Title: Predicting invasiveness of exotic woody species using a traits-based framework

Running title: Predicting invasiveness of exotic plants

Gabriela C. Nunez-Mir[1], Qinfeng Guo[2], Marcel Rejmánek[3], Basil V. Iannone III[4], Songlin Fei[1]*

[1] Department of Forestry and Natural Resources, Purdue University, West Lafayette, IN 47906, USA. Email: gnunezmi@purdue.edu

[2] USDA Forest Service, Eastern Forest Environmental Threat Assessment Center, 3041 East Cornwallis Road, Research Triangle Park, North Carolina 27709, USA. Email: qinfeng.guo@usda.gov

[3] Department of Evolution and Ecology, University of California, Davis, CA 95616, USA. Email: mrejmanek@ucdavis.edu

[4] School of Forest Resources and Conservation, University of Florida, Gainesville, FL 32611, USA. Email: biannone@ufl.edu

* Corresponding author. Email: sfei@purdue.edu. Address: PFEN 221E, 715 West State Street, West Lafayette, IN 47907-2061, USA. Office phone: 765-496-2199.

Abstract

Identifying potentially invasive species and preventing their introduction and establishment are of critical importance in invasion ecology and land management. Although an extensive body of research has been dedicated to identifying traits that confer invasiveness, our current knowledge is still often inconclusive due to limitations in geographic extent and/or scope of traits analyzed. Here, using a comprehensive set of 45 traits, we performed a case study of invasive traits displayed by exotic woody plants in the United States (U.S.) by comparing 63 invasive and 794 non-invasive exotic woody plant species naturalized across the country. We found that invasive woody species often bear the following two key traits: vegetative reproduction and long-distance seed dispersal (via water, birds or mammals). Boosted classification tree models based on these traits accurately predicted species invasiveness (86% accuracy on average). Presented findings provide a generalized understanding of the relative importance of functional traits in identifying potentially invasive woody species in the U.S. The knowledge generated in this study can be used to improve current classification systems of non-native woody plants used by various U.S. governmental agencies and land managers.

Keywords: boosted classification trees, dispersal vectors, invasion screening tools, invasive plants, invasiveness, multivariate statistics, non-invasive plants, vegetative reproduction

Introduction

One of the major threats in the Anthropocene is the increasing rate and impacts of biological invasions (Vitousek et al. 1997, Fei et al. 2014, Bellard et al. 2016). Preventing the introduction and establishment of invasive species is paramount, as eradication is often impossible due to high labor and economic costs, making the identification of potential invaders an important priority (Rejmánek and Pitcairn 2002, Panetta 2015). Effective screening tools have long been sought to assess the potential invasiveness (i.e. potential for spread and/or impact) of exotic species (e.g. the Australian Weed Risk Assessment; Pheloung et al. 1999, He et al. 2018; and the Environmental and Socio-economic Impact Classification of Alien Taxa (EICAT/SEICAT) tools; Blackburn et al. 2014, Bacher et al. 2018). Most existing screening tools, however, are based primarily on *a priori* assumed importance of individual traits and naturalization history beyond native ranges (Gordon et al. 2008, Koop et al. 2012, Conser et al. 2015).

Determining what makes some species more invasive than others continues to be a major challenge in invasion ecology. Ever since the publication of Baker's Law (Baker 1955, Stebbins 1957) (i.e. species capable of uniparental reproduction are more likely to establish after long-distance dispersal than species that rely on suitable mates and pollinators), a considerable amount of research has been dedicated to isolating attributes that characterize successful invaders (Rejmánek 1996, van Kleunen et al. 2010, Rejmánek 2013, van Kleunen et al. 2015, Miller et al. 2017, Klinerová et al. 2018). Unfortunately, previous research on invasive traits made slow progress on identifying key invasion traits, spurring a pessimistic outlook on invader prediction over wide taxonomic groups, such as angiosperms (Williamson 1996, Thompson and Davis 2011). Improved data availability and accessibility through online databases, along with the advancement of computational capabilities and statistical techniques, has opened doors for a

63 new, more promising era of invasive traits research featuring comparative multispecies studies
64 (Gallagher et al. 2015, Heger et al. 2015). Reviews and meta-analytic syntheses of this new
65 research highlight the possibility for case studies capable of providing insights into
66 generalizations for wide groups of organisms (e.g. woody plants) (see Kolar and Lodge 2001,
67 van Kleunen et al. 2010), demonstrating the value of a traits-based approach to predicting
68 invasions.

69 Nevertheless, we still do not have adequate knowledge to generalize the key traits of
70 woody plant species that make certain species invasive. Woody invaders are of particular
71 interest, as forest ecosystems, once thought to be resistant, are now known to be vulnerable to
72 shade-tolerant invasive species that pose a threat to biodiversity (Woods 1993, Hutchinson and
73 Vankat 1997, Silander and Klepeis1999, Mascaro and Schnitzer 2007). Our current knowledge
74 of determinants of woody plant invasiveness is often inconclusive, partially due to the fact that
75 the majority of studies are limited in geographic extent, the number of plant species or genera
76 analyzed, and the breadth of traits included. Moreover, although the most straightforward way to
77 identify invasive attributes is to compare differences in invasiveness among exotic species with
78 overlapping introduced ranges, most previous studies feature comparative approaches that are
79 less effective for this purpose, such as native-exotic comparisons (e.g. van Kleunen et al. 2010).

80 Currently, out of the 887 exotic woody plants found in the U.S., 63 are officially
81 recognized as invasive by governmental agencies. These 63 species satisfy two criteria: 1) being
82 exotic to the United States and 2) being likely to cause economic or environmental harm (Ries et
83 al. 2004). Given this definition, many exotic species are not identified as harmful invasives until
84 quantifiable damage has already been done. Likewise, the potential risks of exotic species may
85 be inaccurately assessed due to the invasion lag phase that many exotic species go through (i.e.

introduced exotic species may persist in low numbers for decades before spreading exponentially) (Crooks 2005). The ability to accurately predict invasiveness through traits may improve the classification systems used by governmental organizations, ensuring that harmful woody plant species are detected before they cause serious impacts and that critical windows for action are not missed. A better understanding of the relative importance of functional traits as determinants of invasiveness can help to improve existing screening tools.

Here, we aim to address the limitations of previous studies by using a dataset containing all woody plants currently defined as "invasive" (i.e. 63 species) and most "non-invasive" (i.e. 794 out of the 824 species) species that are present across the United States based on 45 functional traits and characteristics. The breadth of species and spatial coverage in our dataset provided a unique opportunity to make useful generalizations for this taxonomic group, allowing us to produce an effective model to predict potentially impactful woody invaders. The knowledge produced by this model could guide and improve existing screening tools and current classification system used by U.S. governmental agencies and land managers to define invasive species by demonstrating the utility of traits for identifying invasive woody species.

Methods

Traits database and invasive species list

We compiled a database of 45 quantitative and qualitative traits (Table 1) for 794 non-invasive and 63 invasive exotic woody species (Fei et al. 2019). The traits we compiled fall loosely within the categories of morphological traits, reproductive traits, pollination methods, dispersal vectors, flowering characteristics, and physiological and environmental tolerance. We included multiple traits related to the same ecological/biological function in order to build enough redundancy into

the database to ensure that the signal of any function related to invasiveness is detected. These traits were compiled from multiple sources, including online traits databases, such as the Flora of North America, Flora of China, Centre for Agriculture and Bioscience International (CABI), and grey literature. For more information regarding these traits, including units of measurement and descriptions, please refer to Appendix S1.

