

1 **Virulence evolution via pleiotropy in vector-borne**
2 **plant pathogens.**

3 Elise Woodruff, Nate B Hardy*

4 *Department of Entomology and Plant Pathology, Auburn University*

5 *Corresponding author contact information:

6 Department of Entomology and Plant Pathology

7 301 Funchess Hall

8 Auburn University

9 Auburn, Alabama 36849

10 email: n8@auburn.edu

11 **Abstract**

12 The dynamics of virulence evolution in vector-born plant pathogens can be complex. Here we
13 use individual-based simulations to investigate how virulence evolution depends on genetic
14 trade-offs and population structure in pathogen populations. Although quite generic, the model is
15 inspired by the ecology of the plant-pathogenic bacterium *Xylella fastidiosa*, and we use it to
16 gain insights into possible modes of evolution of virulence in that group. In particular, we aim to
17 sharpen our intuition about how virulence may evolve over short time scales in response to de-
18 creases in vector efficacy. We find that even when pathogens find themselves much more often
19 in hosts than vectors, selection in the vector environment can cause correlational and potentially
20 non-adaptive changes in virulence in the host. The extent on such correlational virulence evolu-
21 tion depends on many system parameters, including the pathogen transmission rate, the relative
22 proportions of the pathogen population occurring in hosts versus vectors, the strengths of selec-
23 tion in host and vector environments, and the extent of virulence per se. But there is a statistical
24 interaction between the strength of selection in vectors and the predominance of pathogens in
25 hosts, such that if within-vector selection is strong enough, the predominance of pathogens
26 within hosts has little effect on the evolution of virulence.

27 **Introduction**

28 In a mixed environment, selection tends to be more efficient in habitat types that are more com-
29 monly-encountered or productive (Via and Lande, 1985; Whitlock, 1996; Draghi, 2021). There-
30 fore, as long as a population is adapting to a common or productive habitat, evolution in less
31 common or productive habitats is expected to be largely correlational (Via and Lande, 1985;
32 Hardy and Forister, 2023). But what if a population's life history entails obligate movements
33 through habitats of different frequency or productivity? And what if a population's evolution can
34 alter the frequency and quality of habitat types? Here we consider these general questions, with a
35 special focus on virulence evolution in vector-born plant pathogens.

36 For such pathogens, the within-host environment is much more commonly-encountered and pro-
37 ductive than the within-vector environment. Thus, without accounting for the details of their life

history, we might expect the evolution of virulence-affecting pathogen phenotypes to be driven by selection in the host, and any evolutionary change in the vector to be largely correlational (Via and Lande, 1985; Hardy and Forister, 2023). But, their life history could change that expectation. The habitat variation experienced by pathogens is largely of the course-grained temporal variety; transmission entails obligate alternations between host and vector environments. Moreover, the evolution of high virulence – that is, pathogen-induced host mortality – can reduce disparities in the frequency and quality of host and vector environments. How such a life history affects asymmetries in selection across habitat types is not clear.

To improve our intuition, we develop and analyze individual-based simulation models. The classical theory of virulence evolution is based on pathogen life history trade-offs. It predicts an optimal level of virulence that balances short- and long-term transmission efficiency (Anderson and May, 1982; Ewald, 1983; Frank, 1996; Alizon *et al.*, 2009; Bull and Lauring, 2014). In general, short-term transmission efficiency increases with within-host pathogen density. But high pathogen density within a host can increase host mortality. This shrinks the time over which an infected host can be the source for pathogen transmission to a new host. Thus, virulence evolution can entail a meta-population-level negative feedback (Alizon *et al.*, 2009; Bull and Lauring, 2014). This theory is based on the epidemiological compartment models that are not explicitly population genetic (Day and Proulx, 2004). They tell us about equilibrium conditions, that is, where a system is ultimately headed. But they tell us nothing about how long a system might take to arrive at equilibrium conditions, or what might happen along the way (Day and Proulx, 2004). Here, our aim is to understand how the non-equilibrium dynamics of virulence evolution depend on genetic architecture and population structure. Our individual-based simulation approach lets us do that.

Although quite generic, our model is inspired by observations of virulence evolution in the insect-vectored plant-pathogenic bacterium *Xylella fastidiosa*. *X. fastidiosa* is associated with a wide range of host species (EFSA, 2016). In most, it is a benign commensalist, but in some cases infections are highly virulent, and *X. fastidiosa* is the causative agent of several important agricultural diseases, such as phony peach disease (Johnson *et al.*, 2023), Olive Quick Decline Syn-

drome (Saponari *et al.*, 2017; Trkulja *et al.*, 2022), and Pierce’s disease in grapevine (Hopkins and Purcell, 2002). The emergence of the latter in California was co-incident with the spread of a new vector species, the glassy-winged sharpshooter (Hemiptera: Cicadellidae: *Homalodisca vitripennis*) (Hopkins and Purcell, 2002). *H. vitripennis* is an exceptionally inefficient vector (Redak *et al.*, 2004), but it became so numerically dominant that most transmission to and from grapevine in California is now via *H. vitripennis*. This is not the only case in which the emergence of high-virulence pathogen genotypes has been casually associated with the emergence of a new vector species or genotype. For example, the global spread of the *Bemisia tabaci* is thought to have repeatedly driven genetic divergence and virulence evolution in begomoviruses, many of which now cause serious diseases problems in crops ranging from okra in western Africa to tomato in Peru and Taiwan (Gilbertson *et al.*, 2015). In the case of Pierce’s disease, the adaptiveness of high virulence for *X. fastidiosa* is uncertain, as vectors prefer asymptomatic host plants, and thus transmission may be inhibited by high within-host density (Daugherty *et al.*, 2011). Can selection for improved within-vector performance cause correlational and potentially non-adaptive evolution of virulence in hosts?

Methods

To address this question, we simulate the evolution of a structured meta-population of pathogen individuals, each of which has a diploid, single-chromosome, 40kb genome. Although loosely inspired by *Xylella fastidiosa*, the details about genomic structure are arbitrary and should not affect our inferences. In the model, individuals reproduce clonally and without recombination or dominance interactions between alleles. Hence, the genome should diversify in a similar manner to a haploid, circular genome of twice the size. Likewise, the simulated genomes are much smaller than the genomes of *X. fastidiosa* (Simpson *et al.*, 2000), but the mutation rate is much higher.

