Journal of Ecology

British Ecological Society

Journal of Ecology 2008, **96**, 869–883 doi: 10.1111/j.1365-2745.2008.01408.x

Threat or invasive status in legumes is related to opposite extremes of the same ecological and life-history attributes

Corey J. A. Bradshaw[1,2,3]*, Xingli Giam[4], Hugh T. W. Tan[4], Barry W. Brook[1] and Navjot S. Sodhi[4]

[1]Research Institute for Climate Change and Sustainability, School of Earth and Environmental Sciences, University of Adelaide, South Australia 5005, Australia; [2]South Australian Research and Development Institute, P.O. Box 120, Henley Beach, South Australia 5022, Australia; [3]School for Environmental Research, Institute of Advanced Studies, Charles Darwin University, Darwin 0909, Northern Territory, Australia; and [4]Department of Biological Sciences, National University of Singapore, 14 Science Drive 4, Singapore 117543, Republic of Singapore

Summary

1. The urgency and scale of the global biodiversity crisis requires the application of generalized predictors of a species' likelihood of going extinct or becoming invasive in non-native areas. A common approach is to correlate species' ecological and life-history characteristics (attributes, traits) with the probability of becoming either threatened (responding negatively to human activity), or invasive (responding positively). The limitation of previous studies is that the fates of becoming threatened or invasive have generally been treated in isolation.

2. Here we consider the problem of threat and invasiveness in unison based on analysis of one of the largest-ever species attributes data bases (8906 species) compiled for a single plant family (Fabaceae). We used generalized linear mixed-effects models (using taxonomic grouping to control for within-family phylogenetic relationships) to correlate species' life-history and ecological traits to three response variables: probability of being (i) threatened or not, (ii) invasive or not, and (iii) threatened or invasive.

3. We found that tall, annual, range-restricted species with tree-like growth forms, inhabiting closed-forest and lowland sites are more likely to be threatened. Conversely, climbing and herbaceous species that naturally span multiple floristic kingdoms and habitat types are more likely to become invasive.

4. *Synthesis*. These results support the idea that at least for one of the richest plant families, species' life-history and ecological traits correlate with a fate response to anthropogenic global change. Our results show that species do demonstrate particular susceptibility to either fate based on their evolved traits, and that traits generally correlated with invasiveness are also those that correlate with a reduced probability of becoming threatened.

Key-words: conservation, extinction, exotic species, Fabaceae, invasion, IUCN Red List, Leguminosae, non-indigenous species, range, threat

Introduction

The contemporary global biodiversity crisis necessitates a better understanding of extinction and invasion biology. These disciplines primarily focus on the identification and ranking of traits and environmental contexts that predict (or at least correlate with) a species' ultimate fate (Purvis *et al.* 2000; Kolar & Lodge 2001; Heger & Trepl 2003; Pimm *et al.* 2006; Sodhi *et al.* 2008). Establishing predictors of extinction risk or invasiveness is important for applying ecological theory to improve management efficiency and prioritize efforts to eradicate harmful invasive non-indigenous species (Rejmánek & Richardson 1996; Buckley, Briese & Rees 2003; Hamilton *et al.* 2005; Pyšek & Richardson 2007) or recover threatened taxa (Pimm *et al.* 2006; Sodhi *et al.* 2008). For example, predictors of a species' predisposition to a particular fate might be especially useful for screening taxa proposed for import to reduce the probability of allowing potentially weedy species into novel areas, or for established alien species

---

*Correspondence author. E-mail: corey.bradshaw@adelaide.edu.au for eradication. On the other hand, effective predictors could be used for selecting potentially sensitive taxa to monitor early detection of population decline. Developing evidence-based rules of thumb for categorizing poorly studied species into proneness groups will aid decision makers in choosing the best ways to allocate finite conservation resources (Duncan & Young 2000; Cadotte, Murray & Lovett-Doust 2006).

