Key determinants of resistance durability to plant viruses: insights from a model linking within- and between-host dynamics

Frederic Fabre, Claude Bruchou, Alain Palloix, Benoît Moury

Key determinants of resistance durability to plant viruses: insights from a model linking within- and between-host dynamics.

Frédéric Fabre[a,*], Claude Bruchou[b], Alain Palloix[c], Benoît Moury[a]

[a]*INRA, UR 407 Unité Pathologie Végétale, F-84140 Montfavet, France*
[b]*INRA, UR 546 Biostatistique et Processus Spatiaux, F-84914 Avignon, France*
[c]*INRA, UR 1052 Unité Génétique et Amélioration des Fruits et Légumes, F-84140 Montfavet, France*

Abstract

The emergence of new genotypes of parasites involves several evolutionary, epidemiological and ecological processes whose individual effects and interactions are difficult to disentangle using experimental approaches. Here, a model is proposed to investigate how these processes lead to the emergence of plant viral genotypes breaking down qualitative resistance genes. At the individual plant scale, selection, drift and mutation processes shape the evolution of viral populations from a set of differential equations. The spatial segregation of virus genotypes in their hosts is also considered. At the host population scale, the epidemiological dynamics is given by an individual-based algorithm. Global sensitivity analyses allowed ranking the ten demo-genetic and epidemiological parameters of the model according to their impact on the mean and variance of the risk of breakdown of a plant resistance. Demo-genetic parameters (number and nature of mutations involved in breakdown, fitness of mutant genotypes) had the largest impact on the mean breakdown risk, whereas epidemiological parameters had more influence on its standard deviation. It is discussed how these results can be used to choose the potentially most durable resistance genes among a pool of candidates. Finally, our analyses point out the parameters which should be estimated more precisely to improve durability predictions.

Keywords: Breeding; Demo-genetic models; Durable resistance; Emergence; Evolutionary epidemiology; Sensitivity analyses.

*Corresponding author. Tel.: +33 4 32 72 28 47; fax: +33 4 32 72 28 42.
E-mail address: frederic.fabre@avignon.inra.fr

1. Introduction

Ecological factors, including migration, climate and agricultural practices are considered to play an important role in the emergence of both animal (Schrag and Wiener, 1995) and plant diseases (Anderson et al., 2004). In contrast, evolutionary factors, including the adaptation of the pathogen to growth within its hosts and to spread between them, was considered to play a lesser role (Schrag and Wiener, 1995; Anderson et al., 2004). According to Schrag and Wiener (1995), this can be because ecological changes are less constrained than evolutionary ones (e.g. mutations are often associated with decreased fitness). Another reason can be the time sequence of their respective interventions during the 2-step processes leading to a successful emergence: (i) the adaptation and initial 'jump' of the pathogen to a new host population and then (ii) its spread and persistence in this population. In any case, a clear understanding of the relative roles of ecology and evolution during these two steps is needed to predict which pathogens are most likely to emerge (Anderson, 1994).

In agriculture, break down of resistances can be considered to be examples of emergence. Cultivation of resistant varieties, when possible, is the most cost-effective, reliable and environmentally friendly method to control losses due to pests and pathogens. This is especially so with plant viruses for which chemotherapy is not available for field use. However, resistance may be broken down by adapted virus populations. This raises the question of resistance durability, that is, how long a resistance will last after its widespread deployment in environments favouring disease development (Johnson, 1979; Lecoq et al., 2004). For qualitative resistances (i.e. resistances that prevent any systemic plant infection by a pathogen), pathogen evolution must occur in susceptible host population. This kind of resistance is widespread in plants (Thompson and Burdon, 1992), and notably against plant viruses (see genes referred to 'immunity' expression in Table 3 of Garcia-Arenal and McDonald (2003)). It exhibits a large range of durability, from very low levels (e.g. the *Tm-1* gene in tomato to *Tomato mosaic virus* (ToMV), genus *Tobamovirus*, family *Tobamoviridae*) to very high levels (e.g. the *Tm-2* gene in tomato and the *N*-gene in tobacco to tobamoviruses) (Parlevliet, 2002). A virus must complete three successive stages to break down such a resistance. (i) At the scale of the cells of a susceptible host, according to the gene-for-gene mechanism (Flor, 1971), mutation and/or recombination events should appear in the avirulence gene of a virus to generate virulent variants. We define 'virulence' as the genetic ability of a pathogen to overcome genetically determined resistance and to cause a compatible interaction leading to disease (Shaner et al., 1992; Sacristán and García-Arenal, 2008). (ii) At the host scale, the virulent variants generated must be sufficiently fit, compared to wild type, to invade (through cell-to-cell and systemic movement) their host and increase their frequency. (iii) At the agro-ecosystem scale, the virulent variants should be transmitted to resistant hosts. The greater the frequency of virulent viruses in susceptible hosts, the greater their chance of being transmitted to resistant ones.

Here, we first introduce a model of this scenario. This nested model links within- and between-host levels of disease dynamics (Mideo et al., 2008). It addresses the features of plant virus population dynamics within plants and those of plant virus epidemiology between plants. All the parameters of this model have biological meaning. Then, analyses aiming to rank the relative importance of the epidemiological, genetic and evolutionary factors involved in the emergence of virulent viruses and the breakdown of qualitative resistance are presented. In addition to their basic interest for acquiring a better understanding of emergence, these results also have an applied interest. Indeed, the definition of resistance durability introduced by Johnson (1979) is a measure a posteriori (e.g. the time necessary to breakdown a resistance after its deployment). This is a problem for breeders who need predictive criteria to identify durable resistance genes 'upstream' in a breeding programme because of the scarcity of sources of resistance genes and the long and costly breeding processes involved. For virus resistances, the first predictor of durability, termed the 'evolutionary potential' of pathogens, was proposed by García-Arenal and McDonald (2003). It focused on processes that govern pathogen population evolution (mainly effective population size, migration and reproduction system). Their approach, which does not account for the molecular events involved in resistance breakdown and the fitness cost associated with virulence, cannot explain why different resistance genes directed toward the same pathogen can display various levels of durability (e.g. *Tm-1* versus *Tm-2* resistance gene to ToMV). Here, we propose another approach to derive a complementary predictive criterion of

durability that is based on the combination of the main factors involved in resistance breakdown weighted by their relative importance.

2. Model description

The model describes the demo-genetic dynamics of a viral population composed of several variants during an epidemic at both the within-plant and between-plant scales. In this section, we introduce successively three sub-models describing: (i) the viral population involved in the breakdown of a qualitative resistance at the cell scale, (ii) the demo-genetic dynamics of this viral population within a host and (iii) the between-host dynamics of the epidemic.

2.1. The viral population involved in resistance breakdown

2.1.1. Genetic structure of the viral population

Consider a situation where the breakdown of a qualitative resistance requires the accumulation of m amino acid substitutions. These m substitutions are conferred by a corresponding number of nonsynonymous point mutations in the viral genome. Among these m mutations, p ($0 \leq p \leq m$) are transitions (i.e. interchanges of purines or of pyrimidines) and m-p are transversions (i.e. interchanges between purine and pyrimidine bases).

At the m nucleotide sites, each genotype may or may not have the nucleotide required for virulence. Mutations other than those directly implicated in virulence are not considered. The virulent genotype is m-point mutations away from the wild type, and for each class of k-point mutants, there are $\binom{m}{k}$ genotypes. For instance, if we consider m=3, then there are three 1-point mutants (001, 010, 100) and three 2-point mutants (011, 101, 110) plus the wild-type genotype (000) and the virulent one (111). N, the total number of genotypes in the virus population considered equals 2^m. This genetic structure is generic for haploid organisms no matter the 'types' of virus genomes (DNA or RNA, single-stranded (ss) or double-stranded (ds), mono- or multipartite).

2.1.2. Mutation rate between genotypes

Let μ_{ij} ($0 \leq i \leq j \leq N$) be the probability matrix of the mutation of genotype j into genotype i during viral replication. Let μ_b be the point mutation rate per replication cycle and per nucleotide in the case of transition and $\mu_b.\kappa^{-1}$ in the case of transversion where κ is the transition to transversion rate ratio. Then, assuming that mutations are independent between nucleotide sites, $\mu_{i,j} = \mu_b^{\Delta ts(i,j)}.[\mu_b.\kappa^{-1}]^{\Delta tv(i,j)}$ for $i \neq j$ and $\mu_{i,i}=1-\Sigma_{i \neq j}\ \mu_{i,j}$ where $\Delta ts(i,j)$ (resp. $\Delta tv(i,j)$) is the number of transitions (resp. transversions) differing between genotype i and j.

2.1.3. Fitness of the genotypes

The fitness of a virus genotype is defined by its rate of offspring production. The fitness of each genotype is defined assuming that (i) the wild type has a growth rate r, (ii) each mutation has a fitness cost C_{mut} (here it is assumed that it is the same cost for all mutations) in plants devoid of the resistance allele and (iii) mutations have independent effects on fitness (i.e. no epistasis among mutations). In a susceptible plant genotype, r_i, the growth rate of genotype i is $r.(1-C_{mut})^{\Delta mut}$ where Δmut is the number of mutations differing between genotype i and the wild-type. In a resistant plant genotype, all genotypes have a null fitness except the virulent one whose fitness is r.

