ORIGINAL PAPER

Identifying risk factors for persistent versus casual establishment to prioritize rapid response to non-indigenous aquarium fish

Lidia Della Venezia · Brian Leung

Abstract Rapid response strategies are necessary to effectively manage non-indigenous species. After detection, only few species persist and potentially cause harm. To help prioritize resources, we employed a multispecies, geographically explicit approach, focusing on non-indigenous aquarium fish establishment in the USA. We modeled casual (i.e. temporary) establishment and persistence separately, to identify which species should be prioritized after detection. To facilitate the usability of quantitative models, we converted our results into simple "rules of thumb", wherein each factor's contribution represents a multiplier. Finally, from a fundamental perspective, separating casual and persistent establishment improved our understanding of the earlier stages of invasions. We identified five species ranking highest for rapid response, if detected in California, New Mexico and Texas. These states, along with Florida and Hawaii, should take precedence in management funding, being those that currently host more persistent species and where more new establishments are forecasted. Expectedly, the important factors differed considerably between sub-stages, with species traits and propagule pressure being most relevant for casual establishment, and the environment being more predictive of persistence. Notably, propagule pressure had no effect on persistence, suggesting that it would not help target eradication for aquarium fish. Our model allows comparisons for > 1000 species across locations to target rapid responses after detection, and can provide guidance for species not currently traded. Our analysis re-evaluates "risky" species in terms of persistence, suggesting that many species which were flagged in the literature actually pose low risk. Conversely, we identify species that, if detected, warrant rapid response.

L. Della Venezia (✉) · B. Leung
Department of Biology, McGill University, 1205 Docteur Penfield, Montreal, QC H3A 1B1, Canada
e-mail: lidiadellavenezia@gmail.com

B. Leung
e-mail: brian.leung2@mcgill.ca

B. Leung
McGill School of Environment, McGill University, Montreal, QC, Canada

Introduction

Invasive species currently represent one of the biggest threats to the environment, and cause enormous

economic losses (e.g., Ehrenfeld 2010; Pyšek et al. 2012; Simberloff et al. 2013). Hence, in recent years, researchers have investigated ways to prevent and reduce the impact that such species might cause (Buckley 2008). Prevention and rapid response have been widely recognized as the most successful management approaches to hamper non-indigenous species in a cost-effective way (Alvarez and Solis 2019). Prevention is usually preferable and represents a priority among management policies (Finnoff et al. 2007; Leung et al. 2002). However, it does not necessarily guarantee success (Vander Zanden et al. 2010), and rapid response remains a critical strategy for management (Wittenberg and Cock 2001), especially when the invader's density is low and measures like eradication are more feasible and cost-effective (Simberloff et al. 2013; Westbrooks 2004). The importance of the appropriateness of intervention has prompted discussion on how resources should be spent (e.g., Leung et al. 2005; Vander Zanden et al. 2010). Prioritizing instances where management would be most needed is essential to direct limited resources (Lohr et al. 2017). Particularly, promptly identifying species and locations of highest concern would allow stakeholders to respond to new introductions and detections more quickly and effectively (e.g., Vander Zanden and Olden 2008). Yet, it remains a major challenge in invasion ecology (Stewart-Koster et al. 2015).

In this study, we provide insight into rapid response for eradication. We focused on the establishment phase, i.e. the process by which a non-indigenous species in a novel location founds a self-sustaining population, with individuals surviving and successfully reproducing (Lockwood et al. 2013). However, even populations that temporarily survive and reproduce can subsequently die off without anthropic intervention (Blackburn et al. 2011). We thus considered that distinguishing between an initial sub-stage wherein a species is found in a novel location temporarily (henceforth termed casual establishment) and the subsequent step wherein the species exists as an enduring self-sustaining population over time (henceforth termed persistent establishment) would be of practical importance for management. In fact, simple detections of a species in the wild are often treated as establishments or a proxy for it, but they arguably relate more closely to casual establishment, overestimating those species which persist to

potentially cause harm. Therefore, when non-indigenous organisms are detected in the wild, distinguishing species that would likely go extinct without human mediation from those that pose a real threat could help determine instances where eradication effort is needed, and provide critical information for prioritizing rapid response resources.

While mathematical models provide rigorous, quantitative means to prioritize management (e.g., Chadès et al. 2011; Kerr et al. 2016), decisions in this field often do not make use of this quantitative evidence, relying instead on expert opinion, and less technical scoring-based approaches (Cook et al. 2010; Leung et al. 2012). Arguably, science should facilitate the usability of technical knowledge for stakeholders (Cassey et al. 2018a), for instance by converting model parameters into rules of thumb. Ideally, such rules of thumb should be easy to apply and should provide useful insights. Similar concepts are often used in medicine to express changes in the odds of an outcome following exposure to a certain factor (e.g., the risk of low weight at birth is 4 times higher in neonates from women exposed to tobacco; Mumbare et al. 2012). Simultaneously, they should still be based on a solid statistical foundation that derives reliable predictions from empirical data. In this manuscript, we also express our models as rules of thumb, in addition to parameter values.

Finally, we consider the fundamental insights from studying the two phases of casual and persistent establishment separately. Currently, few studies have looked into identifying the factors associated with failures at different stages of invasions (Dawson et al. 2009; Marchetti et al. 2004). Despite establishment being the best studied phase in biological invasions (Leung et al. 2012), very few studies specifically concentrated and used information on temporary establishments to more finely characterize this stage (e.g., Essl et al. 2015; Ficetola et al. 2009; Milbau and Stout 2008). For example, Duggan et al. (2006) examined the importance of propagule pressure and other introduction characteristics (e.g., maximum size) on the probability of freshwater fish species from the aquarium trade being introduced (i.e. detected) in the wild and successfully establishing. Here, we analyzed casual and persistent establishment, simultaneously considering propagule pressure, environmental suitability, and species traits (i.e. the main predictors of establishment; Leung et al. 2012),

to assess their relative contributions during each sub-stage. Specifically, despite having been widely recognized as a strong determinant of invasion success (Cassey et al. 2018b; Lockwood et al. 2013), it is unclear whether propagule pressure primarily contributes to casual establishment (e.g., Leung et al. 2012) or it remains important for persistence (e.g., Ficetola et al. 2009). Similarly, environmental conditions, and more importantly climate matching, were shown to be consistently relevant factors influencing establishment across taxa (e.g., Duncan et al. 2014; Hayes and Barry 2008; Mahoney et al. 2015). Nonetheless, local conditions might have a stronger effect on the early survival of the propagules reaching a new location (casual establishment; Essl et al. 2015) or could affect long-term population persistence through vulnerable phases like reproduction (Ficetola et al. 2009). Analogously, studies focusing on the role of species traits found them to be important predictors of establishment (Pyšek et al. 2009), either allowing individuals to cope with novel conditions (Blackburn et al. 2009), or having a stronger effect on persistence, for example favouring future reproduction in spite of early population growth (Sol et al. 2012).

