

Rothamsted Repository Download

A - Papers appearing in refereed journals

Van Den Bosch, F., Mcroberts, N., Van Den Berg, F. and Madden, L. V. 2008. The basic reproduction number of plant pathogens: matrix approaches to complex dynamics. *Phytopathology*. 98 (2), pp. 239-249.

The publisher's version can be accessed at:

- <https://dx.doi.org/10.1094/PHYTO-98-2-0239>

The output can be accessed at: <https://repository.rothamsted.ac.uk/item/89xy5/the-basic-reproduction-number-of-plant-pathogens-matrix-approaches-to-complex-dynamics>.

© Please contact library@rothamsted.ac.uk for copyright queries.

The Basic Reproduction Number of Plant Pathogens: Matrix Approaches to Complex Dynamics

F. van den Bosch, N. McRoberts, F. van den Berg, and L. V. Madden

First and third authors: Department of Biomathematics and Bioinformatics, Rothamsted Research, Harpenden, AL5 2JQ, UK; second author: Scottish Agricultural College, King's Buildings, West Mains Road, Edinburgh EH9 3JG, UK; and fourth author: Department of Plant Pathology, Ohio State University, Wooster, OH 44691-4096.

Accepted for publication 25 September 2007.

ABSTRACT

van den Bosch, F., McRoberts, N., van den Berg, F., and Madden, L. V. 2008. The basic reproduction number of plant pathogens: Matrix approaches to complex dynamics. *Phytopathology* 98:239-249.

The basic reproduction number, R_0 , is defined as the total number of infections arising from one newly infected individual introduced into a healthy (disease-free) host population. R_0 is widely used in ecology and animal and human epidemiology, but has received far less attention in the plant pathology literature. Although the calculation of R_0 in simple systems is straightforward, the calculation in complex situations is challenging. A very generic framework exists in the mathematical and biomathe-

matical literature, which is difficult to interpret and apply in specific cases. In this paper we describe a special case of this general framework involving the use of matrix population models. Leading by example, we explain the existing mathematical literature on this subject in such a way that plant pathologists can apply the method for a wide range of pathosystems.

Additional keywords: Beet necrotic yellow vein virus, comparative epidemiology, cultivar mixture, landscape, nursery, propagator tree, Plum pox virus.

This paper describes a systematic method, originally developed by Diekmann et al. (8), to calculate the basic reproduction number from knowledge of the pathogen's life-cycle and its interactions with the host plant (i.e., from knowledge of the disease cycle). The basic reproduction number, R_0 , is defined as the total number of infections arising from one newly infected individual introduced into a healthy population. Rephrasing this definition, R_0 is also the generation-to-generation multiplication factor of the pathogen population at infinitesimally low pathogen density (i.e., at a density when propagules produced by the pathogen will not come in contact with host individuals previously infected by the pathogen).

R_0 is widely used in ecology and animal and human epidemiology for several reasons. (i) The basic reproduction number defines the threshold, $R_0 = 1$, between an epidemic (increase in pathogen density), $R_0 > 1$, and no epidemic, $R_0 < 1$. When each 'mother' infection causes, on average, more than one 'daughter' infection, R_0 exceeds 1, so the pathogen or disease population density will increase. (ii) It is a useful parameter in comparative epidemiology, summarizing the life-cycle components and the interactions with the host into one metric signifying the pathogen's reproduction capacity. (iii) In many models (but not all), R_0 can be used to determine the ultimate or steady-state value of disease or pathogen density in a host population (10,21,23, 26,27,34), and (iv) R_0 is a simple tool to evaluate different disease control methods within one coherent framework.

The concept behind the basic reproduction number goes back at least to Ross (33), who studied malaria epidemics. Vanderplank (36,37) introduced the metric into the plant pathology literature. In Vanderplank's model, the corrected basic infection rate, R_c , is the number of new infections caused per time-unit by one infective infection, the parameter i is the length of the infectious period, which has the dimension of time. As discussed by Vanderplank, the product of these two, iR_c , thus has the dimension of numbers and is equal to the basic reproduction number, R_0 . Vanderplank uses this quantity, which he terms the progeny-parent ratio as a basis for assessing disease control strategies. Although the work of Vanderplank dates back to the sixties and seventies, R_0 is still not widely used by plant pathologists, in general, despite the intuitive interpretation and application of the parameter. Several studies in botanical epidemiology have addressed the use of R_0 in recent years (13,14,15,22,26,28,30,34), and the recent book by Madden et al. (27) discusses the calculation of R_0 for a range of epidemic models. The present paper seeks to extend the range of techniques available to plant pathological research in this regard.

The calculation of R_0 in cases with complex dynamics is not straightforward. Examples of systems with complex dynamics include: (i) the presence of more than one host cultivar; (ii) large spatial scales (e.g., regions) with multiple fields; and (iii) propagator-tree-nursery systems. A procedure used to calculate R_0 in most cases in the plant pathology literature (13,14,15,28), is to: (i) develop a nonlinear model for pathogen dynamics (usually consisting of a system of nonlinear differential equations); (ii) calculate the steady-state where the pathogen is present (the internal steady-state); and (iii) from the steady-state expression derive a criterion (i.e., an inequality) showing the combination of model parameters in which the pathogen population density is larger than zero. This method, however, does not guarantee that the expression derived actually is R_0 . That the expression is the basic reproduction number can only be ascertained by retrospectively finding a biological interpretation for the components of the ex-

Corresponding author: F. van den Bosch
E-mail address: Frank.vandenBosch@BBSRC.ac.uk

*The e-Xtra logo stands for "electronic extra" and indicates that the online version contains programming scripts for finding the dominant or maximum eigenvalue of an $n \times n$ matrix and some other software tips for starting to use matrix projection models.

pression that convinces the researcher that it is R_0 . In Box 1, we give a specific example clarifying the pitfalls of this approach.

A very generic framework exists in the mathematical and biomathematical literature (8,16) for the calculation of R_0 for virtually any biologically relevant case. Though brilliant in its generality, this framework is abstract and difficult to interpret and apply in specific cases, and has, so far, not been used in plant pathology studies. In this paper, we explain how this approach can be used to derive R_0 for a wide range of applications in plant pathology. This matrix based approach has been used in animal and human epidemiology (7,9,11,32). Motivated by plant disease examples, the main text introduces and generalizes, as appropriate, the matrix-based methods. Some interesting insights from these examples, which can be obtained using the methodology, are discussed in the Boxes. Technical issues on the use of computer packages are described in the Appendix.

We will use the discrete time step of one pathogen (or disease) generation to model the dynamics. The connection between this time step and the basic reproduction number is that the time step of one generation precisely coincides with the interpretation of R_0 as the generation-to-generation multiplication factor. It will be shown that exactly this use of generation as time step enables us to calculate the basic reproduction number. Generation will be denoted with subscript n . So the density of the pathogen (or density of the number of infections) in generation n is denoted by I_n . The life-cycle parameters can have the dimension day, month, or year, as appropriate, for the example under discussion. For example, the latent period modeled in example 1 would be measured in days, and the spore production rate in number of spores per day.

SETTING THE SCENE

Consider the simple case of a foliar plant pathogen causing discrete lesions on wheat leaves, such as, for example, the wheat leaf rust pathogen, *Puccinia triticina*. In a crop consisting of a single cultivar, the pathogen has a basic reproduction number R_0 . If the number of lesions (pustules) in a host population is low (so

that there is zero probability that a produced and disseminated spore is deposited on a previously formed lesion), the density of lesions in generation $n+1$, I_{n+1} , is given by

$$I_{n+1} = R_0 I_n \quad (1)$$

This is because, by definition, each mother lesion produces during its entire infectious period (i.e., lifetime), on average, R_0 daughter lesions in the next generation. For the remainder of the current article, we assume that the density of lesions (or any other units of infection) is low enough that there is no pathogen-imposed limitation to pathogen increase (i.e., that spores or other units of infection do not come in contact with previously formed infections).

This model can easily be solved numerically. Simply start with an initial density of lesions I_0 , substitute this number in equation 1 to find the density of lesions in the first generation, I_1 (for a given R_0). Iterating this process n times (for $n = 2, 3, 4, \dots, N$) will produce the lesion population density for generation n . Figure 1 shows the solutions, with $I_0 = 1$, for different values of R_0 ; the graph makes clear that R_0 is the generation-to-generation multiplication factor. Given a time series of lesion density values (where we note again that time is measured in generations), the basic reproduction number can be calculated from $R_0 = I_{n+1}/I_n$. This simple example shows that the basic reproduction number, R_0 , can be found from (i) deriving a model that connects the density of infections in generation n to the density in generation $n+1$; (ii) solving the model numerically; (iii) calculating the basic reproduction number as the quotient of the density of infections of two successive generations. With one modification, this recipe is also applicable in more complex situations to calculate R_0 . We will however show that the generation-to-generation multiplication factor can be calculated easily from life-cycle components using computer packages, and in some special cases can be calculated explicitly.

