Theor Ecol (2014) 7:87–100
DOI 10.1007/s12080-013-0200-x

ORIGINAL PAPER

Disease in group-defending prey can benefit predators

Andrew M. Bate · Frank M. Hilker

Received: 23 May 2013 / Accepted: 22 August 2013 / Published online: 29 September 2013

Abstract Infectious diseases have the capacity not only to influence the host population but also interacting species like predators. In particular, they can reduce host densities, which can have knock-on effects on predators. Here, we consider how an infectious disease in the prey affects the predator–prey relationship where the prey exhibit some kind of group defence against the predator (using a Holling type IV functional response). We find that the disease can reduce prey densities to levels where the group defence is weaker. This weakened group defence allows predators to survive in many situations where they could not without the disease.

Keywords Eco-epidemiology · Coexistence · Bistability · Homoclinic

Introduction

Group-defending prey pose many difficulties for predators to overcome. Large groups of prey can dazzle and confuse predators, making it difficult for predators to focus on and pick out individual prey from the group. Large groups of prey have many eyes that improve vigilance, reducing the element of surprise often necessary for successful attack. On top of this, large groups of prey may even mob attack, potentially harming predators. There are many examples of group defence (see Krause and Ruxton 2002). As early as 1920, Allen suggests that a school of sardines can confuse Great Northern loons, whereas Miller (1922) suggests a flock of Bush tits has many eyes to spot hawks and will respond with a 'confusion chorus'. More recently, Japanese honeybees have been reported to mob attack foraging hornets by forming a 'hot defensive ball' around the hornet (Ono et al. 1995).

There have been several attempts to mathematically model group defence, the first being Freedman and Wolkowicz (1986). The most common and among the simplest method of incorporating group defence in a predator–prey model is by using Holling type IV functional responses (sometimes called Monod–Haldane functional responses, a term with origins in microbiology, Andrews 1968). Such functional responses behave much like a Holling type II functional response, especially for small prey densities. However, instead of saturating at large prey densities, the functional response will become negatively sloped. That is, the predation rate per predator decreases for larger prey densities as a consequence of group defence. From this, it is worth noting that Holling type IV functional responses usually result in an upper threshold of prey density, beyond which the predator cannot survive. This can be seen as a strong group defence. There are other ways of modelling group defence. For example, Ajraldi et al. (2011) and Venturino (2011) recently suggested a 'square root' functional response for predators of herding prey, particularly for the herding of large mammals. Their argument centres around the idea that predators can only attack those prey along the perimeter of a herd. Such functional responses neglect the other aspects of group defence like predator confusion, but also these functional responses grow particularly high for small prey densities (with infinite gradient at zero).

A. M. Bate (✉) · F. M. Hilker
Centre for Mathematical Biology, Department of Mathematical Sciences, University of Bath, Bath BA2 7AY, UK
e-mail: A.M.Bate@bath.ac.uk

Likewise, Geritz and Gyllenberg (2013) developed a model for group defence where predators capture only individual prey and not those in groups. These individual prey can join and leave groups. This results in a functional response that is proportional to the number of individual prey which increases monotonically (sublinearly) with total prey density.

The combination of group defence and disease has rarely been considered, either theoretically (Venturino 2011, being an exception) or empirically. However, diseases have the capacity to weaken not only the infected individuals, but also the group defences. This weakening can be simply because the disease reduces the size of the group via disease-induced mortality. However, a weaker group defence could also be the result of infected individuals not being as good at contributing to the group defence. For example, Seppälä et al. (2008) show that rainbow trout infected with eye flukes have different shoaling behaviour to those without eye flukes; and although infected and susceptible fish were not mixed, one would expect that infected fish would not co-ordinate well with susceptibles of the shoal, potentially breaking down the whole group defence. In short, there is much prospect for diseases to undermine group defence effort by prey.

Diseases and predators are competing for the same prey hosts. In many models, both the disease and predator can coexist at equilibrium. However, such coexistence between predator and disease is the result of infected prey being more vulnerable to predation than susceptible prey. In particular, equilibrial coexistence cannot occur in models where predators do not discriminate between susceptible and infected prey with respect to the predators' functional response (Siekmann et al. 2010; Hilker and Malchow 2006). This is because discriminate predation is reminiscent of intraguild predation (with susceptible prey as resource, infected prey as intraguild predator and the predator as top predator), whereas indiscriminate predation can be rescaled to exploitative competition (see Sieber and Hilker 2011), where predator and prevalence (the proportion of infected prey in the prey population) do not interact directly but both prey on the same prey host.

In ecology, it has long been established (Gause 1934) that two species competing for a common resource cannot coexist at equilibrium (called the 'principle of competitive exclusion'; Hardin 1960). In short, exploitative competition means extinction of one or more predators. There are factors that undermine this principle; for example, it does not hold if there is any direct interaction between two predators like competition. In particular, coexistence can occur if one of the predators preys on the other predator, i.e. we have intraguild predation. Another counterexample is that coexistence can occur if all populations are oscillating, e.g. due to a Holling type II functional response (McGehee and Armstrong 1977). Likewise, Chesson (2000) demonstrates that coexistence can occur if there is some spatial heterogeneity. Another, often overlooked counterexample is that one or more of the predators are restricted by some sort of density dependence (Gurney and Nisbet 1998, pp.166–167). In this case, coexistence can occur if the density-dependent predator can survive at prey levels set by the other predator.

In this paper, we find that a disease and predator can coexist on the same prey host, contradicting the principle of competitive exclusion. On top of that, if we assume that the prey exhibit some group defence, we find that the disease can benefit the predator by reducing prey densities to more manageable levels for the predator. In particular, we find two cases where an endemic disease can prevent the predator becoming extinct; one case is where the disease reduces the prey density below a critical threshold; the other is that the disease reverses a homoclinic bifurcation, bringing coexistent oscillations from what was the certain extinction of the predator.

Model derivation

In this section, we will construct two models with predators, susceptible prey and infected prey where the prey exhibit group defence. But before we proceed, we need to carefully derive appropriate functional responses.

The functional response

When modelling group defence for the prey, Holling type IV functional responses of the form equivalent to $\frac{aN}{1+bN+cN^2}$ $\left(\text{or the simplification}\frac{aN}{1+cN^2}\right)$ are often used (Freedman and Wolkowicz 1986; Ruan and Xiao 2001; Kot 2001, Chap. 9). Usually, they are used without any mechanistic derivation or justification. Such Holling type IV functional responses can be derived from a Holling type II functional response, $\frac{aN}{1+ahN}$, where a is the attack rate, h is the handling time and N is the prey density. One way of deriving a Holling type IV is by assuming that the attack rate a decreases with respect to N inverse-quadratically, i.e. $a(N) = \frac{a_0}{1+bN^2}$ (Koen-Alonso 2007). Another derivation assumes that the handling time is linearly increasing with respect to N, i.e. $h(N) = h_0 + h_N N$. (There have been a few other attempts to derive a Holling Type IV, for example, Collings (1997) derives it by assuming both a linearly increasing handling time and an inverse-linear attack rate, which is not a simple argument.) The second derivation based on linear handling times will be used here, largely because it is a simpler argument. The handling time formulation is apt if we assume that time taken to attack and catch a prey increases linearly

with respect to prey density. This increased handling time can be considered due to group defence and the additional time it takes to separate and subdue prey at higher prey densities. The time to eat and digest prey is still independent of prey density.

This single-prey Holling type IV functional response does not take into account that the prey is structured because of an infectious disease. We need to derive a two-prey Holling type IV functional response where the two classes of prey are susceptible, S, and infected, I. This can be done by considering the following two-prey Holling type II functional response for susceptible prey, derived using a standard time-management argument (Holling 1959; Murdoch 1972):

$$f_{\mathrm{S}}(S,I)=\frac{a_{\mathrm{S}}S}{1+a_{\mathrm{S}}h_{\mathrm{S}}S+a_{\mathrm{I}}h_{\mathrm{I}}I}.$$

Here, a_{S} and a_{I} are the attack rates on the susceptible and infected preys, respectively. Likewise, h_{S} and h_{I} are the handling times on the susceptible and infected preys, respectively. The infected prey have an equivalent $f_{\mathrm{I}}(S,I)$, which has the numerator $a_{\mathrm{I}}I$.

