Theoretical Ecology
https://doi.org/10.1007/s12080-021-00507-9

ORIGINAL PAPER

Eco-evolutionary dynamics of autotomy

Rohan S. Mehta[1] · **Julie A. Kraus**[2]

Received: 31 July 2020 / Accepted: 31 March 2021

Abstract
The act of deliberately removing a body part–called autotomy–is a behavior that has appeared frequently across the tree of life. Though there are many possible functions for this behavior, it is often thought of first as a mechanism to escape predation. In a predator–prey interaction, autotomy can therefore confer significant benefits to the prey–particularly, the benefit of not being eaten–but it may also incur significant costs, including the energy expenditure required to regrow the body part and any additional consequences of losing the body part in the first place. In addition, the presence of autotomy may affect how predators choose to approach prey. Here, we put these costs and benefits into a game theory framework and analyze the evolutionary, ecological, and eco-evolutionary dynamics of autotomy, considering both predator and prey strategies. We also apply our model to an empirical example using existing data from porcelain crabs. We find a wide range of effects of autotomy on the ecological and evolutionary dynamics of the predators and prey, including the possibility that the prey become locked into performing autotomy by the predators and the possibility that autotomy can rescue predator–prey coexistence.

Keywords Autotomy · Predator–prey interactions · Game theory · Evolutionary dynamics · Ecological dynamics · Eco-evolutionary dynamics

Introduction

Prey employ a wide variety of strategies to avoid or escape predators (Caro 2005; Ruxton et al. 2019), including camouflage, warning coloration (aposematism), mimicry, attempting to startle the predator, toxicity and other secondary defense mechanisms, and attempting to deflect the predation. These traits are generally considered to be a product of coevolution between the predator and the prey; these coevolutionary processes often result in an “arms race” between predator and prey (Cott 1940; Dawkins and Krebs 1979), but this is not the only possible outcome (Abrams 1986; Vermeij 1994).

The prey behavior of autotomy—the self-controlled loss of a body part at a predetermined breakage location—is an example of a prey behavior that is used to deflect predation. Autotomy is widespread throughout the tree of life (Emberts et al. 2019) and is therefore thought to have many independent origins (McVean 1975); there may be many different ways for which autotomy can be advantageous. Autotomy was originally conceived of as a mechanism of predator escape (Fredericq 1883), but several other potential benefits of the behavior have been discovered (Maginnis 2006; Emberts et al. 2019), including getting rid of injured body parts (Emberts et al. 2017), escape from non-predatory entrapment (e.g., (Hodgkin et al. 2014)), and as a method of intrasexual competition (e.g., (Uhl et al. 2010)). Autotomy also comes with serious costs to the individual (Maginnis 2006; Fleming et al. 2007; Emberts et al. 2019), including decreases in locomotive ability (e.g., (Martin and Avery 1998)), disadvantages in sexual selection (e.g., (Smith 1992)), energetic costs of regenerating the lost part (Maginnis 2006), and, ironically, decreased ability to escape from predators in the future (e.g., (Downes and Shine 2001)). The extreme nature of costs and benefits in such a dramatic behavioral trait warrants a thorough analysis of the conditions under which autotomy would be evolutionarily favorable.

Previous theoretical work on the ecology and evolution of autotomy includes several results in Optimal Escape Theory (Ydenberg and Dill 1986; Cooper Jr and Frederick 2007), which considers the problem of the Flight Initiation Distance, the distance a prey should start to try to escape the

---

✉ Rohan S. Mehta
rohan.sushrut.mehta@emory.edu

[1] Department of Physics, Emory University, Atlanta, GA, USA

[2] Department of Aerospace Engineering, Georgia Institute of Technology, Atlanta, GA, USA

predator. Cooper Jr (1998) considered conditions that would favor anticipatory or reactive autotomy (i.e., prior to or after the predator approaches). Cooper Jr and Frederick (2010) incorporated autotomy into an optimal escape model, concluding that the effects of autotomy on flight initiation distance are complex. Recently, Emberts et al. (2019) adapted the theory of Ydenberg and Dill (1986) to the problem of autotomy in a verbal economic model.

Outside of the Optimal Escape Theory framework, Cooper Jr and Vitt (1991) considered the evolution of conspicuousness in the parts of the body used for autotomy. Harris (1989) analyzed two Lotka–Volterra models involving autotomy (or "nonlethal injury"). The first divided the prey population into injured and uninjured portions, with the injury inflicted by a predator who consumes only part of the prey and whose existence depend entirely on this prey consumption. The second model modified the first by instead making the predator population size constant and thus unaffected by the prey consumption. In the first model, Harris (1989) concluded that for predators to survive, the damaged prey must have a higher death rate than reproductive rate. The second model had no steady-state that includes both predators and prey. As a building block within a larger conceptual framework of decoys in predator–prey interactions, Wilkinson (2003) added more realistic assumptions to the model of Harris (1989) and found that the presence of autotomy can dampen otherwise extreme oscillations in the population sizes of predators and prey.

Here, we analyze a game-theoretical model of the coevolution of autotomy in prey and prey choice in a predator, focusing on autotomy as a predator-escape strategy. We determine the conditions under which autotomy is evolutionarily favorable under this model. We also extend our model to incorporate ecological and eco-evolutionary dynamics and demonstrate the wide-ranging, complex effects of autotomy on these dynamics. Finally, we consider the connection of our results to empirically observed autotomy with existing data from porcelain crabs.

Model

We start by formulating a game-theoretical model for a single predator–prey interaction. The game has two players: one predator and one prey. The predator can have one of two strategies: Approach (α) or Avoid (ν) the prey (with the assumption that under the Avoid strategy, the predator instead chooses to approach a different type of prey from its pool of potential prey). The prey can have one of two strategies: Autotomy (A) or No autotomy (N). When it is potentially ambiguous, we refer to the potentially autotomizing prey in question as the "focal" prey, and other prey as "alternative" prey. Note that alternative prey do not play this game.

If the predator avoids the focal prey (ν), their payoff is completely determined by the quality of their alternative pool of potential prey, summarized by the quantity v.

If the predator approaches the focal prey (α), then their payoff is driven by the probability of killing the prey given the encounter. If k is the probability of killing the prey, r is the fraction by which this probability is reduced due to the autotomy strategy, p is the payoff achieved by eating the whole prey, and f is the fraction of the full payoff achieved by eating only the lost part, then the payoff of the predator is kp if the prey does not autotomize (N) and $rkp + (1 - rk)(fp) = p[rk + (1 - rk)f]$ if the prey autotomizes (A). We will generally assume that $r < 1$, that is, autotomy has some effect on the probability of escaping the predator.

For the prey, we can scale the payoffs to be between 0 (death) and 1 (successful escape without incurring any costs due to autotomy). If the prey does not autotomize (N), then the probability of dying (and achieving payoff 0) when a predator approaches (α) is k, and the probability of surviving (and therefore the payoff) is $1 - k$. When a predator does not approach (ν), the probability of surviving (and therefore the payoff) is 1.

If the prey autotomizes (A), and the predator approaches (α), then the probability of survival is $1 - rk$, and the payoff that the prey gets for surviving is decremented by c, the cost of the autotomy. This cost could be, for instance, the metabolic cost of maintaining the regenerative system, or a consequence of the post-autotomy effects of autotomizing individuals. The payoff of autotomizing (A) is therefore $(1 - rk)(1 - c)$. If, on the other hand, the predator does not approach (ν), then the probability of survival is 1, but the prey, having autotomized, still incurs the autotomy cost c. Note that a prey autotomizing against a predator that does not approach makes no practical sense in the context of predator escape (though there are other uses of autotomy that do not involve predator interaction). Hence, the payoff of autotomizing (A) against an avoiding predator (ν) is always strictly less than the payoff for not autotomizing (N) against an avoiding predator.

Table 1 describes the parameters and variables for this model, and Table 2 displays the payoff matrix for this game.

Results

Game theory analysis

We first analyze our autotomy game using the tools of game theory. It is useful to define the following composite quantities of our parameters:

Table 1 Description of parameters and variables for the game-theoretical (Table 2 and Eq. 4) and ecological (Eq. 6) models

Parameter	Description	Range
v	Predator payoff for pursuing alternative prey	$[0, \infty)$
p	Predator payoff for killing focal prey	$[0, \infty)$
c	Cost of autotomy for focal prey	$[0, 1]$
k	Probability of predator killing focal prey with no autotomy	$[0, 1]$
r	Reduction in kill probability due to autotomy	$[0, 1]$
f	Fraction of predator payoff for consuming just an autotomized part	$[0, 1]$
λ	Effect of a single focal prey kill on predator growth rate	$[0, \infty)$
μ	Predator death rate	$[0, \infty)$
β	Focal prey growth rate	$[0, \infty)$
γ	Predation rate	$[0, \infty)$
w	Frequency of avoiding strategy in predator (ν)	$[0, 1]$
x	Frequency of approaching strategy in predator (α)	$[0, 1]$
y	Frequency of non-autotomy strategy in focal prey (N)	$[0, 1]$
z	Frequency of autotomy strategy in prey (A)	$[0, 1]$
p_ν	Population density of avoiding predators (ν)	$[0, \infty)$
p_α	Population density of approaching predators (α)	$[0, \infty)$
p_A	Population density of autotomizing prey relative to carrying capacity (A)	$[0, 1 - p_N]$
p_N	Population density of non-autotomizing prey relative to carrying capacity (N)	$[0, 1 - p_A]$
P	Total predator population density	$[0, \infty)$
n	Total prey population density relative to carrying capacity	$[0, 1]$

$$
\begin{aligned}
K &= \frac{k - rk}{1 - rk} = \frac{(1 - r)k}{1 - rk} \\
V &= \frac{\frac{v}{p} - rk}{1 - rk}.
\end{aligned}
\qquad (1)
$$

K summarizes the effect of the escape aspect of autotomy on the predator with regards to the focal prey (i.e., a larger K means the predator is more affected by the escape aspect of autotomy). K also summarizes the overall intensity of the effect of the escape aspect of autotomy on the prey. V summarizes the effect that the escape aspect of autotomy has on the predator with regards to the alternative prey. The quantities K and V allow us to separate the escape aspects of autotomy from the prey costs (c) and the predator benefits (f). Figure 1 plots K and V (Eq. 1) for a range of parameter values.

For the predator, approaching (α) beats avoiding (ν) when played against non-autotomizing focal prey (N) if non-autotomizing focal prey are better than alternative prey ($kp > v$), and when played against autotomizing focal prey (A) if the autotomy benefit is greater than the benefit gained from pursuing alternative prey ($f > V$). For notational consistency, we note that $kp > v$ is equivalent to $K > V$, so long as $rk < 1$.

For the prey, autotomy (A) beats no autotomy (N) when played against approaching predators (α) if the autotomy cost is less than the escape benefit ($c < K$). As discussed previously, autotomy (A) can never beat no-autotomy (N) when played against avoiding predators (ν).

The effect of combining each of these conditions is described in Fig. 2, which presents the expected equilibrium evolutionary outcomes of this game. In general, autotomy is only such an outcome when costs to prey are low ($c < K$) and benefits to predators are high enough to warrant approach ($f > V$).

The region in Fig. 2 with two possible equilibria ($c < K < V < f$) occurs when it is advantageous for predators to approach (α) when the prey autotomize (A) but not

Table 2 The payoff matrix for the predator–prey game, with predator payoffs listed on top

		Prey	
		No autotomy (N)	Autotomy (A)
Predator	Avoid (ν)	v 1	v 1-c
	Approach (α)	kp $1 - k$	$p\,[rk + (1 - rk)f]$ $(1 - rk)(1 - c)$

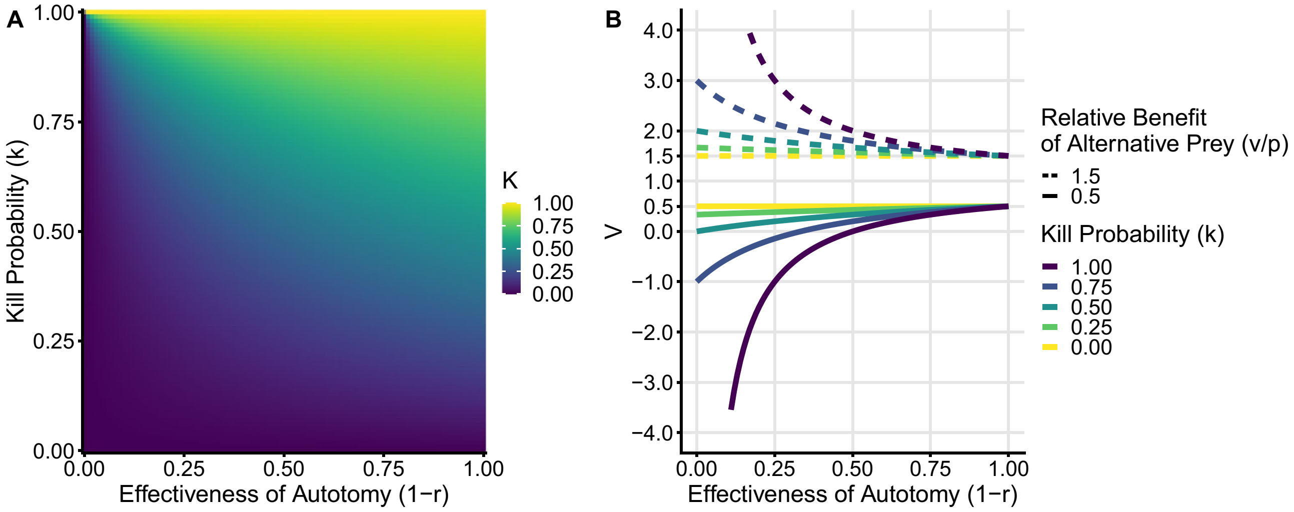

Fig. 1 **(a)** The quantity K increases with increasing kill probability (k) and increasing effectiveness of autotomy $(1-r)$. K summarizes the effect of the escape aspect of autotomy on the predator, as well as the intensity of the effect of autotomy on the prey. **(b)** If alternative prey is better ($\frac{v}{p} > 1$), then the quantity V increases with increasing kill probability (k) and decreasing effectiveness of autotomy $(1-r)$. If alternative prey is worse ($\frac{v}{p} < 1$), then V instead decreases with increasing k and decreasing $1-r$. For $k = 1$ and $r = 1$ in **(a)**, the value of K is 0. If $\frac{v}{p} = 1$ in **(b)**, then $V = 1$ for all r and k. For $k = 1$ in (B), the magnitude of V grows without bound for decreasing $1-r$ and was truncated for this figure

when prey do not autotomize (N), and it is advantageous for prey to autotomize (A) upon predator approach (α) but not upon predator avoidance (ν). Thus, the best response for each player is different based on what the other player plays. For instance, if all prey are non-autotomizing, then the predator should approach alternative prey instead of the focal prey, and if all prey autotomize, then the autotomy benefit is sufficient that predators should approach the focal prey.