It should be noted that equation 1 can also be solved analytically (27), and the solution is given by

$$I_n = (R_0)^n I_0 \quad (2)$$

This solution again shows the role of R_0 as a generation-to-generation multiplication factor. Depending on additional assump-

BOX 1. THE PITFALLS IN CALCULATING THE BASIC REPRODUCTIVE NUMBER FROM NONLINEAR MODELS

Consider a plant population infected with a vector- and seed-transmitted viral pathogen. The dynamics of the density of healthy, $H(t)$, and infected, $I(t)$, plants are described by

$$\begin{aligned} \frac{dH}{dt} &= \left[\begin{array}{l} \text{birth rate} \\ \text{of healthy plants} \end{array} \right]_H - \left[\begin{array}{l} \text{death rate} \\ \text{of healthy plants} \end{array} \right]_H - \left[\begin{array}{l} \text{vectored} \\ \text{virus transmission} \end{array} \right]_H \\ \frac{dI}{dt} &= \left[\begin{array}{l} \text{birth rate} \\ \text{of infected plants} \end{array} \right]_I - \left[\begin{array}{l} \text{death rate} \\ \text{of infected plants} \end{array} \right]_I + \left[\begin{array}{l} \text{vectored} \\ \text{virus transmission} \end{array} \right]_H \end{aligned} \quad (B1.1)$$

Assuming that each plant has a constant probability, μ , per time unit to die and that the rate of vectored transmission is described by the parameter α , we have $[\text{death rate}]_H = \mu H$, $[\text{death rate}]_I = \mu I$ and $[\text{vectored virus transmission}] = \alpha HI$. We assume that the total plant population density, $H + I$, is given by a constant, K , implying that every plant that dies is instantaneously replaced so that $[\text{birth rate of healthy plants}]_H + [\text{birth rate of infected plants}]_I = \mu H + \mu I = \mu K$. A fraction q of the seeds from infected plants carries the virus and thus $[\text{birth rate of infected plants}]_I$ is proportional to qI . This also implies that $[\text{birth rate of infected plants}]_H$ is proportional to $H + (1 - q)I$. Combining this with the assumption on constant plant population density, and substituting all parts in model equation B1.1, we find the following as model equations:

$$\begin{aligned} \frac{dH}{dt} &= \mu(H + (1 - q)I) - \mu H - \alpha HI \\ \frac{dI}{dt} &= \mu q I - \mu I + \alpha HI \end{aligned} \quad (B1.2)$$

The steady-state densities are found by setting $\frac{dH}{dt} = 0$ and $\frac{dI}{dt} = 0$ and rearranging for I and H . The internal steady-state for the density of infected plants is given by $\hat{I} = \frac{1}{\alpha}(\mu q - \mu + \alpha K)$; $\hat{I} > 0$ if $\mu q - \mu + \alpha K > 0$. This inequality can be rearranged, for example, as

$$\frac{\alpha}{\mu(1 - q)} K > 1 \quad \text{or} \quad \frac{\alpha}{\mu} K + q > 1 \quad (B1.3)$$

Both expressions are thresholds for an epidemic, and show that when the parameter combinations on the left-hand side of the inequality sign are larger than 1, the disease has a steady-state density larger than zero (i.e., an epidemic occurs). R_0 must also be larger than 1 for an epidemic. However, the two expressions of equation B1.3 are not the same, and only one of these is equal to R_0 . So, which of these parameter combinations is the real basic reproduction number? We leave this here as an exercise, noting that in a more complex model the authors have struggled for quite a while to understand the difference between these threshold expressions.

tions, R_0 of equation 2 can then be explicitly linked to (and predicted from) disease cycle components such as the mean infectious period, sporulation rate, and probabilities of spores contacting the host (27). For instance, in a simple situation, R_0 can be estimated by: $\alpha\rho\gamma\tau H$, where α is the spore production per unit time of a lesion, ρ is the probability that a produced spore is deposited on a susceptible site, γ is the probability that a deposited spore causes an infection, τ is the length of the infectious period of a lesion, and H is the density of susceptible sites in a host population. For the complex cases to be discussed in this paper, the model equations can also be solved analytically, but this will only in very special cases lead to an *explicit* expression for R_0 .

EXAMPLE 1: MIXTURES OF CULTIVARS

We continue the example of lesion forming foliar pathogens but will include cultivar mixtures. A crop is grown from a random mixture of seeds of two cultivars. A fraction, q , of the seeds is cultivar 1, with a fraction $1 - q$ being cultivar 2. Both cultivars are susceptible to the fungal pathogen, but they have different effects on one or more life-cycle parameters. We consider the following epidemic conditions.

- A sporulating lesion on cultivar i produces, on average, α_i spores per time unit.
- The probability that a spore is deposited on a susceptible

site (site is an area of leaf that can contain a lesion) equals ρ (this probability will depend on crop density, spore transport mechanisms, and environmental circumstances). Given that the spore is deposited on a susceptible site in the crop, a fraction q is deposited on cultivar 1 and a fraction $1 - q$ on cultivar 2. Total site density is denoted by H .

- A spore deposited on a susceptible site of cultivar i will germinate and infect with probability γ_i .
- A lesion on cultivar i has an infectious period of τ_i time units.
- The disease has a negligible latent period. (see the *note* at the end of this section).

From this description of the life-cycle components, it is not immediately obvious what the basic reproduction number in the cultivar mixture is in terms of the parameters. What is simple to calculate, however, is the number of daughter lesions on cultivar i ($i = 1, 2$) resulting from one mother lesion on cultivar j ($j = 1, 2$), which we will denote by R_{ij} (Fig. 2A). That is, new lesions on a given cultivar have been caused by previously existing lesions on either total. For example, to calculate R_{11} , one should realize that the total number of spores produced by a lesion on cultivar 1 equals $\alpha_1\tau_1$, a fraction ρqH of these spores is deposited on a leaf of cultivar 1, and a fraction γ_1 of these deposited spores germinates and forms a daughter lesion. Therefore $R_{11} = \gamma_1\alpha_1\tau_1\rho qH$. Similar reasoning gives $R_{22} = \gamma_2\alpha_2\tau_2\rho(1 - q)H$, $R_{21} = \gamma_2\alpha_1\tau_1\rho(1 - q)H$, and $R_{12} = \gamma_1\alpha_2\tau_2\rho qH$. Note that the latter two expressions account for

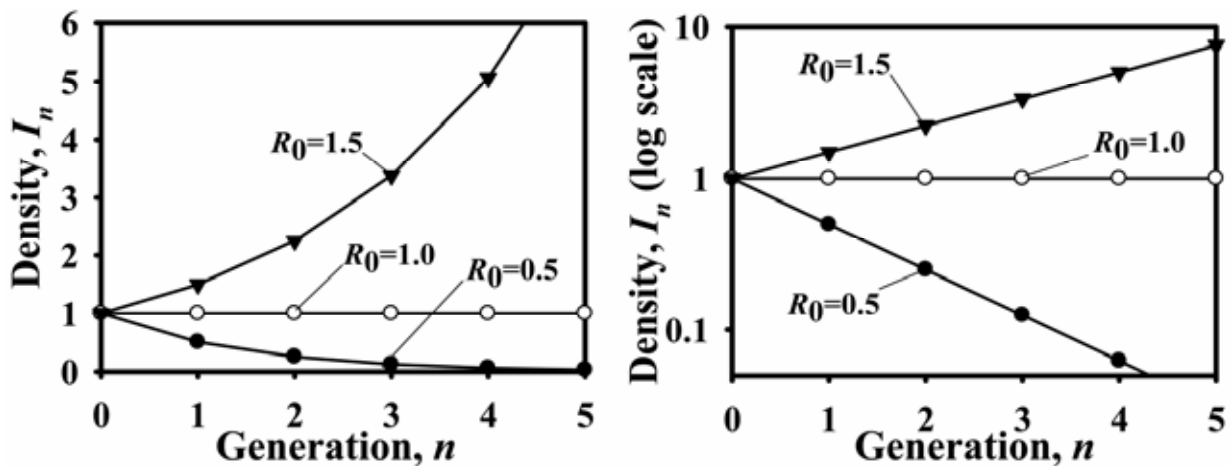


Fig. 1. The density of lesions, I_n , as a function of the pathogen generations, n , elapsed since the start of the experiment for three values of the basic reproduction number, R_0 . Left graph, ordinary scale; right graph, log-scale.

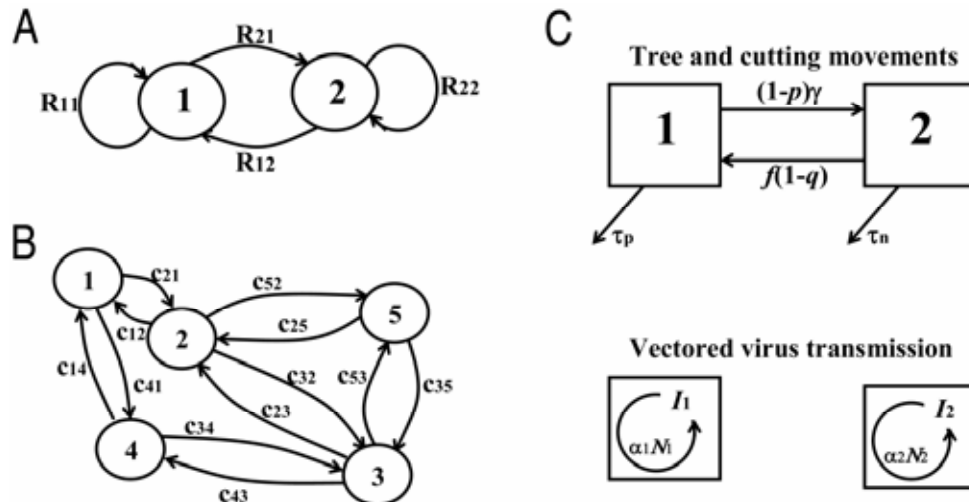


Fig. 2. Schematic representation of the three examples of pathosystems used in the paper.

new lesions developing on the ‘opposite’ cultivar to the one where the spores were produced.