Now, we can assume, like with the one-prey case, that the handling times are density dependent. Thus, we have $h_{\mathrm{S}}(S,I)=h_{\mathrm{S0}}+h_{\mathrm{SS}}S+h_{\mathrm{SI}}I$ and $h_{\mathrm{I}}(S,I)=h_{\mathrm{I0}}+h_{\mathrm{IS}}S+h_{\mathrm{II}}I$, where h_{S0} and h_{I0} are the density-independent handling times of the susceptible and infected prey, respectively; h_{SS} and h_{IS} are the density-dependent (with respect to susceptible prey) handling times of susceptible and infected prey, respectively; whereas h_{SI} and h_{II} are the density-dependent (with respect to infected prey) handling times of susceptible and infected prey, respectively. These formulations take into account that although infected and susceptible prey are seen as different classes of prey, they contribute to the same group defence. In general, all these parameters can be different. For example, imagine a diseased fish that cannot follow the rest of the school, potentially leading to ineffective school movement and compromised group defence, or a diseased meerkat that is not as capable at spotting threats when acting as sentry for the clan, leaving the clan at greater risk. Both of these examples suggest that $h_{\mathrm{SS}}\neq h_{\mathrm{SI}}$. Likewise, infected prey can be easier to catch, subdue and eaten by predator once spotted, suggesting that $h_{\mathrm{SS}}\neq h_{\mathrm{IS}}$ and $h_{\mathrm{SI}}\neq h_{\mathrm{II}}$.

By incorporating these density-dependent handling times, we get the following two-prey Holling type IV functional response for the susceptible prey:

$$f_{\mathrm{S}}(S,I)=\frac{a_{\mathrm{S}}S}{1+a_{\mathrm{S}}h_{\mathrm{S0}}S+a_{\mathrm{I}}h_{\mathrm{I0}}I+a_{\mathrm{S}}h_{\mathrm{SS}}S^2+(a_{\mathrm{S}}h_{\mathrm{SI}}+a_{\mathrm{I}}h_{\mathrm{IS}})SI+a_{\mathrm{I}}h_{\mathrm{II}}I^2}.$$

Likewise, the functional response for the infected prey is:

$$f_{\mathrm{I}}(S,I)=\frac{a_{\mathrm{I}}I}{1+a_{\mathrm{S}}h_{\mathrm{S0}}S+a_{\mathrm{I}}h_{\mathrm{I0}}I+a_{\mathrm{S}}h_{\mathrm{SS}}S^2+(a_{\mathrm{S}}h_{\mathrm{SI}}+a_{\mathrm{I}}h_{\mathrm{IS}})SI+a_{\mathrm{I}}h_{\mathrm{II}}I^2}.$$

If susceptible and infected prey do not contribute to the same group defence, but instead contribute to their own group defence, then we would have that $h_{\mathrm{SI}}=0$ and $h_{\mathrm{IS}}=0$. In this case, we would have a functional response comparable to that of two distinct species under a common predator, both with their own group defence.

Other model assumptions

We consider an SI disease in the prey where disease transmission is either frequency dependent ($\beta(S,I)=\frac{\beta SI}{S+I}$) or density dependent ($\beta(S,I)=\beta SI$), where β is the transmissibility coefficient. All prey are born susceptible, i.e. there is no vertical transmission. We assume (for now at least) that infected preys have different fertility, increased density-independent mortality and different strengths of competition when compared to susceptible prey. Additionally, predators grow linearly with respect to the predation and die at a constant per capita rate.

$$\frac{dS}{dt}=b_{\mathrm{S}}S+b_{\mathrm{I}}I-mS-c_{\mathrm{SS}}S^2-c_{\mathrm{SI}}SI-f_{\mathrm{S}}(S,I)P-\beta(S,I), \quad (1)$$

$$\frac{dI}{dt}=\beta(S,I)-(m+\mu)I-c_{\mathrm{IS}}IS-c_{\mathrm{II}}I^2-f_{\mathrm{I}}(S,I)P, \quad (2)$$

$$\frac{dP}{dt}=(\gamma_{\mathrm{S}}f_{\mathrm{S}}(S,I)+\gamma_{\mathrm{I}}f_{\mathrm{I}}(S,I)-d)P. \quad (3)$$

Here, b_{S} and b_{I} are the per capita birth rates and γ_{S} and γ_{I} are conversion efficiencies from consuming susceptible and infected prey, respectively. c_{SS} and c_{SI} represent density-dependent mortalities that susceptible preys experience when encountering other susceptible and infected prey, respectively. Likewise, c_{IS} and c_{II} represent density-dependent mortalities that infected prey experience when encountering other susceptible and infected prey, respectively. Together, c_{SS}, c_{SI}, c_{IS} and c_{II} represent

intra/interspecific competition. m is the natural per capita (density independent) death rate of the prey, μ is the disease-induced per capita death rate of the prey and d is the per capita death rate of the predator and μ is the disease-induced per capita death rate.

Simplified model

The full model (1–3) is rather complex, with 20 parameters in a three-dimensional system. To mitigate this, we simplify the model as much as possible as a starting point. We can always, in the future, consider more complicated versions once the simpler model is fully understood.

The simplifying assumptions are as follows: $b_S = b_I(:= b)$, $c_{SS} = c_{SI} = c_{IS} = c_{II}(:= c)$, $\gamma_S = \gamma_I(:= \gamma)$, $a_S = a_I(:= a)$, $h_{S0} = h_{I0}(:= h_0)$ and $h_{SS} = h_{SI} = h_{IS} = h_{II}(:= h_N)$. These assumptions essentially can be summarised by saying that infected and susceptible prey only differ by additional mortality for infected prey ($\mu > 0$); that susceptible and infected prey have the same birth rates, are equally good competitors and have equal attack rates, handling times and conversion. By implementing these assumptions, we cannot only gather terms but also collapse the functional responses to a single-prey form:

$$\frac{dS}{dt} = b(S+I) - mS - cS(S+I) - \frac{aSP}{1+ah_0(S+I)+ah_N(S+I)^2} - \beta(S,I), \quad (4)$$

$$\frac{dI}{dt} = \beta(S,I) - (m+\mu)I - cI(S+I) - \frac{aIP}{1+ah_0(S+I)+ah_N(S+I)^2}, \quad (5)$$

$$\frac{dP}{dt} = P\left(\frac{\gamma a(S+I)}{1+ah_0(S+I)+ah_N(S+I)^2} - d\right). \quad (6)$$

Working with total prey $N = S+I$ instead of susceptible prey and prevalence $i = \frac{I}{N}$, i.e. the proportion of infected prey in the prey population, instead of infected prey:

$$\frac{dN}{dt} = (b-m)N - \mu iN - cN^2 - \frac{aNP}{1+ah_0N+ah_NN^2}, \quad (7)$$

$$\frac{di}{dt} = i\left(\left(\frac{\beta(N,i)}{Ni(1-i)} - \mu\right)(1-i) - b\right), \quad (8)$$

$$\frac{dP}{dt} = P\left(\frac{\gamma aN}{1+ah_0N+ah_NN^2} - d\right). \quad (9)$$

For frequency-dependent transmission, $\beta(N,i) = \beta Ni(1-i)$, whereas for density-dependent transmission, $\beta(N,i) = \beta N^2 i(1-i)$.

To reduce the number of parameters further, we non-dimensionalise the system. Let us rescale time such that the predator's death rate becomes one ($t = \frac{1}{d}T$). Predator density is rescaled such that the numerator of the functional response becomes one ($P = \frac{d}{a}y$). Prey density is rescaled such that the numerator of the predator's numerical response is scaled to one ($N = \frac{d}{\gamma a}x$). Then, for frequency-dependent transmission, Eqs. 7–9 become:

$$\frac{dx}{dT} = (b'-m')x - \mu' ix - c'x^2 - \frac{xy}{1+H_0x+H_xx^2}, \quad (10)$$

$$\frac{di}{dT} = i((\beta'-\mu')(1-i) - b'), \quad (11)$$

$$\frac{dy}{dT} = y\left(\frac{x}{1+H_0x+H_xx^2} - 1\right). \quad (12)$$

The new parameters are the scaled prey birth $\left(b' = \frac{b}{d}\right)$ and death $\left(m' = \frac{m}{d}\right)$ rates, scaled disease-induced death rate $\left(\mu' = \frac{\mu}{d}\right)$, scaled density-dependent mortality $\left(c' = \frac{c}{\gamma a}\right)$, scaled transmissibility $\left(\beta' = \frac{\beta}{d}\right)$ and the scaled density-independent $\left(H_0 = \frac{h_0 d}{\gamma}\right)$ and density-dependent $\left(H_x = \frac{h_N d^2}{a\gamma^2}\right)$ handling time.

For density-dependent transmission, the only difference from Eqs. (10–12) is that the prevalence Eq. (11) becomes:

$$\frac{di}{dt} = i((\beta' x - \mu')(1-i) - b'), \quad (13)$$

where $\beta' = \frac{\beta}{\gamma a}$.

For simplicity of notation, we will drop the dashes. From now on, we will only work with the non-dimensionalised parameters, so there should be no confusion of notation.

These models are comparable with existing models; in particular, setting $m = 0$ and $H_x = 0$, we obtain the diseased prey model in Bate and Hilker (2013). Also, with this scaling, we have reduced the model from an intraguild predation model to something resembling exploitative competition, as there is no direct interaction between predators and disease prevalence (cf Sieber and Hilker 2011). In fact, for density-dependent transmission, the model is exploitative competition.