At the start of each simulation, the population is genetically uniform; genomes are essentially empty containers for future mutations. The pathogen population is divided into $n_d=100$ demes, each of which occurs in either a host or vector individual. So as not to be confused with pathogen

93 individuals, we refer each host or vector as a habitat patch. The model parameter $\rho = 0.1$ gives
 94 the proportion of patches that are hosts.

95 **Table 1.** Model parameters and variables.

Parameter	Definition	Range of Values
n_d	Number of pathogen demes	100
ρ	Relative frequency of host environments	0.1
κ	Strength of pleiotropic covariance	-0.8
ω_h	Weakness of selection in host	$1 < \omega_h < 10$
ω_v	Weakness of selection in vector	$1 < \omega_v < 10$
m	Migration rate	$0.001 < m < 0.2$
μ_h	Background rate of host mortality	$1e-4 < \mu_h < 0.01$
μ_v	Rate of vector mortality	0.05
v_{max}	Maximum virulence effect	$0.01 < v_{max} < 0.8$
K_h	Pathogen carrying capacity of one host	$200 < K_h < 2000$
K_v	Pathogen carrying capacity of one vector	20
Variable	Definition	Range of Value
T	Number of generation until mean pathogen within-host-performance phenotype within 10% of optimum	$201 < T < 1e4$
Γ	Degree to which pathogen population's evolutionary path bends towards to the optimum in the vector environment	$-\text{inf} < \Gamma < \text{inf}$
N_h	Number of pathogen individuals in host patches	$0 \leq N_h$
N_v	Number of pathogen individuals in vector patches	$0 \leq N_v$

96 For the sake of simplicity, this is not a multi-species model. Hosts and vectors are not evolvable
 97 entities; they are simply two kinds of pathogen habitat, for example, as in (Holt *et al.*, 2003).
 98 That being said, we do allow for turnover of vector and host patches. In each pathogen genera-
 99 tion, pathogen demes can suffer extirpation at background mortality rates μ_v {0.05} in vector
 100 patches and u_h {0.0001 – 0.01} in host patches. In hosts, this rate is elevated by a virulence
 101 effect. Specifically, $v_i = v_{max} / (1 + \exp(-a \cdot d_i))$, where v_i is the excess risk of mortality experience
 102 by host i , v_{max} is the maximum possible virulence effect {0.01 – 0.8}, d_i is the density of
 103 pathogens within host i , that is n_i/K_i , and $a=5$ controls the steepness of the logistic mapping of
 104 pathogen density to virulence. Patch replacement is instantaneous. Vector and host death
 105 amounts to setting the number of pathogens in that particular patch to zero. The patch is then im-
 106 mediately available for re-colonization in the next pathogen generation.

107 The pathogen life cycle begins with offspring production. To repeat, reproduction is clonal, with
 108 each individual producing one offspring individual. Generations are overlapping. Offspring
 109 genomes are generated by random mutation of the parental genome. The mutation rate is $1e-8$
 110 per site, per genome, per generation. When a mutation occurs, a two-dimensional vector of allele
 111 effects is drawn from a zero-meaned random bivariate normal distribution with variances of 1.0,
 112 and symmetrical covariances, κ , the sign and magnitude of which controls the pleiotropy be-
 113 tween two pathogen quantitative phenotypes. An individual's host-performance phenotype is de-
 114 termined by the sum of the first elements of allele effect vectors. Likewise, an individual's vec-
 115 tor-performance phenotype is the sum of the second elements of allele effect vectors. So, if $\kappa > 0$,
 116 positive pleiotropy prevails and an allele that increase the host-performance phenotype value
 117 tends to also increase the vector-performance phenotype value. Conversely, when $\kappa < 0$, antago-
 118 nistic pleiotropy prevails. Here our goal is to understand how antagonistic pleiotropy between
 119 phenotypes affecting performance in vectors and hosts might drive the evolution of virulence.
 120 Therefore, we focus on the case of $\kappa = -0.8$.

121 The next step in the life cycle is migration, that is, pathogen transmission. This happens at per
 122 capita rate m {0.001 - 0.2} and, in the main version of the model (Fig. 1a), is random between
 123 patches except that migrants from a host patch can only move to a vector patch and vice versa. In
 124 an alternative version of the model (Fig. 1b), we relax this constraint and let migration be unfet-
 125 tered between patches. In other words, we do away with vector transmission, and consider the
 126 evolution of a population in an environment in which there are two kinds of hosts, one being
 127 large, rare and susceptible to infection, and the other being small, abundant and tolerant of infec-
 128 tion. Comparison of the pathogen evolutionary dynamics in this unfettered-migration model to
 129 the main vector-born model, reveals the effects of vector transmission per se.

130 After migration comes selection and population regulation. This entails genotype-environment
 131 matching, and density dependence. In the vector environment, the match between a pathogen's
 132 vector performance phenotype and the local optimum, determines their viability, that is, their
 133 probability of surviving until reproduction. This matching is via a standard Gaussian fitness
 134 function, with variance ω_v {1.0-10.0} setting the weakness of selection. In the host environment,

138 viability is the output of the same kind of Gaussian fitness function, but with variance ω_h {1.0-
 139 10.0}. In vector and host patches, individual-level fitness is also density dependent; in habitat
 140 patch i , each individual's viability is rescaled by the ratio of the patch carrying capacity K_i and
 141 the current local pathogen population size, N_i . As mentioned above, in the host, there is also
 142 group-level selection via a virulence effect. As populations evolve mean host-performance phe-
 143 notype values that more closely match the optimal value for the host environment and their
 144 within-host fitness increases, so too does the rate of host death, that is, deme extirpation. Note
 145 that selection is hard; it affects survival and thus has demographic effects.

146 After selection, the life cycle starts again with offspring production.

147 At the start of each simulation, pathogens are monomorphic, with a value of zero for their vector-
 148 performance and host-performance phenotypes, and the optimal value for each of these pheno-
 149 types is set to 5.0 phenotypic units. During the first 200 generations, the pathogen population is
 150 subject only to density-dependent regulation; selection and virulence effects are not applied, and
 151 so genetic diversity accumulates. Then, starting at generation 201, selection and virulence kick
 152 in. We then observe how the population adapts to its host and vector environments. Our view of
 153 these dynamics is based on two test statistics. The first, T , is simply a long-transformed count of
 154 how many generations it takes to evolve a mean host-performance phenotype within 10% of the
 155 optimum, and thus closely approach their maximum virulence effect on the host. Note that be-
 156 cause of the negative meta-population-level feed-backs induced by high virulence, such proxim-
 157 ity to the optimum host value might not be adaptive for the pathogen population; in other words,
 158 a mean host-performance phenotype value within 10% of the optimum might not be the equilib-
 159 rium state of a pathosystem. Thus, T is best interpreted as the hazard of evolving high virulence,
 160 even if only temporarily.