Despite extensive research, the relative importance of life-history and ecological attributes in predisposing a species to a particular fate remains controversial (Rejmánek & Richardson 1996; Williamson 1999; Kolar & Lodge 2001; Heger & Trepl 2003; Croci, Le Quilliec & Clergeau 2007). Current evidence certainly suggests that particular evolved ecological niches and combinations of life-history and ecological characteristics (organism size, dispersal capacity and native geographic range prior to human interference) and other reproductive, dispersal, morphological and physiological attributes can influence a species' proneness to extinction or to become invasive (Goodwin, McAllister & Fahrig 1999; Duncan & Young 2000; Heger & Trepl 2003; Wu, Chaw & Rejmánek 2003; Hamilton *et al.* 2005; Croci *et al.* 2007; Sodhi *et al.* 2008), with the strength of effect often depending on environmental context (Williamson 1996; Kolar & Lodge 2001; Pimm *et al.* 2006) and the spatial scale of investigation (Stohlgren *et al.* 1999; Stohlgren *et al.* 2002; Pyšek & Richardson 2007). We hypothesize that traits which predispose a species to becoming threatened would also reduce that species' likelihood of responding positively to abrupt change (i.e. the combined probability of being introduced to a novel environment by humans and becoming invasive), and vice versa. Indeed, it has been shown that rare species tend to differ in many life-history and ecological traits from common species – rare species tend to demonstrate lower reproductive effort and dispersal capacity, and more restricted geographic ranges (Kunin & Gaston 1993; Kunin & Gaston 1997; Blackburn & Cassey 2004; Pocock *et al.* 2006). Yet, remarkably, the disciplines of extinction and invasion biology have rarely overlapped (although see Blackburn & Cassey 2004 for a comparison of introduced and re-introduced bird species), with biologists searching mainly instead for correlates of extinction proneness (Duncan & Young 2000; Walker & Preston 2006; Sodhi *et al.* 2008) or invasiveness (Goodwin *et al.* 1999; Heger & Trepl 2003; Hamilton *et al.* 2005; Croci *et al.* 2007) in isolation.

Previous studies of extinction and invasion have been limited by a variety of constraints. Evolved life-history attributes covary between related species, but the confounding effects of sharing evolutionary history in cross-species comparisons (Felsenstein 1985) can be overcome by techniques such as independent contrasts (Purvis *et al.* 2000), mixed-effects models (Brook, Traill & Bradshaw 2006; Sodhi *et al.* 2008) and the comparison of closely related species (e.g. congenerics) (Goodwin *et al.* 1999; Gerlach & Rice 2003; Pyšek & Richardson 2007). Factors limiting effective comparison include insufficient sample size (number of species compared) to detect subtle correlations (Lloret *et al.* 2005), unreliability of control groups (e.g. artificial selection modifying attributes of non-indigenous species), geographic idiosyncrasies, lack of information on unobserved extinctions or failed invasions, mismatched or inappropriate spatial scale, differences in environmental contexts, and global inferences made using taxa from a single region (Duncan & Young 2000; Kolar & Lodge 2001; Pyšek, Richardson & Williamson 2004; Hamilton *et al.* 2005; Lloret *et al.* 2005; Pyšek & Richardson 2007).

To overcome many of these problems, we compiled one of the largest species' attributes data bases from a single taxonomic group that spans the majority of the world's terrestrial niches. The family Fabaceae (also known as the Leguminosae) of the Eudicots is the third-most speciose family of flowering plants, with 19 325 species classified into three sub-families (Papilionoideae, Caesalpinioideae and Mimosoideae) and approximately 727 genera (Lewis *et al.* 2005), representing > 7% of all *c.* 260 000 known flowering plant species (Angiosperm Phylogeny Group 2003). Spread throughout all continents and terrestrial biomes except Antarctica, legumes range in size from dwarf herbs to large tropical trees. The family's high species richness, and its life history, form and functional diversity makes it one of the most important plant groups for humans in terms of food production, fodder, medicines, timber and other commercial products (Harborne 1994).

Our explicit aim was to test the hypothesis that evolved ecological and life-history traits predict a species' propensity to become threatened or invasive, but that the specific outcome depends on opposite extremes of the same traits (a 'fate dichotomy'). Specifically, we tested whether the probabilities of being classed as 'threatened', 'invasive', or 'neutral' (i.e. neither 'threatened' nor 'invasive') were correlated with 15 ecological and life-history attributes. These ecological and life-history attributes describe habitat and range, life-history strategy, reproduction, dispersal, defence mechanisms and leaf characteristics (see Appendix S1 in Supplementary Material for more a detailed justification and description of correlates used and hypotheses related to predicting fate risk). We did not consider phenotypic plasticity, adaptation, or success based on human intervention (e.g. propagule pressure, human affiliation; Hamilton *et al.* 2005; Jeschke & Strayer 2006). Although propagule pressure is considered one of the major determinants of invasion success (Lockwood, Cassey & Blackburn 2005), the detailed history of each species was simply not available, and it was not our explicit objective to predict the particular outcome of each species individually. Rather, our central aim was to examine the propensity of a species to become invasive or threatened given its evolved ecological and life-history template (under the same environmental conditions and histories), and whether statistically identified correlates support the notion of higher susceptibility to either fate rather than two distinct and unrelated processes.