Now, for the frequency-dependent model, by defining functions $f(x) = \frac{x}{h(x)}$ (functional response), $g(x) = b - m - cx$ (per capita growth rate of prey in absence of predators and disease), $h(x) = 1 + H_0x + H_xx^2$ (the denominator of the functional response, or in other words, the total time predators spend searching and handling prey relative to search time) and $p(i) = (\beta - \mu)(1-i) - b$ (per capita growth in prevalence), we get:

$$\frac{dx}{dT} = f(x)[(g(x) - \mu i)h(x) - y], \quad (14)$$

$$\frac{di}{dT} = i\, p(i), \quad (15)$$

$$\frac{dy}{dT} = y(f(x) - 1). \quad (16)$$

For the density-dependent model, $p(i)$ becomes:

$$p(x, i) = (\beta x - \mu)(1 - i) - b. \quad (17)$$

With such functions, we can use analysis similar to that in Kot (2001, Chap. 9) to establish the existence and stability of steady states with relatively clear notation. In the rest of the paper, we will always assume the prey can grow in the absence of predator and disease, i.e. $g(0) > 0$ (equivalently, $b > m$).

Disease-free predator–prey dynamics

Ignoring the disease, the predator–prey model is equivalent to the model in Freedman and Wolkowicz (1986) and Kot 2001, Chap. 9). Since these are existing results, we will summarise and classify them into various Scenarios here. However, for completeness, some of the steady state and nullcline analysis is explained in the Appendix.

There are three different main Scenarios that can be derived from the steady states:

- Scenario 1: There is no coexistent steady state. The prey-only steady state is stable. This can be split into (1A) no real solutions or (1B) only negative solutions for the coexistent steady states. A phase plane of Scenario 1B (top left of Fig. 1) has two vertical predator nullclines that do not intercept the humped prey nullcline in the positive quadrant.
- Scenario 2: One coexistent steady state exists. It is either (2A) stable or (2B) unstable and is the centre of some stable limit cycle. This depends on the slope of the prey nullcline, which is given by the sign of $\frac{\partial y}{\partial x}(x^*) := y'(x)$. Phase planes of Scenarios 2A and 2B show that one of the predator nullclines intercepts the humped prey nullcline in the positive quadrant, resulting in one predator–prey equilibrium and an unstable prey-only steady state. If the interception occurs while the prey nullcline is negatively sloped (i.e. to the right of the maximum in the prey nullcline), the predator–prey equilibrium is stable (Scenario 2A (top middle of Fig. 1)); whereas, if the interception occurs while the prey nullcline is positively sloped (i.e. to the left of the maximum in the prey nullcline), the predator–prey equilibrium is unstable and there is a stable predator–prey limit cycle (Scenario 2B (top right of Fig. 1)).
- Scenario 3: Two coexistent steady states exist. The coexistent steady state with the lower prey density is either (3A) stable or (3B) unstable and is the centre of some limit cycle. Again, this depends on the slope of the prey nullcline, which is given by the sign of $y'(x)$. The stable steady state/limit cycle is bistable with the prey-only steady state, where the higher prey density coexistent steady state forms part of a separatrix. In the phase planes of Scenarios 3A and 3B (Fig. 1 bottom left and middle, respectively), both predator nullclines intercept the humped prey nullcline, resulting in two predator–prey steady states. The prey-only steady state is stable and the 'right' coexistent steady state (i.e. the coexistent steady state with the larger prey density) is always unstable (saddle point). The difference between Scenarios 3A and 3B is the same as the difference between Scenarios 2A and 2B; the stability of the 'left' coexistent steady state (i.e. the coexistent steady state with the smaller prey density) and the existence of a limit cycle depend on where the interception is relative to the maximum of the prey nullcline.

Scenarios 1 and 2 can be said to be the cases where group defence is weak since these Scenarios are also possible for a Holling type II functional response (i.e. the Rosenzweig–MacArthur model). In Scenario 3, group defence is strong enough to dominate dynamics for larger prey densities. This is expressed by the stability of the prey-only equilibrium and the bistability, which is not possible in the Rosenzweig–MacArthur model. Note that the dynamics associated with Holling Type II functional responses are still dominant for smaller prey densities.

This list does not give all the information; there is also a global bifurcation. Freedman and Wolkowicz (1986) and Kot (2001, Chap. 9) demonstrate that the limit cycle in Scenario 3B can collide with the saddle point to form a homoclinic orbit. Beyond this homoclinic bifurcation, the limit cycle disappears and the prey-only steady state is the only stable steady state, like Scenario 1. Consequently, we have another Scenario:

- Scenario 4: Two coexistent steady states exist, neither are stable. No limit cycle exists due to a homoclinic bifurcation. Only the prey-only steady state is stable. In the phase plane of Scenarios 4 (Fig. 1 bottom right), both predator nullclines intercept the humped prey nullcline, resulting in two predator–prey steady states.

Scenario 4 means that there is no stable coexistence. There may be, however, coexistent oscillatory transients dynamics near the homoclinic bifurcation for some initial conditions, meaning that the eventual extinction of the predator would not be apparent in short to medium time scales. Figure 2 demonstrates this homoclinic bifurcation with a phase plane 'before' (left) and 'after' (right) the homoclinic bifurcation. In the left panel, we are in Scenario 3B, with the stable coexistent limit cycle and saddle point are very close. The right panel is in Scenario 4, where the limit cycle has disappeared after colliding with the saddle point, leaving the prey-only steady state as the only stable attractor, despite there being two coexistent steady states.

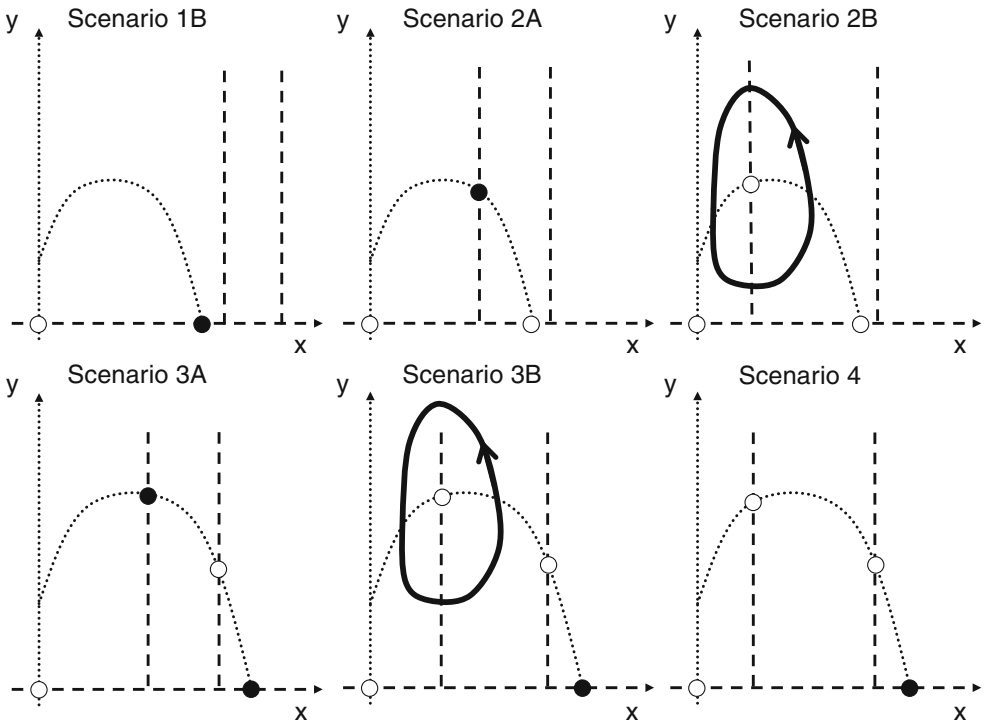

Fig. 1 Sketched phase planes of different Scenarios from the disease-free predator–prey model with group defence. These Scenarios are: a stable prey-only equilibrium with no coexistent equilibrium (Scenario 1B, *top left*); one stable coexistent equilibrium (Scenario 2A, *top middle*); one unstable coexistent equilibrium surrounded by a stable coexistent limit cycle (Scenario 2B, *top right*); bistability between a coexistent equilibrium and prey-only equilibrium (Scenario 3A, *bottom left*); bistability between a coexistent limit cycle surrounding an unstable coexistent equilibrium and a prey-only equilibrium (Scenario 3B, *bottom middle*); and finally a stable prey-only equilibrium with two unstable equilibria and no limit cycle (Scenario 4, *bottom right*). The *dashed lines* represent predator nullclines, the *dotted lines* represent prey nullclines, the *white circles* represent unstable steady states, the *black circles* represent stable steady states and the *loop* represents a stable limit cycle

Scenario 4 essentially means that the usual predator–prey oscillations from the Rosenzweig–MacArthur model cannot be fully contained in the region where prey densities are small enough for group defence to be weak, and instead encroaches into regions where group defence dominates.