The cyclic region in Fig. 2 ($c < K$, $f < V < K$) occurs when it is advantageous for predators to approach (α) non-autotomizing prey (N) but not autotomizing prey (A), and it is advantageous for prey to autotomize (A) upon predator approach (α) but not upon predator avoidance (ν). Thus, assuming we start with avoiding predators and non-autotomizing prey (νN), we would subsequently expect predators to approach the non-autotomizing prey (αN),

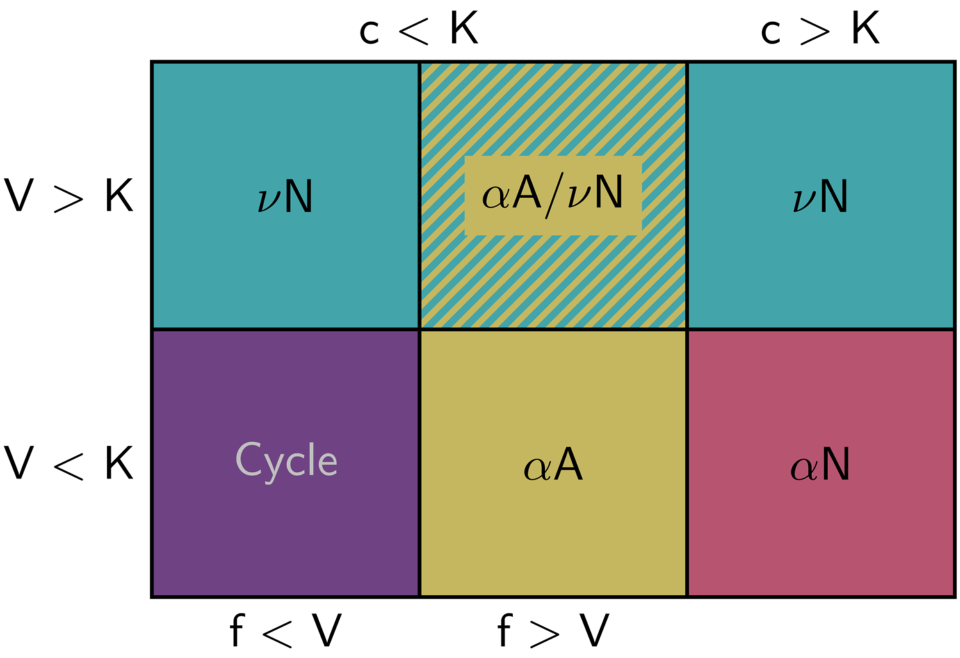

Fig. 2 Expected outcomes (Nash equilibria) for the autotomy game. νN = No predation, no autotomy. αN = Predation, no autotomy. νA = No predation, autotomy. αA = Predation, autotomy. The cycle is νN → αN → αA → νA → νN

which would incentivize prey to autotomize (αA), which would incentivize predators to avoid the focal prey (νA), which would incentivize prey to stop autotomizing (νN), completing the cycle.

Until now, we have only considered fixed, deterministic strategies, but it is possible for predators and prey to employ stochastic, mixed strategies as well. Let w be the probability of predator avoidance (ν), x be the probability of predator approach (α), y be the probability of prey not autotomizing (N), and z be the probability of prey autotomizing (A). We will use the substitutions $w = 1 - x$ and $y = 1 - z$ in order to focus our attention on the probabilities of the two "interesting" strategies, approach (α) and autotomy (A).

The mixed strategy solution for this system is

$$\hat{x} = \frac{c}{k[1 - (1 - c)r]} \tag{2}$$

$$\hat{z} = \frac{V - K}{f - K}. \tag{3}$$

The mixed strategy $(\hat{x}, \hat{z})$ only exists when either αA and νN are both expected equilibrium outcomes or when the system is expected to cycle.

Figure 3 explores how the parameters affect this mixed strategy. The frequency of approach ($\hat{x}$, Eq. 2) increases as parameters change in such a way that K decreases or c increases; these situations all discourage autotomy and so predators must approach more frequently to get the same benefits. The frequency of autotomy ($\hat{z}$, Eq. 3) increases when predators are increasingly encouraged to approach the focal prey (through low V and increasing f) and when predators would be otherwise encouraged to avoid the focal prey (through high V and decreasing f).

Evolutionary dynamics

To analyze the evolutionary dynamics of the autotomy game, we study the replicator game dynamics (Hofbauer and Sigmund 2003; Cressman and Tao 2014) of the four strategies, using the same variables w, x, y, and z that we used to compute the mixed strategy solution. We again use the substitution $w = 1 - x$ and $y = 1 - z$. The replicator equations for our model are

$$\begin{aligned} \frac{dx}{dt} &= x(1 - x)p(1 - rk)[z(f - K) - (V - K)] \\ \frac{dz}{dt} &= z(1 - z)[xk[1 - (1 - c)r] - c]. \end{aligned} \tag{4}$$

The equilibria are (0, 0), (0, 1), (1, 0), (1, 1), and $(\hat{x}, \hat{z})$, where

$$\begin{aligned} \hat{x} &= \frac{c}{k[1 - (1 - c)r]} \\ \hat{z} &= \frac{V - K}{f - K} \end{aligned} \tag{5}$$

Note that $\hat{x}$ and $\hat{z}$ also describe the mixed-strategy equilibrium, which is a general property of replicator game dynamics (Hofbauer and Sigmund 2003; Cressman and Tao 2014).

Details of the stability analysis of Eq. 4 are presented in Appendix B. Table 3 presents stability conditions for each of the equilibria. These stability conditions map directly onto the

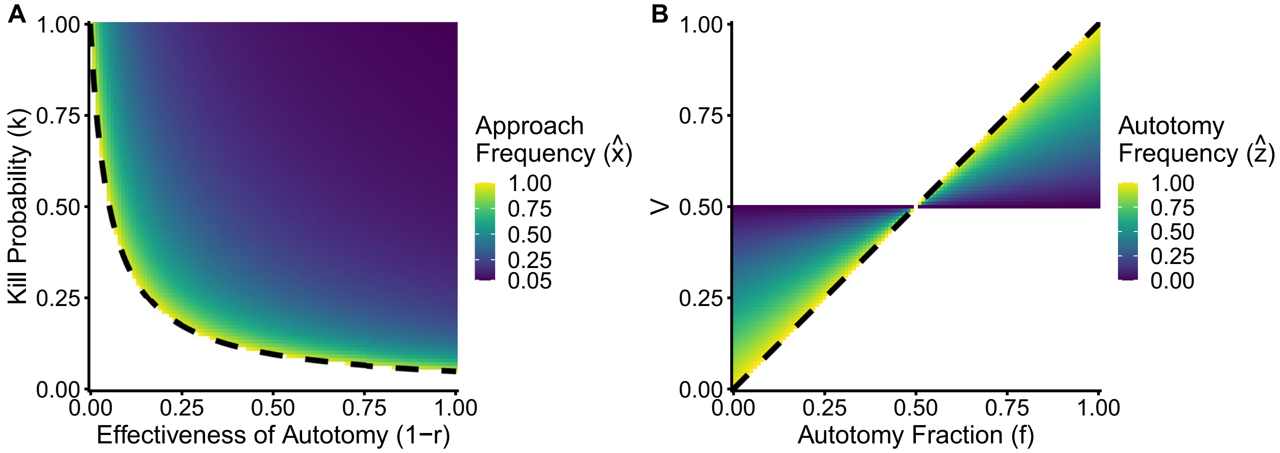

Fig. 3 **(a)** Approach frequency $\hat{x}$ (Eq. 2) increases monotonically with decreasing autotomy effectiveness $1 - r$ and kill probability k, both situations that decrease K, requiring predators to approach more frequently to get the same payoff. The dashed line indicates the boundary $k = \frac{c}{1-r+rc}$, where $c = 0.05$ in this case. Beyond this boundary, the mixed strategy equilibrium does not exist. **(b)** Autotomy frequency $\hat{z}$ (Eq. 3) increases as V goes from K ($K = 0.5$ in this case) to f. $\hat{z}$ also increases with increasing autotomy fraction f if $V < K$ (that is, when predators are encouraged to approach the focal prey, thus putting more emphasis on escape) and decreases otherwise. The dashed line indicates the boundary $f = V$. Beyond this boundary, the mixed strategy equilibrium does not exist

Table 3 Stability conditions for the equilibria of Eq. 4

Equilibrium	Strategy	Stability Conditions
$(0, 0)$	νN	$V > K$
$(1, 0)$	αN	$V < K$ and $c > K$
$(0, 1)$	νA	Never stable
$(1, 1)$	αA	$c < K$ and $f > V$
$(\hat{x}, \hat{z})$	Mixed	Never stable

game-theoretical conditions in Fig. 2, where the cycle occurs when the interior equilibrium is a center, the bistable αA/νN state occurs when the interior equilibrium is a saddle point, αA is an equilibrium when $(1, 1)$ is stable, αN is an equilibrium when $(1, 0)$ is stable, νA is never an equilibrium, and νN is an equilibrium when $(0, 0)$ is stable. This mapping is also a general property of replicator game dynamics (Hofbauer and Sigmund 2003; Cressman and Tao 2014).

Note that for the sake of conciseness we will refer to unstable nodes as "unstable" and saddle points—which are also technically unstable—as simply "saddle points."

We provide these replicator dynamics results primarily to demonstrate the link between our game theory model and the evolutionary dynamics that we combine with an ecological model in following sections. For general understanding of the evolutionary outcomes of this game, however, Fig. 2 is sufficient.

Figure 4 displays phase portraits for each parameter region in Fig. 2. We have labeled Fig. 4 so that the panels correspond to regions in Fig. 2, with the exception that the $c > K$ region in Fig. 4 is now also split up into $f < V$ and $f > V$ parts to demonstrate that there are slightly different transient dynamics for these two regimes that do not appear in the game theory model because that model does not address transient dynamics. Parameters used in Fig. 4 are listed in Table 4. The fact that $(1, 1)$ is either a saddle point or an unstable equilibrium only makes a difference with respect to the dynamics in these cases; in the former case, the population will quickly move toward the $(0, 0)$ stable equilibrium, but in the latter case the population will spend a significant amount of time with a higher frequency of approach (x) due to the increased (if temporary) benefit of autotomy on predators.

Ecological dynamics

In the context of this analysis, autotomy is an ecological interaction; our game theory model is therefore missing this fundamental aspect of the process. As a result, we next consider a purely ecological predator–prey model. Because predators that avoid the focal prey (ν) do not ecologically interact either with the focal prey or with the approaching predators (α), they play no part in the purely ecological model. We can formulate a modified version of the standard Lotka–Volterra system (Hofbauer and Sigmund 1998) (with prey carrying capacity set to 1) that tracks the remaining three strategy population sizes (p_α, p_A, and p_N for approaching predators, autotomizing prey, and non-autotomizing prey, respectively) as follows:

$$\begin{aligned}
\frac{dp_\alpha}{dt} &= \lambda[kp_N + [k + (f - K)(1 - rk)]p_A]p_\alpha - \mu p_\alpha \\
\frac{dp_A}{dt} &= \beta(1 - c)p_A(1 - p_A - p_N) - \gamma rkp_Ap_\alpha \\
\frac{dp_N}{dt} &= \beta p_N(1 - p_A - p_N) - \gamma kp_Np_\alpha,
\end{aligned} \tag{6}$$

where the parameters λk, μ, β, and γk correspond to the standard Lotka–Volterra growth and death rates for predators and prey, respectively. The cost of autotomy decreases prey growth rate through c and autotomy decreases the predation rate through r. The effect of autotomy on the predator is through the decreased kill probability rk and the benefit of eating the lost part f. For a full derivation and analysis of this model, see Appendix C.