Using these four elements we build a model for the dynamics of the lesion density on leaves of cultivar 1 in generation n , $I_{1,n}$, and on cultivar 2, $I_{2,n}$. If there are $I_{1,n}$ lesions on cultivar 1 in generation n these lesions will cause $R_{11}I_{1,n}$ lesions on cultivar 1 in generation $n+1$. Similarly $I_{2,n}$ lesions on cultivar 2 will cause $R_{12}I_{2,n}$ lesions on cultivar 1 in pathogen-generation $n+1$. Our model equation relating lesion numbers in successive pathogen-generations on cultivar 1 thus reads $I_{1,n+1} = R_{11}I_{1,n} + R_{12}I_{2,n}$. Similar arguments can be used to derive the model equation for lesion numbers on cultivar 2. Our model for the development of lesion numbers thus reads

$$\begin{aligned} I_{1,n+1} &= R_{11}I_{1,n} + R_{12}I_{2,n} \\ I_{2,n+1} &= R_{21}I_{1,n} + R_{22}I_{2,n} \end{aligned} \quad (3)$$

A more compact notation of such models is the matrix-vector notation

$$\vec{I}_{n+1} = A\vec{I}_n \quad (4)$$

where

$$\vec{I}_n = \begin{pmatrix} I_{1,n} \\ I_{2,n} \end{pmatrix} \text{ and } A = \begin{pmatrix} R_{11} & R_{12} \\ R_{21} & R_{22} \end{pmatrix} \quad (5)$$

Some further information about matrix vector notation can be found in, for example, the books by Bretcher (3) and Caswell (4, pages 652-668).

The model can be solved numerically in the same way as we solved equation 1. Figure 3 shows results for a specific set of parameter values and initial conditions. These graphs show that, exactly as for equation 1, the density of lesions on each of the cultivars increases from generation to generation, with a constant multiplication factor. This increase is the same factor for I_1 and I_2 (see Discussion) after a few generations. The deviation from this pattern for the first few generations, in this specific case $n < 3$, is due to the chosen initial conditions, whose influence quickly disappears. The multiplication factor, λ_d , is calculated from $\lambda_d = I_{i,n+1}/I_{i,n}$ ($i = 1,2$) for non-small values of n , so that the influence of the initial conditions does not affect λ_d . Now note that since our model has a time step of one generation, λ_d is the generation-to-generation multiplication factor and is thus the basic reproduction number, R_0 .

We conclude that the recipe in the previous section also applies to model equations 3, but that the multiplication factor has to be calculated at values of n for which the effects of the initial condition have disappeared. Box 2 gives some elaborations of the results of this example using the numerical recipe. In the Discus-

sion, we explain in more detail why the solutions are affected by the initial conditions for small values of n .

Models of the structure of equation 4 can be solved analytically (2–4). Given the results above and Figure 3, for large values of n it is not surprising that the solutions are given by

$$\begin{aligned} I_{1,n} &= \lambda_d^n C_1 \\ I_{2,n} &= \lambda_d^n C_2 \end{aligned} \quad (6)$$

where the constants C_1 and C_2 are related to the initial densities of infection. The λ_d can directly be calculated from matrix A in equation 4 (3,4,12), and is called the dominant eigenvalue of A , which is

$$\lambda = \frac{R_{11} + R_{22} + \sqrt{(R_{11} + R_{22})^2 - 4(R_{11}R_{22} - R_{21}R_{12})}}{2} \quad (7)$$

Substituting the expressions for R_{ij} and some algebraic manipulation we find that

$$R_0 = q(\gamma_1\alpha_1\tau_1\rho H) + (1-q)(\gamma_2\alpha_2\tau_2\rho H) \quad (8)$$

Equation 6 shows that the dominant eigenvalue is the growth factor for the density of lesions in each time step. Since the time step equals one generation, this dominant eigenvalue is the generation-to-generation multiplication factor and thus the basic reproduction number.

Box 2 gives some notes on and elaborations of this expression. Note that in cases where the matrix is larger than 2×2 , analytical solutions for the eigenvalue are not generally available, except in very special cases; however, algorithms for finding them are included in many statistical and mathematical packages, as well as in some spread-sheet programs (Appendix).

For two-component systems, we can state that (as illustrated here) the model for the generation-to-generation dynamics of lesion density can be used to calculate R_0 . Subsequently, the expression for R_0 can be used to study the effect of parameter values (relating to the process under study, such as infectious period and sporulation rate) on the basic reproduction number, either numerically or explicitly using equations 7 and 8.

Note on the assumption of negligible duration of the latent period. The method described above is easily extended to include nonnegligible duration of the latent period. Each of the matrix entries R_{ij} then needs to take account of the probability that a lesion survives through the latent period and enters the infectious period. In many models for animal and human pathogens, this probability equals 1, for plant pathogens, however, the probability to survive the latent period can be less than 1, for example due to plant defense responses and/or leaf necrosis/leaf shedding.

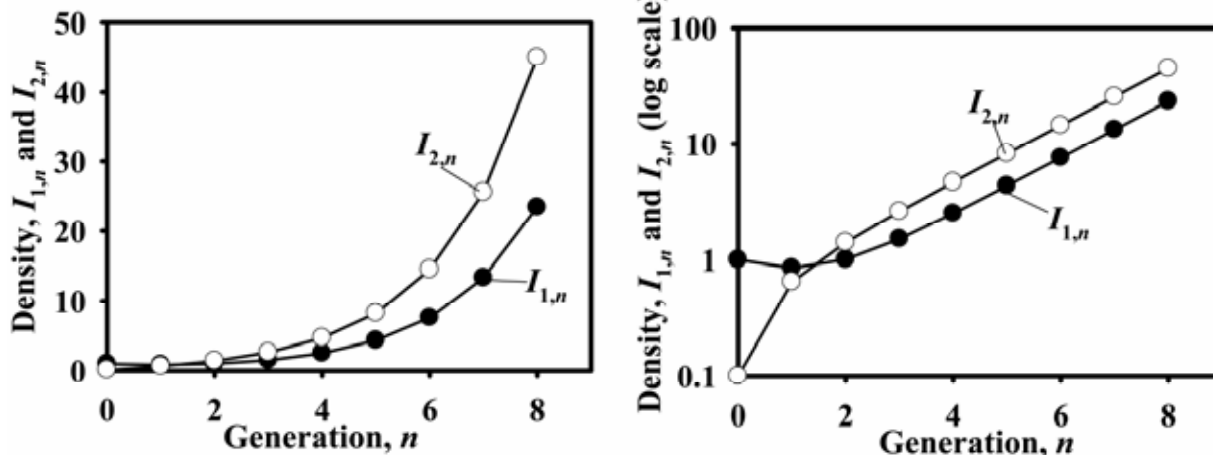


Fig. 3. Pathogen dynamics in a cultivar mixture. The density of lesions on cultivar 1, $I_{1,n}$, and on cultivar 2, $I_{2,n}$, as a function of the pathogen generations, n , elapsed since the start of the epidemic. These are calculated using equation 3, with parameter values $R_{11}=0.8$, $R_{12}=0.5$, $R_{21}=0.5$, $R_{22}=1.5$, and initial conditions $I_{1,n=0}=1.0$ and $I_{2,n=0}=0.1$. Left graph, ordinary scale; right graph, log-scale.

Throughout the paper we assume a negligible duration of the latent period because this simplifies the presentation. We note however again that extensions to include a nonnegligible latent period duration are easily done.

EXAMPLE 2: PATHOGEN DISPERSAL IN A LANDSCAPE

Consider a soilborne pathogen invading a system of fields in a region. The pathogen is dispersed between fields in soil on tractor wheels, farmer boots, etc. As a specific example, one can think of the invasion of *Polymyxa betae*, a soilborne protist that infects fibrous roots of sugar beet plants. The pathogen itself does not cause major damage, but the virus it transmits, *Beet necrotic yellow vein virus* (BNYVV), does cause severe damage to sugar beet crops (2,17,31,35,38). Assume we are interested in the effect of the spatial arrangement of the fields in the region on R_0 . For the sake of simplicity here (although the approach is easily generalized), we assume that all fields are identical and that dispersal depends only on the distance between fields. We model the

density of infected plants, $I_{i,n}$, in each field i in generation n . The epidemic conditions are as follow.

- Each infected plant produces α infectious propagules per time unit.
- A fraction d of these infectious propagules is dispersed in soil on tractor wheels and farmer boots.
- An infectious propagule is dispersed from field j to field i with probability c_{ij} ; c_{ij} will be called the connectivity between field j and field i .
- An infectious unit has a probability γH to infect a plant, where H is the crop density and γ the infection parameter.
- A plant is infectious for τ time units.

