

*Disregarding human pre-introduction
selection can confound invasive crayfish
risk assessments*

**Yiwen Zeng, Kwek Yan Chong, Erin
K. Grey, David M. Lodge & Darren
C. J. Yeo**

Biological Invasions

ISSN 1387-3547

Biol Invasions
DOI 10.1007/s10530-015-0881-8



 Springer

Your article is protected by copyright and all rights are held exclusively by Springer International Publishing Switzerland. This e-offprint is for personal use only and shall not be self-archived in electronic repositories. If you wish to self-archive your article, please use the accepted manuscript version for posting on your own website. You may further deposit the accepted manuscript version in any repository, provided it is only made publicly available 12 months after official publication or later and provided acknowledgement is given to the original source of publication and a link is inserted to the published article on Springer's website. The link must be accompanied by the following text: "The final publication is available at link.springer.com".

Disregarding human pre-introduction selection can confound invasive crayfish risk assessments

Yiwen Zeng · Kwek Yan Chong · Erin K. Grey · David M. Lodge · Darren C. J. Yeo

Received: 2 April 2014 / Accepted: 16 March 2015
© Springer International Publishing Switzerland 2015

Abstract Trait-based risk assessments of invasive species focus on identifying intrinsic biological or ecological traits associated with invasion success, which allows for a new species' invasion risk to be assessed a priori, thus facilitating cost-effective prevention strategies. However, human preferences for species traits—preferences that might affect which species enter into different pathways of invasion—exist for taxa closely associated with people. Disregarding such preferences can confound correlations between species traits and invasion success. Here we develop a risk assessment for crayfish, a group of culturally and ecologically important decapod

crustaceans with numerous harmful invasive species, that explicitly accounts for species traits as well as human preferences as they are expressed in different pathways (e.g., aquaculture, live angling bait for fishing, harvesting). Our results indicate that species traits and human preferences are confounded for introduction and establishment risk models, but subsequent spread risk is not associated with human preferences and can be predicted by clutch size. Although not commonly addressed, this study demonstrates that accounting for human preferences in trait-based risk assessments is important for taxa closely associated with people, as pre-introduction human selection of traits may bias such analyses.

Electronic supplementary material The online version of this article (doi:[10.1007/s10530-015-0881-8](https://doi.org/10.1007/s10530-015-0881-8)) contains supplementary material, which is available to authorized users.

Keywords Conditional inference trees · Random forest · Introduction pathways · Risk assessment · Introduced species

Y. Zeng · K. Y. Chong · D. C. J. Yeo (✉)
Department of Biological Sciences, National University of Singapore, 14 Science Drive 4, Singapore 117543, Republic of Singapore
e-mail: dbsyeod@nus.edu.sg

E. K. Grey · D. M. Lodge
Environmental Change Initiative and Department of Biological Sciences, University of Notre Dame, 100 Galvin Life Sciences Center, Notre Dame, IN 46556, USA

Present Address:

E. K. Grey
Division of Chemistry and Biological Sciences, 1 University Parkway, University Park, IL 60484, USA

Introduction

Predicting the likelihood of species establishment or invasion can provide guidance for voluntary or regulatory practices to prevent the introduction of species likely to be harmful. By enabling the continued importation of species that are likely to be benign (de Moor 2002) while preventing the importation only of harmful species, the net economic benefits to society are maximized (Keller et al. 2007b). Among the more

widely studied prediction methods are trait-based risk assessment approaches, which rely on the propensity for certain species traits to confer an advantage towards establishment or invasion success (Keller and Drake 2009). Unlike other approaches that utilize species-specific distribution or propagule pressure data (Leung et al. 2006; Herborg et al. 2009), which are often not readily available for global taxon-wide studies, trait-based risk assessments can be performed based on relatively more easily available information.

Trait-based risk assessments have been performed on numerous taxa including plants (Pyšek et al. 1995; Thuiller et al. 2006), birds (Forsyth et al. 2004), mammals (Sol and Lefebvre 2000; Sol et al. 2008), fishes (Kolar and Lodge 2002; Drake 2007), and, to a lesser extent, invertebrates (Keller et al. 2007a; Stanzner et al. 2008; Larson and Olden 2010). Best performed independently for different invasion transitions (e.g., introduction, establishment, spread/impact; Kolar and Lodge 2002; Ribeiro et al. 2008), such studies investigate the importance of aspects of life history (Kolar and Lodge 2002; Stanzner et al. 2008), environmental tolerance (Alcaraz et al. 2005; Moyle and Marchetti 2006), and human preference (Alcaraz et al. 2005; Thuiller et al. 2006) that contribute to species becoming invasive.

Although human preferences for particular biological or physiological traits can exist among species in different pathways (e.g., harvesting for consumptive purposes or aquaculture), little attention has been focused on how these human preferences could be a confounding factor in trait-based risk assessments. Considering that species are often chosen for aquaculture based on certain biological characteristics, e.g., high fecundity and environmental tolerance (Ackefors 1999), such traits would naturally be strongly associated with that pathway. These traits might then be credited for conferring establishment or invasion success (Sakai et al. 2001; Alcaraz et al. 2005; Thuiller et al. 2006), when in fact it was human activity that facilitated this process. Failure to consider human preferences could result in the identification of people's preference for species traits rather than actual traits that account for ecological processes (e.g., ability to survive and reproduce in an alien range) as drivers of invasion success. This subjects such risk assessments to spurious correlations between traits and invasion success, making them susceptible to inaccurate predictions or ineffective management

actions should new introduction pathways develop or human preferences for traits change.

This paper explores and disentangles the potential effect of human preferences on trait-based risk assessments for freshwater crayfish (Crustacea: Decapoda: Astacoidea, Parastacoidea)—a group that has a high degree of cultural influence and is distributed in multiple pathways of intentional introductions (sensu Lodge et al. 2006). Freshwater crayfishes are typically large, long-lived decapod crustaceans that are adapted to survive in mainly temperate zone aquatic, terrestrial, and subterranean habitats (Larson and Olden 2010). They play important ecological roles as keystone species (Momot 1995) and ecosystem engineers (Creed and Reed 2004), and are critical components of ecosystems (Hobbs and Lodge 2010). Crayfish are also culturally significant, constituting an important part of local diets, and have been farmed since medieval times (Skurdal and Taugbøl 1994; Gherardi 2011). In terms of the ecosystem services they provide, crayfish are known to support local food webs (Huryn and Wallace 1987; Griffith et al. 1994) and fisheries (Nielsen and Orth 1988; Roell and Orth 1998).