2.2. Model of demo-genetic dynamics of the viral population in a host plant

In the model, (i) the host plant is viewed as a set of patches among which viral genotypes can be more or less spatially separated and (ii) a system of N ordinary differential equations (ODE) simulates the dynamics of the mean number of virus particles of genotype i at time t per patch, $V_i(t)$:

$$\frac{dV_i(t)}{dt} = r_i \xi V_i(t)\left[1 - \frac{1}{K}\left(\Delta_i(t) + 1 + \sum_{j \neq i} \beta_{ij}\Theta_{ij}(t)V_j(t)\right)\right] + \sum_{j=1}^{N} \mu_{ij}\xi(V_j(t) - V_i(t)) \qquad (1)$$

where (i) r_i is the growth rate of viral genotype *i*,

(ii) ξ is the daily number of virus generations ($1/\xi$ is the generation time)

(*iii*) *K* is the carrying capacity of a patch,

(iv) $\Delta_i(t)$ is the mean crowding of genotype *i* as defined by Lloyd (1967). For a given virus particle, this is the expected number of other particles of the same genotype in a patch. In a patchy system, the intensity of intra-genotype competitions (i.e. the average per capita inhibition) is a function of $\Delta_i(t) + 1$ instead of simply $V_i(t)$ for a non-patchy system (Kuno, 1988).

(v) $\beta_{i,j}$ is the Lotka-Volterra competition coefficient. It is the intensity of inter-genotype competition exerted by genotype *j* on genotype *i* and defined, following Solé et al. (1999), as the relative fitness of genotype *j* with respect to genotype *i* (i.e. $\beta_{i,j}=r_j/r_i$).

(vi) $\Theta_{i,j}(t)$ is a symmetric index ($\Theta_{i,j}(t)=\Theta_{j,i}(t)$) describing how many times more (or fewer) virus particles of genotype *i* can expect to share a patch with those of genotype *j*, relatively to the situation where genotypes *i* and *j* are randomly associated. This index adjusts the intensity of the inter-genotype competition according to their distribution in a patchy system.

(vii) $\mu_{i,j}$ is the probability of mutation of genotype *j* into genotype *i*.

The first term of equation (1) is a generalisation of Lotka-Volterra's competition equations for *N* competitors (Bulmer, 1994; Miralles et al., 2001) in a patchy system (Hartley and Shorrocks, 2002). The second term describes the mutation processes occurring between the *N* genotypes.

Values of $\Delta_i(t)$ and $\Theta_{i,j}(t)$ depend on the distribution of virus genotypes between patches. Ives (1988) showed that $\Delta_i(t) = E[y_i(t)] + \frac{Var[y_i(t)]}{E[y_i(t)]} - 1$ and $\Theta_{i,j}(t) = 1 + \frac{Cov[y_i(t), y_j(t)]}{E[y_i(t)].E[y_j(t)]}$

where $y(t)=[y_1(t), y_2(t), \ldots, y_N(t)]$ is a random variable of the number of individuals of each genotype present in a patch at time *t*. In the model, we assume that $y(t)$ possess a Dirichlet-Multinomial (DM) distribution (see Annex 1) with parameters α, $Y(t) = \sum_{i=1}^{N} V_i(t)$, and $\theta_i(t)=V_i(t)/Y(t)$, the mean frequencies of genotype *i* at time *t*. Thus $\Delta_i(t) = V_i(t) + \left(\frac{Y(t)+\alpha}{1+\alpha}\right)\left(1 - \frac{V_i(t)}{Y(t)}\right) - 1$ and $\Theta_{i,j}(t) = 1 - \frac{Y(t)+\alpha}{(1+\alpha).Y(t)}$.

In the DM distribution, parameter α defines the degree of genotype segregation among patches (see Annex 1). The smaller the value of α, the larger the segregation between genotypes, and vice versa. For a given value of α, the degree of genotype segregation depends on *N*. In order to have a degree of genotype segregation independent of *N*, segregation was controlled in the model with the parameter φ ($0 \leq \varphi \leq 1$) as follows: φ is defined such that in 90% of the patches one randomly chosen genotype has a frequency greater than $\frac{1}{N} + \varphi.(0.9 - \frac{1}{N})$, whereas in the plant the overall mean frequencies of all the genotypes remained equal to $1/N$. When φ=0, the genotypes are mixed homogenously: the frequencies of the genotypes are all equal to $1/N$ in all patches. When φ=1, genotypes are highly segregated: in 90% of the patches, one genotype has a frequency exceeding 90%.

In simulations φ and *N* are set first, then a dichotomic search algorithm is used to determine $\alpha(\varphi,N)$ and finally, given $\alpha(\varphi,N)$, r_i, ξ, *K* and $\mu_{i,j}$, the system of ODE is used to simulate $V_i(t)=E(y_i(t)|Y(t))$ in a plant.

2.3. Model for the spread of the viruses between host plants

The epidemiological dynamic is modelled with an Individual-Based Model (IBM) which offers a powerful and flexible way to bridge the gap between individual host and host population scales (DeAngelis and Mooij, 2005). Host plants are the cells of an $n \times n$ matrix. At all times, each plant is labelled as healthy (H) or infected (I) regarding the epidemiological dynamics and as susceptible (S) or resistant (R) regarding their genotype.

At initial time (t=0), all plants are healthy except one. That susceptible plant (genotype S), is infected with λ_{trans} viruses of the wild-type, λ_{trans} being the mean number of virus particles transmitted by a vector to a plant. Therefore, in all simulations, neither the virulent genotype nor the intermediate point mutants occur in the viral population at t=0. Next, the epidemic process runs from t=0 to t=t_{end} and, during a δt time step, the dynamic proceeds in three stages.

Stage 1. At time t, an infected plant k ($1 \leq k \leq n^2$) has a mean number of contacts $P_c(k).\delta t$ with other plants assuming that $P_c(k) = \beta.\left[1 - \exp(-\frac{\Delta t_{inf}(k)}{\tau})\right]$ where (i) β is the contact rate between plants, (ii) $\Delta t_{inf}(k)$ is the time lag between time t and the time at which plant k became infected, and (iii) τ is the mean duration of the latency (i.e. time necessary for an infected plant to become infectious). Underlying this equation is the hypothesis that, at each δt, an infected plant is transmitting (through vector movements not explicitly modelled) some of its viruses at a constant rate β. This rate is modulated by the probability that a vector acquires virus particles, a probability which increases with virus concentration in the plant and thus with time elapsed from infection. The number of contacts of plant k during δt is drawn from a Poisson distribution with parameter $P_c(k).\delta t$. Then, for each efficient contact, steps 2 and 3 are performed.

Stage 2. The total number and composition of the viral population acquired by a vector feeding on a randomly chosen patch of plant k is simulated. First, N_{trans}, the total number of virus particles transmitted, is drawn from a Poisson distribution with parameter λ_{trans}. Second, the number of virus particles of each genotype is drawn from plant k in a DM distribution with parameters $\alpha(\varphi,N)$, $[V_{k,1}(t), V_{k,2}(t),\ldots, V_{k,N}(t)]/Y_k(t)$ and N_{trans}, where $V_{k,i}(t)$ is the overall mean number of virus particles of genotype i at time t in plant k (NB: in the previous part, the subscript k was omitted in ODE equations to avoid overloading the notations).

Stage 3. The viral population sampled in stage 2 is transmitted to a randomly selected plant. When this plant is healthy and its genotype is susceptible, it becomes infected and the demo-genetic of the viral population in this newly infected host is simulated from t to t_{end} with the ODE system set with initial conditions equal to the result of the DM sampling. When this plant is healthy and its genotype is resistant, it becomes infected only if the DM sampling contains at least one virulent virus particle. Initial conditions of the ODE system are set accordingly and the system evolution is simulated from t to t_{end}.

2.4. Global sensitivity analysis of the risk of breakdown of a qualitative resistance

Global sensitivity analysis (GSA) techniques (Saltelli et al., 2008) quantify the relative importance of input factors (i.e. model parameters) by partitioning the variance of an output of interest into that due to main effects of model parameters and that due to their higher order interactions. Here, the output Ψ of interest, termed the durability index, was defined as $\Psi = \log\left(\sum_{k\in\Omega}\int_{t=0}^{t_{end}} \frac{V_{k,N}(u)}{Y_k(u)}du\right)$, where Ω is the subset of susceptible plants in the host population and $V_{k,N}(t)/Y_k(t)$ is the probability to sample a virulent virus in plant k at time t. Ψ measures the cumulative probability of randomly sampling a virulent virus in a field of susceptible plants during the entire duration of an epidemic. Evidence is provided that Ψ is a good proxy of the risk of breakdown of a qualitative resistance in Results section 3.1.

The model being stochastic, GSA were conducted on the mean and standard deviation of Ψ (noted E(Ψ) and SD(Ψ) respectively). An analysis of variance (ANOVA)-based GSA was preferred

in this work after checking that its results are consistent with those derived with Sobol-based GSA (see Annex 2), a more general but highly computational intensive method (Saltelli et al., 2008).

Step 1. A probability distribution must be assigned to each parameter of interest. Ten parameters were included in the analysis (Table 1). As only ranges of variation accounting for the natural variability of the considered pathosystems are known, uniform or log-uniform distributions were assigned to the ten parameters (Table 1). Ranges for μ_b, and λ_{trans} are derived from data acquired on ssRNA viruses transmitted in a non-persistent manner. Many situations have been observed where only one or two mutations are required to obtain a virulent virus variant (Harrison, 2002): m and p were set accordingly. As previously, to date, these data are only available for ssRNA viruses. Ranges of κ, φ, ξ and C_{mut} considered a wide range of possibilities as well as those of the epidemiological parameters τ and β (Table 1). The range of β encompasses epidemics that attain an incidence of only 8% of infected plants at t_{end} to epidemics with all plants infected within one month.