Consider an infected plant in field 1, and assume here there are five fields. During its entire infectious period this plant produces $\alpha\tau$ infectious propagules, of which a fraction $1 - d$ stays in field 1; thus, in the next generation, each infection causes $\gamma H(1 - d)\alpha\tau$ daughter infections. A fraction d is dispersed, and of these propagules, a fraction c_{21} reaches field 2. The total number of daughter infections in the next pathogen generation in field 2 from a mother infection in field 1 thus is $\gamma H d c_{21} \alpha \tau I_{1,n}$. Following

BOX 2. DISEASE DEVELOPMENT IN CULTIVAR MIXTURES

What is the effect of the ratio of the two cultivars on the basic reproduction number? In many cases, disease severity/pathogen density is less in a mixture than the arithmetic mean disease severity/pathogen density in separate pure monocultures (19,20, 29). This phenomenon is often assigned to the presence of pathogen strains differentially adapted to the cultivars in the mixture, causing plants less susceptible to a certain strain to act as filters for spores of this strain, which then do not infect the more susceptible cultivars in the mixture.

Can such effects of mixtures on disease development also be found in cases where there is no differential susceptibility to strains? Our model can describe this situation with effectively one strain in a mixture (equations 7 and 8). Figure B2.1A shows the effect of the fraction of cultivar 1 in the mixture, q , on the basic reproduction number. Other sets of parameter values show qualitatively similar results, and we conclude that for this type of pathosystem, the basic reproduction number is exactly the arithmetic mean in pure stands.

However, most published experimental data show disease/pathogen density after a number of generations have elapsed since

inoculation of the plots. Referring to equations 2 and 6, these studies thus show $C(R_0)^n$ as a function of q , where C is related to the initial density. Because R_0 is a function of q , raising it to powers above 1 results in a nonlinear change in infection density. Figure B2.1B shows that for $n > 1$, the observed pattern does emerge of disease or pathogen density in the mixture being less than the arithmetic mean of pure lines.

The analytical solution for this case, equation 8, shows that for any set of parameter values, there is a linear relation between q and R_0 , demonstrating that the evidence from the specific numerical results hold for any set of parameter values. This analytical expression also shows that for any parameter value and any $n > 1$, pathogen density in the mixture is less than the arithmetic mean in the pure monocultures. We thus conclude that the observed relation between pathogen density and q naturally emerges from the basic reproduction number of a single pathogen strain in a mixture, and that multiple strains of the pathogen (differentially adapted to the different cultivars) are not required for the benefits of multilines to be exhibited.

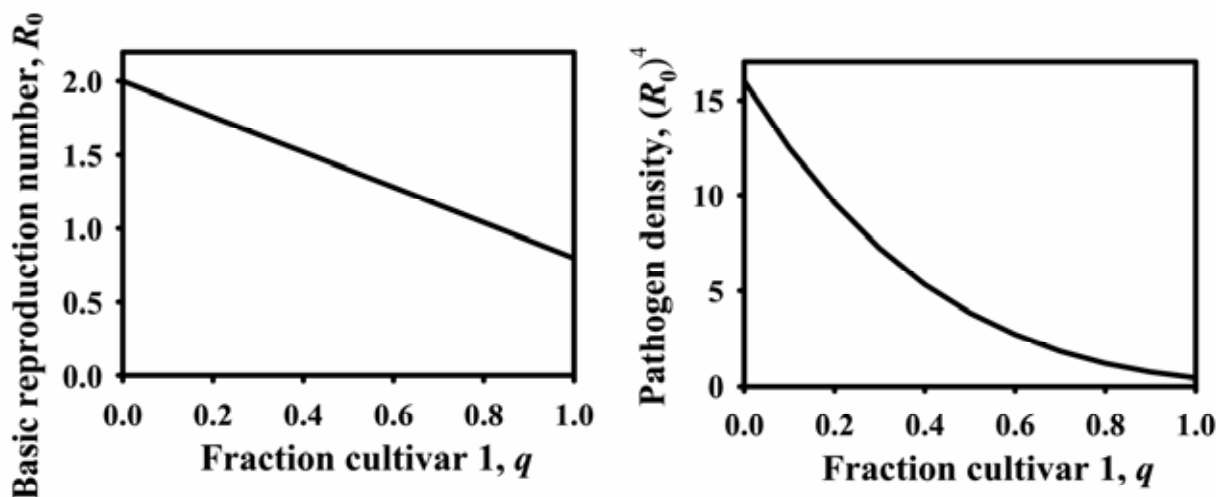


Fig. B2.1. The basic reproduction number, R_0 (left graph), and the pathogen density $C(R_0)^n$ in generation $n = 4$ (right graph, with $C = 1$) as function of the fraction of the crop plants in a mixture that are of cultivar 1, q . R_0 is given by equation 8, with $\gamma_1\alpha_1\tau_1\rho H = 0.8$ and $\gamma_2\alpha_2\tau_2\rho H = 2.0$. The figure shows that R_0 in a cultivar mixture is the arithmetic mean of the R_0 values in pure monocultures, whereas the pathogen density after $n > 1$ generations is less than arithmetic mean density in the pure monocultures.

the same reasoning for the assemblage of fields shown in Figure 2B, where no arrows between fields means no dispersal, we find the model equations:

$$\begin{aligned}
 I_{1,n+1} &= \gamma H(1-d)\alpha\tau I_{1,n} + \gamma Hdc_{12}\alpha\tau I_{2,n} + \gamma Hdc_{13}\alpha\tau I_{3,n} + \gamma Hdc_{14}\alpha\tau I_{4,n} + \gamma Hdc_{15}\alpha\tau I_{5,n} \\
 I_{2,n+1} &= \gamma Hdc_{21}\alpha\tau I_{1,n} + \gamma H(1-d)\alpha\tau I_{2,n} + \gamma Hdc_{23}\alpha\tau I_{3,n} + \gamma Hdc_{24}\alpha\tau I_{4,n} + \gamma Hdc_{25}\alpha\tau I_{5,n} \\
 I_{3,n+1} &= \gamma Hdc_{31}\alpha\tau I_{1,n} + \gamma Hdc_{32}\alpha\tau I_{2,n} + \gamma H(1-d)\alpha\tau I_{3,n} + \gamma Hdc_{34}\alpha\tau I_{4,n} + \gamma Hdc_{35}\alpha\tau I_{5,n} \\
 I_{4,n+1} &= \gamma Hdc_{41}\alpha\tau I_{1,n} + \gamma Hdc_{42}\alpha\tau I_{2,n} + \gamma Hdc_{43}\alpha\tau I_{3,n} + \gamma H(1-d)\alpha\tau I_{4,n} + \gamma Hdc_{45}\alpha\tau I_{5,n} \\
 I_{5,n+1} &= \gamma Hdc_{51}\alpha\tau I_{1,n} + \gamma Hdc_{52}\alpha\tau I_{2,n} + \gamma Hdc_{53}\alpha\tau I_{3,n} + \gamma Hdc_{54}\alpha\tau I_{4,n} + \gamma H(1-d)\alpha\tau I_{5,n}
 \end{aligned}
 \tag{9}$$

or in matrix vector notation $\vec{I}_{n+1} = A\vec{I}_n$, where

$$\vec{I}_n = \begin{pmatrix} I_{1,n} \\ I_{2,n} \\ I_{3,n} \\ I_{4,n} \\ I_{5,n} \end{pmatrix} \quad A = \begin{pmatrix} \gamma H(1-d)\alpha\tau & \gamma Hdc_{12}\alpha\tau & \gamma Hdc_{13}\alpha\tau & \gamma Hdc_{14}\alpha\tau & \gamma Hdc_{15}\alpha\tau \\ \gamma Hdc_{21}\alpha\tau & \gamma H(1-d)\alpha\tau & \gamma Hdc_{23}\alpha\tau & \gamma Hdc_{24}\alpha\tau & \gamma Hdc_{25}\alpha\tau \\ \gamma Hdc_{31}\alpha\tau & \gamma Hdc_{32}\alpha\tau & \gamma H(1-d)\alpha\tau & \gamma Hdc_{34}\alpha\tau & \gamma Hdc_{35}\alpha\tau \\ \gamma Hdc_{41}\alpha\tau & \gamma Hdc_{42}\alpha\tau & \gamma Hdc_{43}\alpha\tau & \gamma H(1-d)\alpha\tau & \gamma Hdc_{45}\alpha\tau \\ \gamma Hdc_{51}\alpha\tau & \gamma Hdc_{52}\alpha\tau & \gamma Hdc_{53}\alpha\tau & \gamma Hdc_{54}\alpha\tau & \gamma H(1-d)\alpha\tau \end{pmatrix}
 \tag{10}$$

Finally, we calculate R_0 by numerically solving model equation 9 and, for large values of n , calculate $R_0 = I_{1,n+1}/I_{1,n}$. Box 3 gives some elaborations of this example.

As mentioned above, in principle, R_0 can be determined analytically as the dominant (first) eigenvalue of the 5×5 matrix A in equation 10. Despite the fact that, no simple formula exists, in general, for the eigenvalues of 5×5 matrices, the dominant eigenvalue can be easily calculated with the use of a mathematical software package. This avoids the necessity of performing the tedious exercise of (i) numerically solving the model system in equation 9, (ii) looking through the output to determine when the effects of the initial conditions on disease dynamics have disappeared, and (iii) calculating R_0 for pairs of generations. The Appendix discusses methods to calculate the dominant eigenvalue of a matrix using Excel, Maple, MATLAB, MATHEMATICA, and Mathcad. In some exceptional cases, useful approximations

BOX 3. THE EFFECT OF SPATIAL ARRANGEMENT OF FIELDS ON R_0

Is the basic reproduction number larger or smaller when fields are clustered in the landscape? We study this question for a slightly simplified example shown in Figure B3.1A, consisting of three fields. To this end we calculate the basic reproduction number as a function of the distance between field 2 and field 1, where we shift field 2 along the dotted line in the figure (so that its distance to fields 1 and 3 varies). The values of the connectivity $c_{12} = c_{21}$ and $c_{23} = c_{32}$ are shown in Figure B3.1B. Figure B3.2 shows R_0 where the distance between field 1 and 2 is shown as fraction, f , of the distance between field 1 and 3.

