepted for publication 21 February 1995.

There have been many studies on relative parasitic fitness in plant pathology since Leonard's (5) pioneering work in 1969. Measurement of relative parasitic fitness offers an efficient way to quantitatively analyze host-pathogen interactions (2,3,6,12,13). The relationship between parasitic fitness and other related concepts also has been discussed (1,9).

The commonly used concept of relative fitness of race A to race B on a host can be described by its relative reproductive success: If p and q are the frequencies of races A and B, respectively, p_0 and q_0 are the initial frequencies, and w is the relative fitness of race A to race B on the host, then

$$w^t = (p/p_0)(q/q_0) = (p/q)(q_0/p_0), \quad (1)$$

in which parameter t is the number of generations or another unit of time, and the more fit race is always defined as race B so w is not larger than 1 (4,7,8,10).

Many plant pathogens have only one generation per growing season. If more than one generation is produced, t is difficult to measure because the pathogen generations overlap. We suggest that the following equation can be used to estimate relative fitness for multiple generations:

$$f = (p/q)(q_0/p_0), \quad (2)$$

in which p and q are frequencies of races A and B, respectively, at the end of a host's growing season and p_0 and q_0 are the initial frequencies for races A and B, respectively. To avoid confusion with w as defined by equation 1, f can be called seasonal fitness. If the pathogens have only one generation per growing season, then $f = w$, otherwise the values of w and f may be different; f also can reflect the relative parasitic ability of race A to race B on the host plant. Regardless of the complexity of the interaction between the two races, the final frequencies, p and q , and the initial frequencies, p_0 and q_0 , express this interaction. In practice, a pathogen population usually consists of several races, and there are usually multiple cultivars grown in an area. Equation 2 can be extended to increase its usefulness. Suppose that m cultivars are grown in a field, with percents c_1, c_2, \dots, c_m , respectively, and there are n races in the pathogen population, we use f_{ijk} to denote the relative seasonal fitness of race i to race j on cultivar k ; i and $j = 1$ to n , and $k = 1$ to m . The frequencies of the n races at the end of an epidemic are p_1, p_2, \dots, p_n , and the initial frequencies are $p_{10}, p_{20}, \dots, p_{n0}$. The overall fitness of the i th race relative to the j th race can be expressed by the following model:

$$(p_i/p_j)(p_{j0}/p_{i0}) = c_1 f_{ij1} + c_2 f_{ij2} + \dots + c_m f_{ijm}, \quad (3)$$

in which i differs from j and $1 \leq i \leq n$.

The race with the greatest fitness, if known, is designated the j th race. For example, if it is race 1, we use the following general equations:

$$(p_i/p_1)(p_{10}/p_{i0}) = c_1 f_{i11} + c_2 f_{i12} + \dots + c_m f_{i1m}, \quad (4)$$

for $i = 2$ to n . Specifically, with $i = 2$ to n , we have:

$$\begin{aligned} (p_2/p_1)(p_{10}/p_{20}) &= c_1 f_{211} + c_2 f_{212} + \dots + c_m f_{21m} \\ (p_3/p_1)(p_{10}/p_{30}) &= c_1 f_{311} + c_2 f_{312} + \dots + c_m f_{31m} \end{aligned} \quad (5)$$

$$\dots \dots \dots$$

$$(p_n/p_1)(p_{10}/p_{n0}) = c_1 f_{n11} + c_2 f_{n12} + \dots + c_m f_{n1m}.$$

In the above equations, the unknown parameter, f_{ik} ($i = 2$ to n ; $k = 1$ to m), can be estimated by arranging field experiments. If there are s sampling fields, c_{ik} denotes the percentage of cultivar k used in sampling field t . In each field we can obtain the value of $(p_i/p_1)(p_{10}/p_{i0})$ for $i = 2$ to n . We use $y_{it} = [(p_i/p_1)(p_{10}/p_{i0})]_t$ to denote the value of $(p_i/p_1)(p_{10}/p_{i0})$ obtained from the field, $t = 1$ to s . For a constant, i , the following linear model applies:

$$y_{it} = c_{i1} f_{i11} + c_{i2} f_{i12} + \dots + c_{im} f_{i1m}, \quad (6)$$

in which $t = 1$ to s .