The species in this database comprise exotic woody species that have been detected across the U.S. Of all the exotic species detected, 63 are officially recognized as "invasive" since they fit the criteria posed by Executive Order 13112, and are therefore prioritized by federal and state programs. Based on the definition postulated by the Executive Order, "invasive species" are not only naturalized exotic species with spreading populations (invaders *sensu* Richardson et al. 2000), but are also assumed to have negative environmental and/or economic impact. The remaining 824 exotic woody species comprise only naturalized species (i.e. exotic species with self-sustaining populations). It is worth noting that some of these species may be undergoing invasional lag, and therefore may become "invasive" in the future. A complete list of all invasive and non-invasive species included in our database can be found in Appendix S2. We chose to use this definition of "invasive" as it is the definition used by all major U.S. governmental agencies, and therefore the most relevant definition to management and conservation in the U.S.

Identifying key invasive traits

The large number of variables in our dataset leads to a high level of multidimensionality, which poses analytical challenges (e.g. noise from unimportant variables). To address this issue, we implemented a systematic, sequential process of multivariate analyses to identify the most important variables for differentiating invasive and non-invasive woody plants. We chose this

multi-step strategy over using a single variable reduction technique in order to minimize human intervention aand subjectivity (i.e. through the selection of parameters) and to ensure the robustness of our results across widely utilized statistical techniques. The multivariate techniques used were selected for their ability to handle mixed data types (i.e. continuous and categorical). These techniques were performed on a dissimilarity matrix constructed by calculating Gower's distances among investigated species. Gower's distance is designed to handle mixed data and accept missing values, which is a common problem with large traits databases (Gower 1971, Pavoine et al. 2009).

Since we did not have information on all traits for all species, we took a subset of the data that only included species with complete data for all categorical traits. Although the distance metric used (Gower's distance) is able to handle missing values, the algorithm to calculate Gower's distance treats missing categorical values as a factor level for these variables, causing undesirable artifacts (i.e. species with the same missing values clustered in trait space). The complete subset included 51 invasive and 109 non-invasive woody species, all displaying complete data for the same 45 traits (the vast majority of excluded species were included in subsequent analyses as described below).

We then performed a series of distance-based redundancy analyses (Legendre and Anderson 1999; R package "vegan", Oksanen et al. 2013) aimed at identifying the subset of traits that best explained the separation between invasive and non-invasive woody species in a multivariate trait space. Distance-based redundancy analysis is a constrained ordination technique that is performed on a distance matrix and therefore can be used on datasets comprised of mixed data types. We executed these analyses in both a manual forward and backward

154 selection manner, by adding (i.e. forward) or removing (i.e. backward) variables sequentially
155 from the dataset before rerunning the redundancy analysis.
156 If adding or removing a variable increased or decreased, respectively, the variation
157 explained by invasive status (R^2), the variable was identified as influential and selected for
158 further analysis. The selected traits, marked in Table 1 with an asterisk, improved the model's R^2
159 (i.e. 24 traits improved the model's R^2). Upon further investigation of these selected variables,
160 we found that three of the 24 variables had the same value for all but one species (e.g. "No" for
161 167 species, "Yes" for 1). These three variables ("animal (other) pollinated", "animal
162 (nonspecific) pollinated", and "animal (other) dispersed seed") were excluded from all further
163 analyses.
164 Using a refined version of the complete subset containing only the remaining 21 traits,
165 we performed a principal coordinates analysis (PCoA) using a Gower's dissimilarity matrix (R
166 package "FD", Laliberté et al. 2014) to determine the major axes of variance in the refined data.
167 We were particularly interested in identifying whether invasion status represented one of these
168 axes. To identify the traits that contributed the highest loadings on the principal coordinates of
169 interest, we performed contingency analyses on the categorical traits and correlations between
170 numerical traits and the scores of the principal coordinate of interest. We assumed that traits
171 having p-values ≤ 0.05 to be most strongly related to the separation among invasive and non-
172 invasive species revealed by our PCoA (sensu Tecco et al. 2013).
173
174 *Evaluating predictive power of key invasive traits*
175 After identifying the six traits that were most strongly related to separation among invasive and
176 non-invasive species (i.e. those having p-values ≤ 0.05), we focused on collecting data to

complete the missing values in our dataset for these six traits, increasing our number of species from 51 (invasive) and 109 (non-invasive), to 63 and 794, respectively. We then performed boosted classification tree models (R package "dismo", Hijmans et al. 2017) to evaluate the predictive power of modelling these traits. Boosted classification trees are improved versions of simple classification tree models that use machine-learning to optimize predictive performance by integrating large numbers of simple tree models in an adaptive manner (i.e. iteratively addressing poorly modelled observations and outliers) (Elith et al. 2008). Our boosted classification trees were trained on 70% of the dataset, and then tested on the remaining 30%. We evaluated the predictive ability of the model using several metrics: sensitivity (i.e. true positive rate—proportion of invasives identified as such), specificity (i.e. true negative rate—proportion of non-invasives identified as such), and total accuracy (i.e. proportion of correctly identified species). The first two metrics, sensitivity and specificity, are derived from the Receiver Operating Characteristic (ROC) curve at the optimal classification threshold (i.e. point at which both sensitivity and specificity are maximized). We accounted for the stochasticity of boosted classification tree models by averaging the results over 100 boosted classification tree models.

Accounting for minimum residence time

Residence time is an important factor of invasion success, as species need to overcome multiple barriers not only to become naturalized, but also to surpass invasion lag phases before becoming widespread (Wilson et al. 2007, Pemberton and Liu 2009, Gallagher et al. 2015). Some of the species classified as non-invasive in our study (i.e. as per Executive Order 13112) could be undergoing invasional lag, and therefore may become "invasive" in the future. To ensure that the

patterns observed were not influenced by any potential misclassifications due to invasional lags, we performed another set of boosted classification tree analyses on the same six traits, but only including non-invasive species that have been in the U.S. for at least one century (i.e. 313 non-invasives and the same 63 invasives). Minimum residence time (Rejmánek et al. 2013) of species in our database was obtained from herbarium records (i.e. year of earliest herbarium record) or year of first introduction for species with well-recorded introduction histories in the United States. Our sources include the plant collections of the Smithsonian Institution and the Consortium of Northeastern Herbaria. See Appendix S3 for the list of invasive and non-invasive species used in this analysis and their respective minimum residence times.

Accounting for environmental context

Since invasions do not occur in standardized environments, the characteristics of recipient systems are likely to play a role in invasion dynamics. Invasion impacts have been found to be context-dependent (Pyšek et al. 2012), suggesting that the association between species traits and invasiveness may also be influenced by the characteristics of the recipient systems. To account for the role of environmental context in the patterns observed, we obtained invasive range data for 650 of the 887 exotic woody plants present in the United States from The Biota of North America Program's Plant Atlas (http://bonap.net/NAPA/Genus/Traditional/County). We then divided the United States into loosely defined regions: Florida, the Southeast, the Southwest, the Great Plains, the Northeast, the Northwest, and California. Florida and California were considered separate regions as a large number of exotic species were exclusively found in those states and the states were distinct enough environmentally. To test if observed patterns of trait invasiveness were applicable across regions, we repeated the same boosted classification tree

223 analyses, as described above, for each region separately (except for the Great Plains and the
224 Southwest, as they did not have enough exotic species for the machine-learning method).
225
226 *Testing for phylogenetic relatedness as a confounding factor*
227 Because of the high relatedness among species with a shared phylogenetic history, species in
228 multispecies comparative approaches, such as the one in this study, should not be assumed to be
229 independent observations (Freckleton 2000). We performed analyses to test the robustness of the
230 patterns observed in our study against the signals of family membership and primary growth
231 form as proxies for phylogenetic relatedness and habit constrains. To test against the signal of
232 family membership, we performed a series of manual stepwise distance-based redundancy
233 analyses on the five families with the largest number of invasive species in our database to
234 identify variables that maximize difference among families (akin to our methodology to identify
235 traits that separated invasives from non-invasives). If the characteristics of invasiveness are
236 robust, we expected invasive species to cluster with each other instead of with their respective
237 families. Using the traits selected (Table 2), we performed a PCoA to see if species cluster
238 according to their family membership or invasive status in an unconstrained ordination. To
239 further confirm the results of these analyses, we repeated the contingency analyses described in
240 "*Identifying key invasive traits,*" this time including Family as a variable.
241 To test against growth form, we performed a PCoA using the same species list and 21
242 traits from the PCoA performed to determine the major axes of variance in the data described in
243 "*Identifying key invasive traits.*" In this analyses however, we visually distinguished species of
244 different growth forms. If the signal of growth form were stronger than that of invasive status,
245 we expected to see clustering by growth form.