Results: frequency-dependent transmission

In the previous section, we set the scene by describing the predator–prey model in the absence of infection. Now, we can incorporate a disease in the prey population. In this section, we will analyse the frequency-dependent model (14–16), and we will then tackle the more complicated case of density-dependent transmission in the next section.

Coexistence between disease and predator

Observing that the prevalence equation (15) is completely independent from both the predator and prey (since $p(i) = (\beta-\mu)(1-i)-b$), we can separate the prevalence equation. From the prevalence equation, we have that the disease-free state ($i^* = 0$) is stable if $p(0) = \beta - \mu - b < 0$. Otherwise, if $p(0) > 0$, the disease-free state is unstable, the disease will be endemic and disease prevalence will approach $i^* = 1 - \frac{b}{\beta-\mu}$.

For the remainder of this section, we will assume that the prevalence is at the equilibrium $i^* = 1 - \frac{b}{\beta-\mu}$. Armed with this quasi-stationary assumption, we can treat prevalence as a constant, reducing the frequency-dependent model (14–16) to the following 2D model:

$$\frac{dx}{dT} = f(x)[(g(x) - \mu i^*)h(x) - y], \tag{18}$$

$$\frac{dy}{dT} = y(f(x) - 1). \tag{19}$$

This model is the same as the disease-free predator–prey model, except that there is an additional disease-induced mortality in the prey. This additional term only alters the 'humped' non-trivial prey nullcline defined by $y(x) = (g(x) - \mu i^*)h(x)$.

Figure 3 demonstrates how this nullcline is changed. Increasing prevalence alters two key points of the humped

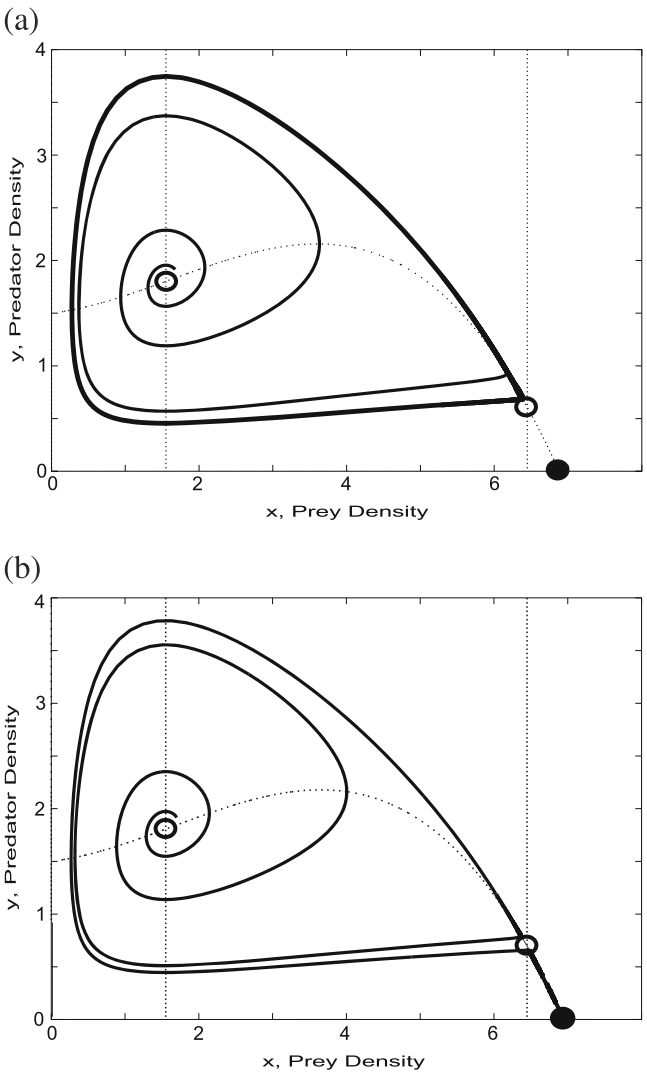

Fig. 2 Phase planes demonstrating the existence of a homoclinic bifurcation and the resulting destruction of the stable limit cycle in the disease-free model. **a** is a phase plane with bistability between a stable predator–prey limit cycle and a prey-only equilibrium (Scenario 3B), where the stable predator–prey limit cycle is close to the predator–prey saddle point ($c = 0.218$); whereas **b** is a phase plane with no stable limit cycle after a homoclinic bifurcation ($c = 0.216$). Here, all trajectories eventually approach the prey-only steady state despite there being two coexistent steady states (Scenario 4). The *dashed lines* represent nullclines. Other parameters: $H_0 = 0.2$, $H_x = 0.1$, $b = 2$, $m = 0.5$

nullcline: (1) the intercept with the horizontal axis (the prey-only steady state) is moved left, i.e. prevalence reduces the prey-only steady state, and (2) the maximum of the nullcline $y(x)$ is moved left, i.e. occurs at lower prey densities.

Loss of stability of the prey-only steady state

As prevalence increases, prey density at the prey-only steady state is reduced. This reduction in prey carrying capacity by the disease can become beneficial for the predator as it can shift the predator–prey system from Scenario 3 to Scenario 2, like in Fig. 3. This shift is important since Scenario 3 means bistability involving a prey-only steady state, whereas Scenario 2 means the predator will always survive. In this case, the disease can help the predator survive under conditions where it cannot survive without the disease due to unmanageable prey densities. The presence of the disease does reduce predator density at the stable coexistent equilibrium, though, but the loss of extinction risk at high prey densities is significant (i.e. the disease can render group defence ineffective).

Stabilisation of limit cycles

The shift of the maximum of the nullcline $y(x)$ to the left reduces or eliminates limit cycles (Fig. 3). In the disease-free predator–prey system, limit cycles only occur if the maximum of the nullcline $y(x)$ is to the right (i.e. at a higher prey density) of the coexistent steady state with the lower prey density. By shifting this maximum beyond the lower steady state, the limit cycle is eliminated and a stable steady state is formed. Thus, we have that Scenario 2B/3B becomes Scenario 2A/3A. This means that increasing prevalence should take Scenario 3B to Scenario 2A via either Scenario 3A or via Scenario 2B.

Disease reversing global bifurcation

As we previously stated, there are significant parameter regions in the predator–prey model where the prey-only steady state is the only attractor despite the existence of two coexistent steady states (Scenario 4). In these regions, the predator cannot survive in the long run, independent of the initial condition. However, the presence of a disease infecting the prey can reverse this homoclinic bifurcation and give rise to a stable predator–prey–disease limit cycle. This means that the disease can facilitate coexistence where it was impossible without the disease.

Overall pattern

Figure 4 demonstrates this reversal of a homoclinic bifurcation. In the absence of the disease ($i^* = 0$), the predator cannot survive, despite there being two predator–prey steady states. As the prevalence increases, the prey-only steady state decreases. If disease-induced mortality is sufficiently high, the disease can bring the prey steady state close to the predator–prey saddle point. At the same time, increased prevalence will reduce the slope of the prey nullcline and shift the maximum to the left. Together, with

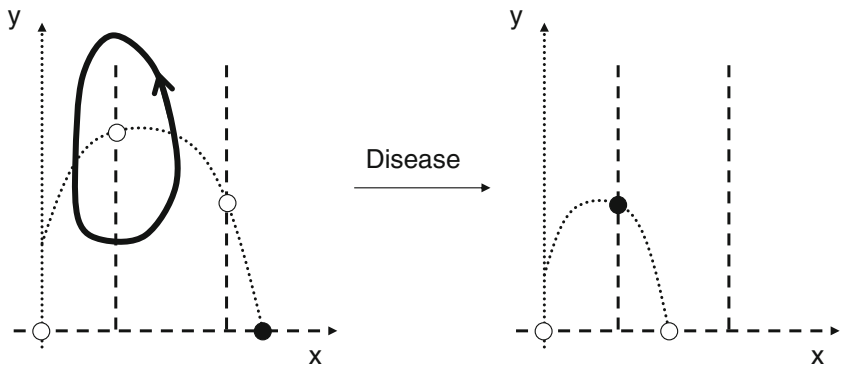

Fig. 3 Impact of disease on group defence in the frequency-dependent model: sketch of the predator–prey phase plane with nullclines and equilibria where x is prey density and y is predator density. *Left-hand figure* is without disease. Here, there is bistability between the prey-only equilibrium and a predator–prey oscillation, where the predator cannot survive 'beyond' the separatrix saddle–point (unstable) equilibrium (Scenario 3B). Including the disease has no effect on the predator nullclines, but it 'lowers' the prey nullcline and moves the maximum to the left and down (*right-hand figure*). These changes stabilise the predator–prey oscillations and result in the prey-only steady state losing stability. Consequently, with the disease, we have a stable predator–prey equilibrium (Scenario 2A). The *lines* and *circles* have the same meaning as Fig. 1