In brief, we focused on fish species introduced through the aquarium trade pathway as our study system to address three-fold objectives: (1) to develop a model to help prioritize rapid response, by separating the sub-stages of establishment (casual versus persistent); (2) to provide "rules of thumb", wherein model parameters are converted into a series of simple multiplicative risk factors; (3) to increase fundamental understanding of the invasion process, by elucidating the relative contributions of propagule pressure, environment, and species traits for each of the two sub-stages of establishment.

Methods

In this study, we focused on aquarium freshwater fish species traded in the United States for which import data were available. The aquarium trade is responsible for importing thousands of individual fish annually and is a significant source of non-indigenous species (Howeth et al. 2016; Rixon et al. 2005). To separate the predictors associated with the early establishment and persistence, we examined spatially referenced records of non-indigenous freshwater fish species in the USA over the past 50 years (United States Geological Survey 2017). We selected all the detection records classified as "aquarium release" and compiled them by state. The USGS categorized each record by status, based on reproduction, persistence and eradication (Table 1). We grouped the species detected in the wild into casually established species (CS) and persistently established species (PS). The CS included all the species by state that were detected at some point in time after 1971 regardless of their present status (59 species and 151 species-location combinations). We chose 1971 as a threshold since previous work from Chapman et al. (1997) and Bradie et al. (2013) had shown that species' popularity remained relatively consistent in the aquarium fish trade after that year. Two occurrences (*Devario malabaricus*, NV, and *Serrasalmus rhombeus*, FL) which were classified as "Eradicated" (Table 1) might have persisted in the absence of human intervention. We found that results were robust to their inclusion and kept these species in the CS group. Since a species that casually establishes can either go extinct or persist, CS included both casual and persistent species. In contrast, the PS group included only those species which avoided extirpation, i.e. species that successfully reproduced and overwintered, and for which

Table 1 Aquarium species status categories, as defined by the USGS Non-indigenous Aquatic Species website and as categorized in this study (CS/PS)

Status	Description	CS/PS
Collected	Species was collected or observed from the site; reproduction is not known	CS
Established	Population is reproducing and overwintering, currently established	CS, PS
Eradicated	Population was eliminated by human activity	CS
Extirpated	Population died out on its own, without human interference	CS
Failed	Population was introduced but died out; failed to reproduce	CS
Unknown	Used when all other categories do not fit	CS

multiple life stages were identified in the wild (21 species and 28 species-location combinations, identified as "Established"; Table 1). All other aquarium fish species imported into the USA that have never been detected in the wild were considered unestablished (neither casually nor persistently). The complete list of occurrences for both CS and PS can be found in Appendix S1 (Table S1.1 and S1.2).

We used the PET modeling framework (Della Venezia et al. 2018) to account for species traits, environment, propagule pressure and density-dependent effects simultaneously. The model defined the probability of a species establishing as:

$$P(E) = 1 - (1 - p)^{N^c} \tag{1}$$

where p was the probability of a single propagule establishing, N was the number of propagules introduced, and c was a shape parameter allowing density-dependent effects. p was modeled as a logistic function of species-specific and location-specific predictors:

$$p_{sl} = \frac{1}{1 + e^{-z_{sl}}} \tag{2}$$

where z_{sl} was defined as:

$$z_{sl} = b_0 + \sum_{w=1}^{W} \left(b_{w1} X_{ws} + b_{w2} X_{ws}^2 \right) \\ + \sum_{m=1}^{M} \left(b_{m1} E_{ml} + b_{m2} E_{ml}^2 \right) \\ + \sum_{w=1}^{W} \sum_{m=1}^{M} (b_{mw} X_{ws} E_{ml}) \tag{3}$$

Each X_{ws} was a trait of species s, for a total of W traits, while each E_{ml} was an environmental condition for location l, for M environmental variables. Both first and second order terms for each of these predictors were included to allow a non-monotonic relation with establishment probability. b denoted the coefficients of the model, for traits (w) and environmental conditions (m), and traits-environment interactions, each described by parameter value b_{mw}, common across all species/location combinations.

We used the PET framework in two different ways. Firstly, we identified the factors that predicted casual establishment (i.e. CS versus unestablished species that were never detected in the wild). We referred to this as the "casual establishment model". Secondly, to

create rules of thumb for rapid response, we identified the factors that predicted PS versus extirpated species (the fraction of CS that later went extinct). We referred to this as the "persistence model".

Variable choice

The PET establishment model incorporated propagule pressure (P), environment (E) and species traits (T) as predictors (Appendix S2). Building from Della Venezia et al. (2018), species traits originated from the online database FishBase (Froese and Pauly 2018) and were minimum and maximum temperature tolerance, northernmost latitude, trophic level and maximum length. Additional traits were removed to avoid multicollinearity, while missing data were imputed using the methodology described in Della Venezia et al. (2018). The environmental variables were obtained from the Bioclim database (www.worldclim.com; Hijmans et al. 2005). Given that the establishment data were at the state level, we estimated mean ($\bar{x}$) and variance (s^2) of each variable for each state. The final set, after removing the highly correlated ones, included variability of the diurnal range ($BIO2_s^2$), mean and variability of the minimum temperature of the coldest month ($BIO6_x$ and $BIO6_s^2$), mean temperature of the warmest quarter ($BIO10_x$) and mean precipitation of the wettest month ($BIO13_x$). All variables were standardized before fitting the models.

Since fish releases from aquarists were virtually impossible to track, as a proxy for propagule pressure we used Canadian import data from Fisheries and Ocean Canada (B. Cudmore and N. Mandrak, unpublished data). Bradie et al. (2013) showed that US aquarium fish imports could accurately be obtained by scaling Canadian imports by population size. The same approach was used to derive geographically explicit estimates of propagule pressure based on population by state. Population size data were available from the United States Census Bureau (https://www.census.gov/en.html).

Model fitting

Our models were fit using maximum likelihood estimation, which allowed finding the best fitting parameters given the observed data. The log-

likelihood function (Della Venezia et al. 2018) was defined as:

$$\log(L) = \sum_{l=1}^{L} \left(\sum_{i_l=1}^{I_l} \log\left(1 - (1 - p_{il})^{N_{il}^c}\right) \right.$$
$$\left. + \sum_{u_l=1}^{U_l} \log(1 - p_{ul})^{N_{ul}^c} \right) \qquad (4)$$

i represented a successful species, i.e. CS or PS depending on the model, and u was a species which failed to establish either temporarily or persistently, for each location l. The sum was iterated over all L states.

We applied a forward selection procedure to identify the most predictive variables (Johnson and Omland 2004). Although issues associated with stepwise variable selection are known (e.g., inflated type I errors) and more robust methods of model selection exist, we opted for a forward selection approach, as other approaches which began with all terms simultaneously in the model resulted in problems of complete separation (Albert and Anderson 1984), given the data available. We selected the Akaike Information Criterion (AIC; Akaike 1974) as an efficient metric to select predictive models (Aho et al. 2017), and each variable was retained in the model when its inclusion decreased AIC by at least 2 units (Anderson and Burnham 2002). Model performance was estimated using the area under the curve (AUC) of the receiver operating characteristic (ROC; Hanley and McNeil 1982).