With their long history of human use, translocation of crayfish has been recorded since the 16th century (Gherardi 2011), and has occurred globally (Hobbs et al. 1989; Lodge et al. 2012). Today, at least 26 species have become established outside their native environments with varying consequences (Lodge et al. 2012). Common reasons for introductions include aquaculture, ornamental trade, and stocking for consumption, all of which provide numerous economic benefits (Lodge et al. 2012). However, many introduced crayfish have also become invasive through their spread and subsequent large negative environmental and economic impacts (Lodge et al. 2012). Invasive crayfish directly affect populations of native crayfish species through competition and predation; as well as indirectly through spread of disease and parasites, and hybridization and introgression (see Gherardi 2007; Lodge et al. 2012). Exotic crayfish also impact entire stream and lake communities, habitats, and ecosystems (Angeler et al. 2001; Geiger et al. 2005) through their habitat modifying capabilities (Angeler et al. 2001), broad diet (Hobbs et al. 1989), and high feeding rates (Bernardo and Ilhéu 1994). Thus, in light of the potential impacts of invasive crayfish, the ability to predict crayfish species' likelihood to establish (or spread) would be an important step in preventing invasion, which would be

highly beneficial for conservation (Lodge et al. 2006; Larson and Olden 2010; Lodge et al. 2012).

To determine if human preference for species traits could have a confounding role in determining crayfish invasion success, we analyze species traits and human uses that predict for species introduction or establishment or spread. We compare models created using three different recursive partitioning tools: classification and regression trees (CART), conditional inference trees (CIT), and random forests (RF). Model comparisons and best model results are presented and compared to results from previous studies. Finally, we suggest how our models could be used to inform efforts to reduce the number of introductions of invasive crayfishes.

Materials and methods

Global species trait database for crayfish

An exhaustive global list of 614 crayfish species (as of 2011) was compiled from online databases and published literature. Taxonomically unresolved species like several *Cambaroides* spp. (see Kawai et al. 2003) and *Austropotamobius italicus* (see Grandjean et al. 2000; Bertocchi et al. 2008) were omitted from the list in order to control for any possible ambiguity related to identities of such species in published biological studies. Thirteen species traits and their definitions were selected based on previously implicated links with invasiveness, and included both categorical and continuous variables (see Online Resource 1): zoogeographical region, native geographic range, habitat preference, substrate preference, burrowing behavior, maximum carapace length, maximum chela length, maximum chela length scaled to maximum carapace length, maximum clutch size, maximum clutch size scaled to maximum carapace length, maximum egg size, minimum reproductive size of females, and brooding season. Eight human-associated traits were also selected, and included both binary and discrete variables (see Online Resource 1): non-commercial harvesting, commercial harvesting, aquaculture, ornamental trade, bait, extirpation history, year of description, and year of first introduction. The five human-use variables (non-commercial harvesting, commercial harvesting, aquaculture, ornamental trade, bait) were included to identify any potential confounding effect of human preference for species traits. The analyses also

explored the possible invasive effects of the trait “number of countries *into which introduction has occurred*”, but these results are presented separately (Online Resource 4) due to the potential for this trait to be a surrogate variable of propagule pressure. To account for the potential effects phylogenetic similarity might have on traits, additional taxonomic factors of “family” and “genus” were included. Trait information was mainly obtained from online databases, published and grey literature (e.g., published theses, technical reports), reliable websites (restricted to those used in previous studies, e.g., Taylor et al. 2007; Faulkes 2010), websites written by established and published crayfish workers (e.g., Lukhaup C, McCormack R), and through personal communication with scientific experts. Aquaculture studies aimed at enhancing survivability and production of captive-bred crayfish, resulting in artificially inflated values of life history traits (fecundity and size), were ignored. Several traits that we would have liked to include (e.g., growth rates, longevity, duration of egg development, age of sexual maturity, time between reproduction events and juvenile survival) were not included because of lack of available data.

Invasion status

The invasive status of crayfish species (within individual countries) was determined from published literature and online databases (see Online Resource 3). Crayfish were classified based on the maximum invasion stage attained in any country, following Blackburn et al. (2011)—*Non-introduced*: when a species is not known to be introduced outside their native range; *Introduced*: when a species is known to have been translocated and released into the wild outside its native range, but has not formed self-sustaining populations; *Established*: when a species has formed self-sustaining but localized populations outside of its native range for at least 5 years, but has not spread into other areas; *Spread*: when a species has formed self-sustaining populations, and is known to have extended its range substantially beyond the initial site of establishment.

Data analyses

To identify pathways and species traits that predict introduction, establishment and spread, we used three decision tree algorithms. The first, Categorical and

Regression Tree (CART), is the most widely used algorithm (e.g., Kolar and Lodge 2002; Larson and Olden 2010) due to its ease of interpretation and ability to deal with missing data and multicollinearity (Speybroeck 2012). Similarly, conditional inference tree (CIT; Hothorn et al. 2006b; Giam et al. 2011; Murray et al. 2011) is another commonly applied technique that can manage relatively smaller sample sizes and trait synergisms (Sodhi et al. 2010). Both methods were performed, but omitted in favor of the better performing random forest (details and results in Online Resource 4).

Random forest (RF), an ensemble machine learning technique (Hothorn et al. 2006b), generates and combines numerous bootstrapped conditional inference trees ($n = 1000$ for this study) based on random re-sampling of variables (Strobl et al. 2007), the outcome of which forms an unbiased, statistically-robust prediction. Owing to the complexity of random forests, an additional step of calculating variable importance values was required to determine the importance of each trait (Strobl et al. 2009). For this we used the Area Under the Receiver Operating Characteristic curve (AUC) metric, which is more robust to imbalanced datasets than other variable importance metrics (Janitzka et al. 2013). The change in AUC was calculated by randomly permutating species trait values among observations and calculating the AUC before and after permutating the species trait averaged over all trees. However, considering the potential for multicollinearity among species traits, the control “mtry” was set at 7, and the conditional method was used in permutation and computing of variable importance (see Strobl et al. 2008).

Predictive accuracy of models was assessed again using AUC and its standard errors (Gibson et al. 2004; Giam et al. 2011). Values of AUC range from 0.5, which indicates a model that performs no better than chance, to 1.0, which has perfect predictive accuracy (Gibson et al. 2004; Giam et al. 2011). Following Robin et al. (2011), we used bootstrapping to compare receiver operating characteristic curves (ROC) across CART, CIT and RF model results to determine if predictive accuracy differed significantly among modeling methods.

