

Invasive plants have broader physiological niches

Steven I. Higgins^{a,1} and David M. Richardson^b

^aDepartment of Botany, University of Otago, Dunedin 9054, New Zealand; and ^bCentre for Invasion Biology and Department of Botany and Zoology, Stellenbosch University, Matieland 7602, South Africa

Edited by Peter M. Vitousek, Stanford University, Stanford, CA, and approved June 6, 2014 (received for review April 1, 2014)

Invasive species cost the global economy billions of dollars each year, but ecologists have struggled to predict the risk of an introduced species naturalizing and invading. Although carefully designed experiments are needed to fully elucidate what makes some species invasive, much can be learned from unintentional experiments involving the introduction of species beyond their native ranges. Here, we assess invasion risk by linking a physiologically based species distribution model with data on the invasive success of 749 Australian acacia and eucalypt tree species that have, over more than a century, been introduced around the world. The model correctly predicts 92% of occurrences observed outside of Australia from an independent dataset. We found that invasiveness is positively associated with the projection of physiological niche volume in geographic space, thereby illustrating that species tolerant of a broader range of environmental conditions are more likely to be invasive. Species achieve this broader tolerance in different ways, meaning that the traits that define invasive success are context-specific. Hence, our study reconciles studies that have failed to identify the traits that define invasive success with the urgent and pragmatic need to predict invasive success.

biological invasion | range size | physiology | ecological niche | tree invasions

Why some species perform better than others when introduced to novel regions is a question of immense theoretical and practical importance. Theoretical ecologists seek to understand the conditions that allow species to invade communities and coexist with other taxa, thereby shaping patterns of biodiversity (1, 2). Applied ecologists need to know which introduced species are most likely to establish, invade, and cause environmental damage (3). More generally, biological invasions are grand natural experiments that provide one of ecology's most profitable avenues for testing our ability to forecast the distribution of species and diversity (4).

Although theoretical ecologists have made impressive progress in understanding the mediators of coexistence using invasion tests (5, 6), applied ecologists bemoan the fact that predicting the next environmental pest seems as intractable now as decades ago, when the global consequences of invasions first became apparent (7). Although many studies have deciphered the attributes of successful invaders (8), others show that invasive organisms do not differ in consistent ways from native taxa and that, if laws do determine invasiveness and invasibility, they are, at best, highly context-specific (9–11). Such uncertainty is perhaps not surprising, because the study of invasions is, by definition, complicated by historical factors. For example, increased propagule pressure and residence time increase invasive success, together often overwhelming any inherent factors that enhance or limit invasiveness (12–14). Hence, invasion biology seems resigned to accepting that a universal definition of invasibility and invasiveness is unlikely and that case-specific solutions are all that can be achieved (9). As a consequence, we still lack a robust, reliable, and universal protocol for screening large numbers of species for invasiveness. However, existing screening schemes do not fully make use of the observation that a match between the target organism's physiology and the target

environment is the most consistent indicator of invasion success (9). Also, screening schemes have not taken full advantage of species distribution modeling (15, 16), a method that allows broad-scale environmental attributes to be matched to indicators of species' physiology (17). Previous applications of species distribution modeling to invasions have been restricted to correlative models with weak links to physiology and small sets of species (18–20).

Here, we test the suitability of species distribution models with clear links to physiology for screening plant species for invasiveness. We use a species distribution model that calibrates a physiological model of plant growth to distribution data (21). This process model is derived from a well-established physiological model (22) that captures the fundamental processes that influence a plant's physiological performance. The model considers carbon, nitrogen, and biomass pools in the roots and shoots of plants and how physiological processes of resource assimilation, allocation, growth, and respiration interact to determine these pools. Each physiological process in the model is constrained by environmental factors. For example, shoot growth is limited by temperature, whereas nitrogen uptake is colimited by temperature, soil moisture, and soil nitrogen levels. The parameters that describe these environmental constraints are estimated by inferring the parameter combinations that best explain the observed distribution of each target species (Fig. S1 shows an overview of the model structure, the environmental dependencies that it considers, and the parameter estimation procedure). The process-based structure of the model has the advantage that it serves to constrain the parameter estimates, because it defines how seasonality in environmental factors and colimitation by environmental factors influence plant performance (21).