The results show that increased clustering of fields (increasing closeness of fields 1 and 2 [left side of Fig. B3.2] or fields 2 and 3 [right side of figure]) increases the basic reproduction number. Of course one specific example does not lend itself for generalizations, but a large number of calculations on specific examples can convince us that a pattern is there.

A very insightful approximation to R_0 in this case was derived by Adler and Neuberger (1). Define $S_i = \sum_j c_{ij}$ to be the immigration potential of field i . It can be thought of as the expected number of infectious propagules which would arrive in field i if one dispersing infectious propagule departs from each field in the landscape (1). Isolated fields will have a low immigration potential and well connected fields will have a high immigration potential. They show that R_0 is proportional to $E+(V/E)$, where E is the mean immigration potential of all fields (i.e., expected value

of S_i) and V is the variance of the immigration potential of the fields. This approximation formula shows again that clustering of fields tends to increase the basic reproduction number of a pathogen in a landscape.

This example shows that the methods presented in this paper can be used to study effects of landscape structure on disease dynamics in complex and realistic scenarios.

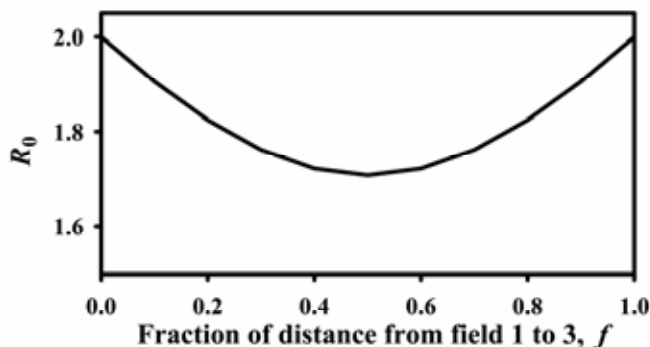


Fig. B3.2. The basic reproduction number, R_0 for the three-field pathosystem, as a function of f as defined in Figure B3.1. The other parameter values are: $d = 0.5$ and $\pi\alpha H = 2$ (equations 9 and 10).

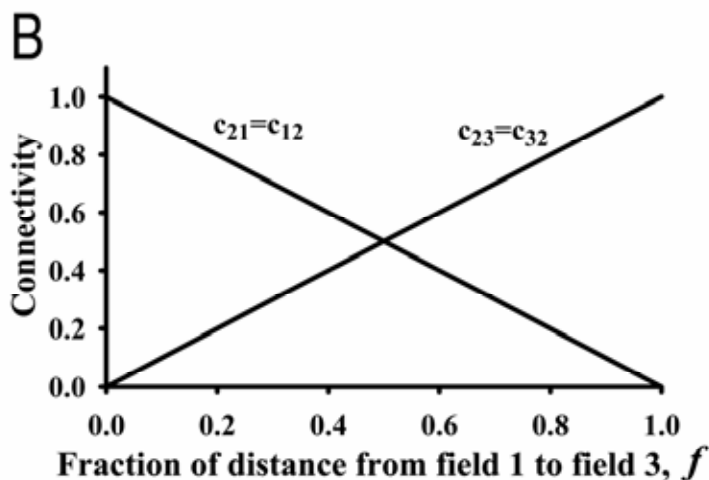
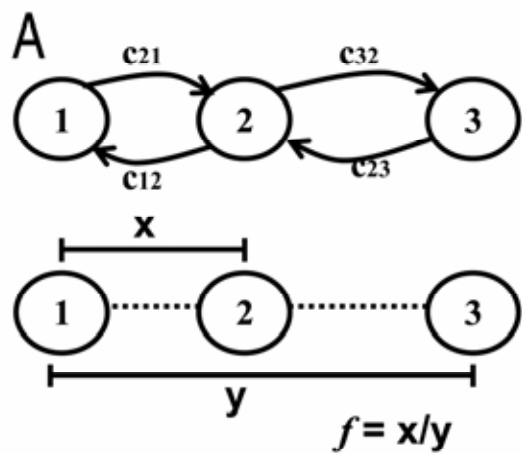


Fig. B3.1. A, The spatial arrangement of fields in the simplified example. The dotted line is the distance over which field 2 is shifted from overlaying field 1 to overlaying field 3. When field 2 is close to field 1 or field 3, clustering of fields is large; when field 2 is midway between the two fields, clustering is minimal. B, The connectivity between fields 1 and 2, and 2 and 3, as function of the position of field 2 along the dotted line shown in graph A. The distance between field 1 and 2 is quantified as the fraction, f , of the distance between field 1 and 3. The lines are $c_{12} = c_{21} = 1 - f$ and $c_{23} = c_{32} = f$.

to the dominant eigenvalue can be developed, as exemplified in Box 3.

EXAMPLE 3: PROPAGATOR TREES AND NURSERIES: COPING WITH TYPE CHANGE

Introduction. We first generalize our terminology and call an I_j plant in examples 1 and 2 ‘an individual of type j ’. That is, in example 1, a type 1 individual is a lesion on a plant of cultivar 1, and in example 2, it is an infected plant in field 1. In both examples infected plants do not change type during their infectious period. An infected plant of cultivar 1 (or in field 1) remains so during its entire infectious period. Obviously there are many circumstances in which infected individuals can change type. For example, garden plants are often grown in a nursery, then distributed to garden centers and subsequently sold for planting in homeowner gardens. A plant that becomes infected in the nursery (a type 1, I_1 , individual) changes type when transported to a garden center (and becomes a type 2, I_2 , individual), and changes type again (into type 3, I_3) when sold to a garden owner.

We introduce a system with one type-change and derive the equations for this specific example. Next we generalize and allow for any type change to occur. We will see that this generalization actually simplifies the model derivation through the introduction of a matrix describing the pathogen transmission rates from type j to type i , and a matrix describing the amount of time an individual infected while type j spends being type i during its infectious period.

The special example. Fruit trees, such as plum, are multiplied by taking cuttings from propagator trees and growing them for some time in a nursery. The majority of the established cuttings are sold to orchard/garden owners, and a small fraction is used to establish new propagator trees. We consider here, as an example, the pathogen *Plum pox virus* (PPV), a member of the genus *Potyvirus*, causing plum pox disease (Sharka) in plum and some other tree species (5,25). This virus is transmitted by aphids and through cuttings taken from infected propagator trees.

Denote the density of infected propagator trees in generation n by $I_{1,n}$ and the density of infected cuttings in the nursery by $I_{2,n}$. The ‘type change’ that occurs in this system is that individuals of type 2 can change into individuals of type 1, when they are used to establish a new propagator tree. The life-cycle components in this pathosystem are (Fig. 2C) as follows.

- Cuttings stay τ_c time units in the nursery, after which a fraction f is used to establish new propagator trees. The remaining fraction, $1 - f$, is sold to orchards/garden owners.
- Before selling or being used as new propagator trees, the cuttings are screened and an infected tree is detected and discarded with probability q .
- A propagator tree is used for τ_p time units and then removed.
- In each time unit, γ cuttings are taken from each propagator tree.
- Before planting in the nursery the cuttings are screened and infected cuttings are detected and discarded with probability p .
- Due to insect-vectored transmission, an infected propagator tree causes $\alpha_1 N_1$ new infected propagator trees per time unit, where N_1 is the density of healthy propagator trees and α_1 is the transmission coefficient.
- Due to vectored transmission, an infected cutting in the nursery causes $\alpha_2 N_2$ new infected cuttings in the nursery per time unit, where N_2 is the density of healthy cuttings in the nursery and α_2 is the transmission coefficient.

We refer to equations 3, 4, and 5 for the model structure, and develop the matrix elements R_{ij} . To this end, we follow a freshly infected propagator tree and calculate how many daughter infections it causes in the field with propagator trees and in the nursery during its entire infectious period; the process is repeated for a freshly infected cutting in the nursery.

- An infected propagator tree causes $\alpha_1 N_1$ new infected propagator trees per time unit and does this for τ_p time units giving $R_{11} = \alpha_1 N_1 \tau_p$. From an infected propagator tree, cuttings are taken (per time unit) and after screening, a fraction $(1 - p)$ is planted in the nursery. This is done for τ_p time units giving $R_{21} = (1 - p)\gamma\tau_p$ for infections in the nursery arising from infected propagator trees.
- An infected cutting in the nursery causes $\alpha_2 N_2$ new infected cuttings per time unit and does this for τ_n time units, giving a total of $\alpha_2 N_2 \tau_c$ daughter infections. In addition, after screening, a fraction, $f(1 - q)$, of the infected cuttings changes type by being used to establish new propagator trees, and from these propagator trees $(1 - p)\gamma\tau_p$ infected cuttings are taken and planted in the nursery. The two sources of infected cuttings together give us $R_{22} = \alpha_2 N_2 \tau_c + f(1 - q)(1 - p)\gamma\tau_p$. Note that the $(1 - p)\gamma\tau_p$ term of R_{22} is the same as R_{21} .
- For the final matrix element, R_{12} , the multiplication factor for infected propagator plants arising from infected nursery plants, again realize that a cutting is used to establish a new propagator tree with probability $f(1 - q)$, and that an infected propagator tree causes $\alpha_1 N_1 \tau_p$ new infected propagator trees. This gives us $R_{12} = f(1 - q) \alpha_1 N_1 \tau_p$.