To determine the importance of traits for predicting outcomes based on invasion stages (Blackburn et al. 2011), we modeled each transition separately—with each subsequent transition being a subset of the previous: *Transition 1: Non-introduced to Introduced*

(including *Established* and *Spread*); *Transition 2: Introduced but not established to Established* (including *Spread*); and *Transition 3: Established but not spread to Spread* (Fig. 1). These analyses aimed to provide a clear idea of which traits predict the spread of a species and which traits merely predict for establishment or introduction.

A separate series of analyses was then performed to determine if human preference for species traits could confound trait-based risk assessments. Whenever any of the five human-use variables were identified as predictors, re-analyses would be carried out excluding these five variables. Then an additional series of analyses was performed using human uses (e.g., aquaculture or bucket bait) as response variables and species traits as explanatory variables. Applying human uses as response variables would indicate species traits that correlate with introduction pathways. If traits identified in the absence of human uses (diagonal solid arrow in Fig. 1) reflect the determinants of the pathways themselves (vertical dotted arrow in Fig. 1), it would imply a pre-introduction selection for these traits.

All analyses were performed on R version 2.14.1 (R Development Core Team 2011) with the package ‘party’ version 0.9-99996 (Hothorn et al. 2006a, 2006b; Strobl et al. 2007, 2008) and ‘pROC’ (Robin et al. 2011).

Results

Of the 614 species of extant and recently extinct crayfish known (Crandall and Buhay 2008), only 133 species possessed full sets of trait information (Online Resource 2; Online Resource 5). Among these, 98 species were non-introduced, nine species were introduced but non-established, seven were established but not known to have spread, and 19 were known to have substantial spread after establishment (Online Resource 3).

Comparisons between the three tree-based algorithms also revealed that RF had the highest AUC for every transition, and was significantly better than both CART and CIT models for all transitions (Table 1). We therefore report only the results from the RF models, and make CART and CIT results available in Online Resource 4.

Results from *Transition 1* and *Transition 2* identified particular human-uses variables as the most

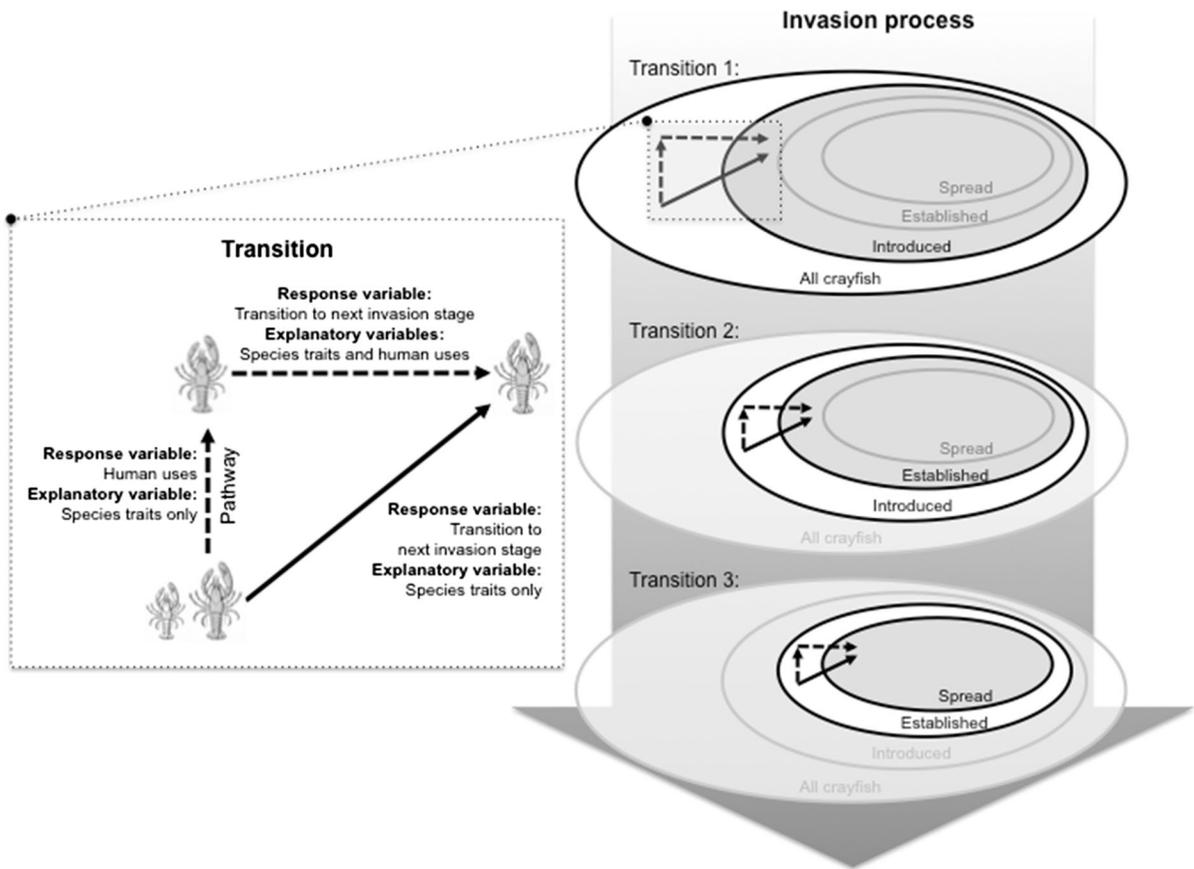


Fig. 1 Schematic representation of the analyses performed in this study. The diagram (right) depicts datasets (*nested ellipses*) used to determine important traits predicting an outcome (*shaded in grey*) of each transition of the invasion process. Between these datasets, *horizontal* and *diagonal arrows* across *shaded* and *non-shaded* datasets represents *Transitions 1, 2, and 3* (species introduction, establishment, and spread). Within each transition (*left*), *solid arrows* represent the analyses that do not take into account human preference, while *dotted arrows* represent the sequence of analyses that account for human preference. The *horizontal arrow* represents analyses with species traits and human uses as explanatory variables, while the

diagonal arrow indicates occasions when only species traits were used. To determine which traits were associated with important pathways (*vertical arrow*), species traits alone were correlated to human use response variable. Human preference refers to human selection of species with a particular set of traits, whereas human use refers to the actual purpose of these species (e.g., aquaculture, ornamental trade; used as a response or explanatory variable). Traits associated with a particular pathway (*vertical dotted arrow*) on the other hand, were determined by analyzing human uses as the response variable and species traits as the explanatory variable

important predictors of species introduction and establishment. Hence we first excluded these variables, and then examined traits predicting for these pathways (when response variables were human uses; vertical dotted arrow in Fig. 1) in *Transition 1* (bait and non-commercial harvesting) and *Transition 2* (aquaculture).