sufficiently large prevalence, a stable limit cycle will appear as the homoclinic bifurcation is reversed. We have suddenly moved from Scenario 4 to 3B. In this region, the predator can survive with the right initial condition. However, if we increase prevalence further, the prey-only steady state will lose stability as it collides with the predator–prey saddle point in a transcritical bifurcation (like in Fig. 3). After this transcritical bifurcation, we will move to Scenario 2B where the predator will survive no matter what the initial condition. The next transition occurs when the predator–prey limit cycle is stabilised by a Hopf bifurcation (like in Fig. 3), leading to Scenario 2A. Increasing prevalence further ($i^* > 0.6$), the predator–prey steady state will collide with the prey-only steady state in a transcritical bifurcation, resulting in the loss of the predator–prey steady state and a stable prey-only steady state (Scenario 1). And finally, if prevalence (and disease-induced mortality, μ) is sufficiently high ($i^* > 0.75$), the disease can wipe out the prey population (i.e. if $b < m + \mu$). This host extinction is a trademark of frequency-dependent diseases and cannot happen in density-dependent diseases (see next section).

Summary

For a frequency-dependent disease, the prevalence equation is independent of prey or predator densities. Consequently, the prevalence can be assumed to be fixed. With this in mind, we find that the disease can coexist with the predator (Scenarios 2 and 3). On top of this, the disease can help the predator by (a) keeping prey densities below densities where prey group defence is strong; (b) stabilising predator–prey cycles (preventing large booms and busts of predator and prey populations); and (c) reversing the homoclinic bifurcation, thus preventing the eventual extinction of the predator. In particular, Fig. 4 demonstrates that with increasing prevalence, we can go from a prey-only steady state (Scenario 4) to bistability between the prey-only steady state and a predator–prey limit cycle (Scenario 3B) to a predator–prey limit cycle (Scenario 2B) to a predator–prey steady state (Scenario 2A) to prey-only steady state (Scenario 1) to diseased-induced extinction of the prey (and predator, of course).

Results: density-dependent transmission

Unlike in the frequency-dependent model, we cannot separate the disease from the predator–prey dynamics in the density-dependent model (14), (16) and (17). This means that 2D phase plane analysis used in the disease-free and frequency-dependent models cannot give the whole story. In particular, it does not provide much insight into the existence of more complex dynamics like chaos and quasi-periodic dynamics. However, such phase plane analysis is still very enlightening as a similar pattern of progressing from Scenario 4 to Scenario 1 occurs.

Firstly, both oscillatory and equilibrial coexistence between predator and disease prevalence also occur in the density-dependent model. This coexistence is more interesting and complex than in the frequency-dependent model as the predator–prevalence–prey equations are in the form of exploitative competition; thus this coexistence contradicts the principle of competitive exclusion.

The coexistence is facilitated by the mixture of density-dependent terms (i.e. the '$1 - i$' terms) and density-independent terms (in this case, 'b') in the per capita growth rate for prevalence $p(x, i)$. This means that the prevalence

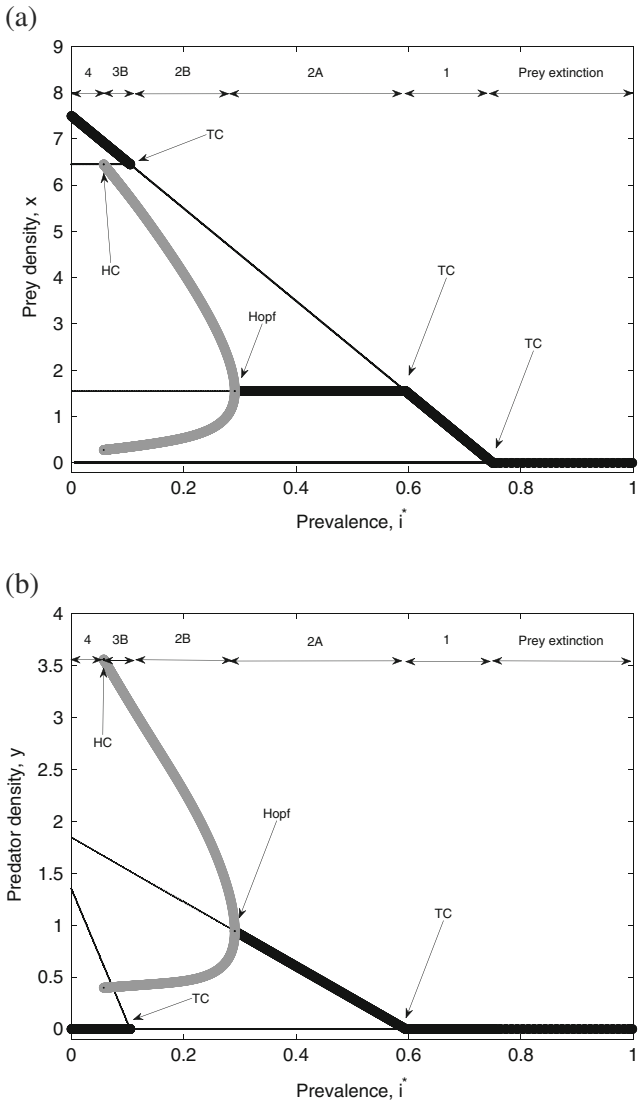

Fig. 4 Frequency-dependent model: Bifurcation diagrams of **a** prey density and **b** predator density, with respect to prevalence equilibrium i^*, showing the progression of Scenarios as prevalence increases. As prevalence is assumed to be static, we can treat it as a control parameter. In the absence of disease ($i^* = 0$), only the prey-only steady state is stable but two predator–prey steady states exist (Scenario 4). However, as we increase prevalence, we go from Scenario 4 to Scenario 3B (bistability between predator–prey oscillations and prey-only steady state) to Scenario 2B (only the predator–prey oscillations are stable) to Scenario 2A (only the predator–prey steady state is stable) to Scenario 1 (only the prey-only steady state is stable) to prey extinction. *Thick black lines* represent stable equilibria, *thick grey lines* represent stable oscillations and *thin black lines* represent unstable equilibria. '*TC*', '*HC*' and '*Hopf*' stand for transcritical, homoclinic and Hopf bifurcation, respectively. Other parameters: $H_0 = 0.2$, $H_x = 0.1$, $b = 2$, $m = 0.5$, $c = 0.2$ $\mu = 2$. Figures produced using MATLAB, using data from continuation software XPPAUT

nullplane (the values of (x, p, i) such that $p(x, i) = 0$) is not fixed to a particular value of prey density but instead exists for a range of prey densities. Since the predator nullplanes have fixed prey densities, if one or more of these