Finally, to improve our fundamental understanding of the establishment process, we characterized the importance of the three categories of predictors in analysis by comparing the full PET framework to each submodel (i.e. excluding propagule pressure, environment or species traits). We ranked them based on their AIC and the percentage of deviance explained, to evaluate the relative importance of species traits, environment and propagule pressure during each substage of establishment.

Multiplicative risk factors

Once the best model for each sub-stage was identified, we derived "rules of thumb" to easily compare species and prioritize resources. We determined the effect of each important predictor on the likelihood of a species successfully establishing (casually or persistently) versus failing. To do so, we converted the parameters of the fitted models (Eqs. 2, 3) and expressed each factor as a multiplier, increasing or decreasing risk of establishment. This was accomplished by calculating the odds ratios (OR), i.e. the ratio between the odds (probability of successfully establishing versus failing) for varying values of each significant predictor and the odds at a reference value (Appendix S3):

$$OR_k = \frac{Odds_k}{Odds_{ref_k}} \qquad (5)$$

For each significant predictor k, its mean value across species or locations was chosen as reference. Odds ratios are the simplest way of interpreting the results of a logistic model and have been used extensively in epidemiology and medicine (e.g., Bland and Altman 2000; Cummings 2009). Each OR_k represented a multiplicative risk factor, i.e. a measure of the relative change in risk of establishing versus failing, relative to the odds of average species and locations. Based on logistic regression, odds ratios had the advantage of being estimated independently for each variable, so that the relative contribution of each predictor could be assessed separately. Also, each OR_k mathematically corresponded to a multiplier, so that the cumulative effect of all predictors (OR_c) was the product across all OR_k.

$$OR_C = \prod_{k=1}^{K} OR_k \qquad (6)$$

For instance, if a species s had a value for trait k = 1 corresponding to $\Delta OR_1 = 2$ (Fig. 1a), the likelihood of successfully persisting versus failing for species s would be twice as high as that of a species with a mean value for trait 1. If species s also had $OR_2 = 1.25$ for trait k = 2 (Fig. 1b), then the overall $OR_c = 2 \times 1.25$, and, altogether, species s would be 2.5 times more likely to establish versus not establishing than the average species. Odds ratios could be used to compare different species as well, where the ratio $OR_{s=1}/OR_{s=2}$ represents the relative odds of success of species 1 over species 2.

Based on logistic regression, probabilities can be obtained from the odds ($p = \frac{Odds}{1+Odds}$). Knowing OR_C and the $Odds_{ref}$ when all predictors were set to the reference value, actual probabilities for each species/location combination would correspond to:
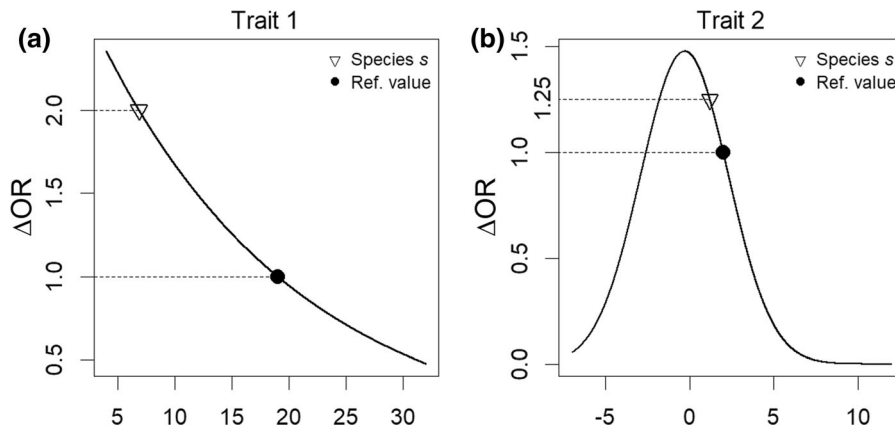
Fig. 1 Illustrative example of rules of thumb, i.e. odds ratios (OR), derived for two species traits. Each dark dot represents the reference point, i.e. the mean trait value across species in the dataset, while each triangular dot corresponds to the species s of interest

$$p = \frac{Odds_{ref}OR_C}{1 + Odds_{ref}OR_C} \qquad (7)$$

All data manipulations, model fitting and analyses were conducted in the R statistical programming environment (R Core Team 2018).

Results

Five factors distinguished species likely to become naturally extirpated from those that persisted (AUC =

0.946, $\approx$ 52% deviance explained; Table 2). The minimum temperature of the coldest month ($BIO6_s^2$) was among the strongest predictors of persistence, likely determining whether the species could over-winter (Table 3). In particular, its variability ($BIO6_s^2$) was the best determinant for persistence (Table 3), presumably because strong fluctuations could drive species to extirpation, even those that were able to initially survive the low minima. Persistence also appeared to be favoured at quite high mean temper-atures of the warmest quarter ($BIO10_x$) and for intermediate species with high maximum temperature

Table 2 AIC, AIC difference from the best model (ΔAIC), AUC, fitted $\hat{c}$ parameter and percentage of deviance explained (%dev.exp) by the full casual establishment and persistence model, and their respective submodels

The best model for each dataset is indicated in bold

	AIC	ΔAIC	AUC	$\hat{c}$	%dev.exp
Casual establishment model					
PET	**1359.89**	**0**	**0.957**	**0.546**	**35.59**
sp. traits—pp	1504.97	145.08	0.926	0.542	28.07
Environment—pp	1753.45	393.56	0.838	0.360	13.93
sp. traits—environment	1585.52	225.63	0.897	23.32	
Propagule pressure	1889.50	529.61	0.795	0.378	8.67
sp. traits	1773.85	413.96	0.806	–	14.95
Environment	1896.87	536.98	0.719	–	6.84
Null model	2066.59	706.70	0.500	–	0
Persistence model					
PET	85.91	2	0.946	1×10^{-8}	51.60
sp. traits—pp	117.20	33.29	0.841	1×10^{-12}	26.87
Environment—pp	106.34	22.43	0.840	1×10^{-8}	34.85
sp. traits—environment	**83.91**	**0**	**0.946**	**–**	**51.52**
Propagule pressure	148.82	64.91	0.439	4×10^{-9}	0
sp. traits	115.20	31.29	0.840	–	26.87
Environment	104.34	20.43	0.862	–	34.86
Null model	146.82	62.91	0.500	–	0

Table 3 Parameter values for the predictors retained in each sub-stage best model

Parameter	Casual establishment model			Persistence model		
	Rank	1st	2nd	Rank	1st	2nd
Species traits						
Maximum temperature tolerance (°C)	4	0.411	–	3	0.360	0.834
Minimum temperature tolerance (°C)	5	− 0.512	− 0.404	NA	–	–
Northernmost latitude	3	0.917	− 0.589	NA	–	–
Trophic level	6	− 0.166	–	NA	–	–
Maximum length (cm)	1	1.795	− 1.844	2	1.382	− 7.679
Environmental conditions						
Minimum temperature coldest month (°C; $BIO6_x$)	7	0.112	0.212	NA	–	–
Mean temperature warmest quarter (°C; $BIO10_x$)	NA	–	–	4	0.504	− 2.670
Precipitation wettest month (mm; $BIO13_x$)	2	0.053	0.479	NA	–	–
Minimum temperature coldest month (°C; $BIO6_s^2$)	NA	–	–	1	4.971	− 4.114
Interactions						
Min. temp. coldest month ($BIO6_x$) and Max. length	8	− 0.129	–	NA	–	–
Min. temp. coldest month ($BIO6_s^2$) and Max. length	NA	–	–	5	− 3.172	–

Rank indicates the entry order of each important variable in the corresponding model, either as a first order term only (1st) or including an additional second order term (2nd)

tolerances (Table 3). Finally, the probability of persistence appeared to change with interacting minimum temperature of the coldest month and maximum length, with big species being favored in more stable cold environments (Table 3).