Within *Transition 1* (*Non-Introduced* vs. *Introduced*), species used as bait and involved in non-commercial harvesting, greater maximum clutch size and larger native range were associated with higher

likelihood of species introduction (AUC values: 0.016–0.131). When human-use variables were excluded (diagonal solid arrow in Fig. 1), larger maximum clutch size, maximum chela length, native range and earlier year of description were predictors (AUC values: 0.016–0.30; Table 2). These variables (maximum clutch size, maximum chela length, native range, and year of description) were similarly identified when bait or non-commercial harvesting were used as binary response variables (AUC values: 0.013–0.079; Table 2). This signifies that these traits

Table 1 Model accuracy (measured as AUC) for each transition model created using CART, CIT and RF methods

Method	<i>Transition 1</i>				<i>Transition 2</i>			<i>Transition 3</i>	
	With human uses	Without human uses	Bait	Non-commercial harvest	With human uses	Without human uses	Aquaculture	With human uses	Without human uses
CART	0.58	0.57	0.63	0.54	0.58	0.58	0.52	0.57	0.57
CIT	0.94	0.88	0.72	0.91	0.5	0.5	0.82	0.5	0.5
RF	0.99	0.96	0.96	0.98	0.91	0.89	0.95	0.91	0.94

Random forest AUC values (highlighted in bold) were significantly higher than that of the both CART and CIT (determined by a bootstrap test comparing two ROC curves), across all analyses of transitions. Columns “Bait”, “Non-commercial Harvest”, and “Aquaculture” refer to model accuracies when these variables were used as the response instead of the predictors

are associated with particular pathways, thus indicating the potential for a pre-introduction selection of traits.

Within *Transition 2 (Introduced vs. Established)*, species with larger clutch size (scaled) and used in aquaculture had higher establishment risk according to RF (AUC values: 0.013–0.053). When human-use variables were excluded, scaled clutch size was again identified as a predictor of establishment risk (diagonal solid arrow in Fig. 1), as well as aquaculture (horizontal dotted arrow in Fig. 1; Table 3), indicating that it could be confounded with human preference in assessments of establishment risk.

For *Transition 3 (Established vs. Spread)*, measurements of fecundity (maximum and scaled clutch size) were found to be the most important variables predicting spread risk when human uses were excluded and included (AUC values: 0.018–0.068) (Table 4). Specifically, spread risk was high if the species had >354 eggs per clutch and scaled clutch of greater than 5.8.

Although the number of countries into which introduction had occurred previously was important for establishment and spread (Online Resource 4), it was excluded from the results due to its potential to be a reflection of propagule pressure.

Discussion

Species introduction and the pre-introduction selection of traits

Based on the RF results, which have greater predictive accuracy when compared to CART and CIT (Table 1), bait and non-commercial harvesting were key predictors in *Transition 1*. As such, it is apparent that human

uses are important determinants of species introduction (Table 2). These two human uses intrinsically influence the introduction success of species as they are likely the key introduction pathways. They influence the purpose, means and intensity of introduction events, hence affecting the ability of species to survive and overcome the first geographic barrier to become introduced species.

Species used as live bait (for angling) are typically purchased in bait shops, or collected from neighboring and endemic waterways (DiStefano et al. 2009). They are usually transported alive (Litvak and Mandrak 1993; Ludwig and Leitch 1996) over shorter distances than in the ornamental trade or in aquaculture (Puth and Allen 2004; DiStefano et al. 2009), and are often discarded directly into fishing sites if unused by anglers. These actions increase the chances of crayfish bait species surviving the transportation process, overcoming geographic barriers when the anglers travel across basin boundaries (Ludwig and Leitch 1996). Furthermore, as wild-caught species are commonly used as bait, anglers can also be vectors for secondary transport of alien species (DiStefano et al. 2009). These factors are compounded by the prevalent use of crayfish species as bait in areas like North America (DiStefano et al. 2009) and Australia (Doupé et al. 2004), and the typically lower level of regulation and enforcement when compared with aquaculture or ornamental trade (DiStefano et al. 2009).

The other important introduction pathway (in *Transition 1*) was non-commercial harvesting. Species used in this manner represent those that are harvested, but not for the immediate purpose of trading (although some have been known to trade such species). These species include those exploited non-commercially for food like *Paranephrops planifrons* (Parkyn 2010) or

Table 2 Transition 1: variable importance of species traits in the Random Forest models using the response variables—species introduction success [*Non-introduced* (n = 98) vs.*Introduced* (n = 35)] including and excluding human uses, and the pathways bait and non-commercial harvesting

Variable	Introduction success including human uses	Introduction success excluding human uses	Bait	Non-commercial harvest
Family	0.001	0.000	0.003	0.013
Year of description	0.010	0.021	0.032	0.025
Region	0.000	0.000	0.003	0.022
Range	0.016	0.025	0.024	0.000
Substrate	−0.001	−0.001	−0.001	0.001
Habitat	0.002	0.003	0.001	0.004
Burrow	0.003	0.007	0.030	0.050
Carapace length (max)	0.005	0.003	−0.001	0.051
Chela length (max)	0.006	0.016	0.005	0.026
Chela length (scaled)	0.001	0.001	0.010	0.002
Clutch size (max)	0.025	0.030	0.006	0.079
Clutch size (scaled)	0.005	0.004	0.015	0.008
Egg size (max)	0.001	0.000	0.000	0.001
Reproduction size (min)	0.001	−0.001	0.006	0.000
Brooding season	0.001	0.003	0.000	0.002
Non-commercial harvest	0.018	–	–	–
Commercial harvest	0.004	–	–	–
Aquaculture	0.008	–	–	–
Ornamental	0.001	–	–	–
Bait bucket	0.131	–	–	–
Extirpation	0.000	–	–	–

Variable importance values indicate a reduction in AUC of the model when permuted. Variables with importance values ≥ 0.01 are in bold. Dashes represent traits not included in an analysis

recreation like *Cherax cainii* or *Cherax quadricarinatus* (Beatty 2005; Ah Yong and Yeo 2007), and those collected by hobbyists for personal aquariums like *Cambarellus montezumae* (see Kamita 1973). Such species often have cultural significance to people, as collection of crayfish for consumptive purposes (i.e., non-commercial use on a local scale) and recreation dates back centuries in some cultures (Gherardi 2011). This indicates a high degree of affiliation to people, and an increased likelihood of introduction as such species are transported to new environments through human travel, hence representing a major factor when dealing with species introductions (Jeschke and Strayer 2008).