For ease of notation, we define

$$k' = k + (f - K)(1 - rk). \tag{7}$$

and

$$\begin{aligned}
\beta_0 &= \frac{\beta}{k\gamma} \\
n_0 &= \frac{\mu}{\lambda k} \\
n_1 &= \frac{\mu}{\lambda k'}.
\end{aligned} \tag{8}$$

Table 4 Parameters used in Fig. 4

Parameter	A	B	C	D	E	F	G	H
v	2	0.5	2	0.5	0.2	0.2	0.2	0.2
p	1	1	1	1	1	1	1	1
c	0.1	0.1	0.5	0.5	0.1	0.1	0.5	0.5
k	0.5	0.3	0.5	0.3	0.3	0.3	0.3	0.3
r	0.3	0.3	0.3	0.3	0.3	0.3	0.3	0.3
f	0.2	0.7	0.2	0.7	0.05	0.2	0.05	0.2

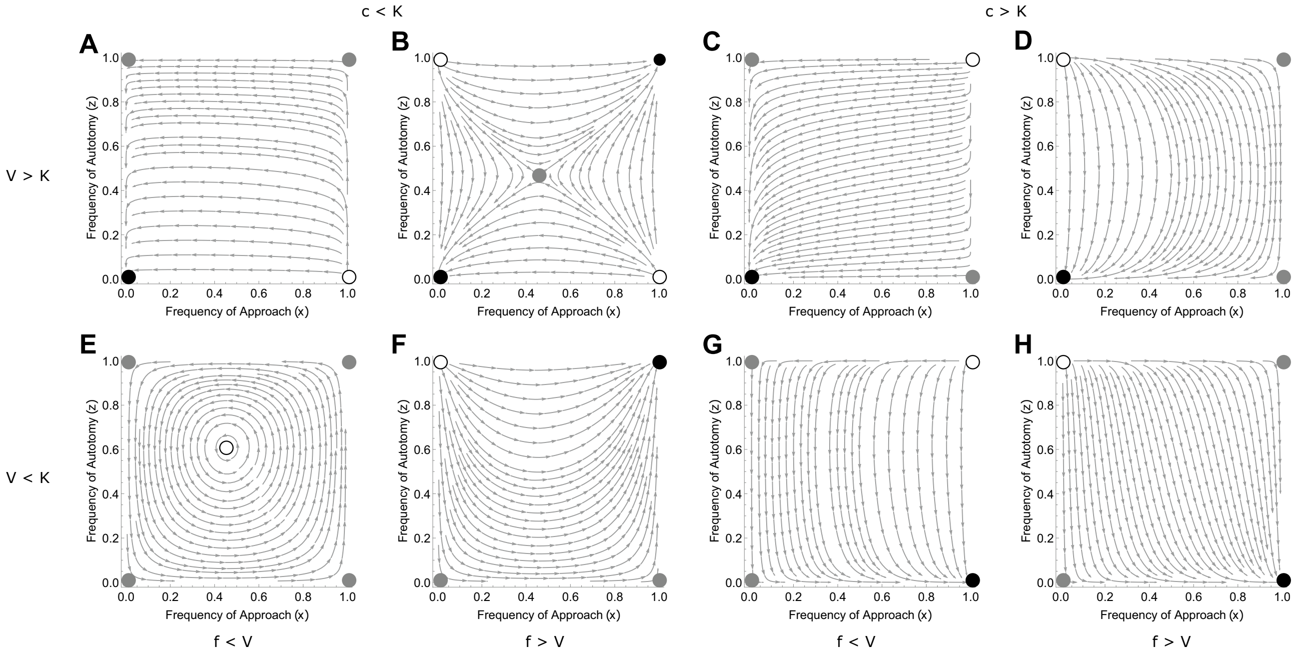

Fig. 4 Examples of phase portraits for the replicator dynamics (Eq. 4), where each curve represents a possible trajectory. **(a)** (0, 0) (νN) stable, frequency of approach quickly decreases, but frequency of autotomy transiently increases. **(b)** (0, 0) (νN) and (1, 1) (αA) stable. **(c)** (0, 0) (νN) stable, monotonic decrease in both quantities. **(d)** (0, 0) (νN) stable, frequency of autotomy quickly decreases, but frequency of approach transiently increases. **(e)** Cycle. **(f)** (1, 1) (αA) stable, frequency of approach quickly increases, but frequency of autotomy transiently decreases. **(g)** (1, 0) (αN) stable, frequency of autotomy quickly decreases, but frequency of approach transiently decreases. **(h)** (1, 0) (αN) stable, monotonic decrease in frequency of autotomy and monotonic increase in frequency of approach. Black dots are stable, gray dots are saddle points, and white dots are unstable. Panels A, B, E, and F correspond to the four panels under the $c < K$ label (left and middle columns) in Fig. 2. Panels C and D are subcases of the $c > K$, $V > K$ panel (top right) in Fig. 2. Panels G and H are subcases of the $c > K$, $V < K$ panel (bottom right) in Fig. 2. Parameters used for this figure are listed in Table 4

There are five equilibria of this system, three of which are solvable without initial conditions: one with total extinction ($\hat{p}_\alpha = \hat{p}_A = \hat{p}_N = 0$), one with coexistence and only autotomizing prey:

$$\begin{aligned}\hat{p}_\alpha &= \beta_0 \frac{1-c}{r}(1-n_1)\\ \hat{p}_A &= n_1\\ \hat{p}_N &= 0,\end{aligned} \tag{9}$$

one with coexistence and only non-autotomizing prey:

$$\begin{aligned}\hat{p}_\alpha &= \beta_0(1-n_0)\\ \hat{p}_A &= 0\\ \hat{p}_N &= n_0,\end{aligned} \tag{10}$$

and two underdetermined equilibria where initial conditions matter: an only-prey equilibrium with $p_\alpha = 0$ and $p_A + p_N = 1$, and a coexistence between all three strategies that can only exist if $1 - c = r$.

Figure 5 depicts the parameter regimes where each equilibrium is stable. There are two qualitatively different situations that depend on f vs K; if $f < K$, then autotomy does not sufficiently benefit predators and so for coexistence to occur, the predator death rate must be lower than it would have to be in the absence of autotomy. On the other hand, if $f > K$, then autotomy provides sufficient benefit to the predators that coexistence can occur with a higher predator death rate than it would in the absence of autotomy. In this latter case, autotomy "rescues" predator–prey coexistence by allowing it to occur for situations where the predator would otherwise go extinct.

Some characteristics of this system do not depend on f, however. First, given a nonzero kill rate k, there is always a low enough predator death rate μ such that coexistence occurs. Second, so long as autotomy is neither completely ineffective nor completely effective ($0 < r < 1$), there is always a low enough predator death rate such that coexistence with autotomizing focal prey occurs.

Eco-evolutionary dynamics

In the ecological model (Eq. 6), strategy frequencies of both predators and prey change due to the ecological dynamics alone. To combine evolution and ecology in this system, we turn our ecological system from one that focuses on changes

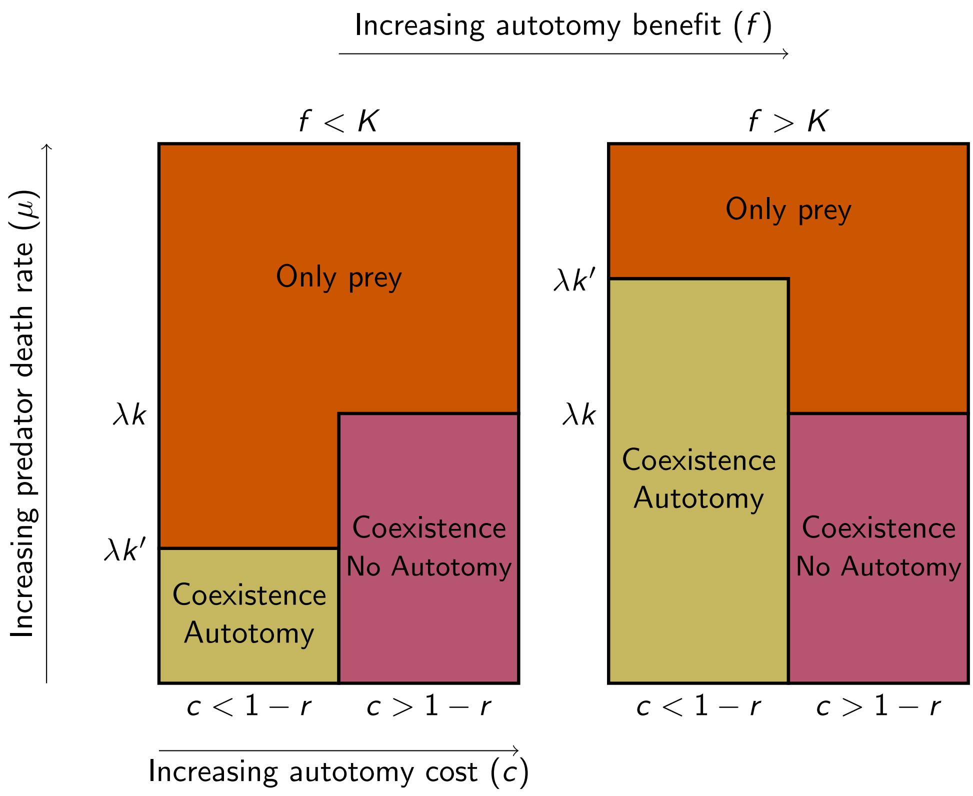

Fig. 5 Long-term outcomes of the ecological dynamics (Eq. 6). When autotomy benefit is small ($f < K$), coexistence is more difficult when autotomy is favored (at low cost c compared to effectiveness $1 - r$). When autotomy benefit is large ($f > K$), coexistence is easier when autotomy is favored. Difficulty of coexistence here is measured by the minimum value of predator death rate μ that allows for coexistence. k' is defined in Eq. 7

in strategy frequency to one that focuses on changes in population density and bring back as the evolutionary part a modified version of the replicator dynamics where payoffs are now dependent on population density as well as strategy frequency:

$$\begin{aligned}
\frac{dP}{dt} &= P[[k + (f - K)(1 - rk)z][\lambda - (1 - x)p]n + v(1 - x) - \mu] \\
\frac{dn}{dt} &= \beta(1 - cz)n(1 - n) - \gamma k[1 - z(1 - r)]nxP \\
\frac{dx}{dt} &= x(1 - x)[np[k + (f - K)(1 - rk)z] - v] \\
\frac{dz}{dt} &= z(1 - z)[kxP(1 - r) - c],
\end{aligned} \tag{11}$$

where $P = p_\alpha + p_v$ and $n = p_A + p_N$. A derivation and an analysis of this model can be found in Appendix D.

The equation for $\frac{dP}{dt}$ is not obvious and is constructed by noting that in our ecological model we have an equation for $\frac{dp_\alpha}{dt} = \frac{d(xP)}{dt}$, and we compute $\frac{dP}{dt}$ from that using the chain rule (see Appendix D). This formulation enables us to put no particular assumptions on the dynamics of the avoid strategy except for its evolutionary competition with the approach strategy.

There are eleven equilibria of this model and one non-equilibrium long-term solution. All five equilibria that can be stable and their stability conditions are listed in Table 5, where for ease of notation we substitute

$$\mu_V = \lambda k + p(1 - rk)(V - K). \tag{12}$$

Table 5 Descriptions and stability conditions for the five equilibria that can possibly be stable in the eco-evolutionary model (Eq. 11). The expressions for β_0, n_0, and n_1 are in Eq. 8, and the expression for μ_V is Eq. 12. The expression for $\tilde{P}$ is Eq. 15. The expression for $\tilde{n}$ is Eq. 16. The expression for $\tilde{z}$ is Eq. 56. Finally, c_0^* and c_1^* are defined in Eqs. 18 and 19, respectively

Description	$\hat{P}$	$\hat{n}$	$\hat{x}$	$\hat{z}$	Stability Conditions
Prey-only, no approach	0	1	0	0	$V > K$ and $\mu > \mu_v$
Prey-only, approach	0	1	1	0	$V < K$ and $\mu > \lambda k$
Coexistence, no autotomy	$\beta_0(1 - n_0)$	n_0	1	0	$n_0 < 1, \mu > \lambda \frac{v}{p}, c > c_0^*$
Coexistence, autotomy	$\beta_0(1 - n_1)(\frac{1-c}{r})$	n_1	1	1	$n_1 < 1, \mu > \lambda \frac{v}{p}$, and $c < c_1^*$
Coexistence, some autotomy	$\tilde{P}$	$\tilde{n}$	1	$\tilde{z}$	$\tilde{n} < 1, 0 < \tilde{z} < 1, \mu > \lambda \frac{v}{p}, c_0^* > c$, and $c > c_1^*$.

There are two potentially stable prey-only equilibria ($\hat{P} = 0$, $\hat{n} = 1$) that are differentiated by their equilibrium approach frequency $\hat{x}$; both have $\hat{z} = 0$ (as autotomy is unfavored when there are no predators). The first has no predator approach ($\hat{x} = 0$) and is stable only when alternative prey are valuable ($V > K$). The second has predator approach ($\hat{x} = 1$) and is stable only when alternative prey are not valuable ($V < K$). Both of these equilibria require the predator death rate to be relatively high, which makes sense because they are equilibria where the predators go extinct. The fact that we have a predator strategy frequency when we do not actually have any predators may seem a bit strange, but this result should be interpreted as a "ghost" of the transient dynamics leading to the approach toward the equilibrium. For instance, in the case where $\hat{x} = 1$, the predators eventually go extinct, but, during that process, the approach strategy is favored because alternative prey are not valuable ($V < K$).

The other possible long-term situation where there is no coexistence of predator and prey involves extinction of the prey population and runaway growth of the predator population, where the predator population is focused entirely on alternative prey. The runaway growth is a side-effect of our minimal-assumption dynamics for the predators using the avoid strategy; had we imposed a carrying capacity on them or explicitly modeled the alternative prey, the predators would not grow without bound. Thus, it is best to consider this situation as the case where the avoid strategy wins, with the specific dynamics irrelevant to the scope of this model. Such runaway growth occurs if the predator death rate is low, and this particular threshold is higher (i.e., easier to obtain) if alternative prey are valuable ($V > K$; note the large swaths of space labeled "no approach" in Fig. 6 for $V > K$).

