Epidemiological and ecological consequences of virus manipulation of host and vector in plant virus transmission

Nik J. Cunniffe[1], Nick P. Taylor[1], Frédéric M. Hamelin[2], Michael J. Jeger[3]

[1]Department of Plant Sciences, University of Cambridge, Cambridge CB2 3EA, UK
[2]IGEPP, INRAE, Institut Agro, Univ Rennes, Rennes, France
[3]Department of Life Sciences, Imperial College London, Ascot SL5 7PY, UK

ABSTRACT

Many plant viruses are transmitted by insect vectors. Transmission can be described as persistent or non-persistent depending on rates of acquisition, retention, and inoculation of virus. Much experimental evidence has accumulated indicating vectors can prefer to settle and/or feed on infected versus noninfected host plants. For persistent transmission, vector preference can also be conditional, depending on the vector's own infection status. Since viruses can alter host plant quality as a resource for feeding, infection potentially also affects vector population dynamics. Here we use mathematical modelling to develop a theoretical framework addressing the effects of vector preferences for landing, settling and feeding – as well as potential effects of infection on vector population density – on plant virus epidemics. We explore the consequences of preferences that depend on the host (infected or healthy) and vector (viruliferous or nonviruliferous) phenotypes, and how this is affected by the form of transmission, persistent or non-persistent. We show how different components of vector preference have characteristic effects on both the basic reproduction number and the final incidence of disease. We also show how vector preference can induce bistability, in which the virus is able to persist even when it cannot invade from very low densities. Feedbacks between plant infection status, vector population dynamics and virus transmission potentially lead to very complex dynamics, including sustained oscillations. Our work is supported by an interactive interface https://plantdiseasevectorpreference.herokuapp.com/. Our model reiterates the importance of coupling virus infection to vector behaviour, life history and population dynamics to fully understand plant virus epidemics.

INTRODUCTION

Plant diseases have impacts on crops, affecting yield, and on natural plant ecosystems and landscapes, affecting population and community structure (Boyd et al. 2013; Savary et al. 2019; Ristaino et al. 2021). Plant diseases caused by viruses can be particularly damaging (Jones and Naidu 2019). Most plant viruses are transmitted by vectors, which both acquire virus from infected plants and inoculate virus to healthy plants. These vectors can come from a range of insect families, although aphids, whiteflies and other hemipterans with piercing mouthparts are particularly important (Bragard et al. 2013). Transmission plays a critical part in disease epidemiology and depends to a large extent on vector life history and behaviour concerning movement, landing, settling, feeding, and reproduction on plants. These aspects are directly linked to the mode of transmission, whether non-persistent, semi-persistent, or persistent (Whitfield, 2015). In non-persistent transmission the virus is restricted to the insect's stylet, in semi-persistent transmission the virus enters the insect's foregut, and in persistent transmission the virus passes through the gut to the haemolymph and then to the salivary glands. These modes of transmission can be characterised in large part by the rates of vector

acquisition and inoculation and the retention time within the vector (Jeger et al., 2009). The distinction between non-persistent (transient) and persistent (intimate) transmission is particularly marked. Hence, attempts to determine what drives a plant virus epidemic should not only aim to characterise the plant-virus interaction, but also the plant-vector and virus-vector interactions.

Epidemiological models have often assumed that vector transmission occurs essentially at random (Cunniffe et al., 2015). However, given the ongoing accumulation of experimental evidence suggesting more complexity, such an assumption when used in developing mathematical models is often an over-simplification (Jones 2014). Vectors show preferences in their interactions with plants that can directly influence the rate of transmission. Preference can take many forms, with much past work focusing on arthropod vectors of vertebrate-host parasites (Kingsolver, 1987; Chamchod & Britton, 2011; Gandon, 2018; Cozzarolo et al., 2020). Focusing on plant viruses in particular, vector preference can be expressed in different ways: (i) host preference within a vector's host range; (ii) preference for the host phenotype, restricted in the sense here to whether the host is infected or healthy; and (iii) conditional preference, whereby the preference for infected or healthy hosts depends on whether the vector is viruliferous or non-viruliferous.

Plant viruses are often termed as generalist or specialist in relation to their host range. Yet, such classification makes little sense for insect-transmitted plant viruses unless account is taken of the vector life history, host preferences (for landing, settling, feeding and reproduction) and transmission characteristics. Host range preferences have been well characterised for many insect vectors, but there is not always a simple relationship with insect life history (Fariña et al. 2019) or plant traits (Lagos-Kutz et al. 2020). The second aspect of vector preference we identify, host phenotype preference, refers to the vector's preference for infected or healthy hosts, irrespective of whether the vector is non-viruliferous or viruliferous. Preferences for landing and feeding are dependent on sensory cues which can be olfactory as well as visual (Bosque-Perez & Eigenbrode (2011)). However, as pointed out by Sisterson (2008), it important to distinguish landing, or orientation, preference from feeding preference, since both can potentially affect pathogen spread differently.

For non-persistent transmission, it seems likely that inhibition of settling while allowing probing would encourage transmission, whereas prolonged settling would retard transmission. However, mathematical modelling indicates that both inhibition and prolonged settling can contribute to epidemic development, for example via density dependent production of alate (winged) forms on crowded host plants (Donnelly et al., 2019; Carr et al. 2020). For persistent transmission, thrips species have long been reported to have settling and feeding preferences for tospovirus-infected plants. Such preferences can affect thrips life history traits. Virus acquisition occurs with early instars and the developmental period to adults was shortened when *Thrips palmi* had acquired virus from groundnut bud necrosis virus-infected plants (Daimei et al. 2017). Such a shortening may facilitate the early inoculation of healthy plants by inoculative adults. The effects of volatile emissions on vector behaviour cannot always be extrapolated from one species to another (Claudel et al., 2018). More *Bemisia tabaci* whiteflies landed and settled on plants infected with squash vein yellowing virus than non-infected squash, but the opposite was found with watermelon (Shresthra et al. 2017). Whitefly behaviour differed between the two cucurbit hosts but integrating the various life history traits into a comparison of potential disease dynamics suggested a more rapid spread of the virus in watermelon fields.

Where a virus modifies plant quality sufficiently to improve it as a feeding resource, this can lead to an accompanying increase in vector abundance termed as pathogen-mediated insect superabundance (Donnelly and Gilligan, 2020). However, vector preferences are not always consistent across different hosts, even when virus-infected plants are generally higher quality hosts (Tungadi et al., 2017). The effect of insects feeding on plants infected with viruses they do not vector has also been studied. When *B. tabaci* preferentially feeds on different hosts infected with tomato spotted wilt virus (which is thrips-vectored), body size, longevity and fecundity were all reduced (Chen et al., 2017) indicating that the initial preference for a virus-infected plant was induced by host volatiles and not by subsequent performance. However, host volatiles are not always the cue for settling and feeding preference. For the beetle-transmitted bean pod mottle virus, the beetles are more attracted to infected soybean plants, which have dramatically higher sucrose levels, and although beetles consumed less leaf foliage per plant, they fed on more plants per unit of time when they had high levels of sucrose (Smith et al. 2017). Pea aphid clones adapted to either pea or alfalfa were tested to see how bean leafroll virus affected their performance and preference (Davis et al. 2017). Aphid clone x host plant species x virus status interactions and unique virus-association phenotypes were found.

Effects on vector fitness can also be seen, for example, on squash infected with papaya ringspot virus, transmitted by *Aphis gossypii* (Gadhave et al. 2019). The overall performance of *A. gossypii* was substantially higher, with an extended settling (arrestment), on infected plants when compared to on healthy plants. This effect was not present with the non-vector *B. tabaci*. Hence there was a positive fitness effect on the vector in what would be considered a transient vector-virus interaction. Equally, effects on plant fitness can be negative in the longer term (Shi et al. 2021). It was found that initial feeding of *Myzus persicae* on *Nicotiana tabacum* infected with cucumber mosaic virus reduced the reproduction rate and longevity of aphids subsequently introduced to the previously foraged plants.

As well as affecting feeding, virus-induced changes in host phenotype can also affect vector dispersal and disease spread, depending on transmission mode and efficiency (Chesnais et al. 2019). In studies of turnip yellows virus on *Montia perfoliata* transmitted by *Brevicoryne brassicae* in the persistent mode (Chesnais et al. 2020), it was found that non-viruliferous aphids showed greater fecundity only on infected plants and had reduced dispersal and activity levels. However, viruliferous aphids showed greater dispersal and activity levels with a greater fecundity and efficiency in feeding irrespective of plant infection status.

Just as plant viruses can modify the host phenotype and affect insect vector behaviour and population dynamics, modification of the vector phenotype so that preference is conditional on whether the vector is non-viruliferous or viruliferous can also affect disease dynamics. For conditional vector preference, the question arises: whether, and under what circumstances, the virus is manipulating the plant host and the vector to its own advantage? If this is indeed the case and it can be shown that there is a genetic basis to such manipulation, then the question takes on an evolutionary as well as epidemiological dimension. The “Virus manipulation hypothesis” that changes in vector behaviour are induced by both host and vector phenotype was proposed by Ingwell et al. (2012). Accordingly, non-persistent and persistent transmission of viruses will have different effects on vector preferences for landing, settling, feeding, and dispersal from infected and healthy plants (Fereres and Moreno, 2009, Mauck et al. 2012, 2016). Any difference in reported effect seems to depend on whether probing or settling was evaluated in studies (Carmo-Souza et al. 2014; Rajabaskar et al. 2014). Overall, the evidence supports the view that viruses manipulate both host and vector to enhance transmission

(Mauck et al. 2018) and hence fitness. A substantial body of evidence was reviewed by Eigenbrode et al. (2018), albeit with some inconsistency across findings (see Tables 1, 2 and Figure 1 in Eigenbrode et al. 2018). This inconsistency may be due to a lack of occurrence or, despite the large number of studies reviewed, a lack of information. There appears to be a remarkable degree of convergence among unrelated viruses with similar transmission characteristics (Mauck et al. 2018, 2019). Predictions were tested for non-persistent, semi-persistent and persistent transmission, with adaptive manipulation most apparent in the latter case. However, although the main factors influencing transmission and selection for manipulative traits have been identified, there are important gaps in linking findings with evolutionary processes (Mauck & Chesnais 2020), especially the molecular and environmental constraints on virus manipulation. A further constraint is the difficulties of exploring the consequences of virus manipulation, where this has been shown in laboratory or microcosm studies, in field settings in crops or wild plant populations. Of course, mathematical modelling presents one way of dealing with this constraint (Cunniffe & Gilligan 2020; Fabre et al. 2021).

Mathematical modelling has contributed to analysis of the epidemiological consequences of vector preferences as seen in the model of Sisterson (2008). Previously, in modelling the spread of barley yellow dwarf virus, it was found that with a low incidence of infected plants, disease spread is favoured by vectors preferring infected plants, whereas with a high incidence, spread is favoured by vectors favouring healthy plants (McElhany et al. 1995). However, this initial model did not distinguish the preferences of viruliferous and non-viruliferous vectors. Preferences of vectors for healthy or infected host plants will affect both aggregation and dispersal of vectors (Hebert and Allen, 2016). Assumptions on birth and death rates of vectors based on their infectivity status can have important effects on both vector population dynamics and should be included in epidemiological models (Sisterson, 2009). Including conditional preference depending on vector phenotype into models, showed that a switch in preference once a vector acquires virus from infected plants can enhance spread (Roosien et al. 2013).

A more comprehensive model with conditional vector preferences but also including more vector life history traits, especially dispersal, has also been developed (Shaw et al. 2017; as corrected in Shaw et al. 2018). Traits including intrinsic growth rate, population carrying capacity, and landing and departure rates, were introduced conditional upon whether the host is healthy or infected and whether the vector is viruliferous or non-viruliferous. The form in which preference was introduced in this model, different to Roosien et al. (2013), used a phenomenological response in which the fraction of infected plants was raised to a power characteristic of the vector preference. This required an arbitrary correction factor (their parameter ξ) to be included "to keep the model well-behaved", since the model equations do not include any constraint stopping the fraction of infected plants from going outside the range [0,1]. It is also unclear that recovery of viruliferous vectors is modelled correctly, since both the model equations and the associated schematic indicate that viruliferous vectors on healthy hosts "recover" to become non-viruliferous vectors on infected hosts (since the γV_h term links the V_h and N_i compartments in the model equations). The model was parameterised for barley yellow dwarf virus and potato virus Y, although – given the interaction between density and rate in the transmission terms – it is difficult to follow from what is presented in the paper how this was done, precisely. The key result from numerous simulations indicated that vector population growth rates overall had the greatest effect on virus spread, with rates of vector dispersal from hosts of the same virus status as the vector also important. These interpretations were based mainly on numerical

simulations of the model with a global sensitivity analysis based on a partial rank correlation coefficient technique for the time taken to reach 80% of hosts infected.

A further model describing vector preferential behaviour and how this affects transmission was developed specifically for the tospovirus tomato spotted wilt virus vectored by the thrips species *Frankliniella occidentalis* (Ogada et al., 2016). In this system there is persistent-propagative transmission which is also transstadial: acquisition only occurs at the larval stages, the virus replicates within the vector during the development stages, and mobile adults are then able to inoculate virus, as previously modelled (Jeger et al., 2015). Ogada et al. (2016) included a linear relative preference term to reflect that viruliferous adults prefer healthy plants while non-viruliferous adults prefer infected plants. Thus, eggs which are laid on infected plants would likely result in larvae which acquire virus. In general, in this paper, we ignore the direct effects of the virus on vector birth and death rates, which are most apparent with viruses that propagate in the vector (Higashi and Bresson, 2013; Chen et al., 2016).

The objective in this paper is to develop a model specific to viral pathogens of plants vectored by insects, building on previous epidemiological models. The model is complex but includes biologically relevant parameters that can be estimated from experimental data, although these data are mostly obtained in microcosm studies rather than from field observations. The distinction between our model and the earlier models is that, following appropriate parameterisation, the model can be used for both non-persistently transmitted and persistently transmitted viruses. A further distinction is that we use analytical methods to derive epidemiological quantities that depend on the model parameters, and to show how model behaviours depend on parameter values. We derive the basic reproduction number and endemic equilibria analytically and use these results to show how transmission type and vector life-history and behaviour interact with conditional vector preference, as well as to determine under which conditions different long-term outcomes are possible depending on initial conditions. To illustrate the commonalities and differences in approaches with previous models of conditional vector preference, the variables and parameters defined are compared in Supplementary Table S1. We also introduce a user-friendly interactive online interface to our model, which offers readers of our paper an opportunity to understand for themselves how changes to underlying epidemiological parameters, as well as those relating to aspects of vector preference, lead to different epidemiological outcomes.

METHODS

Overview of epidemiological approach

A model of plant-virus-vector interactions has been previously introduced to account for the range of transmission classes in plant-virus-vector interactions (Jeger et al. 1998, Madden et al. 2000). Here, we modify that model to account for the effects of conditional vector preference on life history from landing, settling, and feeding, to reproduction and flight. In addition, more detail on vector population dynamics is included as an important element in virus epidemiology (Strauss et al. 2020). The following assumptions were made:

1. Vectors are attracted by plant cues (visual or olfactory) to land on infected plants to a greater or lesser extent.
2. Whether or not vectors then settle and feed for an extended period, or only probe and then depart, also depends on the plant's infection status.

3. The strength of vector preference can differ for viruliferous and non-viruliferous vectors, i.e. preference is conditional on vector status as well as plant infection status. We note that in our analysis of the model we assume that conditional preference is relevant only for persistent transmission; for non-persistent transmission we assume that viruliferous and non-viruliferous vectors behave similarly.
4. The proportion of probes that leads to vectors settling for an extended feed affects the number of plants visited by vectors per unit of time, and so the overall transmission rate.
5. Whether vectors probe or feed has different effects on transmission for non-persistent vs persistent viruses.
6. The fecundity of the vector can be affected by the plants it feeds on, with vectors that predominantly feed on infected plants potentially having either a higher or lower birth rate.
7. The loss rate of the vector, whether from additional mortality or the chance of movement away from the plant population, may be affected by the number of distinct plants visited per extended feed, reflecting the possibility that vectors must travel between plants more often if the number of plants visited per feed is increased.
8. The flight pattern of a vector (duration of flights) may depend on whether it is viruliferous or non-viruliferous.

In our model we account for the effects of different vector preferences for healthy susceptible vs. infected plants, potentially distinguishing different preferences of non-viruliferous vs. viruliferous vectors, as well as preference for probing plants vs. settling for an extended feed. As we describe below, much of the complexity of our model comes from properly accounting for the accumulative effects of these behaviours on: i) numbers of contacts between vectors and hosts per unit time, and ii) vector population dynamics.

The basic model

We assume that there is no latent period in the plant or vector population, no propagative or transovarial transmission of the virus, and no immigration or emigration of vectors. The underlying model is (see also Figure 1):

$$\begin{aligned}
\frac{dS}{dt} &= \rho N - \Lambda(S,I,Z) - \rho S, \\
\frac{dI}{dt} &= \Lambda(S,I,Z) - (\rho + \mu)I, \\
\frac{dX}{dt} &= g(S,I,X,Z) + \tau Z - \Omega(S,I,X) - h_{-}(S,I)X, \\
\frac{dZ}{dt} &= \Omega(S,I,X) - \tau Z - h_{+}(S,I)Z.
\end{aligned} \tag{1}$$

The variables S and I in Equation (1) are the densities (per unit area) of healthy susceptible and infected plants, respectively, with N the plant population density in the absence of disease. The variables X and Z are the densities of non-viruliferous and viruliferous vectors respectively, with κ the vector population density in the absence of disease. A list of all model variables and parameters in the model is given in Table 1. Plants are removed by harvesting or natural mortality at rate ρ irrespective of disease status. Infected plants are also subject to an additional removal at rate μ, which could

correspond to disease-induced mortality or alternatively to roguing of infected plants (i.e. systematic removal of symptomatic plants). Losses in the plant population are replenished by replanting with susceptible plants at a fixed rate ρN, i.e proportional to the population density in the absence of disease. The flexibility of the parameters ρ and μ mean that the model could refer to a managed crop population or a natural plant population. The rate at which viruliferous vectors cease to be inoculative is τ. The model includes birth of vectors at rate $g(S, I, X, Z)$ with all vectors born non-viruliferous and in which the vector birth rate is averaged over the vector population and so depends on the fraction of plants that are healthy S or infected I. Death of non-viruliferous and viruliferous vectors occurs at per capita rates $h_{-}(S, I)$ and $h_{+}(S, I)$ respectively, where these rates again account for the infection statuses of the plant population. The S and I dependences in vector birth and death rates allow the consequent effects of vector preference on vector population dynamics to be tracked.

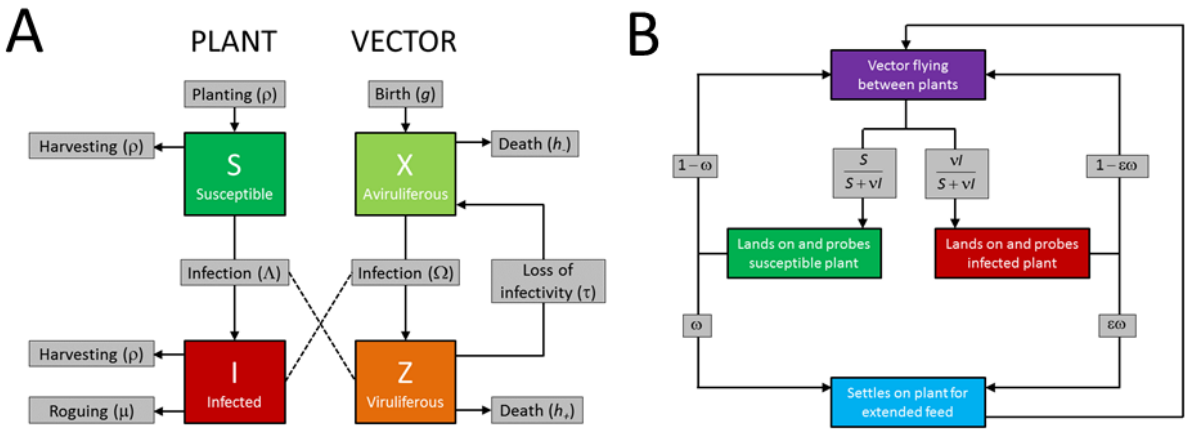


Figure 1. Model structure. (A) Compartments in the epidemiological model. All arrows corresponding to transitions causing hosts or vectors to leave a compartment are labelled with per capita rates, apart from those corresponding to infection of hosts or acquisition of the virus by vectors (i.e. from $S \to I$ and from $X \to Z$) which are labelled with net rates. The rate of planting of hosts and birth of vectors are also labelled with net rates. **(B)** Schematic showing how the movement, settling, and feeding behaviours of vectors are modelled. These affect the infection terms (i.e. Λ and Ω) in the epidemiological model by controlling the probability of transmission (Equation (26)), as well as the average rate at which vectors visit plants (Equations (11) and (13)). For persistent transmission (PT), the values of ν, ω and ϵ can be different for viruliferous and non-viruliferous vectors (e.g. for a PT virus, the bias to land on infected plants is ν_{-} for non-viruliferous vectors, ν_{+} for viruliferous vectors). For non-persistent transmission (NPT), the values of these parameters do not depend on the vector's infection status. Note that in the schematic "flying" refers to any movement between plants: for vectors such as aphids that also – or exclusively – produce non-winged forms, this corresponds to movement between plants by crawling.