We apply this method to two taxa, where large numbers of closely related species have been introduced by people outside of their native continent and the proportions that have been naturalized

Significance

Invasibility (the vulnerability of a receiving environment to invasion) and invasiveness (the capacity of an organism to invade) are fundamental metrics in theoretical ecology. However, attempts to apply these concepts to the question of whether an introduced species will invade have failed. Here, we use a model that characterizes the physiological niche of plants to successfully predict invasive success of 749 species. We find that, although invasive success is dependent on different factors in different ecological contexts, the modeled global potential range of species integrates these contingencies to produce a reliable and interpretable indicator of the likelihood that a species will reproduce and spread.

Author contributions: S.I.H. and D.M.R. designed research; S.I.H. performed research; and S.I.H. and D.M.R. wrote the paper.

The authors declare no conflict of interest.

This article is a PNAS Direct Submission.

¹To whom correspondence should be addressed. Email: steven.higgins@otago.ac.nz.

This article contains supporting information online at www.pnas.org/lookup/suppl/doi:10.1073/pnas.1406075111/-DCSupplemental.

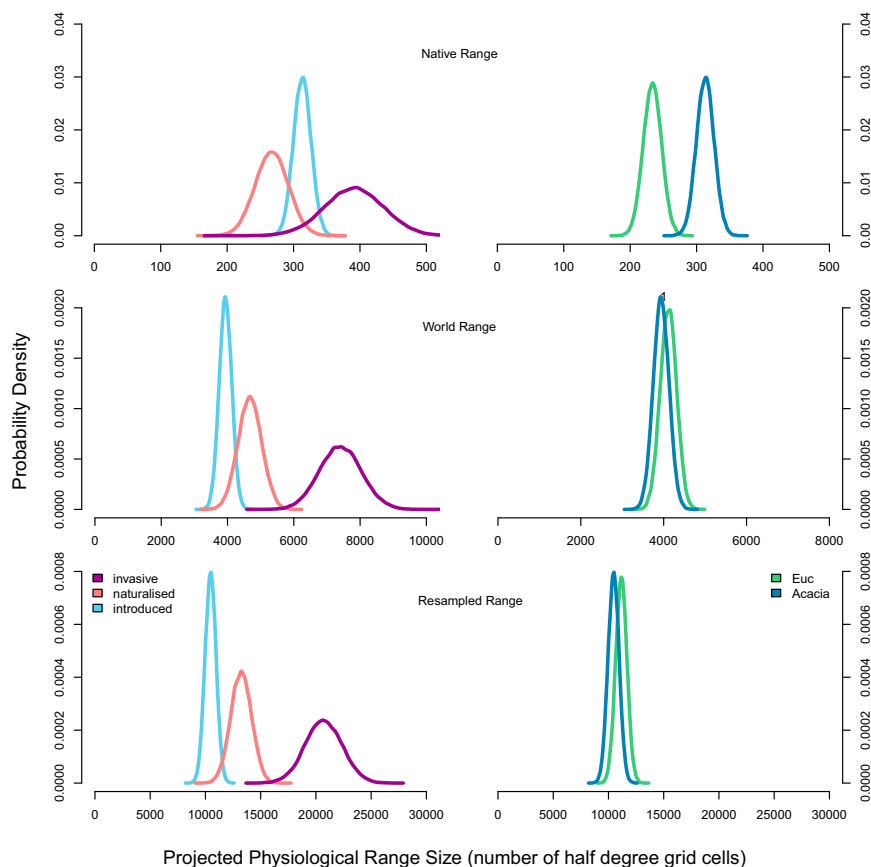


Fig. 2. Posterior densities of the potential range sizes of invasive, naturalized, and introduced acacia and eucalypt (Euc) species in their native Australia, the world, and a resampled world, where each of the world's environmental zones is equally common. The posterior densities are calculated from the regression coefficients of a Bayesian regression model.