Despite the logical connections between human preference and species introduction, studies of species traits and invasiveness seldom account for these factors (e.g., Marchetti et al. 2004; Statzner et al. 2008). Such efforts aim to determine intrinsic biological traits that

predispose species to an ability to invade— independent of extrinsic factors like human interest. Consequently, this raises the question of whether human preference could obscure any correlation between species traits and introduction likelihood.

In the absence of human uses, species with larger clutch size, species range, and chela length, along with species that were described earlier (year of description) appear more likely to be introduced. The usual interpretation would be that environmental adaptability associated with larger species ranges could increase the likelihood of surviving transportation events (Via et al. 1995), while higher fecundity (Drake and Lodge 2006; Lockwood et al. 2005) and larger chela sizes (Taylor 1988; Flinders and Magoulick 2003; Nakata and Goshima 2003) indicate an ostensible advantage towards being introduced alive into a new habitat. However, from our results, the effects of identified traits, when mediated through human

Table 3 Transition 2: variable importance of species traits in the Random Forest models using the response variables—species establishment success [Introduced (n = 10) vs.

Established (n = 25)] including and excluding human uses, and the pathway aquaculture

Variable	Establishment success including human uses	Establishment success excluding human uses	Aquaculture
Family	0.000	0.000	0.005
Genus	0.004	0.008	0.018
Year of description	−0.001	−0.001	0.009
Region	0.000	0.000	0.001
Range	0.000	0.000	0.000
Substrate	−0.001	−0.001	0.000
Habitat	−0.001	−0.001	0.006
Burrow	−0.002	−0.004	0.014
Carapace length (max)	−0.001	−0.003	0.010
Chela length (max)	−0.006	−0.013	0.005
Chela length (scaled)	−0.002	−0.001	−0.001
Clutch size (max)	0.001	0.003	0.059
Clutch size (scaled)	0.013	0.025	0.017
Egg size (max)	−0.005	−0.008	−0.001
Reproduction size (min)	−0.016	−0.018	−0.002
Brooding season	−0.002	−0.004	0.000
Year of 1st introduction	−0.001	0.003	–
Non-commercial harvest	0.000	–	–
Commercial harvest	0.002	–	–
Aquaculture	0.053	–	–
Ornamental	0.004	–	–
Bait bucket	−0.001	–	–
Extirpation	0.000	–	–

Variable importance values indicate a reduction in AUC of the model when permuted. Important variables, defined as those ≥ 0.01 are in bold. Dashes represent traits not included in an analysis

selection in particular pathways, appear to be compounded and greater in magnitude than just a straightforward biological advantage to the process of invasion. Models indicate that predictors of use such as bait include year of description, range size and fecundity, while non-commercial harvesting is predicted by year of description, fecundity, and chela length. Species that were described earlier have a longer history of human association, and hence are more likely to be experimented on for use, for example, as bait. Species used for bait could have been selected based on their environmental tolerance (Via et al. 1995; Cabrero-Sañudo and Lobo 2003), behavioral advantage (i.e., competitiveness; Flinders and Magoulick 2003; Nakata and Goshima 2003), and reproductive capabilities, as these would confer breedability and resilience in captivity (Reynolds 2002). This is

especially important for species like *Orconectes rusticus*, and *Orconectes virilis*, which are often bred and sold in bait shops (DiStefano et al. 2009). Additionally, crayfish collected from the wild for consumptive and recreational purposes tend to be of larger sizes (Lodge and Hill 1994) and have a higher reproductive potential (Reynolds 2002). The preference towards species exhibiting particular desirable traits illustrates an important pre-introduction selection bias that could influence similar trait-based assessments.

These findings are supported by several other studies, which identified human preference to be important in the invasion of plants (Thuiller et al. 2006) and fishes (Alcaraz et al. 2005). Yet, many recent studies still fail to account for this and perform risk assessments without explicitly considering the influence of human preferences (e.g., Keller et al. 2011; Van

Table 4 *Transition 3*: variable importance of species traits in the Random Forest models using the response variables—species spread success [*Established* (n = 6) vs. *Spread* (n = 19)] including and excluding human uses

Variable	Spread success including human uses	Spread success excluding human uses
Family	0.000	0.000
Genus	0.003	-0.002
Year of description	0.003	0.003
Region	0.000	0.000
Range	0.000	0.000
Substrate	0.001	-0.002
Habitat	0.002	-0.001
Burrow	-0.003	-0.003
Carapace length (max)	0.004	0.009
Chela length (max)	-0.002	-0.002
Chela length (scaled)	-0.002	-0.003
Clutch size (max)	0.042	0.068
Clutch size (scaled)	0.018	0.023
Egg size (max)	-0.001	-0.002
Reproduction size (min)	-0.002	-0.001
Brooding season	-0.002	-0.001
Year of 1st introduction	-0.001	0.007
Non-commercial harvest	-0.001	-
Commercial harvest	0.000	-
Aquaculture	0.002	-
Ornamental	0.000	-
Bait bucket	0.000	-
Extirpation	0.000	-

Important variables, defined as those ≥ 0.01 are in bold. Dashes represent traits not included in an analysis

Wilgen and Richardson 2012). Considering that such studies can include taxa that have a high degree of human association (Van Wilgen and Richardson 2012), this represents potential bias in trait-based analyses. Pre-introduction selection of traits could potentially exist in numerous other taxa that have human association due to the selection criteria. With the traits already correlated with pathways in a stage preceding species introduction (e.g., chela length and year of description), such bias appears to be present in freshwater crayfish. This bias has also been observed in other taxa (e.g., ornamental plants), whereby more fecund cultivars have been selected for prior to introduction and correlated to increased invasiveness (Kitajima et al. 2006). Failure to account for such mechanisms could lead to traits identified in risk assessments being actually more related to human preferences than to direct ecological processes. Forming predictions based

on such traits could then subject risk assessment accuracy to changes in human preference and developments of new pathways.

Predicting for establishment and spread

When comparing between established and non-established alien species (*Transition 2*), aquaculture and maximum clutch size (scaled) predicted for successful establishment. The importance of these human associations and traits could stem from the associated propagule pressure (e.g., Duggan et al. 2006; Consuegra et al. 2011; Chucholl 2013), and their innate environmental tolerance and reproductive potential (e.g., Ackefors 1999; Skurdal et al. 1999). These would predispose species to overcome survival and reproduction barriers, greatly increasing the probability of forming self-sustaining populations. However,

when such traits were excluded, maximum clutch size was still associated with establishment likelihood. This indicates that, similar to previous studies, fecundity is a good predictor of establishment risk independent of the influence of pathways (e.g., Kolar and Lodge 2001; Drake 2007; Larson and Olden 2010). However, considering that fecundity was found to be associated with aquaculture, human preference could still have an underlying confounding effect, which could compound existing innate traits, increasing the risk of establishment.