Results

Invasive traits

The results of the PCoA on the refined dataset (i.e. the dataset containing 21 traits for 51 invasive and 109 non-invasive woody species) showed a diagonal separation in ordination space between invasive and non-invasive woody species (Fig. 1). This separation occurred along principal coordinates 1 and 2 (i.e., PCo1 and PCo2), the axes along which the highest level of variation in the data occurs. Although the percentages of variation along PCo1 and PCo2 appear to be low (7.1% and 5.8%, respectively), these percentages are noteworthy given the high dimensionality of the data (i.e. high number of variables). Therefore, these results indicate that invasive and non-invasive exotics do differ in traits.

Because the separation between invasive and non-invasive species did not occur along a single principal coordinate axis (i.e. separation occurred across PCo1 and PCo2), we performed contingency analyses on all 21 variables, instead of performing correlations between quantitative variables and principal coordinate scores. Six out of the 21 selected traits significantly separated invasive from non-invasive species (marked with a cross symbol in Table 1). These traits, which included vegetative regeneration, growth form (i.e. trees, shrubs or lianas), and seed dispersal vectors (i.e. water, mammal, bird or insect), indicate important differences between invasive and non-invasive woody species (Table 3). Vegetative regeneration (i.e. plants that demonstrate layering, suckering, root/stump resprouting, runners, or rhizomes as methods of natural spread or persistence) was the strongest contributor to PCo1 (χ^2=33.2), followed by bird-borne seed dispersal (χ^2=13.9).

A closer look into the values of these traits in the data showed a clear difference between invasive and non-invasive woody species in their ability to reproduce vegetatively and in seed

dispersal vectors (Table 3). Invasive woody species appear to be more capable of vegetative regeneration than non-invasive species (78% of invasives display vegetative regeneration vs 28% of non-invasives). Invasives were also more likely to be bird (63% of invasives vs 30% of non-invasives), water (41% of invasives vs 16% of non-invasives) and mammal (29% of invasives vs 9% of non-invasives) dispersed than non-invasive species. On the other hand, seeds of non-invasive woody species are more likely to be insect-dispersed than those of invasive woody species (29% of non-invasive species vs 10% of invasives). Invasive and non-invasive plants also differed in growth forms: invasives are much more likely to be lianas (16% of invasives are lianas, while only 3% of non-invasives display this growth form), conversely non-invasives are more likely to be trees (32% of non-invasives vs 22% of invasives).

Predictive power of key invasive traits

The boosted regression tree model was very effective in predicting the invasive status of exotic species, with an average accuracy of 0.86 (i.e. species were correctly identified as invasive or non-invasive 86% of the time on average across 100 models). The models displayed an average specificity of 0.86 (i.e. true negative rate—non-invasives were correctly identified 86% of the time on average across 100 models) and an average sensitivity of 0.83 (i.e. true positive rate—invasives were correctly identified 83% of the time on average across 100 models) (Table 4). The reported relative contribution of each variable to the model, based on the number of times a variable is used to split the data and how much these splits improve the model (Elith et al. 2008), further confirmed our preliminary findings regarding the importance of each variable. Natural vegetative regeneration was the most important trait in separating invasives from non-invasives, with a relative contribution of 39.3% to our models on average. The second most important trait

was dispersal of seeds by water accounting for 20.1% of variable contributions, followed by bird seed dispersal (15.9%), primary growth form (14.6%), mammal dispersal (8.8%), and insect dispersal (in the opposite direction, 1.3%). For more details on the results of these boosted classification trees, see Appendix S4, which contains a set of partial dependency plots for a single boosted classification tree model (i.e. one iteration of the 100 models that were averaged).

The boosted classification trees performed with non-invasive species that have been in the United States for more than a century complemented the results from the set of boosted classification trees performed with all non-invasive species. Natural vegetative regeneration was the most important determinant of invasiveness, accounting for 32.6% of the variation, followed by water seed dispersal (25.6%), bird seed dispersal (20.3%), primary growth form (10.3%), mammal seed dispersal (8.5%), and lastly insect seed dispersal (in the opposite direction, 2.7%). This model was also efficient in discriminating between invasive and non-invasive exotic plants with an average accuracy of 84% (mean: 0.84, st. dev.: 0.06), 82% sensitivity (mean: 0.82, st. dev.: 0.09), and 85% specificity (mean: 0.85, st. dev.: 0.07).

The region-specific boosted classification tree analyses showed slight differences in the predictive power of the key traits identified (Table 4). The US-level model displayed the highest accuracy (86%), followed closely by the Northwest, Northeast, and California (all 84% accurate). The model was least accurate in Florida (79%). The models also varied slightly in sensitivity and specificity. The regional models were better at correctly identifying invasive species than the full model (i.e. higher sensitivity), but were less effective at correctly identifying non-invasives (i.e. lower specificity). There were also some differences in the relative contribution of each trait among the regional models and the full model. Natural asexual regeneration was the most important discriminant for all models, despite varying in importance

among regions (e.g. it was least important in Florida when compared to other regions – 34%, and most important in the Midwest – 53%). Water dispersal was found to be the second most important trait, except in the Midwest where it was not very important (<1%). Insect seed dispersal, the least important trait in most models including the full model, was much more important in Florida than in other regions (8.7% versus 1.3% in the full model). The remaining traits displayed very slight variations among the regional and full models.

Phylogenetic relatedness

The results of the PCoA testing the signal of family membership indicated that most invasive species tend to cluster in the lower left quadrant of the biplot, regardless of family membership (Fig. 2a). A number of non-invasive Rosaceae species also clustered in this quadrant. Invasive Fabaceae species, the exception to this pattern, cluster with non-invasive Fabaceae species on the right side of PCo1, separating from the other four families. Fabaceae species (legumes) are a large, distinct family of nitrogen-fixing trees, shrubs and herbaceous species, many of which have been found to be invasive in different parts of the world. Being a monophyletic taxonomic group, legumes display a high level of interrelation among species; therefore, they share many characteristics that separate these species from other families in trait space. The contingency analyses including Family as a variable, showed that Family was not a significant factor, further evidencing that family membership is not a confounding variable in this study. The results for the primary growth PCoA also showed no major clustering among species of the same growth form, indicating that the invasive status signal is stronger than the growth form signal (Fig. 2b).

Discussion

Through a combination of multivariate and machine-learning methods, we were able to identify six key traits that can consistently predict invasiveness of exotic woody species across the United States. The ability to reproduce vegetatively in the wild and long-distance dispersal (via water, birds, mammals) were traits consistently associated with invasiveness in exotic woody species. Invasive and non-invasive woody species also differed in primary growth form, with invasive species displaying a higher proportion of lianas and a lower proportion of trees than non-invasive species. The boosted classification tree models created using these traits effectively predicted invasive status of exotic woody species at the regional and U.S. level, demonstrating the ability of these key traits to discriminate invasive from non-invasive species in various environmental contexts. The results of the analyses performed to account for minimum residence time and phylogenetic relatedness further evidenced the robustness of the predictive power of these key traits.