result implies that invasive species do not have larger potential ranges because of preferences for more common environmental zones. Together, these findings suggest that it is not so much the native range size but the potential global range that is informative of invasive success. The current native geography may thus not provide the full suite of environments in which a species can persist. The global stage on which invasions play out, however, offers many more environment types and therefore, provides a less biased arena for inferring the range of environmental conditions to which a species is preadapted.

The estimated model parameters describe dependencies of physiological rates (growth, respiration, carbon, and nitrogen uptake) on environmental factors (temperature, soil moisture, solar radiation, and soil nitrogen). Each dependency is described by a response function, the parameters of which can be interpreted as traits that define the species' environmental tolerances. These traits, when interpreted by the model, reveal that invasive status is clearly associated with projected global range size (Fig. 2). Although valuable, this analysis does not identify if particular traits allow invasive species to occupy larger ranges.

We, therefore, now turn to the question of whether invasiveness is associated with distinct traits or trait syndromes. A multivariate analysis of the physiological niche parameters used to project these ranges revealed clusters of species. However, each cluster included all status groups, rendering this classification useless as a tool for screening invasive risk (Fig. S3). Because the parameters in this classification describe the species' environmental tolerances, the implication is that it is immaterial for invasion success whether species achieve large potential range sizes through, for example, a broad temperature or a broad

moisture tolerance. Indeed, plotting the average niche breadths of the nine niche dimensions used in the species distribution model (Fig. 3) revealed that invasive species, on average, did not have broader niche dimensions. On the contrary, for most niche dimensions, the average niche breadth decreased with invasive success; the exception here was that invasive success was associated with a capacity to take up soil nitrogen over a broader range of soil nitrogen conditions (Fig. 3).

Taken together, these results suggest that invasive success reveals itself only as a high-dimension interaction between plant physiology and environmental conditions. For example, a higher temperature tolerance for carbon uptake may result in a dramatic increase in the potential range size of a species that has a tolerance of dry soils but only a modest increase in the potential range size of a species that cannot tolerate dry soils. Furthermore, a unit increase in temperature tolerance may open up a smaller area of geographic range in a topographically diverse region but a larger geographic area in a flat region. Our analysis suggests that the most appropriate way to quantify the outcome of these interactive and context-specific factors is the global potential range size. This view is further supported by our finding that there was no clear relationship between invasive status group and whether a species tended to have positive or negative range associations with other species (Fig. 1, inner rings and Fig. S2, inner rings). The implication is that invasive status group does not represent a syndrome defined by a preference for a particular environment type.

Conclusion

Our analysis of invasive success clearly shows that invasive species have larger potential global ranges than naturalized species,

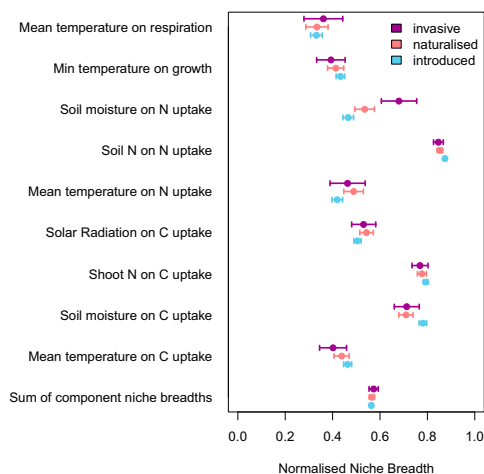


Fig. 3. Average normalized niche breadth for the sum of all nine niche axes in the species distribution model and each individual niche axis separated by invasive status. The points indicate the means, and the bars span the 95% credible intervals estimated using a Bayesian regression model.