Apart from the effects of vector preference on the vector's population dynamics, changes relative to the model of Jeger et al. (1998) largely focus on altering the transmission rates to account for the preferences of non-viruliferous vs viruliferous vectors to land, probe and/or feed on uninfected vs. infected plants. In Equation (1), the composite parameters $\Lambda(S, I, Z)$ and $\Omega(S, I, X)$ control inoculation of healthy plants by viruliferous vectors and acquisition of the virus from infected plants to non-viruliferous vectors, respectively. In the absence of vector preference, the forms we adopt would be

$$\Lambda(S,I,Z) = \phi\gamma Z \frac{S}{S+I}, \qquad \Omega(S,I,X) = \phi\eta X \frac{I}{S+I}. \tag{2}$$

Symbol	Meaning	Value
S	Density of susceptible and healthy plants	-
I	Density of infected plants	-
X	Density of non-viruliferous vectors	-
Z	Density of viruliferous vectors	-
N	Density of plants in the absence of disease	1,000
ρ	Rate at which plants die naturally/are harvested	0.01 d^{-1}
μ	Rate of roguing (or additional disease-induced death)	0.01 d^{-1}
α	Rate at which vectors die in the absence of any effect due to number of flights per feed	0.12 d^{-1}
τ	Rate at which vectors lose infectivity	4 d^{-1} (NPT) 0.05 d^{-1} (PT)
σ	Per capita rate at which vectors are born as the vector population density tends to zero	0.18 d^{-1}
ζ	Vector population density at which density dependence reduces the birth rate to zero	1977.6 (NPT) 163.2 (PT)
Γ	Average time spent feeding on a plant when vector chooses to settle for extended feed	2 d
δ	Controls any increase in the death rate as more plants are visited per feed	Varied (Default = 0)
β	Proportionate change in birth rate on infected plants	Varied (Default = 1)
ν_-	Bias of non-viruliferous vector to land on infected plants	Varied (Default=1)
ν_+	Bias of viruliferous vector to land on infected plants	Varied (Default=1)
ω_-	Probability that a non-viruliferous vector settles to feed on a susceptible plant	Varied (Default=0.5)
ω_+	Probability that viruliferous vector settles to feed on a susceptible plant	Varied (Default=0.5)
ϵ_-	Bias of non-viruliferous vector to settle to feed on an infected plant	Varied (Default=1)
ϵ_+	Bias of a viruliferous vector to settle to feed on an infected plant	Varies (Default=1)
ϕ_-	The average number of plants visited by a non-viruliferous vector per unit of time	Derived (Equation (10))
ϕ_+	The average number of plants visited by a viruliferous vector per unit of time	Derived (Equation (12))
γ	Probability that an uninfected plant is inoculated by a single visit from a viruliferous vector	Derived (Equation (6))
η	Probability that a non-viruliferous vector acquires the virus in a single visit to an infected plant	Derived
$\Lambda(S,I,Z)$	Overall rate at which uninfected plants become infected	Derived
$\Omega(S,I,X)$	Overall rate at which non-viruliferous vectors become viruliferous	Derived
$g(S,I,X,Z)$	Rate at which vectors are born	Derived
$h_-(S,I)$	Rate (per capita) at which non-viruliferous vectors die	Derived
$h_+(S,I)$	Rate (per capita) at which viruliferous vectors die	Derived
κ	Total density of vectors in the absence of disease	Derived
R_0^{PV}	Average number of vectors directly acquiring virus by the introduction of a single infected plant	Derived
R_0^{VP}	Average number of plants directly inoculated by the introduction of a single viruliferous vector	Derived
R_0	Basic reproduction number	Derived ($R_0^2 = R_0^{PV} R_0^{VP}$)

Table 1. Summary of symbols and their meanings, with default values used in numerical work. NPT indicates non-persistent transmission, PT persistent transmission.

The overall rate at which susceptible plants are inoculated, $\Lambda(S, I, Z)$, is therefore the product of: ϕ, the average number of plants visited by each vector per unit of time; γ, the probability that an uninfected plant is inoculated with the virus during a single visit by a viruliferous vector; Z, the density of vectors that are currently viruliferous; and $S/(S+I)$, the probability that a single visit by a vector is made to an uninfected plant. Similarly, the overall rate at which non-viruliferous vectors acquire infection, $\Omega(S, I, X)$, depends on: ϕ, the rate at which plants are visited; η, the probability that an non-viruliferous vector acquires the virus from an infected plant; X, the density of non-viruliferous vectors; and $I/(S+I)$, the probability that an individual visit is made to an infected plant.

Vector landing preference

We extend the basic model by including vector preference. We define ν_- to be the degree to which non-viruliferous vectors prefer to land on infected plants. Non-viruliferous vectors (-) can then be biased either in favour of ($\nu_- > 1$), or against ($\nu_- < 1$), landing on infected plants (*cf.* Figure 1)

$$P(\text{non-viruliferous vector lands on a susceptible plant})=\frac{S}{S+\nu_- I},$$
$$P(\text{non-viruliferous vector lands on an infected plant})=\frac{\nu_- I}{S+\nu_- I}. \quad (3)$$

Hence for $\nu_- \in [0,1)$ a non-viruliferous vector prefers to land on healthy, susceptible plants; for $\nu_- \in (1, \infty)$, it prefers to land on infected plants. For $\nu_- = 1$, there is no bias in landing. An analogous parameter ν_+ controls the landing behaviour of viruliferous vectors (+), with

$$P(\text{viruliferous vector lands on a susceptible plant})=\frac{S}{S+\nu_+ I},$$
$$P(\text{viruliferous vector lands on an infected plant})=\frac{\nu_+ I}{S+\nu_+ I}. \quad (3)$$

Vector feeding preference

We assume that an individual vector probes a plant directly after landing, but then chooses between settling for an extended feed or immediately moving off to probe a different plant. The probability of feeding potentially depends on both the state of the vector (non-viruliferous or viruliferous) and of the plant (healthy or infected). The parameters ω_- and ω_+ are the probabilities that non-viruliferous (-) or viruliferous (+) vectors settle to feed on healthy susceptible plants. The parameters ϵ_- and ϵ_+ then control any bias for ($\epsilon_\pm > 1$) or against ($\epsilon_\pm < 1$) settling on infected plants. For non-viruliferous vectors the probabilities of settling and feeding vs. probing and moving elsewhere are therefore

$$P(\text{non-viruliferous vector settles for extended feed on a susceptible plant})=\omega_-,$$
$$P(\text{non-viruliferous vector probes a susceptible plant but moves elsewhere})=1-\omega_-,$$
$$P(\text{non-viruliferous vector settles for extended feed on an infected plant})=\varepsilon_-\omega_-,$$
$$P(\text{non-viruliferous vector probes an infected plant but moves elsewhere})=1-\varepsilon_-\omega_-. \quad (4)$$

The corresponding probabilities for viruliferous vectors are

$$
\begin{aligned}
P(\text{viruliferous vector settles for extended feed on a susceptible plant}) &= \omega_+, \\
P(\text{viruliferous vector probes a susceptible plant but moves elsewhere}) &= 1-\omega_+, \\
P(\text{viruliferous vector settles for extended feed on an infected plant}) &= \varepsilon_+\omega_+, \\
P(\text{viruliferous vector probes an infected plant but moves elsewhere}) &= 1-\varepsilon_+\omega_+.
\end{aligned} \tag{5}
$$

As well as the obvious constraint that $\omega_\pm \leq 1$, we note that parameters must also be constrained such that $\epsilon_\pm\omega_\pm \leq 1$ in order for all probabilities in Equations (5) and (6) to be well-posed.

Effect of transmission type on the probability of transmission per visit

We account for the different effects of probing vs. feeding on the spread of viruses that are non-persistently transmitted (NPT) vs. persistently transmitted (PT) (Mauck et al. 2012). We assume the probability that uninfected plants are inoculated with the virus during a single visit by a viruliferous vector (γ) depends on the fraction of visits in which it probes or feeds, with

$$
\gamma = \begin{cases} 1 & \text{for NPT viruses} \\ \omega_+ & \text{for PT viruses} \end{cases}, \tag{6}
$$

since PT viruses can only be inoculated by viruliferous vectors if they settle for extended feeds. For NPT viruses we assume that a single visit is sufficient for inoculation.

However, the probability of acquisition by a non-viruliferous vector during a single visit to an infected plant is given by

$$
\eta = \begin{cases} 1-\varepsilon_-\omega_- & \text{for NPT viruses} \\ \varepsilon_-\omega_- & \text{for PT viruses} \end{cases}, \tag{7}
$$

because virions of NPT viruses are detached from the vector's stylet back into the same plant during an extended feed, whereas extended feeding is required to acquire PT viruses (Mauck et al. 2012).

Number of plants visited per vector per unit of time

The distinction between probing and feeding will affect the length of the average visit to a plant, which in turn affects the overall rate at which plants are visited by vectors. We assume the duration of an extended feed by a given vector is fixed at Γ, and that probing is instantaneous compared to feeding. Defining Δ_{--} and Δ_{-+} as the average lengths of visits by non-viruliferous vectors to uninfected and infected plants, respectively

$$
\begin{aligned}
\Delta_{--} &= \omega_- \times \Gamma + (1-\omega_-)\times 0 = \omega_-\Gamma, \\
\Delta_{-+} &= \varepsilon_-\omega_- \times \Gamma + (1-\varepsilon_-\omega_-)\times 0 = \omega_-\varepsilon_-\Gamma.
\end{aligned} \tag{8}
$$

Given the probabilities that non-viruliferous vectors land on uninfected vs. infected plants, the length of the average visit by a non-viruliferous vector is therefore

$$
\Delta_- = \frac{S}{S+\nu_- I}\times\omega_-\Gamma + \frac{\nu_- I}{S+\nu_- I}\times\omega_-\varepsilon_-\Gamma = \frac{\omega_-\Gamma\left(S+\nu_-\varepsilon_- I\right)}{S+\nu_- I}. \tag{9}
$$

If we furthermore assume the time spent by vectors in moving between pairs of plants is very small compared to Γ (where we note in support of this assumption that in our default parameterisation Γ is taken to be of the order of days), then ϕ_-, the average number of plants visited by a non-viruliferous vector per unit of time, is

$$\phi_- = \frac{1}{\Delta_-} = \frac{S + \nu_- I}{\omega_- \Gamma\left(S + \nu_- \varepsilon_- I\right)}. \tag{10}$$

An analogous calculation shows the length of an average visit by a viruliferous vector is

$$\Delta_+ = \frac{S}{S + \nu_+ I} \times \omega_+ \Gamma + \frac{\nu_+ I}{S + \nu_+ I} \times \omega_+ \varepsilon_+ \Gamma = \frac{\omega_+ \Gamma\left(S + \nu_+ \varepsilon_+ I\right)}{S + \nu_+ I}, \tag{11}$$

and the average number of plants visited by a viruliferous vector per unit of time, ϕ_+, is

$$\phi_+ = \frac{1}{\Delta_+} = \frac{S + \nu_+ I}{\omega_+ \Gamma\left(S + \nu_+ \varepsilon_+ I\right)}. \tag{12}$$

Vector population dynamics

Plant host-virus-vector models often consider only the simplified situation in which the population size of the vector was fixed, with births/deaths and immigration/emigration always held balanced by density-dependent factors. However, this assumption makes it impossible to consider any effects on epidemics of conditional vector preference that depend on vector population dynamics (Shaw et al. 2017). We therefore include births and deaths of vectors explicitly in the model.

We assume a general form of logistic growth controls the net birth rate of vectors

$$g(S, I, X, Z) = (\sigma_- X + \sigma_+ Z)\left(1 - \frac{X + Z}{\zeta}\right), \tag{13}$$

in which σ_- is the per capita rate of reproduction of non-viruliferous vectors when there is no density dependence, σ_+ is the same rate for viruliferous vectors, and ζ is the density of the vector population at which the birth rate is zero. The assumption is made that the maximum density of the vector population ζ does not depend on the infection status of the reproducing vectors; instead, we assume the birth rate depends on the infection status of the plants on which vectors feed.

We denote the fraction of feeds on infected plants by $f_-(S, I)$ for non-viruliferous vectors, and $f_+(S, I)$ for viruliferous vectors, where both quantities depend on the proportions of infected vs. uninfected plants in the host population. We assume the maximum rates of reproduction are given by

$$\begin{aligned} \sigma_- &= \sigma\left(1 + \left(\beta - 1\right) f_-(S, I)\right), \\ \sigma_+ &= \sigma\left(1 + \left(\beta - 1\right) f_+(S, I)\right), \end{aligned} \tag{14}$$

where σ is the maximum rate for vectors that feed entirely on uninfected plants and $\sigma\beta$ is the corresponding maximum rate if only infected plants were fed upon (i.e. when $f_\pm = 1$). The parameter β may therefore be interpreted as the relative reproduction rate on infected plants, and so in general values $\beta < 1$ and $\beta > 1$ are possible (as well as the default assumption, $\beta = 1$).

Specification of the model of the vector birth rate is completed by calculating $f_\pm(S, I)$. For non-viruliferous vectors, combining Equations **Error! Reference source not found.** and (4) indicates

$$p_{--} = P(\text{next visit leads to an extended feed on a susceptible plant}) = \frac{\omega_- S}{S + \nu_- I},$$
$$p_{-+} = P(\text{next visit leads to an extended feed on an infected plant}) = \frac{\omega_- \varepsilon_- \nu_- I}{S + \nu_- I}. \tag{15}$$

The fraction of feeds by non-viruliferous vectors that are on infected plants is therefore

$$f_-(S,I) = \frac{p_{-+}}{p_{--} + p_{-+}} = \frac{\omega_- \varepsilon_- \nu_- I}{\omega_- S + \nu_- \omega_- \varepsilon_- I} = \frac{\nu_- \varepsilon_- I}{S + \nu_- \varepsilon_- I}. \tag{16}$$

The corresponding quantity for viruliferous vectors is

$$f_+(S,I) = \frac{\nu_+ \varepsilon_+ I}{S + \nu_+ \varepsilon_+ I}. \tag{17}$$

The full model of the birth rate of vectors is therefore given by

$$g(S,I,X,Z) = \sigma\left((X+Z) + (\beta - 1) I \left(\frac{\nu_- \varepsilon_- X}{S + \nu_- \varepsilon_- I} + \frac{\nu_+ \varepsilon_+ Z}{S + \nu_+ \varepsilon_+ I} \right) \right) \left(1 - \frac{X+Z}{\zeta} \right), \tag{18}$$

where σ is the maximum per capita rate of reproduction for vectors that feed entirely on uninfected plants, β is the relative rate of vector reproduction on infected plants, and ζ is the size of the vector population at which the birth rate is zero.

To model an overall rate of loss of vectors from the system, we assume the per capita rate of death of vectors is also affected by the number of plants visited per vector per extended feed. Moving between plants is energy consuming and puts the vector at additional risk, e.g. of predation. Losses also potentially occur because vectors making flights may simply move away from the plant population, especially non-colonising vectors. The same argument as was used to derive Equation (10) indicates that for non-viruliferous vectors

$$\theta_- = \text{Average number of plants visited by a non-viruliferous vector per feed,}$$
$$= \frac{S + \nu_- I}{\omega_- (S + \nu_- \varepsilon_- I)} = \Gamma \phi_-. \tag{19}$$

The analogous quantity for viruliferous vectors is

$$\theta_+ = \text{Average number of plants visited by a viruliferous vector per feed,}$$
$$= \frac{S + \nu_+ I}{\omega_+ (S + \nu_+ \varepsilon_+ I)} = \Gamma \phi_+. \tag{20}$$

We assume that the per capita vector death rate can include a component that increases linearly with the number of plants visited per feed, taking

$$\mathrm{h}_-(S,I) = \alpha\left(1 + \delta(\theta_- - 1)\right) = \alpha\left(1 + \delta(\Gamma \phi_- - 1)\right),$$
$$\mathrm{h}_+(S,I) = \alpha\left(1 + \delta(\theta_+ - 1)\right) = \alpha\left(1 + \delta(\Gamma \phi_+ - 1)\right). \tag{21}$$

In Equation (21) α is the underlying rate at which vectors die, and δ controls any increase in the death rate as more plants are visited per feed. Since the number of plants visited per feed is constrained via $\theta_\pm \geq 1$, the form we have taken corresponds to an additional death rate for all values $\delta > 0$.

The vector population in the absence of disease is a derived quantity in our model. Denoting this population size by κ, it follows by seeking an equilibrium value of the model when $I = 0$ that

$$0 = g(N,0,\kappa,0) - h_-(N,0)\kappa = \kappa\left[\sigma\left(1-\frac{\kappa}{\zeta}\right) - \alpha\left(1+\delta\left(\frac{1}{\omega_-}-1\right)\right)\right], \tag{22}$$

meaning the vector carrying capacity when the pathogen is absent is given by

$$\kappa = \zeta\left(1-\frac{\alpha}{\sigma}\left(1+\delta\left(\frac{1}{\omega_-}-1\right)\right)\right). \tag{23}$$

The vector population can only be self-sustaining in the absence of disease (i.e. $\kappa > 0$) if the maximum birth rate on uninfected plants is sufficiently large, i.e. when

$$\sigma > \alpha\left(1+\delta\left(\frac{1}{\omega_-}-1\right)\right). \tag{24}$$

Summary of the model

The model is

$$\begin{aligned}
\frac{dS}{dt} &= \rho N - \Lambda(S,I,Z) - \rho S,\\
\frac{dI}{dt} &= \Lambda(S,I,Z) - (\rho+\mu)I,\\
\frac{dX}{dt} &= g(S,I,X,Z) + \tau Z - \Omega(S,I,X) - h_-(S,I)X,\\
\frac{dZ}{dt} &= \Omega(S,I,X) - \tau Z - h_+(S,I)Z,
\end{aligned} \tag{25}$$

in which transmission is controlled by

$$\begin{aligned}
\Lambda(S,I,Z) &= \phi_+\gamma Z\frac{S}{\omega_+\Gamma\left(S+\nu_+\varepsilon_+ I\right)}, \qquad \gamma = \begin{cases} 1 & \text{for NPT viruses} \\ \omega_+ & \text{for PT viruses} \end{cases},\\
\Omega(S,I,X) &= \phi_-\eta X\frac{\nu_- I}{\omega_-\Gamma\left(S+\nu_-\varepsilon_- I\right)}, \qquad \eta = \begin{cases} 1-\varepsilon_-\omega_- & \text{for NPT viruses} \\ \varepsilon_-\omega_- & \text{for PT viruses} \end{cases},
\end{aligned} \tag{27}$$

and where for PT we allow the "plus" preference parameters to differ from the "minus" ones, whereas for NPT related pairs of parameters must be identical (i.e. parameters are constrained such that $\nu_- = \nu_+ = \nu$; $\omega_- = \omega_+ = \omega$ and $\epsilon_- = \epsilon_+ = \epsilon$).

The population dynamics of the vector depend on

$$\begin{aligned}
g(S,I,X,Z) &= \sigma\left((X+Z) + (\beta-1)I\left(\frac{\nu_-\varepsilon_- X}{S+\nu_-\varepsilon_- I} + \frac{\nu_+\varepsilon_+ Z}{S+\nu_+\varepsilon_+ I}\right)\right)\left(1-\frac{X+Z}{\zeta}\right),\\
h_-(S,I) &= \alpha\left(1+\delta(\Gamma\phi_- - 1)\right),\\
h_+(S,I) &= \alpha\left(1+\delta(\Gamma\phi_+ - 1)\right).
\end{aligned} \tag{28}$$

However, when $\beta = 1$ and $\delta = 0$ (default values), i.e. when vector preference does not interact with vector population dynamics, the functions controlling population dynamics become simpler, with

$$
\begin{aligned}
g(S,I,X,Z) &= \sigma\left(X+Z\right)\left(1-\frac{X+Z}{\zeta}\right), \\
\mathrm{h}_{-}(S,I) &= \alpha, \\
\mathrm{h}_{+}(S,I) &= \alpha.
\end{aligned}
\tag{29}
$$

Model analysis and numerical approach

The invasion threshold and expressions determining the endemic equilibria can be derived analytically from Equation (26), with stability of equilibria found via the linearisation of the system in the vicinity of equilibrium. Qualitative and quantitative results are obtained analytically and numerically based on parameter ranges taken from the literature, with extensive numerical scans used to characterise the behaviour of the model, in a range of biologically relevant cases.

Online interface

An interactive online interface was created, allowing model behaviour for different parameterisations to be visualised. It is available at https://plantdiseasevectorpreference.herokuapp.com/ (Figure 2). This interface allows its user to explore the results of a single model run, or to perform a one-way parameter scan. In both cases, the user has the flexibility to choose parameterisation of the model, either using the default settings for PT or NPT, or by selecting a custom set of parameters. Whenever the user changes values of parameters, the interface triggers a callback function, which calculates the model outputs or equilibria before returning the results in the form of tables and interactive figures. The interface was built using Python Dash (Plotly, 2015), which is open source. The source code for the interface is at https://github.com/nt409/vector-interface; this Python code also acts as a reference implementation of our mathematical model.

RESULTS

Invasion threshold

We are interested in whether the virus can invade a compatible vector-plant system following first introduction and the extent to which vector preference contributes to successful invasion. The plant population density (N) in the absence of the virus is S plants per unit area as all plants are healthy and susceptible. The vector population density (κ) in the absence of the virus is X vectors per unit area as all vectors are non-viruliferous. We therefore assume that in the absence of virus the average population density of vectors per plant (κ/N) is at the carrying capacity per plant.

Suppose that when the virus is introduced there is a very small number of infected plants, $I \ll S \approx N$. From Equation (3), the probability a non-viruliferous vector lands on a susceptible plant is given by $S/(S+\nu_I)$, whereas the probability a non-viruliferous vector lands on an infected plant is $\nu_I/(S+\nu_I)$. Any infected plant would remain infectious for $1/(\mu+\rho)$ time units. The population density of non-viruliferous vectors is κ. From Equation (26), acquisition of virus from infected plants by the non-viruliferous vector population would therefore occur at rate

$$\phi_\eta X \frac{\nu_I}{S+\nu_I} \approx \frac{\eta\kappa\nu_}{\omega_\Gamma N}, \tag{30}$$

per infected plant, since $\phi_ \approx 1/(\omega_\Gamma)$ and $(\nu_I)/(S+\nu_I) \approx \nu_I/N$ when I is small.