Step 2. The model is run for a set of samples of the parameter space in a field of susceptible plants (n^2=400). Because we are interested in deriving GSA results for E(Ψ) and SD(Ψ), 1000 replications were performed for each sample of the parameter space. The sampling scheme used was a fractional factorial design in which a common number of fixed levels was chosen equally distributed over the range of each parameter. Ranges for m and p ($0 \leq p \leq m \leq 2$) necessitate a design with five levels. In a full factorial design all combinations of levels are used. But, assuming that higher-order interactions are negligible, the fractional design requires a relatively small number of model runs (Ginot et al., 2006; Saltelli et al., 2008) without loss of information on main effects and lower order interactions. An important characteristic of such design is their resolution. A V resolution fractional design was chosen to avoid confounding main effects and 2-way interactions. That design with 10 factors and 5 levels per factor required 3125 parameter combinations and, moreover, allows dealing with the factor p nested within m.

Step 3. Sensitivity indices are estimated from the simulation results. Two sensitivity indices are of particular interest. First order (or main) indices estimate the average influence of the factor alone and should be used for ranking purposes. Total order indices estimate the influence of a factor alone, or involved in interactions with all others factors and should be used to set non-influential factors. In ANOVA-based GSA, decomposition of variance derived from the classical analysis of variance table was used to compute indices as the part of variance explained by a factor (or its second order interactions) relative to the total variance.

2.5. One-at-a-time analysis of the effect of single parameters on the risk of breakdown

One-at-a-time analyses were realized to investigate more precisely how the main factors revealed by GSA impact on the durability index. This analysis involves plotting the relationship between the output of interest versus a set of values of a given model parameter, all others parameters being fixed to their reference values (Table 1). For continuous parameters, E(Ψ) and SD(Ψ) were assessed for 50 values regularly sampled in the corresponding probability distribution (Table 1), 1000 replicates being done for each parameter value. The effects of discrete parameters (m and p) were explored similarly but all their possible values were used. For E(Ψ), to simplify comparison between sets of model parameters, a mean relative durability index E(Ψ_r) was estimated by comparison with a situation where one transition with no fitness cost was required for virulence ($E[\Psi_r(C_{mut},m,p)] = E[\Psi(C_{mut},m,p)] - E[\Psi(C_{mut}=0,m=1,p=1)]$).

2.6. Model implementation and statistical analyses

The model was implemented in Fortran 90 and compiled with Intel® Fortran Compiler 9.1. Random number generator and routines for sampling in the probability density functions were from Intel® Math Kernel Library 8.0. The module RK_Suite90 (Brankin et al., 1992) was used to solve ODE systems. Parameters not included in the GSA were set to 10^6 for K, 120 days for t_{end} and 1 for r.

Sensitivity analyses were done using the package "Sensitivity" of the R software environment (http://cran.r-project.org/) and a Quasi Monte Carlo sequence generated with the package "fOptions".

Factorial simulated designs were generated with the factex procedure of SAS (vers. 8, SAS Inc., Cary, NC).

Before using Ψ as a proxy for resistance durability, it is checked, by running simulations in a field composed of equal numbers of susceptible and resistant plants (n^2=900), that Ψ is highly correlated with the risk of breakdown of a qualitative resistance. Then, all others simulations done for GSA and one-at-a-time analyses were run in field composed of only susceptible plants (n^2=400).

Table 1: Probability density functions used to characterize the range of variation of the ten parameters included in global sensitivity analyses.

Parameter	Parameter designation (unit)	PDF[a]	Reference value	Reference
μ_b	Mutation rate (nucleotide^{-1}.generation^{-1})	Log-U [10^{-4}, 10^{-6}]	10^{-5}	Drake and Holland, 1999 Malpica et al., 2002
κ	Transition/transversion rate ratio	U [2, 10]	7	Yang et al., 2000 Moury et al., 2002
C_{mut}	Relative fitness cost (in terms of replicative capacity) of a mutation (mutation^{-1})	Log-U [0, 0.8]	0.05	Jenner et al., 2002 Desbiez et al., 2003 Ayme et al., 2006
m	Number of mutations required for virulence	U_d [1, 2]	1	Harrison, 2002 Kang et al., 2005
p	Number of transitions required for virulence	U_d [0, m]	1	
λ_{trans}	Mean number of virus particles transmitted by vectors	Log-U [1, 1000]	10	Ali et al., 2006 Moury et al., *2007*
β	Infection rate (plants.day^{-1})	U [0.03, 0.25]	0.14	
τ	Mean length of the latent period (day^{-1})	U [2, 12]	7	
φ	Intensity of spatial segregation of virus variants	U [0, 1]	0	Dietrich and Maiss, 2003 Jridi et al., 2006 Takahashi et al., 2007
ξ	Daily number of virus generations (generation.day^{-1})	U [2, 24]	12	Cuevas et al., 2005[b]

[a] PDF: Probability Density Function. U: Uniform distribution. U_d: Discrete uniform distribution. Log-U: Log-uniform distribution. [b] Information about plant virus generation times are still lacking (Pita and Roossinck, 2008). Cuevas et al. (2005) provide an estimate of the product $r\xi$ for an animal RNA virus, the *Vesicular stomatitis virus*.

3. Results

3.1. The durability index is highly correlated with the risk of breakdown

The breakdown of a qualitative resistance is a random process involving the elementary event VrT '*a virulent virus is acquired on a susceptible plant and then transmitted to a resistant plant*'. A resistance is broken down, event RB, if '*VrT is observed at least one time during an epidemic*'. Setting a field with even proportions of susceptible and resistant plants, we investigated the relationship between the durability index Ψ and (i) p(RB), the probability of RB and (ii), for p(RB)>0, the number of times VrT was observed. The model was run for 2000 random draws of each of the 10 parameters (sampled from the distributions of Table 1). One thousand replications were performed for each random draw, p(RB) being estimated as the proportion of replicates where a breakdown was observed. Results are displayed in Figure 1. The mean of Ψ highly correlated both with p(RB) (Mac Fadden r^2=0.94 ; Fig. 1A) and with the mean number of times VrT is observed (r^2= 0.96; Fig. 1B).

Additional simulations were realized to estimate the waiting time of the first VrT event by relaxing the constraint on t_{end}. In order to have a reasonable duration of simulation, m was set to 1 and the model run for 1000 random draws of the nine remaining parameters, 100 replications being done. Results indicate that E(Ψ) is also highly correlated to the time necessary for the first breakdown of resistance (r^2= 0.94). In all, Ψ provides a good proxy of the probability, intensity and waiting time of

the epidemiological processes leading to resistance breakdown. It is achieved at a lower computational cost than simulations involving resistant plants.

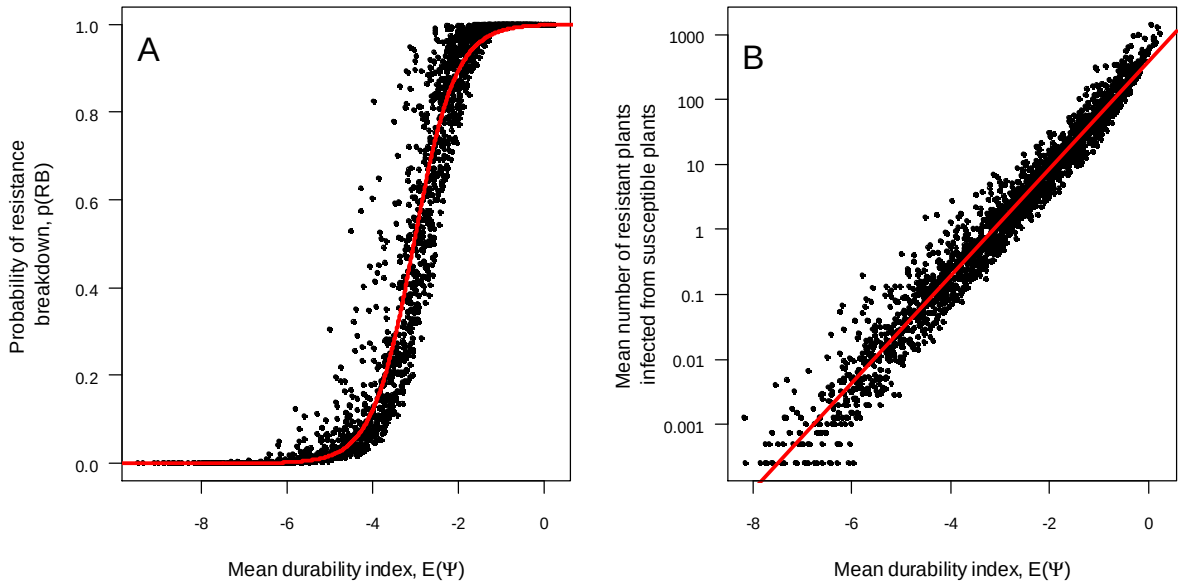


Fig. 1. A. Relationship between the mean durability index, E(Ψ), and the probability of resistance breaking, *p*(RB). The Mac Fadden coefficient of determination is 0.94. **B.** Relationship between E(Ψ) and, for the case where *p*(RB)>0, the mean number of times a virulent virus is acquired on a susceptible plant and then transmitted to a resistant one (event VrT). The coefficient of determination is 0.96. For both plots, simulations were done assuming a field of 900 plants of which half were susceptible and half resistant. The model was run for 2000 random draws of all model parameters (sampled in the distribution of Table 1) and 1000 replicates done for each draw. The probability *p*(RB) is estimated as the proportion of replicates where a breakdown was observed.

3.2. Major determinants of the risk of breakdown (Global sensitivity analyses)

The fit of the ANOVA model being high (r^2=0.99 for E(Ψ) and r^2=0.94 for SD(Ψ)), sensitivity indices could be derived appropriately and used to reveal key determinants of the durability index.