which in turn, have larger potential ranges than species that were introduced but failed to naturalize. The fact that introduced, naturalized, and invasive species differed in potential range size supports the proposition (26) that species that invade, naturalize, and fail to naturalize represent three ecologically distinct groups. The linkage that we detected between potential range size and invasiveness is intuitive, because it suggests that species pre-adapted to a larger range of environmental conditions are more likely to be successful invaders. This explanation of our findings is consistent with the theory of fitness homeostasis (27), which proposes that species that maintain fitness over a broader range of environments are more likely to be invasive (12). This theory is supported by studies that show that invasive success is positively associated with the genotypic and phenotypic diversity of founder populations (14). Fitness homeostasis is, however, difficult to measure directly (12). We suggest that our method for estimating global potential range size may provide a readily inferable proxy for fitness homeostasis.

Our analysis provides a rapid and economical method for ranking the invasive risk of plant species. It failed to identify traits that are associated with invasive success, a finding that is supported by the many studies that have likewise failed to identify the traits that define invasive success, further supporting the conclusion that the traits that define invasive success are context-specific (9). The species distribution modeling approach that we use here should be applicable to other groups of invasive plants, including grasses and herbaceous plants. Although the mechanistic nature of the approach that we use is tailored to plants and therefore, cannot be applied to animals, analogous methods could be developed for animals. In this regard, dynamic energy balance theory models (28) could provide the mechanistic core—the role that the Thornley transport resistance (TTR) model adopted in this study—for estimating the physiological niches of animals from distribution data (29).

Our findings have implications for species distribution modeling, a method routinely criticized for making questionable assumptions (17, 30). We showed that, despite being an incomplete descriptor of the factors that define species' ranges, our species distribution models, nonetheless, predicted 92% of the combined occurrences derived from an independent dataset of 749 species included in this study. Hence, species distribution models, despite making strong assumptions, are capable of making useful predictions, supporting a view that species' ranges

are primarily defined by climate factors and species climate niches are often conservative (20, 31).

Materials and Methods

Species Distribution Data and Climate Data. We used 379 *Acacia* and 374 *Eucalypt* (members of the *Angophora*, *Corymbia*, and *Eucalyptus* genera) species known to be introduced outside of Australia (24). For the acacias, 47 and 22 of these species are reported as naturalized and invasive, respectively. Our usage of these terms follows definitions provided in ref. 23: introduced species have been transported by humans across a major geographical barrier, naturalized species can survive and reproduce in the introduced environment, and invasive species produce reproductive offspring distant from the sites of introduction. For the eucalypts, 72 species are reported to be naturalized, and 8 species are reported to be invasive. Our list differs slightly from the list in ref. 24 because of unavailable data for a minority of species (Dataset S1 shows the list of species used). For each species, we extracted occurrence data from the AVH (<http://avh.chah.org.au>) on December 3, 2013. For the GBIF records, we extracted data directly from www.gbif.org/ on December 2 and 3, 2013.

The species distribution model (21) requires data on soil nitrogen (32), soil water, solar radiation (monthly estimates from ref. 33), and mean, maximum, and minimum temperatures (monthly estimates from ref. 34). All variables (except soil nitrogen) are available at 1-km-grid resolution; for soil nitrogen, the resolution was 0.5°.

Species Distribution Modeling. We use the TTR model (22) as implemented in the work in ref. 21 to project the distributions of species (Fig. S1). The TTR, as developed in the work in ref. 22, is an ordinary differential equation model that considers how plant growth is influenced by carbon uptake, nitrogen uptake, and the allocation of carbon and nitrogen between roots and shoots. It explicitly separates the physiological processes of resource uptake from biomass growth. The implementation described in ref. 21 relates the uptake and growth processes to environmental forcing variables. Specifically, the model considers how carbon uptake might be limited by temperature, soil moisture, solar radiation, and shoot nitrogen; nitrogen uptake might be limited by temperature, soil moisture, and soil nitrogen; and growth and respiration might be influenced by temperature. The model runs on a monthly time step, which allows it to explicitly consider how seasonal fluctuations in the forcing variables interactively influence plant resource uptake and growth. In this study, we use exactly the same model version that was used in ref. 21 (ref. 21 has a full description of the model and its assumptions).