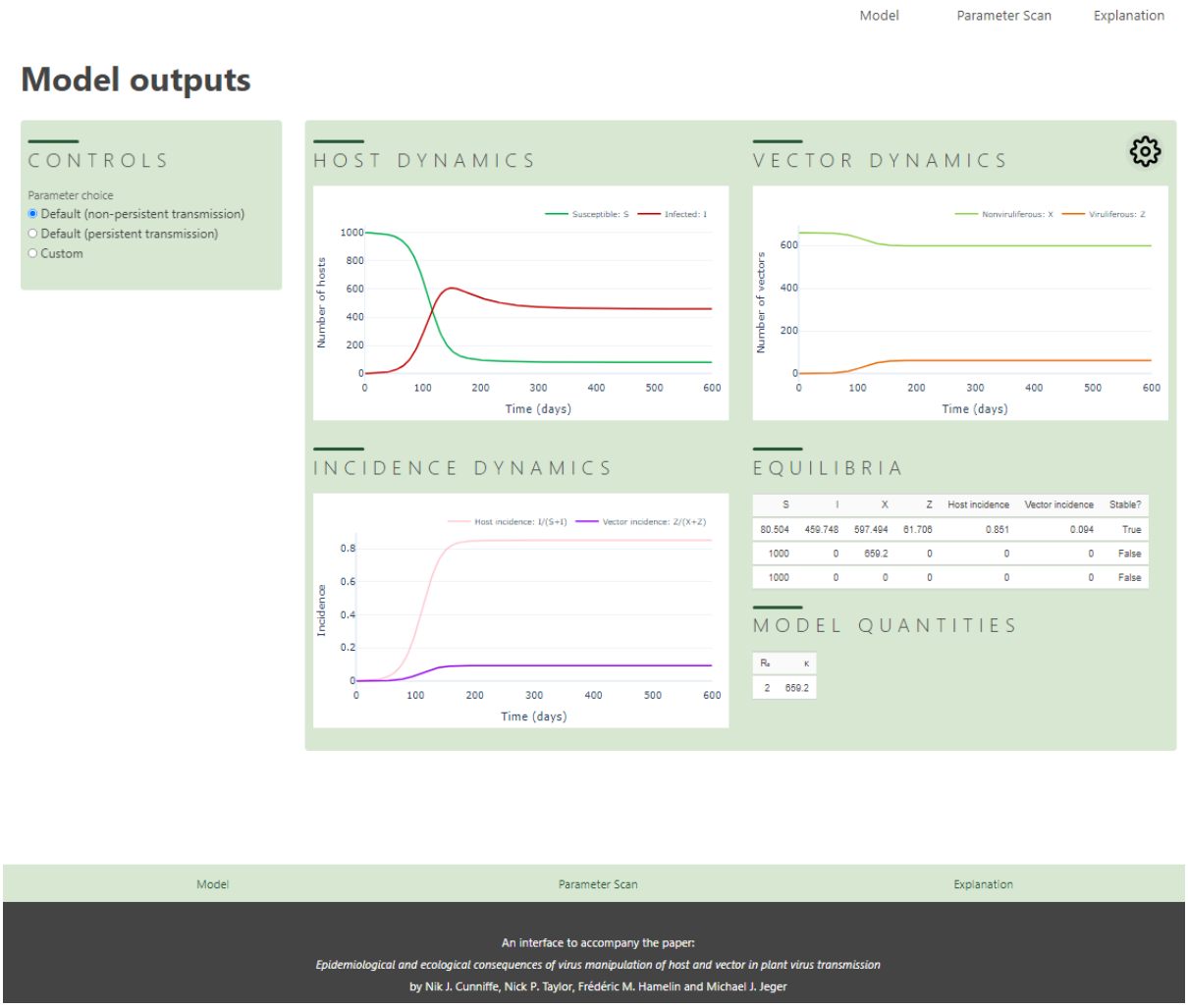


Figure 2. Online interface to the model. A screenshot showing the results of the "Model" tab of the interactive interface, available online at https://plantdiseasevectorpreference.herokuapp.com/. The user can select the model parameterisation to investigate (left hand side), and – by selecting between the "Model" and "Parameter Scan" tabs on the top-right – can choose to be presented with the dynamics of the model in time, or with a bifurcation diagram showing how the number, numeric value(s) and stability of equilibria depend upon the value of a model parameter. The cog-wheel icon at the top right of the main panel allows the user to select the number of columns used to display graphical results. The screenshot in the figure shows the results in "Two column" mode.

The average number of non-viruliferous vectors directly acquiring virus from a single infected plant over its infectious period is therefore

$$R_0^{PV} = \frac{\kappa}{N} \cdot \frac{\eta\nu_}{\omega_\Gamma(\mu+\rho)}. \tag{31}$$

Suppose for each newly viruliferous vector, at the time of introduction, $Z \ll X \approx \kappa$. From Equation (4), the probability a viruliferous vector lands on a susceptible plant is $S/(S+\nu_+ I)$, and the probability a viruliferous vector lands on an infected plant is $\nu_+ I/(S+\nu_+ I)$. As the population density of infected plants (I) is small at the time of introduction, then from Equation (22), the per capita death rate of viruliferous vectors, $h_+(S,I)$ is approximately $\alpha(1+\delta(1/\omega_+ - 1))$. These vectors additionally lose infectivity at rate τ. A viruliferous vector therefore remains inoculative for an average of

$$\frac{1}{\tau+\alpha\left(1+\delta\left(\frac{1}{\omega_+}-1\right)\right)} \tag{32}$$

time units. For each viruliferous vector in the plant population which, apart from the initially introduced infected plant, is otherwise totally susceptible, inoculation occurs at rate

$$\phi_+\gamma\frac{S}{S+\nu_+ I} \approx \frac{\gamma}{\omega_+\Gamma}, \tag{33}$$

since $\phi_+ \approx 1/(\omega_+ \Gamma)$ and $S/(S+\nu_+ I) \approx 1$ when I is small.

The average number of plants per unit area that would become directly infected for each viruliferous vector infected by the initially infected plant in an otherwise uninfected population is therefore

$$R_0^{VP} = \frac{\gamma}{\omega_+\Gamma\left(\tau+\alpha\left(1+\delta\left(\frac{1}{\omega_+}-1\right)\right)\right)}. \tag{34}$$

We note that ν_+ does not appear in this expression.
The pathogen's basic reproduction number is obtained by the product of R_0^{VP} and R_0^{PV}, leading to

$$R_0^2 = \frac{\kappa}{N}\cdot\frac{\eta\nu_-\gamma}{\omega_-\omega_+\Gamma^2(\mu+\rho)\left(\tau+\alpha\left(1+\delta\left(\frac{1}{\omega_+}-1\right)\right)\right)}. \tag{35}$$

To ease interpretation, we back substitute $\phi_- \approx 1/(\omega_- \Gamma)$ and $\phi_+ \approx 1/(\omega_+ \Gamma)$, and rearrange the order of terms, leading to

$$R_0^2 = \frac{\kappa}{N}\cdot\frac{1}{\mu+\rho}\cdot\phi_-\eta\nu_-\cdot\frac{1}{\left(\tau+\alpha\left(1+\delta\left(\frac{1}{\omega_+}-1\right)\right)\right)}\cdot\phi_+\gamma. \tag{36}$$

Successive terms in this expression allow an intuitive interpretation as follows:

> *Average number of vectors per plant in absence of virus × average infectious period (time units) of a single infected plant × average number of plants visited by a vector (per unit time) × probability of virus acquisition by a single vector during a single visit × average period (time units) a vector remains viruliferous × average number of plants visited per vector (per unit time) × probability of inoculation by a single viruliferous vector.*

The heuristic derivation of the basic reproduction number outlined above matches that can be obtained more formally from the spectral radius of the Next Generation Matrix (Supplementary Appendix S2). The squared formulation reflects the fact that there are two cycles involved in transmission, from plant to vector and from vector to plant (Heesterbeek & Roberts, 2007). Clearly, $R_0 \geq 0$ for non-negative values of the parameters (including κ, which can be positive or negative depending on the values of the other parameters; cf. Equation (24)). The virus invades the system whenever $R_0 > 1$, but note that the parameters γ (Equation (7)), η (Equation (8)) and κ (Equation (24)), depend on other parameters and/or the transmission type (NPT or PT) of the virus in question. As noted above ν_- affects the numerical value of R_0, but ν_+ does not. Any effect of ν_+ depends on viruliferous vectors being attracted to infected plants, and during initial invasion this "second order" effect is small. However, the probabilities that a non-viruliferous (ω_-) or a viruliferous (ω_+) vector settles and feeds on a plant are both retained in the expression for R_0.

A basic reproduction number was derived from an earlier version of the current model in which a conditional vector preference parameter ν_Z, consistent with ν_+ here, appeared in the term $N + \nu_Z$ (Jeger 2020, Appendix A). However, that term is not valid dimensionally since N is a population density (plants per unit area). When the appropriate corrections are made to the past modelling work, the basic reproduction number is essentially the same as the derived R_0^2 here.

Default parameterisation of the model

The default parameterisation (Table 1) assumes that infected plants are no better or worse than susceptible plants for vector reproduction ($\beta = 1$) and that there is no additional mortality associated with increased flights between plants ($\delta = 0$). The default density of vectors at which the birth rate becomes zero, ζ, was selected – after fixing the other parameter values as described below – to ensure $R_0 = 2$ for both NPT and PT, giving Γ = 2 days and $\zeta = 1977.6$ (NPT) and $\zeta = 163.2$ (PT). We note that with these changes $\kappa = 659.2$ (NPT) and $\kappa = 54.4$ (PT). The plant population density in the absence of disease is $N = 1000$ per unit area. The vector population size is therefore between about 0.2 and 2 vectors/plant.

Other parameter values are informed, where possible, by previous modelling studies (e.g. Jeger et al. 1998, 2004, 2011; Madden et al. 2000; Holt et al. 1997). Parameters set in this way include the rate at which plants are harvested ($\rho = 0.01\ d^{-1}$), the additional rate of disease-induced death and/or roguing ($\mu = 0.01\ d^{-1}$), the underlying per capita vector death rate ($\alpha = 0.12\ d^{-1}$), the maximum per capita vector birth rate ($\sigma = 0.18\ d^{-1}$), and, assuming the rate at which vectors lose infectivity depends strongly on the transmission mode, $1/\tau = 0.25\ d$ (NPT) and $1/\tau = 20\ d$ (PT). There is no direct analogue of Γ, the length of a feed, in previous modelling work. We therefore arbitrarily assume a value of 2 d, irrespective of the transmission type, to set the number of plants visited per unit time in the model. Assuming, at least by default, that there is no vector preference for landing based on infection status, we take the values $\nu_- = \nu_+ = 1$, and for feeding we take $\omega_- = \omega_+ = 0.5$, with $\epsilon_- = \epsilon_+ = 1$. By default, therefore, individual vectors move between plants once every two days, and on average probe and move away from one plant before settling on the second plant they visit.

Time plots for baseline parameterisation

Densities per unit area of healthy susceptible plants, infected plants, non-viruliferous vectors, and viruliferous vectors, following the introduction of one infected plant per unit area into an otherwise

healthy plant population and one viruliferous vector per unit area into an otherwise non-viruliferous vector population, are shown in Figure 3. The models were run for 600 days to ensure equilibrium, but this was approached in less than 200 days for NPT and a little more than 200 days for PT. The density of infected hosts and disease incidence in the plant population was always higher with NPT, with a much lower density of viruliferous vectors than with PT. The proportion of viruliferous vectors in the vector populations was much higher with PT.

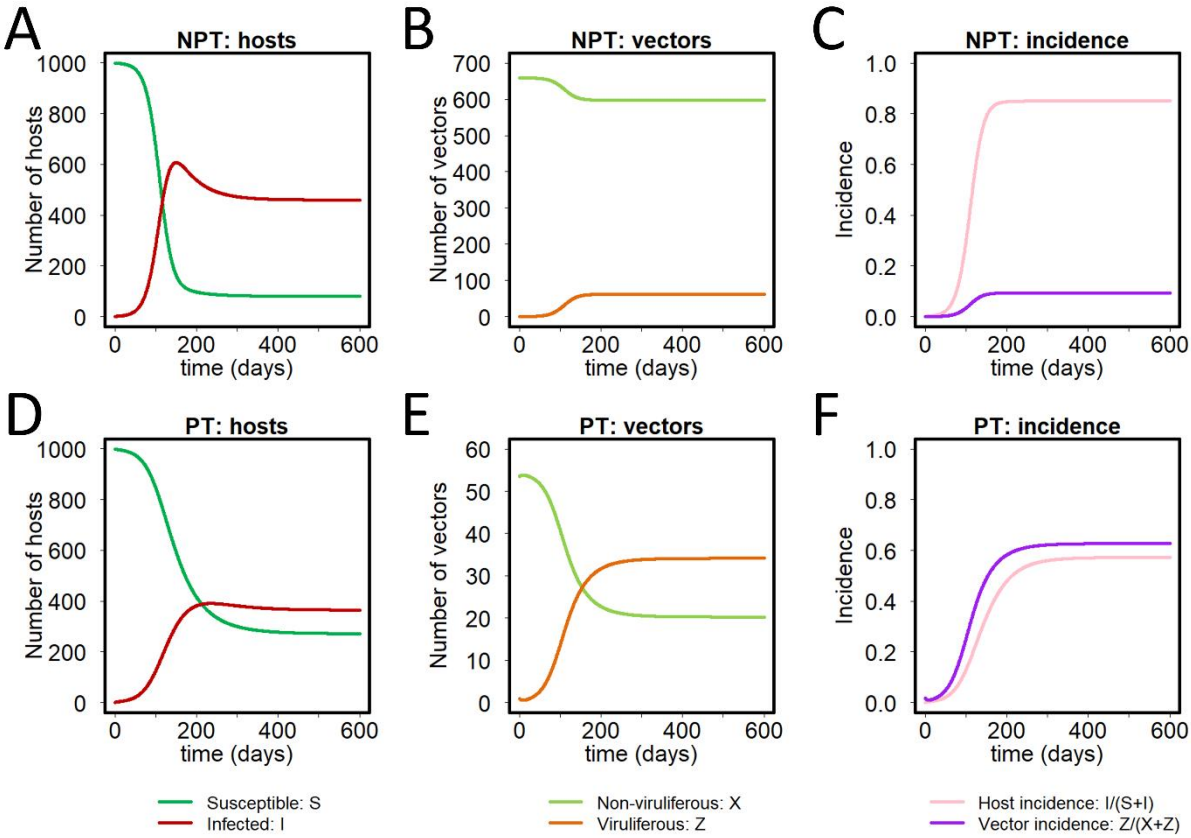


Figure 3. Baseline parameterisation of the model. Models were parameterised using available literature, select parameters to ensure a default value of $R_0 = 2$ for both the NPT and PT models. **(A)** Numbers of susceptible and infected hosts; **(B)** numbers of non-viruliferous and viruliferous vectors; and **(C)** incidences (i.e. proportions infected) for both hosts and vectors, as functions of time, using the default parameterisation of the NPT model. **(D)-(F)** The same for the PT model. All disease progress curves are shown for the case in which a single infected host and viruliferous vector are simultaneously introduced at time $t = 0$, and the models were run for 600 days (a value selected to be sufficiently large for equilibrium to be attained for both parameterisations).

Sensitivity to life history and epidemiological parameters

The sensitivities of R_0, the equilibrium density of infected hosts I_∞, and the equilibrium disease incidence $I_\infty/(S_\infty + I_\infty)$ to changes in model parameters are shown in Figure 4. In these plots, the vector preference parameters for landing and feeding are set to 1; hence sensitivity to the derived inoculation γ and acquisition η probabilities are not shown. For R_0, the responses to vector death rate (α) show the greatest divergence when NPT (Figure 4A) and PT (Figure 4D) are compared. The R_0 value is higher for PT than NPT at low death rates. The sensitivity to other parameters is similar for both transmission types, especially when parameters are varied above the default values. For I_∞, the

greatest divergence between NPT and PT occurs for the host roguing rate (μ), with higher values at lower rates for NPT. There are also differences in responses to the initial plant population size (N), with an almost linear increase in I_∞ as N increases above the default for NPT, whereas it levels off for PT. The equilibrium disease incidences for NPT and PT largely reflect the fact that, at least for our default model parameterisation, overall disease incidence was higher for NPT (*cf.* Figure 3).

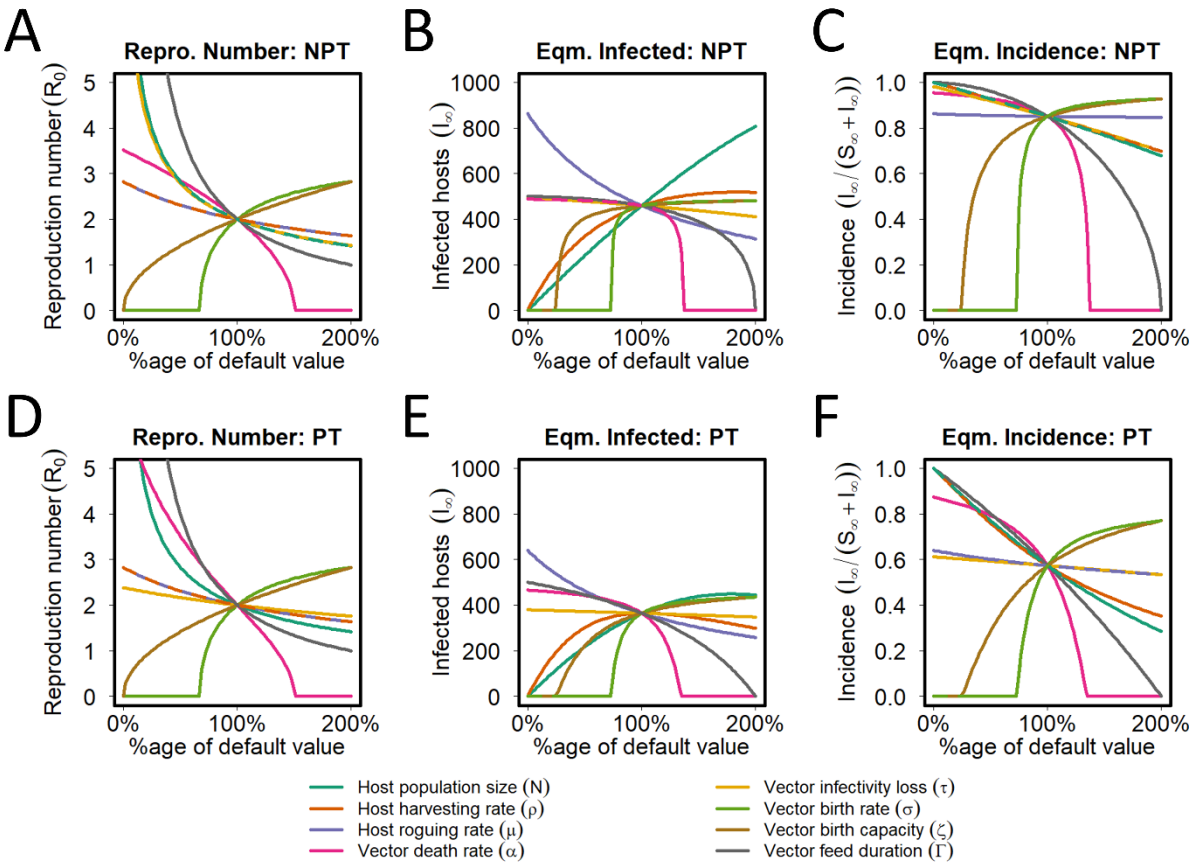


Figure 4. Sensitivity of underlying model to life history and epidemiological parameters. Responses of: **(A)** the basic reproduction number, R_0; **(B)** the equilibrium number of infected hosts, I_∞; and **(C)** the equilibrium incidence, $I_\infty/(S_\infty + I_\infty)$, as the life history and epidemiological parameters in the model were varied between 0% and 200% of their default values for the NPT model. **(D)-(F)** The same for the PT model.

Model equilibria

There are two disease-free equilibria, $(S, I, X, Z) = (N, 0,0,0)$, and $(S, I, X, Z) = (N, 0, \kappa, 0)$. If the parameters are such that $\kappa < 0$ then the equilibrium in which $X = 0$ is always stable. If $\kappa > 0$, and so if the vector population can persist in the absence of disease, then the equilibrium in which $X = 0$ is always unstable. The stability of the equilibrium in which $X = \kappa > 0$ is controlled by the basic reproduction number: if $R_0 \leq 1$ then this disease-free equilibrium is stable, otherwise it is unstable (Supplementary Appendix S3).

Explicit expressions determining the endemic equilibria in which $I \neq 0$ can be determined by routine – albeit long-winded – calculation (Supplementary Appendix S3). In general, equilibrium values of I are given by the solutions of a certain quartic equation, with the values of the other state variables fully specified in terms of these values of I. There are therefore up to four non-zero equilibria,

although for any given parameter set, some or even all of these might not take biologically meaningful values (i.e. for some equilibria generated by the quartic equation, at least one state variable is negative at equilibrium).

In the case when $\beta = 1$ and $\delta = 0$, i.e. when vector preference does not interact with vector population dynamics, further simplification is possible (Supplementary Appendix S3). In this case it can be shown that the equilibrium values of I are given by solutions to a certain quadratic equation. In this case there is therefore a maximum of two non-zero equilibria, although again it is possible that one or both equilibria do not take biologically meaningful values.

For both of our default parameterisations of the model, there is a total of three biologically meaningful equilibria (the two disease-absent equilibria, as well as a single equilibrium with $I \neq 0$; the other disease-present equilibrium from the relevant quadratic has at least one negative state variable). Since $\kappa > 0$, only the disease-present equilibrium is locally stable, and so the long-term behaviour of the model is unambiguous and does not depend on initial conditions.

However, it is clearly of interest to understand whether multiple disease-present equilibria can be attained, and if so, whether they are stable or not. To assess this, we did an extensive randomisation scan (Supplementary Appendix S4), finding the number and stability of equilibria as parameters were randomly varied. Bistability between one of the vector-present/disease-absent or the vector-absent/disease-absent equilibrium and a single disease-present equilibrium was relatively commonplace (around 25-30% of cases). There was also a very small proportion (<<1% of cases) with bistability between two disease-present equilibria (with or without an additional stable disease-absent equilibrium); examples are given in Supplementary Appendix S4. In the remainder of the results as presented in the main paper, parameters were such that only the simpler bistability involving a single disease-present equilibrium was seen.

Bistability

A particular example showing the effect of bistability between the disease-present and vector-present/disease-absent equilibrium was investigated in some detail (Figure 5), focusing upon the effects of μ, the host roguing/additional disease-induced mortality parameter. Note that for both NPT and PT, the basic reproduction number is a decreasing function of μ, which means $R_0 < 1$ for large values of μ (Figures 5A and 5D). For both models, at the default parameterisation ($\mu = 0.01\ d^{-1}$; shown by a dot) there is a single biologically meaningful disease-present equilibrium, as well as the disease-free equilibrium (at which the vector is present since $\kappa > 0$). The other equilibrium for which $I \neq 0$ has a negative value of I. As μ increases, for both NPT and PT, there is a transcritical bifurcation at $\mu \approx 0.07\ d^{-1}$ (i.e. at the threshold value $R_0 = 1$). This leads to an exchange of stability between the negative equilibrium – which becomes positive as μ is further increased – and the disease-absent equilibrium, meaning that for larger values of μ both the other disease-present equilibrium and the disease-absent equilibria are stable. The model then exhibits bistability; the equilibrium that is attained in the long term depends on the initial condition, with different basins of attraction (Figures 5B and 5E) leading to different results in the long term (Figures 5C and 5F). The only qualitative difference between the results for the two parameterisations is that, for PT, only the disease-free equilibrium exists for very large roguing rates. In particular, there is a saddle node bifurcation at $\mu \approx 0.1868\ d^{-1}$, in which the stable and unstable disease-present equilibria collide and annihilate each other, meaning that for sufficiently large values of μ only the disease-free equilibrium exists. This

actually also occurs for NPT, but at a much larger value of μ, well outside the range of values shown in Figure 5.