Multiplicative risk factors

Using the parameters from the fitted models (Table 3), we calculated the odds for varying values of the predictors (Appendix S3), and we compared them against the mean across locations and species as our reference value ($Odds_{ref} = 0.002184$, corresponding to $p_{ref} = 0.002180$) to obtain our rules of thumb (Fig. 2; Table 4). Some predictors were substantially more important. For example, maximum temperature tolerance appeared more influential than maximum length for persistence, as shown by the range of corresponding ORs (Fig. 2a, b). Temperature tolerance sharply increased persistence, with species with tolerances higher than 35 °C being at least 50 times more likely to successfully persist than failing (Fig. 2a). On the other hand, persistent establishment appeared likely for a relatively narrow range of lengths, with species about 60 cm in length being 4 times more likely to persist than to fail compared to the

mean, and very big species being extremely unlikely to cause concern in the long term (Fig. 2b). Similarly, the effect of minimum temperature variability ($BIO6_s^2$; Fig. 2d) appeared substantially stronger than the mean warmest temperature ($BIO10_x$; Fig. 2c), increasing the odds up to ≈ 200 times with respect to average conditions.

Odds ratios allowed us to combine the contribution of different predictors by simple multiplication, to obtain their overall expected effect on persistence risk. To illustrate, the species expected to have the highest likelihood of persistent establishment in the USA was the jaguar guapote (*Parachromis managuensis*; for a list of traits, see Appendix S1). The estimate of OR was ≈ 150 for maximum temperature tolerance, and ≈ 4 for size, respectively (Fig. 2a, b), making this species about ($150 \times 4 =$) 600 times more likely to establish than the average fish in our persistence dataset. Similarly, based on local environmental conditions, the state where persistent establishment was more likely to occur was Hawaii, with OR of about 1 and 55, for mean temperature of the warmest quarter and minimum temperature variance, respectively (Fig. 2c, d), making the insular state ≈ 55 times more suitable to persistent species than the average state. For the jaguar guapote, the calculated overall
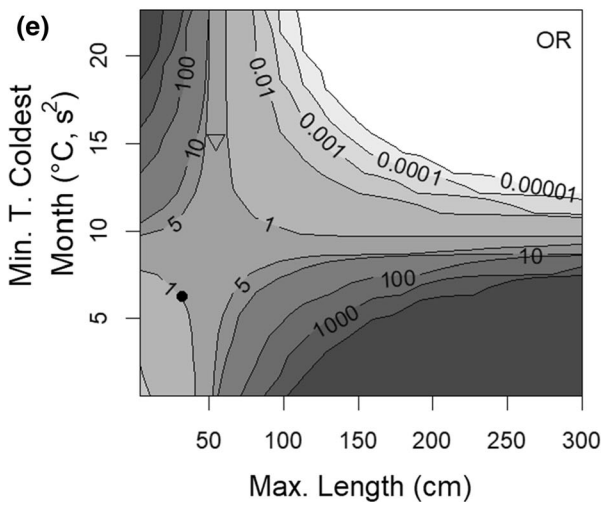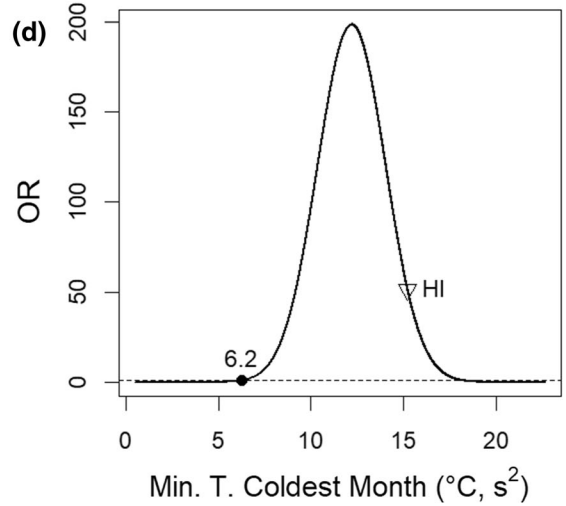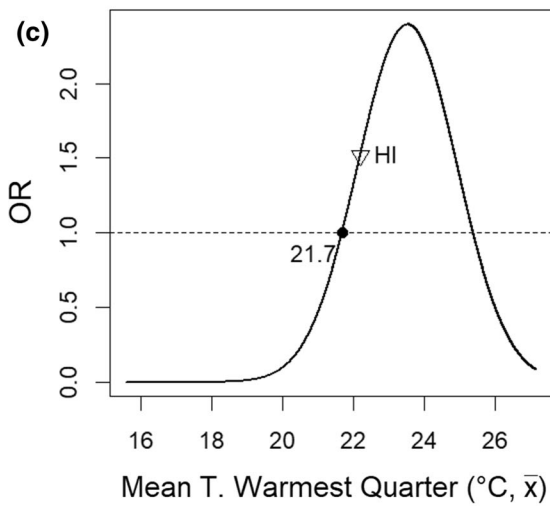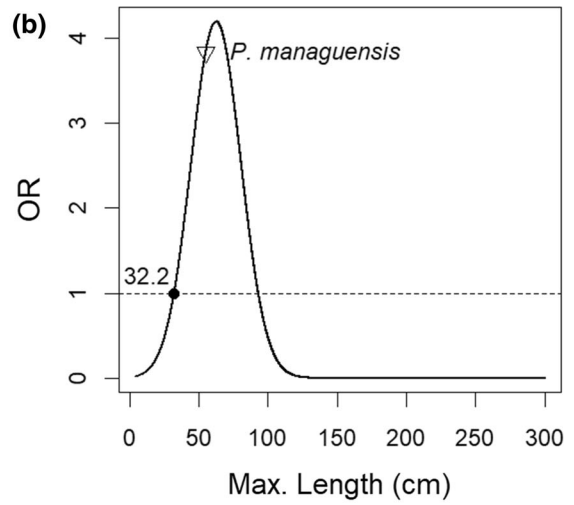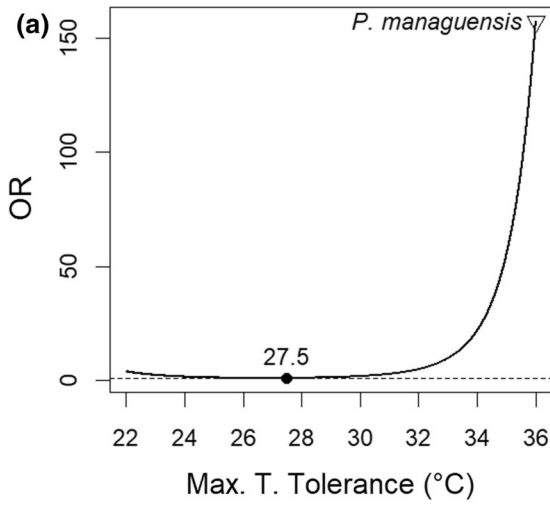
◄**Fig. 2** Effect of each significant predictor on the likelihood of persistently establishing versus failing, expressed as odds ratio (OR), when gradually varying each predictor. OR equals 1 (dashed line) at each variable's mean value, reported by the corresponding point. The mean values for the interaction plot (**e**) correspond to those of the respective main terms (**b**, **d**). The triangles indicate the OR of *P. managuensis* (traits) and Hawaii (environmental conditions), and their interaction (**e**). Very high values in (**e**) coincide with areas of extremely low absolute probability values, so that little probability increases determine very high OR