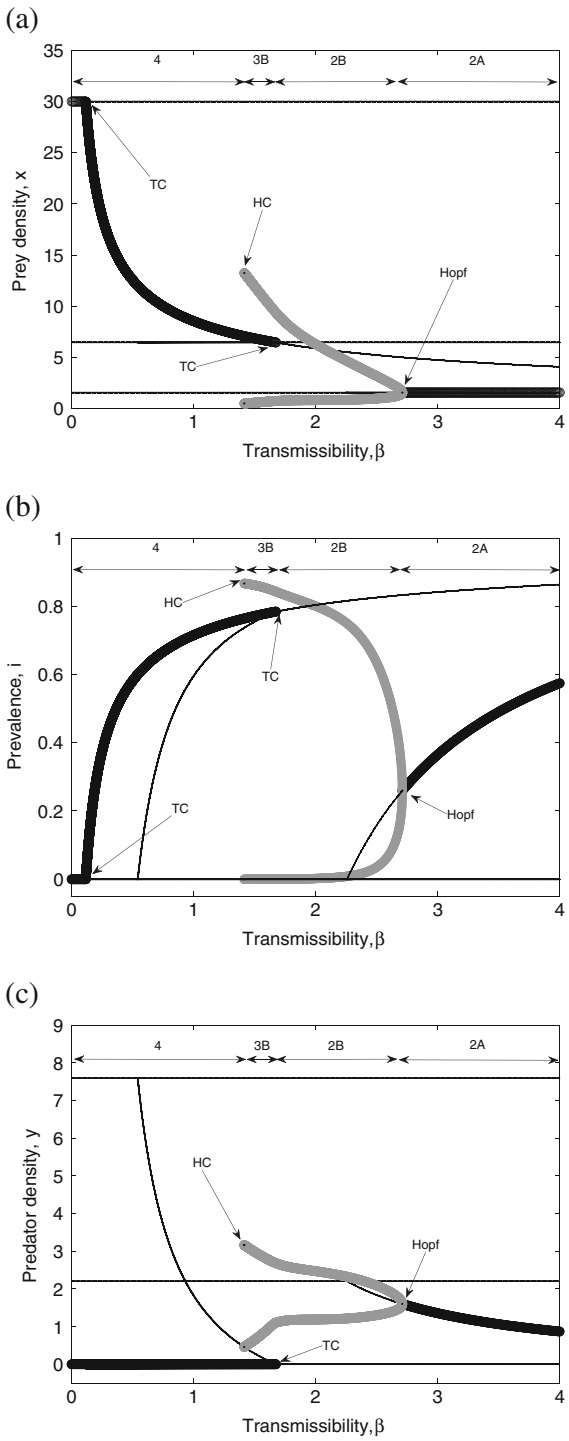

Fig. 5 Density-dependent model: Bifurcation diagrams of **a** prey density, **b** prevalence and **c** predator density, with respect to transmissibility β. Together they show the progression from a stable prey-only (or prey–disease) steady state with two other predator–prey (or predator–prey–disease steady states) (Scenario 4); to bistability between a coexistent limit cycle and prey–disease equilibrium (Scenario 3B); to a coexistent limit cycle (Scenario 2B); to a coexistent steady state (Scenario 2A). The stable limit cycle numerically breaks down at 'HC.' The *labels* and *lines* have the same meaning as Fig. 4. The trivial (no prey) steady state has been omitted. Parameter values $\mu = 1.5$, $c = 0.05$, $b = 2$, $m = 0.5$, $H_0 = 0.2$ and $H_x = 0.1$

prey densities lie within the range of prey densities for the prevalence nullplane, coexistence will occur (subject to positive predator densities and prevalence).

Secondly, the same Scenarios and transitions occur in the density-dependent model as in the frequency-dependent model. For example, Fig. 5 demonstrates that increasing transmissibility (as a proxy for prevalence and thus Fig. 5 is equivalent to Fig. 4) goes through the same transitions, from Scenario 4 to Scenario 3B to Scenario 2B to Scenario 2A, as Fig. 4 (except for Scenario 1, which occurs for levels of transmissibility well beyond the range of Fig. 5, and disease-induced extinction, which cannot happen in the density-dependent model).

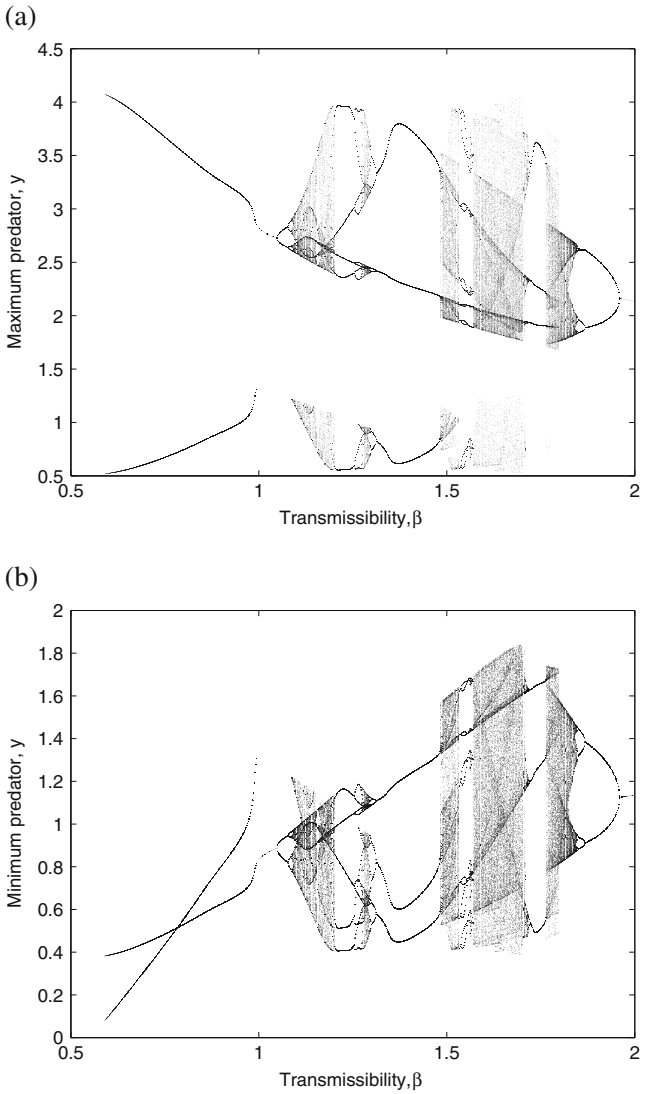

Fig. 6 Complex dynamics in the density-dependent model: Bifurcation diagrams of **a** (local) maximum predator density and **b** (local) minimum predator density, with respect to transmissibility. For $\beta \lesssim 0.6$, we are in Scenario 4 and the predator cannot survive. At $\beta \approx 0.6$, a two-cycle (with respect to the predator) appears (i.e. there are two local maxima and minima). At around $\beta = 1$, one of the local maxima collides with a local minimum, resulting in the loss of both. Soon afterwards, a period doubling cascade occurs, resulting in chaos. After this, the second branch of maxima and minima reappears, but this time as a chaotic attractor. In the interval $\beta \in (1, 2)$, a series of attractor crises occur. Parameter values: $\mu = 1.5$, $c = 0.2$, $b = 2$, $m = 0.5$, $H_0 = 0.2$ and $H_x = 0.1$. Using the initial condition $(x, y, i) = (0.5, 0.5, 0.1)$, we find the numerical solution (by using MATLAB's 'ode45' and the log-transform of equations (14–16), subject to Eq. 17, to avoid numerical errors around zero) for time up to $T = 7,000$ and then discard transients (all data up to $T = 4,000$)

One novelty is that prevalence does not always increase with transmissibility (Fig. 5c). In particular, the loss of stability for the disease–prey steady state at the transition between Scenarios 3B and 2B results in massive reduction of prevalence (although the predator–prey–disease limit cycle will have short periods where prevalence is higher than the disease–prey steady state).

Lastly, complex dynamics can occur. In the 2D predator–prey and frequency-dependent models, the possible stable dynamics are limit cycles and equilibria only. In 3D systems like the density-dependent model, many more phenomena can be found within regions of Scenarios 2B and 3B. An example of such complex dynamics is Fig. 6.

In Fig. 6, there are several complex dynamics. After the reversal of the homoclinic orbit (at approximately $\beta = 0.6$), the species coexist on a 'two-cycle' (note that both predators and prey exhibit two local maxima and minima each, whereas the prevalence, not shown here, exhibits only one local maximum and minimum each). At approximately $\beta = 1$, one branch of the attractor suddenly disappears as one of the maxima collides with one of the minima. Note that this branch emerges again in form of a chaotic attractor, as the remaining branch has undergone a cascade of period doubling bifurcations. At around $\beta = 1.8$, the system stabilises via a period halving cascade. But for parameter values in between, the bifurcation diagram displays a number of different attractor crises, in which branches of the attractor merge and split, or significantly change in size out of the blue. This suite of attractor crises is indicative of global bifurcations and in some way a more complex analogue of the homoclinic bifurcation known from the disease-free 2D predator–prey model. The non-local phenomena characteristic of the Holling type IV predator–prey model therefore persist, in increased variety, also in the 3D model with disease. Hence, group defence tends to induce sudden catastrophic changes in the qualitative dynamics.

Discussion

In this paper, we consider how an infectious disease in the prey affects the predator–prey relationship where the prey exhibits some kind of group defence against the predator. We find that the disease can reduce prey densities to levels where the group defence is not as strong. This allows

predators to survive in situations where they could not without the disease.

In the absence of the disease, there are three Scenarios where the predator cannot survive: prey-only steady with no other unstable steady states (Scenario 1), bistability between prey-only steady state and predator–prey steady state/oscillations (Scenario 3, survival depends on the starting point) and a prey-only steady state with two unstable predator–prey steady states (Scenario 4). The disease can help the predator survive in the latter two cases. Firstly, the disease can reduce the prey carrying capacity to densities more manageable for the predator, moving from bistability between a prey-only steady state and a predator–prey steady state/limit cycle to where only the predator–prey steady state/limit cycle is stable. On top of this, the disease can reverse a homoclinic bifurcation, going from just a prey-only steady state to bistability between the prey-only steady state and the predator–prey limit cycle. This is due to the disease dampening the predator–prey oscillations, keeping prey densities too small for group defence to dominate. Combining these two phenomena together, we do have cases that, with the disease, only the predator–prey steady state/limit cycles are stable, whereas in the absence of the disease, only the prey-only steady state exists. In this case, the disease is helping the predator survive for all initial conditions where it could not survive in the diseases absence.