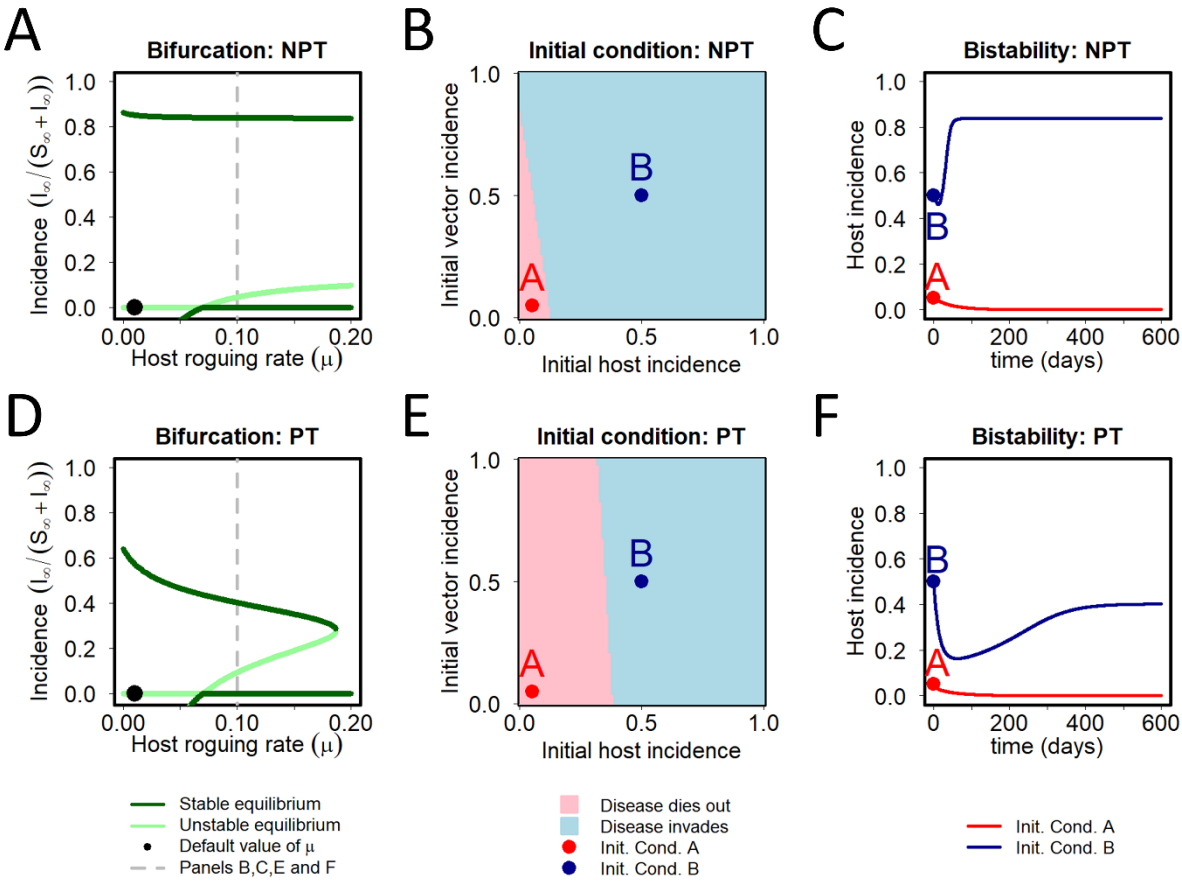


Figure 5. Bistability. For certain sets of parameters, the model has multiple stable biologically plausible equilibria; we illustrate this here by considering how the behaviour of the model depends upon the rate at which hosts are rogued. **(A)** A bifurcation diagram for the NPT model, using the host roguing rate (μ) as the bifurcation parameter. When $\mu = 0.07$ d^{-1} (i.e. seven times larger than the default value of 0.01 d^{-1}) the basic reproduction number $R_0 = 1$. Because R_0 is a decreasing function of the roguing rate (Equation (36)), it follows that $R_0 < 1$ for larger values of μ, and so the disease-free equilibrium becomes locally stable. However, there a locally stable equilibrium corresponding to disease being present in the system is still present (as well as a further unstable equilibrium). **(B)** When $\mu = 0.1$ d^{-1}, which of the NPT model's two locally stable equilibria – the disease-free equilibrium or the disease-present equilibrium at which the incidence is around 0.8 – is eventually attained depends on the initial conditions. Qualitatively similar results occur for all values of $\mu > 0.07$ d^{-1}. **(C)** Disease severity in the NPT model as a function of time starting at initial conditions marked A $(I_0/(S_0+I_0), Z_0/(X_0+Z_0)) = (0.05,0.05)$ and B $(I_0/(S_0+I_0), Z_0/(X_0+Z_0)) = (0.5,0.5)$ in (B), with all other parameters equal to the defaults (apart from $\mu = 0.1$ d^{-1}). **(D)-(F)** The same for the PT model. The only qualitative difference between the results of the two versions of the model is that, in the PT model, only the disease-free equilibrium exists for very large roguing rates (there is in fact a saddle node bifurcation at $\mu \sim 0.1868$ d^{-1} in which the stable and unstable disease-present equilibria collide and annihilate each other).

Vector preference

The expression for R_0 (Equation (36)) indicates invasion for both NPT and PT viruses is promoted by increasing $\nu_- > 1$, since it increases the probability of non-viruliferous vectors landing on infected plants, and so the number of plant-vector contacts that could lead to virus acquisition. Increasing ν_+, the bias of viruliferous vectors for infected plants, would effectively be "wasting" contacts as when

the virus is rare, contacts between infected plants and viruliferous vectors are very unlikely. Therefore, the value of ν_+ does not affect invasion. Increasing ϵ_+, thereby making viruliferous vectors more likely to feed on infected plants, also has no effect on invasion, for the same reason.

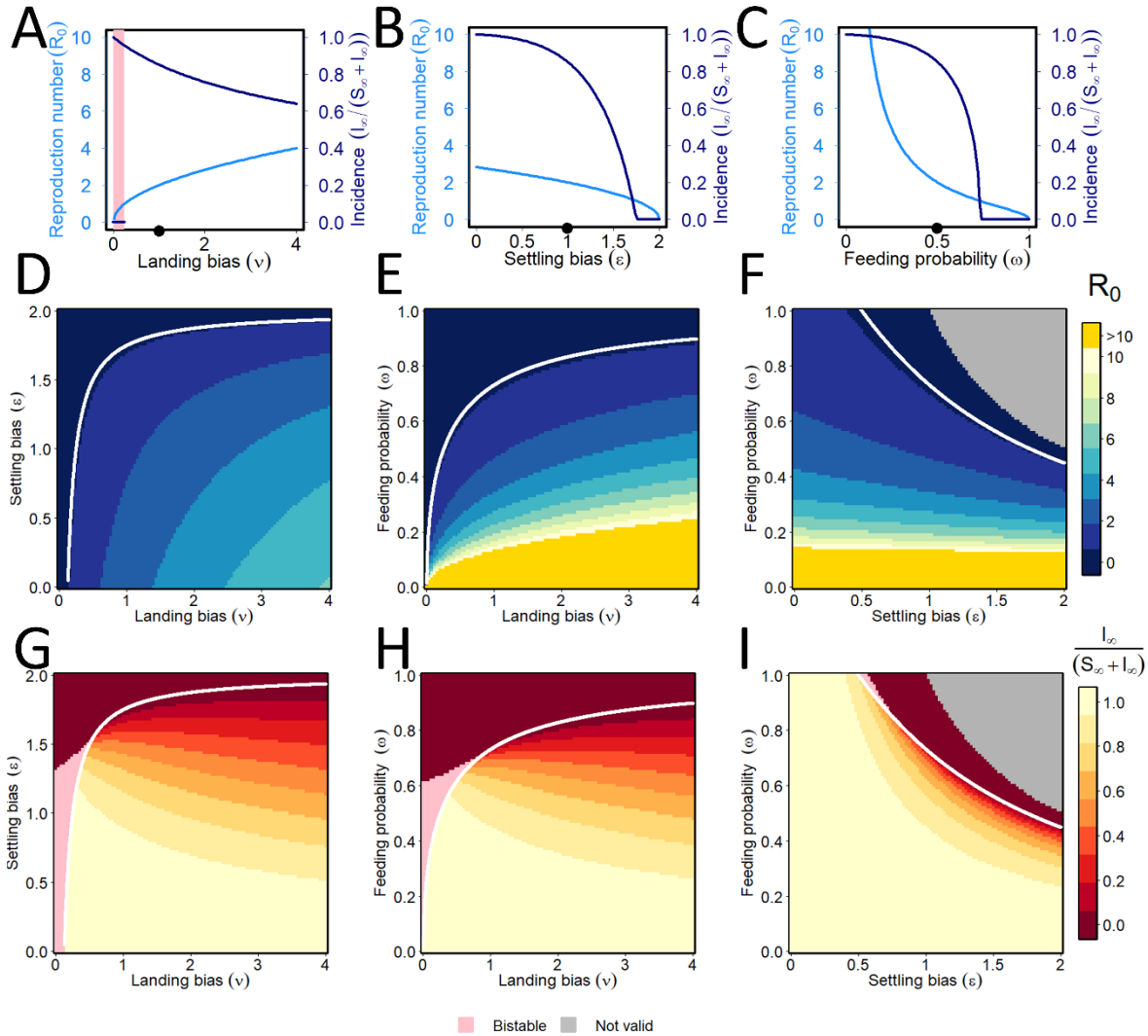


Figure 6. Non-persistent transmission: effect of vector preference. (A)-(C) Responses of the basic reproduction number (light blue) and equilibrium disease incidence (dark blue) to changes in the vector preference parameters: (A) landing bias (ν); (B) settling bias (ϵ) (C) feeding probability (ω). For NPT there can be no conditional preference, so we introduce $\nu = \nu_- = \nu_+, \epsilon = \epsilon_- = \epsilon_+$ and $\omega = \omega_- = \omega_+$. Default values $\nu = \epsilon = 1, \omega = 0.5$ are marked with black dots on each plot's x-axis. As the parameter ν is altered in a one-way sensitivity analysis around the default values of other parameters, the model exhibits bistability whenever $R_0 < 1$. Therefore, which of the stable disease-free equilibrium and the stable disease-present equilibrium is attained in the long-term depends on the initial conditions. The relevant range of the parameter ν is marked by pink shading. **(D)-(I)** Responses of the basic reproduction number (top row: (D)-(F)) and the equilibrium incidence (bottom row: (G)-(I)) as pairs of vector preference parameters are simultaneously altered. Pairs of parameters for which $R_0 = 1$ are marked with a white curve. The grey regions in panels (F) and (I) correspond to invalid combinations of parameters for which $\omega\epsilon > 1$. Pairs of parameters for which the model exhibits bistability are again marked in pink. In panels (E) and (F) values of $R_0 > 10$ are shown in orange so as not to skew the colour-ramp and so obscure the main features of the results.

However, the vector preference parameters, ω_-, ω_+ and ϵ_- have opposite effects on invasion depending on the transmission type, NPT or PT. Since non-viruliferous vectors must feed on infected plants to acquire PT viruses, increasing ϵ_- promotes invasion in this case. However, invasion of NPT viruses would be reduced, since feeding stops the virus being acquired by the vector as virus is lost from the stylet (Mauck et al. 2012). Increasing the underlying probability of feeding by either uninfected (ω_-) or infected (ω_+) vectors is also deleterious for NPT, essentially because feeding "wastes" time. However, for PT viruses, feeding is required for both acquisition and inoculation.

In our numerical work, we first considered vector preference for NPT viruses (Figure 6), for which there is guaranteed to be no conditional preference based on the vector's infective status, i.e. in the model $\nu = \nu_- = \nu_+$, $\epsilon = \epsilon_- = \epsilon_+$ and $\omega = \omega_- = \omega_+$. For NPT, this is assumed to always be the case due to the transient nature of virus acquisition, retention, and inoculation. For increasing values of the landing bias ν, the basic reproduction number, R_0, increases while final disease incidence, $I_\infty/(S_\infty + I_\infty)$, decreases (Figure 6A). With increasing values of the settling bias ϵ, R_0 decreases and final disease incidence drops considerably beyond the default value (Figure 6B), an effect even more pronounced with feeding probability, ω (Figure 6C). As the parameter ν decreases from the default value, the model exhibits bistability for $R_0 < 1$ (shown in pink). As the settling bias and feeding probability for NPT increase for any landing bias, both R_0 (Figures 6D-6F) and final disease incidence (Figures 6G-6I) decrease.

Similar plots are shown for PT in Figure 7, in the non-conditional case in which vector preference does not depend on vector infection status, i.e. when again $\nu = \nu_- = \nu_+, \epsilon = \epsilon_- = \epsilon_+$ and $\omega = \omega_- = \omega_+$. As for NPT, R_0 increases and final disease incidence decreases with increased landing bias, ν (Figure 7A). However, unlike NPT, for PT R_0 increases with settling bias, ϵ, and the drop in final disease incidence is less pronounced (Figure 7B). The model exhibits bistability in one-way sensitivity analyses to both landing bias and settling bias. However, and in contrast to NPT, for which the model was bistable for all values of $R_0 < 1$, there is instead a saddle node bifurcation at the value $0 < R_0^* < 1$ at which the stable and the unstable (not shown on the graphs) endemic equilibria collide. This means there is only the single disease-free equilibrium whenever $R_0 < R_0^*$. For PT, both R_0 and final incidence are irresponsive to the feeding probability, ω, since feeding is required for acquisition/inoculation, and any additional probability of feeding is offset by fewer plants visited per unit time (note how the factors of ω_- and ω_+ both cancel for PT in Equation (27)). The patterns for Figures 7A and 7B are recapitulated in the two-way sensitivity analyses shown in Figures 7D and 7E.

Conditional vector preference

Sensitivity scans showing the effect(s) of conditional vector preference – which can only be expressed by PT viruses – are in Figure 8. There is a rapid rise in R_0 as both the biases that can potentially be shown by non-viruliferous vectors for infected plants increase; however, there is no change as the corresponding biases shown by viruliferous vectors increase. This is as would be expected from the derivation of R_0. The final disease incidence increases but soon levels off with increases in the biases of non-viruliferous vectors for infected plants. With an increasing preference of viruliferous vectors for infected plants, the final disease incidence decreases from a saturation level of 1.0 when this preference is close to 0, i.e. the bias of viruliferous vectors is for healthy plants is increasingly important to reach high disease incidence. Again, two-way sensitivity scans (Figures 8C-8I) reinforce patterns from the one-way analyses, while allowing the effects of interactions between pairs of parameters to be seen. If non-viruliferous vectors are attracted and land or settle upon infected plants

(ν_- and ϵ_- are both high), but viruliferous vectors are repulsed from infected plants (ν_+ and ϵ_+ both low), this is, unsurprisingly, optimal from the point of view of the virus.

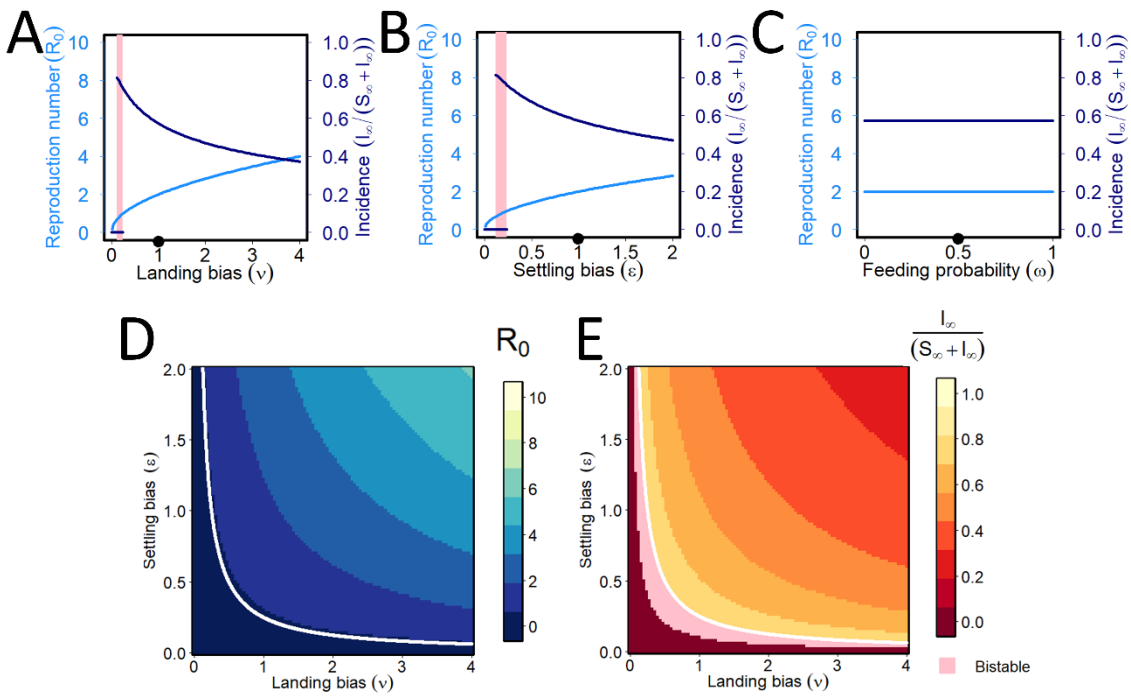


Figure 7. Persistent transmission: effect of vector preference. (A)-(C) Responses of the basic reproduction number (light blue) and equilibrium incidence (dark blue) to changes in the vector preference parameters: (A) landing bias (ν); (B) settling bias (ϵ) (C) feeding probability (ω). For PT, conditional vector preference is biologically meaningful (shown in Figure 8), but in its absence, the model exhibits bistability in one-way sensitivity analyses to both the landing bias (ν) and the settling bias (ϵ) parameters (panels A and B). There is no sensitivity in response to feeding probability (panel C). In contrast to NPT, the model is not bistable whenever $R_0 < 1$ for a one-way sensitivity analysis altering ν (or ϵ). Instead, there are saddle node bifurcations in which the stable and the unstable disease-present equilibria collide and annihilate each other, at around $\nu = \epsilon = 0.1016$ (note the unstable equilibria are not shown on this Figure, so the bifurcation is represented by the disease-present equilibrium abruptly "stopping" as the parameter values get smaller). **(D)-(E)** Responses of the basic reproduction number (D) and the equilibrium incidence (E) as landing bias and settling bias are simultaneously altered. Pairs of parameters for which $R_0 = 1$ are again marked with a white curve. Pairs of parameters for which the model exhibits bistability are again marked in pink.

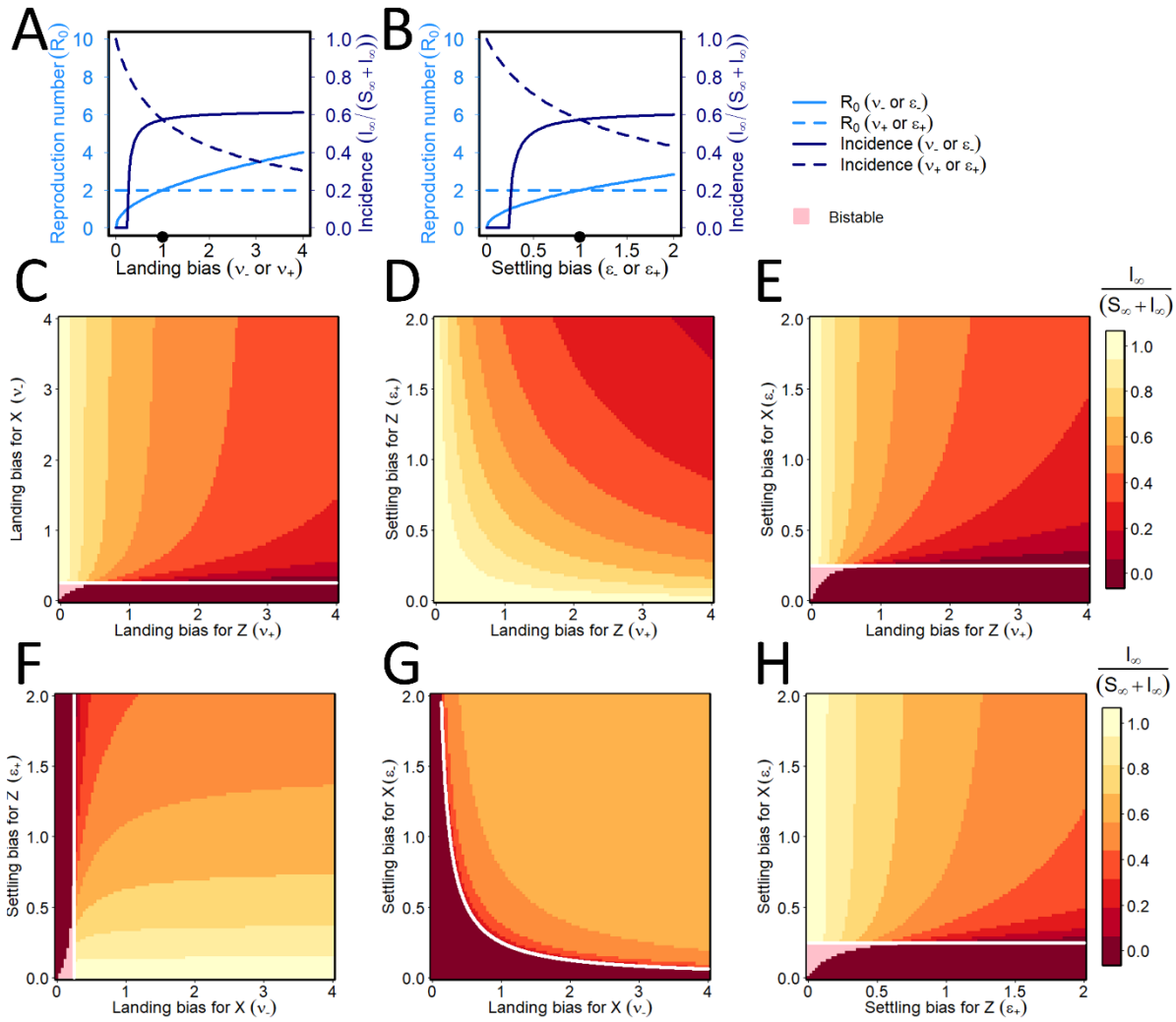


Figure 8. Persistent transmission: effect of conditional vector preference. (A)-(B) Responses of the basic reproduction number (light blue) and final disease incidence (dark blue) to changes on the conditional landing (A; ν_- and ν_+) and settling (B; ϵ_- and ϵ_+) biases of non-viruliferous (solid lines) and viruliferous (dotted lines) vectors. As suggested by Figure 7, both the reproduction number and the disease incidence are irresponsive to changes in both ω_- and ω_+ for PT, and so these sensitivity analyses are not shown. **(C)-(H)** Responses of the final disease incidence as pairs of conditional vector preference parameters are simultaneously altered. Again, pairs of parameters for which $R_0 = 1$ are marked with a white curve, with bistability shown by pink shading.