Through an ordination using 21 influential traits, we found a separation between invasive and non-invasive species along the primary axes of variation (i.e. PCo1 and PCo2). Although the variation captured by these axes was relatively small, these results are particularly meaningful given the high number of traits included in the ordination, considerably evidencing the separation of traits between invasive and non-invasive exotic plant species. In fact, this separation of traits between invasive and non-invasive species has been observed in previous research. For example, a comparison of invasive and non-invasive pine species showed a separation between the two pine groups due to mean seed mass, minimum juvenile period, and mean interval between large seed crops, signaling the existence of an r-K selection continuum, on which invasive pines fall on the r-selected end (Rejmánek and Richardson 1996). In fact, many studies indicate that these

361 invasive and non-invasive exotic plant species fall at opposite sides of the acquisitive-
362 conservative continuum, with invasive exotics displaying traits generally linked to resource
363 acquisition strategies, such as high relative growth rate, specific leaf area, maximum height, and
364 shoot biomass allocation (Grotkopp et al. 2002, van Kleunen et al. 2010, Tecco et al. 2013,
365 Gallagher et al. 2015, Erskine-Ogden et al. 2016). Because these studies focused mostly on traits
366 related to resource capture/allocation and fitness in disturbed habitats, the findings of our study
367 make an important advancement by identifying other determining traits of invasiveness in a wide
368 range of habitats.

369 The strongest determinant of invasiveness at the regional and U.S.-level in our study was
370 vegetative regeneration. Species classified under "regenerates vegetatively" include plants that
371 demonstrate layering, suckering, resprouting (from root fragments or root crown/stump), runners,
372 or rhizomes as methods of natural spread or persistence—albeit not exclusively (i.e. they may
373 also reproduce sexually). An exceedingly larger proportion of invasives displayed vegetative
374 regeneration as defined above (82% of invasives versus 28% of non-invasives). Vegetative
375 regeneration has been identified as a major driver of invasiveness and a major hurdle to control
376 and eradicate not only for woody plants (Reichard and Hamilton 1997, Rolim et al. 2015), but
377 also for herbaceous plants (Burns 2006, Marco et al. 2010, Rolim et al. 2015, Klinerová et al.
378 2018). In fact, prolific resprouting is one of the most challenging traits for invasive control, as it
379 is a major driver of reinvasion and persistence, prompting a body of research dedicated to the
380 management of resprouting through specialized chemical treatments (Witkowski and Garner
381 2008, Coffman et al. 2010, Constán-Nava et al. 2010, Enloe et al. 2015, Espeland et al. 2017).
382 On the other hand, the importance of uniparental reproduction for the establishment of exotic
383 species outside their range was first recognized by Baker (1955). This association is expected

384 since vegetative regeneration provides introduced propagules with an advantage to increase their
385 abundance rapidly and occupy vacant niches when suitable pollinators or mates are not available
386 (Baker 1955, Lloret et al. 2005, Van Kleunen et al. 2015).

387 The results of our study also highlight the importance of long-distance modes of dispersal
388 for invasiveness. Interestingly, while Reichard and Hamilton (1996) correctly concluded that
389 vegetative reproduction is an important attribute of invasive woody species, they did not
390 recognize vertebrate dispersal as an important trait. However, the importance of vertebrate
391 dispersal has been stressed by many other researchers (e.g. Binggeli 1996, Rejmánek and
392 Richardson 1996, Widrlechner et al. 2004). Efficient dispersal of propagules is essential to
393 advance from the naturalization/establishment stage to the invasion/spread stage (Gibson et al.
394 2011, Richardson and Rejmánek 2011, Pyšek et al. 2014). The general expectation is that long-
395 distance vectors of dispersal enhance invasiveness by facilitating spread farther from the site of
396 introduction (Richardson et al. 2000, Trakhtenbrot et al. 2005). In line with expectations, our
397 U.S. level results show that invasive woody species distinguish themselves from non-invasives in
398 their use of birds (65% invasives versus 28% of non-invasives), water (38% vs 12%), and
399 mammals (33% vs 13%) as agents of seed dispersal. This pattern was also observed at the
400 regional level, although there were some differences in the relative contribution of each vector
401 (i.e. the importance of birds versus water versus mammals), suggesting an interaction between
402 the importance of the long-distance dispersal vector and the environmental context. Birds, among
403 the most efficient long-distance dispersal agents (Vittoz and Engler 2007), have been identified
404 as the most prevalent mode of dispersal among invasive trees and shrubs at the global level (43%
405 of invasive trees and 61% of invasive shrubs globally) (Richardson and Rejmánek 2011).
406 Likewise, seeds can be transported long distances down streams and rivers, or along coastal

currents. The vast majority of invasive woody plants in our database that display water seed dispersal thrive along bodies of water (96% of water-dispersed invasives, 34% of all invasives). Examples include Russian olive (*Elaeagnus angustifolia*) along streams, or Australian pine (*Casuarina equisetifolia*) along coasts. On the other hand, non-invasive species predominantly displayed insect-borne dispersal of seeds (e.g. myrtle wattle─*Acacia myrtifolia*). This pattern appeared to be strongest in Florida. Insects, particularly ants, are attracted to seeds with elaiosomes, and will therefore transport seeds before consuming the elaiosome and dropping the rest of the viable seed. Insect-borne transport of seeds is short-distance, rarely exceeding 10 meters (Bossard 1991, Vittoz and Engler 2007).

Our results also indicated a difference in the predominance of growth forms between invasive and non-invasive species. Although most woody exotic species in our database (invasive or non-invasive) are shrubs (63% and 54% respectively), non-invasive woody plants at the U.S. level were more likely to be trees than invasives (42% versus 19%), while invasive woody plants were more likely to be lianas than non-invasives (17% versus 4%). The relatively high proportion of non-invasive trees may be a result of invasional lag (i.e. period of time in which recently introduced invasive species display slow rates of population growth or spread before spreading explosively and/or becoming an environmental nuisance). Trees generally have longer generation times than shrubs and lianas. Therefore, it may take longer for long-lived tree species to be identified as "invasive" given the criteria of Executive Order 13112 (i.e. spread rapidly and cause environmental or economic harm), despite their potential to become harmful in the future (Iannone et al. 2014). However, the results of our boosted classification tree models performed on non-invasives with minimum residence times greater than a century do not support

429 this argument, as primary growth form was also found to be an important determinant of
430 invasiveness in our analysis.

431 Nonetheless, there are a few considerations that must be made when interpreting these
432 results. Future climate change scenarios may favor some non-invasive species and make them
433 invasive, while adversely affecting current invasives making them non-invasive. However, the
434 invasive traits identified in this study are likely to allow current invasives to adapt to new
435 environments or expand their ranges to newly suitable habitats (Hellman et al. 2008).
436 Furthermore, invasive status, as defined here, is a human-made designation based on Executive
437 Order 13112. Governmental agencies in the U.S. use this definition to prioritize their efforts
438 towards harmful invasive exotics. However, the definition and classification of invasive species
439 is a highly debated topic. In Appendix S5 (Figure S1), we present the ordination from Figure 1,
440 but we visually marked species that have been classified as invasive by Rejmánek and
441 Richardson (2013) despite being considered "non-invasive" by U.S. governmental agencies.
442 Rejmánek and Richardson's definition of "invasive" woody taxa is strictly ecological and does
443 not include any judgments about their impacts.

444 The juxtaposition of these classification systems in Appendix S5 provides some
445 interesting insights. There are two visually distinctive groups of species labeled invasive by
446 Rejmánek and Richardson (2013) — those that overlapped with species labelled as "invasive" by
447 both classification systems, and those that overlapped with species labelled as "non-invasive" by
448 the governmental definition (i.e. the definition used in this study). The first group of species is of
449 particular interest, because it highlights species that could be invasive by the definition of
450 Executive Order 13112, but could potentially have been misclassified by governmental agencies.
451 These species may be relatively recent introductions that are currently in invasion lag phase, and

452 therefore are not recognized as harmful at the moment. Since these species display similar traits
453 to current invasives (as per the governmental definition), it is likely that they will become
454 invasive in the future.