The importance of fecundity is further evident in the identification of maximum clutch size and scaled clutch size as predictors of spread (when analyzed with and without pathways). High fecundity could aid the establishment and spread of alien species because of the high propagule pressure or competitive advantage over native species (see Sakai et al. 2001). This innate biological trait could be more beneficial in overcoming the ecological process of dispersal (spread) more than the extrinsic human traits.

Yet, this series of analyses suggests that the artificial selection of species for human purposes affects introduction and establishment stages, and that attributes commonly related to invasion success were already associated with pathways prior to species introduction. This highlights the importance of considering human preferences while performing a multi-step analysis within the invasion process, as emphasized by Sakai et al. (2001). However, while this sequence of analyses is important for taxa in commerce-related (intentional) pathways like crayfish, it might be less so for other taxa in transport-related (unintentional) pathways such as species accidentally introduced through vectors like ballast water (Frischer et al. 2005). Unlike species that are closely affiliated to uses by people, species transported accidentally by ballast water, hull fouling, or as by-catch in fishing, would not be selected by people for a particular purpose, and hence would not be subjected to pre-introduction human preference.

Owing to the increasing risk and problems posed by invasive species, many preventive measures have been suggested to stop or limit their spread (Lodge et al. 2006; Reaser et al. 2008). Among these, risk screening, pathway analysis, and monitoring present cost-effective means of preventing species invasions (Keller et al. 2009). This study highlights important traits to predict which species are likely to be harmful,

and identifies pathways on which management activities might be most effectively focused. With fecundity associated with establishment and spread (*Transition 2* and *3* in Fig. 1), invasion risk of newly translocated species can be assessed. The pathways of bait use, non-commercial harvesting, and aquaculture are also of significance globally. These pathways represent those correlated with higher likelihood of introduction and establishment, where voluntary or regulatory actions might be more cost efficient in lowering the probability of release or escape of crayfish. This is especially important considering the high cost and difficulty of eradication or control of invasive or established species (Lodge et al. 2012).

Ultimately, while assessing the traits important for the invasion process (via a multi-step analysis), it was determined that pathways were key predictors of introduction and establishment, while spread was predicted by intrinsic traits or introduction history. Additional analyses also revealed a pre-introduction selection of traits for crayfish. This highlights the potential confounding effect of human preferences in risk assessments, and illustrates the need for multi-step analyses to include such factors to account for this bias (at least for taxa that are closely associated with humans). Finally, this study identifies important species traits (fecundity), as well as the key pathways of introduction, and indicates a priority for which authorities can monitor and coordinate specific actions.

Acknowledgments We thank Xingli Giam, Reuben P. Keller, and L. Roman Carrasco for their advice and help; Tadashi Kawai (Wakkanai Fisheries Research Institute) and Tyler Black (North Carolina Wildlife Resources Commission) for information on *Cambaroides schrenckii* and *Cambarus* species, respectively. We also thank the reviewers who contributed greatly to the improvement of this paper. The University of Notre Dame Environmental Change Initiative supported Grey and Lodge. Support from the Ah Meng Memorial Conservation Fund (National University of Singapore Grant No. R-154-000-617-720), an AcRF Tier 1 grant from the Singapore Ministry of Education (NUS Grant Number R-154-000-465-133), North Carolina Survey Group, ECI, and a US EPA GLRI grant are also acknowledged.

References

- Ackefors H (1999) The positive effects of established crayfish introductions in Europe. In: Gherardi F, Holdich DM (eds) Crayfish in Europe as alien species: how to make the best out of a bad situation. A.A. Balkema Publishers, Rotterdam, pp 49–63