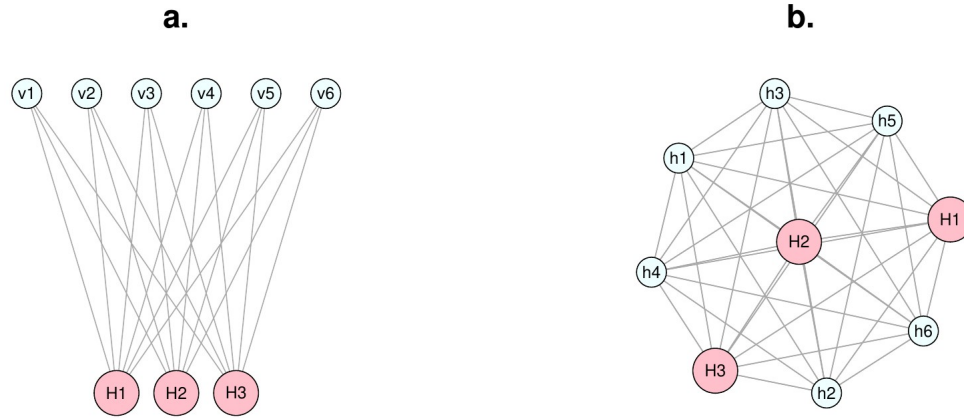
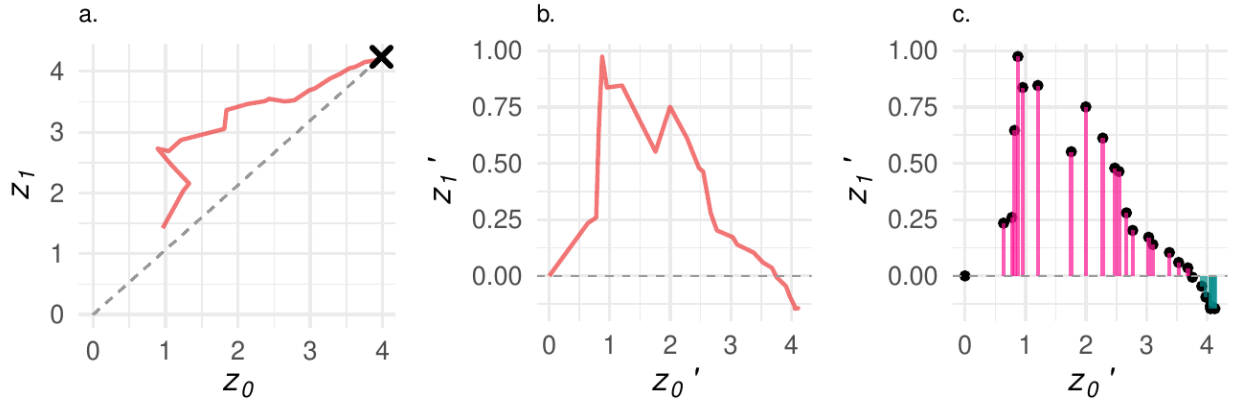


Figure 1. Alternative population structures. Network vertices represent specific host (H_i or h_i) or vector (v_i) patches. Larger vertices have higher pathogen carrying capacities. Edges represent possible paths for pathogen migration. **(a)** In the main model, the pathogen population is vector-borne; migration is only possible between trophic levels, that is, from a host (H_i) to a vector (v_i), or vice versa. **(b)** In an alternative version of the model, migration is unfettered; thus, rather than consisting of a mix of hosts and vectors, the environment consists of a few large and susceptible hosts (H_i) and several small and tolerant hosts (h_i).

The second statistic we track, Γ , is a measure of the degree to which, until a population evolves a within-host performance phenotype within the 10% threshold of the optimum, the population's evolutionary path bends towards or away from the optimum in the vector environment. In other words, we look at the extent to which adaptation in the pathogen population is dominated by the vector or host habitat type (See Fig. 2 for an example).

To calculate Γ we use a little trigonometry. First, we translate a population's evolutionary path through the phenotypic space so as to start at the origin. We do this by subtracting the first post-burnin (generation 201) mean value for each phenotype (z_0 , and z_1) from the mean phenotype value for each subsequent generation. Since the optimal value for each phenotype is 5.0, and pathogen populations start out with phenotype values of 0.0, a straight evolutionary path to the

178 joint optimal phenotype would have a slope of one. Therefore, for each simulation, we rotate the
 179 translated evolutionary path D radian degrees about the origin, where D is the inverse tangent of
 180 one. This rotation is done as follows: $z_{0i}' = z_{0i} \cos(D) + z_{1i} \sin(D)$; $z_{1i}' = z_{1i} \cos(D) - z_{0i} \sin(D)$,
 181 where (z_{0i}, z_{1i}) is point i along the simulated post-burnin translated evolutionary path, z_{0i} is the
 182 population's mean value for the host performance phenotype, z_{1i} is the mean value for the vector
 183 performance phenotype, and (z_{0i}', z_{1i}') is that same point in the rotated coordinate space. We can
 184 then calculate the degree to which the evolutionary path bends towards what is optimal in the
 185 vector environment as $\text{sum}(z_{1i}')/T$, in other words, the per generation average deviation from the
 186 ideal evolutionary path.



187 **Figure 2.** Calculation of the Γ statistic. (a) An example evolutionary path through the phenotype
 188 space. For this simulation $\kappa=-0.2$, $v_{max}=0.3$, $m=0.1$, $K_v=20$, $K_h=2000$, $\omega_v=3.0$, $\omega_h=3.0$, $\mu_h = 1e-4$,
 189 $\mu_v=0.1$, and $\rho = 0.1$. (b) That same path translated to start at the origin and rotated so that the
 190 ideal path from the origin to the joint phenotypic optimum lies along the x-axis. (c) Γ is calcu-
 191 lated as the sum of deviations from the ideal path, scaled by the length of the path in generations.
 192 When Γ is positive, the evolutionary path bends mostly towards the vector environment; con-
 193 versely when it is negative, the host environment dominates.

194 A total of 200 simulations were performed for each version of the model, that is, the vector-
 195 borne transmission model, and the unfettered-migration model. For each run, a value for each

196 free model parameter was drawn from a random uniform distribution with ranges as given in Ta-
 197 ble 1.