Typically, both the predator and disease are in competition for prey hosts. In several models, this competition leads to only one of the predator or disease persisting, i.e. the predator/disease manages to keep prey/host density low enough that the disease/predator population will eventually die out (for example, Hilker and Malchow (2006) and Siekmann et al. (2010), although coexistence can occur if all populations oscillate). Here, in both the density-dependent and frequency-dependent models, there is a stable predator–prey–disease equilibrium. This was also true in the diseased prey models in Bate and Hilker (2013) and several extension models in Table 6 of Anderson and May (1986), although this was not elaborated in either paper. This is novel in itself, especially for the density-dependent model, since the principle of competitive exclusion states that two consumers cannot share a resource. Previously, counterexamples are the result of temporal heterogeneity (Armstrong and McGehee (1980), for example, via predator–predator–prey oscillations or spatial heterogeneity (Chesson 2000)). Here, we have steady state coexistence, which is largely independent of the choice of functional response (for example, using linear and Holling type II functional responses would also have steady state coexistence). In particular, it is independent of group defence; however, with group defence, we find that the disease not only coexists with predators, it can also help predators survive where they could not without the disease.

The counterexample of the principle of competitive exclusion found in the density-dependent model occurs because there is a mix of density-dependent and density-independent terms in the prevalence equation (17). Gurney and Nisbet (1998, pp.166–167) found that adding a density-dependent mortality (a quadratic term) to one of the predators allowed both predators to coexist at equilibrium. This can be generalised to other forms of density dependence like predator interference (by using a Beddington–DeAngelis functional response) in one or both predators. The reason that density dependence defies competitive exclusion is that it gives the predator a range of prey densities under which it can be at equilibrium, and if the other consumer can also survive at steady state in this range, coexistence can occur. Without density dependence, the range is a point which means coexistence generally cannot occur. The same density dependence argument occurs in the density-dependent model, in the prevalence equation (17), since the mixture of density-dependent $((\beta x-\mu)(1-i))$ and density-independent (b) terms means that prevalence can be static for a range of x.

There are several key assumptions in this model that lead to coexistence of both the disease and predator. For example, if infected prey are completely sterile, then the b term in prevalence equation (17) becomes $b(1-i)$. With this, the prevalence equation can only be static for $i^*=0,1$ unless prey density is $x^*=\frac{\mu+b}{\beta}$, which is generally not true. Since $i^*=1$ means all prey are infected and sterile (leading to the extinction of the prey and predator), equilibrial coexistence between predator and disease cannot occur in general. Likewise, the lack of vertical transmission also allows for coexistence (for example, in Hilker and Malchow 2006 and Siekmann et al. 2010, there is perfect vertical transmission, an assumption that leads to the lack of equilibrial coexistence).

For the frequency-dependent model, coexistence of predator and disease is not as profound as is the case in the density-dependent model. The prevalence equation 15 shows that the prevalence–prey equations follow amensalism (disease prevalence has a negative effect on prey growth but disease prevalence does not gain from higher prey densities) and not exploitation. Consequently, the principle of competitive exclusion does not apply for frequency-dependent transmission.

Venturino (2011) tackled group defence from a different perspective, leading to significantly different result. Instead of a non-monotonic functional response like the Holling Type IV used in this paper, he uses square root functional response. This choice of functional response is based on the idea that predators can only take prey on the outskirt of the herd and thus the functional response should be proportional to the perimeter of the herd. However, Venturino (2011) assumes this only applies to susceptible prey since

infected prey are assumed to leave the herd and thus experience a linear functional response. The resulting dynamics are less complicated in their model, only equilibria and limit cycles seem to occur with no bistability. Coexistence between predator and prey can occur as well as cases where the disease helps the predator survive. In this paper, bistability occurs in Scenario 3 and more complex dynamics can occur in the density-dependent model.

Previous eco-epidemiological models have demonstrated equilibrial coexistence between predator and disease for the prey host invariably, but those models cannot be simplified to a exploitative competition model. Instead, they can only be simplified to an intraguild predation or food chain model (in particular, Venturino (2011)). As such, coexistence between predator and disease is expected. There is one model that looks like exploitative competition and has coexistence (Das et al. 2009), but this coexistence occurs because the predator grows logistically in the absence of prey, so implicitly the predator has another resource.

For brevity, we have not looked into the case where the prey nullcline has both a maximum and a minimum (see Appendix). In this case, steady states with low prey density are stable, likewise for high prey density (i.e. Scenarios 2A and 3A), but for moderate densities, the steady state is unstable (i.e. Scenarios 2B, 3B and 4). Given this nullcline will probably flatten, move to the left and eventually lose both extrema as we increase prevalence/virulence, it seems plausible that there may be some prevalence region where we are in Scenario 4 whereas without the disease we would be in Scenario 3A or 3B. However, further increases in prevalence/virulence would reverse this and go through the usual pattern from Scenario 4 to Scenario 3 to Scenario 2 and so forth.

In this paper, we derive a general 'two species' Holling type IV functional response incorporating a handling time that is linear with respect to prey density to a Holling type II functional response. This formulation, although straightforward, seems novel as multispecies Holling type IV functional responses are rarely considered and single species Holling type IV functional responses are usually stated and not derived and explained. In particular, assuming that handling time is a linear function of prey density seems to be the simplest assumption in deriving a single species Holling type IV functional response.

For simplicity, we assumed that the infected prey and susceptible prey are equivalent. Although the full model is cumbersome, future investigations could relax some of these simplifying assumption. For example, we could assume that infected prey may contribute less to the group defence. The authors suspect that if a disease does weaken group defence by more than just reducing host density, the disease could even further benefit the predator by increasing predator density and not just by eliminating extinction risk. In particular, if the disease is trophically transmitted (we have direct transmission in this paper), it may be beneficial for the disease if the infected prey break down group defence to aid transmission to predators. This should depend on relative importance for the disease of the effect on prey to predator transmission as well as the greater predator numbers and lower prey numbers caused by the breakdown of group defence. However, by doing so, the resulting eco-epidemiological system would almost certainly result in more complicated intraguild predation.

To conclude, we find that predator and disease can coexist at steady state, contradicting the principle of competition. On top of this, in some cases where group defence in the prey is prominent, coexistence between prey and predator can often benefit from the presence of the disease, either by reversing a homoclinic bifurcation or by reducing the prey density below a group defence threshold.

Appendix: steady state analysis

Disease-free model

From steady state analysis, we have the following conditions for each Scenario (assuming $b > m$):

- Scenario 1: There is no coexistent steady state. Prey-only steady state is stable. (1A)$H_0 > 1$ or $(H_0-1)^2 - 4H_x < 0$ (no real solutions), (1B)$H_0 < 1$, $(H_0-1)^2 - 4H_x > 0$, $\frac{b-m}{c} < \frac{(1-H_0)\pm\sqrt{(H_0-1)^2-4H_x}}{2H_x}$ (two negative solutions).
- Scenario 2: One coexistent steady state exists. It is either (2A) stable or (2B) the centre of some stable limit cycle (depending on the sign of $y'(x^*)$) ($H_0 < 1$, $(H_0-1)^2-4H_x > 0$, $\frac{(1-H_0)-\sqrt{(H_0-1)^2-4H_x}}{2H_x} < \frac{b-m}{c} < \frac{(1-H_0)+\sqrt{(H_0-1)^2-4H_x}}{2H_x}$ (one positive and one negative solution))
- Scenario 3: Two coexistent steady state exists. The coexistent steady state with the lower prey density is either (3A) stable or (3B) the centre of some limit cycle (depending on the sign of $y'(x^*)$). This is bistable with the prey-only steady state, where the higher prey density coexistent steady state acting as a separatrix. ($H_0 < 1$, $(H_0-1)^2 - 4H_x > 0$, $\frac{(1-H_0)\pm\sqrt{(H_0-1)^2-4H_x}}{2H_x} < \frac{b-m}{c}$ (two positive solutions))

Frequency-dependent model

The conditions are the same for the frequency-dependent model as for the disease-free model except you must substitute m with $m + \mu i^*$, where $i^* = \max\left(0, 1 - \frac{b}{\beta-\mu}\right)$. Note

that if $b < m+\mu i^*$, then the disease will cause the extinction of both predator and prey.