Methods

GLOBAL DATA BASE OF THE FAMILY FABACEAE

We compiled a representative list of 8906 Fabaceae species from online data bases, published Floras, taxonomic Manuals, checklists and monographic revisions (data available from the authors upon request). Several data bases containing information describing aspects of Fabaceae biology have been assembled over the last 20 years, such as the International Legume Data base and Information Service (ILDIS) (www.ildis.org) and Germplasm Resources Information Network (GRIN) (www.ars-grin.gov). Both the data bases contain valuable information on taxonomic synonymy, global distribution and root nodulation data. However to date, there has been no data base that provides simultaneous information on distribution, ecology and life-history attributes of a large proportion of the legumes. We therefore combined elements of the various data bases, published Floras, Manuals, checklists and monographic revisions together to construct a globally representative data base. We used this to test the hypothesis that evolved characteristics (described below) can be used to predict a species' particular fate.

To ensure ample representation of every geographical region in which legumes are found, we recorded species from all three subfamilies (Papilionoideae, Caesalpinioideae and Mimosoideae) from at least one full Flora/Manual from each floristic kingdom (Takhtajan 1986). In the geographically larger Holoarctic and Paleotropic kingdoms, nine and 13 Floras, respectively, were consulted. Floras consulted for at least one full Subfamily are listed in Supplementary Appendix S2 (List S1), and the complete list of Floras and Manuals referred to in this assessment is presented in List S2 (Supplementary Appendix S2). In areas where complete Flora treatments have not been published, species were recorded from every available publication that contained descriptive information (e.g. monographic revisions and papers). A complete list of publications consulted is presented in List S3 (Supplementary Appendix S2). In cases where > 1 subspecies were listed as separate entries with different ecological responses ('invasive', 'threatened' or neither), we included only the subspecies considered 'threatened' or 'invasive'. We omitted two species listing both 'invasive' and 'threatened' subspecies (see complete subspecific filtering criteria and species removed or retained in Supplementary Appendix S1, Table S1).

THREAT AND INVASIVENESS

'Threatened' species were identified according to IUCN Red List criteria (*Critically Endangered*, *Endangered*, *Vulnerable* and *Near Threatened*). Nearly all IUCN Red-Listed species were included in our data set, so the sample is unbiased with respect to that classification system. Invasive species were identified using the definition of Pyšek, Prach and Smilauer (1995) as 'alien species increasing in both anthropogenic and natural habitats'. Invasive species were obtained from the global checklist of weeds compiled by Randall (2002) and the invasive plant guide by Weber (2003). In line with the adopted definition, only species and subspecies classified as 'weed', 'environmental weed' and 'noxious weed' in Randall (2002) were considered 'invasive'. Species classified as 'quarantine weed', 'native weed', 'naturalized', 'sleeper weed' or 'casual alien' were not considered 'invasive'. Species classified as 'quarantine weed' are species prohibited entry under a country's quarantine regulations, and hence, do not indicate invasiveness. Species classified under 'native weed' were excluded because they do not satisfy the criterion of being alien. Species classified as 'naturalized', 'sleeper weed', or 'casual alien' were also excluded because these terms do not reflect an increase in population size.

To validate the results of the global analysis that may have biases associated with underrepresented species in the IUCN Red List (Lughadha *et al.* 2005), we also compiled Fabaceae species from a single, well-studied region – Canada and the USA. We compiled 1158 species classed from GH to G5 from the NatureServe rare and endangered species data base (www.natureserve.org/explorer). We classed all GH (Possibly Extinct), G1 (Critically Imperilled) and G2 (Imperilled) legume species in the NatureServe data base as 'threatened', and all others as 'not threatened'. Species with range ranks across our definitions (e.g. G2G3 and G2G4) of 'threatened' and 'not threatened' indicate uncertainty in status and were removed prior to analysis. We were thus left with 491 legume species for which we had trait information. Threat status was correlated with our compiled ecological and life-history attributes (see below).