Thus, the system of equations (for constant i , $k = 1$ to m , and $t = 1$ to s) can be represented by the following matrix operation:

$$\begin{pmatrix} c_{11} & c_{12} & \dots & c_{1m} \\ c_{21} & c_{22} & \dots & c_{2m} \\ \dots & \dots & \dots & \dots \\ c_{s1} & c_{s2} & \dots & c_{sm} \end{pmatrix}_{s \times m} \begin{pmatrix} f_{i11} \\ f_{i12} \\ \dots \\ f_{i1m} \end{pmatrix}_{m \times 1} = \begin{pmatrix} y_{i1} \\ y_{i2} \\ \dots \\ y_{is} \end{pmatrix}_{s \times 1}$$

or $C \cdot F_{i1} = Y_i$, in which F_{i1} is the $m \times 1$ vector, C is the $(c_{ik})_{s \times m}$ matrix; and Y_i is the $s \times 1$ vector.

When $s \geq m$, the least squares estimator (LSE) of F_{i1} is

$$\hat{F}_{i1} = (C' C)^{-1} C' Y_i, \quad (7)$$

in which symbol C' denotes the transpose of matrix C , $\hat{F}_{i1} = (\hat{f}_{i11}, \hat{f}_{i12}, \dots, \hat{f}_{i1m})'$, \hat{f}_{i1k} denotes the LSE of unknown parameter f_{i1k} for $k = 1$ to m .

For $i = 2$ to n , by substituting C , C' , and Y_i into equation 7, we

can obtain the corresponding LSE of F_{i1} . Our purpose is to obtain the nonnegative LSE of f_{i1k} ($i = 2$ to n ; $k = 1$ to m) because fitness is nonnegative. A related method was reported by Sun and Zeng (11).

When $m = 1$ and $n = 2$, i.e., only one cultivar is used and two races exist, equation 4 becomes

$$y_{21} = (p_2/p_1)(p_{10}/p_{20}) = f_{21}.$$

This is the case defined by equation 2, in which f_{21} is the relative seasonal fitness of race 2 to race 1 on the cultivar. Equation 4, therefore, can be considered a logical extension of equation 2. When $m = 1$ and $n > 2$, $y_{i1} = (p_i/p_1)(p_{10}/p_{i0}) = f_{i1}$, for $i = 2$ to n . If there are s observations, the mean $(1/s)(y_{i1} + y_{i2} + \dots + y_{is})$ is the LSE of f_{i1} , the relative seasonal fitness of race i to race 1 on that cultivar for $i = 2$ to n .

The following is a numerical example for $n = m = 2$ for estimating unknown parameters f_{211} and f_{212} and the application of regression analysis after f_{211} and f_{212} have been estimated. The corresponding model is as follows:

$$(p_2/p_1)(p_{10}/p_{20}) = c_1 f_{211} + (1 - c_1) f_{212}. \quad (8)$$

Suppose four experiments ($s = 4$) were conducted to estimate the relative seasonal fitness of the two races ($n = 2$) on the two cultivars ($m = 2$). Then the model is

$$y_{2t} = c_{1t} f_{211} + (1 - c_{1t}) f_{212},$$

in which $y_{2t} = (p_2/p_1)(p_{10}/p_{20})$, for $t = 1$ to 4.

For simplicity, suppose the initial frequencies of races 1 and 2 are 0.5 and 0.5 and also suppose that the proportions of cultivar 1 in the four fields are 0, 0.4, 0.7, and 0.9, respectively, and the final ratios of p_2/p_1 in the corresponding fields are 0.95, 0.55, 0.35, and 0.25, respectively. Then Y_2 is a 4×1 vector, the transpose of Y_2 is $Y_2' = (y_{21}, y_{22}, y_{23}, y_{24}) = (0.95, 0.55, 0.35, 0.25)$,

$$C = \begin{pmatrix} 0 & 1 \\ 0.4 & 0.6 \\ 0.7 & 0.3 \\ 0.9 & 0.1 \end{pmatrix}, \quad C' = \begin{pmatrix} 0 & 0.4 & 0.7 & 0.9 \\ 1 & 0.6 & 0.3 & 0.1 \end{pmatrix}.$$

and the LSE of F_{21} is as follows:

$$\hat{F}_{21} = (C'C)^{-1} C'Y_2, \quad (9)$$

in which $\hat{F}_{21} = (\hat{f}_{211}, \hat{f}_{212})'$ and \hat{f}_{21k} is the LSE of the relative seasonal fitness of race 2 to race 1 on cultivar k for $k = 1$ or 2.