455 This pattern underscores the importance of using traits-based knowledge, such as that
456 produced by this study, to inform classification systems of invasive status, instead of exclusively
457 using the criteria posed by Executive Order 13112. The fact that our model was able to
458 effectively discriminate between invasives and non-invasives regardless of minimum residence
459 time, as evidenced by the results of our boosted classification tree models using only non-
460 invasive species with a minimum residence time of one century, further demonstrates the utility
461 of our findings to inform classification systems and preventative measures.
462
463 **Future research**
464 Although our study provided new insigths into invasive traits of exotic woody plants, more
465 broad-scale studies are needed to understand woody plant invasiveness in other geographical
466 contexts. Woody invaders that have been successful in other parts of the world may display
467 different traits than the ones identified in this study. For instance, exotic conifers that have
468 invaded various countries in the Southern Hemisphere (e.g. New Zealand, South Africa, and
469 Argentina) have winged seeds, and benefit from long-distance wind dispersal once reproductive
470 individuals are established on ridgetops or hilltops (Simberloff et al. 2010). As mentioned above,
471 invasive status is a human-made designation with many inconsistent interpretations (Colautti and
472 MacIsaac 204, Catford et al. 2016). Here, we provide analyses based on the definition used by
473 U.S. agencies. However, perhaps it would be useful to consider alternative approaches to
474 defining invasiveness. For instance, Catford et al. (2016) suggested determining which functional

traits are associated to different population characteristics by which invasives are identified, such as spread rate, local abundance and environmental range. Although using this approach may allow for broader generalizations, availability of such data is still limited.

Future studies should also consider that invasive species vary in the level of impact they have on recipient ecosystems. Therefore, the traits associated to highly destructive species (i.e. "transformers" or "ecosystem engineers"; Fei et al. 2014) may differ from those of less harmful species, representing a useful distinction for prioritization of management. In this study, we attempted to test the robustness of the patterns observed against the signals of phylogenetic relatedness and habit, using family membership and primary growth form as proxies. Future studies would benefit from incorporating phylogeny into their analyses at a higher resolution.

Conclusion

Through a comparative analysis of 63 invasive and 794 non-invasive exotic woody species on 45 traits, we developed a statistical model able to predict with considerable accuracy the invasiveness of exotic woody plant species found across the United States. Our results suggest that the strongest determinants of invasiveness are vegetative reproduction and long distance dispersal vectors, such as birds, bodies of water and mammals. Invasive and non-invasive woody plants also differed in the predominance of certain growth forms, with lianas being much more predominant in the invasive pool than in the non-invasive. The findings from this study present an innovative contribution to the field, not only reaffirming existing notions of invasive traits at a macroscale level, but also elucidating the importance of less explored traits and their utility in predicting invasiveness. Furthermore, our findings can greatly improve existing screening tools and current classifications of invasive status.

Acknowledgements

We would like to thank Christian Houser and Kimberly Ordonez for assisting in the completion of the traits database used in this study. This research was partially supported by the NSF Macrosystems Biology Program (Grant No.1241932 and 1638702) and USDA Mcintire-Stennis program to S. Fei.

References

Bacher, S., Blackburn, T. M., Essl, F., Genovesi, P., Heikkilä, J., Jeschke, J. M., ... & Martinou, A. F. 2018. Socio-economic impact classification of alien taxa (SEICAT). Methods in Ecology and Evolution **9**:159-168.

Baker, H. G. 1955. Self-compatibility and establishment after 'long-distance' dispersal. Evolution **9**:347-349.

Bellard, C., P. Cassey, and T. M. Blackburn. 2016. Alien species as a driver of recent extinctions. Biology Letters **12**:20150623.

Binggeli, P. 1996. A taxonomic, biogeographical and ecological overview of invasive woody plants. Journal of Vegetation Science **7**:121-124.

Blackburn, T. M., Essl, F., Evans, T., Hulme, P. E., Jeschke, J. M., Kühn, I., ... & Pergl, J. 2014. A unified classification of alien species based on the magnitude of their environmental impacts. PLoS biology **12**:e1001850.

Bossard, C. C. 1991. The role of habitat disturbance, seed predation and ant dispersal on establishment of the exotic shrub *Cytisus scoparius* in California. American Midland Naturalist **126**:1-13.

Burns, J. H. 2006. Relatedness and environment affect traits associated with invasive and noninvasive introduced Commelinaceae. Ecological Applications **16**:1367-1376.

Coffman, G. C., R. F. Ambrose, and P. W. Rundel. 2010. Wildfire promotes dominance of invasive giant reed (*Arundo donax*) in riparian ecosystems. Biological Invasions **12**:2723-2734.

Conser, C., Seebacher, L., Fujino, D. W., Reichard, S., & DiTomaso, J. M. 2015. The development of a plant risk evaluation (PRE) tool for assessing the invasive potential of ornamental plants. PloS one, **10**:e0121053.

Constán-Nava, S., A. Bonet, E. Pastor, and M. J. Lledó. 2010. Long-term control of the invasive tree *Ailanthus altissima*: Insights from Mediterranean protected forests. Forest Ecology and Management **260**:1058-1064.

Crooks, J. A. 2005. Lag times and exotic species: the ecology and management of biological invasions in slow-motion. Ecoscience **12**:316-329.

Elith, J., J. R. Leathwick, and T. Hastie. 2008. A working guide to boosted regression trees. Journal of Animal Ecology **77**:802-813.

Enloe, S. F., N. J. Loewenstein, D. Streett, and D. K. Lauer. 2015. Herbicide treatment and application method influence root sprouting in Chinese tallowtree (*Triadica sebifera*). Invasive Plant Science and Management **8**:160-168.

Erskine-Ogden, J., E. Grotkopp, and M. Rejmánek. 2016. Mediterranean, invasive, woody species grow larger than their less-invasive counterparts under potential global environmental change. American Journal of Botany **103**:613-624.

Espeland, E. K., J. M. Muscha, J. Scianna, R. Kilian, N. M. West, and M. K. Petersen. 2017. Secondary invasion and reinvasion after Russian-olive removal and revegetation. Invasive Plant Science and Management **10**:340-349.

Fei, S., Q. Guo; G. C. Nunez-Mir, and B. V. Iannone. 2019. Functional traits of non-native woody plants in the United States. DOI: 10.4231/1YW7-DE10.

Fei, S., J. Phillips, and M. Shouse. 2014. Biogeomorphic impacts of invasive species. Annual reviews of Ecology, Evolution, and Systematics **45**:69-87

548 Freckleton, R. P. 2000. Phylogenetic tests of ecological and evolutionary hypotheses: checking
549 for phylogenetic independence. Functional Ecology **14**:129-134.
550 Gallagher, R. V., R. P. Randall, and M. R. Leishman. 2015. Trait differences between
551 naturalized and invasive plant species independent of residence time and phylogeny.
552 Conservation Biology **29**:360-369.
553 Gibson, M. R., D. M. Richardson, E. Marchante, H. Marchante, J. G. Rodger, G. N. Stone, M.
554 Byrne, A. Fuentes-Ramírez, N. George, and C. Harris. 2011. Reproductive biology of
555 Australian acacias: important mediator of invasiveness? Diversity and Distributions
556 **17**:911-933.
557 Gordon, D. R., D. A. Onderdonk, A. M. Fox, R. K. Stocker, and C. Gantz. 2008. Predicting
558 invasive plants in Florida using the Australian weed risk assessment. Invasive Plant
559 Science and Management **1**:178-195.
560 Gower, J. C. 1971. A general coefficient of similarity and some of its properties. Biometrics
561 **27**:857-871.
562 Grotkopp, E., M. Rejmánek, and T. L. Rost. 2002. Toward a causal explanation of plant
563 invasiveness: seedling growth and life-history strategies of 29 pine (Pinus) species. The
564 American Naturalist **159**:396-419.
565 He, S., L. Yin, J. Wen, and Y. Liang. 2018. A test of the Australian Weed Risk Assessment
566 system in China. Biological Invasions **20**:2061–2076.
567 Heger, T., S. Haider, W.C. Saul, and J. M. Jeschke. 2015. Species from different taxonomic
568 groups show similar invasion traits. Immediate Science Ecology **3**:1-13.
569 Hellmann, J. J., J. E. Byers, B. G. Bierwagen, and J. S. Dukes. 2008. Five potential consequences
570 of climate change for invasive species. Conservation Biology **22**:534-543.