ECOLOGICAL AND LIFE-HISTORY ATTRIBUTES

For each of the species, we compiled attributes hypothesized to relate to a species' propensity to become threatened/extinct (Thompson 1994; Turner *et al.* 1994; McIntyre, Lavorel & Tremont 1995; Duncan & Young 2000; Walker & Preston 2006; Sodhi *et al.* 2008), invasive (Pyšek *et al.* 1995; Lake & Leishman 2004; Hamilton *et al.* 2005; Cadotte *et al.* 2006), and general plant survival strategies (Westoby 1998). These included information on geographic distribution and range, altitudinal range, habitat, life strategy, growth habit, height, presence of armaments, presence of indumentum (hairs), length of the smallest leaf unit, floral display, presence of hooks on the fruit, fruit dehiscence and seed size. More specific detail on the justification and particular hypotheses related to the attributes chosen (or omitted) can be found in Supplementary Appendix S1. All attributes were either uncorrelated or weakly inter-correlated (Spearman's ρ and Kendall's τ) with the exception of *habit* and *height* ($\rho = -0.746$, $\tau = -0.687$; Table S2); regardless, moderately correlated variables do not compromise model ranking procedures when based on estimates of relative parsimony (see Methods section).

STATISTICAL ANALYSES

To determine the relationships between the ecological and life-history traits and the threat and invasiveness risks of the compiled legume species, we fitted generalized linear mixed-effect models (GLMM) to the data using the lmer function implemented in the *R* Package v 2.4 (R Development Core Team 2004). All GLMM constructed had a binomial response variable (see below for details) and a logit link function.

Species are phylogenetic units with shared evolutionary histories and are therefore not statistically independent (Felsenstein 1985). Even though we sampled species from only a single plant family (Fabaceae), it was still necessary to decompose the variance across species by coding the random-effects error structure of the GLMM as a hierarchical taxonomic effect (Blackburn & Duncan 2001b). We had adequate replication to use the nested random effect of Subfamily/Tribe, but insufficient replication within genera to include Genus (there were many instances of monospecific genera). We also included a second set of models using the nested taxonomic effect of Tribe within nominal phylogenetic group (NPG) to examine whether our choice of taxonomic association affected our conclusions (few differences arise – see Results section – so we base the majority of our analyses on the Subfamily/Tribe random effect). Our method is more appropriate than the independent-contrasts approach (Purvis *et al.* 2000) in situations where a complete phylogeny of the study taxon is unavailable, when categorical variables are included in the analysis (as was the case in this study), and when model selection, rather than Neyman–Pearson hypothesis testing, is the statistical paradigm being used for inference (Blackburn & Duncan

2001a). The amount of variance in the threat and invasiveness probability response variable captured by each model considered (see below) was assessed as the percent deviance explained (%DE), expressed relative to the deviance of a null model with no fixed effects but retaining the hierarchical random effect (Brook *et al.* 2006).

Our model-building strategy used existing knowledge from other studies, ecological theory and logic to construct a plausible set of *a priori* hypotheses to examine the predictors of threat and invasiveness risk in plants. We split the modelling approach into three phases to examine related questions: (i) **Phase 1** examined threat risk, which was defined as whether a species was IUCN Red-Listed (threatened = all categories except 'Least Concern' and Data Deficient') or not (including neutral (neither threatened nor invasive) and invasive species). The ecological and life-history correlates height, range, habitat, altitudinal range, life cycle, growth habit, presence/absence of hooks on fruit, presence/absence of fruit dehiscence, seed size, presence/absence of armaments, presence/absence of hairs on vegetative parts, and maximum length of leaf lamina were modelled as linear predictors (fixed factors) of the binomial response. The variables height, range, growth habit and maximum length of leaf lamina were coded as ordered-level (ordinal) factors, with all others coded as nominal factors. We further considered models where height and maximum length of leaf lamina were coded as $\log_{10}$-scaled continuous variables to account for potential bias associated with multi-level ordinal factors (see Supplementary Appendix S1). No differences were detected using covariates in the initial threat analysis (see Results section), so they were not considered further. No interactions were considered. Various combinations of the ecological life-history terms were built under themes which represent broad hypotheses about the dominant determinant of threat/invasiveness: allometry, habitat models, life-history strategy models, reproduction and dispersal models, defence models and various combinations of these themes ($n = 65$ models; Supplementary Table S3); (ii) **Phase 2** examined the response variable invasiveness as a binary coded variable (invasive or not invasive, with the latter category including all neutral and threatened species) using the same model set as *threat*; (iii) **Phase 3** used the same model set as above, but the response variable was coded as either *threatened* or *invasive* (i.e. after removing all 'neutral' species).