likelihood of successfully persisting versus failing in Hawaii was more than 60 thousand times higher than the average species/location, including the contribution of the interaction terms (OR interaction corresponding to ≈ 2, Fig. 2e; overall OR = 150 × 4 × 1 × 55 × 2). Thus, while the mean probability of persistence was low across all species and locations (p_{ref} = 0.002180), for the jaguar guapote in Hawaii it was 0.9964. Similar estimates can be easily obtained and compared for any species/location combination, using the parameter values from each model, and the information provided in Appendix S1 and Appendix S3.

Re-evaluating "risky" species in terms of persistence

Additionally, we looked at species within our dataset which had been flagged as potentially invasive in the literature. As a case example, we focused on the Great Lakes and looked at species previously identified as potentially invasive in this area, including the European weatherfish (*Misgurnus fossilis*), the spined loach (*Cobitis taenia*), the white cloud mountain minnow (*Tanichthys albonubes*), the clown loach (*Chromobotia macracanthus*), the silver arowana (*Osteoglossum bicirrhosum*) and the glass catfish (*Kryptopterus bicirrhis*). We found that none of the species listed appeared to be likely to persist in the Great Lakes region, although some of them would be troubling in other states, being tens to hundreds of times more likely to succeed than on average. These were the glass catfish in Hawaii, and the white cloud mountain minnow and the clown loach in both Florida and Hawaii (Appendix S4). By comparison, the silver arowana was never able to persist, nor was it predicted to pose a substantial threat across the USA, despite having been casually established in 9 states.

Table 4 Selected multiplicative risk factors to quickly quantify risk of casual and persistent establishment

	Mean	+ 1000%	+ 100%	+ 50%	+ 25%	− 25%	− 50%
Casual establishment							
Max. temperature tolerance (°C)	26.7	39	30.3	28.8	27.8	25.2	23.2
Min. temperature tolerance (°C)	22.5	–	–	13.2; 20.2	12; 21.4	9.7; 23.7	8.3; 25
Northernmost latitude	3.4	–	–	12.4; 33.5	7.7; 38.3	− 1; 47	− 6.1; 52.1
Trophic level	3.1	–	–	–	2.4	4	–
Max. length (cm)	21.6	61.7; 225.3	31.7; 255.3	27.4; 259.6	24.8; 262.2	17.7; 269.3	12.3; 274.7
Min. temp. coldest month (°C; $BIO6_x$)	− 7.5	–	6.3	− 23.7; 2.5	− 20.5; − 0. 7	–	–
Precipit. wettest month (mm; $BIO13_x$)	111.1	–	55.9; 160.3	68.1; 148.1	78.4; 137.8	–	–
Persistence							
Max. temperature tolerance (°C)	27.5	33.1	23.3; 30.3	24.1; 29.5	24.7; 28.8	–	–
Max. length (cm)	32.2	–	40.8; 84.5	36.9; 88.4	34.7; 90.6	29.3; 96	25.7; 99.7
Mean temp. warmest quarter (°C; $BIO10_x$)	21.7	–	22.7; 24.4	22.2; 24.9	21.9; 25.1	21.4; 25.7	21.1; 26
Min. temp. coldest month (°C; $BIO6_s^2$)	6.2	7.8; 16.6	6.6; 17.8	6.5; 18	6.4; 18.1	6.1; 18.3	5.9; 18.6

The first column reports the mean value of each relevant predictor in the equivalent model, while the other columns identify the variable values corresponding to an OR change of + 1000%, + 100%, + 50%, + 25%, − 25% and −50%. The means of variables that were significant for both models differ, because the persistence model dataset represents a subset of the casual establishment model dataset (e.g., maximum length)

On the other hand, among the species/location combinations for which casual establishment had already occurred, our model predictions for persistence suggested different species of highest concern. These included the red-bellied piranha (*Pygocentrus nattereri*), the lowland cichlid (*Herichthys carpintis*), the Rio pearlfish (*Nematolebias whitei*), the banded leporinus (*Leporinus fasciatus*), and the climbing perch (*Anabas testudineus*), particularly in California, Florida, Hawaii, New Mexico and Texas. All had a likelihood of successfully persisting versus failing tens to hundreds of times higher than average species/locations. If these species were detected again in their casual occurrence sites, they should be prioritized for rapid response.

Comparing establishment sub-stages: casual versus persistent

The combination of propagule pressure, environment, and traits (i.e. PET) performed better than any of the submodels for casual establishment, both in goodness of fit (AIC) and prediction accuracy (AUC_{PET} = 0.957, $\approx$ 36% deviance explained; Table 2; Appendix S5). In contrast, for persistence, the best model included only species traits and environmental conditions (AUC_{ET} = 0.946, $\approx$ 52% deviance explained; Table 2). Particularly, the large difference in AIC between each best model and its alternatives ($\Delta AIC > 22$ for the casual establishment model; $\Delta AIC > 145$ for the persistence model; Table 2) suggested that our results should be robust to specific approaches for model selection (i.e., our use of forward selection).

For casual establishment, species traits were more explanatory than the environment (either with or without propagule pressure; Table 2), and indeed all traits examined were important (Table 3; Appendix S5). More specifically, casual establishment risk was inversely related to trophic level, seemingly favouring herbivorous species. Casual establishment risk also increased with latitude for species whose northern-most distribution limit ranged up to $\approx 23°N$, and then decreased at higher latitudes (Table 3; Appendix S5). Some traits were also predictive of persistent species, with maximum length and, to a lesser extent, maximum temperature tolerance being important for both sub-stages. Interestingly, risk was unimodally related to maximum length, with intermediate to big species

having an advantage for casual establishment, and relatively smaller ones being favored for persistence. Optimal physiological ranges of temperature were significant both for casual and persistent establishment, with minimum temperature tolerance losing relevance at the latter sub-stage (Table 3; Appendix S5).