There are three coexistence equilibria. The first has:

$$\begin{aligned}\hat{P} &= \beta_0(1-n_0)\\ \hat{n} &= n_0\\ \hat{x} &= 1\\ \hat{z} &= 0,\end{aligned} \tag{13}$$

and is the no-autotomy coexistence equilibrium.

Next, we have:

$$\begin{aligned}\hat{P} &= \beta_0\left(\frac{1-c}{r}\right)(1-n_1)\\ \hat{n} &= n_1\\ \hat{x} &= 1\\ \hat{z} &= 1,\end{aligned} \tag{14}$$

which is the autotomy coexistence equilibrium.

Finally, we have a coexistence equilibrium with an intermediate frequency of autotomy:

$$\hat{P} = \tilde{P} = \beta_0 \frac{1 - c\tilde{z}}{1-(1-r)\tilde{z}}(1-\tilde{n}) \tag{15}$$

$$\hat{n} = \tilde{n} = \frac{\mu}{\lambda[k + (f-K)(1-rk)\tilde{z}]}, \tag{16}$$

and $\hat{z} = \tilde{z}$ is the positive real root of the following quadratic equation:

$$\begin{aligned}0 &= a_2\tilde{z}^2 + a_1\tilde{z} + a_0\\ a_2 &= c\lambda(\beta_0 k - 1)(f-K)(1-r)(1-rk)\\ a_1 &= \beta_0 ck(1-r)[\mu - k\lambda] + \lambda[ck(1-r)\\ &\quad - c(f-K)(1-rk) + \beta_0 k(f-K)(1-r)(1-rk)]\\ a_0 &= \beta_0(k\lambda - \mu)(1-r) + ck\lambda.\end{aligned} \tag{17}$$

The full expression for $\tilde{z}$ can be found in Eq. 56 in Appendix D.

Stability conditions for these equilibria are shown in Table 5. The quantities c_0^* and c_1^* are

$$c_0^* = \beta_0(1-n_0)k(1-r) \tag{18}$$

$$c_1^* = \frac{\beta_0(1-n_1)k(1-r)}{\beta_0(1-n_1)k(1-r) + r}, \tag{19}$$

and are discussed in detail in Appendix D.

Figure 6 depicts the expected long-term outcomes of our eco-evolutionary model (Eq. 11). As with the ecological model (Eq. 6 and Fig. 5), predator–prey coexistence requires a sufficiently low predator death rate (μ). However, in this case, if the predator death rate is too low, we reach the no-approach runaway predator growth situation. As expected, autotomy is favored for high autotomy benefit (f) and low autotomy cost (c). There are also situations where both autotomy and no-autotomy coexistence equilibria are stable, as well as situations where other equilibria are stable at the same time as the autotomy coexistence equilibrium. Note that there are several other possibilities for the diagrams in Fig. 6, with different orderings of c_0^* and c_1^* (for $V < K$) or different orderings of μ_V, $\lambda k'$, and $\lambda \frac{v}{p}$ (for $K < V < f$). These other possibilities merely trade regions of multistability for regions where only one equilibrium is stable. We chose the particular options displayed in Fig. 6 to best display the general patterns without getting overwhelmed by detail.

The eco-evolutionary model provides two different effects on the existence of autotomy. First, it allows for a situation where autotomy is only present in part of the population, by contrast to the game-theoretical model, for which the mixed-strategy equilibrium was never stable. Second, and more importantly, it allows for autotomy that was too costly for the game-theory model ($c > K$) to exist, given certain ecological

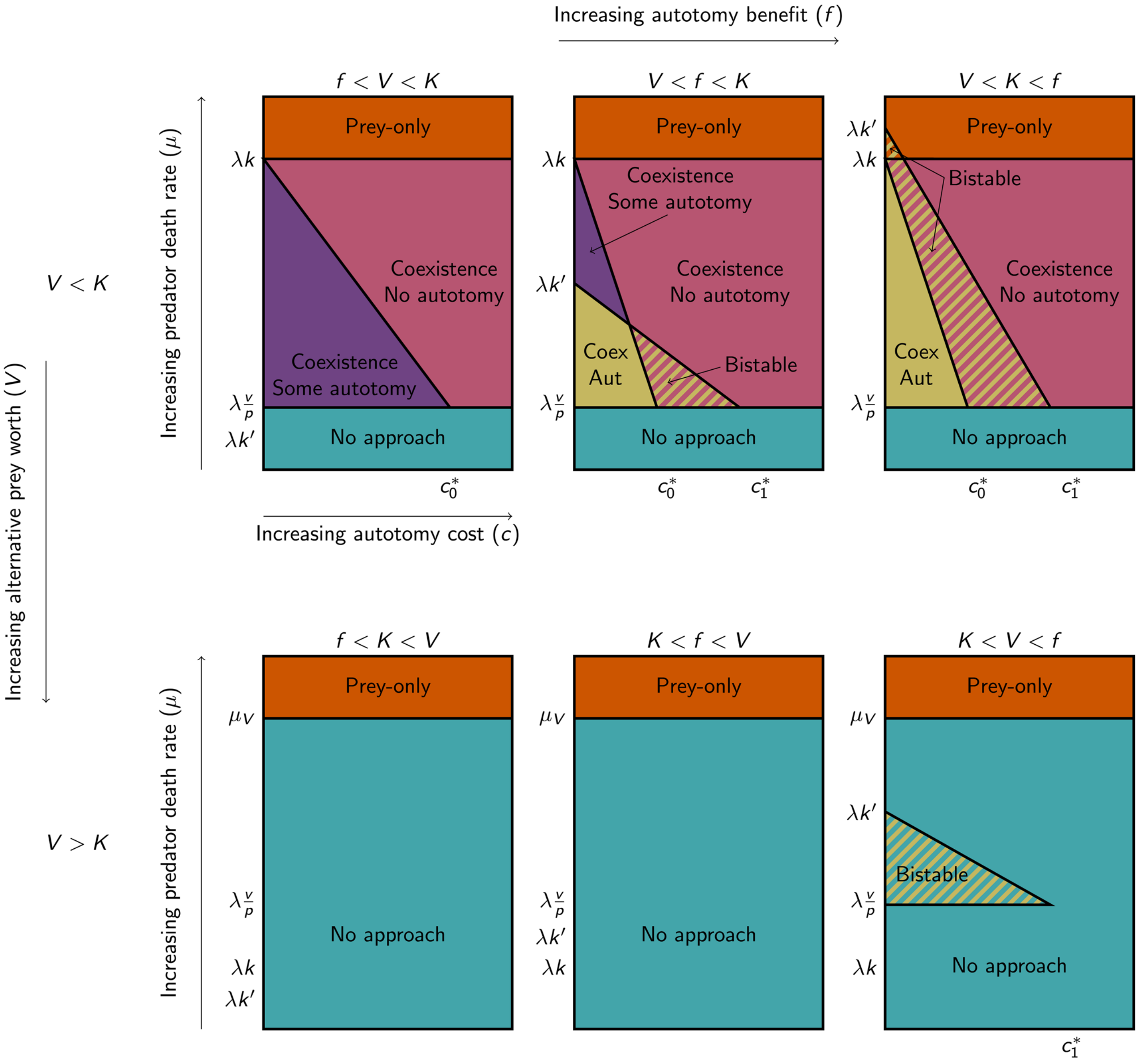

Fig. 6 Expected equilibrium outcomes for the eco-evolutionary dynamics (Eq. 11). As predator death rate (μ) increases, the system moves from a no-approach runaway predator nonequilibrium dynamic to predator–prey coexistence equilibria to a prey-only equilibrium. Coexistence equilibria with autotomy occur for low autotomy cost (c) and high autotomy benefit (f). If alternative prey are highly valuable ($V > K$), coexistence is only possible for high autotomy benefit and low autotomy cost. The quantities k', c_0^*, c_1^*, and μ_V are defined in Eqs. 7, 18, 19, and 12, respectively

parameters. In particular, we predict that costly autotomy is more likely to be present in prey with high growth rates or, perhaps non-intuitively, low predation rates. This conclusion arises from the fact that the cost-based stability condition for the autotomy coexistence equilibrium is

$$\begin{aligned}\hat{P} &> \frac{c}{k(1-r)}\\ \frac{\beta}{\gamma k}\left(\frac{1-c}{r}\right)\left[1-\frac{\mu}{\lambda[k+(f-K)(1-rk)]}\right] &> \frac{c}{k(1-r)},\end{aligned} \tag{20}$$

and so a sufficiently large β or sufficiently small γ will make this true. See Appendix D for details.

Comparison of model results

Our game-theoretical model outlined three "axes" that determine the predator–prey outcomes: cost of autotomy to the prey, benefit of autotomy to the predators, and the relative value of focal and alternative prey to the predators. Autotomy is favored when the prey cost is low, approach is favored when focal prey are more valuable than alternative prey, and autotomy + approach is favored when in particular autotomizing prey are more valuable than alternative prey.

While the game-theoretical model assumed that there were always predators and prey that could have either strategy, the ecological model approached the problem from a different angle: will there be predators or prey at all? In particular, under what circumstances can predators that approach the focal prey coexist with the prey? The ecological model revealed first that, given a low enough predator death rate, coexistence (and therefore the approach strategy) can be guaranteed. Second, coexistence with autotomy is always possible so long as autotomy has intermediate effectiveness. Finally, autotomy can "rescue" predator–prey coexistence by allowing predators with relatively higher death rates to exist; these results add a new element to the results obtained from the game-theoretical model: if the presence of autotomy can force the maintenance of the approach strategy through ecology alone, then the game-theoretical model underestimates the evolutionary viability of autotomy.

Combining our game-theoretical and ecological models into our eco-evolutionary model resulted in dynamics that both parallel our purely game-theoretical results and incorporate the additional insights from the ecological model.

In our game-theoretical model, autotomy was only a stable outcome for relatively high predator benefit f; lower predator benefit could at best result in a cycle. The main effect of adding ecology to this model turns this cycle into a stable equilibrium with an intermediate frequency of autotomy, which is in general a more viable approach for long-term maintenance of autotomy as it does not involve periods of extinction of the strategy. Otherwise, the predator benefit axis is unchanged.

The alternative prey benefit axis remains largely unchanged once ecology is added; autotomy is still viable mostly when alternative prey benefit V is low, and there is still the possibility of bistability when alternative prey benefit is high, but in the eco-evolutionary model, additional—relatively finicky—conditions related to ecological parameters and the initial conditions need to be met for this to be the case.

The biggest difference between the game-theoretical and eco-evolutionary models is the condition of the viability of autotomy from the prey-side, which was only possible for low cost ($c < K$) in the game-theoretical model. In the eco-evolutionary model, this condition is replaced by one that also incorporates ecological parameters in such a way that given any cost $c < 1$, we can freely choose a large enough prey growth rate β or small enough predation rate γ to allow for autotomy to be favored. Prey that grow fast or have a low predation rate can accommodate the cost of autotomy to their population density to an extent that predators are still motivated to approach them, which then creates a situation where autotomy is favored.

Empirical application

We close with an application to porcelain crabs (*Petrolisthes* sp. and *Pachycheles* sp.), who lose their limbs to avoid predation (Wasson et al. 2002; Wasson and Lyon 2005; Knope and Larson 2014).

In an experiment involving challenging porcelain crabs of species *Petrolisthes* with predatory crabs, Wasson et al. (2002) found that the probability of killing a non-autotomizing crab given predatory approach was $k = \frac{55}{114} = 0.48$. Given autotomy, 58 out of 59 prey escaped, resulting in an effective kill rate of $rk = \frac{1}{59} = 0.017$, so $r = \frac{0.017}{0.48} = 0.035$. The value of K in this system is therefore 0.47. Multiple studies reviewed in Juanes and Smith (1995) found anywhere from $c = 0.2 - 0.5$ in crustaceans. Almost all of this range would satisfy $c < K$ in this system, meaning that autotomy would be favorable for these crabs against predatory crabs.

In a different study that examined the stomach contents of four species of rockfish (*Sebastes* sp.) for both porcellanid crabs (*Petrolisthes* sp. and *Pachycheles* sp.) and brachyuran crabs (*Cancer* sp., *Lophopanopeus* sp., and *Paraxanthias* sp.), Knope and Larson (2014) found that out of 400 rockfish stomachs, there were 231 complete and 14 limb-only brachyuran samples and 22 complete and 12 limb-only porcellanid samples. Let us assume that a full skeleton is due to a failed escape by any prey, and a limb-only skeleton is due to a successful escape by autotomizing prey. For simplicity, let us also assume that the previously stated kill probability of $k = 0.48$ and autotomy efficacy of $r = 0.035$ are consistent across these crab species. We know that the total number of partial porcellanid skeletons P_P and full porcellanid skeletons F_P given the total number of approaches of porcellanids T_P must equal

$$\begin{aligned} P_P &= T_P(z_P)(1 - rk) = 12 \\ F_P &= T_P k(1 - (1 - r)z_P) = 22, \end{aligned} \tag{21}$$

so we can solve for z_P, the autotomy frequency in porcellanid crabs, to be

$$\begin{aligned} \frac{F_P}{k(1 - (1 - r)z_P)} &= \frac{P_P}{(1 - rk)z_P} \\ z_P &= \frac{kP_P}{F_P + kP_P - rk(F_P + P_P)} = 0.21. \end{aligned} \tag{22}$$

We can use the same approach to find that the autotomy frequency in brachyuran crabs is $z_B = 0.029$, which is consistent with the idea that brachyuran crabs autotomize much less readily than do porcellanid crabs (Knope and Larson 2014).

We can then back-calculate the number of approaches of porcellanids as

$$P_P = T_P(z_P)(1 - rk)$$
$$T_P = \frac{P_P}{z_P(1 - rk)} = 58. \tag{23}$$

We can use the same approach to find that the total number of approaches of brachyurans is $T_B = 491$, which means that $x = \frac{T_P}{T_B + T_P} = 0.11$. Note that it is not necessary that each fish make only one approach, so having more approaches than fish is not a problem.