198 We analyzed simulation model outputs by fitting mutli-variate linear models, using the R pack-
 199 age *lme4* (Bates *et al.*, 2003). In one model, the response variable was a log transformation of T .
 200 In the other, the response was Γ . For both models the fixed predictor variables were (1) ω_h , the
 201 weakness of selection in hosts, (2) ω_v , the weakness of selection in vectors, (2) m , the migration
 202 rate, (3) μ_h , the background rate of host mortality, (4) v_{max} , the maximum extent of virulence, that
 203 is, pathogen-induced host mortality, (5) N_h/N_v , the log-transformed ratio of the number of
 204 pathogens in hosts to those in vectors, and (6) the interaction between N_h/N_v and ω_v . This inter-
 205 action term is what we are most curious about; it gives us the clearest view of the possibility that
 206 selection for improved within-vector performance could drive the correlational evolution of viru-
 207 lence in hosts, even if most of the pathogen populations occurs within hosts.

208 To get a better sense for what could be complex causal links in the system, we also analyzed
 209 model outputs by fitting a structural equation model, using the R package *laavan* (Rosseel, Y *et*
 210 *al.*, 2017). Here again, we focused on the case of strong and negative pleiotropic covariances,
 211 $\kappa=-0.8$, where vectors are more abundant than hosts, $\rho=0.1$, but have much smaller carrying ca-
 212 pacities, and much higher background mortality rates.

213 To sum up, we examined how, when vectors are more abundant but less productive than hosts,
 214 and there is strong antagonistic pleiotropy between within-host and within-vector performance
 215 phenotypes (i) the time it takes a population to evolve a host-performance phenotype close to the
 216 optimal value, and (ii) the degree to which a population's evolutionary path through the pheno-
 217 typic space bends towards or away from the vector environment depends on (a) the relative car-
 218 rying capacities of host and vector demes, (b) the relative strengths of selection in host and vec-
 219 tor demes, (c) the migration rate, and (d) the maximum virulence effect of high density in host
 220 demes. We also considered how all of this is affected by doing-away with vector-based transmis-
 221 sion, and allowing for completely random migration.

Simulations were performed with SLiM 4 (Haller and Messer, 2023). SLiM models are specified with codes written in the Eidos language. The Eidos code for the model described here is provided in Supplementary File S1.

Results and Discussion

Let us start by considering the the variance in T – that is, the time it takes for the population to evolve a host-performance phenotype that is close to the optimum – when there is vector-based transmission. In the linear regression model (Fig. 3a.), the predictors explained a considerable portion of its variance (adjusted- $R^2 = 0.79$). Each of the inferred primary effects is intuitive; thus the statistical model provides us with some assurance that our evolutionary model is behaving itself. Prior to analysis, the data were mean-centered and variance-scaled, so effects are expressed in units of standard deviations (SD). Three of the primary fixed effects significantly decrease T . The first is ω_v , that is, the weakness of selection in vectors (coefficient -0.15 SD, p-value = $2.3e-5$); evolution in hosts is faster when selection in vectors is weaker. The second is m , that is migration rate (coefficient = -0.18 SD; p-value = $2.2e-7$); this resonates with previous theoretical work showing that high migration reduces the meta-population-level cost of evolving high virulence (Bull and Lauring, 2014). The third is N_h/N_v , that is, the predominance of pathogen individuals in host patches (coefficient = -0.39 SD, p-value = $3.6e-8$). On the other hand, two parameters significantly increase T . This first is v_{max} , the maximum additional host mortality than can be caused by an infection (coefficient = 0.27 SD, p-value = $3.9e-10$); as v_{max} rises, so does the meta-population-level fitness cost of evolving a within-host performance phenotype that closes matches the optimum, hence, the negative feedback on virulence evolution increases in strength. The variable T also tends to increase with larger values for ω_h , that is, with weaker selection in host patches (coefficient = 0.69 SD; p-value < $2e-16$); simply put, adaptation to the host environment is slower when the within-host fitness consequences of maladaptation are less pronounced. The interaction between the weakness of selection in vectors, and the predominance of pathogens in hosts, $\omega_v:N_h/N_v$, was not significant (coefficient = -0.063, p-value = 0.084) (Fig 4a.). To summarize, even with obligate migration between habitat types, selection tends to be more efficient in a particular habitat when it is more commonly encountered (Whitlock, 1996; Hardy and Foris-

ter, 2023). But the effect of selection in the less productive/frequent habitat type is of a similar magnitude of the effect of the disparity in between-habitat type productivity/frequency.

To reiterate, because of negative meta-population-level feed-backs a close match between the mean host-performance phenotype and the optimum, can be nonadaptive. In that case, hitting the host-habitat-match threshold could be largely dictated by stochastic processes. Because virulence effects complicate the interpretation of variance in T , our alternative statistic Γ , is especially useful.

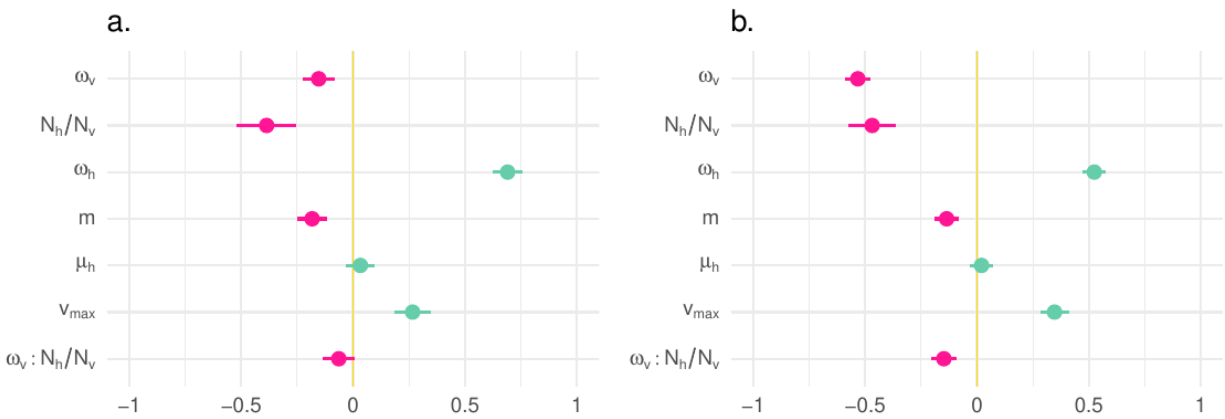
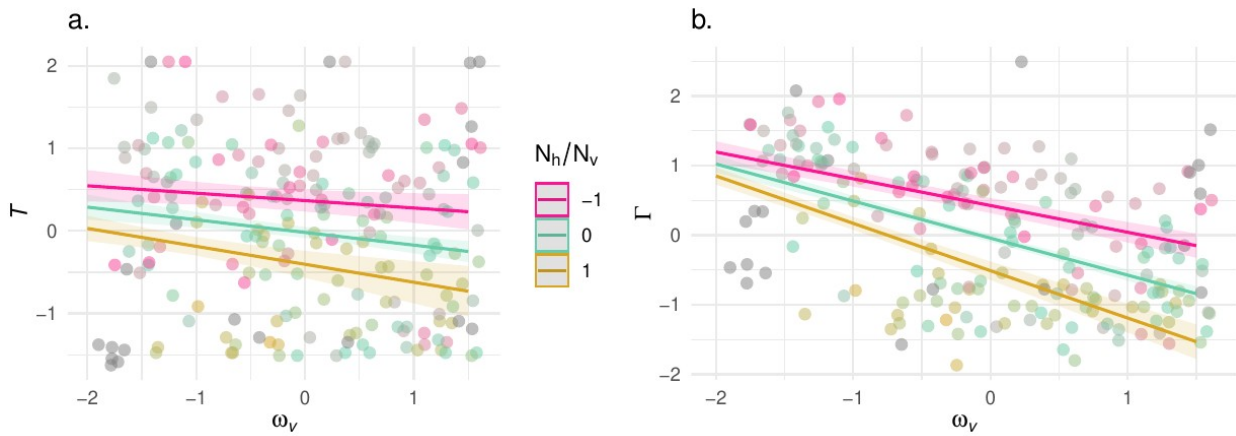