Interactions between vector population dynamics and disease

Our model allows the maximum vector reproduction rate to depend on the disease status of plants upon which they predominantly feed, via an assumed relative rate of reproduction on infected plants (β). It also allows us to assume that that more vector flights between plants leads to an increased loss rate of vectors, $\delta > 0$. The paired effects of varying these two parameters simultaneously on final disease incidence is shown in Figure 9A (NPT) and Figure 9B (PT). The overall pattern is similar for both transmission modes. Bistability between the disease-free equilibrium with the vector fixed at carrying capacity and another equilibrium in which the disease has invaded, is again possible. However, for larger values of δ vectors would die out in the absence of disease (light blue), although sufficiently large values of β can then "rescue" both vector and virus populations, allowing disease to persist.

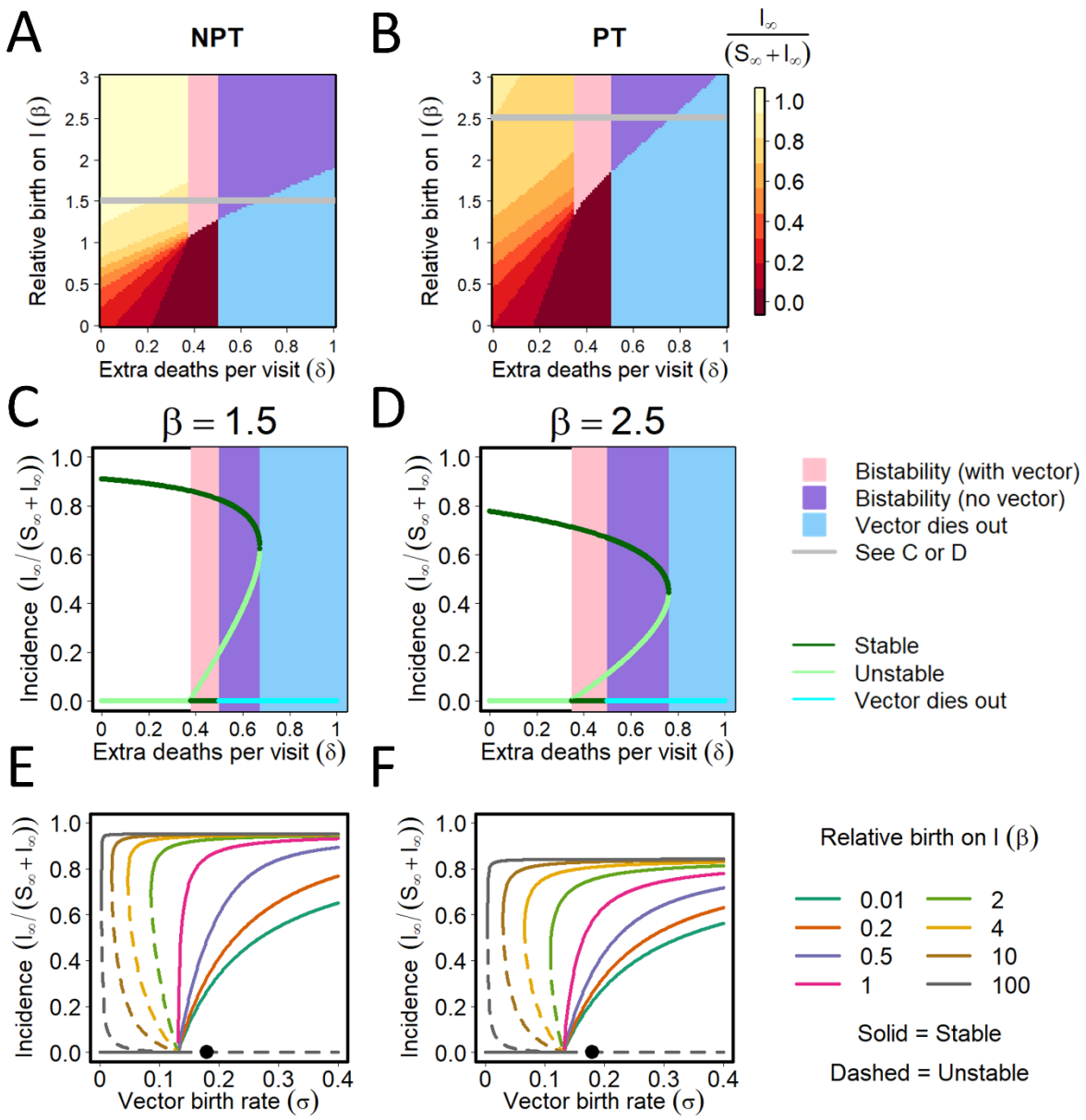


Figure 9. Interactions between vector population dynamics and disease. (A)-(B): Final disease incidence in relation to the relative birth rate of vectors on infected plants (β) and the increase in the vector loss rate with the number of plants visited between each successive feed (δ) for NTP (A) and PT (B). The region of bistability between the disease-free equilibrium with the vector fixed at its carrying capacity and an endemic equilibrium in which the disease has invaded is shaded pink. In the light blue region, for larger values of δ the vector would die out in the absence of disease, although sufficiently large values of β can "rescue" both vector and virus populations. **(C)-(D):** Final disease incidence in relation to the increase in the vector death rate with the number of plants visited between each successive feed, δ, for a fixed value of the relative birth rate of vectors on infected plants, β. Responses are shown for different values of β for NPT ($\beta = 1.5$) vs. PT ($\beta = 2.5$) to show the full range of behaviour; note these values of β correspond to the grey horizontal lines in (A) and (B). **(E)-(F):** Final disease incidence in relation to the underlying birth rate of vectors for NPT (E) and PT (F), for different values of the relative birth rate of vectors on infected plants, β.

The response of the final disease incidence to the increase in the vector loss rate with the number of plants visited between each successive feed, δ, is shown for fixed values of β: 1.5 for NPT (Figure 9C) and 2.5 for PT (Figure 9D). These values of β were chosen to allow the full range of behaviour to be

shown for each mode of transmission and correspond to the horizontal "slices" through Figures 9A and 9B marked by grey lines. The responses of the final disease incidence to the underlying birth rate (σ) of vectors for different values of the relative birth rate on infected plants (β) is shown for NPT (Figure 9E) and PT (Figure 9F). In both cases, final disease increases with increasing relative birth rates on infected plants. However, the final disease incidence becomes unstable when a low underlying birth rate is combined with high relative birth rate on infected plants. Bistability is possible for values of $\beta > 1$, where sufficiently large initial disease levels can allow the disease to persist even when the disease-free equilibrium is stable (i.e. when $R_0 < 1$).

Interactions between conditional vector preference and population dynamics

For PT viruses, conditional vector preference and population dynamics can interact (Figure 10). For $\nu_- = 1$, the final disease incidence increases as β increases (with diminishing effect for $\beta > 1$); as ν_- increases above 1 there is little change in the final disease incidence for $\beta > 1$ with some decreases for $\beta < 1$ (Figure 10A). For $\nu_+ = 1$, the final disease incidence increases as β increases (again with diminishing effect for $\beta > 1$); as ν_+ increases above 1 there is a reduction in final disease incidence for $\beta < 1$. There is an unstable equilibrium as ν_- approaches 0 for $\beta < 1$ (Figure 10B). The effects of the settling bias (i.e. ϵ_- or ϵ_+) on final disease incidence for different values of β are qualitatively similar and are not shown.

The effect on the final disease incidence of the landing biases is shown for different values of the loss rate of vectors for each additional flight between different plants (δ) for non-viruliferous vectors (ν_-; Figure 10C) and viruliferous vectors (ν_+; Figure 10D), as well as the setting bias for non-viruliferous vectors (ϵ_-; Figure 10E) and viruliferous vectors (ϵ_+; Figure 10F). Landing and settling biases for non-viruliferous vectors (Figures 10C and 10E) give quantitatively similar final disease incidences for different increased loss rates of vectors (δ), but differ for viruliferous vectors (Figures 10D and 10F). As ν_+ increases there is a decreasing trend in final disease incidence for all values of δ; however, as ϵ_+ increases there is no trend above $\delta \approx 0.2$ with an indication of a nonmonotonic trend for lower values of δ.

Interactions between vector preference and population dynamics

The effect on the final disease incidence of the settling bias ($\epsilon_- = \epsilon_+$; A) and landing bias ($\nu_- = \nu_+$; B) in relation to relative vector birth rates on infected plants (β) and for NPT is shown in Figure 11. As there can be no conditional vector preference for NPT, the bias parameters do not depend on the infection status of the vector. At low levels of landing bias close to zero, the final disease incidence is unstable but stable equilibria then decline from a maximum value for $\nu < 1$ as ν increases in a similar way to that seen for conditional preference in PT (ν_+, Figure 10B). The effect of β in reducing the final disease incidence as landing bias increases is most apparent for β < 1. The effect of the settling bias ($\epsilon_- = \epsilon_+$, Figure 11A) shows a continuous decrease in final disease incidence as settling bias increases. Similar plots are shown for settling bias (Figure 11C) and landing bias (Figure 11D) in relation to different values of the additional rate of death for vectors of each additional flight between different plants (δ). At sufficiently low values of the settling bias $\epsilon_- = \epsilon_+$ and with $\delta > 0$ the final disease incidence loses stability via a Hopf bifurcation, leading to a sustained periodic solution oscillating around the steady state (Figures 11E and 11F).

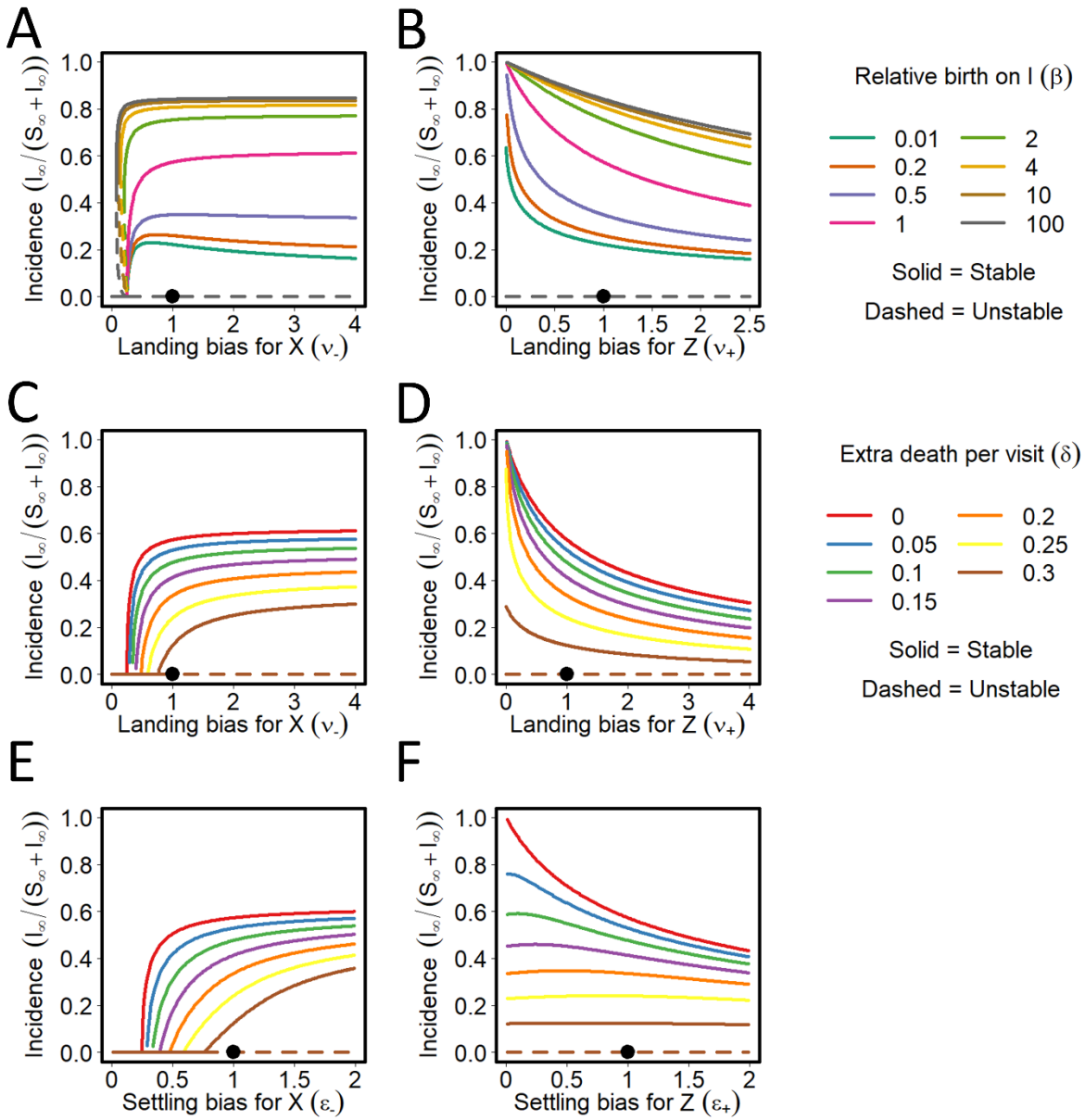


Figure 10. Persistent transmission: interactions between conditional vector preference and vector population dynamics. (A)-(B) Final disease incidence in relation to the landing bias for non-viruliferous vectors (ν_+; A) and viruliferous vectors (ν_+; B) for different relative vector birth rates on infected plants (β). Solid lines indicate a stable equilibrium incidence; dotted lines indicate values for which the equilibrium is unstable; the solid black dot marks the default parameterisation in which there is no bias. **(C)-(F)** Final disease incidence of the landing biases for non-viruliferous vectors (ν_+; C) and viruliferous vectors (ν_+; D) and the setting bias for non-viruliferous vectors (ϵ_-; E) and viruliferous vectors (ϵ_+; F), for different values of the loss rate for vectors of each additional flight between different plants (δ).

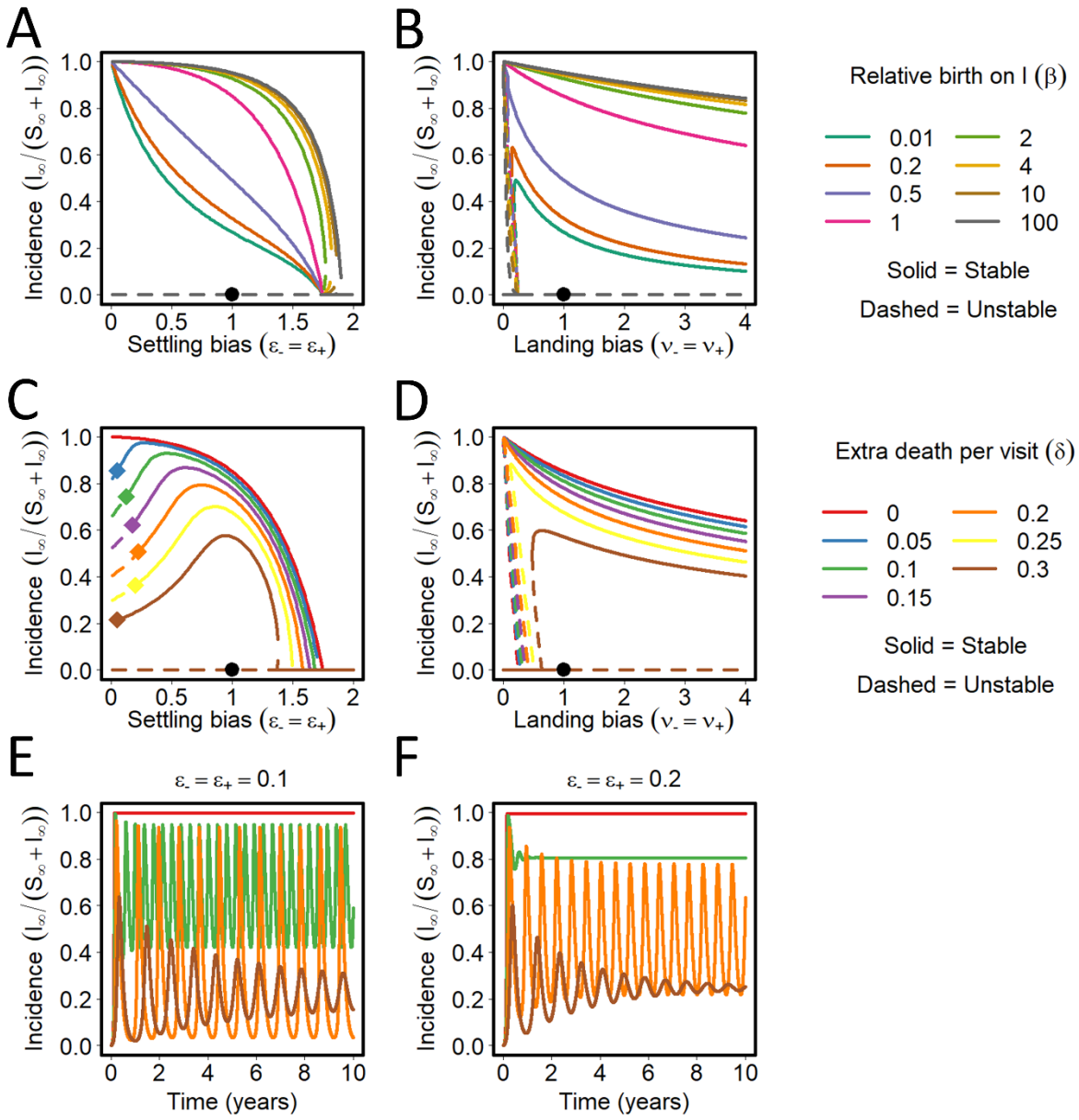


Figure 11. Non-persistent transmission: interactions between vector preference and vector population dynamics (A)-(B) Effect on the equilibrium incidence of the settling bias ($\epsilon_- = \epsilon_+$; A) and landing bias ($\nu_- = \nu_+$; B) for different relative vector birth rates on infected plants (β) for NPT. Solid lines again indicate a stable equilibrium incidence; dotted lines indicate values for which the equilibrium is unstable; the solid black dot marks the default parameterisation in which there is no bias. **(C)-(D)** Effect on the equilibrium incidence of the settling bias ($\epsilon_- = \epsilon_+$; C) and landing bias ($\nu_- = \nu_+$; D) for different values of the additional rate of death for vectors of each additional flight between different plants (δ). At sufficiently low values of the settling bias $\epsilon_- = \epsilon_+$ the disease-present equilibrium loses stability via a Hopf bifurcation, leading to a periodic solution oscillating around the steady state. The bifurcation point is marked with a diamond ($\epsilon_- = \epsilon_+ = 0.058$ when $\delta = 0.05$; $\epsilon_- = \epsilon_+ = 0.134$ when $\delta = 0.10$; $\epsilon_- = \epsilon_+ = 0.197$ when $\delta = 0.15$; $\epsilon_- = \epsilon_+ = 0.236$ when $\delta = 0.20$; $\epsilon_- = \epsilon_+ = 0.222$ when $\delta = 0.25$ and $\epsilon_- = \epsilon_+ = 0.062$ when $\delta = 0.30$). **(E)-(F)** Incidence over time for $\epsilon_- = \epsilon_+ = 0.1$ (E) and $\epsilon_- = \epsilon_+ = 0.2$ (F), showing responses for $\delta = 0, 0.1, 0.2$ and 0.3. In cases corresponding to dotted responses in (C) and (D) sustained periodic oscillations are obtained.

DISCUSSION

Epidemiological analysis

The model introduced is quite complex, involving parameters describing epidemiological processes, vector life history, and behaviour. However, the complexity of the model is justified in terms of integrating these diverse aspects of the plant-virus-vector interaction. The approach taken in analysing the effects of vector preference on disease epidemiology is novel for plant viruses, although Roosien et al. (2013) followed a somewhat similar approach using a simpler mathematical model. Our analysis stresses the importance of the basic reproduction number R_0 and the equilibrium or final disease incidence, and how these quantities depend on values of key parameters. It then examines the stability of these equilibria and whether alternative equilibria can be attained depending on initial conditions. Taking an analytical supplemented by a numerical approach enables qualitative as well as quantitative insight. It also enables a clear comparison of differences between NPT and PT. We also provide a user-friendly interface (https://plantdiseasevectorpreference.herokuapp.com/), since in the past we have found this can be very helpful in aiding engagement by non-modellers with modelling results (Cunniffe et al. (2015)).

Life history and epidemiological parameters

In comparing the basic reproduction number, R_0, for NPT and PT, when there is no vector preference, there is a clear divergence for vector death rate (α) which gives a higher R_0 at low death rates for PT, other parameters held at their default values (Figure 4A vs. Figure 4D). Hence, the need for prolonged settling and feeding, typical of PT, leads to a higher R_0 when vector longevity is long. Vector mortality was not explicitly included earlier models (Roosien et al. 2013; Shaw et al. 2017). For final disease incidence (infected plants as a proportion of total plants at equilibrium) the major divergence between the two modes of transmission arises with the host roguing rate (μ), which can equally be interpreted as additional mortality induced by disease (Figure 4C vs. Figure 4F). A higher final incidence is seen for NPT when μ is low, other parameters held at their default values. Hence minimal roguing of infected plants in a system affected by an NPT virus would not have a major effect in reducing disease incidence. Another key feature of the parameter μ is that changing it can induce bistability (Figure 5), potentially making the effects of roguing more unpredictable under certain circumstances. Host mortality – either natural, or additional with roguing or disease-induced – was not included in earlier models targeting vector preference (Roosien et al. 2013; Shaw et al. 2017), despite its effects on disease epidemiology having been shown to be important in past models of virus vector systems without vector preference (Jeger et al., 1998; Madden et al., 2000), as well as in plant disease models more generally (Cunniffe et al., 2014).

Importance of vector preference parameters

A key feature of our model is that it distinguishes vector preferences for the bias to land (ν) from that to settle (ϵ) on infected plants, as well from that for the probability that the vector feeds (ω). Our basic analysis focused on the case in which the preference does not depend on whether the vector is viruliferous or non-viruliferous (in fact this is the only possible case for NPT viruses). For NPT, the one-way sensitivity analysis showing responses to the bias to land (ν) recapitulates an already well-known result, i.e. vectors preferring infected plants aids initial invasion of disease by increasing R_0, but reduces final disease incidence (Figure 6A). Essentially this is because visits by vectors concentrate on

plants that are already infected at high prevalence (McElhany et al. 1995). However, by separating out the different components of vector preference, our model shows that, as the settling bias and feeding probability for NPT increase for any given landing bias, both R_0 and the final disease incidence decrease. This is because vectors which have settled for an extended feed are not able to transmit. An NPT virus able to induce what has been dubbed an "attract-and-deter" plant phenotype in which vectors are attracted to infected plants, but then deterred from settling for an extended feed (Carr et al., 2020) is therefore expected to be most epidemiologically successful.