We can plug these results into the mixed-strategy equilibrium using $\hat{x} = 0.11$ (Eq. 2) and $\hat{z} = 0.21$ (Eq. 3), assuming $f = 0.37$ (estimated from Lawton (1989)), and obtain $c = 0.051$ and $\frac{v}{p} = 0.46$, indicating an extremely low cost to the crab for autotomy and that the brachyuran crabs are worth a bit under half as much to the predator as are the porcelain crabs. The low cost result is consistent with empirical observations (Wasson et al 2002), and the relative worth result would require additional study to evaluate.

Discussion

We have derived and analyzed three separate models for the ecological and evolutionary dynamics of autotomy. First, we have studied a game-theoretical model where the prey can choose whether or not to employ autotomy and the predators can choose whether or not to pursue the prey. By studying the equilibria and the replicator dynamics of this system, we discovered three situations where autotomy is maintained in the population in the long term. First, if the predator is always motivated to approach the prey, then autotomy is favored as long as it provides some benefit to the prey. Second, if the predator is not motivated to approach non-autotomizing prey, but autotomy provides sufficient benefit to the predator, then the presence of autotomy encourages predators to approach when they would otherwise not do so. The prey can therefore "lock themselves" into performing autotomy. Finally, if the predator is motivated to approach non-autotomous prey, but autotomy does not provide sufficient benefit to the predator, and autotomy provides sufficient benefit to the prey, then the system will cycle through all possible strategy pairs.

Our second model is a purely ecological Lotka–Volterra model where the parameters are modified by terms that reflect the effects of autotomy on the predators and on the prey. Ignoring extreme cases, there are always ecological parameters that allow predator–prey coexistence in general and coexistence with autotomy in particular, which suggests that ecological dynamics can facilitate the existence of autotomy. Autotomy can also "rescue" coexistence by allowing predators to exist when they would have gone extinct without autotomy.

Our third model combines these first two models to produce eco-evolutionary dynamics of autotomy. This model has two major differences from either of the first two models alone. First, the cycle in the game-theoretical model is converted into an equilibrium with an intermediate frequency of autotomy, rendering it easier for autotomy to be maintained at an intermediate level over long timescales. Second, and more importantly, the cost of autotomy to the prey can in most circumstances be mitigated purely by ecological parameters. In particular, we expect costly autotomy to more likely occur in predator–prey systems with high prey growth rates or low predation rates, i.e., systems where the prey population density remains sufficiently high in the face of a high autotomy cost that predators are still incentived to approach the prey.

In applying our theoretical results to empirical data from porcelain crabs, we expect that autotomy should be common in these crabs due to the low cost compared to the benefit of autotomy for the crabs themselves, and we predict that porcelain crabs are significantly more valuable to some of their predators than alternative prey are to those predators.

In general, we find that the effects of autotomy on the ecological and evolutionary dynamics of predator–prey systems are highly variable, and that the ultimate outcomes depend very strongly on the specific parameter values of the system. Future work applying this model to empirical systems could involve measuring the values of these parameters and classifying existing autotomizing systems as likely or unlikely under our models, with the unlikely ones of particular interest for future study.

A decades-old problem in the study of autotomy in lizards is the ability to distinguish between predation pressure (xP in our model), predator efficiency (k in our model), and escape efficiency ($1 - r$ in our model) by looking at the prevalence of autotomy (Schoener 1979; Bateman and Fleming 2009). Models like ours that explicitly take into account all of these factors can help address these difficulties by allowing explicit predictions relating to all three factors.

In our game-theoretical model, the only time a purely distracting autotomy ($f = 0$) can be maintained in the long term as a result of natural selection is as part of a cycle. In real systems, however, prey may be playing the autotomy game with multiple different sets of predators, and so if each of these games would individually fall into this cycling regime but are out of phase with one another, it would be possible for purely distracting autotomy to be maintained in the long term. Considering multiple games with the same prey and different predators would therefore be an useful extension of our model that would allow exploration of the multi-predator hypothesis (Blumstein 2006) in the context of autotomy.

Our eco-evolutionary model adds to a rich tradition of models that combine ecological and evolutionary dynamics (Berryman 1992; Abrams 2000; Fussmann et al. 2007). Many of these models use a separation-of-time-scales assumption (e.g., (Marrow et al 1996; Cortez and Ellner 2010)), placing either the ecological or the evolutionary dynamics on a fast timescale and the other one on a slow timescale. One benefit of the simplicity of our eco-evolutionary model is the lack of need for this assumption. Of course, ecological and evolutionary dynamics can actually operate at different timescales, and it would be interesting to see how separation-of-timescales approaches yield different results from our approach.

Crucially, our model differs from previously studied models of antipredator behavior (e.g., (Ives and Dobson 1987; Lima 1998)) because, unlike strictly antipredator behaviors, autotomy has the possibility of conferring a benefit on the predator through ingestion of the lost part. It is this potential benefit that leads to the diverse evolutionary dynamics in our model. In fact, setting $f = 0$ would yield a general model of costly, strictly antipredator behavior.

Autotomy is known to increase prey vulnerability to subsequent predation (Maginnis 2006). Our model folds this effect into the effectiveness parameter r, which can be interpreted as an average across all population predation events for individuals that would choose to perform autotomy, whether or not they have actually done so prior to the predation event. One limitation of this method of parameterization is that cases where enough prey have performed autotomy for there to be an effective kill rate post-autotomy that is actually higher than the kill rate without autotomy require $r > 1$, which we do not consider in our analysis. Incorporating the approaches of Harris (1989) and Wilkinson (2003) that explicitly track autotomized vs. intact individuals would be useful to study the effect of increased vulnerability in more detail. Tracking autotomized vs. intact individuals would also allow for the incorporation of predator learning into this model, whereby predators may be able to learn either to approach the prey in general (if autotomy happens with sufficient frequency and benefit to be worthwhile for the predator) or at the very least learn to approach autotomized individuals in particular. Predator learning in this sense could provide additional pressure to reach an approach-autotomy equilibrium, as predators are further incentivized to approach the autotomized prey and the prey are further incentivized to develop a useful, minimally costly autotomy mechanism.

Another of the choices we made in formulating our models was incurring a cost of autotomy c even without predator approach. There are two possible justifications for this. First, there may be some sort of metabolic cost to maintaining the whole system of autotomy and regeneration ("allocation costs" in Maginnis (2006)). This justification would apply under all interpretations of the autotomy strategy (A), including "individuals that have the potential to choose whether or not to autotomize in the presence of a predator." Second, individuals who have autotomized incur well-documented costs on their future survival and reproduction that do not relate to future predator escape (Maginnis 2006). This justification would apply under all interpretations of the autotomy strategy (A) in a population-average sense similar to that discussed above for r, but it would especially apply in a more explicit sense if the strategy is interpreted as "individuals must autotomize in a potential encounter even if the predator chooses to avoid the prey." Our model is agnostic to these interpretations. Future models may consider making the cost of autotomy differ under varying levels of predation pressure.

Our models use the simplest possible biologically relevant ecological dynamics framework (Berryman 1992): a Lotka–Volterra model with a prey carrying capacity and a linear predator functional response. We chose this simple structure to establish baseline results for our model of autotomy. Future extensions of this work could include more complex ecological models, including the use of a nonlinear predator functional response. In addition, future work could explore the evolution of the parameters of the autotomy game.

Appendix

A Game theory analysis

Let $P(A|B)$ be the payoff of strategy A played against strategy B. The approach strategy (α) beats the avoid strategy (v) against the autotomy strategy (A) when

$$\begin{aligned} P(\alpha|A) &> P(v|A) \\ rkp + (1-rk)(fp) &> v \\ f &> \frac{\frac{v}{p} - rk}{1 - rk} = V, \end{aligned} \tag{24}$$

and against the no autotomy strategy (N) when

$$\begin{aligned} P(\alpha|N) &> P(v|N) \\ kp &> v \\ k &> \frac{v}{p} \\ \frac{k - rk}{1 - rk} &> \frac{\frac{v}{p} - rk}{1 - rk} \\ K &> V, \end{aligned} \tag{25}$$

with both of these calculations requiring $rk < 1$. If $rk = 1$ (that is, autotomy is completely ineffective and the kill is guaranteed), then both conditions collapse to $p > v$.

The autotomy strategy (A) beats the no autotomy strategy (N) against the approach strategy (α) when

$$\begin{aligned}(1-rk)(1-c) &> 1-k\\ c &< 1-\frac{1-k}{1-rk}=\frac{1-rk-1+k}{1-rk}=\frac{k-rk}{1-rk}=K.\end{aligned}$$

When $rk = 1$, this condition is impossible to satisfy.

The autotomy strategy can never beat the no autotomy strategy against the avoid strategy (v), as this would require

$$1-c>1, \tag{26}$$

which is never true.

A Nash equilibrium is defined as a pair of states (A, B) such that no player benefits from switching states (Myerson 2013). In our case, this means that $P(A|B) > P(C|B)$ and $P(B|A) > P(D|A)$, where C and D are the alternative strategies to A and B, respectively. Combining the results from Eqs. 24–26, we obtain that vN is a Nash equilibrium when $V > K$, vA is never a Nash equilibrium, αN is a Nash equilibrium when $V < K$ and $c > K$, and αA is a Nash equilibrium when $f > V$ and $c < K$. There is no pure Nash equilibrium when $f < V$ and $V < K$.

To compute a mixed strategy Nash equilibrium (x, z), where x is the frequency of playing α over v and z is the frequency of playing A over N, we must choose x to equalize the payoffs the prey gets from playing N or A:

$$\begin{aligned}P(N|x)&=P(A|x)\\ (1-x)(1)+x(1-k)&=(1-x)(1-c)+x(1-rk)(1-c)\\ x&=\frac{c}{k[1-(1-c)r]}.\end{aligned}\tag{27}$$

This solution must be between 0 and 1, and it is always greater than 0 so long as $c > 0$. We therefore only need

$$\begin{aligned}\frac{c}{k[1-(1-c)r]}&<1\\ c&<k[1-(1-c)r]\\ c&<k-rk+crk\\ c(1-rk)&<k-rk\\ c&<K.\end{aligned}\tag{28}$$

We must also choose z to equalize the payoffs the predator gets from playing v or α:

$$\begin{aligned}P(v|z)&=P(\alpha|z)\\ v&=(1-z)kp+z[rkp+(1-rk)fp]\\ z&=\frac{k-\frac{v}{p}}{k-kr-f(1-rk)}\\ z&=\frac{K-V}{K-f}=\frac{V-K}{f-K}.\end{aligned}\tag{29}$$

For z to be between 0 and 1, we require $V-K$ and $f-K$ to have the same sign and if that sign is positive, $f > V$, whereas if that sign is negative, $f < V$. These conditions yield $f > V > K$ or $f < V < K$.

B Replicator Dynamics

To obtain the replicator equations, we must compute mean fitnesses (payoffs) of the predators and prey. First, for the predator:

$$\begin{aligned}W_{\text{pred}}(x,z)&=(1-x)P(v|x,z)+xP(\alpha|x,z)\\ &=(1-x)v+x[kp(1-z)+[rkp+(1-rk)fp]z]\\ &=v+x(kp-v)+xz[p[f(1-rk)-(k-rk)]]\\ &=v+x[p(1-rk)(K-V)]+xz[p(1-rk)(f-K)]\\ &=v+p(1-rk)[x(K-V)+xz(f-K)]\\ &=v+xp(1-rk)[K-V+z(f-K)].\end{aligned}\tag{30}$$

For the prey, we have:

$$\begin{aligned}W_{\text{prey}}(x,z)&=(1-z)P(N|x,z)+zP(A|x,z)\\ &=(1-x)(1-z)(1)+(1-x)z(1-c)+x(1-z)\\ &\quad(1-k)+xz(1-rk)(1-c)\\ &=1-cz-kx+xz(k-rk+crk)\\ &=1-cz-kx+kxz[1-r(1-c)].\end{aligned}\tag{31}$$

Our system is two-dimensional, so we will use x and z as before to obtain the full replicator dynamics.

$$\begin{aligned}\frac{dx}{dt}&=x[P(\alpha|x,z)-W_{\text{pred}}(x,z)]\\ &=x[kp(1-z)+[rkp+(1-rk)fp]z-W_{\text{pred}}(x,z)]\\ &=x(1-x)[kp-v-zp[k-rk-f(1-rk)]]\\ &=x(1-x)p(1-rk)[z(f-K)+(K-V)]\\ \frac{dz}{dt}&=z[P(A|x,z)-W_{\text{prey}}(x,z)]\\ &=z[(1-rk)(1-c)x+(1-x)(1-c)-W_{\text{prey}}(x,z)]\\ &=z(1-z)[kx[1-(1-c)r]-c].\end{aligned}\tag{32}$$

The equilibria are obtained by setting each equation to 0 and solving for x and z. Setting $\frac{dx}{dt}=0$ yields $\hat{x}=0$, $\hat{x}=1$, or $\hat{z}=\frac{V-K}{f-K}$. Setting $\frac{dz}{dt}=0$ yields $\hat{z}=0$, $\hat{z}=1$, or $\hat{x}=\frac{c}{k[1-(1-c)r]}$. The equilibria can be obtained by choosing one result from the first set and a result for the other variable from the second set.

To obtain the stability of these equilibria, we must compute the Jacobian matrix of this system and calculate its eigenvalues, evaluated at each equilibrium. The equilibrium is stable if and only if all eigenvalues have negative real part (Guckenheimer and Holmes 1990).