To avoid potential circularity in correlations between IUCN threat status based on restricted range size criteria and native range, we repeated the threat-risk analysis (Phase 1) by removing all 'threatened' species (as defined above) that were designated as such due to small range sizes (IUCN Red List Criterion B, D2 or both – www.iucnredlist.org). All models used in the full analysis were repeated with this reduced set to determine whether conclusions reached using the full data set were affected by potential circularity. The restricted data set was also used to verify the conclusions reached in Phase 3 (*threatened* or *invasive*).

We used an index of Kullback–Leibler (K-L) information loss to assign relative strengths of evidence to the different competing models (Burnham & Anderson 2002), Akaike's Information Criterion corrected for small sample size (AIC_c), as well as the dimension-consistent Bayesian information criterion (BIC), an approximation of the Bayes factor given no informative prior information on relative model support (Burnham & Anderson 2002). These indices of model parsimony identify the relative evidence of model(s) from a set of candidate models. The relative likelihoods of candidate models were calculated using AIC_c and BIC weights (Burnham & Anderson 2002), with the weight ($w\text{AIC}_c$ and wBIC) of any particular model varying from 0 (no support) to 1 (complete support) relative to the entire model set. However, the K-L prior used to justify AIC_c weighting can favour more complex models when sample sizes are large (Burnham & Anderson 2004; Link & Barker 2006), as was the case for our data set). As such, we considered BIC weighting as most appropriate for determining the contribution of the most important correlates of extinction risk (major effects only), and AIC_c weighting to identify the most parsimonious models for maximizing prediction accuracy (major terms plus tapering effects; Burnham & Anderson 2004; Link & Barker 2006). The sample size for each model set was reduced in most cases owing to some missing data for some species in some of the hypothesized correlates (updated sample sizes given in Results); however, all model comparisons within a set were made using the same number of species so that AIC_c and BIC were comparable. Consistency in results among the analysis phases (see Results section) suggests that conclusions remained robust to reductions in sample size necessitated by missing data for some species' traits.

Results

The 8906 legume species we sampled reliably (and after filtering – Supplementary Appendix S1) mirrored Subfamily diversity: 63% were from the most diverse Subfamily Papilionoideae, 24% from the Mimosoideae and 13% from the Caesalpinioideae (Table 1), compared to the worldwide distribution of all Fabaceae species (71%, 17% and 12%, respectively based on 19 325 species; Supplementary Fig. S1a) (Lewis *et al.* 2005). All 36 known tribes were represented in our data base (Lewis *et al.* 2005), and species were sampled from all six floristic kingdoms (Takhtajan 1986). There was considerable variation in the proportion of threatened, neutral and invasive legumes among kingdoms (Supplementary Fig. S2). Likewise, our data base closely mimicked the world distribution of species according to life cycle and major growth habit categories (Supplementary Fig. S1b,c). Of all species sampled, 46.6% were missing data for seed size (4756 of 8906 species), so we removed this term from the model set for the main analysis (thus reducing the total number of models considered from 65 to 57; see Supplementary Table S3), but we did also consider the reduced data set with seed size separately (see below).