Density-dependent model

The steady states (x, y, i) are:

- $(0, 0, 0)$ always exists and is stable if $b < m$
- $(x^*, 0, 0)$, where $x^* = \frac{b-m}{c}$. This exists when $b > m$ and is stable when $f(x^*) < 1$ (i.e. predators cannot survive) and $\frac{\beta x^*}{\mu+b} < 1$ (i.e. disease cannot spread)
- $(x^*, 0, i^*)$ where x^*, i^* solve $p(x^*, i^*) = 0$ and $g(x^*) = \mu i^*$. This exists when $x^* > 0$ and $i^* > 0$ (i.e. $b > m$ and $\frac{\beta x^*}{\mu+b} > 1$). It is stable if $f(x^*) < 1$ (i.e. predators cannot survive)
- $(x^*, y^*, 0)$, where x^* solves $f(x^*) = 1$ (i.e. $x^* = \frac{(1-H_0)\pm\sqrt{(H_0-1)^2-4H_x}}{2H_x}$) and $y^* = g(x^*)h(x^*)$. This exists if $x^* > 0$ and $g(x^*) > 0$ (i.e. $H_0 < 1$, $(H_0-1)^2 - 4H_x > 0$ and $x^* < \frac{b-m}{c}$). This means there can be up to two such steady states . It is stable if $\frac{\beta x^*}{\mu+m} < 1$ (i.e. disease cannot invade), $f'(x^*) > 0$ and $h(x^*)g'(x^*) + h'(x^*)g(x^*) := y'(x^*) < 0$. If $f'(x^*) < 0$, then this steady state is a saddle point, whereas if $f'(x^*) > 0$ and $y'(x^*) > 0$, we have that the steady state in unstable and is surrounded by a stable limit cycle. The sign of $f'(x^*)$ depends on the relative values of x^* (when two steady states occur); the smaller x^* has $f'(x^*) > 0$, whereas the larger x^* has $f'(x^*) < 0$.
- (x^*, y^*, i^*), where x^* solves $f(x^*) = 1$ (i.e. $x^* = \frac{(1-H_0)\pm\sqrt{(H_0-1)^2-4H_x}}{2H_x}$), i^* solves $p(x^*, i^*) = 0$ and $y^* = (g(x^*) - \mu i^*)h(x^*)$. This exists if $x^* > 0$ (i.e. $H_0 < 1$ and $(H_0-1)^2 - 4H_x > 0$), $i^* > 0$ (i.e. $\frac{\beta x^*}{\mu+m} > 1$) and $y^* > 0$ (i.e. $g(x^*) > \mu i^*$). This means that there can be up to two steady states. By using qualitative stability criteria on the Jacobian at these steady states, we have that the system is definitely stable when $f'(x^*) > 0$ and $\frac{\partial y^*}{dx^*}(x^*, i^*) < 0$. Likewise, if $f'(x^*) < 0$, then the Jacobian has a positive determinant which means the steady state is unstable. If $\frac{\partial y^*}{dx^*}(x^*, i^*) > \frac{i(\beta x^*-\mu)}{f(x^*)}$, then the Jacobian has a positive trace which means the steady state is unstable. Consequently, we only do not know the stability for the region $f'(x^*) > 0$ and $0 < \frac{\partial y^*}{dx^*}(x^*, i^*) < \frac{i(\beta x^*-\mu)}{f(x^*)}$, presumably there is a Hopf bifurcation within this region (like the disease-free case). Like the predator–prey case, there can be up to two steady states.

This steady state analysis can be summarised into the same Scenarios as before, but some of the criteria have not been fully analysed. In particular, the Hopf bifurcation separating Scenario 2A/3A and 2B/3B has not been found.

Phase plane analysis

To complement the steady state analysis, we can use phase plane analysis to derive and demonstrate the different Scenarios (Figs. 1 and 2). For simplicity, we will use nullclines to refer to both the nullclines of the predator–prey system and nullplanes of both the frequency and density-dependent models.

There are up to three different predator nullclines. The predator-free nullcline ($y = 0$) always exists. The other two nullclines are the roots (if they exist) of the quadratic equation derived from $f(x) = 1$. These roots are always positive when they exist.

There are two different prey nullclines; one is the prey-free nullcline ($x = 0$), the other nullcline is derived from $y = h(x)(g(x) - \mu i)$. The latter nullcline is in fact cubic with respect to x. Assuming that $b > m + \mu i$, then the intercept at $x = 0$ is positive, and there is one intercept with $y = 0$ at $g(x) = \mu' i$. Given that the nullcline is cubic with respect to x, there can be up to two local extrema. Thus the nullcline can have:

- no realistic (positive) extrema ($y'(0) < 0$ and $y'(x)$ has no positive (or real) roots).
- two realistic (positive) extrema ($y'(0) < 0$ and $y'(x)$ has two positive roots). These extrema are one local minimum and one local maximum, the minimum occurs at lower prey density than the maximum. The region between these two extreme has a positive slope $y'(x) > 0$.
- only one positive local maximum ($y'(0) > 0$)

For simplicity, we will consider the third type of (disease-free) nullclines. The first case will not have a limit cycle, as $y'(x) < 0$ for all $x > 0$. This means only Scenarios 1, 2A and 3A can occur. The second case is a little more complex than the third case, but the same arguments still apply. In fact, the only difference is that for small prey densities (lower than the local minimum), $y'(x) < 0$ and thus steady states can be stable here. In between the maximum and minimum, limit cycles are likely to occur. This formulation does not add any new Scenarios but may change the order of Scenario changes when we increase prevalence. In particular, it seems plausible that the disease may destabilise the predator–prey equilibrium if the disease moves the minimum to a lower prey density than the lower predator nullcline (i.e. going from Scenario 2A to Scenario 2B or from Scenario 3A to Scenario 3B or 4).

There are at most two disease nullclines, the disease-free nullcline $i = 0$ and the endemic nullcline $p(i, x) = 0$. In the frequency-dependent model, the endemic nullcline is $i = 1 - \frac{b}{\beta-\mu}$.

References

Ajraldi V, Pittavino M, Venturino E (2011) Modelling herd behaviour in population systems. Nonlinear Anal R World Appl 12:2319–2338

Allen WE (1920) Behavior of loon and sardines. Ecology 1:309–310

Anderson RM, May RM (1986) The invasion, persistence and spread of infectious disease with animal and plant communities. Abstr Pap Print Philos Trans R Soc Lond B Biol Sci 314:533–570

Andrews JF (1968) A mathematical model for the continuous culture of microorganisms utilizing inhibitory substrates. Biotechnol Bioeng 10:707–723

Armstrong RA, McGehee R (1980) Competitive exclusion. Am Nat 115:151–170

Bate AM, Hilker FM (2013) Predator–prey oscillations can shift when diseases become endemic. J Theor Biol 316:1–8

Chesson P (2000) General theory of competitive coexistence in spatially-varying environments. Theor Popul Biol 58:211–137

Collings JB (1997) The effects of the functional response on the bifurcation behavior of a mite predator-prey interaction model. J Math Biol 36:149–168

Das K, Roy S, Chattopadhyay J (2009) Effect of disease-selective predation on prey infected by contact and external sources. Biosystems 95:188–199

Freedman HI, Wolkowicz GSK (1986) Predator–prey systems with groups defence: the paradox of enrichment revisited. Bull Math Biol 48:493–508

Gause G (1934) The struggle for existence. Williams and Wilkins, Baltimore

Geritz SAH, Gyllenberg M (2013) Group defence and the predator's functional response. J Math Biol 66:705–717

Gurney WSC, Nisbet RM (1998) Ecological dynamics. Oxford University

Hardin G (1960) The competitive exclusion principle. Science 131:1292–1297

Hilker FM, Malchow H (2006) Strange periodic attractors in a prey–predator system with infected prey. Math Popul Stud 13:119–134

Holling CS (1959) Some characteristics of simple types of predation and parasitism. Can Entomol 91:385–398

Koen-Alonso M (2007) A process-oriented approach to the multispecies functional response. In: From energetics to ecosystems: the dynamics and structure of ecological systems, Springer, p 1–36

Kot M (2001) Elements of mathematical ecology. Cambridge University

Krause J, Ruxton GD (2002) Living in groups. Oxford University

McGehee R, Armstrong RA (1977) Some mathematical problems concerning the ecological principle of competitive exclusion. J Differ Equ 23:30–52

Miller RC (1922) The significance of gregarious habit. Ecology 3:122–126

Murdoch WW (1972) The functional response of predators. Biol Control 15:237–40

Ono M, Igarashi T, Ohno E, Sasaki M (1995) Unusual thermal defence by a honeybee against mass attack by hornets. Nature 377:334–336

Ruan S, Xiao D (2001) Global analysis in a predator–prey system with nonmonotonic functional response. SIAM J Appl Math 61:1445–1472

Seppälä O, Karvonen A, Valtonen ET (2008) Shoaling behaviour of fish under parasitism and predation risk. Anim Behav 75:145–150

Sieber M, Hilker FM (2011) Prey, predators, parasites: intraguild predation or simpler community models in disguise? J Anim Ecol 80:414–421

Siekmann I, Malchow H, Venturino E (2010) On competition of predators and prey infection. Ecological Complexity 7:446–457

Venturino E (2011) A minimal model for ecoepidemics with group defense. J Biol Syst 19:763–785