571 Hijmans, R. J., S. Phillips, J. Leathwick, J. Elith, and M. R. J. Hijmans. 2017. Package
572 'dismo'. Circles **9**:1-68.
573 Hutchinson, T. F., and J. L. Vankat. 1997. Invasibility and Effects of Amur Honeysuckle in
574 Southwestern Ohio Forests. Conservation Biology **11**:1117-1124.
575 Iannone, B. V., M. L. Zellner, and D. H. Wise. 2014. Modeling the impacts of life-history traits,
576 canopy gaps, and establishment location on woodland shrub invasions. Ecological
577 Applications **24**:467-483.
578 Klinerová, T., K. Tasevová, and P. Dostál. 2018. Large generative and vegetative reproduction
579 independently increases global success of perennial plants from Central Europe. Journal
580 of Biogeography **45**:1550–1559
581 Kolar, C. S., and D. M. Lodge. 2001. Progress in invasion biology: predicting invaders. Trends
582 in Ecology & Evolution **16**:199-204.
583 Koop, A. L., L. Fowler, L. P. Newton, and B. P. Caton. 2012. Development and validation of a
584 weed screening tool for the United States. Biological Invasions **14**:273-294.
585 Kumschick, S., and D. M. Richardson. 2013. Species-based risk assessments for biological
586 invasions: advances and challenges. Diversity and Distributions **19**:1095-1105.
587 Laliberté, E., P. Legendre, B. Shipley, and M. E. Laliberté. 2014. Package 'FD'. Measuring
588 functional diversity from multiple traits, and other tools for functional ecology.
589 Legendre, P., and M. J. Anderson. 1999. Distance-based redundancy analysis: Testing
590 multispecies responses in multifactorial ecological experiments. Ecological Monographs
591 **69**:1-24.

592 Lloret, F., F. Médail, G. Brundu, I. Camarda, E. V. A. Moragues, J. Rita, P. Lambdon, and P. E.
593 Hulme. 2005. Species attributes and invasion success by alien plants on Mediterranean
594 islands. Journal of Ecology **93**:512-520.
595 Marco, A., S. Lavergne, T. Dutoit, and V. Bertaudiere-Montes. 2010. From the backyard to the
596 backcountry: how ecological and biological traits explain the escape of garden plants into
597 Mediterranean old fields. Biological Invasions **12**:761-779.
598 Mascaro, J., and S. A. Schnitzer. 2007. *Rhamnus cathartica* L. (Common Buckthorn) as an
599 Ecosystem Dominant in Southern Wisconsin Forests. Northeastern Naturalist **14**:387-
600 402.
601 Miller, J. T., C. Hui, A. H. Thornhill, J. J. Le Roux, and D. M. Richardson. 2017. Is invasion
602 success of Australian trees mediated by their native biogeography, phylogenetic history,
603 or both? AoB PLANTS **9**: plw080.
604 Nunez-Mir, G., A. Liebhold, Q. Guo, E. Brockerhoff, I. Jo, K. Ordonez, and S. Fei. 2017. Biotic
605 resistance to exotic invasions: its role in forest ecosystems, confounding artifacts, and
606 future directions. Biol Invasions. **19**:3287–3299.
607 Oksanen, J., F. G. Blanchet, R. Kindt, P. Legendre, P. R. Minchin, R. B. O'hara, G.L. Simpson,
608 P. Solymos, M.H.H. Stevens, H. Wagner, and M. J. Oksanen. 2013. Package
609 'vegan'. Community ecology package, version 2.
610 Panetta, F. D. 2015. Weed eradication feasibility: lessons of the 21st century. Weed Research
611 **55**:226-238.
612 Pavoine, S., J. Vallet, A. B. Dufour, S. Gachet, and H. Daniel. 2009. On the challenge of treating
613 various types of variables: application for improving the measurement of functional
614 diversity. Oikos **118**:391-402.

615 Pemberton, R. W., and H. Liu. 2009. Marketing time predicts naturalization of horticultural
616 plants. Ecology **90**:69-80.
617 Pheloung, P. C., P. A. Williams, and S. R. Halloy. 1999. A weed risk assessment model for use
618 as a biosecurity tool evaluating plant introductions. Journal of Environmental
619 Management **57**:239-251.
620 Pyšek, P., Jarošík, V., Hulme, P. E., Pergl, J., Hejda, M., Schaffner, U. and M. Vilà. 2012. A
621 global assessment of invasive plant impacts on resident species, communities and
622 ecosystems: the interaction of impact measures, invading species' traits and environment.
623 Global Change Biology **18**:1725-1737.
624 Pyšek, P., V. Jarošík, J. Pergl, L. Moravcová, M. Chytrý, and I. Kühn. 2014. Temperate trees and
625 shrubs as global invaders: the relationship between invasiveness and native distribution
626 depends on biological traits. Biological Invasions **16**:577-589.
627 Razanajatovo, M., N. Maurel, W. Dawson, F. Essl, H. Kreft, J. Pergl, P. Pyšek, P. Weigelt, M.
628 Winter, and M. Van Kleunen. 2016. Plants capable of selfing are more likely to become
629 naturalized. Nature communications **7**:13313.
630 Reichard, S. H., and C. W. Hamilton. 1997. Predicting invasions of woody plants introduced into
631 North America. Conservation Biology **11**:193-203.
632 Rejmánek, M. 1996. A theory of seed plant invasiveness: the first sketch. Biological
633 conservation **78**:171-181.
634 Rejmánek, M. 2013. Invasiveness. In: D. Simberloff and M. Rejmánek (eds.) Encyclopedia of
635 Biological Invasions. University of California Press, Berkeley, pp. 379-385.

636 Rejmánek, M. 2015. Biological invasions in forests and forest plantations. Pages 452-469 *in* K.
637 S.-H. Peh, R. T. Corlett, and Y. Bergeron, editors. Routledge Handbook of Forest
638 Ecology, Routledge, NY.
639 Rejmánek, M., and M. J. Pitcairn. 2002. When is eradication of exotic pest plants a realistic
640 goal? In: *Turning the Tide: The Eradication of Invasive Species* (eds C.R. Veitch & M.N.
641 Clout), pp. 249–253. IUCN, Gland, Switzerland and Cambridge, UK.
642 Rejmánek, M., and D. M. Richardson. 1996. What Attributes Make Some Plant Species More
643 Invasive? Ecology **77**:1655-1661.
644 Rejmánek, M., and D. M. Richardson. 2013. Trees and shrubs as invasive alien species–2013
645 update of the global database. Diversity and Distributions **19**:1093-1094.
646 Rejmánek, M., Richardson, D. M., and P. Pyšek. 2013. Plant invasions and invasibility of plant
647 communities. In: E. van der Maarel and J. Franklin (eds.) *Vegetation Ecology,* Second
648 Edition, Wiley-Bleckwell, Chichester, 387-424.
649 Richardson, D. M., P. Pyšek, M. Rejmánek, M. G. Barbour, F. D. Panetta, and C. J. West. 2000.
650 Naturalization and invasion of alien plants: concepts and definitions. Diversity and
651 Distributions **6**:93-107.
652 Richardson, D. M., and M. Rejmánek. 2011. Trees and shrubs as invasive alien species–a global
653 review. Diversity and Distributions **17**:788-809.
654 Ries, P., M. E. Dix, M. Lelmini, and D. Thomas. 2004. National strategy and implementation
655 plan for invasive species management. Washington, DC: US Department of Agriculture,
656 Forest Service. FS-805.