- Ahyong S, Yeo DCJ (2007) Feral populations of the Australian red-claw crayfish (*Cherax quadricarinatus* von Martens) in water supply catchments of Singapore. *Biol Invasions* 9:943–946
- Alcaraz C, Vila-Gispert A, García-Berthou E (2005) Profiling invasive fish species: the importance of phylogeny and human use. *Divers Distrib* 11:289–298
- Angeler DG, Sánchez-Carrillo S, García G, Alvarez-Cobelas M (2001) The influence of *Procambarus clarkii* (Cambaridae, Decapoda) on water quality and sediment characteristics in a Spanish floodplain wetland. *Hydrobiologia* 464:9–98
- Beatty S (2005) Translocations of freshwater crayfish: contributions from life histories, trophic relations and diseases of three species in Western Australia. PhD thesis, Murdoch University
- Bernardo JM, Ilhéu M (1994) Red swamp crayfish (*Procambarus clarkii*): contribution to material cycling. *Verh Internat Verein Limnol* 25:2447–2449
- Bertocchi S, Brusconi S, Gherardi F, Grandjean F, Souty-Grosset C (2008) Genetic variability in the threatened crayfish *Austropotamobius italicus* in Tuscany: implications for its management. *Fund Appl Limnol* 173:153–164
- Blackburn TM, Pyšek P, Bacher S et al (2011) A proposed unified framework for biological invasions. *Trends Ecol Evol* 26:333–339
- Cabrero-SaÑudo FJ, Lobo JM (2003) Estimating the number of species not yet described and their characteristics: the case of Western Palearctic dung beetle species (Coleoptera, Scarabaeoidea). *Biodivers Conserv* 12:147–166
- Chucholl C (2013) Invaders for sale: trade and determinants of introduction of ornamental freshwater crayfish. *Biol Invasions* 15:125–141
- Consuegra S, Phillips N, Gajardo G, de Leaniz CG (2011) Winning the invasion roulette: escapes from fish farms increase admixture and facilitate establishment of non-native rainbow trout. *Evol Appl* 4:660–671
- Crandall KA, Buhay JE (2008) Global diversity of crayfish (Astacidae, Cambaridae, and Parastacidae—Decapoda) in freshwater. *Hydrobiologia* 595:295–301
- Creed RP Jr, Reed JM (2004) Ecosystem engineering by crayfish in a headwater stream community. *J N Am Benthol Soc* 23:224–236
- de Moor I (2002) Potential impacts of alien freshwater crayfish in South Africa. *Afr J Aquat Sci* 27:125–139
- DiStefano RJ, Litvan ME, Horner PT (2009) The bait industry as a potential vector for alien crayfish introductions: problem recognition by fisheries, agencies and a Missouri evaluation. *Fisheries* 24:586–597
- Doupe RG, Morgan DL, Gill HS, Rowland AJ (2004) Introduction of redclaw crayfish *Cherax quadricarinatus* (von Martens) to Lake Kununurra, Ord River, Western Australia: prospects for a ‘yabby’ in the Kimberley. *J Proc R Soc West Aust* 87:187–191
- Drake JM (2007) Parental investment and fecundity, but not brain size, are associated with establishment success in introduced fishes. *Funct Ecol* 21:963–968
- Drake JM, Lodge DM (2006) Forecasting potential distributions of nonindigenous species with a genetic algorithm. *Fisheries* 31:9–16
- Duggan IC, Rixon CAM, MacIsaac HJ (2006) Popularity and propagule pressure: determinants of introduction and establishment of aquarium fish. *Biol Invasions* 8:377–382
- Faulkes Z (2010) The spread of the parthenogenetic marbled crayfish, *Marmorkrebs* (*Procambarus* sp.), in the North American pet trade. *Aquat Invasion* 4:447–450
- Flinders CA, Magoulick DD (2003) Effectiveness of stream permanence on crayfish community structure. *Am Midl Nat* 149:134–147
- 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
- Frischer ME, McGrath BR, Hansen AS, Vescio PA, Wyllie JA, Wimbush J, Nierzwicki-Bauer SA (2005) Introduction pathways, differential survival of adult and larval Zebra Mussels (*Dreissena polymorpha*), and possible management strategies, in an Adirondack Lake, Lake George, NY. *Lake Reserv Manag* 21:391–402
- Geiger W, Alcorlo P, Baltanás A, Montes C (2005) Impact of an introduced Crustacean on the trophic webs of Mediterranean wetlands. *Biol Invasions* 7:49–73
- Gherardi F (2007) The impact of invasive crayfish. In: Gherardi F (ed) *Biological invaders in inland waters: profiles, distribution and threats*. Springer, Dordrecht, pp 507–542
- Gherardi F (2011) Towards a sustainable human use of freshwater crayfish (Crustacea, Decapoda, Astacidea). *Knowl Manag Aquat Ecosyst* 401:1–22
- Giam X, Ng TH, Lok AFSL, Ng HH (2011) Local geographic range predicts freshwater fish extinctions in Singapore. *J Appl Ecol* 48:356–363
- Gibson LA, Wilson BA, Cahill DM, Hill J (2004) Spatial prediction of rufous bristlebird habitat in a coastal heathland: a GIS-based approach. *J Appl Ecol* 41:213–223
- Grandjean F, Harris DJ, Souty-Grosset C, Crandall KA (2000) Systematics of the European endangered crayfish species *Austropotamobius pallipes* (Decapoda: Astacidae). *J Crustac Biol* 20:522–529
- Griffith MB, Perry SA, Perry WB (1994) Secondary production of macroinvertebrate shredders in headwater streams with different baseflow alkalinity. *J N Am Benthol Soc* 13:345–356
- Herborg L-M, Drake JM, Rothlisberger JD, Bossenbroek JM (2009) Identifying suitable habitat for invasive species using ecological niche models and the policy implications. In: Keller RP, Lodge DM, Lewis MA, Shogren JF (eds) *Bioeconomics of invasive species: integrating ecology, economics, policy, and management*. Oxford University Press, New York, pp 44–62
- Hobbs HH III, Lodge DM (2010) Decapoda. In: Thorp JH, Covich AP (eds) *Ecology and classification of North American freshwater invertebrates*, 3rd edn. Elsevier Inc, London, pp 901–968
- Hobbs HH III, Jass JP, Huner JV (1989) A review of global crayfish introductions with particular emphasis on two North American species (Decapoda, Cambaridae). *Crustaceana* 56:299–316
- Hothorn T, Buehlmann P, Dudoit S, Molinaro A, Van Der Laan M (2006a) Survival ensembles. *Biostatistics* 7:355–373