For E(Ψ), main effect indices alone explained 94% (sum of first-order indices) of the overall variance: factors have thus mainly additive effects (Fig. 2A). Results for E(Ψ) (Fig. 2A and 2B), showed that three parameters were pre-eminent. The number of mutations required for virulence (m) was the most influential factor, alone accounting for 41% of the variance of the mean durability index, and 45% when adding the nature of mutations (p). The next two factors, the mutation rate (μ_b) and the relative fitness cost (C_{mut}), accounted for 21% and 17% of the variance, respectively. In all, the main effects of these four factors (m, p, μ_b and C_{mut}) accounted for 83% of the variance. The infection rate β was the least factor to have a noticeable effect (7%). The effects of the five other factors (φ, intensity of genotypes segregation; ξ, daily number of virus generations; κ, transition/transversion rate ratio; τ, mean length of the latent period and λ_{trans}, mean number of virus particles transmitted by vectors) were negligible.

Regarding SD(Ψ), first-order indices indicate that three parameters are very important. The first is the infection rate β (37% of explained variance), followed by C_{mut} (27%) and by φ, the intensity of genotype segregation (8%) (Fig. 2D). The seven remaining factors have sensitivity indices below 2%. The analysis also highlighted the importance of some interactions (main effect indices alone explained 78% of the overall variance). Two interactions appeared of significant importance: C_{mut} x β (5% of explained variance) and C_{mut} x φ (5% of explained variance).

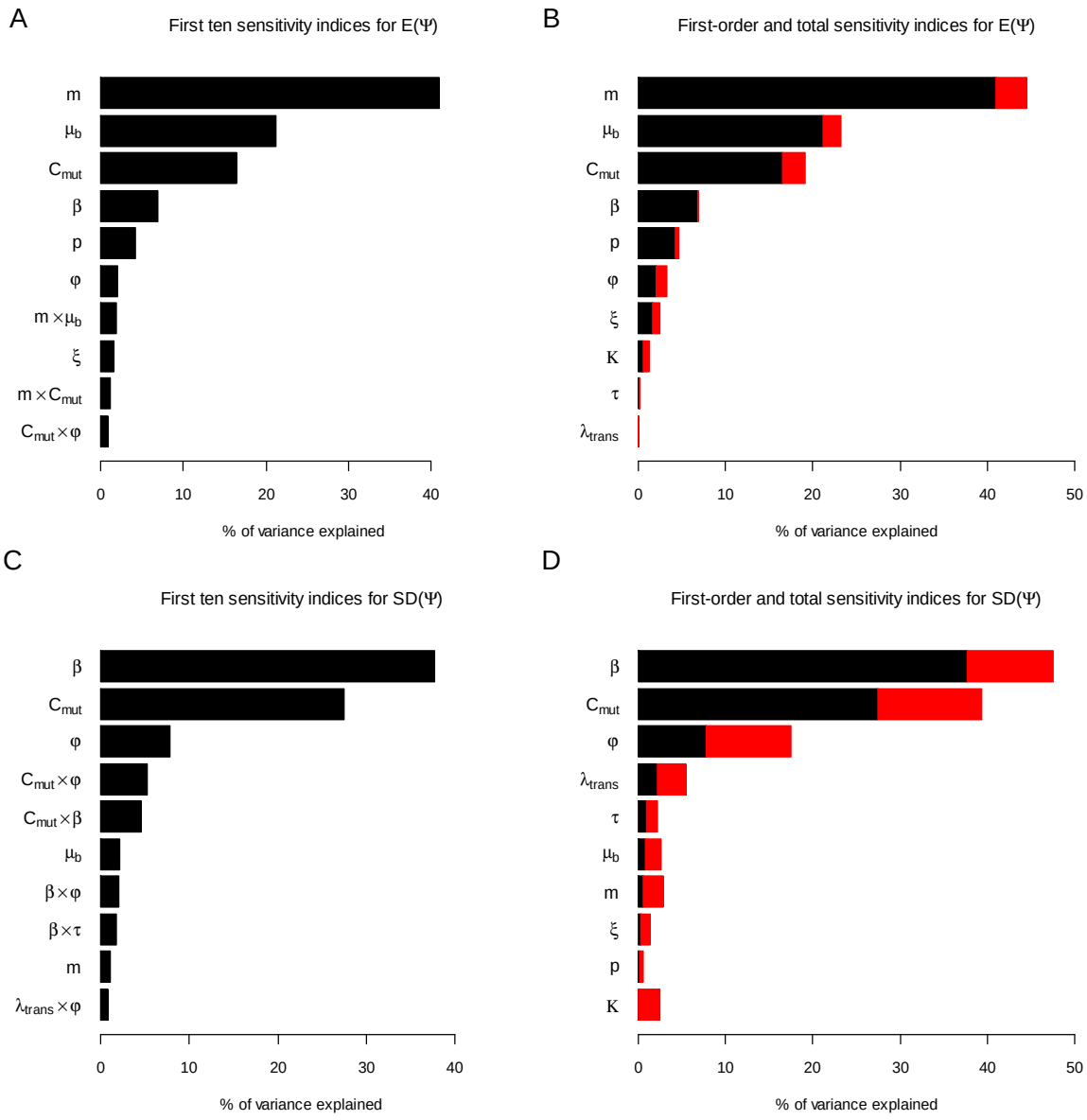


Fig. 2. Sensitivity indices of the durability index derived from an ANOVA on a fractional factorial design of resolution V. The ten factors analysed are: μ_b :mutation rate, κ: transition/transversion rate ratio, C_{mut}: relative fitness cost, m: number of mutations required for virulence, p: number of transitions (among the m mutations) required for virulence, λ_{trans}: mean number of virus particles transmitted by vectors, β: infection rate, τ: mean length of the latent period, φ: intensity of genotypes spatial segregation and ξ: daily number of virus generations. The explained variance was 99% for E(Ψ) and 94% for SD(Ψ) with an ANOVA model including second-order interactions **A**: First ten sensitivity indices for E(Ψ). **B**: Main and total sensitivity indices for E(Ψ). **C**: First ten sensitivity indices for SD(Ψ). **D**: Main and total sensitivity indices for E(Ψ). In graphs **B** and **D**, the first part of bars (black) correspond to main indices (effect of the factor alone) and full bars (black and red parts) correspond to total indices (red parts correspond to the effect of the factor in interaction with all others factors).

3.3. Effects of the main factors on the breakdown risk (One-at-a-time analyses)

We first explored the relationships between E(Ψ) and three important factors revealed by GSA, m, p, and C_{mut}. By setting to 1 the mean durability index of a resistance imposing one transition with no fitness cost for virulence, a mean relative durability index E(Ψ_r) was assessed as a function of

C_{mut} for six genetic combinations of m and p (one transition, one transversion, two transitions, one transition + one transversion, two transversions and three transitions). $E(\Psi_r)$ quantifies how many times lower is the durability index for resistance genes requiring others values of m, p, and C_{mut}. Results are displayed in Figure 3.

Parameters m and p clearly impact on $E(\Psi_r)$: for example, with very low fitness cost (e.g. 10^{-4}), the relative mean durability index is 2 (resp. 4) orders lower for a determinant of virulence requiring two (resp. three) transitions compared to a situation where only one transition is needed. This difference increases with C_{mut}: with a fitness cost of 0.8, the relative index is 4 orders lower for a 2-transition situation compared to a 1-transition one. The effect of C_{mut} on $E(\Psi_r)$ was also clear: for a situation where one transition is required for a breakdown of resistance, the relative durability index is decreased by 3 orders when the fitness cost increases from 10^{-4} to 0.8. Finally, the C_{mut} effect on $E(\Psi_r)$ was not linear: $E(\Psi_r)$ remained quite constant for slightly deleterious mutations ($C_{mut} < 10^{-3}$) and then decreased linearly (on a log-scale) for higher fitness costs.