Rolim, R. G., P. M. A. de Ferreira, A. A. Schneider, and G. E. Overbeck. 2015. How much do we know about distribution and ecology of naturalized and invasive alien plant species? A case study from subtropical southern Brazil. Biological Invasions **17**:1497-1518.

Silander, J. A., and D. M. Klepeis. 1999. The Invasion Ecology of Japanese Barberry (*Berberis thunbergii*) in the New England Landscape. Biological Invasions **1**:189-201.

Simberloff, D., Nunez, M.A., Ledgard, N.J., Pauchard, A., Richardson, D.M., Sarasola, M., Van Wilgen, B.W., Zalba, S.M., Zenni, R.D., Bustamante, R. and E. Peña. 2010. Spread and impact of introduced conifers in South America: lessons from other southern hemisphere regions. Austral Ecology **35**:489-504.

Stebbins, G. L. 1957. Self fertilization and population variability in the higher plants. The American Naturalist **91**:337-354.

Tecco, P. A., C. Urcelay, S. Diaz, M. Cabido, and N. Pérez-Harguindeguy. 2013. Contrasting functional trait syndromes underlay woody alien success in the same ecosystem. Austral Ecology **38**:443-451.

Thompson, K., and M. A. Davis. 2011. Why research on traits of invasive plants tells us very little. Trends in Ecology & Evolution **26**:155-156.

Trakhtenbrot, A., R. Nathan, G. Perry, and D. M. Richardson. 2005. The importance of long-distance dispersal in biodiversity conservation. Diversity and Distributions **11**:173-181.

van Kleunen, M., W. Dawson, and N. Maurel. 2015. Characteristics of successful alien plants. Molecular Ecology **24**:1954-1968.

van Kleunen, M., E. Weber, and M. Fischer. 2010. A meta-analysis of trait differences between invasive and non-invasive plant species. Ecology Letters **13**:235-245.

679 Vitousek, P. M., H. A. Mooney, J. Lubchenco, and J. M. Melillo. 1997. Human domination of
680 Earth's ecosystems. Science **277**:494-499.
681 Vittoz, P., and R. Engler. 2007. Seed dispersal distances: a typology based on dispersal modes
682 and plant traits. Botanica Helvetica **117**:109-124.
683 Widrlechner, M. P., J. R. Thompson, J. K. Iles, and P. M. Dixon. 2004. Models for predicting the
684 risk of naturalization of nonnative woody plants in Iowa. Journal of Environmental
685 Horticulture **22**:23-31.
686 Williamson, M. 1996. Biological Invasions. Chapman & Hall, London.
687 Wilson, J. R. U., D. M. Richardson, M. Rouget, Ş. Procheş, M. A. Amis, L. Henderson, and W.
688 Thuiller. 2007. Residence time and potential range: crucial considerations in modelling
689 plant invasions. Diversity and Distributions **13**:11-22.
690 Witkowski, E. T. F., and R. D. Garner. 2008. Seed production, seed bank dynamics, resprouting
691 and long-term response to clearing of the alien invasive *Solanum mauritianum* in a
692 temperate to subtropical riparian ecosystem. South African Journal of Botany **74**:476-
693 484.
694 Woods, K. D. 1993. Effects of Invasion by *Lonicera tatarica* L. On Herbs and Tree Seedlings in
695 Four New England Forests. The American Midland Naturalist **130**:62-74.

Tables

Table 1: Functional traits and characteristics obtained for each invasive and non-invasive species in the traits database

Morphology	**Primary growth form*†**
	Ovate leaf shape*
	Leaf arrangement*
	Max. leaf width (cm)*
	Max. leaf area (cm^2)*
	Elliptic leaf shape*
	Oblong leaf shape*
	Lanceolate leaf shape*
	Number of growth forms*
	Max. leaf length (cm)
	Max. height (m)
	Leaf type
Pollination	Mammal pollinated*
	Animal (other) pollinated*
	Animal (nonspecific) pollinated*
	Bird pollinated
	Insect pollinated
	Self-pollinated
	Wind pollinated
Flowers	Flower color number*
	Flower color primary
	Flower type
	Flower description
Dispersal	**Water dispersed seed*†**

	Mammal dispersed seed*†
	Bird dispersed seed*†
	Insect dispersed seed*†
	Animal (other) dispersed seed*
	Self-dispersed seed*
	Wind dispersed seed
	Animal (nonspecific) dispersed seed
	Seed weight (g/seed)
	Fruit type
Regeneration	**Vegetative regeneration*†**
	Sexual regeneration*
Physiology and environmental tolerance	Length of life cycle*
	Min. elevation (m)*
	Max. elevation (m)*
	Photosynthetic pathway
	Chromosome number
	Min. pH
	Max. pH
	Max. hardiness zone
	Cotyledon number
Other	Means of introduction

699 * Traits that increased R^2 when added or decreased R^2 when removed in a series of distance-based
700 redundancy analyses executed in a manual stepwise manner to identify traits that separate invasives from
701 non-invasives.
702 † Traits that produced the highest amount of variation between invasives and non-invasives in an
703 unconstrained ordination.

Table 2: Important traits in differentiating species from different families as identified through a series of distance-based redundancy analyses executed in a manual stepwise manner

Number of growth forms
Min. elevation (m)
Max. elevation (m)
Length of life cycle
Natural asexual regeneration
Wind pollinated
Mammal pollinated
Bird pollinated
Insect pollinated
Wind dispersed seed
Water dispersed seed
Seed weight (g/seed)
Mammal dispersed seed
Bird dispersed seed
Insect dispersed seed
Fruit type
Leaf arrangement
Max. leaf width (cm)
Leaf type
Group
Max. height (m)
Max. leaf length (cm)

Table 3: Pearson's chi-squared and percentage of invasives and non-invasives of the most influential traits on PCo1. These traits produce the highest amount of variation between invasives and non-invasives in an unconstrained ordination. P-values for all traits fell under 0.05.

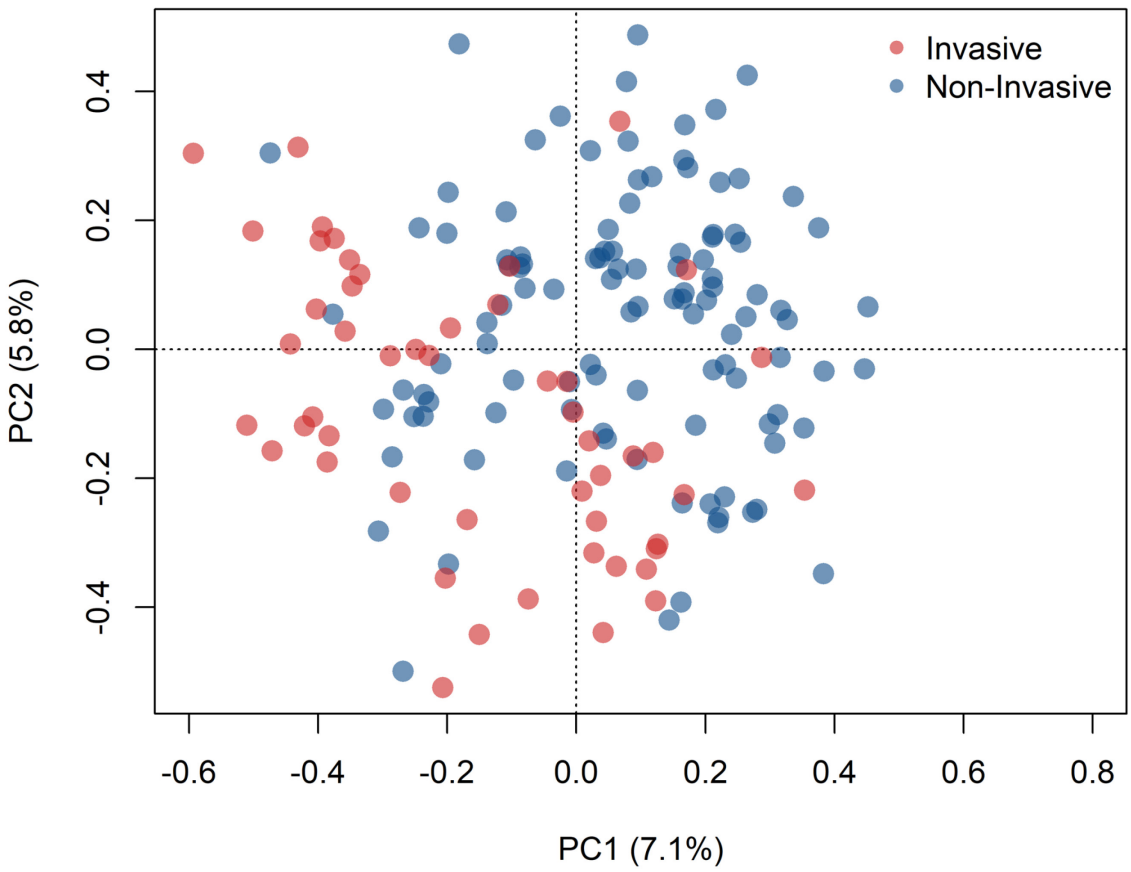
Trait	X^2	% of invasives	% of non-invasives
Vegetative regeneration	33.2	78	30
Bird dispersed seed	13.9	61	32
Water dispersed seed	10.2	43	16
Primary growth form	9.8	-	-
Mammal dispersed seed	9.3	30	10
Insect dispersed seed	6.4	11	31