The matrix A in equation 5 thus has the form

$$A = \begin{pmatrix} \alpha_1 N_1 \tau_p & \alpha_1 N_1 f(1 - q) \tau_p \\ (1 - p) \gamma \tau_p & \alpha_2 N_2 \tau_c + (1 - p) \gamma f(1 - q) \tau_p \end{pmatrix} \quad (11)$$

and the basic reproduction number can be calculated by substituting the matrix elements in equation 6. Box 4 discusses the effect of disease control efforts in the propagator trees and in the nursery on the basic reproduction number.

Generalization. The two matrix components affected by type change of infected cuttings in the nursery into infected propagator trees are R_{12} and R_{22} , the top right and bottom right element of the matrix, respectively. First consider R_{12} . As required, the dimension of R_{12} is number (i.e., number of new infected individuals of type 1 arising from infected individuals of type 2) because the parameter combination $\alpha_1 N_1$ is a rate (number of new infections per time unit) and it is multiplied by $f(1 - q)\tau_p$, which has the dimension of time. What exactly is the interpretation of $f(1 - q)\tau_p$? The term relates to infected cuttings in the nursery that change type and become infected propagator trees. A fraction $f(1 - q)$ of the infected cuttings actually change type, the others, $1 - f(1 - q)$, are removed (sold or detected infected) from the system. One could say that this last group spends 0 time units as an infected propagator tree and the first group spends τ_p time units as infected propagator trees. This means that, on average, an infected cutting in the nursery will change type and spend $0 \cdot (1 - f(1 - q)) + \tau_p f(1 - p) = \tau_p f(1 - q)$ time units as an infected propagator tree. We can thus write

$$R_{12} = \begin{bmatrix} \text{number of new infections} \\ \text{per time unit of type 1 from} \\ \text{an individual of type 1} \end{bmatrix} \bullet \begin{bmatrix} \text{amount of time spent as type 1} \\ \text{when the individual became} \\ \text{infected being type 2} \end{bmatrix} = r_{11} T_{12} \quad (12)$$

The matrix element R_{22} has a similar structure with $\alpha_1 N_1$ and $(1 - p)\gamma$ being rates (vectored transmission between cuttings within the nursery and vertical transmission with cuttings taken from propagator trees) and τ_c and $f(1 - q)\tau_p$ being time units. This term can thus be interpreted as

$$R_{22} = \begin{bmatrix} \text{number of infections per} \\ \text{time unit of type 2 from an} \\ \text{infected individual of type 2} \end{bmatrix} \bullet \begin{bmatrix} \text{amount of time spent as type 2} \\ \text{when the individual became} \\ \text{infected being type 2} \end{bmatrix} + \begin{bmatrix} \text{number of infections per} \\ \text{time unit of type 2 from an} \\ \text{infected individual of type 1} \end{bmatrix} \bullet \begin{bmatrix} \text{amount of time spent as type 1} \\ \text{when the individual became} \\ \text{infected being type 2} \end{bmatrix} = \quad (13)$$

$$= r_{22} T_{22} + r_{21} T_{12}$$

In the most general case, all possible pathogen transmissions between types and all possible type-changes would be possible. In that case, we can define: r_{ij} as the number of infected individuals of type i caused by infected individuals of type j , per time unit; and T_{ij} as the amount of time an individual that became infected being type j spends being infectious as type i . We can then write

$$A = \begin{pmatrix} r_{11}T_{11} + r_{12}T_{21} & r_{11}T_{12} + r_{12}T_{22} \\ r_{21}T_{11} + r_{22}T_{21} & r_{22}T_{22} + r_{21}T_{12} \end{pmatrix} \quad (14)$$

In general, equation 14 is not necessarily directly intuitive, but its structure can be derived by considering each term individually and constructing the equation from the logical interconnection of terms. For the specific example, some of the terms are zero, for example r_{12} equals zero because cuttings in the nursery do not infect propagator trees via vectored disease transmission and T_{21} equals zero because propagator trees do not change type (i.e., they are not transferred to a nursery).

Readers familiar with matrix multiplication will recognize that equation 14 can be written as

$$A = \begin{pmatrix} r_{11} & r_{12} \\ r_{21} & r_{22} \end{pmatrix} \begin{pmatrix} T_{11} & T_{12} \\ T_{21} & T_{22} \end{pmatrix} \quad (15)$$

Equation 15 can also be written in the more compact form $A = RT$, with R and T denoting the matrices of rates and times, respectively. Readers not familiar with matrix multiplication can either take equations 14 and 15 as defining matrix multiplication or consult, for example, the books by Bretcher (3) and Caswell (4) or any other text book about linear algebra. Developing a model to calculate the basic reproduction number is, thus, using equation 15, a matter of finding expressions for all r_{ij} and T_{ij} in terms of the life-cycle components of the pathogen in the pathosystem under consideration. Equations 14 and 15 of course generalize to systems with three or more types.

DISCUSSION

The value of R_0 in understanding and quantifying population dynamic processes is well documented in various fields (13,16,19,27,32). Even though Vanderplank (36,37) nicely related dis-

BOX 4. DISEASE CONTROL IN PROPAGATOR-NURSERY SYSTEMS

In some (many) situations, model parameters are not independent. In particular, in example 3, the conservation of trees stipulates that there is a relation between N_1 , N_2 , γ , τ_c , and τ_p . Consider a system without the pathogen. The N_1 propagator trees produce $N_1\gamma$ cuttings per time unit, which should equal the number of cuttings leaving the nursery per time unit, N_2/τ_c . Further, N_2/τ_c cuttings from the nursery are used each time unit to establish new propagator trees, which should equal the number of propagator trees removed per time unit, N_1/τ_p . Rewriting these consistency relations, we have $N_1 = N_2/(\gamma\tau_c)$ and $f = 1/(\gamma\tau_p)$.

Parameter values: (values are based on 18,19,24,25) Cuttings are grown in the nursery for 2 years, propagator trees are productive for around 20 years, and each produces approximately 100 cuttings per year. We model a producer that sells 1,000 trees from the nursery per year, implying $N_2 = 2,000$ as $\tau_c = 2$. Using the equations above we calculate N_1 and f . We put the screening efficiency parameters, p and q , to zero as default. The vectored transmission rates are assumed to be 0.0002.

Results: Figure B4.1A shows the effect of screening accuracy, p and q , on R_0 . Figure B4.1B shows the effect of vector control, through variation in the transmission rates α_1 and α_2 (new infections per time unit per infected tree) on the basic reproduction number.

duction number. Figure B4.1C shows the effect of roguing (infected tree removal), parameterized as the fractional reduction in lifetime τ_c and τ_p of infected trees, on the basic reproduction number.

Conclusions: (i) Substituting equation 11 in 7, we see that the screening efficiencies p and q only appear in the basic reproduction number formula as product $(1-p)(1-q)$, which implies that they have the same effect on R_0 . This is also shown in Figure B4.1A. Screening all cuttings is likely to be more costly than screening the trees used to establish new propagator trees. This implies that the screening of all cuttings from a propagator tree is less cost effective than the screening of trees to be used as propagator trees. (ii) Figure B4.1B shows that control of the vector population in the nursery (i.e., reducing α_2) has a bigger impact on R_0 than vector control in the propagator-tree orchard (i.e., reducing α_1). (iii) Figure B4.1C shows that roguing to reduce the lifetime of an infected propagator tree, τ_p , has approximately the same effect on R_0 as reducing the lifetime of an infected cutting in the nursery, τ_c , when considered on a fractional basis (a proportion of the lifetimes in the two systems). Given the difference in life span between nurseries and propagator-tree fields, it is more cost effective to rogue out infected propagator trees than infected nursery trees.