We used all available AVH records for the species distribution data. The AVH data include only presence observations, but the species distribution model requires both presence and absence observations. For the AVH dataset, Australia has been well-sampled, and the absence of presence data is indicative that the species does not occur. A variety of procedures exist for simulating absence points in such situations, and the resulting absence points are referred to as pseudoabsence points. Analyses with simulated data suggest that the method used to generate such pseudoabsences is one of the smallest sources of error in species distribution modeling (35). We follow a procedure shown to be less biased than alternative methods (35) to generate pseudoabsence data: we create a mask of 0.1° surrounding presence points and then randomly select 10,000 points not in this mask but in Australia (including Tasmania).

The model predicts the potential biomass of an individual plant as a function of the environmental forcing variables at a site. Following the work in ref. 21, we assume that p_i , the probability of a species occurring at site i , is described by the complementary log-log of the modeled plant biomass at site i and that the likelihood of observing the presence absence data (y_i) at site i is described by the Bernoulli distribution. To estimate the parameters, we used the differential evolution optimization algorithm (36) to find the set of parameters that maximizes this likelihood over all sites. The algorithm mutates and recombines a population of candidate solutions over generations. We ran the algorithm for each species using populations of 100 candidate solutions for 500 generations. This procedure was repeated five times. Each subsequent iteration through the species list used solutions from the previous iterations to initiate the algorithm. We used the same number of presence points as absence points for the model fitting (Dataset S1 provides a list of the number of presence data points used for each species). Each call to the optimization algorithm used a random subsample of the absence points.

Evaluation of the Species Distribution Models. We create a confusion matrix (a matrix of the number of true positives, true negatives, false positives, and false negatives) for each species distribution model estimated from the AVH distribution data. We examined whether the projected global distribution model matched the GBIF data for the species outside of their native Australia. Despite the awareness of invasive plants and the fact that our study species are conspicuous, there were few GBIF records for most study species. A one-sided binomial test was used to evaluate whether the model predicted more of the observed distributional records than would be expected by chance for each species, and a combined probability test of these binomial tests (z test) was used to summarize the individual tests. Trials were possible for 163 species. All data were used for the statistical testing, but only trials with $n > 15$ are plotted in Fig. 1 and Fig. S2.

To avoid projecting distributions into a statistical domain not found in the native continent, we formally define the domain within which global projections could be made. We used the environmental data used in the species distribution models to characterize clusters of environmental conditions (environmental zones) using discriminant analysis of principle components (DAPC) (37). We subsequently identified which of these environmental zones does not occur in Australia. These zones were considered to be beyond the domain of the training dataset, and such sites were masked in our global projections. This procedure, in effect, excluded the Boreal and Arctic Tundra regions.

Evaluation of the Physiological Niche of Invasive and Noninvasive Species. We calculated the niche volume in three ways: projecting species ranges for Australia, projecting species ranges for the world, and using a resampled dataset that assumes that the world's environmental zones are equally common. This third method allows us to test whether species have broader ranges because of a preference for more common environment types. To create a dataset where each environmental zone is equally common, we created a resampled dataset of the environmental data. The first step in this resampling procedure is to define the environmental zones. DAPC (37) was used to classify the environmental data into 100 environmental zones. The second step is to sample a finite number (1,000 in our case) of locations from each of 100 environmental zones, which produces an environmental dataset

where each environment zone is equally represented. We then project whether species would occur in each of the resulting $100 \times 1,000$ locations. This resampled dataset retains the covariance structure of the original environmental data.

A Bayesian linear regression model (38) was used to describe the effects of group (acacia or eucalypt) and invasive status (invasive, naturalized, or introduced) on potential native, global, and resampled global range sizes. We plot the posterior estimates of the mean estimated range sizes of these groups in Fig. 2. The significance of the regression coefficients was tested using a z test (38).