Using *threat* as the binary response (i.e. 'threatened' or 'not', where the 'not' category represented all 'neutral' and invasive species; $n = 2737$), the two most highly supported GLMM, as ranked by the BIC, demonstrated important contributions of native geographic range, plant height, habitat, altitudinal distribution, life cycle and growth habit (combined $w\text{BIC} = 0.999$) explaining > 16% of the deviance (Table 2). These were also the most parsimonious models according to AIC (corrected for small sample size) weights ($w\text{AIC}_c$), suggesting a lack of tapering effects. The term growth habit alone accounted for nearly 12% of the deviance (%DE), with important contributions of height (single-term model with %DE = 4.9%). In general, taller (Fig. 1a), range-restricted (Fig. 1b) species were more threatened, as well as species occupying closed forests (Fig. 1c) and those spanning only one altitudinal band (Fig. 1d). Annual species were more likely to be classed as threatened than perennial species

Table 1. Taxonomic and species attribute summary for the 8906 sampled threatened (Thr), invasive (Inv) and neutral (Ntr) Fabaceae species (Subfamilies, Tribes, Genera, and nominal phylogenetic groups [NPG – see Supplementary Appendix S1 for description]) examined. Note that not all attribute level sample sizes sum to 8906 due to missing data

Taxonomy/status	Number of spp/gr	Attribute	Level	Number of spp	Proportion of Thr	Proportion of Ntr	Proportion of Inv
Papilionoideae	5591	Height	≤ 1 m	3274	0.003	0 875	0.121
Caesalpinioideae	1207		> 1 and ≤ 10 m	2935	0.080	0.801	0.119
Mimosoideae	2108		> 10 m	1293	0.210	0.715	0.075
Number of tribes	36	Altitudinal range	Lowland	3399	0.136	0.759	0.106
Number of genera	537		Montane	1382	0.046	0.915	0.039
			Both	978	0.015	0.724	0.261
'threatened'	688	Habitat	closed forest	1703	0.258	0.733	0.008
'neutral'	7298		Open	4673	0.042	0.851	0.107
'invasive'	920		Both	1080	0.028	0.688	0.284
NPG1-cercidae	183	Range	1 floristic kingdom	8599	0.080	0.826	0.094
NPG2-detarieae	402		≥ 2 floristic kingdoms	304	0.007	0.625	0.368
NPG3-cassieae	50	Life cycle	Annual	918	0.004	0.732	0.264
NPG4-mimosoideae	2675		Perennial	7960	0.086	0.829	0.085
NPG5-papilionoideae	5596	Habit	Tree	2126	0.260	0.656	0.084
			Shrub	2873	0.037	0.890	0.073
			Climber	723	0.015	0.860	0.124
			Herb	3168	0.005	0.856	0.139
		Floral display	Solitary	598	0.030	0.865	0.105
			Inflorescence	8223	0.077	0.819	0.104
		Hooks on fruit	Present	455	0.009	0.741	0.251
			Absent	7129	0.072	0.830	0.098
		Fruit dehiscence	Present	4515	0.056	0.829	0.115
			Absent	1937	0.138	0.749	0.114
		Seed size	≤ 10 mm	3958	0.032	0.819	0.149
			> 10 and ≤ 20 mm	545	0.169	0.719	0.111
			> 20 mm	253	0.174	0.787	0.040
		Armaments	Present	692	0.053	0.645	0.302
			Absent	8110	0.074	0.839	0.087
		Hairs	Present	6690	0.061	0.828	0.111
			Absent	2102	0.110	0.813	0.077
		Maximum leaf lamina length	≤ 4 cm	4753	0.039	0.841	0.120
			> 4 and ≤ 8 cm	1582	0.093	0.794	0.113
			> 8 and ≤ 12 cm	906	0.113	0.819	0.068
			> 12 and ≤ 16 cm	613	0.131	0.796	0.073
			> 16 and ≤ 20 cm	365	0.132	0.830	0.038
			> 20 cm	323	0.186	0.789	0.025

(Fig. 1e; although see below), and trees were more threatened than shrubs, climbers and herbs (Fig. 1f). The NatureServe threatened classification (219 species remaining after removing all species with missing attribute data) analysis revealed strongest support for the saturated model based on $w\mathrm{AIC}_c$ (Supplementary Table S4), with the greatest contribution to %DE from life cycle, height, range and altitude. This confirms that the results based on the global data base of IUCN Red-Listed species appear valid.

Including the term seed size reduced the IUCN Red List *threat* analysis data set to 1693, and there was only extremely weak support for this term's contribution to the model weights ($w\mathrm{AIC}_c << 0.001$). Treating height and maximum length of leaf lamina as covariates altered goodness-of-fit and model ranking only slightly (Supplementary Table S5), with height having a positive coefficient with increasing threat probability (data not shown). Substituting nominal phylogenetic group (NPG) for Subfamily in the nested taxonomic random effect altered model ranking and goodness-of-fit only marginally (Supplementary Table S6), so we used the Subfamily/Tribe random effect in all subsequent analyses.