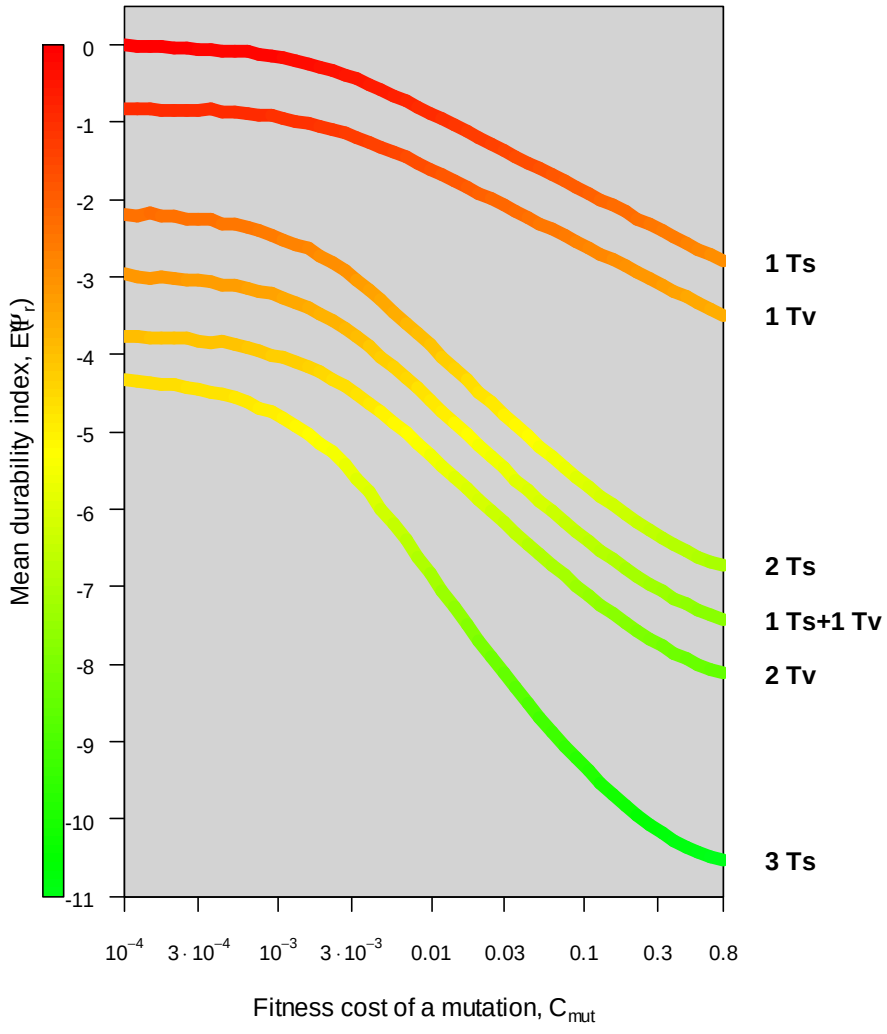


Fig. 3. Mean relative durability index as a function of the fitness cost of a mutation (C_{mut}) for six determinants of virulence (i.e. combination of m and p): (i) one transition (1 Ts), (ii) one transversion (1 Tv), (iii) two transitions (2 Ts), (iv) one transition and one transversion (1 Ts + 1 Tv), (v) two transversions (2 Tv) and (vi) three transitions (3 Ts). The mean relative durability index, $E(\Psi_r)$, is a comparison of $E(\Psi)$ to $E(\Psi)$ for a situation where one transition with no fitness cost was required for virulence (i.e. $E(\Psi_r) = E[\Psi(C_{mut},m,p))] - E[\Psi(C_{mut}=0,m=1,p=1))]$). Other parameters were set to their reference values (Table 1).

We then explored the relationships between SD(Ψ) and C_{mut}, β and φ, the three most important factors according to the GSA results. Results are displayed in Figure 4. C_{mut} and φ have both monotonic, but opposite and non-linear, effects on SD(Ψ). The standard deviation of the durability index decreases with increasing fitness cost (C_{mut}) and increases with increasing intensity of viral genotype segregation (φ). Moreover, for both parameters, an asymptote is reached for $C_{mut} > 0.01$ and $\varphi > 0.5$. The figure is different for the infection rate β. SD(Ψ) is increasing with β in the range [0.03, 0.07], and then, for $\beta > 0.07$ SD(Ψ) decreased. However, no asymptote is reached in the range of variation explored. For $\beta = 0.07$, the mean proportion of infected susceptible plants at t_{end} is 0.5 and thus the standard deviation of the proportion of infected plants is the highest (this is so for binomial-like processes).

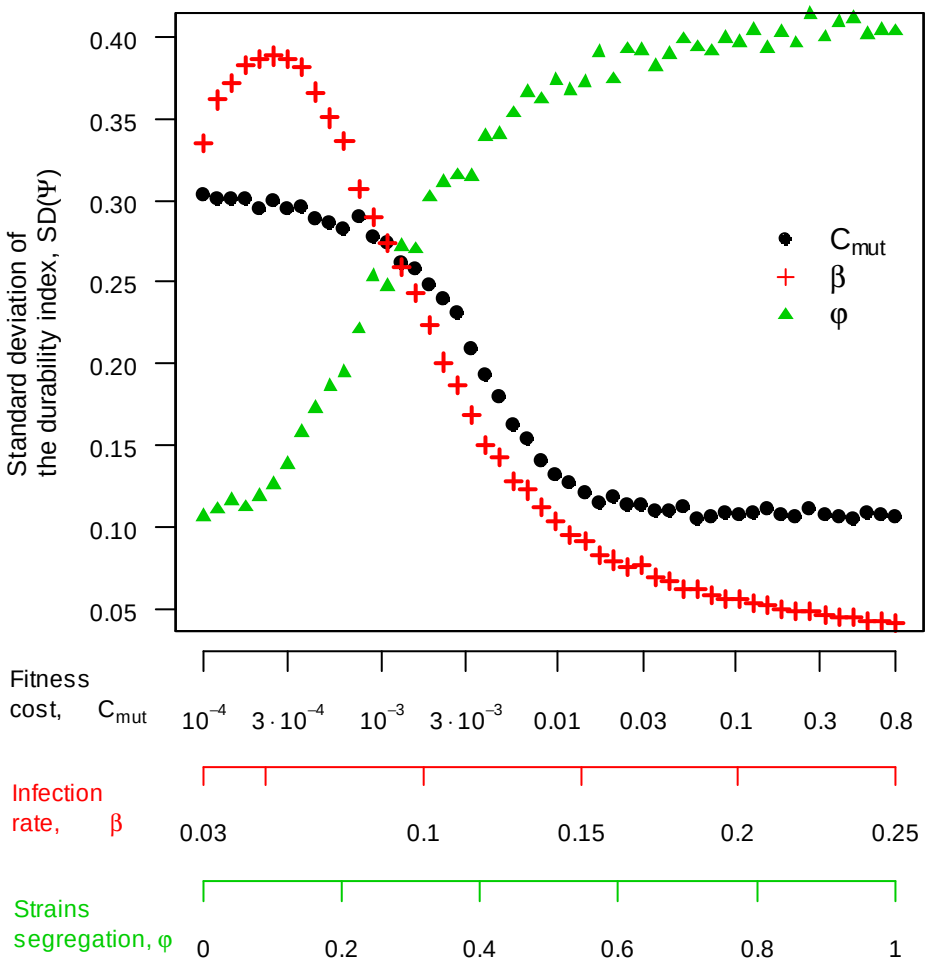


Fig. 4. Standard deviation of durability index, SD(Ψ), as a function of (i) the fitness cost of a mutation (C_{mut}), (ii) the infection rate (β) and (iii) the intensity of virus segregation (φ). SD(Ψ) was assessed for 50 values regularly sampled in the distribution of each these parameters (Table 1) using 1000 replicates while setting the nine others to their reference values.

4. Discussion

Understanding the interplay between the evolutionary and epidemiological processes acting on a virus population both at the host and population scales is fundamentally pertinent to the study of the emergence of new diseases and to the design of efficient control strategies (Jeger et al., 2006; Scherm 2006). However, using reductionist approaches to measure all the processes acting on a large scale is difficult because of the many factors involved and, consequently, the difficulty in designing controlled experiments. Even at the scale of individual host plants, this remains difficult for plant virus populations, as reflected by the scarcity of estimates of these factors (Table 1). In this context, mathematical modelling is essential in making it possible to bridge the scales at which data can be acquired (e.g. molecular processes generating virulent variants, infection of individual plants by a virus population) and those at which the evolutionary epidemiology processes of interest occur and which cannot be studied experimentally (or only with difficulty) (Matthews and Haydon, 2007). This is illustrated by our holistic approach which combined sensitivity analyses, computer modelling and parameter estimates provided by laboratory experiments. By scaling up from the plant cell scale to the host population scale, insights were obtained into the relative contribution of the epidemiological, genetic and evolutionary factors, and their interactions, to the emergence of virulent plant viruses. The model and its analysis apply equally to the different types of plant viruses (DNA or RNA, ss or ds) that are transmitted horizontally. However, it should be appreciated, especially when considering sensitivity analysis results, that for some parameters (m, p, μ_b, λ_{trans}) published data concern only RNA viruses transmitted in a non-persistent manner. Whether the ranges of variation defined are specific of these virus remain unknown as no estimates are available for other types of virus or modes of vector transmission.

Compared to the models of Van den Bosh and Gilligan (2003) and Pietravalle et al. (2006) on these issues, we investigated the effects of additional factors (e.g. the number and nature of mutations required for virulence) while addressing some specificities of plant viruses (e.g. occurrence of severe 'bottlenecks' during horizontal transmission, segregation of virus genotypes in their hosts). Like these authors, we assumed that the virulent genotype did not pre-exist. This made it possible to consider measures of durability that take into account the time necessary for the virulent genotype to appear in the population. Previous models studying virulence dynamics in plant pathogens classically assumed that virulence pre-existed (Van den Bosh and Gilligan, 2003). While these authors questioned how farmers can best deploy genetic resources in time and space to prolong resistance durability, we questioned which criteria should be used by plant breeders to better predict durability and select resistance genes and genotypes.

The focus was on the breakdown of qualitative resistances because these can be easily introduced into cultivars via breeding practices and, from a theoretical point of view, the mechanisms of their breakdown are best understood. For a qualitative type of resistance, viral evolution is unlikely to occur in the resistant plant. Hence, resistance breakdown is a chance process involving the acquisition, by a vector, of a virulent genotype from a susceptible plant. Sensitivity analysis revealed the key determinants of this process. Despite integration across scales, the model reveals few significant interactions between factors as key determinants of the risk of breakdown. The most important factor by far was genetic: the number of mutations required for virulence. This is consistent with the trend, noted by Harrison (2002), of a noticeable increase in durability of resistance with the number of mutations needed for virulence. This result is also consistent with the more general finding that the number of evolutionary steps required greatly reduces the probability of infectious disease emergence (Antia et al., 2003). The mutation rate and the fitness cost of mutations were the next two most important factors and explained similar amounts of variance of $E(\Psi)$. What is expected is that the mutation rate of plant DNA viruses is some orders of magnitude smaller than that of RNA viruses (García-Arenal et al., 2001). Thus, all other things being equal, the durability of a resistance is expected to be substantially longer for DNA than for RNA viruses. Beyond this point, μ_b will be an important factor to consider when ranking the durability of several resistances if the mutation rate of the targeted virus varies according to factors such as the host plant. That remains largely unknown. It will also be important to consider μ_b for ranking the durability of resistances controlling both DNA and RNA viruses. However, currently, no such gene is known. The importance of the fitness penalty imposed on pathogens by the breakdown of a resistance gene was emphasized by Leach et al. (2001)

for fungal and bacterial avirulence genes and proved for some plant viruses (Jenner et al., 2002; Desbiez et al., 2003; Lanfermeijer et al., 2003). This factor is certainly more important, because it is more variable for viruses than for other micro-organisms. Indeed, for fungi, most fitness costs are small, whereas they are more frequent and often quite high for viruses (Carrasco et al. 2007; Sacristán and García-Arenal, 2008). This is due to the small genome size of RNA and DNA plant viruses which imposes overlapping of coding and regulatory sequences as well as the multi-functionality of proteins. In contrast, the complexity of fungal genomes allows high levels of redundancy and alternative metabolic pathways (Sacristán and García-Arenal, 2008). Finally, despite their quite large ranges of variation, epidemiological factors were only of secondary importance in our analysis. The infection rate β had a surprisingly low impact on the durability index (7 % of explained variance) and the length of the latent period (τ) was negligible for E(Ψ). In agreement with Jeger et al. (2004, 2006), this result emphasizes that epidemiology must integrate the evolutionary processes occurring in virus populations if new sustainable strategies of disease management are be developed. More generally, this emphasizes the importance of evolutionary over epidemiological (ecological) factors as key determinants of the initial 'jump' of a pathogen to a new host. However, a successful emergence also requires the subsequent spread and persistence of this newly evolved pathogen within its new host population. Reasons and ways for dealing with this additional step are discussed later.