The Jacobian matrix of this system is

$$\mathbf{J} = \begin{bmatrix} p(1-rk)(1-2x)[z(f-K)+(K-V)] & p(1-rk)x(1-x)(f-K) \\ z(1-z)k[1-(1-c)r] & (1-2z)[kx[1-(1-c)r]-c]. \end{bmatrix}$$

For (0, 0), the eigenvalues are $-c$ and $p(1-rk)(K-V)$, so (0, 0) is stable when $V > K$ and a saddle point otherwise.

For (1, 0), the eigenvalues are $k - rk - c(1-rk)$ and $p(1-rk)(V-K)$, so (1, 0) is stable when $V < K$ and $c > K$, a saddle point when $V < K$ and $c < K$ or $V > K$ and $c > K$, and unstable otherwise.

For (0, 1), the eigenvalues are c and $p(1-rk)(f-V)$, so (0, 1) is a saddle point if $f < V$ or unstable otherwise.

For (1, 1), the eigenvalues are $c(1-rk)-(k-rk)$ and $p(1-rk)(V-f)$, so (1, 1) is stable when $c < K$ and $f > V$. If $c > K$ and $f > V$, or if $c < K$ and $f < V$, then (1, 1) is a saddle point, and if $c > K$ and $f < V$, then it is unstable.

For the interior equilibrium, it is easier to look at the trace and determinant instead of the eigenvalues. The trace is 0. The determinant is

$$\frac{-pc(1-rk)[k-rk-c(1-rk)](f-V)(V-K)}{k(f-K)[1-(1-c)r]}. \tag{33}$$

Now, $\hat{z}$ exists under either $f > V > K$ or $f < V < K$. In the former case, the determinant is negative, and so we have a saddle point. In the latter case, the determinant is positive, so we have a center.

Our autotomy game is a specific example of a two-player, bimatrix game, the general analysis of which can be found in, e.g., Hofbauer and Sigmund (1998) and Cressman et al. (2003).

C Ecological Dynamics

Consider a standard Lotka–Volterra predator–prey model (Hofbauer and Sigmund 1998):

$$\begin{aligned} \frac{dP}{dt} &= \lambda nP - \mu P \\ \frac{dn}{dt} &= \beta n(1-n) - \gamma nP, \end{aligned} \tag{34}$$

where P is predator population size, n is prey population size as a fraction of carrying capacity, λ is the predation effect for the predator, μ is predator death rate, β is prey growth rate, and γ is predation effect for the prey.

First of all, we note that because there is no strategy-switching through mutation in a purely ecological model, we treat each strategy as a separate population. Also, the avoiding predators are completely resource-partitioned with the approaching predators, and so there is no ecological interaction between the two strategies. We can therefore ignore avoiding predators for now.

Tracking only approaching predators, autotomizing prey, and non-autotomizing prey as three separate populations results in the following system:

$$\begin{aligned} \frac{dp_\alpha}{dt} &= \lambda[kp_N + [k+(f-K)(1-rk)]p_A]p_\alpha - \mu p_\alpha \\ \frac{dp_A}{dt} &= \beta(1-c)p_A(1-p_A-p_N) - \gamma rkp_Ap_\alpha \\ \frac{dp_N}{dt} &= \beta p_N(1-p_A-p_N) - \gamma kp_Np_\alpha, \end{aligned} \tag{35}$$

where for the predators we factor out the kill rate k from the growth rate λ and then modify the effective kill rate for an interaction with an autotomizing predator $(k+(f-K)(1-rk) = rk+(1-rk)f)$, where the contribution to the growth rate for a kill is 1 and for a failed kill is f. For the autotomizing prey, the growth rate is decreased by the autotomy cost c and the predation rate is decreased by the reduction in kill rate due to autotomy r.

Setting $\frac{dp_\alpha}{dt} = 0$ yields $\hat{p}_\alpha = 0$ or $\mu = \lambda k\hat{p}_N + \lambda[k+(f-K)(1-rk)]\hat{p}_A$. Setting $\frac{dp_A}{dt} = 0$ yields $\hat{p}_A = 0$ or $\hat{p}_\alpha = \frac{\beta(1-c)(1-\hat{p}_A-\hat{p}_N)}{\gamma rk}$. Setting $\frac{dp_N}{dt} = 0$ yields $\hat{p}_N = 0$ or $\hat{p}_\alpha = \frac{\beta(1-\hat{p}_A-\hat{p}_N)}{\gamma k}$.

There are therefore five possible equilibria of this model: total extinction ($\hat{p}_\alpha = \hat{p}_A = \hat{p}_N = 0$), prey-only ($\hat{p}_\alpha = 0$, $\hat{p}_A + \hat{p}_N = 1$, and initial conditions matter), coexistence with all-autotomy ($\hat{p}_\alpha = \frac{\beta}{\gamma k}\frac{1-c}{r}[1 - \frac{\mu}{\lambda[k+(f-K)(1-rk)]}]$, $\hat{p}_A = \frac{\mu}{\lambda[k+(f-K)(1-rk)]}$, $\hat{p}_N = 0$), coexistence with no autotomy ($\hat{p}_\alpha = \frac{\beta}{\gamma k}[1-\frac{\mu}{\lambda k}]$, $\hat{p}_A = 0$, $\hat{p}_N = \frac{\mu}{\lambda k}$), and coexistence with some autotomy. This last equilibrium requires $r = 1-c$, and that makes the latter two equations the same in Eq. 35 w/r/t equilibria, so the system is underdetermined and initial conditions matter.

The Jacobian of this system is

$$\begin{bmatrix} \lambda[kp_N + [k+(f-K)(1-rk)]p_A] - \mu & \lambda[k+(f-K)(1-rk)]p_\alpha & \lambda kp_\alpha \\ -\gamma rkp_A & \beta(1-c)(1-2p_A-p_N) - \gamma rkp_\alpha & -\beta(1-c)p_A \\ -\gamma kp_N & -\beta p_N & \beta(1-p_A-2p_N) - \gamma kp_\alpha. \end{bmatrix}$$

The eigenvalues of the Jacobian at the total extinction equilibrium are $-\mu$, β, and $\beta(1-c)$, so it is always a saddle point.

The eigenvalues of the no-autotomy coexistence equilibrium are

$$\begin{aligned}
\lambda_1 &= \frac{\beta}{\lambda k}(\lambda k - \mu)(1 - c - r) \\
\lambda_2 &= \frac{-\beta\mu - \sqrt{\beta^2\mu^2 - 4\lambda k(\lambda k - \mu)}}{2\lambda k} \\
\lambda_3 &= \frac{-\beta\mu + \sqrt{\beta^2\mu^2 - 4\lambda k(\lambda k - \mu)}}{2\lambda k}
\end{aligned} \tag{36}$$

Note that this equilibrium exists only when $\mu < \lambda k$, so $\lambda_1 < 0 \iff c > 1 - r$, and $\lambda_2, \lambda_3 < 0$. Thus, this equilibrium is stable when it exists and $c > 1 - r$.

The eigenvalues of the autotomy coexistence equilibrium are

$$\begin{aligned}
\lambda_1 &= \frac{\beta}{\lambda r[f + (1-f)rk]}[\mu - \lambda[f + (1-f)rk]](1 - c - r) \\
\lambda_2 &= \frac{-\beta\mu(1-c) - \sqrt{\beta^2\mu^2(1-c)^2 - 4\beta(1-c)\lambda\mu[f + (1-f)rk][\lambda[f + (1-f)rk] - \mu]}}{2\lambda[f(1 - rk) + rk]} \\
\lambda_3 &= \frac{-\beta\mu(1-c) + \sqrt{\beta^2\mu^2(1-c)^2 - 4\beta(1-c)\lambda\mu[f + (1-f)rk][\lambda[f + (1-f)rk] - \mu]}}{2\lambda[f(1 - rk) + rk]}
\end{aligned} \tag{37}$$

Note that this equilibrium exists only when $\mu < \lambda[rk + (1-f)rk] = [f + (1-f)rk]$, so $\lambda_1 < 0 \iff c < 1 - r$, and $\lambda_2, \lambda_3 < 0$. Thus, this equilibrium is stable when it exists and $c < 1 - r$.

The eigenvalues of the Jacobian for a prey-only equilibrium $(0, \hat{p}_A, 1 - \hat{p}_A)$ are $-\beta(1 - c\hat{p}_A)$, $\lambda[k + \hat{p}_A(f - K)(1 - rk)] - \mu$, and 0.

If $\mu < \lambda[k + \hat{p}_A(f - K)(1 - rk)]$, then this equilibrium is unstable. If $\mu > \lambda[k + \hat{p}_A(f - K)(1 - rk)]$, then the leading eigenvalue of this equilibrium is 0, which indicates (in this case) null stability, as initial conditions matter. Note that this condition leaves the possibility for bistability between a coexistence equilibrium and a prey-only equilibrium for $\lambda k' < \mu < \lambda k$ when $f < K$ or $\lambda k < \mu < \lambda k'$ when $f > K$, depending on the equilibrium value of $\hat{p}_A$ in the prey-only equilibrium. For simplicity, we leave this possibility out of Fig. 5.

D Eco-evolutionary Dynamics

Incorporating the evolutionary dynamics into this model cannot entail merely importing Eq. 4 into Eq. 35, as that system assumes a constant population size for both predator and prey and is unable to incorporate density-dependence as it is currently formulated, and it is this density-dependence that allows the ecological model to interact with the evolution of the system in the first place.

We first modify Eq. 4 by noting that from the prey perspective, we have modeled predator interaction at a fixed rate, with differing outcomes based on the different encountered predator strategies. Increasing the total number of predators present will increase the encounter rate, but the relative fraction of each predator strategy will still determine the evolution of the prey strategy. This is unsatisfying because of the change in perspective from our game-theoretic per-encounter model to our eco-evolutionary continuous-time model. In particular, under a per-encounter model, a predator density of 0 would mean that autotomy would be neutral compared to no autotomy, because there are no interactions, where "Pa" refers to "payoff":

$$\begin{aligned}
W_{\text{prey}}(x, z, n, P) &= (1 - z)\text{Pa}(N|x, z, n, P) + z\text{Pa}(A|x, z, n, P) \\
&= (1 - z)(1 - x) \\
&\quad P(1) + (1 - z)xP(1 - k) + z(1 - x)P(1 - c) \\
&\quad + zxP(1 - rk)(1 - c) \\
&= P[1 - cz - kx + kxz[1 - r(1 - c)]] \\
\frac{dz}{dt} &= z[\text{Pa}(A|x, z, n, P) - W_{\text{prey}}(x, z, n, P)] \\
&= z[(1 - rk)(1 - c)xP + (1 - c)(1 - x)P \\
&\quad - W_{\text{prey}}(x, z, n, P)] \\
&= z(1 - z)P[kx[1 - (1 - c)r] - c],
\end{aligned} \tag{38}$$

and so $P = 0 \rightarrow \frac{dz}{dt} = 0$.

A more biologically reasonable approach would be to note that costs of autotomy are often incurred over time rather than per-encounter (with examples including metabolic costs of maintaining the autotomy/regeneration system and downstream effects of losing the body part and of regeneration that are not related to predator–prey encounters). Thus, we derive the equation for $\frac{dz}{dt}$ in a slightly different way to accommodate this continuous-time perspective. Formally, if $p_N(t)$ is the number of non-autotomizing individuals over time and $p_A(t)$ is the number of autotomizing individuals over time with per-capita mortality rates μ_N and μ_A, so that in isolation $p_i(t) = p_i(0)e^{-\mu_i t}$ for $i = N, A$, then if $z = \frac{p_A}{p_A + p_N}$, we have

$$\frac{dz}{dt} = z(1 - z)(\mu_N - \mu_A). \tag{39}$$

We have to slightly depart from our game (Table 2) in order to accommodate costs accrued over time with predation still occurring per-encounter. Non-autotomizing prey still experience a mortality rate of k per encounter, so that the per-capita mortality rate over time is kxP. However, autotomizing prey now experience a mortality rate of rk per-encounter, with the effect of the fixed cost c applied over time, so that the per-capita mortality rate over time is $rkxP + c$. Plugging these results into Eq. 39 yields

$$\frac{dz}{dt} = z(1 - z)[kxP(1 - r) - c]. \tag{40}$$

Under this model, $P = 0$ would lead to a loss of autotomy, which makes biological sense.

For the predators, the effect of density is due to different pools of prey. Only the payoff of the approach strategy is affected by focal prey density. Incorporating prey density amounts to multiplying the per-interaction prey value p by the prey density n (where "Pa" refers to "payoff"):

$$\begin{aligned} W_{\text{pred}}(x, z, n, P) &= (1 - x)\text{Pa}(v|x, z, n, P) + x\text{Pa}(\alpha|x, z, n, P) \\ &= (1 - x)v + x[kp(1 - z)n \\ &\quad + [rkp + (1 - rk)fp]zn] \\ &= v + x[kpn - v] + xzpn(1 - rk)(f - K) \\ \frac{dx}{dt} &= x[\text{Pa}(\alpha|x, z, n, P) - W_{\text{pred}}(x, z, n, P)] \\ &= x[kp(1 - z)n + [rkp + (1 - rk)fp]zn \\ &\quad - W_{\text{pred}}(x, z, n, P)] \\ &= x(1 - x)[np[f(1 - rk)z + k[1 \\ &\quad - (1 - r)z]] - v]. \end{aligned} \tag{41}$$

To obtain the dynamics of the prey population, we note that $n = p_A + p_N$ so $\frac{dn}{dt} = \frac{dp_A}{dt} + \frac{dp_N}{dt}$, so we add together the two equations from Eq. 6 into a single equation:

$$\frac{dn}{dt} = \beta(1 - cz)n(1 - n) - \gamma k[1 - z(1 - r)]np_\alpha, \tag{42}$$

where the fraction of the population that performs autotomy ($z = \frac{p_A}{p_A+p_N}$) becomes a trait of the population.