We calculated Yule's measure of association between species pairs and report the mode of each species association measure. Yule's association is defined as $(ad - bc)/((a + b)(c + d)(a + c)(b + d))^{0.5}$, where a is the number of plots with both species present, b is the number of plots with only species 1 present, c is the number of plots with only species 2 present, and d is the number of plots with both species absent.

We use DAPC (37) to identify species that have similar physiological niche parameters and examine the extent to which invasive, naturalized, and introduced species are associated with particular clusters in the multivariate physiological niche space. This multivariate analysis identified three clusters, and a 3×3 (three DAPC clusters \times three invasive status groups) contingency table was constructed (Fig. S3). We additionally calculated the niche breadth of each species on each of nine niche axes in the species distribution model (21). A Bayesian linear regression model (38) was used to test whether average niche breadth was influenced by invasive status (invasive, naturalized, or introduced); group (acacia or eucalypt) was considered a random effect in this model (Fig. 3).

ACKNOWLEDGMENTS. We thank Hadley O'Sullivan for setting up the computing facilities used for these analyses. S.I.H. acknowledges funding support from the University of Otago. D.M.R. acknowledges support from the DST-NRF (Department of Science and Technology - National Research Foundation) Centre of Excellence for Invasion Biology and National Research Foundation Grant 85417.