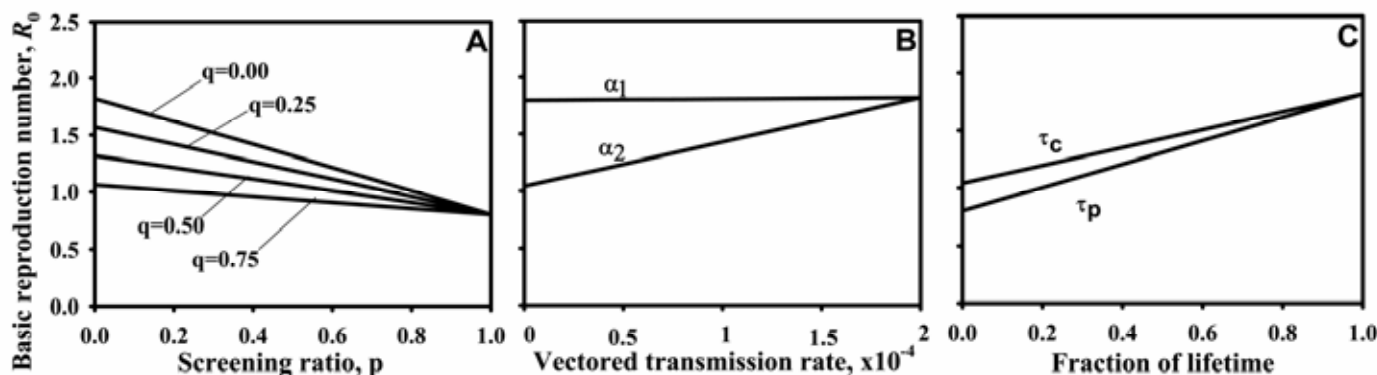


Fig. B4.1. The effect of disease control on the basic reproduction number, R_0 , for the nursery-propagator-tree system (equation 11). **A**, Effect of the probability to recognize and remove an infected cutting during screening, p , before it is planted in the nursery on R_0 . Each line is for a different value of the probability to recognize and remove a full grown tree in the nursery before being used to establish a new propagator tree, q . **B**, Effect of vector control (in terms of the transmission rate) on R_0 . The transmission rate in the group of propagator trees is α_1 and that in the nursery is α_2 . **C**, Effect of roguing of infected plants (sanitation) on R_0 , where roguing reduces the lifetime of the infected tree/cutting. The fraction of the lifetime of a cutting in the nursery, τ_c , and the fraction of the life time of a propagator tree, τ_p , is plotted on the horizontal axis.

ease control strategies to R_0 , utilization of this metric has remained uncommon in plant disease epidemiology, apart from several exceptions. Models formulated in terms of nonlinear differential equations are useful for many purposes, but as demonstrated in Box 1, thresholds derived for epidemics from such models may not, in fact, be equivalent to R_0 . For simple systems (e.g., one host population, one pathogen strain), it is relatively straightforward to derive an expression for R_0 based on first principles (27). However, for more complex systems, such derivations also become complex. The methods originally developed by Diekmann et al. (8) and Heesterbeek (16) provide the framework for dealing with complexity of pathosystems in the derivation of the metric.

In this paper, we built on the work of Diekmann and Heesterbeek to derive R_0 for several complex scenarios of relevance in plant pathology. We need to stress that the methods introduced in this paper are not meant to be a means of simulating epidemic progress curves. The only reason we showed some epidemic progress curves (Figs. 1 and 3), and discussed the curves, was to help make a convincing argument of how to calculate R_0 from a model that has the time step of a pathogen generation. We encourage investigators to apply the matrix methods to obtain R_0 by either using one of the computer programs discussed in the Appendix, or in the 2×2 case, to calculate the largest eigenvalue of the matrix (which then equals R_0), by using equation 7.

In most applications of the methods, the final step will be to plot graphs of R_0 as a function of life-cycle parameters. This can only be done by assigning numerical values to all parameters in the model. This implies that these parameter values need to be estimated from published or unpublished data. While parameter estimation is an important topic and potentially challenging in practice, it is outside the scope of this paper. Some of the methods that are used are introduced in Madden et al. (27), where references for further reading are suggested. Caswell (4) and the literature cited therein provide a good introduction to the issues of estimation for life cycle parameters in matrix models.

Several issues related to the applicability of the methodology in this paper, not discussed so far in an effort to avoid unnecessary complexity, are covered individually below. *First, the methods apply to pathogens with discrete generations as well as to pathogens with overlapping generations.* Although the methods using matrix approaches introduced in this paper use the pathogen generation time as basic time step, they are not only applicable to pathogens with generations that are separated in time (so that at each point in time only one generation of the pathogen is alive or active). The methods are equally well applicable to pathogens where a (large) number of generations occur in overlapping time-windows. This follows from the general theory (8). Recall that R_0 is the number of secondary infections from one mother infection at infinitesimally low pathogen density. At such low densities, so-called density-dependent interactions between pathogen individuals (e.g., lesions) are negligible. Thus, each pathogen individual produces offspring without any interaction with other pathogen individuals, and the sum of the number of offspring produced is not affected by whether the pathogen generations are separated in time or overlapping.

Second, formulation of the matrix on a generation-to-generation basis is essential for proper derivation of R_0 . It is very important to bear in mind that we have been able to equate the largest eigenvalue of the matrix A to R_0 because the time step in the models was the same as the generation time. If a model is constructed that calculates the densities of infected individuals in the various categories using a time step not equal to the pathogen's generation time, the largest eigenvalue of the matrix does measure the growth of the population over the time step, but the eigenvalue is not equal to R_0 .

Third, although matrices have more than one eigenvalue the largest eigenvalue is the basic reproduction number, R_0 . Solutions of models of the form of equations 3 and 4 and 9 and 10 are sums

of terms of the form of equation 6. For example the full solution of equation 3 is

$$\begin{aligned} I_{1n} &= C_1 \lambda_d^n + C_3 \lambda_2^n \\ I_{2n} &= C_2 \lambda_d^n + C_4 \lambda_2^n \end{aligned} \quad (16)$$

where λ_d is the largest eigenvalue of A and is given by equation 7, and λ_2 is the other eigenvalue of this matrix and is given by equation 7 when we replace the + sign in the square root by a - sign. The λ_d components of this solution grow faster than the λ_2 components and therefore for larger values of n the contribution of λ_2 is negligible, leading to the conclusion that λ_d is the generation to generation multiplication factor. The contribution to the solution of the eigenvalue λ_2 causes the deviations at small n as discussed above. (For those familiar with complex eigenvalues: the largest eigenvalue will always be real because the number of offspring cannot be a complex number.)

Fourth, multiple infection transformations are not a component of the basic reproduction number. The above mentioned infinitesimally low pathogen density has another effect on methods simplifying the construction of the model to calculate R_0 . Consider a lesion-forming, spore-producing, fungal leaf pathogen. At infinitesimally low pathogen density, the probability that a spore lands on an existing lesion is negligible, and thus its probability to germinate, infect, and form a lesion is not affected by other lesions. Thus, no corrections are needed for the build-up (increase) in density of infections. Besides the use of R_0 as a generation to generation multiplication factor, it also has some uses in situations where density dependence does play a role. In many models (but not all!), R_0 can be used to determine the ultimate or steady-state value of disease or pathogen density in a host population (10,21,23,26,27,34).

Fifth, initial conditions do not affect R_0 . As we have shown from the numerical approach (Fig. 3), the density of the various types of infected individuals can increase or decrease in the first few generations, depending on the initial conditions. This is because the early population dynamics depend on the initial (i.e., in generation $n = 0$) density of individuals in different categories and on the matrix structure. However, R_0 determines whether in the long run the pathogen density will increase, $R_0 > 1$, or the pathogen will die out, $R_0 < 1$. In other words, R_0 is only dependent on the matrix structure and not on the initial conditions (i.e., R_0 is defined strictly from the matrix of parameters and not from the initial conditions involving infection density).

Sixth, R_0 is a number and not a rate, by definition. However, it might be confusing to see in, for example, Caswell's book that the largest eigenvalue (λ_d) is often called the "population growth rate" (4). The models Caswell discusses usually have time steps of a length other than the generation time of the organism under consideration. In the methods explained in this paper, the time step is one generation and the largest eigenvalue measures the growth of the density of infected individuals per this generation step. This choice of time step provides the direct link between the largest eigenvalue and the generation-to-generation multiplications factor, R_0 (a number).

To conclude, the aim of this paper was to translate some of the existing mathematical framework (8) for the calculation of R_0 into a language palatable to plant epidemiologists. The key reason for this is that using the framework, it is possible to calculate R_0 in a relatively simple way based on a description of the life-cycle components and the interactions with the host in complex situations. Virtually any level of detail deemed necessary for the specific case under consideration can be incorporated in the approach. In contrast, the calculation of R_0 for such complex situations from a model formulated as a set of nonlinear differential (or difference) equations can easily lead to a threshold expression which is not equal to R_0 . Although nonlinear differential equations are invaluable tools for studying epidemics, they should not serve as the primary basis for the determination of R_0 .

APPENDIX

Numerical recipes to calculate the dominant eigenvalue. This appendix provides programming scripts for finding the dominant or maximum eigenvalue of an $n \times n$ matrix and some other software tips for starting to use matrix projection models. Scripts are available for several commonly used programming languages.

The specific numerical values of the matrix entries used in this appendix do not relate to the examples discussed in the main text. They are only given so that the reader can check whether he/she finds the correct eigenvalue when using one of the programs.

Excel. The data analysis tools included in Excel do not include eigenvalue calculations. However, a number of freeware and commercial add-ins to Excel do include these functions. Here we highlight one of these which has been developed specifically to add tools for matrix population modeling to Excel. This is the PopTools add-in developed by Greg Hood at CSIRO, Australia. Instructions for obtaining and installing PopTools can be found at <http://www.cse.csiro.au/poptools/>. Once PopTools is loaded, an extra menu item appears on the Excel menu bar allowing the user to call the various Poptools functions. The steps required to obtain the value of R_0 are as follow.