Obtaining the dynamics of the predator population is a bit more involved because we only have ecological dynamics for the approaching predators, but we want to model the strategy distribution of the predators ($x = \frac{p_\alpha}{p_\alpha+p_v} = \frac{p_\alpha}{P}$). We can start by noting that $p_\alpha = xP$, and so our equation for $\frac{dp_\alpha}{dt}$ in Eq. 6 is tracking the rate of change of the composite quantity xP over time. Because in the eco-evolutionary model, we are tracking the evolution of x separately, we must disentangle the rate of change of P from the rate of change of x:

$$\begin{aligned} \frac{d(xP)}{dt} &= x\frac{dP}{dt} + P\frac{dx}{dt} \\ \frac{dP}{dt} &= \frac{1}{x}\left[\frac{d(xP)}{dt} - P\frac{dx}{dt}\right] \\ &= \frac{1}{x}[\lambda[kp_N + [rk + (1 - rk)f]p_A]xP - \mu xP \\ &\quad - P[x(1 - x)[np[k + (f - K)(1 - rk)z] - v]]] \\ &= P[\lambda n[k + (f - K)(1 - rk)z] - \mu - [(1 - x) \\ &\quad [np[k + (f - K)(1 - rk)z] - v]]] \\ &= P[[k + (f - K)(1 - rk)z][\lambda - (1 - x)p]n \\ &\quad + v(1 - x) - \mu]. \end{aligned} \tag{43}$$

Note that this transformation does not assume anything about the avoiding predator dynamics (except for their ecological distinctness from the approaching predators) and instead allows those dynamics to flow naturally from the interactions between their payoffs and the payoffs of the approaching predators.

Our full system of equations is therefore:

$$\begin{aligned} \frac{dP}{dt} &= P[[k + (f - K)(1 - rk)z][\lambda - (1 - x)p]n + v(1 - x) - \mu] \\ \frac{dn}{dt} &= \beta(1 - cz)n(1 - n) - \gamma k[1 - z(1 - r)]nxP \\ \frac{dx}{dt} &= x(1 - x)[np[k + (f - K)(1 - rk)z] - v] \\ \frac{dz}{dt} &= z(1 - z)[kxP(1 - r) - c]. \end{aligned} \tag{44}$$

Setting $\frac{dP}{dt} = 0$ yields $\hat{P} = 0$, or $\hat{n} = \frac{\mu - v(1-\hat{x})}{[\lambda-(1-\hat{x})p][k+(f-K)(1-rk)\hat{z}]}$. Setting $\frac{dn}{dt} = 0$ yields $\hat{n} = 0$ or $\hat{P} = \frac{\beta}{\gamma k}\frac{1-c\hat{z}}{1-(1-r)\hat{z}}\frac{1-\hat{n}}{\hat{x}}$. Setting $\frac{dx}{dt} = 0$ yields $\hat{x} = 0$, $\hat{x} = 1$, or $\hat{z} = \frac{\frac{v}{\hat{n}p} - k}{(f-K)(1-rk)}$. Finally, setting $\frac{dz}{dt} = 0$ yields $\hat{z} = 0$, $\hat{z} = 1$, or $\widehat{xP} = \frac{c}{k(1-r)}$.

The Jacobian matrix of this system is

$$\begin{bmatrix} k'_z[\lambda - (1 - x)p]n + v(1 - x) - \mu & Pk'_z[\lambda - (1 - x)p] & P[k'_z pn - v] & P[[(f - K)(1 - rk)][\lambda - (1 - x)p]n] \\ -\gamma k[1 - z(1 - r)]nx & \beta(1 - cz)(1 - 2n) - \gamma k[1 - z(1 - r)]xP & -\gamma k[1 - z(1 - r)]nP & \beta(-c)n(1 - n) + \gamma k(1 - r)nxP \\ 0 & x(1 - x)pk'_z & [npk'_z - v](1 - 2x) & x(1 - x)[np(f - K)(1 - rk)] \\ z(1 - z)kx(1 - r) & 0 & z(1 - z)kP(1 - r) & [kxP(1 - r) - c](1 - 2z) \end{bmatrix}$$

Let us first dispense with the case where $\hat{P} = \hat{n} = 0$. Eigenvalues of the Jacobian for these equilibria are shown in Table 6.

All equilibria with $\hat{P} = \hat{n} = 0$ are saddle points, and none are stable, which makes sense because under our prey

Table 6 Eigenvalues of the Jacobian for all equilibria with $\hat{P} = \hat{n} = 0$

$\hat{P}$	$\hat{n}$	$\hat{x}$	$\hat{z}$	λ_1	λ_2	λ_3	λ_4
0	0	0	0	β	$-c$	$-v$	$v - \mu$
0	0	1	0	β	$-c$	v	$-\mu$
0	0	0	1	$\beta(1-c)$	c	$-v$	$v - \mu$
0	0	1	1	$\beta(1-c)$	c	v	$-\mu.$

dynamics the absence of predation means we should expect the prey to reach their carrying capacity.

The only other option for when $\hat{P} = 0$ is $\hat{n} = 1$. Eigenvalues of the Jacobian for all equilibria with $\hat{P} = 0$ and $\hat{n} = 1$ are shown in Table 7.

There is one other non-coexistence case, where $\hat{P} > 0$ but $\hat{n} = 0$. It is actually more straightforward to analyze this situation from a less rigorous perspective. Setting $\hat{n} = 0$ yields a closed system of equations for P and x:

$$\begin{aligned} \frac{dP}{dt} &= P[v(1-x) - \mu] \\ \frac{dx}{dt} &= -vx(1-x). \end{aligned} \tag{45}$$

that can be solved as:

$$\begin{aligned} P(t) &= P_0 e^{-\mu t}[x_0 + (1-x_0)e^{vt}] \\ x(t) &= \frac{x_0}{x_0 + (1-x_0)e^{vt}}, \end{aligned} \tag{46}$$

where $x \to 0$ as $t \to \infty$ and $P \to 0$ if $\mu > v$ and $P \to \infty$ if $\mu < v$. The latter case is the no-approach dynamic with runaway predator growth.

So the relevant equilibrium here is $\hat{P} = 0$ and $\hat{x} = 0$, which we have already discussed above as a total-extinction equilibrium. Note that because in general $n > 0$ during this process, we cannot necessarily use $\mu < v$ as a condition for this situation as the effect of p is artificially eliminated by setting n to 0.

We can now study the coexistence equilibria where $\hat{P} > 0$ and $0 < \hat{n} \leq 1$.

First, we note that $\hat{x} = 0 \to \hat{n} = 1$ and $\hat{z} = 0$ or $\hat{z} = 1$. Equilibrium then requires $\mu - v = (\lambda - p)[k + (f - K)(1 - rk)\hat{z}]$, which is a boundary case with free choice of $\hat{P}$. So no equilibrium with $\hat{x} = 0$ exists that leads to coexistence.

Now, if $\hat{x} = 1$ and $\hat{z} = 0$, we have as eigenvalues

$$\begin{aligned} \lambda_1 &= -c + \hat{P}k(1-r) \\ \lambda_2 &= v - \frac{\mu p}{\lambda} \\ \lambda_3 &= \frac{-\beta\mu + \sqrt{(\beta\mu)^2 - 4\beta\mu\lambda k(\lambda k - \mu)}}{2k\lambda} \\ \lambda_4 &= \frac{-\beta\mu - \sqrt{(\beta\mu)^2 - 4\beta\mu\lambda k(\lambda k - \mu)}}{2k\lambda} < 0. \end{aligned} \tag{47}$$

The condition for $\lambda_1 < 0$ is

$$\begin{aligned} 0 &> \lambda_1 \\ c &> \hat{P}k(1-r) \\ \hat{P} &< \frac{c}{k(1-r)}. \end{aligned} \tag{48}$$

Because existence of this equilibrium requires $\mu < \lambda k$, $\lambda_3 < 0$.

The other condition required for stability of this equilibrium is $\lambda_2 < 0 \iff \mu > \lambda\frac{v}{p}$.

If $\hat{x} = 1$ and $\hat{z} = 1$, we have as eigenvalues

$$\begin{aligned} \lambda_1 &= c - \hat{P}k(1-r) \\ \lambda_2 &= v - \frac{\mu p}{\lambda} \\ \lambda_3 &= \frac{-\beta\mu(1-c) + \sqrt{(\beta\mu(1-c))^2 - 4\beta\mu(1-c)\lambda[k + (f-K)(1-rk)][\lambda[k + (f-K)(1-rk)] - \mu]}}{2\lambda[k + (f-K)(1-rk)]} \\ \lambda_4 &= \frac{-\beta\mu(1-c) - \sqrt{(\beta\mu(1-c))^2 - 4\beta\mu(1-c)\lambda[k + (f-K)(1-rk)][\lambda[k + (f-K)(1-rk)] - \mu]}}{2\lambda[k + (f-K)(1-rk)]} < 0. \end{aligned} \tag{49}$$

Table 7 Eigenvalues of the Jacobian for all equilibria with $\hat{P} = 0$ and $\hat{n} = 1$. μ_V is defined in Eq. 12

$\hat{P}$	$\hat{n}$	$\hat{x}$	$\hat{z}$	λ_1	λ_2	λ_3	λ_4	Stability Conditions
0	1	0	0	$-\beta$	$-c$	$kp - v$	$v - \mu + k(\lambda - p)$	$V > K$ and $\mu > \mu_V$
0	1	1	0	$-\beta$	$-c$	$v - kp$	$k\lambda - \mu$	$V < K$ and $\mu > k\lambda$
0	1	0	1	$-\beta(1-c)$	c	$p[k + (f-K)(1-rk)] - v$	$v - \mu + (\lambda - p)[k + (f-K)(1-rk)]$	Always saddle point
0	1	1	1	$-\beta(1-c)$	c	$v - p[k + (f-K)(1-rk)]$	$-\mu + \lambda[k + (f-K)(1-rk)]$	Always saddle point

The condition for $\lambda_1 < 0$ is $\hat{P} > \frac{c}{k(1-r)}$. The condition for $\lambda_2 < 0$ is $\mu > \lambda\frac{v}{p}$. Because the existence condition for this equilibrium is $\mu < \lambda[k + (f - K)(1 - rk)]$, $\lambda_3 < 0$ always.

There are two critical values of c that are obtained from the cost-based conditions for the two previous coexistence equilibria. First, for the coexistence, no-autotomy equilibrium:

$$\begin{aligned} \hat{P} &< \frac{c}{k(1-r)} \\ \beta_0(1-n_0) &< \frac{c}{k(1-r)} \\ c &> k(1-r)\beta_0(1-n_0) = c_0^*. \end{aligned} \tag{50}$$

This condition is only possible if $c_0^* < 1$, which occurs iff

$$\beta_0 < \frac{c}{k(1-r)(1-n_0)}. \tag{51}$$

Note that $c_0^* \geq 0$ is always true.

Second, for the coexistence, autotomy equilibrium:

$$\begin{aligned} \hat{P} &> \frac{c}{k(1-r)} \\ \beta_0(1-n_1)\frac{1-c}{r} &> \frac{c}{k(1-r)} \\ c &< \frac{\beta_0(1-n_1)k(1-r)}{\beta_0(1-n_1)k(1-r)+r} = c_1^*. \end{aligned} \tag{52}$$

If $\beta_0 = 0$, $c_1^* = 0$, and as $\beta_0 \to \infty$, $c_1^* \to 1$.

First of all, under what circumstances can we choose β_0 such that $c < c_1^*$ (in which case ecological characteristics can overcome the cost of autotomy?) Rewriting the equation for $c < c_1^*$ in terms of β_0 yields

$$\beta_0 > \frac{rc}{k(1-r)(1-c)(1-n_1)}. \tag{53}$$

So we can always choose a β_0 large enough that $c < c_1^*$.

Next, under what circumstances can both these equilibria coexist? We would need both conditions for β_0 (Eqs. 51 and 53) to be true, which requires

$$\frac{r}{1-c} < \frac{1-n_1}{1-n_0}. \tag{54}$$

If $f < K$, $n_1 > n_0$, so we need $c < 1 - r$ for bistability over a range of β_0. If $f > K$, $n_1 < n_0$, so we can have bistability for some values of $c > 1 - r$ as well. Under circumstances where bistability is impossible, we would expect the following intermediate autotomy equilibrium to take over: $\hat{x} = 1$ but $0 < \hat{z} < 1$. In this case, setting $\frac{dP}{dt} = 0$ yields $\hat{n} = \frac{\mu}{\lambda[k+(f-K)(1-rk)\hat{z}]}$. Setting $\frac{dz}{dt} = 0$ yields $\hat{P} = \frac{c}{k(1-r)}$. Setting $\frac{dn}{dt} = 0$ yields $\hat{P} = \frac{\beta}{\gamma k}\frac{1-c\hat{z}}{1-(1-r)\hat{z}}[1 - \frac{\mu}{\lambda[k+(f-K)(1-rk)\hat{z}]}]$, which we can use to solve for $\hat{z}$.

Before we solve for $\hat{z}$, let us consider the function

$$g(z) = \beta_0\frac{1-cz}{1-(1-r)z}\left(1 - \frac{\mu}{\lambda k'_z}\right), \tag{55}$$

noting that $\hat{z}$ is the intersection of $g(z)$ with the line $z = \frac{c}{k(1-r)} = C$.