Figure 3. Summary of multi-variate linear models decomposing the variance in (a) T , the number of generations it takes for the pathogen population to evolve a within-host performance phenotype within 10% of the optimum, and (b) Γ , the degree to which, before T , the pathogen population's evolutionary path bends towards ($\Gamma > 0$) or away from ($\Gamma < 0$) the vector environment. Points show positive (teal) and negative (pink) estimated coefficients, and horizontal bars show 95% confidence intervals. All predictors have been centered and scaled to units of standard deviations (the units of the x-axis).

Let us turn then to the linear regression of the variance in Γ , the degree to which simulated evolutionary paths bend toward ($\Gamma > 0$) or away from ($\Gamma < 0$) the vector environment (adjusted $R^2 = 0.86$) (Fig 3b.). Two parameters significantly increase Γ : (1) ω_h , the weakness of selection in hosts (coefficient = 0.57, p-value < 2e-16), and (2) v_{max} , the maximum virulence effect (coeffi-

268 cient = 0.35; p-value < 2e-16). On the other hand, three primary fixed effects significantly de-
 269 crease Γ : (1) ω_v , the weakness of selection in the vector (coefficient = -0.53, p-value < 2e-16);
 270 (2) m , the migration rate, m (coefficient = -0.14; p-value = 2.0e-6), and N_h/N_v , predominance of
 271 the host environment (coefficient = -0.47; p-value = 3.8e-15). These effects are consistent with
 272 those estimated on T . But with Γ , we recover significance for the interaction between the weak-
 273 ness of selection in the vector and the predominance of the host, $\omega_v:N_h/N_v$ (coefficient = -0.15
 274 SD, p-value = 1.7e-6; Fig 4b.).

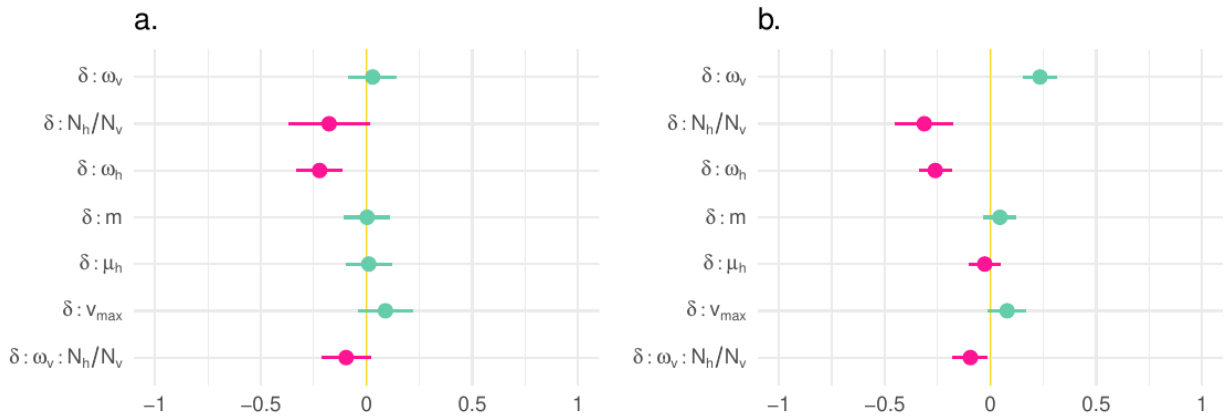
275 To sum up, the predominance of hosts might not have a strong influence on how selection in vec-
 276 tors affects the rate of adaptation in hosts, and if hosts are sufficiently abundant, even very strong
 277 selection in vectors is expected to have little effect on the rate of pathogen adaptation to the host
 278 environment. But same cannot be said for Γ , the extend to which the vector causes correlational
 279 evolution in the host environment; if selection is strong enough in vectors, just how much more
 280 of the pathogen populations occurs in hosts matters little.



281 **Figure 4.** Marginal effect estimates of the interaction between the weakness of selection in vec-
 282 tors, ω_v , and the predominance of hosts N_h/N_v . Each point and linear interpolation shows the pre-
 283 dicted value for T (a) and Γ (b) for a given value of ω_v in combination with a level of N_h/N_v ,
 284 where pink is for -1 SD, teal is for the mean, and gold is for +1 SD. Interpretation: For Γ , as the
 285 host environment becomes more abundant, the $\omega_v:N_h/N_v$ interaction becomes steeper, so that
 286 when selection vectors is strong, the disparity in Γ between levels of host predominance is di-

287 minished. If selection in the vector is strong enough, the abundance and productivity of hosts is
 288 irrelevant.