719 **Table 4:** Mean and standard deviation of the evaluative measures of predictive performance for
720 100 boosted classification trees performed by region and on all 63 invasive species and 794 non-
721 invasive species (full model).

	Accuracy		Sensitivity		Specificity	
	Mean	St. Dev.	Mean	St. Dev.	Mean	St. Dev.
California	0.84	0.06	0.88	0.10	0.84	0.07
Florida	0.79	0.08	0.91	0.09	0.78	0.09
Midwest	0.80	0.10	0.90	0.12	0.78	0.12
Northeastern	0.84	0.06	0.85	0.10	0.84	0.09
Southeastern	0.80	0.06	0.92	0.08	0.76	0.08
Northwest	0.84	0.07	0.85	0.10	0.84	0.09
Full model	**0.86**	**0.05**	**0.83**	**0.09**	**0.86**	**0.06**

722

723



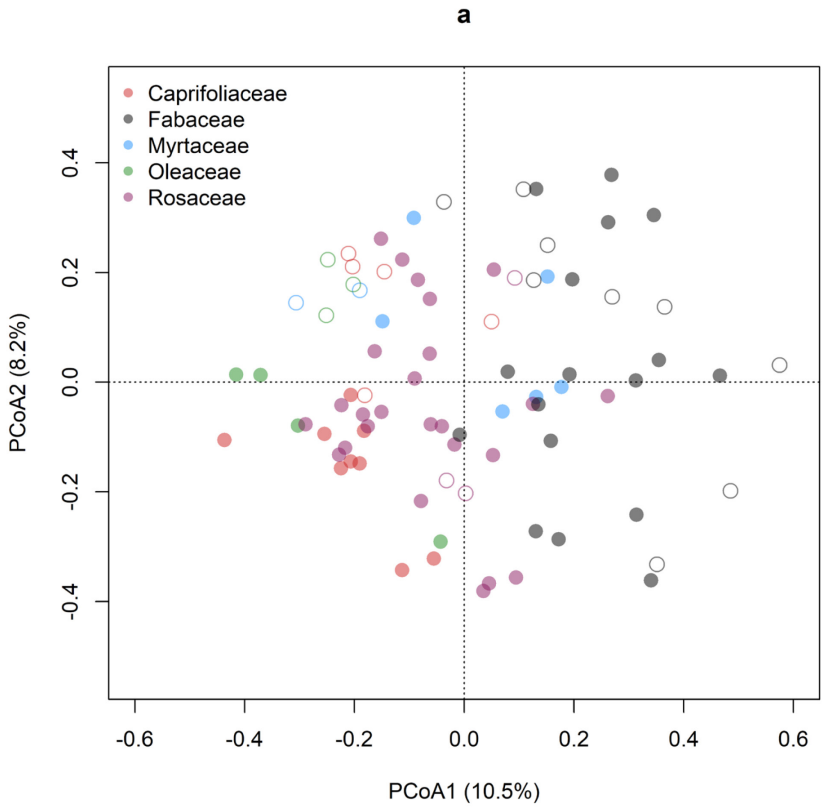
724

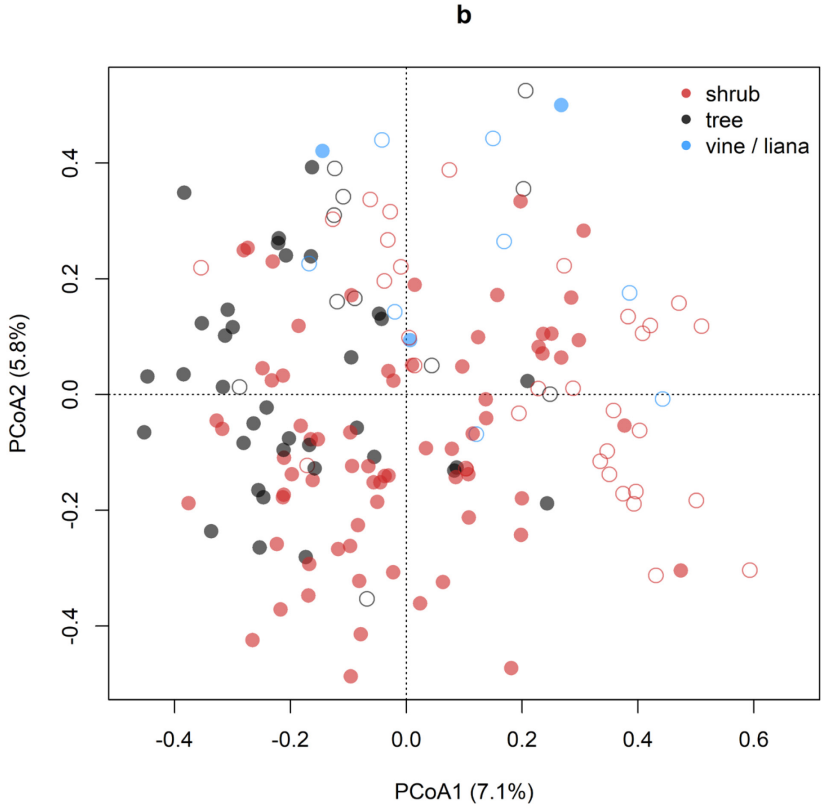
725 **Figure 1:** Biplot of principal coordinates 1 and 2. PCoA of invasive (n=51) and non-invasive

726 (n=109) woody species using the 21 traits selected through manual stepwise distance-based

727 redundancy analyses.

728


a
Caprifoliaceae
Fabaceae
Myrtaceae
Oleaceae
Rosaceae
PCoA2 (8.2%)
PCoA1 (10.5%)
-0.6
-0.4
-0.2
0.0
0.2
0.4
0.6


b
shrub
tree
vine / liana
PCoA2 (5.8%)
PCoA1 (7.1%)
-0.4
-0.2
0.0
0.2
0.4
0.6

729

Figure 2: (a) Biplot of principal coordinates 1 and 2. PCoA of invasive and non-invasive species of four families (Caprifoliaceae (N: 9/I: 5), Fabaceae (N: 16/I: 10), Rosaceae (N: 25/I: 3), Myrtaceae (N: 6/I: 2), and Oleaceae (N: 4/I: 3)). (b) Biplot of principal coordinates 1 and 2. PCoA of invasive (n=51) and non-invasive (n=109) exotic woody species using 21 functional traits. Colors distinguish species that display different primary growth forms. In both figures, filled symbols are non-invasive, empty symbols are invasive.