In contrast, environmental conditions appeared to be more relevant for persistence, explaining alone $\approx$ 35% of deviance (Table 2). While precipitations and mean minimum temperatures were determinants of casual establishment (Table 2; Appendix S5), persistence was predicted by mean temperatures of the warmest quarter and variability in minimum temperature of the coldest month (Table 2). Nonetheless, the submodel combining species traits and environment was the best in distinguishing persistent species from those that were subsequently extirpated, while propagule pressure was generally not predictive of persistence for aquarium fish (Table 2). Propagule pressure appeared to favour casual establishment and its inclusion in the model added considerably to the final percentage of explained deviance (Table 2), with values of propagule pressure higher than the median increasing risk (Appendix S2). However, propagule pressure was not important for persistence of aquarium fish species (Table 2; Appendix S2). This was further corroborated by the fitted $\hat{c}$ value being consistently very close to zero for persistence models (at zero, propagule pressure would have no effect; Table 2).

Discussion

Strides have been made in the field of invasion biology, aiming at identifying priorities for management and guiding strategies for prevention and early response. However, prioritization remains a challenging task, and geographically explicit, multispecies risk assessment frameworks could help make the process more efficient. Here, we focused on establishment and rapid response. Among the casual fish species considered in this study, more than 80% later went extinct without human intervention, and would have resulted in unnecessary allocation of valuable resources, if funds had been spent on their eradication. Instead, by separating establishment into two sub-stages and pinpointing the factors associated with successful casual and persistent establishment, we have provided

analyses to target species likely to persist, after they have been detected.

Although we recognize that non-indigenous species could generate local impact even during casual establishment, being able to prioritize species that pose a lasting threat and to redirect (often scarce) resources toward instances where their investment would be necessary, would help maximize the efficacy of management strategies (Jenkins 2013; Keller and Perrings 2011). When detections take place early, the five predictors identified and the multiplicative risk factors derived from the persistence model would allow a quick simultaneous assessment of multiple species and locations. In an exemplificative case, the pirapitinga (*Piaractus brachypomus*) has managed to casually establish in as many as 44 states, but it has never persisted due to unfavorable local conditions. However, its likelihood of persistence is non-negligible in some states like Florida and Nevada, where measures should be taken if detected. Further, we have listed *Pygocentrus nattereri, Herichthys carpintis, Nematolebias whitei, Leporinus fasciatus* and *Anabas testudineus* as the most likely to persist among species that have already casually established. These species are ranked as highest concern and they should be eradicated if detected again, especially in Texas, New Mexico, California, Florida or Hawaii. Instead, when we looked at species that had already been classified as potential threats in the Great Lakes area (e.g., Howeth et al. 2016; Kolar and Lodge 2002; Rixon et al. 2005), our results suggested that even if they were detected, they would have a high likelihood of extirpation without intervention. Although casual establishment might occur for some of these species, persistence was predicted to be very unlikely, and the states bordering the Great Lakes should prioritize the investment of management resources on different pathways of introduction of potentially invasive species or alternative environmental concerns.

On the other hand, species traits and local environmental conditions that are advantageous in the earliest phases of establishment might also represent useful filtering criteria to drive restrictions in the aquarium market and to define targets to invest resources for early detection (Mehta et al. 2007). Although we focused primarily on informing rapid response, preventing casual establishment could be necessary in specific cases, e.g., for species that can cause substantial temporary impact, or for which eradication would hardly be feasible (e.g., Dogliotti et al. 2018; Simberloff 2003). In such cases, a reduction in the number of commercialized individuals should be considered, as it might be sufficient to make establishment risk virtually inexistent. However, targeting propagule pressure after detection would not reduce persistence risk for freshwater aquarium fish, based on our results. Overall, our model and the associated multiplicative risk factors provide quantitative support to decision making that would help target which species might pose a risk and in which locations, and thus where rapid response strategies should be developed. The specific strategies for rapid response (eradication, containment, etc.; Britton et al. 2011), and their feasibilities (i.e., risk management) have not been explicitly examined here as they were beyond the scope of this manuscript, and would require a separate set of models and analyses (e.g., Peterson et al. 2008).

Separating establishment into sub-stages also provided additional fundamental knowledge about this phase of invasions. Generally, our results reflect the importance of distinct predictors during separate phases of the invasion process (e.g., Dawson et al. 2009; Essl et al. 2015; Milbau and Stout 2008). For instance, propagule pressure has been recognized as the most consistent predictor of establishment success across taxa (Cassey et al. 2018b; Lockwood et al. 2013). Looking at sub-stages, our results suggested that propagule pressure was very important for casual establishment, but not for persistence, confirming previous observations on freshwater fish (Marchetti et al. 2004). However, studies on vascular plants had found that continual propagules contribution could enhance establishment success at later stages (e.g., Essl et al. 2015), suggesting that establishment dynamics vary across taxa.

In contrast to propagule pressure, both species-specific and location-specific characteristics played an important role during both sub-stages of establishment. However, species traits were more important for casual establishment, while location-specific variables were more important for persistence. As expected, the relevant predictors differed between stages. For example, trophic level was retained as a predictor of casual establishment, but did not have an effect on persistence. Even for traits that were relevant for both stages, their relationship with risk changed. For example, very large species were generally disadvantaged across stages, in agreement with findings in the

literature (Ribeiro et al. 2008; Ruesink 2005). While mid-range species were favored for early establishment, potentially due to an initial survival advantage, only relatively small ones appeared to find a suitable environment for persistence. Such species were often detected in the wild across northern states (e.g., *Pygocentrus nattereri*; Appendix S1), but coming from tropical or subtropical regions, they could only overwinter in the mild climate of the southernmost USA (Bennett et al. 1997).

The environment, on the other hand, seemed to play a greater role for persistence, similarly to what has been observed for bryophytes (Essl et al. 2015). Our results supported observations from previous studies about the importance of climatic variables at different stages of invasion in fish (Bomford et al. 2010; Howeth et al. 2016) and other vertebrates (Duncan et al. 2001; Forsyth et al. 2004; Mahoney et al. 2015). Expectedly, the minimum temperature of the coldest month was one of the strongest environmental predictors of casual establishment and persistence, both as mean and variability. After overcoming low temperatures during the earlier phase, aquarium species seemed to favor locations that are relatively steady in winter, in accordance with previous studies (Bradie and Leung 2017; Drake and Lodge 2004).