From the breeder's standpoint, the outputs of the model can help in establishing more general guidelines for resistance gene selection because, unlike most previous studies, our approach is quantitative and multi-factorial. Combining the three major factors for E(Ψ), m, p, and C_{mut} provides new criteria to select durable resistance. In Figure 3, it can be see, for example, that resistance genes requiring (i) one transition and one transversion with slightly deleterious mutations (fitness cost of about 10^{-3}), (ii) two transitions with a 5.10^{-3} fitness cost, (iii) one transversion with a 0.25 fitness cost and (iv) one transition with a 0.8 fitness cost have similar mean durability indices. In general, the effect of an additional mutation required for virulence is equivalent to the effect of a 10^2 to 10^3 increase of the fitness cost imposed by the mutation. However, the standard deviation SD(Ψ) decreases with C_{mut} (Fig. 4). Thus, from a disease management perspective, situation (iv) is the most desirable because it leads to predictions with the same mean durability index and lower uncertainty. However, as discussed later, it should be appreciated that fitness cost may be a transitory state and can decrease in time due to compensatory mutations. More generally, the objective of decision makers is to minimize the average durability index while avoiding its large variation (e.g. Hazell et al., 1971). Depending on the relative importance attributed to the expected benefit and to the uncertainty of its true value, a breeder may prefer a resistance gene characterized by a somewhat higher mean durability index but a lower variance. Accordingly, epidemiological studies of virus incidence will provide valuable information as the infection rate β was the most important factor explaining SD(Ψ). Nevertheless from the breeder's standpoint, the outputs of the model also indicate that combining resistance genes corresponding to distinct avirulence genes in a single genotype will theoretically increase the number of mutations required for virulence. This impacts strongly on E(Ψ), and also increases the chances of accumulating fitness costs, that will both decrease E(Ψ) and SD(Ψ). Thus 'pyramiding' resistance genes is a recommendable strategy assuming that these genes have not previously been overcome.

The model includes some of the stochastic effects occurring in a viral population during epidemics. First, the random sampling done by vectors when acquiring viruses on a plant was taken into account because the size of the 'bottlenecks' imposed by vectors during transmission is known to shape the evolution of virus populations (Moury et al., 2007). Second, to our knowledge, this work is the first to account for the more or less heterogenous distribution of virus variants within their host (Dietrich and Maiss, 2003; Jridi et al., 2006; Takahashi et al., 2007). Patterns of virus segregation in their hosts should (i) increase intra-genotype competition while decreasing inter-genotype competition, a fundamental feature of the 'aggregation model of species coexistence' (see Hartley and Shorrocks, 2002 and references therein) and (ii) impose genetic drift on virus populations during horizontal transmission. The analysis revealed that the intensity of genotype segregation (φ) is the third most important factor explaining the variance of the breakdown process. The size of the bottlenecks imposed by vectors (λ_{trans}) also increases (in absolute values) genetic drift and thus SD(Ψ) but, relatively to β, C_{mut} and φ its effect remains low. Whether other stochastic processes, like the

variability in the viral population size due to the probabilistic events of birth or mutations occurring in populations at low density (demographic stochasticity), are also important remains to be studied.

The model ignores recombination which also shapes virus populations. To our knowledge, no example of resistance breakdown due to recombination between avirulent genotypes has been reported. Whether recombination is an important evolutionary process for emergence has been questioned (Holmes and Drummond, 2007). Its importance as a factor reducing the time necessary to combine mutations required for virulence depends basically on its frequency and on the effective population size of the viral population (Althaus and Bonhoeffer, 2005). The simultaneous presence of various virus genotypes in individual cells is also a prerequisite. Several studies have revealed that closely related viruses tend to be spatially separated in plants (Dietrich and Maiss, 2003; Jridi et al., 2006; Takahashi et al., 2007), but whether or not this is a common phenomenon remains largely unknown. Finally, the likelihood of combining mutations by recombination depends on their location in the virus genome (Rhodes et al. 2005). When mutations are located a few bases apart in the same avirulence gene, recombination should not play a major role. However, recombination should be a more important process when considering the combination of resistance genes targeting distinct avirulence factors. Recombination should also help viruses to recover from the fitness cost of the mutations required for virulence, the second most important factor revealed by our analysis. Similarly, compensatory mutations can reduce, or even suppress, these fitness penalties (García-Arenal et al., 2001; Wijngaarden et al., 2005). Moreover, complementation, the process by which a viral function affected by a mutation is provided by a fully competent virus genotype co-infecting the same cell, can sometimes assist the maintenance of mutants less fit than the wild type (Fraile et al., 2008). In the absence of co-infection of the same cell by several genotypes, its efficiency is likely to be higher for RNA than for DNA viruses due to their higher mutation rate. Finally, the model assumed that no epistasis for fitness occurred between the mutations required for virulence when epistatic effects are supposed to be common for viruses (Holmes, 2003). Following Sanjuan et al. (2004), two cases are possible for deleterious mutations. While antagonistic epistasis between fitness-costly mutations could assist the emergence of virulent genotypes, synergistic epistasis puts them at a disadvantage. Although examples have been reported in the literature, antagonism seems much more common than the synergism (Burch and Chao, 2004; Sanjuan et al., 2004). These possibilities should be investigated in further developments of the model.

In addition to these facets, the model will be useful in testing the effect of agricultural practices on the durability of resistances to plant viruses, and more particularly, the effect of resistance deployment strategies (McDonald and Linde, 2002). To do so, the second step of the process leading to emergence, i.e. the spread and long-term persistence of the virulent virus genotype, must be modelled. For annual crops, the first requirement is to model the within-year spread of the viral population in a metapopulation of hosts in which patches are fields of susceptible and resistant varieties. Such metapopulations will be easy to handle in the IBM used here and between-field virus transmission events can be described with usual dispersal kernels. Secondly, to model the between-year persistence of the viral population, it is necessary to include the seasonal dynamics of the disease and particularly the survival of the viral population in alternative hosts during the "crop-free" season. Given these changes, it will be possible to derive optimal strategies for deploying resistances genes in space and time using, for example, graph-based Markov decision processes (Peyrard et al., 2007). The model could also have a broader interest as, for example, investigating other pathogens such as fungi. The demo-genetic dynamics of the fungal population at the plant scale and its spread at the field scale could be described using the same modelling framework. However, some additional processes should be modelled regarding the evolution toward virulence. Although substitutions are most often required in virus avirulence genes to gain virulence, a wider diversity of mechanisms of molecular evolution, such as insertions and deletions, have been reported in fungi (Sacristán and García-Arenal, 2008).

Acknowledgments

The authors are grateful to Jean-Paul Bouchet for his skill with the Fortran compiler and to Rachid Senoussi and Fabien Halkett for useful comments on an earlier draft of this manuscript. We also thank Cindy E. Morris, Mike Jeger and Mike Thresh for their comments and help with the English.

References

Ali, A., Li, H., Schneider, W.L., Sherman, D.J., Gray, S., Smith, D., Roossinck, M.J., 2006. Analysis of genetic bottlenecks during horizontal transmission of *Cucumber mosaic virus*. J. Virol. 80, 8345-8350.

Althaus, C.L., Bonhoeffer, S., 2005. Stochastic interplay between mutation and recombination during the acquisition of drug resistance mutations in Human immunodeficiency virus type 1. J. Virol. 79, 13572–13578.

Anderson, P.K., Cunningham, A.A., Patel, N.G., Morales, F.J., Epstein, P.R., Daszak, P., 2004. Emerging infectious diseases of plants: pathogen pollution, climate change and agrotechnology drivers. Trends Ecol. Evol. 19, 535–544.

Anderson, P.K., 1994. A conceptual framework of integrated epidemiology for application to emerging diseases. Ann. N.Y. Acad. Sci. 740, 439–444.

Antia, R., Regoes, R.R., Koella, J.C., Bergstrom, C. T., 2003. The role of evolution in the emergence of infectious diseases. Nature 426, 658-661.

Archer, G., Saltelli, A., Sobol, I. M., 1997. Sensitivity measures, ANOVA-like techniques and the use of the bootstrap. J. Stat. Comput. Simul. 58, 99–120.

Ayme, V., Souche, S., Caranta, C., Jacquemond, M., Chadœuf, J., Palloix, A., Moury, B., 2006. Different Mutations in the Genome-Linked Protein VPg of Potato virus Y Confer Virulence on the pvr23 Resistance in Pepper. Mol. Plant-Microbe Interact. 19, 557-563.