- Hothorn T, Hornik K, Zeileis A (2006b) Unbiased recursive partitioning: a conditional inference framework. *J Comput Graph Stat* 15:651–674
- Huryh AD, Wallace JB (1987) Production and litter processing by crayfish in an Appalachian mountain stream. *Freshw Biol* 18:277–283
- Janitzka S, Strobl C, Boulesteix A-L (2013) An AUC-based permutation variable importance measure for random forests. *BMC Bioinformatics* 14:119
- Jeschke JM, Strayer DL (2008) Are threat status and invasion success two sides of the same coin? *Ecography* 31:124–130
- Kamita T (1973) Crayfishes brought to Japan from Mexico. *Zool Mag* 83:123–126
- Kawai T, Machino Y, Ko HS (2003) Reassessment of *Cambaroides dauricus* and *C. schrenckii* (Crustacea: Decapoda: Cambaridae). *Korean J Biol Sci* 7:191–196
- Keller RP, Drake JM (2009) Trait-based risk assessment for invasive species. In: Keller RP, Lodge DM, Lewis MA, Shogren JF (eds) *Bioeconomics of invasive species: integrating ecology, economics, policy, and management*. Oxford University Press, New York, pp 44–62
- Keller RP, Drake JM, Lodge DM (2007a) Fecundity as a basis for risk assessment of nonindigenous freshwater molluscs. *Conserv Biol* 21:191–200
- Keller RP, Lodge DM, Finnoff DC (2007b) Risk assessment for invasive species produces net bioeconomic benefits. *PNAS* 104:203–207
- Keller RP, Lodge DM, Lewis MA, Shogren JF (2009) *Bioeconomics of invasive species: integrating ecology, economics, policy, and management*. Oxford University Press, New York
- Keller RP, Kocov D, Džeroski S (2011) Trait-based risk assessment for invasive species: high performance across diverse taxonomic groups, geographic ranges and machine learning/statistical tools. *Divers Distrib* 17:451–461
- Kitajima K, Fox AM, Sato T, Nagamatsu D (2006) Cultivar selection prior to introduction may increase invasiveness: evidence from *Ardisia crenata*. *Biol Invasions* 8:1471–1482
- Kolar CS, Lodge DM (2001) Progress in invasion biology: predicting invaders. *Trends Ecol Evol* 16:199–204
- Kolar CS, Lodge DM (2002) Ecological predictions and risk assessment of alien fishes in North America. *Science* 298:1233–1236
- Larson ER, Olden JD (2010) Latent extinction and invasion risk of crayfishes in the southeastern United States. *Conserv Biol* 24:1099–1110
- Leung B, Bossenbroek JM, Lodge DM (2006) Boats, pathways, and aquatic biological invasions: estimating dispersal potential with gravity models. *Biol Invasions* 8:241–254
- Litvak MK, Mandrak NE (1993) Ecology of freshwater baitfish use in Canada and the United States. *Fisheries* 18:6–13
- Lockwood JL, Cassey P, Blackburn T (2005) The role of propagule pressure in explaining species invasions. *Trends Ecol Evol* 20:223–228
- Lodge DM, Hill AM (1994) Factors governing species composition, population size, and productivity of cool-water crayfishes. *Nord J Freshw Res* 69:111–136
- Lodge DM, Williams S, MacIsaac H et al (2006) Biological invasions: recommendations for U.S. policy and management. *Ecol Appl* 16:2035–2054
- Lodge DM, Deines A, Gherardi F et al (2012) Global introductions of crayfishes: evaluating the impact of species invasions on ecosystem services. *Annu Rev Ecol Evol Syst* 43:449–472
- Ludwig HR Jr, Leitch JA (1996) Interbasin transfer of aquatic biota via anglers' bait buckets. *Fisheries* 25:21–23
- 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
- Momot WT (1995) Redefining the role of crayfish in aquatic ecosystems. *Res Fish Sci* 3:33–63
- Moyle PB, Marchetti MP (2006) Predicting invasion success: freshwater fishes in California as a model. *Bioscience* 56:515–524
- Murray KA, Rosauer D, McCallum H, Skerratt LF (2011) Integrating species traits with extrinsic threats: closing the gap between predicting and preventing species declines. *Proc R Soc B* 278:1515–1523
- Nakata K, Goshima S (2003) Competition for shelter of preferred species between the native crayfish species *Cambaroides japonicus* and the alien crayfish species *Pacifastacus leniusculus* in Japan in relation to prior residence, sex difference, and body size. *J Crustac Biol* 23:897–907
- Nielsen LA, Orth DJ (1988) The hellgrammite-crayfish bait fishery of the New River and its tributaries, West Virginia. *N Am J Fish Manag* 8:317–324
- Parkyn S (2010) *Paraneohrops planifrons*. IUCN 2013. IUCN Red List of Threatened Species. Version 2013.2. <http://www.iucnredlist.org/details/153750/0>. Accessed 10 Feb 2014
- Puth LM, Allen TFH (2004) Potential corridors for the rusty crayfish, *Orconectes rusticus*, in northern Wisconsin (USA) lakes: lessons for exotic invasions. *Landscape Ecol* 20:567–577
- Pyšek P, Prach K, Smilauer P (1995) Relating invasion success to plant traits: an analysis of the Czech alien flora. In: Pyšek P, Prach K, Rejmanek M, Wade M (eds) *Plant invasions—general aspects and special problems*. SPB Academic, Amsterdam, pp 39–60
- R Development Core Team (2011) R: a language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. <http://www.R-project.org/>. Assessed 23 March 2013
- Reaser JK, Meyerson LA, Von Holle B (2008) Saving camels from straws: how propagule pressure-based prevention policies can reduce the risk of biological invasion. *Biol Invasions* 10:1085–1098
- Reynolds JD (2002) Growth and reproduction. In: Holdich DM (ed) *Biology of freshwater crayfish*. Blackwell Science Ltd, Oxford, pp 152–191
- 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
- Robin X, Turck N, Hainard A, Tiberti N, Lisacek F, Sanchez JC, Müller M (2011) pROC: an open-source package for R and S+ to analyze and compare ROC curves. *BMC Bioinformatics* 12:77
- Roell MJ, Orth DJ (1998) Indirect effects of fishery exploitation and pest control in a riverine food web. *N Am J Fish Manag* 18:337–346

- Sakai AK, Allendorf FW, Holt JS et al (2001) The population biology of invasive species. *Annu Rev Ecol Syst* 32:305–332
- Skurdal J, Taugbøl T (1994) Minimum size regulation as a tool in crayfish management practice. *Nord J Freshw Res* 69:144–148
- Skurdal J, Taugbøl T, Burba A, Edsman L, Söderback B, Styrshave B, Tuusti J, Westman K (1999) Crayfish introductions in the Nordic and Baltic countries. In: Gherardi F, Holdich DM (eds) *Crayfish in Europe as alien species: How to make the best out of a bad situation*. A. A. Balkema Publishers, Rotterdam, pp 193–220
- Sodhi NS, Wilcove DS, Lee TM et al (2010) Deforestation and avian extinction on tropical landbridge islands. *Conserv Biol* 24:1290–1298
- Sol D, Lefebvre L (2000) Behavioural flexibility predicts invasion success in birds introduced to New Zealand. *Oikos* 90:599–605
- Sol D, Bacher S, Reader SM, Lefebvre L (2008) Brain size predicts the success of mammal species introduced into novel environments. *Am Nat* 172:63–71
- Speybroeck N (2012) Classification and regression trees. *Int J Public Health* 57:243–246
- Statzner B, Bonada N, Dolédec S (2008) Biological attributes discriminating invasive from native European stream macroinvertebrates. *Biol Invasions* 10:517–530
- Strobl C, Boulesteix AL, Zeileis A, Hothorn T (2007) Bias in random forest variable importance measures: illustrations, sources and a solution. *BMC Bioinformatics* 8:25
- Strobl C, Boulesteix AL, Kneib T, Augustin T, Zeileis A (2008) Conditional variable importance for random forests. *BMC Bioinformatics* 9:307
- Strobl C, Hothorn T, Zeileis A (2009) Party on! A new, conditional variable importance measure for random forests available in the party Package. University of Munich Department of Statistics, Technical Report Number 50
- Taylor RC (1988) Population dynamics of the crayfish *Procambarus spiculifer* observed in different-sized streams in response to two droughts. *J Crustac Biol* 8:410–419
- Taylor CA, Schuster GA, Cooper JE et al (2007) A reassessment of the conservation status of crayfishes of the United States and Canada after 10+ years of increased awareness. *Fisheries* 32:372–389
- Thuiller W, Richardson DM, Rouget M, Procheş Ş, Wilson JRU (2006) Interactions between environment, species traits, and human uses describe patterns of plant invasions. *Ecology* 87:1755–1769
- Van Wilgen NJ, Richardson DM (2012) The roles of climate, phylogenetic relatedness, introduction effort and reproductive traits in the establishment of non-native reptiles and amphibians. *Conserv Biol* 26:267–277
- Via S, Gomulkiewicz R, Jong GD, Scheiner SM, Schlichting CD, Tienderen PHV (1995) Adaptive phenotypic plasticity: consensus and controversy. *Trends Ecol Evol* 10:212–217