The approach used here can be applied to other suits of organisms across different pathways of introduction, to derive geographically explicit, pathway-specific risk factors. Clearly, when more information is available, the methodology can be extended to accommodate alternative predictors among species traits and environment, e.g., invasion history or native range size (Kolar and Lodge 2002; Peoples and Midway 2018). Analogously, all available information on both successful and failed establishments should be incorporated as it becomes available, while data on escapes and intentional releases of non-indigenous individuals in the wild would represent ideal propagule pressure estimates and help exclude species that were never released. Ideally, density data could be integrated in the analysis and related to casual and persistent establishment, as well as impact. The scale of our study and the types of data available also prevented us from obtaining more finely resolved predictions to inform management with higher geographical resolution. Although data might be particularly limiting for certain pathways, information-rich

datasets exist (e.g., trade in birds, reptiles and amphibians; Abellán et al. 2016; Herrel and van der Meijden 2014; Robinson et al. 2015), and advancements are being made to improve pathway-level knowledge compiled from diverse sources (e.g. Saul et al. 2017). Despite these limitations, in the context of invasions by aquarium fish, this work suggests a small number of predictors can differentiate species and locations likely to establish and persist. Moreover, our model suggests different species of most concern, from the perspective of persistence, and thus different targets of rapid response, once detection of a species in the wild has occurred.

Authors' contributions Both LDV and BL conceived the project. LDV built and analysed the models, and derived the multiplicative risk factors for rapid response. LDV led the preparation of the manuscript, and both LDV and BL contributed to the final version.

References

Abellán P, Carrete M, Anadón JD, Cardador L, Tella JL (2016) Non-random patterns and temporal trends (1912–2012) in the transport, introduction and establishment of exotic birds in Spain and Portugal. Divers Distrib 22:263–273

Aho K, Derryberry D, Peterson T (2017) A graphical framework for model selection criteria and significance tests: refutation, confirmation and ecology. Methods Ecol Evol 8:47–56

Akaike H (1974) A new look at the statistical model identification. IEEE Trans Autom Control 19:716–723

Albert A, Anderson JA (1984) On the existence of maximum likelihood estimates in logistic regression models. Biometrika 71:1–10

Alvarez S, Solis D (2019) Rapid response lowers eradication costs of invasive species: evidence from Florida. Choices 33(4):1–9

Anderson DR, Burnham KP (2002) Avoiding pitfalls when using information-theoretic methods. J Wildl Manag 66:912–918

Bennett WA, Currie RJ, Wagner PF, Beitinger TL (1997) Cold tolerance and potential overwintering of the red-bellied piranha *Pygocentrus nattereri* in the United States. Trans Am Fish Soc 126:841–849

Blackburn TM, Cassey P, Lockwood JL (2009) The role of species traits in the establishment success of exotic birds. Glob Change Biol 15:2852–2860

Blackburn TM, Pyšek P, Bacher S, Carlton JT, Duncan RP, Jarošík V, Wilson JR, Richardson DM (2011) A proposed unified framework for biological invasions. Trends Ecol Evol 26:333–339

Bland JM, Altman DG (2000) The odds ratio. BMJ 320:1468

Bomford M, Barry SC, Lawrence E (2010) Predicting establishment success for introduced freshwater fishes: a role for climate matching. Biol Invasions 12:2559–2571

Bradie J, Leung B (2017) A quantitative synthesis of the importance of variables used in MaxEnt species distribution models. J Biogeogr 44:1344–1361

Bradie J, Chivers C, Leung B (2013) Importing risk: quantifying the propagule pressure–establishment relationship at the pathway level. Divers Distrib 19:1020–1030

Britton JR, Gozlan RE, Copp GH (2011) Managing non-native fish in the environment. Fish Fish 12:256–274

Buckley YM (2008) The role of research for integrated management of invasive species, invaded landscapes and communities. J Appl Ecol 45:397–402

Cassey P, Delean S, Lockwood JL, Sadowski J, Blackburn TM (2018a) Dissecting the null model for biological invasions: a meta-analysis of the propagule pressure effect. PLoS Biol 16:e2005987

Cassey P, García-Díaz P, Lockwood JL, Blackburn TM (2018b) Invasion biology: searching for predictions and prevention, and avoiding lost causes. Invasion Biol Hypotheses Evid 9:1

Chadès I, Martin TG, Nicol S, Burgman MA, Possingham HP, Buckley YM (2011) General rules for managing and surveying networks of pests, diseases, and endangered species. Proc Natl Acad Sci 108:8323–8328

Chapman FA, Fitz-Coy SA, Thunberg EM, Adams CM (1997) United States of America trade in ornamental fish. J World Aquac Soc 28:1–10

Cook CN, Hockings M, Carter RB (2010) Conservation in the dark? The information used to support management decisions. Front Ecol Environ 8:181–186

Cummings P (2009) The relative merits of risk ratios and odds ratios. Arch Pediatr Adolesc Med 163:438–445

Dawson W, Burslem DF, Hulme PE (2009) Factors explaining alien plant invasion success in a tropical ecosystem differ at each stage of invasion. J Ecol 97:657–665

Della Venezia L, Samson J, Leung B (2018) The rich get richer: invasion risk across North America from the aquarium pathway under climate change. Divers Distrib 24:285–296

Dogliotti A, Gossn J, Vanhellemont Q, Ruddick K (2018) Detecting and quantifying a massive invasion of floating aquatic plants in the río de la plata turbid waters using high spatial resolution ocean color imagery. Remote Sens 10:1140

Drake JM, Lodge DM (2004) Effects of environmental variation on extinction and establishment. Ecol Lett 7:26–30

Duggan IC, Rixon CA, MacIsaac HJ (2006) Popularity and propagule pressure: determinants of introduction and establishment of aquarium fish. Biol Invasions 8:377–382

Duncan RP, Bomford M, Forsyth DM, Conibear L (2001) High predictability in introduction outcomes and the geographical range size of introduced Australian birds: a role for climate. J Anim Ecol 70:621–632

Duncan RP, Blackburn TM, Rossinelli S, Bacher S (2014) Quantifying invasion risk: the relationship between establishment probability and founding population size. Methods Ecol Evol 5:1255–1263

Ehrenfeld JG (2010) Ecosystem consequences of biological invasions. Annu Rev Ecol Evol Syst 41:59–80

Essl F, Dullinger S, Moser D, Steinbauer K, Mang T (2015) Macroecology of global bryophyte invasions at different invasion stages. Ecography 38:488–498

Ficetola GF, Thuiller W, Padoa-Schioppa E (2009) From introduction to the establishment of alien species: bioclimatic differences between presence and reproduction localities in the slider turtle. Divers Distrib 15:108–116

Finnoff D, Shogren JF, Leung B, Lodge D (2007) Take a risk: preferring prevention over control of biological invaders. Ecol Econ 62:216–222

Forsyth DM, Duncan RP, Bomford M, Moore G (2004) Climatic suitability, life-history traits, introduction effort, and the establishment and spread of introduced mammals in Australia. Conserv Biol 18:557–569

Froese R, Pauly D (eds) (2018) FishBase. World Wide Web electronic publication. www.fishbase.org, version (02/2018)

Hanley JA, McNeil BJ (1982) The meaning and use of the area under a receiver operating characteristic (ROC) curve. Radiology 143:29–36

Hayes KR, Barry SC (2008) Are there any consistent predictors of invasion success? Biol Invasions 10:483–506

Herrel A, van der Meijden A (2014) An analysis of the live reptile and amphibian trade in the USA compared to the global trade in endangered species. Herpetol J 24:103–110