For PT, since extended feeds are required to acquire and inoculate the virus, any increase to the settling bias ϵ leads to the basic reproduction number increasing (Figure 7B), although the final incidence again decreases. The numerical value of the feeding probability ω is unimportant for PT (Figure 7C), since the formulation of our model properly accounts for the idea that vectors which are engaged in feeding are not able to move between plants and transmit the disease (cf. Equation (27), in which both ω parameters cancel out of the transmission terms). For both NPT and PT, bistability is possible for landing biases $\nu < 1$ and for a range of settling biases $\epsilon <> 1$ (the pink areas on Figures 6A,6G,6H,7A,7B and 7E). These bistabilities are between a stable disease-free equilibrium and a stable disease-present equilibrium, i.e. occur for $R_0 < 1$.

The distinction between landing and settling was made in a previous modelling study by Sisterson (2008), although the model referred only to orientation (rather than settling) bias, and all results presented were based on simulations of an agent-based model. The point has been made (Carmo-Souza et al. 2014; Rajabaskar et al. 2014) that interpretation of experimental results depends on whether probing or settling has been reported. In past modelling studies, Roosien et al. (2013) did not distinguish between landing and feeding, and Shaw et al. (2017) used landing and settling entirely interchangeably. Donnelly et al. (2019) did distinguish the two elements of vector preference, but in a model formulation that is heavily conditioned upon the NPT that is that focus of that work. We concur with Fereres and Moreno (2009) and Mauck et al. (2012, 2016), that the mode of transmission has different effects on landing, settling, feeding and movement between plants. Our model and results provide – for the first time – a framework to understand the distinct effects of these biases when comparing NPT and PT.

Conditional vector preference

Conditional vector preference – in which preferences depend on vectors' own infection status – is only relevant for PT. Both R_0 and the final disease incidence depend on the landing and settling biases of viruliferous and non-viruliferous vectors, with – again – there being no effect of the probability with which individual vectors feed (Figure 8). There is a rapid rise in R_0 as ν_- and ϵ_-, the landing and settling biases shown by non-viruliferous vectors for infected plants, increase. There is no change in R_0 with the biases shown by viruliferous vectors (i.e. ν_+ and ϵ_+), since these parameters do not appear in the R_0 expression (Figures 8A and 8B). Disease incidence increases but soon levels off with increases in the landing and settling biases of non-viruliferous vectors for infected plants. With increasing landing and settling biases of viruliferous vectors for infected plants, however, the final disease incidence decreases from a saturation level of 1.0 when these biases are close to 0. The bias of viruliferous vectors for healthy plants is increasingly important to reaching a high disease incidence.

Our results are ostensibly rather similar those of Roosien et al. (2013) although while those authors did account for conditional vector preference, distinct aspects of vector choice and settling behaviour

were confounded into a single "preference" parameter in their model. As noted by Roosien et al. (2013), there are also similarities to those obtained previously (e.g. McElhany et al. 1995; Sisterson 2008), but those earlier results were based on uniform preferences for infected plants when these are rare and uniform preferences for healthy plants when infected plants are abundant. Thus, their results are based on the rarity or abundance of infected plants, not on a switching of preferences following virus acquisition as in conditional preference.

Interactions with vector population dynamics

A key feature of our model is that effect(s) of virus infection on vector fecundity is accounted for explicitly by introducing β, the proportionate change in the vector birth rate on infected plants (relative to that on healthy plants). Fecundity is a measure of vector performance often used in experimental studies and can increase or decrease on infected plants. Mauck et al. (2012) analysed some 224 experimental studies from 55 published papers on vector attraction, settling and feeding, and performance on virus-infected and healthy plants. For vector performance, they found a bias towards positive outcomes for PT viruses and a bias towards negative outcomes for NPT viruses. However, as noted in the Introduction, evidence for this remains mixed across different vector groups, host species, and transmission classes (Chen et al., 2017; Davis et al. 2017; Tungadi et al. 2017; Smith et al. 2017; Gadhave et al. 2019; Shi et al. 2021). A notable example of how an increase in vector fecundity on virus-infected plants can lead to a marked increase in population size was given by Jiu et al. (2007). The invasive B biotype of the *B. tabaci* complex increased its fecundity on tobacco by 12-fold when infected with tobacco curly shoot virus and by 18-fold when infected with tomato yellow leaf curl virus, whereas the indigenous biotype performed equally on virus-infected and healthy plants. The B biotype displaced the indigenous biotype and was associated with the accelerated spread of the viruses in southern China.

There is no equivalent to the parameter β in Roosien et al. (2013). In Shaw et al. (2017) vector intrinsic growth rates and carrying capacities were specified separately for healthy and infected hosts, but their results were inconsistent with those of Shaw et al. (2019). Another parameter we introduce is a measure of vector losses from the plant population (δ), arising for example from additional mortality (or other forms of loss) due to the number of flights taken per individual feed. There is no equivalent to this parameter in both of Roosien et al. (2013) or Shaw et al. (2017). We assume that additional mortality applies irrespective of the mode of transmission or whether the vector is viruliferous or non-viruliferous.

The pattern in the parameter space defined by β and δ is similar for both NPT and PT viruses (Figure 9). Bistability between the disease-free equilibrium with the vector fixed at carrying capacity and another equilibrium in which the disease has invaded is once again possible. For larger values of δ the vector would die in the absence of disease. However, interestingly, sufficiently large values of the relative birth rate on infected plants can then "rescue" both vector and virus populations, allowing disease to persist for certain initial conditions.

The responses of the final disease incidence to the underlying birth rate of vectors for different values of the relative birth rate on infected plants are also quite similar for NPT and PT, with final disease incidence increasing with the relative birth rate on infected plants in both cases. For PT, responses to certain conditional preference parameters were non-monotone for different values of β and δ (e.g. Figures 10A and 10F), indicating that effects on population dynamics of preferential reproduction on

infected plants, or of vector losses due to additional mortality from flights, can have relatively complex effects. A similar analysis for NPT – in which vector preference parameters were systematically altered, but in which as before numerical values of "plus" and "minus" parameters were tied together for viruliferous and non-viruliferous vectors – led to rich dynamics. For certain combinations of parameters, we noted the model exhibited sustained oscillations (Figures 11E and 11F). However, we did not analyse the dynamical complexity of the model in any further detail, noting that these periodic solutions required extreme values of underlying parameters.

Ecological context

The host range of a plant virus reflects considerable ecological complexity (McLeish et al. 2018). It is clearly difficult to isolate the epidemiological consequences of vector preference from the ecological circumstances found in the field. It is difficult to see how hypotheses on epidemic and evolutionary dynamics given different forms of vector preference could be tested in the field, and this is one reason why modelling can further assist in creating and analysing different scenarios. The example we pursue here is in differentiating the consequences of non-conditional and conditional vector preferences given the other ecological noise around. As with many other examples, this stresses the need to resolve and integrate where possible epidemiology with ecological insights (Jeger, 2020).

Dispersal is a key element of insect ecology especially at the local scale where dispersal characteristics can depend upon plant characteristics (Lombaert et al. 2006). Most transmission events occur in such a setting and the movement of insect vectors could be considered in this context. Dispersal strategies are likely to be an important feature in fragmented plant populations (Cote et al. 2016) where the role of vector preferences in determining non-random movement of vectors in response to host infections status may be critical in disease spread. Plant characteristics are also affected by environmental conditions such as nutrient resources, which can affect vector dispersal and spread of virus (Strauss et al. 2020). However, whether these characteristics interact with host phenotypes of healthy or infected has not been studied. Dispersal is not modelled explicitly in this study, but rather implicitly through the number of flights made by a vector for each individual feed. We assume that consequentially there will a loss of vectors from the system, whether due to random movement of non-colonising vectors, additional predation, or additional mortality. The movement of non-colonising vectors out the system is likely to occur as the number of flights increases. Additional mortality may occur due to predation or parasitism, although it may be the well-known alarm signal of vectors under attack that may have caused movement in the first place (Jeger et al. 2011). The energetic costs associated with number of flights per individual feed have not specifically been studied with respect to virus transmission, although the energetics of pollinating insects in relation to the food resource is well-studied (McCallum et al., 2013),

The model developed is essentially for one virus, one host, and one vector. Co-infection with several viruses is often observed and the interactions between them may affect both the volatile and visual signals picked up by vectors (Penaflor et al. 2016; Gautam et al. 2020). Co-infection is acknowledged to affect underpinning epidemiological dynamics, sometimes in subtle ways (Hilker et al. 2017, Allen et al. 2019, Hamelin et al. 2019). Accounting for different plant hosts might also be important. A further model updating the model of Shaw et al. (2017) has been developed (Schoemaker et al. 2019), in which conditional vector preference was analysed in multi-host plant communities. Results indicated that non-random associations of viruliferous vectors between preferred and non-preferred host species would occur. This model showed that host diversity led to reduced virus spread,

demonstrating the interplay between host diversity and vector behaviour. However, the competitive effects that may occur with multiple vector species have not been addressed, except from a theoretical standpoint (van den Bosch and Jeger, 2017; Zaffaroni et al. (2020)).

Interactions between vectors and non-vector herbivores and pollinators, and other tritrophic interactions can also affect virus spread (Jeger et al. 2011, Groen et al. 2016, Jeger et al. 2018, Chisholm et al. 2019, Crowder et al. 2019, Pradit et al. 2019). Indeed, this may even swamp any signal that comes from vector preference. Links with vector-natural enemy associations have also been confirmed (Liu et al. 2017). Tomato infected with tomato yellow leaf curl virus changed the host preference of the parasitoid *Encarsia formosa* between Q- and B-biotypes of *B. tabaci*: on infected but not healthy plants the Q-biotype was more attractive to the parasitoid than the B-biotype, due to quantitative differences in volatile profiles. A key question in future research is does a link between vector preference, transmission type, and natural enemies lead to increased virus fitness?

Insect vectors of plant viruses may also be herbivore pests. As such, any preference for healthy or diseased plants may affect the extent of damage as direct pests. A model was developed in which herbivory was linked with a simple epidemiological model (Nakazawa et al. 2012). Analysis showed that the virus was better able to persist when the herbivore fed on healthy rather than infected plants. Further, whether the disease establishes an endemic equilibrium depended on the initial conditions, with disease persistence sometimes possible even when the basic reproduction number was under 1. As a result, infected plants were able to persist in the system due to reduced herbivory on these plants but still allowing transmission. Further experimental and modelling studies are required to disentangle direct pest effects from those resulting from virus infection (Zaffaroni et al., 2019). A general model investigated the relationship of the basic reproduction number with a further threshold parameter, the consumption number (Duffy & Collins 2019). The behaviour of the model was complex but yielded results showing that "consumer-resource" co-existence could limit the spread of infectious disease.

Evolutionary implications

An outstanding question to be asked is what are the evolutionary determinants of vector preference? For example, adaptive dynamics (Diekmann et al. 2002) could be used to determine the values of preference parameters that maximise the basic reproduction number. This approach was followed in modelling the evolutionary response of plant viruses to different mechanisms of host plant resistance and disease control (van den Bosch et al. 2006, van den Bosch et al. 2007); and modelling the evolution of virulence/mutualism where there is both horizontal (vector) and vertical (seed/pollen) transmission of plant viruses (Hamelin et al. 2016, Hamelin et al. 2018).

A key issue that has not received much theoretical consideration is how vector preference traits, whether non-conditional or conditional have evolved and the extent to which this is due to the vector, the virus, the host, or the parameters defined by their epidemiological interaction. Gandon (2018) proposed a theoretical framework applicable to general vectored host-parasite interactions, as a framework for analysing the evolution of preference traits, but with most emphasis on vertebrate hosts. Preferences for infected hosts were introduced as ratios in the searching efficiencies of vectors carrying, or not carrying, the parasite on infected and noninfected hosts. Although examples of conditional vector preference in insect-vectored plant viruses were introduced, there was no consideration of the differences in NPT and PT transmission, although some of the obtained results for alternative invasion thresholds may be interpreted in terms of transmission mode. The

evolutionary analysis used adaptive dynamics and gave insight into why some pathogens have evolved manipulative strategies involving both the vector and the host while others have not.

For plant diseases, such an analysis may prove particularly useful for a scenario in which there is a "new encounter" between a virus and an established host-vector association. Very few studies appear to have looked at this scenario. Porath-Krauss et al. (2020) introduce barley yellow dwarf virus into hosts and vectors that the virus had not encountered for thousands of viral generations (estimated as some 250 days), without significant reduction in disease prevalence. In the case of a truly new encounter, we can assume that, prior to first encounter, there would be no vector bias with landing occurring at random. Following Antia et al. (2003), we could then consider invasion of the virus as initially unsuccessful at the first encounter because $R_0 < 1$. Subsequently, after several attempts, the virus could "learn" to manipulate the vector and host through manipulation of the landing biases by promoting virus acquisition by non-viruliferous vectors from infected hosts until $R_0 > 1$. Such an analysis, using a simplified version of the current model, will be made in future work by the authors.

ACKNOWLEDGEMENTS

NJC thanks John Carr and Ruairi Donnelly for very helpful discussions during initial formulation of the mathematical model. NPT acknowledges the Biotechnology and Biological Sciences Research Council of the United Kingdom (BBSRC) for support via a University of Cambridge DTP PhD studentship.

REFERENCES

Allen, L.J.S., Bokil, V.A., Cunniffe, N.J., Hamelin, F.M., Hilker, F.M., Jeger, M.J. (2019) Modelling vector transmission and epidemiology of co-infecting plant viruses. Viruses. 11:1153.

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.

Bosque-Pérez, N.A., Eigenbrode, S.D. (2011) The influence of virus-induced changes in plants on aphid vectors: Insights from luteovirus pathosystems. Virus Res. 159:201–205.

Boyd, I.L., Freer-Smith, P.H., Gilligan, C.A., Godfray, H.C.J. (2013) The consequences of tree pests and diseases for ecosystem services. Science. 342:1235773.

Buonomo, B., Vargas-De-León, C. (2013) Stability and bifurcation analysis of a vector-bias model of malaria transmission. Math. Biosci. 242:59–67.

Carmo-Souza, M., Moreno, A., Garzo, E., Fereres, A. (2014) A non-persistently transmitted-virus induces a push-pull strategy in its aphid vector to optimize transmission and spread. Virus Res. 186:38–46.

Carr, J.P., Tungadi, T., Donnelly, R., Bravo-Cazar, A., Rhee, S.-J., Watt, L.G., Mutuku, J.M., Wamonje, F.O., Murphy, A.M., Arinaitwe,W., et al. (2020) Modeling and manipulation of aphid-mediated spread of non-persistently transmitted viruses. Virus Res. 277:197845.

Chamchod, F., Britton, N. F. (2011) Analysis of a vector-bias model on malaria transmission. Bull. Math. Biol. 73:639–57.

Chen, G., Su, Q., Shi, X.B., Liu, X., Peng, Z.K., Zheng, H.X., Xie, W., Xu, B.Y., Wang, S.L., Wu, Q.J., et al. (2017) Odor, not performance dictates Bemisia tabaci's selection between healthy and virus infected plants. Front. Physiol. 8:146.

Chen, Y., Lu, C., Li, M.,Wu, W., Zhou, G., Wei, T. (2016) Adverse effects of *Rice gall dwarf virus* upon its insect vector *Recilia dorsalis* (Hemiptera: Circadellidae). Plant Disease. 100:784-790.

Chesnais, Q., Couty, A., Uzest, M., Brault, V., Ameline, A. (2019) Plant infection by two different viruses induce contrasting changes of vector fitness and behavior. Insect Sci. 26:86–96.

Chesnais, Q., Vidal G.C., Coquelle, R., Ivon, M., Mauck, K. E., Brault, V., Ameline, A. (2020) Post-acquisition effects of viruses on vector behaviour are important components of manipulation strategies. Oecologia. 194:429-440.

Chisholm, P. J., Eigenbrode, S. D., Clark, R.E., Basu, S., Crowder, D. W. (2019) Plant-mediated interactions between a vector and non-vector herbivore promote the spread of a plant virus. Proc. R. Soc. B. 286:20191383.

Clark, R. E., Basu, S., Lee, B. W., Crowder, D. W. (2019) Tri-trophic interactions mediate the spread of a vector-borne plant pathogen. Ecology. 100:e02879.

Claudel, P., Chesnais, Q., Fouché, Q., Krieger, C., Halter, D., Bogaert, F., Meyer, S., Boissinot, S., Hugueney, P., Ziegler-Graff, V., Ameline, A., Brault, V. (2018) The aphid-transmitted *Turnip yellow virus* differentially affects volatiles emission and subsequent vector behaviour in two *Brassicaceae* plants. Int. J. Mol. Sci. 19:2316.

Cote, J., Bestion, E., Jacob, S., Travis, J., Legrand, D., and Baguette, M. (2017) Evolution of dispersal strategies and dispersal syndromes in fragmented landscapes. Ecography. 40:56-73.

Cozzarolo, C.-S., Glaizot, O., Christe, P., Pigeault, R. (2020)Enhanced attraction of arthropod vectors to infected vertebrates: a review of empirical evidence. Front. Ecol. Evol. 8:568140.

Crowder, D. W., Li, J., Borer, E. T., Finke, D, L., Sharon, F., Pattemore, D. E., Medlock, J. (2019) Species interactions affect the spread of vector-borne plant pathogens independent of transmission mode. Ecology. 100:e02782.

Cunniffe, N.J., Gilligan, C.A. 2020 Use of mathematical models to predict epidemics and to optimise disease detection and management. Chapter 12 in Emerging Plant Diseases and Global Food Security, APS Press.

Cunniffe, N.J., Koskella, B., Metcalf, C.J.E., Parnell, S., Gottwald, T.R., Gilligan, C.A. (2015) Thirteen challenges in modelling plant diseases. Epidemics. 10:6-10.

Cunniffe, N.J., Laranjeira, F.F., Neri, F.M., DeSimone, R.E., Gilligan, C.A. (2014) Cost-effective control of plant disease when epidemiological knowledge is incomplete: modelling Bahia bark scaling of citrus. PLOS Computational Biology. 10:e1003753.

Cunniffe, N.J., Stutt, R.O.J.H., DeSimone, R.E., Gottwald, T.R., Gilligan, C.A. (2015) Optimising and communcating options for the control of invasive plant pathogens when there is epidemiological uncertainty. PLOS Computational Biology 11, e1004211.

Cunniffe, N.J., Stutt, R.O.J.H., van den Bosch, F. & Gilligan, C.A. (2012) Time-dependent infectivity and flexible latent and infectious periods in compartmental models of plant disease. Phytopathology. 102:365–380.

Daimei, G., Raina, H.S., Devi, P.P., Saurav, G.K., Renukadevi, P., Malathi, V.G., Senthilraja, C., Mandel, B., Rajagopal, R. (2017) Influence of Groundnut bud necrosis on the life history traits and feeding preferences of its vector, *Thrips palmi*. Phytopathology. 107:1440–1445.

Davis, T.S., Wu, Y., Eigenbrode, S.D. (2017) The effects of bean leafroll virus on life history traits and host selection behaviour of specialized pea aphid (*Acyrthosiphon pisum*, Hemiptera: Aphididae) genotypes. Environ. Entomol. 46:68–74.

Donnelly, R., Gilligan, C.A. (2020) What is pathogen-mediated insect superabundance. J. R. Soc. Interface. 17: 20200229.

Donnelly, R., Cunniffe, N.J., Carr, J.P., Gilligan, C.A. (2019) Pathogenic modification of plants enhances long-distance dispersal of non-persistently transmitted viruses to new hosts. Ecology 100:e02725.

Duffy, K.J., Collins, O.C. (2019) Consumer-resource co-existence as a means of reducing infectious disease. Journal of Biological Dynamics 13:177-191.

Eigenbrode, S.D., Bosque-Pérez, N.A., Davis, T.S. (2018) Insect-Borne Plant Pathogens and Their Vectors: Ecology, Evolution, and Complex Interactions. Annu. Rev. Entomol. 63:169–191.

Fabre, F., Coville, J., Cunniffe, N.J. (2021) Optimising reactive disease management using spatially explicit models at the landscape scale. Chapter 4 in Plant Diseases and Food Security in the 21st Century. Springer.

Fereres, A., Moreno, A. (2009) Behavioual aspects influencing plant virus transmission by homopteran insects. Virus Research. 141:158-168.

Gadhave, K.R., Dutta, B., Coolong, T., Srinivasan R. (2019) A non-persistent aphid-transmitted *Potyvirus* differentially alters the vector and non-vector biology through host plant quality manipulation. Scientific Reports. 9:2503.

Gandon, S. (2018) Evolution and manipulation of vector host choice. Am. Nat. 192:23-34.

Gautam, S., Gadhave, K. R., Buck, J. W., Dutta, B., Coolong, T., Adkins, S., Srinivasan, R. (2020) Virus-virus interactions in a plant host and in a hemipteran vector: implications for vector fitness and virus epidemics. Virus Res. 286:198069.

Groen S. C., Jiang, S., Murphy, A. M., Cunniffe, N. J., Westwood, J. H., Davey, M. P., et al. (2016) Virus infection of plants alters pollinator preference: a payback for susceptible hosts? PLoS Pathog. 12:e1005790.

Hamelin, F. M., Allen, L. J. S., Prendeville, H. R., Hajimorad, M. R., Jeger, M. J. (2016) The evolution of plant virus transmission pathways. J. Theor. Biol. 396:75-89.

Hamelin, F.M., Hilker, F.M., Sun, T.A., Jeger, M.J., Hajimorad, R., Allen, L.J.S., Prendeville, H.R. (2017) The evolution of parasitic and mutualistic plant-virus symbioses through transmission-virulence trade-offs. Virus Res. 241:77–87.

Hamelin, F.M., Allen, L.J.S., Bokil, V.A., Gross, L.J., Jeger, M.J., Manore, C.A., Power, A.G., Rua, M.A., Cunniffe, N.J. (2019) Coinfections by noninteracting pathogens are not independent and require new tests of interaction. PLOS Biology. 17:e300551.

Hebert, M.P., Allen, L.J.S. (2016) Disease outbreaks in plant-vector-virus models with vector aggregation and dispersal. Discr. Contin, Dynam. Syst. Ser. B. 21:2169-2191.

Heesterbeek, J.A.P., Roberts, M.G. (2007) The type-reproduction number T in models for infectious disease control. Math. Biosci. 206:3–10.