- MacArthur R (1972) *Geographical Ecology: Patterns in the Distribution of Species* (Princeton Univ Press, Princeton).
- Chesson P (2000) Mechanisms of maintenance of species diversity. *Annu Rev Ecol Syst* 31(3):343–366.
- Mack RN, et al. (2000) Biotic invasions: Causes, epidemiology, global consequences, and control. *Ecol Appl* 10(3):689–710.
- Richardson DM, et al. (2011) Human-mediated introductions of Australian acacias - a global experiment in biogeography. *Divers Distrib* 17(5):771–787.
- Tilman D (2004) Niche tradeoffs, neutrality, and community structure: A stochastic theory of resource competition, invasion, and community assembly. *Proc Natl Acad Sci USA* 101(30):10854–10861.
- Scheffer M, van Nes EH (2006) Self-organized similarity, the evolutionary emergence of groups of similar species. *Proc Natl Acad Sci USA* 103(16):6230–6235.
- Kolar CS, Lodge DM (2001) Progress in invasion biology: Predicting invaders. *Trends Ecol Evol* 16(4):199–204.
- Rejmánek M, Richardson DM (1996) What attributes make some plant species more invasive? *Ecology* 77(6):1655–1661.
- Hayes K, Barry S (2008) Are there any consistent predictors of invasion success? *Biol Invasions* 10(4):483–506.
- van Kleunen M, Weber E, Fischer M (2010) A meta-analysis of trait differences between invasive and non-invasive plant species. *Ecol Lett* 13(2):235–245.
- Hulme PE (2012) Weed risk assessment: A way forward or a waste of time? *J Appl Ecol* 49(1):10–19.
- Rejmánek M, Richardson DM, Higgins SI, Pitcairn MJ, Grotkopp E (2005) *Invasive Alien Species: A New Synthesis*, eds Mooney HA, McNeely JA, Neville L, Schei PJ, Waage JK (Island, Washington, DC), pp 104–161.
- Simberloff D (2009) The role of propagule pressure in biological invasions. *Annu Rev Ecol Syst* 40:81–102.
- Forsman A (2014) Effects of genotypic and phenotypic variation on establishment are important for conservation, invasion, and infection biology. *Proc Natl Acad Sci USA* 111(1):302–307.
- Peterson AT, Soberon J, Pearson R, Anderson RP, Martinez-Meyer E, Nakamura Miguel, Araujo MB (2011) *Ecological Niches and Geographic Distributions* (Princeton Univ Press, Princeton).
- Franklin J (2013) Species distribution models in conservation biogeography: Developments and challenges. *Divers Distrib* 19(10):1217–1223.
- Higgins SI, O'Hara RB, Roemerann C (2012) A niche for biology in species distribution models. *J Biogeogr* 39(12):2091–2095.
- Peterson AT (2003) Predicting the geography of species' invasions via ecological niche modeling. *Q Rev Biol* 78(4):419–433.
- Thuiller W, et al. (2005) Niche-based modelling as a tool for predicting the risk of alien plant invasions at a global scale. *Glob Change Biol* 11(12):2234–2250.
- Petitpierre B, et al. (2012) Climatic niche shifts are rare among terrestrial plant invaders. *Science* 335(6074):1344–1348.
- Higgins SI, et al. (2012) A physiological analogy of the niche for projecting the potential distribution of plants. *J Biogeogr* 39(12):2132–2145.
- Thornley JHM (1998) Modelling shoot: Root relations: The only way forward? *Ann Bot (Lond)* 81(2):165–171.
- Richardson DM, et al. (2000) Naturalization and invasion of alien plants: Concepts and definitions. *Divers Distrib* 6:93–107.
- Hui C, Richardson DM, Visser V, Wilson JR (2014) Macroecology meets invasion ecology: Performance of Australian acacias and eucalypts around the world revealed by features of their native ranges. *Biological Invasions* 16(3):565–576.
- Gaston KJ (1998) Species-range size distributions: Products of speciation, extinction and transformation. *Philos Trans R Soc Lond B Biol Sci* 353(1366):219–230.
- Richardson DM, Pyšek P (2012) Naturalization of introduced plants: Ecological drivers of biogeographical patterns. *New Phytol* 196(2):383–396.
- Hoffman AA, Parsons PA (1991) *Evolutionary Genetics and Environmental Stress* (Oxford Univ Press, Oxford).
- Kooijman SALM (2010) *Dynamic Energy Budget Theory for Metabolic Organisation* (Cambridge Univ Press, Cambridge, UK).
- Kearney MR, Simpson SJ, Raubenheimer D, Kooijman SALM (2013) Balancing heat, water and nutrients under environmental change: A thermodynamic niche framework. *Funct Ecol* 27(4):950–966.
- McInerney GJ, Etienne RS (2012) Ditch the niche—is the niche a useful concept in ecology or species distribution modelling? *J Biogeogr* 39(12):2096–2102.
- Crisp MD, et al. (2009) Phylogenetic biome conservatism on a global scale. *Nature* 458(7239):754–756.
- Batjes N (2005) ISRIC-Wise—Global Data Set of Derived Soil Properties on a 0.5 by 0.5 Degree Grid (Oak Ridge National Laboratory Distributed Active Archive Center, Oak Ridge, TN), Version 3.0, 10.3334/ORNLDAAAC. Available at www.daac.ornl.gov.
- Trabucco A, Zomer R (2010) Global soil water balance geospatial Database, Consortium for Spatial Information (CGIAR-CSI) (Washington, DC). Available at www.cgiar-csi.org.
- Hijmans RJ, Cameron SE, Parra JL, Jones PG, Jarvis A (2005) Very high resolution interpolated climate surfaces for global land areas. *Int J Climatol* 25(15):1965–1978.
- Barbet-Massin M, Jiguet F, Albert CH, Thuiller W (2012) Selecting pseudo-absences for species distribution models: How, where and how many? *Methods Ecol Evol* 3(2):327–338.
- Mullen K, Ardia D, Gil D, Windover D, Cline J (2011) DEoptim: An R package for global optimization by differential evolution. *J Stat Softw* 40(6):1–26.
- Jombart T (2008) adegenet: A R package for the multivariate analysis of genetic markers. *Bioinformatics* 24(11):1403–1405.
- Gelman A, Su YS (2013) arm: Data Analysis Using Regression and Multilevel/Hierarchical Models, R Package Version 1.6-10 R Foundation for Statistical Computing, Vienna, Austria.