289 How much of this depends on vector-borne dispersal as opposed to other facets of the pathogen
 290 population structure and ecology? To find out, we ran simulations in which migration was unfet-
 291 tered between habitat patches. Then, we combined the outputs of both model types, and fit linear
 292 models which included terms for the interaction between dispersal mode, δ , and each other pre-
 293 dictor. Here, as above, the effects are more pronounced for Γ than T (Fig. 5). For T , only the in-
 294 teraction between δ and ω_h is significant (coefficient = -0.22 SD, p-value = 7.1e-5), and the inter-
 295 action between δ and N_h/N_v is just shy of significant (coefficient = -0.18, p-value = 0.073). In
 296 contras, for Γ , δ has a significant interaction with four other predictors: (1) $\delta:\omega_h$ (coefficient = -
 297 0.26 SD, p-value = 1.6e-10); (2) $\delta:\omega_v$ (coefficient = 0.23 SD, p-value = 5.22e-8); (3) $\delta:N_h/N_v$ (co-
 298 efficient = -0.31 SD, p-value = 1.1e-5); (4) $\delta:\omega_v:N_h/N_v$ (coefficient -0.096, p-value = 0.026). The
 299 interaction between δ and v_{max} just is short of significant (coefficient = 0.079 SD, p-value =
 300 0.094).



301 **Figure 5.** Interactions between dispersal mode (δ) and other predictors of (a) T , the number of
 302 generations until a pathogen population evolves a within-host performance phenotype within
 303 10% of the optimum, and (b) Γ , before T , the degree with which correlation evolution in the
 304 pathogen population is dominated by the vector ($\Gamma > 0$) or host ($\Gamma < 0$) environment. Points show

305 the positive (teal) and negative (pink) coefficients, and horizontal lines give 95% confidence in-
 306 tervals.

307 So, for Γ , doing away with vector-borne dispersal reduces the importance of selection within and
 308 outside of large and susceptible hosts. Conversely, it increases the importance of the predomi-
 309 nance of the large and susceptible hosts, $\delta:N_h/N_v$, and steepens the interaction between that quan-
 310 tity and the weakness of selection. It also seems to augment the maximum virulence effect. In
 311 sum, when we remove constraints on dispersal that force pathogens to move between trophic lev-
 312 els, we make selection in each trophic level less consequential, and we make demography more
 313 consequential. Even without vector-bourne dispersal – and perhaps because of the volatility of
 314 large and susceptible host resources (Olofsson, Ripa and Jonzén, 2009), which is exacerbated by
 315 high virulence – the weakness of selection outside of large and susceptible hosts can still affect
 316 correlational evolution within such hosts. But vector-bourne dispersal makes such effects more
 317 significant.

318 To wrap up the analysis, let us put everything together in a structural equation model (Fig 6.).
 319 Can selection in vectors on antagonistically pleiotropic loci affect correlational and potentially
 320 non-adaptive evolution of virulence in the host? It seems so. Weakening selection in the vector
 321 significantly – that is, increasing ω_v – reduces T (coefficient = -0.15 SD; p-value < 1e-4) and Γ
 322 (coefficient = -0.5 SD; p-value < 1e-4). Conversely, weakening selection in the host environment
 323 tends to increase T (coefficient = 0.56 SD; p-value < 1e-4) and Γ (coefficient = 0.43 SD; p-value
 324 < 1e-4). But the frequency at which pathogen genotypes encounter host or vector environments
 325 is also important; Larger values for N_h tend to reduce T (coefficient = -0.19 SD; p-value < 1e-4)
 326 and Γ (coefficient = -0.22 SD; p-value < 1e-4). Conversely, increasing the total number of
 327 pathogens in vectors, N_v significantly increases Γ (coefficient 0.32 SD, p-value < 1e-4) and T
 328 (coefficient 0.39, p-value < 1e-4). So, the structural equation models is telling us that habitat type
 329 frequency (i.e. the values for N_h and N_v) certainly has a powerful effect on the evolution of the
 330 pathogen population, as per previous theoretical work (Via and Lande, 1985; Whitlock, 1996;
 331 Hardy and Forister, 2023), but these effects are of a smaller magnitude to the strength of selec-
 332 tion in each habitat type, and selection in vectors is just about as important as selection in hosts.

333 The evolution of virulence in hosts also depends strongly on pathogen transmission rate, m , and
 334 the upper limit of the virulence effect, v_{max} . Transmission rate has a negative effect on T (coeffi-
 335 cient = -0.23; p-value < 1e-4), and Γ (coefficient = 0.18; p-value < 1e-4). Again, this is in keep-
 336 ing with previous theoretical work that has shown that in simple pathosystems the optimal level
 337 of virulence increases with pathogen transmission rate, as it attenuates the cost of increased host
 338 mortality (Bull and Luring, 2014). Before fitting the model, we hypothesized that the v_{max} pa-
 339 rameter could affect pathogen evolution in two ways. It could affect T and Γ directly by changing
 340 the adaptive landscape, to wit, by reducing the maximum productivity of host patches. Or, it
 341 could affect T and Γ indirectly, by reducing N_h . The model shows that both are important; v_{max}
 342 has a strong negative effect on N_h (coefficient = -0.54; p-value < 1e-4) as well as strong positive
 343 direct effects on T and Γ (coefficient for T = 0.21; p-value < 1e-4; coefficient for Γ = 0.31; p-
 344 value < 1e-4).

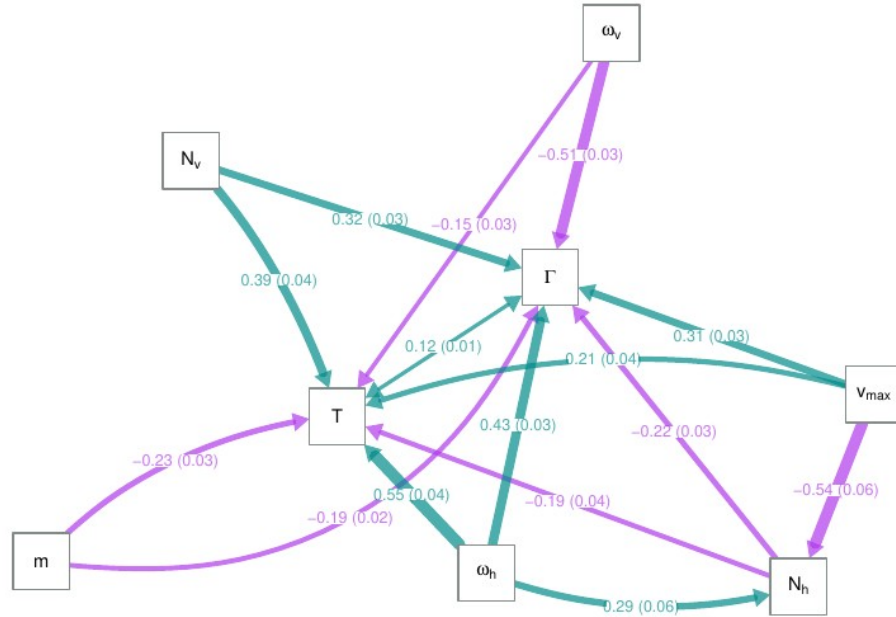


Figure 6. Structured equation analysis of model of the evolution of virulence in a vector-borne pathogen. Edges show positive (teal) and negative (purple) causal relationships among model parameters (ω_h , ω_v , m , v_{max}) and variables (N_h , N_v , T , Γ). The width of each edge is in proportion to the magnitude of its effect. Effect coefficients are printed on each edge, followed by its standard error in parentheses.