Brankin, R.W., Gladwell, I., Shampine, L.F., 1992. RKSUITE: a suite of Runge-Kutta codes for the initial value problem for ODEs. Softreport 92-S1, Department of Mathematics, Southern Methodist University, Dallas, Texas, U.S.A.

Bulmer, M., 1994. Theoretical Evolutionary Ecology. Sinauer, Sunderland, Massachusetts.

Burch, C. L., Chao, L., 2004. Epistasis and its relationship to canalization in the RNA virus Φ6. Genetics 167, 559-567.

Carrasco, P., de la Iglesia F. and Elena, S.F. 2007. Distribution of fitness and virulence effects caused by single-nucleotide substitutions in Tobacco etch virus. J. Virol. 81, 12979-12984.

Cuevas, J.M., Moya, A., Sanjuan, R., 2005. Following the very initial growth of biological RNA viral clones. J. Gen. Virol. 86, 435-443.

DeAngelis, D.L., Mooij, W.M., 2005. Individual-based modelling of ecological and evolutionary processes. Ann. Rev. Ecol. Evol. Syst. 36, 147-168.

Desbiez, C., Gal-On, A., Girard, M., Wipf-Scheibel, C., Lecoq, H., 2003. Increase in Zucchini yellow mosaic virus symptom severity in tolerant zucchini squash cultivar is related to a point mutation in P3 protein and is associated with a loss of relative fitness on susceptible plants. Phytopathology 93, 1478–1484.

Dietrich, C., Maiss, E., 2003. Fluorescent labeling reveals spatial separation of potyvirus populations in mixed infected Nicotiana benthamiana plants. J. Gen. Virol. 84, 2871-2876.

Drake, J.W., Holland, J.J., 1999. Mutation rates among RNA viruses. P. Natl. Acad. Sci. U.S.A. 96, 13910–13913.

Flor, H.H., 1971. Current status of the gene-for-gene concept. Ann. Rev. Phytopathol. 9, 275-296.

Fraile, A., Sacristán, S., García-Arenal, F., 2008. A quantitative analysis of complementation of deleterious mutants in plant virus populations. Spanish J. Agr. Res., 6, 195-200.

García-Arenal, F., McDonald, B.A., 2003. An analysis of the durability of resistance to plant viruses. Phytopathology 93, 941-952.

García-Arenal, F., Fraile, A., Malpica, J.M., 2001. Variability and genetics structure of plant virus populations. Ann. Rev. Phytopathol. 39, 157-186.

Ginot, V., Gaba, S., Beaudouin, R., Aries, F., Monod, H., 2006. Combined use of local and ANOVA-based global sensitivity analyses for the investigation of a stochastic dynamic model: Application to the case study of an individual-based model of a fish population. Ecol. Mod. 193, 479-491.

Harrison, B.D., 2002 Virus variation in relation to resistance breaking in plants. Euphytica 124, 181–192.

Hartley, S., Shorrocks, B., 2002. A general framework for the aggregation model of coexistence. J. Anim. Ecol. 71, 651-662.

Hazell, P.B.R., 1971. A linear alternative to quadratic and semivariance programming for farm planning under uncertainty. Am. J. Agric. Econ. 53:53-62.

Holmes, E.C., 2003. Error thresholds and the constraints to RNA virus evolution. Trends Microbiol. 11, 543-546.

Holmes, E.C., Drummond, A., 2007. The evolutionary genetics of viral emergence, In: Childs, J.E., Mackenzie, J.S., Richt, J.A. (Eds.), Wildlife and Emerging Zoonotic Diseases: The Biology, Circumstances and Consequences of Cross-Species Transmission. Springer Berlin Heidelberg, pp. 51-66.

Ives, A.R., 1988. Aggregation and the coexistence of competitors. Ann. Zool. Fennici 25, 75-88.

Jeger, M.J., Holt, J., Van den Bosch, F., Madden, L.V., 2004. Epidemiology of insect-transmitted plant viruses: modelling disease dynamics and control interventions. Phys. Entomol. 29, 291–304.

Jeger, M.J., Seal, S.E., Van den Bosch, F., 2006. Evolutionary epidemiology of plant virus disease. Adv. Vir. Res. 67, 163–204.

Jenner, C.E., Wang, X., Ponz, F., Walsh, J.A., 2002. A fitness cost for Turnip mosaic virus to overcome host resistance. Virus Res. 86, 1–6.

Johnson, R., 1979. The concept of durable resistance. Phytopathology 69, 198-199.

Jridi, C., Martin, J.F., Marie-Jeanne, V., Labonne, G., Blanc, S., 2006. Distinct viral populations differentiate and evolve independently in a single perennial host plant. J. Virol. 80, 2349-2357.

Kang, B.C., Yeam, I., Jahn, M.M., 2005. Genetics of plant virus resistance. Ann. Rev. Phytopathol. 43, 581-621.

Kitakado, T., Kitada, S., Kishino, H., Skaug, H.J., 2006. An integrated-likelihood method for estimating genetic differentiation between populations. Genetics 173, 2073–2082.

Kuno, E., 1988. Aggregation pattern of individuals and the outcomes of competition within and between species: differential equation models. Res. Pop. Ecol. 30, 69– 82.

Lanfermeijer, F.C., Dijkhuis, J., Sturre, M.J.G., Haan, P., Hille, J., 2003. Cloning and characterization of the durable Tomato mosaic virus resistance gene Tm-22 from Lycopersicon esculentum. Plant Mol. Biol. 52, 1039–51.

Leach, J.E., Vera Cruz, C.M., Bai, J., Leung, H., 2001. Pathogen fitness penalty as a predictor of durability of disease resistance genes. Ann. Rev. Phytopathol. 39, 187-224.

Lecoq, H., Moury, B., Desbiez, C., Palloix, A., Pitrat, M., 2004. Durable virus resistance in plants through conventional approaches: a challenge. Virus Res. 100, 31-39.

Leonard, T., Hsu, S.J., 1999. Bayesian Methods. Cambrigde, University Press.

Lloyd, M., 1967. Mean crowding. J. Anim. Ecol. 36, 1– 30.

Malpica, J.M., Fraile, A., Moreno, I., Obies, C.I., Drake, J.W., Garcia-Arenal, F., 2002. The rate and character of spontaneous mutation in an RNA viruses. Genetics 162, 1505-1511.

Matthews, L., Haydon, D., 2007. Introduction. Cross-scale influences on epidemiological dynamics : from genes to ecosystems. J. R. Soc. Interface 4, 763-765.

McDonald, B.A., Linde, C., 2002. Pathogen population genetics, evolutionary potential, and durable resistance. Ann. Rev. Phytopathol. 40, 349-379.

Mideo, N., Alizon, S., Day, T., 2008. Linking within- and between-host dynamics in the evolutionary epidemiology of infectious diseases. Trends Ecol. Evol. 23, 511-517.

Miralles, R., Ferrer, R., Solé, R.V., Moya, A., Elena, S.F., 2001. Multiple infection dynamics has pronounced effects on the fitness of RNA viruses. J. Evol. Biol. 14, 654– 662.

Moury, B., Fabre, F., Senoussi, R., 2007. Estimation of the number of virus particles transmitted by an insect vector. P. Natl. Acad. Sci. U.S.A. 104, 17891-17896.

Moury, B., Morel, C., Johansen, E., Jacquemond, M., 2002. Evidence for diversifying selection in Potato virus Y and in the coat protein of other potyviruses. J. Gen. Virol. 3, 2563-2573.

Parlevliet, J.E., 2002. Durability of resistance against fungal, bacterial and viral pathogens: present situation. Euphytica 124, 147-156.

Pietravalle, S., Lemarié, S., Van den Bosch, F., 2006. Durability of resistance and cost of virulence. Eur. J. Plant Pathol. 114, 107-116.

Pita, J.S., Roossinck, M.J., 2008. Virus Populations, Mutation Rates and Frequencies, in: Roossinck, M.J. (Ed.), Plant Virus Evolution. Springer Berlin Heidelberg Ed., pp. 109-121.

Peyrard, N., Sabbadin, R., Pelzer, E., Aubertot, J., 2007. A Graph-based Markov decision process framework applied to the optimization of strategies for integrated management of diseases. Phytopathology 97, S92.

Rhodes, T.D., Nikolaitchik, O., Chen, J.B., Powell, D., Hu, W.S. 2005. Genetic recombination of Human immunodeficiency virus type 1 in one round of viral replication: Effects of genetic distance, target cells, accessory genes, and lack of high negative interference in crossover events. J. Virol. 79, 1666–1677.

Sacristán, S., García-Arenal, F., 2008. The evolution of virulence and pathogenicity in plant pathogen populations. Mol. Plant Pathol. 9, 369-384.

Saltelli, A., Ratto, M., Andres, T., Campolongo, F., Caraboni, J., Gatelli, D., Saisona, M., Tarantola, S. 2008. Global Sensitivity Analysis: The primer. John Wiley and Sons, West Sussex, England.

Sanjuan, R., Moya, A., Elena, S.F., 2004. The contribution of epistasis to the architecture of fitness in an RNA virus. P. Natl. Acad. Sci. U.S.A. 101, 15376–15379.

Schrag, S., Wiener, P., 1995. Emerging infectious disease: what are the relative roles of ecology and evolution? Trends Ecol. Evol. 10, 319–324.

Scherm, H., Ngugi, H.K., Ojiambo, P.S., 2006. Trends in theoretical plant epidemiology. Eur. J. Plant Pathol. 115, 61-73.

Shaner, G., Stromberg, E., Lacy, G., Barker, K., Pirone, T., 1992. Nomenclature and concepts of pathogenicity and virulence. Ann. Rev. Phytopathol. 30, 47–66.