Higashi, C.H.V., Bressan A. (2013) Influence of a propagative plant virus on the fitness and wing dimorphism of infected and exposed insect vectors. Oecologia. 172:847-856.

Hilker, F.M., Allen, L.J.S., Bokil, V.A., Briggs, C.J., Feng, Z., Garrett, K.A., Gross, L.J., Hamelin, F.M., Jeger, M.J., Manore, C.A., Power, A.G., Redinbaugh, M.G., Rua, M.A., Cunniffe, N.J. (2017) Modeling virus coinfection to inform management of maize lethal necrosis in Kenya. Phytopathology. 107:1095-1108.

Holt, J., Jeger, M., Thresh, J., Otim-Nape, G. (1997) An Epidemiological Model Incorporating Vector Population Dynamics Applied to African Cassava Mosaic Virus Disease. J. Appl. Ecol. 34:793–806.

Ingwell, L.L., Eigenbrode, S.D., Bosque-Pérez, N.A. (2012) Plant viruses alter insect behavior to enhance their spread. Sci. Rep. 2:578.

Jeger, M.J. (2020) The epidemiology of plant virus disease: towards a new synthesis. Plants. 9:1768.

Jeger, M., van den Bosch, F., Madden, L., Holt, J. (1998) A model for analysing plant-virus transmission characteristics and epidemic development. Math. Med. Biol. 1:1–18.

Jeger, M.J., van den Bosch, F., McRoberts, N. (2015) Modelling transmission characteristics and epidemic development of the tospovirus-thrip interaction. Arthropod-Plant Interact. 9:107-120.

Jeger, M.J., Chen, Z., Powell, G., Hodge, S., van den Bosch, F. (2011) Interactions in a host plant-virus-vector-parasitoid system: modelling the consequences for virus transmission and disease dynamics. Virus Res. 159:183–93.

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. Physiol. Entomol. 29:291–304.

Jeger, M.J., Madden, L.V., van den Bosch, F. (2018) Plant virus epidemiology: application and prospects for mathematical modeling and analysis to improve understanding and disease control. Plant Dis. 102:837-854.

Jiu, M., Zhou, Z-P., Tong, L., Xu, J., Yang, X., Wan, F-H., Liu, S-S. (2007) Vector-virus mutualism accelerates population increase of an invasive whitefly. PLoS ONE 2: e182.

Jones, R.A.C., Naidu, R.A. (2019) Global dimensions of plant virus dieases: current status and future perspectives. Ann. Rev. Virol. 6:387-409.

Kingsolver, J.G. (1987) Mosquito host choice and the epidemiology of malaria. Am. Nat. 130:811-827.

Lagos-Kutz, D., Pawlowski, M.L., Haudenshield, J., Han, J., Domier, L.L., Hartman, G.L. (2020) Evaluation of soybean for resistance to *Neohyadatothrips variabilis* (Thysanoptera: Thripidae) noninfected and infected with soybean vein necrosis virus. J. Econ. Entomol. 113:949-955.

Liu, X., Chen, G., Zhang, Y.J., Xie, W., Wu, Q.J., Wang, S.L. (2017) Virus-infected plants altered the host selection of Encarsia formosa, a parasite of whiteflies. Front. Physiol. 8:937.

Lombaert, E. Boll, R., Lapchin, L. (2006) Dispersal strategies of phytophagous insects at a local scale: adaptive potential of aphids in an agricultural environment. BMC Evolutionary Biology. 6:75.

Madden, L.V, Jeger, M.J., van den Bosch, F. (2000) A theoretical assessment of the effects of vector-virus transmission mechanism on plant virus disease epidemics. Phytopathology. 90:576–594.

Mauck, K.E., Chesnais, Q. (2020) A synthesis of virus-vector associations reveals important deficiencies in studies on host and vector manipulation by plant viruses. Virus Res. 285:197957.

Mauck, K., Bosque-Pérez, N.A., Eigenbrode, S.D., De Moraes, C.M., Mescher, M.C. (2012) Transmission mechanisms shape pathogen effects on host-vector interactions: evidence from plant viruses. Funct. Ecol. 26:1162–1175

Mauck, K.E., Smyers, E., De Moraes, C.M., Mescher, M.C. Virus infection influences host plant interactions with non-vector herbivores and predators. Funct. Ecol. 2015, 29, 662–673.

Mauck, K.E., De Moraes, C.M., Mescher, M.C. (2016). Effects on pathogens on sensory-mediated interactions between plants and insect vectors. Current Opinion in Plant Biology 32, 53-61.

Mauck, K.E., Chesnais, Q., Shapiro, L.R. (2018) Evolutionary determinants of host and vector manipulation by plant viruses. Adv. Virus Res. 101:189–250.

Mauck, K. E., Kenney, J., Chesnais, Q. (2019) Progress and challenges in identifying molecular mechanisms underlying host and vector manipulation by plant viruses. Curr. Opin. Ins. Sci. 33:7–18.

McCallum, K.P., McDougall, F.O., Seymour, R.S. (2013) A review of the energetics of pollination biology. Journal of Comparative Physiology B. 183:867-876.

McElhany, P, Real, L.A., Power, A.G. (1995) Vector preference and disease dynamics: a study of barley yellow dwarf virus. Ecology. 76:444-457.

McLeish, M. J., Fraile, AS., García-Arenal, F. (2018) Ecological complexity in plant virus host range evolution. Adv. Vir. Res. 101:293-339.

Nakazawa, T., Yamanaka, T., Urano, S. (2012) Model analysis for plant disease dynamics co-mediated by herbivory and herbivore-borne phytopathogens. Biology Letters. 8:685-688.

Penaflor, M.F., Mauck, K.E., Alves, K.J., de Moraes, C.M., Mescher, M.C. (2016) Effects of single and mixed infections of Bean pod mottle virus and Soybean mosaic virus on host-plant chemistry and host-vector interactions. Funct. Ecol. 30:1648–1659.

Ogada, P.A., Moualeu, D.P., Poehling, H.-M. (2016) Predictive models for *Tomato spotted wilt virus* spread dynamics, considering *Frankliniella occidentalis* specific life processes as influenced by the virus. PLOS ONE. 11:e0154533.

Plotly. 2015. Collaborative data science. https//plot.ly

Porath-Krause, A., Campbell, R., Shoemaker, L., Sieben, A., Strauss, A.T., Shaw, A.K., Seabloom, E.W., Borer, E.T. (2021) Pliant pathogens: Estimating viral spread when confronted with new vector, host, and environmental conditions. Ecology and Evolution. 11:1877-1887.

Pradit, N., Mescher, M.C., Wang, Y., Vorsa, N., Rodriguez-Saona, C. (2019) Phytoplasma infection of cranberries benefits non-vector phytophagous insects. Front. Ecol. Evol. 7:181.

Rajabaskar, D., Bosque-Pérez, N.A., Eigenbrode, S.D. (2014) Preference for a virus vector for infected plants is reversed after virus acquisition. Virus Res. 186:32–37.

Ristaino, J.B., Anderson, P.K, Bebber, D.P, Brauman, K.A, Cunniffe, N.J., Fedoroff, N.V., Finegold, C., Garrett, K.A., Gilligan, C.A., Jones, C.M., Martin, M.D., MacDonald, G.K., Neenan, P., Records, A., Schmale, D.G., Tateosian, L., Wei, Q. (2021) The persistent threat of emerging plant disease pandemics to global food security. Proceedings of the National Academy of Sciences. 118:e2022239118.

Roosien, B.K., Gomulkiewicz, R., Ingwell, L.L., Bosque-Pérez, N.A., Rajabaskar, D., Eigenbrode, S.D. (2013) Conditional vector preference aids the spread of plant pathogens: results from a model. Environ. Entomol. 42:1299–308.

Savary, S., Willocquet, L., Pethybridge, S.J., Esker, P., McRoberts, N., Nelson, A. (2019) The global burden of pathogens and pests on major crops. Nat. Ecol. & Evol. 3:430-439.

Shaw, A.K., Peace, A., Power, A.G., Bosque-Pérez, N.A. (2017) Vector population growth and condition-dependent movement drive the spread of plant pathogens. Ecology. 98:2145–2157.

Shaw, A.K., Peace, A., Power, A.G., Bosque-Pérez, N.A. 2018. (Correction to Shaw et al. 2017). Ecology. 99:2904.

Shaw, A. K., Igoe, M., Power, A. G., Bosque-Pérez, N.A., Peace, A. (2019) Modelling approach influences dynamics of a vector-borne pathogen system. Bull. Math. Biol. 81:2011-2018.

Shi, X., Deng, J., Zhang, Z., Yan, S., Zheng, L., Sun, S., Gao, Y., Zhou, X., Zhang, D., Liu, Y. (2021) Initial ingestion of CMV-infected plants reduces subsequent aphid performance. Arthropod-Plant Interactions. 15:153-160.

Shoemaker, L.G., Hayhurst, E., Weiss-Lehman, C.P., Strauss, A.T., Porath-Krause, A., Borer, E.T., Seabloom, E.W., Shaw, A.K. (2019) Pathogens manipulate the preference of vectors, slowing disease spread in a multi-host system. Ecol. Lett. 22:1115-1125.

Shrestha, D., McAuslane, H.J., Adkins, S.T., Smith, H.A., Dufault, N., Colee, J.,Webb, S.E. (2017) Host-mediated effects of semipersistently transmitted Squash vein yellowing virus on Sweetpotato whitefly (Hemiptera: Aleyrodidae) behaviour and fitness. J. Econ. Entomol. 110:1433–1441.

Sisterson, M.S. (2008) Effects of insect-vector preference for healthy or infected plants on pathogen spread: insights from a model. J. Econ. Entomol. 101:1-8.

Sisterson, M.S. (2009) Transmission of insect-vectored pathogens:effects of vector fitness as a function of infectivity status. Environ. Entomol. 38:345-355.

Smith, C.M., Gedling, C.R., Wiebe, K.F., Cassone, B.J. (2017) A sweet story: Bean pod mottle virus transmission dynamics by Mexican bean beetles (*Epilachna varivestis*). Genome Biol. Evol. 9:714–725.

Strauss, A.T., Henning, J.A., Porath-Krause, A., Asmus, A.L., Shaw, A.K., Borer, E., Seabloom, E.W. (2020) Vector demography, dispersal and the spread of disease: Experimental epidemics under elevated resource supply. Funct. Ecol. 34:2560–2570.

Tungadi, T., Groen, S.C., Murphy, A.M., Pate, A.E., Iqbal, J., Bruce, T.J.A., Cunniffe, N.J., Carr, J.P. (2017) Cucumber mosaic virus and its 2b protein alter emission of host volatile organic compounds but not aphid vector settling in tobacco. Virol. J. 14:91.

van den Bosch, F., Akudibillah, G., Seal, S.E., Jeger, M.J. (2006) Host resistance and the evolutionary response of plant viruses. J. Appl. Ecol. 43:506-516.

van den Bosch, F., Jeger, M.J. (2017) The basic reproduction number of vector-borne plant virus epidemics. Virus Res. 241:196-202.

van den Bosch, Jeger, M.J., Gilligan, C.A. (2007) Disease control and its selection for damaging plant virus strains in vegetatively propagated staple food crops, a theoretical assessment. Proc. R. Soc. B. 274:11-18.

van den Driessche, P. (2017) Reproduction numbers of infectious disease models. Infectious Disease Modelling. 2:288-303.

Zaffaroni, M., Cunniffe, N.J., Bevacqua, D. (2020). An ecophysiological model of plant-pest interactions: the role of nutrient and water availability. J. Roy. Soc. Interface. 17:20200356.

Zaffaroni, M., Rimbaud, L., Maileret, L., Cunniffe, N.J., Bevacqua, D. (2021) The role of interference between vectors in control of plant diseases. bioRxiv https://doi.org/10.1101/2021.06.08.447498

Zeilinger, A.R., Daugherty, M.P. (2014) Vector preference and host defense against infection interact to determine disease dynamics. Oikos. 123:613–622.

Supplementary Table S1: Comparison with previous models of vector preference

A table showing a comparison between the current model and the models of Roosien et al. (2013), Shaw et al. (2017) and Gandon (2018) is given below.

(a) Host plant and vector variables and parameters

	Roosien et al. 2013		Shaw et al. 2017		Gandon 2018		Present model	
	Description	Symbol	Description	Symbol	Description	Symbol	Description	Symbol
Host plant variables	Fraction of noninfected plants	1-I	Fraction of hosts that are healthy	H	Density of susceptible hosts	S	Density of susceptible and healthy plants	S
	Fraction of infected plants	I	Fraction of hosts that are infected	I	Density of infected hosts	I	Density of infected plants	I
	Plant population size = 1		Number of hosts/ha Minimum number of hosts/ha needed to support the vectors	F ξ	Host density	S + I	Density of plants in the absence of virus	N
Host plant population dynamic parameters	Mortality rate of hosts not modelled				Mortality rate of infected hosts	d	Rate at which plants die naturally Rate of roguing (or disease-induced mortality	ρ μ
Vector variables	Fraction of non-inoculative vectors	1-Z	Number of non-viruliferous vectors on healthy hosts/ha Number of non-viruliferous vectors on infected hosts/ha	N_h N_i	Density of susceptible vectors	V_s	Density of non-viruliferous vectors	X
	Fraction of inoculative vectors	Z	Number of viruliferous vectors on healthy hosts/ha Number of viruliferous vectors on infected hosts/ha	V_h V_i	Density of infected vectors	V_i	Density of viruliferous vectors	Z
	Vector population size = 1		Number of vectors/ha	$N_h + N_i + V_h + V_i$	Density of vectors	$V_s + V_i$	Density of vectors in the absence of virus	κ
Vector population dynamic parameters	Vector population dynamics not modelled		Intrinsic vector growth rate on healthy hosts Intrinsic growth rate on infected hosts	r_h r_i	Per capita fecundity of uninfected vectors Per capita fecundity of infected vectors Quality of the infected host relative to the susceptible host	f_s f_i ϕ	Per capita birth rate as vector population density tend to 0 Proportionate change in birth rate on infected plants	σ β
			Vector carrying capacity/healthy host Vector carrying capacity/infected host	K_h K_i	Intensity of density dependence on vector fecundity Maximal fecundities	κ F_s, F_i	Vector population density as density dependence reduces the birth rate to 0	ζ
			Mortality (and birth) rate not modelled explicitly		Mortality rate of susceptible vectors Mortality rate of infected vectors	δ_{Vs} δ_{Vi}	Rate at which vectors die in the absence of any effect due to number of flights Increase in death rate as more plants are visited per feed	α δ

(b) Transmission and vector preference parameters

	Roosien		Shaw		Gandon		Cunniffe	
	Description	Symbol	Description	Symbol	Description	Symbol	**Description**	symbol
Transmission	Compound parameter: average number of vectors per plant × frequency of vector movement × probability that an inoculative vector transmits the virus to a noninfected plant	ϕ_p	Maximum vector departure rate from infected plants Rate healthy hosts become infected	a_h β_i	Handling time Probability that a susceptible host gets infected by an infected vector Probability of superinfection of an infected host	τ β ϕ_H	Average number of plants visited by viruliferous vector per day Probability that an uninfected plant is inoculated by a viruliferous vector in a single visit	ϕ_+ γ
	Compound parameter: average number of vectors per plant × frequency of vector movement × probability that a noninoculative vector acquires the virus from an infected plant	ϕ_v	Maximum vector departure rate from healthy plants Rate vector on infected host becomes virulent	a_i β_v	Handling time Probability that a susceptible vector gets infected on an infected host Probability of superinfection of an infected vector	τ b ϕ_V	Average number of plants visited by non-viruliferous vector per day Probability that non-viruliferous vector acquires the virus in a single visit to an infected plant	ϕ_- η
	Dispersal not modelled		Dispersal loss (for each status) Same status half-departure constant Same status half-departure constant	μ c_1 c_2	Dispersal not modelled			
Vector preference	Preference parameter of inoculative vectors for infected plants	ρ_+	Preference of virulent vectors for settling on infected hosts	ε	Searching efficiency of infected vectors on susceptible hosts Searching efficiency of infected vectors on infected hosts Preference for infected hosts in infected vectors (derived)	α_s α_i π	Bias of viruliferous vectors to land on infected plants	ν_+
	Preference parameter of noninoculative vectors for infected plants	ρ_-	Preference of nonvirulent vectors for settling on infected hosts	δ	Searching efficiency of susceptible vectors on susceptible hosts Searching efficiency of susceptible vectors on infected hosts Preference for infected hosts in susceptible vectors (derived)	a_s a_i p	Bias of non-viruliferous vectors to land on infected plants	ν_-
					Cost associated with a bias in vector preference	ρ		
Feeding preference							Average time spent feeding on a plant when vector chooses to settle after landing	Γ

	Feeding not modelled					Probability that non-viruliferous vector settles to feed on a susceptible plant Bias of non-viruliferous vector to feed on an infected plant Probability that viruliferous vector settles to feed on a susceptible plant Bias of viruliferous vector to feed on an infected plant	ω_- ε_- ω_+ ε_+
Recovery rates	Recovery rate of infected plants/rate of replacement of infected plants with noninfected plants	γ	Not modelled			Replacement rate of infected plants with healthy susceptible plants	ρ
	Recovery rate of inoculative vectors/rate of replacement of inoculative vectors with noninoculative vectors	τ	Vector recovery rate by feeding	γ	Not modelled	Rate at which vectors lose infectivity	τ
Invasion criterion		derived	Not derived		3 versions derived	derived	

Supplementary Appendix S2. Deriving the Basic Reproduction Number via the Next Generation Matrix approach

Focusing on the virus-carrying compartments in Equation (26) leads to

$$\frac{dI}{dt} = \frac{\gamma ZS}{\omega_+ \Gamma(S + \nu_+ \epsilon_+ I)} - (\rho + \mu)I$$

$$\frac{dZ}{dt} = \frac{\eta \nu_- XI}{\omega_- \Gamma(S + \nu_- \epsilon_- I)} - (\tau + \alpha_+)Z$$

The disease-free equilibrium is $(S, I, X, Z) = (N, 0, \kappa, 0)$. Following the standard notation for Next Generation calculations (van den Driessche, 2017), take the rates at which new infections are caused

$$F = \begin{pmatrix} \dfrac{\gamma ZS}{\omega_+ \Gamma(S + \nu_+ \epsilon_+ I)} \\ \dfrac{\eta \nu_- XI}{\omega_- \Gamma(S + \nu_- \epsilon_- I)} \end{pmatrix}$$

and the rates at which infections are removed

$$V = \begin{pmatrix} \rho + \mu \\ \tau + \alpha_+ \end{pmatrix}$$

The Jacobians of these matrices are

$$J_F = \begin{pmatrix} -\dfrac{\gamma ZS \nu_+ \epsilon_+}{\omega_+ \Gamma(S + \nu_+ \epsilon_+ I)^2} & \dfrac{\gamma S}{\omega_+ \Gamma(S + \nu_+ \epsilon_+ I)} \\ \dfrac{\eta \nu_- XS}{\omega_- \Gamma(S + \nu_- \epsilon_- I)^2} & 0 \end{pmatrix}$$

and

$$J_V = \begin{pmatrix} \rho + \mu & 0 \\ 0 & \tau + \alpha_+ \end{pmatrix}$$

Evaluating $M = J_F J_V^{-1}$ at the disease-free equilibrium gives

$$M = \begin{pmatrix} 0 & \dfrac{\gamma}{\omega_+ \Gamma(\tau + \alpha_+)} \\ \dfrac{\eta \nu_- \kappa}{\omega_- \Gamma N(\rho + \mu)} & 0 \end{pmatrix}$$

The basic reproduction number, R_0, is then the spectral radius of the matrix M, i.e.

$$R_0 = \sqrt{\frac{\gamma \eta \nu_- \kappa}{\omega_- \omega_+ \Gamma^2 N(\tau + \alpha_+)(\rho + \mu)}}$$

which matches the result given in Equation (36) of the main text.

Supplementary Appendix S3. Finding model equilibria

Disease free equilibria

These correspond to equilibria for which $I = Z = 0$. There are two such equilibria:

- E1: $(S, I, X, Z) = (N, 0,0,0)$
- E2: $(S, I, X, Z) = (N, 0, \kappa, 0)$

where the derived parameter κ is as defined in Equation (24), i.e.

$$\kappa = \zeta\left(1 - \frac{\alpha}{\sigma}\left(1 + \delta\left(\frac{1}{\omega_-} - 1\right)\right)\right)$$

We note that the equilibrium E1 is always biologically meaningful, whereas E2 can only be attained in practice if $\kappa > 0$.

At either equilibrium, the Jacobian of Equation (26) is given by

$$J = \begin{pmatrix} -\rho & 0 & 0 & * \\ 0 & -(\rho + \mu) & 0 & {}^{\partial\Lambda}/_{\partial \mathrm{Z}} \\ * & * & {}^{\partial \mathrm{g}}/_{\partial \mathrm{X}} - h_- & * \\ 0 & {}^{\partial\Omega}/_{\partial \mathrm{I}} & 0 & -(\tau + h_+) \end{pmatrix}$$

in which * denotes values not needed in the subsequent analysis, and where partial derivatives, and the two functions h_- and h_+ are understood to be evaluated at the equilibrium under consideration.

The local stability of an equilibrium is controlled by the eigenvalues of J, i.e. by the characteristic equation $\det(J - \lambda I) = 0$. Expanding down the third column indicates that one eigenvalue is given by

$$\lambda = {}^{\partial \mathrm{g}}/_{\partial \mathrm{X}} - h_- = \sigma\left(1 - \frac{2X}{\zeta}\right) - \alpha\left(1 + \delta\left(\frac{1}{\omega_-} - 1\right)\right)$$

in which X is the value at the equilibrium in question.

For E1, $X = 0$ meaning that $\lambda = \frac{\sigma\kappa}{\zeta}$ (and so that this equilibrium is definitely unstable whenever $\kappa > 0$ – i.e. whenever E2 is biologically meaningful – but potentially stable if $\kappa < 0$, depending on the other three eigenvalues).

For E2, $X = \kappa$ and so $\lambda = -\frac{\sigma\kappa}{\zeta}$ (meaning this equilibrium is potentially stable whenever $\kappa > 0$, again depending on the rest of the eigenvalues, but must definitely be unstable if $\kappa < 0$).

The remaining three eigenvalues are controlled by the eigenvalues of the 3x3 matrix

$$J^* = \begin{pmatrix} -\rho & 0 & * \\ 0 & -(\rho + \mu) & {}^{\partial\Lambda}/_{\partial \mathrm{Z}} \\ 0 & {}^{\partial\Omega}/_{\partial \mathrm{I}} & -(\tau + h_+) \end{pmatrix}$$

Expanding the characteristic equation of this second matrix down the first column indicates there is always a negative eigenvalue $\lambda = -\rho$.