Conclusions

Let us recap. Our goal was to better understand how the population structure and life history of vector-borne plant pathogens could amend the general rule that selection is more efficient in more common and productive habitats (Whitlock, 1996). More specifically, we wanted to evaluate the plausibility of the hypotheses that the evolution of virulence in hosts could be a correlational response driven by selection for improved performance in vectors. With statistical analyses of the dynamics of individual-based simulation models, we were unable to reject this hypoth-

esis; even when pathogens are much more often found within hosts than vectors, if selection in vectors is sufficiently strong, it can drive correlational evolution in hosts, especially as parameterized by our statistic Γ . On the other hand, we also recovered several strong effects consistent with governing roles for demography, asymmetries in habitat productivity, and virulence per se consistent with previous theory (Via and Lande, 1985; Day and Proulx, 2004; Alizon *et al.*, 2009). Strong selection in vectors is one among many factors that can drive the evolution of virulence in hosts.

Of course these inferences are contingent of the many simplifying assumptions of our model. Here we stress two of the most liberal. The first is the assumptions of instantaneous replacement of host and vector patches. Relaxation of this assumption, and allowing for a more realistic recruitment, would reduce the effective abundance of host patches, and therefore tend to reduce the demographic disparities that counterbalance the selection in vectors. Hence, we doubt that this would bias our analysis against the rejection of our hypothesis. The second is the assumption that within-host and within-vector performance phenotypes evolve exclusively via pleiotropic alleles. Indeed, this was integral to our premise; we wondered if strong antagonistic pleiotropy could suffice to drive correlational evolution of virulence in hosts. Nevertheless relaxation of this assumption could shed light what other genetic architectural contingencies could be important; indeed, many general questions have yet to be answered about how adaptation to heterogeneous environments depends on the genetic architectures of the traits under selection (Kimbrell and Holt, 2007; Kawecki, 2008; Bridle, Kawata and Butlin, 2019). Here, suffice it to say that our insights into the evolution of virulence in vector-borne plant pathogens depend on assumptions, the relaxation of which could yield a richer and understanding of virulence evolution.

To close, let us reconsider the evolution of virulence in *Xylella fastidiosa*. In California vineyards, the emergence of new highly-virulent genotypes closely followed the establishment of a new, markedly inefficient vector species (Hopkins, 1989). Could this have been because of negative genetic correlations between traits affecting performance within vectors and hosts (Gilbertson *et al.*, 2015)? Our simulation model suggests the answer is yes, possibly. Strong selection for improved performance in the vector does seem capable, in at least certain situations, of causing

correlational evolution of pleiotropic host-performance traits. (And in theory this could also arise via tight linkage rather than pleiotropy per se (Via and Lande, 1985)). Moreover, if such correlational evolution in the host causes a non-adaptive increase in virulence, that is, increased v_{max} , the demographic consequences could further bend the evolutionary path towards the vector optimum. And this could further interfere with overall life-history optimization.

Of course using this hypothesis to explain the evolution of virulence in *Xylella* presupposes that there are some strong antagonistic pleiotropies affecting performance in hosts and vectors. But this is quite likely. Indeed, much of the virulence of *Xylella* infections has been attributed to the plastic induction of “sticky” cell phenotypes which can clog xylem vessels, but also increase the efficiency of acquisition by vectors (Chatterjee, Wistrom and Lindow, 2008; Killiny and Almeida, 2014). That being said, there are other tenable hypotheses for the evolution of increased virulence in *X. fastidiosa* in Californian vineyards. In particular, in addition to be an especially poor vector, *H. vitripennis* is also exceptionally polyphagous. Hence, the story of virulence evolution in Californian populations of *X. fastidiosa* likely also entails changes in their population structure, perhaps increasing the alpha diversity of pathogen communities and the potential for phenotypic evolution via recombination (Gilbertson *et al.*, 2015). Nevertheless, we cannot yet reject the hypothesis that much of the evolution of within-host virulence can be traced back to selection in vectors.

Data Accessibility

This work is based on analyses of the dynamics of evolutionary simulation models. The code for these models is provided as a supplementary document.

Competing Interests

The authors have none to declare.

408 **Acknowledgments**

409 This work was supported by the Alabama Agricultural Experiment station and the Research Sup-
 410 port Program of the Auburn University Office of the Senior Vice President for Research and
 411 Economic Development.

412 **Literature cited**

Alizon, S. *et al.* (2009) ‘Virulence evolution and the trade-off hypothesis: history, current state of affairs and the future’, *Journal of Evolutionary Biology*, 22(2), pp. 245–259. Available at: <https://doi.org/10.1111/j.1420-9101.2008.01658.x>.

Anderson, R.M. and May, R.M. (1982) ‘Coevolution of hosts and parasites’, *Parasitology*, 85(2), pp. 411–426. Available at: <https://doi.org/10.1017/S0031182000055360>.

Bates, D. *et al.* (2003) ‘lme4: Linear Mixed-Effects Models using “Eigen” and S4’. Available at: <https://doi.org/10.32614/CRAN.package.lme4>.

Bridle, J.R., Kawata, M. and Butlin, R.K. (2019) ‘Local adaptation stops where ecological gradients steepen or are interrupted’, *Evolutionary Applications*, 12(7), pp. 1449–1462. Available at: <https://doi.org/10.1111/eva.12789>.

Bull, J.J. and Luring, A.S. (2014) ‘Theory and Empiricism in Virulence Evolution’, *PLOS Pathogens*, 10(10), p. e1004387. Available at: <https://doi.org/10.1371/journal.ppat.1004387>.

Chatterjee, S., Wistrom, C. and Lindow, S.E. (2008) ‘A cell–cell signaling sensor is required for virulence and insect transmission of *Xylella fastidiosa*’, *Proceedings of the National Academy of Sciences*, 105(7), pp. 2670–2675. Available at: <https://doi.org/10.1073/pnas.0712236105>.