Hijmans RJ, Cameron SE, Parra JL, Jones PG, Jarvis A (2005) Very high resolution interpolated climate surfaces for global land areas. Int J Climatol 25:1965–1978

Howeth JG, Gantz CA, Angermeier PL, Frimpong EA, Hoff MH, Keller RP, Mandrak NE, Marchetti MP, Olden JD, Romagosa CM, Lodge DM (2016) Predicting invasiveness of species in trade: climate match, trophic guild and fecundity influence establishment and impact of non-native freshwater fishes. Divers Distrib 22:148–160

Jenkins PT (2013) Invasive animals and wildlife pathogens in the United States: the economic case for more risk assessments and regulation. Biol Invasions 15:243–248

Johnson JB, Omland KS (2004) Model selection in ecology and evolution. Trends Ecol Evol 19:101–108

Keller RP, Perrings C (2011) International policy options for reducing the environmental impacts of invasive species. Bioscience 61:1005–1012

Kerr NZ, Baxter PW, Salguero-Gómez R, Wardle GM, Buckley YM (2016) Prioritizing management actions for invasive populations using cost, efficacy, demography and expert opinion for 14 plant species world-wide. J Appl Ecol 53:305–316

Kolar CS, Lodge DM (2002) Ecological predictions and risk assessment for alien fishes in North America. Science 298:1233–1236

Leung B, Lodge DM, Finnoff D, Shogren JF, Lewis MA, Lamberti G (2002) An ounce of prevention or a pound of

cure: bioeconomic risk analysis of invasive species. Proc R Soc Lond B Biol Sci 269:2407–2413

Leung B, Finnoff D, Shogren JF, Lodge D (2005) Managing invasive species: rules of thumb for rapid assessment. Ecol Econ 55:24–36

Leung B, Roura-Pascual N, Bacher S, Heikkilä J, Brotons L, Burgman MA, Dehnen-Schmutz K, Essl F, Hulme PE, Richardson DM, Sol D (2012) TEASIng apart alien species risk assessments: a framework for best practices. Ecol Lett 15:1475–1493

Lockwood JL, Hoopes MF, Marchetti MP (2013) Invasion ecology. Wiley, London

Lohr CA, Hone J, Bode M, Dickman CR, Wenger A, Pressey RL (2017) Modeling dynamics of native and invasive species to guide prioritization of management actions. Ecosphere 8:e01822

Mahoney PJ, Beard KH, Durso AM, Tallian AG, Long AL, Kindermann RJ, Nolan NE, Kinka D, Mohn HE (2015) Introduction effort, climate matching and species traits as predictors of global establishment success in non-native reptiles. Divers Distrib 21:64–74

Marchetti MP, Moyle PB, Levine R (2004) Invasive species profiling? Exploring the characteristics of non-native fishes across invasion stages in California. Freshw Biol 49:646–661

Mehta SV, Haight RG, Homans FR, Polasky S, Venette RC (2007) Optimal detection and control strategies for invasive species management. Ecol Econ 61:237–245

Milbau A, Stout JC (2008) Factors associated with alien plants transitioning from casual, to naturalized, to invasive. Conserv Biol 22:308–317

Mumbare SS, Maindarkar G, Darade R, Yenge S, Tolani MK, Patole K (2012) Maternal risk factors associated with term low birth weight neonates: a matched-pair case control study. Indian Pediatr 49:25–28

Peoples BK, Midway SR (2018) Fishing pressure and species traits affect stream fish invasions both directly and indirectly. Divers Distrib 24:1158–1168

Peterson DP, Rieman BE, Dunham JB, Fausch KD, Young MK (2008) Analysis of trade-offs between threats of invasion by nonnative brook trout (*Salvelinus fontinalis*) and intentional isolation for native westslope cutthroat trout (*Oncorhynchus clarkii lewisi*). Can J Fish Aquat Sci 65:557–573

Pyšek P, Jarošík V, Pergl J, Randall R, Chytrý M, Kühn I, Tichý L, Danihelka J, Chrtek Jun J, Sádlo J (2009) The global invasion success of Central European plants is related to distribution characteristics in their native range and species traits. Divers Distrib 15:891–903

Pyšek P, Jarošík V, Hulme PE, Pergl J, Hejda M, Schaffner U, Vilà M (2012) A global assessment of invasive plant impacts on resident species, communities and ecosystems: the interaction of impact measures, invading species' traits and environment. Glob Change Biol 18:1725–1737

R Core Team (2018) R: a language and environment for statistical computing. R Foundation for Statistical Computing, Vienna

Ribeiro F, Elvira B, Collares-Pereira MJ, Moyle PB (2008) Life-history traits of non-native fishes in Iberian watersheds across several invasion stages: a first approach. Biol Invasions 10:89–102

Rixon CA, Duggan IC, Bergeron NM, Ricciardi A, Macisaac HJ (2005) Invasion risks posed by the aquarium trade and live fish markets on the Laurentian Great Lakes. Biodivers Conserv 14:1365–1381

Robinson JE, Griffiths RA, John FAS, Roberts DL (2015) Dynamics of the global trade in live reptiles: shifting trends in production and consequences for sustainability. Biol Cons 184:42–50

Ruesink JL (2005) Global analysis of factors affecting the outcome of freshwater fish introductions. Conserv Biol 19:1883–1893

Saul WC, Roy HE, Booy O, Carnevali L, Chen HJ, Genovesi P, Harrower CA, Hulme PE, Pagad S, Pergl J, Jeschke JM (2017) Assessing patterns in introduction pathways of alien species by linking major invasion data bases. J Appl Ecol 54:657–669

Simberloff D (2003) How much information on population biology is needed to manage introduced species? Conserv Biol 17:83–92

Simberloff D, Martin JL, Genovesi P, Maris V, Wardle DA, Aronson J, Courchamp F, Galil B, García-Berthou E, Pascal M, Pyšek P (2013) Impacts of biological invasions: what's what and the way forward. Trends Ecol Evol 28:58–66

Sol D, Maspons J, Vall-Llosera M, Bartomeus I, García-Peña GE, Piñol J, Freckleton RP (2012) Unraveling the life history of successful invaders. Science 337:580–583

Stewart-Koster B, Olden JD, Johnson PT (2015) Integrating landscape connectivity and habitat suitability to guide offensive and defensive invasive species management. J Appl Ecol 52:366–378

U.S. Geological Survey (2017) Nonindigenous aquatic species database. http://nas.er.usgs.gov Gainesville, FL

Vander Zanden MJ, Olden JD (2008) A management framework for preventing the secondary spread of aquatic invasive species. Can J Fish Aquat Sci 65:1512–1522

Vander Zanden MJ, Hansen GJ, Higgins SN, Kornis MS (2010) A pound of prevention, plus a pound of cure: early detection and eradication of invasive species in the Laurentian Great Lakes. J Great Lakes Res 36:199–205

Westbrooks RG (2004) New approaches for early detection and rapid response to invasive plants in the United States. Weed Technol 18:1468–1471

Wittenberg R, Cock MJ (eds) (2001) Invasive alien species: a toolkit of best prevention and management practices. CABI, Abingdon