Because the only singularity of $g(z)$ is at $z = \frac{1}{1-r} > 1$, $g(z)$ is a continuous function over [0, 1]. The intermediate value theorem tells us that $g(z) = C$ for some z over the interval [0, 1] if and only if either $g(0) < C < g(1)$ or $g(0) > C > g(1)$. The first case implies that both other coexistence equilibria are unstable, if they exist, and the second case implies that both are stable, if they exist. In the first case, we should expect this intermediate equilibrium to be stable, and in the second case, we should expect this intermediate equilibrium to be unstable and separate the basins of attraction of the two stable equilibria. These conditions are equivalent to saying $c_0^* < c < c_1^*$ or $c < c_0^*$ and $c > c_1^*$, respectively.

Setting the two values of $\hat{P}$ equal to each other leads to the quadratic equation in Eq. 17, the roots of which are

$$\begin{aligned} \tilde{z} &= \frac{-W \pm \sqrt{X}}{Y} \\ W &= -\beta_0 ck\mu(1-r) - \lambda[\beta_0 k(1-r)[(f-K)(1-rk) - ck] \\ &\quad + c(f-K)(1-rk)] \\ X &= \beta_0\lambda k(f-K)(1-r)(1-rk) + 2kY[c\lambda + \beta_0(\mu - k\lambda) \\ &\quad (1-r)] + c(1-rk)[\beta_0(\mu - k\lambda K)) - \lambda f] \\ Y &= 2\lambda c(f-K)(1-rk)(1-r)(1-\beta_0 k). \end{aligned} \tag{56}$$

The value of $\tilde{z}$ is the positive one of the two roots that is also less than 1, provided such a root exists.

The eigenvalues of the Jacobian for this equilibrium are

$$\lambda_1 = v - \frac{\mu p}{\lambda}, \tag{57}$$

and the roots of a cubic polynomial that is somewhat intractable, and so we shall restrict our analysis of this equilibrium to numerical explanation, aside from the following condition for $\lambda_1 < 0$: $\mu > \lambda\frac{v}{p}$.

Finally, substituting $\hat{z} = \frac{\frac{v}{\hat{n}p} - k}{(f-K)(1-rk)}$ (which is the case when $0 < \hat{x} < 1$) into the expression for $\hat{n}$ yields:

$$\begin{aligned} \hat{n} &= \hat{n}\frac{p[\mu - v(1-\hat{x})]}{v[\lambda - p(1-\hat{x})]} \\ \mu &= \lambda\frac{v}{p} \end{aligned} \tag{58}$$

so this equilibrium is a boundary case that occurs in the transition between two stable equilibria.

Combining all of these stability conditions into a coherent picture requires some additional work. We start by considering the case where $V < K$. In this case, there are three quantities which are important: λk, $\lambda k'$, and $\lambda \frac{v}{p}$.

For $f < V < K$, we get $\lambda k' < \lambda \frac{v}{p} < \lambda k$. We also know that the coexistence with autotomy equilibrium requires $\lambda \frac{v}{p} < \mu < \lambda k'$, which is impossible. The coexistence with no autotomy equilibrium requires $\lambda \frac{v}{p} < \mu < \lambda k$, which is possible. Finally, the prey-only equilibrium requires $\mu > \lambda k$. So we can divide the μ axis neatly into three sections for these two equilibria (with the intermediate-autotomy equilibrium taking over when the no-autotomy equilibrium is unstable) and then for the runaway no-approach situation for $\mu < \lambda \frac{v}{p}$.

For $V < f < K$, we get $\lambda \frac{v}{p} < \lambda k' < \lambda k$. It is now possible for both the autotomy and non-autotomy coexistence equilibria to be stable at the same time. If $c_0^* < c < c_1^*$, we can have bistability of these two equilibria. If $c_1^* < c < c_0^*$, the intermediate autotomy equilibrium is instead stable. The same situation applies for $V < K < f$.

Note that in both of these cases, it is important to incorporate the fact that c_0^* and c_1^* are both functions of μ. First, we note that c_0^* is a strictly decreasing function of μ:

$$\frac{d}{d\mu} c_0^* = -\frac{\beta_0 (1-r)}{\lambda} < 0, \tag{59}$$

and c_0^* at $\mu = 0$ is $\beta_0 k(1-r)$, while at $\mu = \lambda k$, $c_0^* = 0$. Thus, as μ increases from 0 to λk, c_0^* decreases monotonically from $\beta_0 k(1-r)$ to 0.

Note that it is clear by inspection of (Eq. 19) that $c_1^* \geq 0$, that the only value of μ such that $c_1^* = 0$ is $\mu = \lambda k'$, and that when $\mu = 0$, $c_1^* = \frac{\beta_0 k(1-r)}{\beta_0 k(1-r)+r}$.

The derivative of c_1^* with respect to μ is

$$\frac{d}{d\mu} c_1^* = -\frac{\beta_0 rk(1-r)\lambda k'}{(\lambda rk' - \beta_0 k(1-r)(\mu - \lambda k'))^2} < 0, \tag{60}$$

so as μ increases from 0 to $\lambda k'$, c_1^* decreases monotonically from $\frac{\beta_0 k(1-r)}{\beta_0 k(1-r)+r}$ to 0.

For the values of c_0^* and c_1^* to "cross" as μ increases, we require either $k' < k$ (so $f < K$) and $c_1^* < c_0^*$ at $\mu = 0$ or $f > K$ and $c_1^* > c_0^*$ at $\mu = 0$. The former case is satisfied with $f < K$ and $\beta_0 > \frac{1}{k}$. The latter case is satisfied with $f > K$ and $\beta_0 < \frac{1}{k}$. Thus, we can always choose β_0 to force c_0^* and c_1^* to "cross" as μ increases, which allows for both bistable and intermediate-autotomy regimes to be present. We represent this situation in Fig. 6 for $V < f < K$, but not for $V < K < f$ to demonstrate the existence of both these possibilities.

For $V > K$, the prey-only equilibrium instead requires $\mu > \mu_V$. If both f and K are less than V, then neither coexistence equilibrium is stable at any point and $\mu = \mu_V$ transitions between the prey-only equilibrium and the runaway no-approach situation.

If $f > V > K$, then the only possible coexistence equilibrium is the autotomy one, as $\lambda k' > \lambda \frac{v}{p} > \lambda k$. If $\mu_V > \lambda \frac{v}{p}$, then there is a bistable regime between the coexistence equilibrium and the runaway situation (note that $0 \leq c_1^* \leq 1$ always so there will always be a parameter regime where this is the case). If $\mu_V < \lambda \frac{v}{p}$, then there is instead a bistable regime between the prey-only and runaway situations.

When $V > K$, the condition for $\mu_V > \lambda \frac{v}{p}$ is:

$$\begin{aligned} \lambda \frac{v}{p} &< k\lambda + p(1 - rk)(V - K) \\ \lambda \frac{v}{p} &< k\lambda + p(\frac{v}{p} - k) \\ \lambda(\frac{v}{p} - k) &< p(\frac{v}{p} - k), \end{aligned} \tag{61}$$

which is $\lambda < p$ if $V > K$.

Acknowledgements We thank Michael Cortez and four anonymous reviewers for helpful comments on the manuscript, including providing a better derivation of one equation in the eco-evolutionary model.

Author contributions All authors contributed to study conception and design. Analysis was performed by Rohan Sushrut Mehta. The first draft of the manuscript was written by Rohan Sushrut Mehta, and all authors commented on previous versions of the manuscript. All authors read and approved the final manuscript.

Funding This study was funded by a Simons Investigator Award in the Mathematical Modeling of Living Systems to Daniel Weissman and NSF award 1806833 (PoLS SRN).

Declarations

Conflicts of interest The authors have no conflicts of interest to declare that are relevant to the content of this article.

References

Abrams PA (1986) Is predator-prey coevolution an arms race? Trends Ecol Evol 1(4):108–110

Abrams PA (2000) The evolution of predator-prey interactions: theory and evidence. Annu Rev Ecol Syst 31(1):79–105

Bateman P, Fleming P (2009) To cut a long tail short: a review of lizard caudal autotomy studies carried out over the last 20 years. J Zool 277(1):1–14

Berryman AA (1992) The orgins and evolution of predator-prey theory. Ecology 73(5):1530–1535

Blumstein DT (2006) The multipredator hypothesis and the evolutionary persistence of antipredator behavior. Ethology 112(3):209–217

Caro T (2005) Antipredator defenses in birds and mammals. University of Chicago Press

Cooper WE Jr (1998) Conditions favoring anticipatory and reactive displays deflecting predatory attack. Behav Ecol 9(6):598–604

Cooper WE Jr, Frederick WG (2007) Optimal flight initiation distance. J Theor Biol 244(1):59–67

Cooper WE Jr, Frederick WG (2010) Predator lethality, optimal escape behavior, and autotomy. Behav Ecol 21(1):91–96

Cooper WE Jr, Vitt LJ (1991) Influence of detectability and ability to escape on natural selection of conspicuous autotomous defenses. Can J Zool 69(3):757–764

Cortez MH, Ellner SP (2010) Understanding rapid evolution in predator-prey interactions using the theory of fast-slow dynamical systems. Am Nat 176(5):E109–E127

Cott HB (1940) Adaptive coloration in animals. Methuen & Co., Ltd, London

Cressman R, Tao Y (2014) The replicator equation and other game dynamics. Proc Natl Acad Sci USA 111(Supplement 3):10,810–10,817

Cressman R, Ansell C, Binmore K (2003) Evolutionary dynamics and extensive form games, vol 5. MIT Press

Dawkins R, Krebs JR (1979) Arms races between and within species. Proc R Soc London, Ser B 205(1161):489–511

Downes S, Shine R (2001) Why does tail loss increase a lizard's later vulnerability to snake predators? Ecology 82(5):1293–1303

Emberts Z, Miller CW, Kiehl D, St Mary CM (2017) Cut your losses: self-amputation of injured limbs increases survival. Behav Ecol 28(4):1047–1054

Emberts Z, Escalante I, Bateman PW (2019) The ecology and evolution of autotomy. Biol Rev 94(6):1881–1896

Fleming PA, Muller D, Bateman PW (2007) Leave it all behind: a taxonomic perspective of autotomy in invertebrates. Biol Rev 82(3):481–510

Fredericq L (1883) About autotomy or mutilation by reflex as a defense mean of animals. Archives de Zoologie Expérimentale et Générale 1:413–426

Fussmann GF, Loreau M, Abrams PA (2007) Eco-evolutionary dynamics of communities and ecosystems. Funct Ecol 21(3):465–477

Guckenheimer J, Holmes P (1990) Nonlinear Oscillations, Dynamical Systems, and Bifurcations of Vector Fields. Springer-Verlag, New York, NY, USA

Harris RN (1989) Nonlethal injury to organisms as a mechanism of population regulation. Am Nat 134(6):835–847

Hodgkin J, Clark LC, Gravato-Nobre MJ (2014) Worm-stars and half-worms: Novel dangers and novel defense. Worm, Taylor & Francis 3:2157–61

Hofbauer J, Sigmund K (1998) Evolutionary games and population dynamics. Cambridge University Press

Hofbauer J, Sigmund K (2003) Evolutionary game dynamics. Bull Am Math Soc 40(4):479–519

Ives AR, Dobson AP (1987) Antipredator behavior and the population dynamics of simple predator-prey systems. Am Nat 130(3):431–447

Juanes F, Smith LD (1995) The ecological consequences of limb damage and loss in decapod crustaceans: a review and prospectus. J Exp Mar Biol Ecol 193(1–2):197–223

Knope ML, Larson RJ (2014) Autotomy in porcelain crabs is an effective escape mechanism from rockfish predation. Mar Ecol 35(4):471–477

Lawton P (1989) Predatory interaction between the brachyuran crab *Cancer pagurus* and decapod crustacean prey. Mar Ecol Prog Ser 52(2):169–179

Lima SL (1998) Nonlethal effects in the ecology of predator-prey interactions. Bioscience 48(1):25–34

Maginnis TL (2006) The costs of autotomy and regeneration in animals: a review and framework for future research. Behav Ecol 17(5):857–872

Marrow P, Dieckmann U, Law R (1996) Evolutionary dynamics of predator-prey systems: an ecological perspective. J Math Biol 34(5–6):556–578

Martin J, Avery R (1998) Effects of tail loss on the movement patterns of the lizard *Psammodromus algirus*. Funct Ecol 12(5):794–802

McVean A (1975) Autotomy. Comp Biochem Physiol A Physiol 51(3):497–505

Myerson RB (2013) Game Theory. Harvard University Press

Ruxton GD, Allen WL, Sherratt TN, Speed MP (2019) Avoiding attack: the evolutionary ecology of crypsis, aposematism, and mimicry. Oxford University Press

Schoener TW (1979) Inferring the properties of predation and other injury-producing agents from injury frequencies. Ecology 60(6):1110–1115

Smith LD (1992) The impact of limb autotomy on mate competition in blue crabs *Callinectes sapidus* rathbun. Oecologia 89(4):494–501

Uhl G, Nessler SH, Schneider JM (2010) Securing paternity in spiders? a review on occurrence and effects of mating plugs and male genital mutilation. Genetica 138(1):75

Vermeij GJ (1994) The evolutionary interaction among species: selection, escalation, and coevolution. Annu Rev Ecol Syst 25(1):219–236

Wasson K, Lyon BE (2005) Flight or fight: flexible antipredatory strategies in porcelain crabs. Behav Ecol 16(6):1037–1041

Wasson K, Lyon BE, Knope M (2002) Hair-trigger autotomy in porcelain crabs is a highly effective escape strategy. Behav Ecol 13(4):481–486

Wilkinson MH (2003) Decoys in predation and parasitism. Comments on Theoret Biol 8(2–3):321–338

Ydenberg RC, Dill LM (1986) The economics of fleeing from predators. Adv Study Behav 16(C):229–249