1. Enter the values of the projection matrix into a suitable $n \times n$ grid of cells.
2. Highlight the matrix.
3. From the PopTools menu choose Matrix tools: eigenanalysis (symmetric).
4. A dialogue box appears with the cell range specifying the matrix already entered.
5. Enter a suitable cell reference for the output.
6. Click Go.
7. The eigenvalues and eigenvectors for the matrix will be entered into the worksheet. The eigenvectors are printed on the left in an $n \times n$ block. The eigenvalues are printed on the right of the output with a different cell background color. The eigenvalues are printed in descending order; the first one is the estimated value of R_0 .

Mathcad template. Although Mathcad includes programming controls, its user interface is in the form of a WYSIWIG mathematics notebook. The steps required to obtain the value of R_0 are as follows.

1. Define the projection matrix, **A**, containing the multiplication and transfer rates.

$$A := \begin{pmatrix} 1.1 & 0.05 \\ 0.05 & 1.075 \end{pmatrix}$$

2. Use the built in "eigenvals" function to get the numerical values of the eigenvalues of **A**.

$$\lambda_A := \text{eigenvals}(A)$$

3. Print the eigenvalues of **A** by using the = key.
4. Mathcad automatically sorts the eigenvalues with the largest first.

$$\lambda_A = \begin{pmatrix} 1.139 \\ 1.036 \end{pmatrix}$$

MATLAB script. The code given below can either be run directly as separate lines or in its entirety when collated into an .m file. Note that the file name cannot contain spaces. The steps required to obtain the value of R_0 are as follows.

1. Define matrix, **A**. The letters within the matrix structure represent the matrix elements, whereby matrix rows are separated by ";".

$$A = \begin{bmatrix} a & b & c & \dots \\ d & e & f & \dots \\ g & h & i & \dots \end{bmatrix};$$

2. Calculate the eigenvalues of **A** and store them in "lambda"
lambda = eig(A)
3. Determine the number of elements in vector "lambda"
[a,b] = size(lambda)
4. Determine which elements of "lambda" are noncomplex and return 1 to "NonComplex" if the element is not complex and 0 when the element is complex
for i = 1:a
NonComplex(i) = isreal(lambda(i))
end
5. Determine the maximum real eigenvalue
MaxLambda = max(NonComplex.*lambda)

ACKNOWLEDGMENTS

F. van den Bosch thanks DEFRA for financial support of the project "Development and testing of a model epidemiological framework to optimize the detection and intervention strategies for plant pathogens of statutory concern." Rothamsted research received support from the Biotechnology and Biological Research Council (BBSRC) of the United Kingdom. F. van den Bosch thanks H. "R₀" Heesterbeek for useful discussions on the basic reproduction number. Scottish Agricultural College received financial support from the Scottish Executive Environment and Rural Affairs Department. Ohio State University received funding from state and federal agencies.

LITERATURE CITED

1. Adler, F. R., and Nuernberger, B. 1994. Persistence in patchy irregular landscapes. *Theor. Popul. Biol.* 45:41-75.
2. Asher, M. J. C. 1993. Rhizomania. Pages 313-346 in: *The Sugar Beet Crop: Science Into Practice*. D. Cooke and R. Scott, eds. Chapman and Hall, London, UK.
3. Bretcher, O. 1997. *Linear Algebra with Applications*. Prentice Hall, Upper Saddle River, NJ.
4. Caswell, H. 2001. *Matrix Population Models*. 2nd ed. Sinauer Associates, Inc. Publishers. Sunderland, MA.
5. Chan, M.-S., and Jeger, M. J. 1994. An analytical model of plant virus disease dynamics with rouging and replanting. *J. Appl. Ecol.* 31:413-427.
6. De Jong, M. C. M., Diekmann, O., and Heesterbeek, J. A. P. 1994. The computation of R_0 for discrete-time epidemic models with dynamic heterogeneity. *Math. Biosci.* 119:97-114.
7. de Koeijer, A., Heesterbeek, H., Schreuder, B., Oberthur, R., Wilesmith, J., van Roermund, H., and de Jong, M. 2004. Quantifying BSE control by calculating the basic reproduction ratio R-0 for the infection among cattle. *J. Math. Biol.* 48:1-22.
8. Diekmann, O., Heesterbeek, J. A. P., and Metz, J. A. J. 1990. On the definition and the computation of the basic reproduction ratio R_0 in models for infectious-diseases in heterogeneous populations. *J. Math. Biol.* 28:365-382.
9. Diekmann, O., Dietz, K., and Heesterbeek, J. A. P. 1991. The basic reproduction ratio for sexually-transmitted diseases. 1. Theoretical considerations. *Math. Biosci.* 107:325-339.
10. Diekmann, O., and Heesterbeek, J. A. P. 2000. *Mathematical Epidemiology of Infectious Diseases. Model Bulding, Analysis and Interpretation*. John Wiley & Sons, Chichester, England.
11. Dietz, K., Heesterbeek, J. A. P., and Tudor, D. W. 1993. The basic reproduction ratio for sexually-transmitted diseases 2. Effects of variable HIV infectivity. *Math. Biosci.* 117:35-47.
12. Edelstein-Keshet, L. 2005. *Mathematical Models in Biology*. Society for Industrial and Applied Mathematics, Philadelphia.
13. Gilligan, C. A. 2002. An epidemiological framework for disease management. *Adv. Bot. Res.* 38:1-64.
14. Gubbins, S., Gilligan, C. A., and Kleczkowski, A. 2000. Thresholds for invasion in plant-parasite systems. *Theor. Popul. Biol.* 57:219-234.
15. Gubbins, S., and Gilligan, C. A. 1999. Invasion thresholds for fungicide resistance. *Proc. R. Soc. Ser. B* 266:2539-2549.
16. Heesterbeek, J. A. P. 2002. A brief history of R-0 and a recipe for its calculation. *Acta Biotheor.* 50:375-376.
17. Heijbroek, W. 1988. Dissemination of rhizomania by soil, beet seeds and stable manure. *Neth. J. Plant Pathol.* 94:9-15.
18. Jackson, D. I., and Looney, N. E. 1999. Temperate and subtropical fruit production. CABI Publishing, Wallingford, UK.
19. Jeger, M. J. 2000. Theory and plant epidemiology. *Plant Pathol.* 49:651-658.

20. Jeger, M. J., Jones, D. G., and Griffiths, E. 1981. Disease progress of non-specialized fungal pathogens in intraspecific mixed stands of cereal cultivars. II. Field experiments. *Ann. Appl. Biol.* 98:199-210.
21. Jeger, M. J., and van den Bosch, F. 1994. Threshold criteria for model plant disease epidemics. I. Asymptotic results. *Phytopathology* 84:24-27.
22. Jeger, M. J., van den Bosch, F., and Dutmer, M. Y. 2002. Modeling plant virus epidemics in a field-nursery system. *IMA J. Math. Appl. Med.* 19:75-94.
23. Kermack, W. O., and McKendrick, A. G. 1927. A contribution to the mathematical theory of epidemics. *Proc. R. Soc. Lond. Ser. A* 115:700-721.
24. Lamb, K., Kelly, J., and Bowbrick, P. 1995. Nursery stock Manual. Growers Manual No. 1. 2nd ed. Grower Books, Nexus Media Limited, Kent.
25. Levy, L., Damsteegt, V., Scorza, R., and Kolber, M. 2000. Plum Pox Potyvirus Disease of Stone Fruits. *APSnet Feature*. Published online by The American Phytopathological Society, St. Paul, MN.
26. Madden, L. V., and van den Bosch, F. 2002. A population-dynamics approach to assess the threat of plant pathogens as biological weapons against annual crops. *BioScience* 52:65-74.
27. Madden, L. V., Hughes, G., and van den Bosch, F. 2007. The Study of Plant Disease Epidemics American Phytopathological Society, St. Paul, MN.
28. Madden, L. V., Jeger, M. J., and van den Bosch, F. 2000. A theoretical assessment of the effects of vector-virus transmission mechanisms on plant virus disease epidemics. *Phytopathology* 90:576-594.
29. Mundt, C. C. 2002. Use of multiline cultivars and cultivar mixtures for disease management. *Annu. Rev. Phytopathol.* 40:381-410.
30. Park, A. W., Gubbins, S., and Gilligan, C. A. 2001. Invasion and persistence of disease in a spatially structured metapopulation. *Oikos* 94:162-174.
31. Richard-Molard, M. 1985. Rhizomania: A world-wide danger to sugar beet. *Span* 28:92-94.
32. Roberts, M. G., and Heesterbeek, J. A. P. 2003. A new method for estimating the effort required to control an infectious disease. *P. Roy. Soc. Lond. B-Bio.* 270:1359-1364.
33. Ross, R. 1911. *The Prevention of Malaria*. 2nd ed. Murray, London.
34. Segarra, J., Jeger, M. J., and van den Bosch, F. 2001. Epidemic dynamics and patterns of plant diseases. *Phytopathology* 91:1001-1010.
35. Stacey, A. J., Truscott, J. E., Asher, M. J. C., and Gilligan, C. A. 2004. A model for the invasion and spread of Rhizomania in the United Kingdom: Implications for disease control strategies. *Phytopathology* 94:209-215.
36. Vanderplank, J. E. 1963. *Plant Diseases: Epidemics and Control*. Academic Press, NY.
37. Vanderplank, J. E. 1975. *Principles of Plant Infection*. Academic Press, New York.
38. Webb, C. R., Gilligan, C. A., and Asher, M. J. C. 1999. A model for the temporal buildup of *Polymyxa betae*. *Phytopathology* 89:30-38.