Solé, R.V., Ferrer, R., González-García, I., Quer, J., Domingo, E., 1999. Red Queen dynamics, competition, and critical points in a model of RNA virus quasispecies. J. Theor. Biol. 198, 47– 59.

Takahashi, T., Sugawara, T., Yamatsuta, T., Isogai, M., Natsuaki, T., Yoshikawa, N., 2007. Analysis of the spatial distribution of identical and two distinct virus populations differently labelled with cyan and yellow fluorescent proteins in coinfected plants. Phytopathology 97, 1200-1206.

Thompson, J.N., Burdon, J.J., 1992. Gene-for-gene coevolution between plants and parasites. Nature 360, 121-325.

Van den Bosch, F., Gilligan, C.A., 2003. Measures of durability of resistance. Phytopathology 93, 616-625.

Wijngaarden, P.J., Van den Bosch, F., Jeger, M.J., Hoekstra, R.F., 2005. Adaptation to the cost of resistance: a model of compensation, recombination and selection in a haploid organism. Proc. R. Soc. B 272, 85-89.

Yang, Z., Nielsen, R., Goldman, N., Pedersen, A.M.K., 2000. Codon-substitution models for heteregeneous selection pressure at amino acid sites. Genetics 155, 431-449.

Annex 1. Notes on the dirichlet-multinomial distribution

The beta-binomial model is a generalisation of the binomial model when the probability of 'success' varies between trials and follow a beta distribution. Similarly, the dirichlet-multinomial (DM) distribution is a generalization of the multinomial distribution (Leonard and Hsu, 1999). For example, this distribution is used in population genetics to describe the number of alleles sampled at each locus in demes of a metapopulation (Kitakado et al., 2006). The dirichlet distribution is the multivariate beta distribution.

Let $y=(y_1, y_2, \ldots, y_N)$ be the vector of the number of virus particles of each genotype (N, total number of genotypes) observed in a patch of a host plant and $Y=\Sigma y_i$. Given the genotype frequencies in that patch is $\pi=(\pi_1, \pi_2, \ldots, \pi_N)$, y follows a multinomial distribution of parameters π and Y. Now, let $\theta=(\theta_1, \theta_2, \ldots, \theta_N)$ be the overall mean frequencies of the genotype in the plant. Assuming that the prior distribution of π is Dirichlet with parameters $\alpha\theta_1, \alpha\theta_2, \ldots, \alpha\theta_N$ where α is an over-dispersion parameter ($\alpha>0$), the distribution of y is obtained by integrating out the π_i. This results in the DM distribution which probability density function is:

$$p(y/\alpha,\theta,Y)=\frac{Y!\Gamma(\alpha)/\Gamma(Y+\alpha)}{\prod_{j=1}^{N}\left[y_j!\Gamma(\alpha.\theta_j)\right]}\prod_{j=1}^{N}\Gamma(y_j+\alpha.\theta_j)$$

The moments of the DM distribution are $E(y_i|Y)=Y.\theta_i$, $Var(y_i|Y)=\rho.Y.(\theta_i-\theta_i^2)$ and $Cov(y_i,y_j|Y)=-\rho.Y.\theta_i.\theta_j$ where $\rho=\frac{Y+\alpha}{1+\alpha}$. The smaller the value of α, the more diverse are the samples between patches. Conversely, when α tends to infinity, the DM distribution converges to the multinomial distribution.

Annex 2. Sobol- and ANOVA-based Global Sensitity Analyses (GSA) perform similarly with the model developed.

In ANOVA-based GSA, each factor takes a limited number of levels (five in our case study), whereas other GSA methods allow levels to be sampled over the full range of any quantitative parameter (Ginot et al., 2006). This is particularly so for the Sobol-based GSA which, moreover, does not require any assumptions regarding the relationship between input parameters and model output (model-free method) (Saltelli et al., 2008). For our purpose, major drawbacks of this method were (i) the many model runs required and (ii) the impossibility of dealing with nested factors.

A Sobol GSA was realized in order to check the validity of ANOVA-based GSA results in our case study. The sampling scheme, which involved drawing levels of parameters according to their true probability distribution, was generated using low discrepancy sequences (also termed quasi-random sequences) that have better properties to fill the space of parameters than pseudo-random sequences (at least for dimension below 30). The more general model-free variance decomposition of Sobol was used to estimate sensitivity indices. Their estimation required $L.(s+2)$ runs of the model, where s is the number of factors and L is the number of rows of the two independent samples required by the method. The precision of the sensitivity estimates was evaluated using the bootstrap method (Archer et al., 1997). To achieve a reliable precision, L was set to 2000, and s was equal to nine (for Sobol-based GSA, factors m and p should be merged in a single factor $m(p)$, representing the number and nature of the mutations required for virulence). In all, the method required 22000 runs of the model whereas only 3125 were needed for the ANOVA-based GSA.

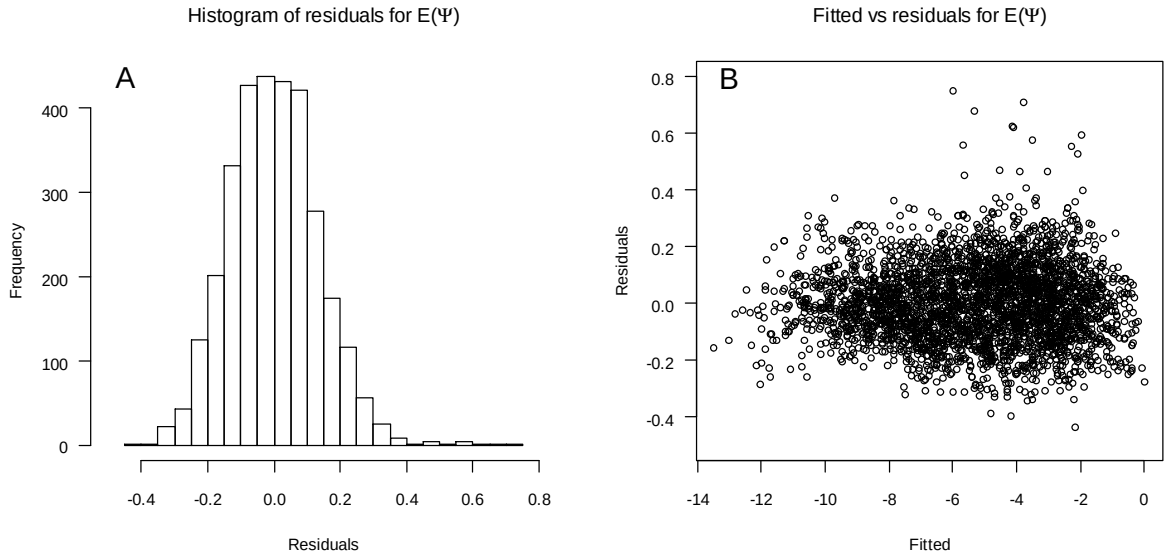
Sobol-based and ANOVA-based sensitivity indices were compared for $E(\Psi)$ using 100 replicates for each parameter space sample. ANOVA residuals were distributed normally in both cases

(Supplementary Fig. 1). Moreover, the fit of the ANOVA model being very high (99% of explained variance), sensitivity indices could be derived appropriately.

Seventeen of 18 (≈ 95%) first order and total order sensitivity indices derived with an ANOVA-based method were included in the 95% confidence interval of sensitivity indices derived with Sobol (Supplementary Table 1). Only the total sensitivity indice for the factor number and nature of mutations required for virulence [$m(p)$] obtained with ANOVA was slightly higher than the upper limit of the 95% confident interval of the corresponding indices obtained with Sobol (0.49 versus 0.48). Thus, in our case study, despite the strong non-linearity of the model, the two sensitivity analysis methods gave similar results with a seven times lower computational cost for the ANOVA.

Supplementary Table 1: First-order and total sensitivity indices for the mean durability index, E(Ψ), issued from ANOVA-based and Sobol-based variance decomposition methods. For Sobol decomposition, 95% confidence intervals (CI) estimated by bootstrap are provided. Factors are ranked according to Sobol first-order indices, as suggested by Saltelli et al. (2008) for prioritization of factors. As Sobol indices are estimated numerically, it is possible to obtain negative estimates near zero (Archer et al., 1997).

Sensitivity indices	First-order indices		Total indices	
Parameter	Sobol (95% CI)	ANOVA	Sobol (95% CI)	ANOVA
$m(p)$, number and nature of mutations required for virulence	0.37 (0.27,0.46)	0.45	0.38 (0.28,0.48)	0.49
C_{mut}, relative fitness cost of a mutation	0.19 (0.12, 0.27)	0.17	0.22 (0.14, 0.29)	0.19
μ_b, mutation rate	0.18 (0.11,0.25)	0.21	0.16 (0.1, 0.23)	0.23
β, infection rate	0.06 (0.02, 0.09)	0.07	0.05 (0.01, 0.08)	0.07
φ, intensity of spatial segregation of virus variants	0.03 (-0.01, 0.06)	0.02	0.05 (0.02, 0.09)	0.03
ξ, daily number of virus generations	0.01 (-0.003,0.03)	0.02	0.008 (-0.008, 0.02)	0.02
κ, transition/transversion rate ratio	0.01 (-0.007, 0.03)	0.006	0.003 (-0.01, 0.02)	0.01
τ, mean length of the latent period	0.005 (-0.001, 0.01)	0.001	-0.001 (-0.006, 0.005)	0.002
λ_{trans}, mean number of virus particles transmitted by vectors	0.003 (0, 0.006)	10^{-5}	-0.002 (-0.005, 0.001)	10^{-4}



Supplementary Fig. 1. A posteriori visual check of the assumptions of ANOVA for E(Ψ). **A**: Histogram of residuals issued from ANOVA. **B**: Relation between residuals and fitted values of ANOVA.