The remaining pair of eigenvalues come from the 2 x 2 matrix

$$J^{**} = \begin{pmatrix} -(\rho+\mu) & \partial\Lambda/\partial Z \\ \partial\Omega/\partial I & -(\tau+h_+) \end{pmatrix}$$

This third matrix always has a negative trace, $Tr(J^{**}) = -(\rho+\mu) - (\tau+h_+)$, and has determinant

$$\Delta(J^{**}) = (\rho+\mu)(\tau+h_+) - \partial\Omega/\partial I \, \partial\Lambda/\partial Z = (\rho+\mu)(\tau+h_+) - \frac{\gamma\eta\nu_- X}{N\omega_-\omega_+\Gamma^2}$$

in which again X is either $X = 0$ (E1) or $X = \kappa$ (E2).

For E1, $\Delta(J^{**}) = (\rho+\mu)(\tau+h_+) > 0$, and since $Tr(J^{**}) < 0$, both eigenvalues of J^{**} are negative.

For E2, the determinant is

$$\Delta(J^{**}) = (\rho+\mu)(\tau+h_+)\left(1 - \frac{\gamma\eta\nu_-\kappa}{N\omega_-\omega_+\Gamma^2(\rho+\mu)(\tau+h_+)}\right) = (\rho+\mu)(\tau+h_+)(1-R_0^2)$$

and so if $R_0^2 < 1$ then both eigenvalues are negative, whereas if $R_0^2 > 1$ then one is positive. This means that $R_0^2 > 1$ ensures this equilibrium is unstable.

Collating the conclusions of the results above

$\boldsymbol{\kappa > 0}$	$\boldsymbol{R_0 > 1}$	**Stability of E1 (Disease- and vector-free equilibrium)**	**Stability of E2 (Disease-free but vector-present equilibrium)**
✓	✓	Unstable	Unstable
✓	×	Unstable	Stable
×	✓	Stable	Not biologically meaningful
×	×	Stable	Not biologically meaningful

Disease present equilibria

This corresponds to equilibria for which $I \neq 0$.

Case I. Vector preference does not interact with population dynamics ($\boldsymbol{\beta = 1, \delta = 0}$)

Adding the equations for dS/dt and dI/dt from Equation (26) and equating to zero indicates

$$0 = \rho N - \rho S - (\rho+\mu)I$$

Adding the equations for dX/dt and dZ/dt indicates (at least when $\beta = 1$ and $\delta = 0$) that either $X + Z = 0$, which is guaranteed to generate no biologically meaningful equilibria, or that

$$X + Z = \zeta\left(1 - \frac{\alpha}{\sigma}\right) = \kappa$$

Using these expressions to eliminate S and X from Equation (26), then eliminating Z, noting throughout that $\delta = 0$ means that $h_+ = \alpha$, eventually leads to

$$I\left(\frac{\gamma\left(N - \left(1+\frac{\mu}{\rho}\right)I\right)\eta\kappa\nu_-}{\left((\tau+\alpha)\omega_-\Gamma\left(N - \left(1+\frac{\mu}{\rho}\right)I + \nu_-\epsilon_- I\right) + \eta\nu_- I\right)\left(\omega_+\Gamma\left(N - \left(1+\frac{\mu}{\rho}\right)I + \nu_+\epsilon_+ I\right)\right)} - (\rho+\mu)\right) = 0$$

Ignoring the equilibrium at $I = 0$, which has been covered above, further equilibria are given by solutions to

$$\frac{\gamma\left(N-\left(1+\frac{\mu}{\rho}\right)I\right)\eta\kappa\nu_-}{\left((\tau+\alpha)\omega_-\Gamma\left(N-\left(1+\frac{\mu}{\rho}\right)I+\nu_-\epsilon_-I\right)+\eta\nu_-I\right)\left(\omega_+\Gamma\left(N-\left(1+\frac{\mu}{\rho}\right)I+\nu_+\epsilon_+I\right)\right)}=(\rho+\mu)$$

Simple, albeit long-winded, algebraic manipulations reduce this to the quadratic

$$a_2I^2+a_1I+a_0=0$$

in which

$$a_2=\left(\nu_-\epsilon_--\left(1+\frac{\mu}{\rho}\right)+\frac{\eta\nu_-}{(\alpha+\tau)\omega_-\Gamma}\right)\left(\nu_+\epsilon_+-\left(1+\frac{\mu}{\rho}\right)\right)$$

$$a_1=N\left(\nu_-\epsilon_-+\nu_+\epsilon_++\frac{\eta\nu_-}{(\alpha+\tau)\omega_-\Gamma}+\left(1+\frac{\mu}{\rho}\right)(R_0^2-2)\right)$$

$$a_0=N^2(1-R_0^2)$$

The values of the other state variables can then be obtained by back-substitution into

$$S=N-\left(1+\frac{\mu}{\rho}\right)I$$

$$Z=\frac{\eta\nu_-\kappa I}{(\alpha+\tau)\omega_-\Gamma\left(N-\left(1+\frac{\mu}{\rho}\right)I+\nu_-\epsilon_-I\right)+\eta\nu_-I}$$

$$X=\kappa-Z$$

Case II. Vector preference interacts with population dynamics ($\boldsymbol{\beta}\neq\mathbf{1}$ and/or $\boldsymbol{\delta}>\mathbf{0}$)

Similar calculations indicate that again

$$S=N-\left(1+\frac{\mu}{\rho}\right)I$$

but in this case the values of X and Z can only be found implicitly in terms of the equilibrium value of X, with

$$X=\frac{(\rho+\mu)\omega_-\Gamma(S+\nu_-\epsilon_-I)((\tau+\alpha-\alpha\delta)\omega_+\Gamma(S+\nu_+\epsilon_+I)+\alpha\delta\Gamma(S+\nu_+I))}{\gamma\eta\nu_-S}$$

and

$$Z=\frac{(\rho+\mu)\omega_+\Gamma(S+\nu_+\epsilon_+I)I}{\gamma S}$$

Extensive algebraic manipulations eventually indicate that the solutions for I are given by roots to the quartic

$$a_4I^4+a_3I^3+a_2I^2+a_1I+a_0=0$$

in which the coefficients are given by

$$a_0=\sigma P_0T_0-\alpha\zeta Q_0N$$

$$a_1=\sigma(P_0T_1+P_1T_0)+\alpha\zeta(Q_0L-Q_1N)$$

$$a_2=\sigma(P_0T_2+P_1T_1+P_2T_0)+\alpha\zeta(Q_1L-Q_2N)$$

$$a_3 = \sigma(P_1T_2 + P_2T_1) + \alpha\zeta Q_2 L$$

$$a_4 = \sigma P_2 T_2$$

and where the coefficients are defined in terms of the quantities

$$L = 1 + \frac{\mu}{\rho}$$

$$P_0 = \zeta N - C\omega_- J_0 N$$

$$P_1 = -\zeta L - C[\omega_- J_0(\nu_-\epsilon_- - L) + \omega_+ J_1 N + \eta\nu_-\omega_+ N]$$

$$P_2 = -C[\omega_- J_1(\nu_-\epsilon_- - L) + \eta\nu_-\omega_+(\nu_+\epsilon_+ - L)]$$

$$Q_0 = (1-\delta)\omega_- N J_0 + \delta N J_0$$

$$Q_1 = (1-\delta)[\omega_- N J_1 + \omega_-(\nu_-\epsilon_- - L)J_0 + \eta\nu_-\omega_+ N] + \delta[N J_1 + (\nu_- - L)J_0 + \eta\nu_- N]$$

$$Q_2 = (1-\delta)[\omega_-(\nu_-\epsilon_- - L)J_1 + \eta\nu_-\omega_+(\nu_+\epsilon_+ - L)] + \delta[(\nu_- - L)J_1 + \eta\nu_-(\nu_+ - L)]$$

$$T_0 = \omega_- N J_0$$

$$T_1 = \omega_- N J_1 + \omega_-(\nu_-\epsilon_- - L)J_0 + \eta\nu_-\omega_+ N + (\beta - 1)\nu_-\omega_-\epsilon_- J_0$$

$$T_2 = \omega_-(\nu_-\epsilon_- - L)J_1 + \eta\nu_-\omega_+(\nu_+\epsilon_+ - L) + (\beta - 1)[\nu_-\epsilon_-\omega_- J_1 + \eta\nu_-\nu_+\epsilon_+\omega_+]$$

in which

$$A = (\tau + \alpha - \alpha\delta)\omega_+\Gamma$$

$$B = \alpha\delta\Gamma$$

$$C = \frac{(\mu + \gamma)\Gamma}{\gamma\eta\nu_-}$$

$$J_0 = (A + B)N$$

$$J_1 = A\nu_+\epsilon_+ + B\nu_+ - (A + B)L$$

Supplementary Appendix S4. Number and stability of equilibria

We performed a randomisation scan to understand how frequently multiple biologically relevant equilibria could be obtained. We selected sets of parameters by sampling uniformly at random from the following ranges, chosen so that the full range of biologically plausible behaviours were examined.

Parameter	Minimum value	Maximum value
N	1	2,000
ρ	0.001	0.05
μ	0.001	0.2
α	0	0.6
σ	0.001	0.4
Γ	0.001	4
δ	0	1
β	0	3
τ	0	8 (NPT); 1 (PT)
ζ	1,000 (NPT); 1 (PT)	2,400 (NPT); 300 (PT)
ν_-	0.001	4
ν_+	0.001	4
ω_-	0.001	1
ω_+	0.001	1
ϵ_-	0.001	2
ϵ_+	0.001	2

In all cases, we calculated the values of the derived parameters γ and η according to the transmission type (i.e. NPT vs PT). We constrained the parameters for NPT such that $\nu_- = \nu_+$, $\omega_- = \omega_+$ and $\epsilon_- = \epsilon_+$, and – for both transmission types – immediately discarded any samples for which $\omega_\pm \epsilon_\pm > 1$.

For each of $12{,}500{,}000$ samples for each type of transmission, we then used the quartic equation in Supplementary Appendix S3 to calculate the (up to) four equilibrium values for which $I \neq 0$. Solving the equations as a polynomial was computationally convenient, since we then knew in advance how many solutions to expect (i.e. four solutions, if those involving complex numbers are counted).

For certain random sets of parameters, the quartic had coefficients of vastly differing sizes (indeed, there was at least one case in which the biggest coefficient was 17 orders of magnitude larger than the smallest coefficient). This led to potential inaccuracies in the numerical solution of the quartic polynomial (for which we used the NumPy solver "roots"). To check the accuracy of the calculated roots, potential equilibria were back substituted into right hand side of the system of differential equations. Any candidate equilibria that did not lead to values sufficiently close to zero for all four equations were used as initial estimates for a direct numerical solution of the nonlinear system. In almost all cases, this two-step process found relevant solutions to sufficient accuracy. In a small number of cases the second step of the procedure either failed to identify an equilibrium or converged on a different root. These parameter sets were discarded, before replacement with valid parameter sets, thus ensuring there were $12{,}500{,}000$ samples for each type of transmission. This happened relatively infrequently: 51,472 times for NPT and 43,133 times for PT (i.e. in <0.5% of cases).

For each of the $12{,}500{,}000$ valid independent set of parameters the remained, we characterised:

- whether the vector carrying capacity (κ) was positive or negative;
- whether the basic reproduction number (R_0) was larger or smaller than one;
- the number of biologically meaningful equilibria generated by the quartic equation;
- the number of these equilibria that were also locally stable.

We defined a biologically meaningful equilibrium to be any for which the values of all state variables were non-negative at equilibrium. Stability of equilibria was determined using standard techniques based on analytic calculation of the Jacobian of Equation (26), using numerically calculated equilibrium values in the calculation of the eigenvalues of this matrix. Stability was then tested by considering the sign of the real parts.

Counts and proportions of each type of result, for each type of transmission, are tabulated below. Note that when $\kappa < 0$, it is necessarily the case that $R_0 < 1$. Also note that when rows which are in principle possible do not appear in the table, we have not proved such cases do not exist, this is simply because no such cases were found in our – very extensive – scan (an example for NPT would be when $\kappa < 0$ and $R_0 < 1$ but with two biologically meaningful equilibria which were both locally stable; we did not find a case like this in our numerical work). Finally, note that the first % column shows percentages falling into each category as a proportion of all parameter sets, whereas the second % column shows as a proportion of parameter sets for which at least one biologically meaningful equilibrium was found (i.e. after cases in which there were no biologically meaningful equilibria were filtered out).

Results are tabulated in the tables overleaf.

We identify several cases in which there are multiple stable biologically relevant equilibria:

- **Case One (shown in blue):** bistability between a locally stable disease-free equilibrium (i.e. $R_0 < 1$) at which the vector is present (i.e. $\kappa > 0$) and a single locally stable disease-present equilibrium;
- **Case Two (shown in green):** bistability between a locally stable disease-free equilibrium (i.e. $R_0 < 1$) at which the vector is absent (i.e. $\kappa < 0$) and a single locally stable disease-present equilibrium;
- **Case Three (shown in pink):** bistability between a pair of locally stable disease-present equilibria, with the disease-free equilibrium being unstable (i.e. $R_0 > 1$)
- **Case Four (shown in grey):** tristability between a locally stable disease-free equilibrium (i.e. $R_0 < 1$) at which the vector is absent (i.e. $\kappa < 0$) and a pair of locally stable disease-present equilibria
- **Case Five (shown in gold):** tristability between a locally stable disease-free equilibrium (i.e. $R_0 < 1$) at which the vector is present (i.e. $\kappa > 0$) and a pair of locally stable disease-present equilibria

Exemplars of model behaviour in Cases One and Two are already shown in the main text (e.g. Figures 5 and 9, respectively). Examples of Cases Three to Five are detailed below

<u>Non-persistent transmission</u>

$\kappa > 0$	$R_0 > 1$	Number of biologically relevant equilibria ($I \neq 0$)	Of which how many are stable?	Number of times found in the scan	%	% (valid)
×	×	0	0	9,506,452	76.05%	-
×	×	1	0	5	<0.01%	<0.01%
×	×	1	1	17	<0.01%	<0.01%
×	×	2	0	1	<0.01%	<0.01%
×	×	2	1	486,423	3.89%	21.06%
✓	×	0	0	684,243	5.47%	-
✓	×	1	0	7	<0.01%	<0.01%
✓	×	2	0	19,775	0.16%	0.86%
✓	×	2	1	234,040	1.87%	10.13%

✓	×	3	2	1	<0.01%	<0.01%
✓	✓	0	0	2	<0.01%	-
✓	✓	1	0	59,203	0.47%	2.56%
✓	✓	1	1	1,509,831	12.08%	65.38%

Persistent transmission

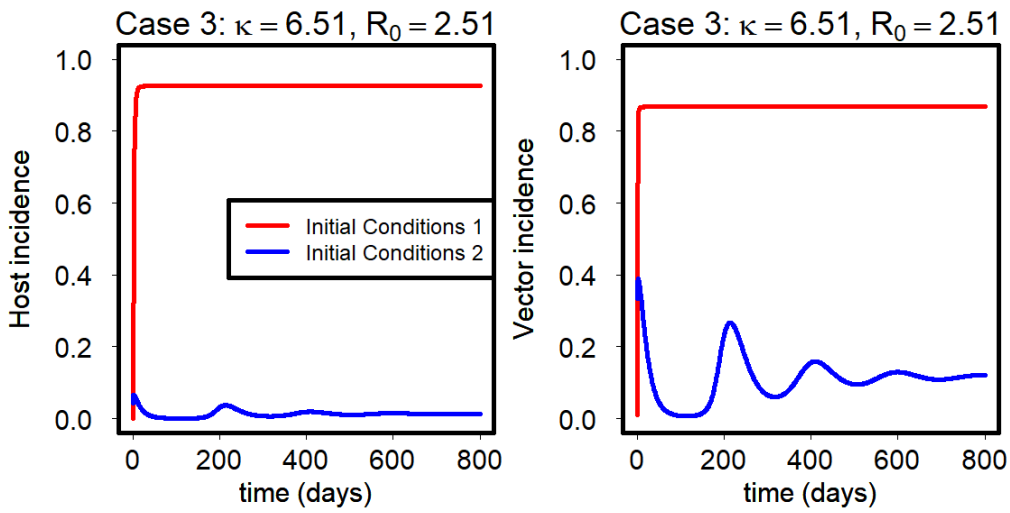
$\kappa > 0$	$R_0 > 1$	Number of biologically relevant equilibria ($I \neq 0$)	Of which how many are stable?	Number of times found in the scan	%	% (valid)
×	×	0	0	9,797,225	78.38%	-
×	×	1	0	1	<0.01%	<0.01%
×	×	1	1	5	<0.01%	<0.01%
×	×	2	0	33	<0.01%	<0.01%
×	×	2	1	196,498	1.57%	16.57%
×	×	4	2	1	<0.01%	<0.01%
✓	×	0	0	1,516,919	12.14%	-
✓	×	1	0	1	<0.01%	<0.01%
✓	×	1	1	1	<0.01%	<0.01%
✓	×	2	0	1,598	0.01%	0.13%
✓	×	2	1	127,778	1.02%	10.78%
✓	×	4	1	11	<0.01%	<0.01%
✓	×	4	2	11	<0.01%	<0.01%
✓	✓	1	0	3,806	0.03%	0.32%
✓	✓	1	1	852,607	6.82%	71.90%
✓	✓	3	1	96	<0.01%	0.01%
✓	✓	3	2	3,409	0.03%	0.29%

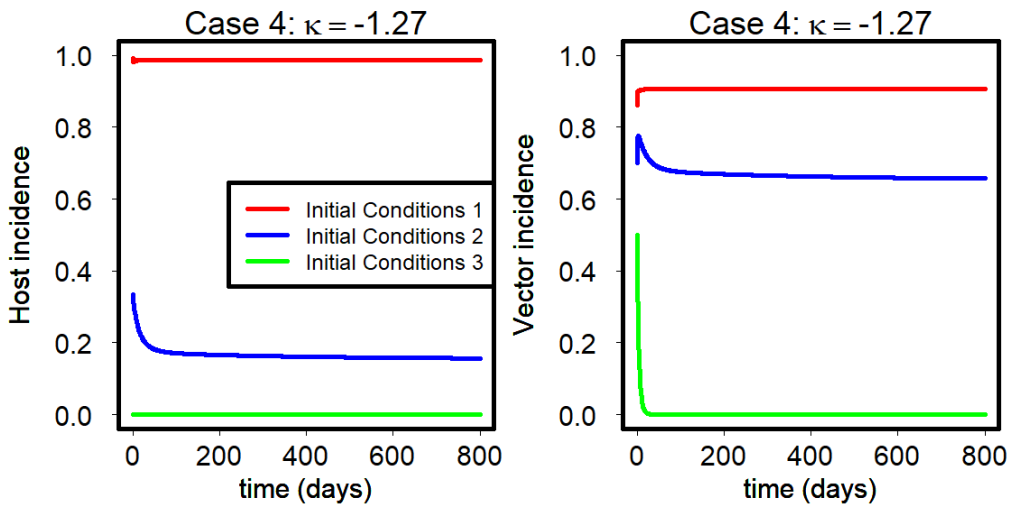
Examples of Cases Three, Four and Five

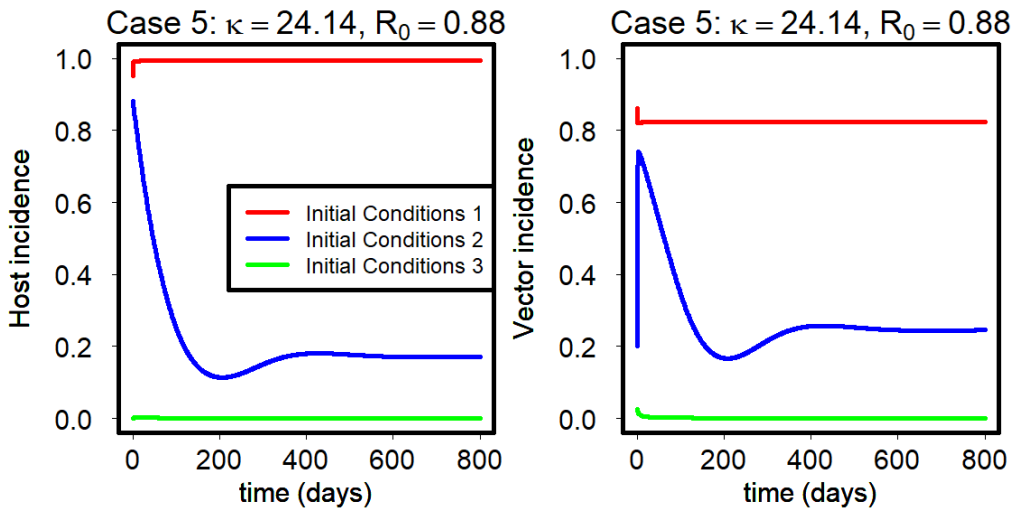
All examples are for PT and were selected at random. The parameter sets considered are below (shown to the very large number of significant figures as generated by the randomisation procedure).

Parameter	Case Three	Case Four	Case Five
N	1357.78471380456	1319.4498303372	758.422673812394
ρ	0.0203851207824684	0.0405860211146373	0.00196785718199866
μ	0.070862116192852	0.0523757168662303	0.0423162927283986
α	0.364490774736621	0.172566453141838	0.0835683097305965
σ	0.390441941465403	0.172671738699192	0.351101323665888
Γ	0.0817319970600838	0.431801010900859	1.58321791034575
δ	0.154354048987494	0.266497509538162	0.498826196145645
β	2.90534775065342	2.57881567083572	1.553702775786
τ	0.898696022660075	0.0612631271055258	0.0554984526167504
ζ	132.645389890262	145.151797209077	244.713013585564
ν_-	3.37903844254511	2.73611880606566	2.67039673107236
ν_+	1.96477227364072	0.0248878129250102	0.107380027200243
ω_-	0.892493274544967	0.966105712055771	0.151813191157394
ω_+	0.200137868270559	0.964307280496935	0.887821927785777
ϵ_-	0.352802530096919	0.474819355540881	0.144803724097339
ϵ_+	0.500524147329523	0.975518138714481	1.3507622398246

Graphs showing the range of outcomes for different initial conditions in the different cases are shown below.







The initial conditions used in the model runs shown above were as follows.

Case	**Initial Conditions**	S_0	I_0	X_0	Z_0
Three	1	1000	1	100	1
	2	8.8	0.4	0.34	0.17
Four	1	5	550	8	50
	2	600	300	3	7
	3	1300	1	0.001	0.001
Five	1	5	100	8	50
	2	20	150	4	1
	3	760	0.5	20	0.5