Daugherty, M.P. *et al.* (2011) ‘Vector preference for hosts differing in infection status: sharp-shooter movement and *Xylella fastidiosa* transmission’, *Ecological Entomology*, 36(5), pp. 654–662. Available at: <https://doi.org/10.1111/j.1365-2311.2011.01309.x>.

Day, T. and Proulx, S.R. (2004) ‘A General Theory for the Evolutionary Dynamics of Virulence.’, *The American Naturalist*, 163(4), pp. E40–E63. Available at: <https://doi.org/10.1086/382548>.

Draghi, J.A. (2021) ‘Asymmetric Evolvability Leads to Specialization without Trade-Offs’, *The American Naturalist*, 197(6), pp. 644–657. Available at: <https://doi.org/10.1086/713913>.

EFSA (2016) ‘Update of a database of host plants of *Xylella fastidiosa*: 20 November 2015’, *EFSA Journal*, 14(2). Available at: <https://efsa.onlinelibrary.wiley.com/doi/abs/10.2903/j.efsa.2016.4378> (Accessed: 24 June 2024).

- Ewald, P.W. (1983) 'Host-Parasite Relations, Vectors, and the Evolution of Disease Severity', *Annual Review of Ecology and Systematics*, 14(1), pp. 465–485. Available at: <https://doi.org/10.1146/annurev.es.14.110183.002341>.
- Frank, S.A. (1996) 'Models of Parasite Virulence', *The Quarterly Review of Biology*, 71(1), pp. 37–78. Available at: <https://doi.org/10.1086/419267>.
- Gilbertson, R.L. *et al.* (2015) 'Role of the Insect Supervectors *Bemisia tabaci* and *Frankliniella occidentalis* in the Emergence and Global Spread of Plant Viruses', *Annual Review of Virology*, 2(Volume 2, 2015), pp. 67–93. Available at: <https://doi.org/10.1146/annurev-virology-031413-085410>.
- Haller, B.C. and Messer, P.W. (2023) 'SLiM 4: Multispecies Eco-Evolutionary Modeling', *The American Naturalist*, 201(5), pp. E127–E139. Available at: <https://doi.org/10.1086/723601>.
- Hardy, N.B. and Forister, M.L. (2023) 'Niche Specificity, Polygeny, and Pleiotropy in Herbivorous Insects', *The American Naturalist*, 201(3), pp. 376–388. Available at: <https://doi.org/10.1086/722568>.
- Holt, R.D. *et al.* (2003) 'Parasite establishment in host communities', *Ecology Letters*, 6(9), pp. 837–842. Available at: <https://doi.org/10.1046/j.1461-0248.2003.00501.x>.
- Hopkins, D.L. (1989) 'Xylella fastidiosa: xylem-limited bacterial pathogen of plants', *Annual Review of Phytopathology*, 27, pp. 271–290.
- Hopkins, D.L. and Purcell, A.H. (2002) 'Xylella fastidiosa: Cause of Pierce's Disease of Grapevine and Other Emergent Diseases', *Plant Disease*, 86(10), pp. 1056–1066. Available at: <https://doi.org/10.1094/PDIS.2002.86.10.1056>.
- Johnson, K.A. *et al.* (2023) 'Prevalence and Distribution of Phony Peach Disease (Caused by *Xylella fastidiosa*) in the United States', *Plant Disease*, 107(2), pp. 326–334. Available at: <https://doi.org/10.1094/PDIS-03-22-0653-RE>.
- Kawecki, T.J. (2008) 'Adaptation to Marginal Habitats', *Annual Review of Ecology, Evolution, and Systematics*, 39(1), pp. 321–342. Available at: <https://doi.org/10.1146/annurev.ecolsys.38.091206.095622>.
- Killiny, N. and Almeida, R.P.P. (2014) 'Factors Affecting the Initial Adhesion and Retention of the Plant Pathogen *Xylella fastidiosa* in the Foregut of an Insect Vector', *Applied and Environmental Microbiology*, 80(1), pp. 420–426. Available at: <https://doi.org/10.1128/AEM.03156-13>.
- Kimbrell, T. and Holt, R.D. (2007) 'Canalization Breakdown and Evolution in a Source-Sink System.', *The American Naturalist*, 169(3), pp. 370–382. Available at: <https://doi.org/10.1086/511314>.

Olofsson, H., Ripa, J. and Jonzén, N. (2009) ‘Bet-hedging as an evolutionary game: the trade-off between egg size and number’, *Proceedings of the Royal Society B: Biological Sciences*, 276(1669), pp. 2963–2969. Available at: <https://doi.org/10.1098/rspb.2009.0500>.

Redak, R.A. *et al.* (2004) ‘The biology of xylem fluid-feeding insect vectors of *Xylella Fastidiosa* and their relation to disease epidemiology’, *Annual Review of Entomology*, 49(Volume 49, 2004), pp. 243–270. Available at: <https://doi.org/10.1146/annurev.ento.49.061802.123403>.

Rosseel, Y *et al.* (2017) ‘Package “Ivaan”’.

Saponari, M. *et al.* (2017) ‘Isolation and pathogenicity of *Xylella fastidiosa* associated to the olive quick decline syndrome in southern Italy’, *Scientific Reports*, 7(1), p. 17723. Available at: <https://doi.org/10.1038/s41598-017-17957-z>.

Simpson, A.J.G. *et al.* (2000) ‘The genome sequence of the plant pathogen *Xylella fastidiosa*’, *Nature*, 406(6792), pp. 151–157. Available at: <https://doi.org/10.1038/35018003>.

Trkulja, V. *et al.* (2022) ‘*Xylella fastidiosa* in Europe: From the Introduction to the Current Status’, *The Plant Pathology Journal*, 38(6), pp. 551–571. Available at: <https://doi.org/10.5423/PPJ.RW.09.2022.0127>.

Via, S. and Lande, R. (1985) ‘Genotype-environment interaction and the evolution of phenotypic plasticity’, *Evolution*, 39(3), pp. 505–522. Available at: <https://doi.org/10.1111/j.1558-5646.1985.tb00391.x>.

Whitlock, M.C. (1996) ‘The Red Queen Beats the Jack-Of-All-Trades: The Limitations on the Evolution of Phenotypic Plasticity and Niche Breadth’, *The American Naturalist*, 148, pp. S65–S77. Available at: <https://doi.org/10.1086/285902>.