By substituting matrix C , C' , and vector Y_2 into equation 9 and solving the equations, we obtain $\hat{f}_{211} = 0.134$, $\hat{f}_{212} = 0.916$ and the regression equation:

$$(p_2/p_1)(p_{10}/p_{20}) = 0.134c_1 + 0.916(1 - c_1). \quad (10)$$

Equation 10 can be used to predict the effect of host composition on race frequencies. Assuming the initial frequencies of the two races are equal ($p_{10} = p_{20} = 0.5$), equation 10 becomes:

$$(p_2/p_1) = 0.134c_1 + 0.916(1 - c_1). \quad (11)$$

For $0 \leq c_1 \leq 1$, we can estimate p_2/p_1 by equation 11. For example, if $c_1 = 0$, i.e., only cultivar 2 was used, $(p_2/p_1) = \hat{f}_{212} = 0.916$; if $c_1 = 1$, i.e., only cultivar 1 was used, $(p_2/p_1) = \hat{f}_{211} = 0.134$.

If p_2/p_1 needs to be fixed, i.e., we want to control the ratio of the two races, we can calculate c_1 by rearranging equation 11:

$$c_1 = [0.916 - (p_2/p_1)] / (0.916 - 0.134). \quad (12)$$

As long as $0.134 \leq (p_2/p_1) \leq 0.916$, we can find a c_1 , $0 \leq c_1 \leq 1$, that satisfies equation 12. When n and m are larger than 2, the equations can be solved by matrix algebra.

In estimating relative seasonal fitness, it is important to carefully define the meaning of "the end of a growing season" because it is closely related to the value of the final frequencies of races. We think "the end of a growing season" should be defined as the time when propagation of the pathogen population has stopped. For example, in the study of wheat stripe rust in China, a daily average temperature of 22°C is set as an indicator of the end of the epidemic season because no further infection and sporulation of the fungus can occur when the daily temperature is above 22°C. In some winter-wheat regions in China, the end of the epidemic is about 2 weeks before harvest, whereas in some mountain regions the end of the epidemic is at harvest when the daily average temperature is seldom higher than 22°C.

Relative parasitic fitness is a measure of the relative success of one pathogenic genotype in a given host population. We have presented a method to estimate fitness in the case of multiple races and cultivars. Although our models need to be tested by field-plot experiments, we foresee potential application of the approach in quantitative analysis of host-pathogen interactions and in predicting the effects of host mixtures on pathogen population.

LITERATURE CITED

- Andriveau, D. 1993. Nomenclature for pathogenicity and virulence: The need for precision. *Phytopathology* 83:889-890.
- Barrett, J. A. 1983. Estimating relative fitness in plant parasites: Some general problems. *Phytopathology* 73:510-512.
- Bronson, C. R. 1981. The influence of unnecessary virulence genes on the reproductive fitness of *Erysiphe graminis* f. sp. *tritici*. Ph.D. thesis. Michigan State University, East Lansing.
- Groth, J. V., and Barrett, J. A. 1980. Estimating parasitic fitness: A reply. *Phytopathology* 70:840-842.
- Leonard, K. J. 1969. Selection on heterogeneous populations of *Puccinia graminis* f. sp. *avenae*. *Phytopathology* 59:1851-1857.
- Leonard, K. J., and Czochoz, R. J. 1980. Theory of genetic interaction among populations of plants and their pathogens. *Annu. Rev. Phytopathol.* 18:237-258.
- MacKenzie, D. R. 1978. Estimating parasitic fitness. *Phytopathology* 68:9-13.
- MacKenzie, D. R. 1980. Estimating parasitic fitness: Epilogue. *Phytopathology* 70:842-843.
- Shaner, G., Stromberg, E. L., Lacy, G., Bark, K. R., and Pirone, T. P. 1992. Nomenclature and concepts of pathogenicity and virulence. *Annu. Rev. Phytopathol.* 30:47-66.
- Skylakakis, G. 1980. Estimating fitness of plant pathogenic fungi: A theoretical contribution. *Phytopathology* 70:696-698.
- Sun, P., and Zeng, S. M. 1993. Statistical model to study the relationship between the wheat variety composition and the race composition of wheat stripe rust. *Ecol. Modell.* 66:277-287.
- Sun, P., and Zeng, S. M. Modeling the cultivar-race interactions in inter-regional disease spread. *J. Plant Dis. Prot.* In press.
- Zeng, S. M. 1991. PANCRIN, a prototype model of the pandemic cultivar-race interaction of yellow rust on wheat in China. *Plant Pathol.* 40:287-295.