Removing Criteria B and D2 threatened species (approximately one third of species designated as 'threatened') produced nearly identical model rankings (Table 3), with an even higher goodness-of-fit (%DE ≅ 19%). The directions of the relationships for prediction threat risk based on the reduced data set were nearly identical to the full data set (Fig. 2), with one notable difference – the reduced data set produced a higher predicted threat risk for perennial compared to annual species (Fig. 2e). This was opposite to the trend revealed by the full data set (cf. Fig. 1e), mainly because so few annuals were present in the reduced data set. We also considered only the IUCN-listed species in a separate analysis where *Critically Endangered*, *Endangered* and *Vulnerable*

Table 2. The five most parsimonious generalized linear mixed-effects models investigating the life-history correlates of threat risk for legumes (full data set, $n = 2737$ species, 310 'threatened') according to (a) the Bayesian Information Criterion (BIC) and (b) Akaike's Information Criterion corrected for small sample size (AIC_c). The five most highly BIC-ranked models accounted for > 99% of the information-theoretic weight (*w*BIC) of the total of 57 models considered (see Table S3). Terms shown are HT = height, RG = range, HBT = habitat, ALT = altitude, LC = life cycle, HB = habit. Also shown are the number of parameters (*k*), the maximum log-likelihood (*LL*), the difference in BIC and AIC_c for each model from the most parsimonious model (ΔBIC and ΔAIC_c), AIC_c weight, and the percent deviance explained (%DE) in the response variable by the model under consideration

Model	*k*	*LL*	ΔBIC	*w*BIC	ΔAIC_c	$wAIC_c$	%DE
(a) BIC-ranked							
~RG + HBT + ALT + LC + HB	13	–675.486	0.000	0.857	2.468	0.223	16.37
~HT + RG + HBT + ALT + LC + HB	15	–671.939	3.576	0.143	0.000	0.767	16.81
~RG + HBT + LC + HB	11	–688.202	14.942	< 0.001	23.345	< 0.001	14.80
~RG + ALT + LC + HB	11	–689.791	18.191	< 0.001	26.514	< 0.001	14.60
~HT + RG + HBT + LC + HB	13	–684.634	18.507	< 0.001	20.731	< 0.001	15.24
(b) AIC_c-ranked							
~HT + RG + HBT + ALT + LC + HB	15	–671.939	3.576	0.143	0.000	0.767	16.81
~RG + HBT + ALT + LC + HB	13	–675.486	0.000	0.857	2.468	0.223	16.37
saturated model	25	–663.932	40.56	< 0.001	8.761	0.010	17.80
~HT + RG + HBT + LC + HB	13	–684.634	18.507	< 0.001	20.731	< 0.001	15.24
~HT + RG + ALT + LC + HB	11	–688.202	14.942	< 0.001	23.345	< 0.001	14.80

Table 3. The five most parsimonious generalized linear mixed-effects models investigating the life history correlates of threat risk for legumes after removing species designated as 'range-restricted' as the principal reason for IUCN Red Listing (Categories B and D2) (full data set, $n = 2426$ species, 112 'threatened') according to the Bayesian Information Criterion (BIC) and Akaike's Information Criterion corrected for small sample size (AIC_c). The five most highly BIC- and AIC_c-ranked models accounted for 99% of the information-theoretic weight (*w*BIC and $wAIC_c$) of the total of 57 models considered (see Supplementary Table S3). Terms shown are HT = height, RG = range, HBT = habitat, ALT = altitude, LC = life cycle, HB = habit. Also shown are the number of parameters (*k*), the maximum log-likelihood (*LL*), the difference in BIC and AIC_c for each model from the most parsimonious model (ΔBIC and ΔAIC_c), AIC_c weight, and the percent deviance explained (%DE) in the response variable by the model under consideration

Model	*k*	*LL*	ΔBIC	*w*BIC	ΔAIC_c	$wAIC_c$	%DE
BIC-ranked							
~RG + HBT + ALT + LC + HB	13	–322.442	0.000	0.698	0.000	0.583	19.26
~RG + ALT + LC + HB	11	–328.457	3.286	0.135	6.491	0.023	17.75
~HT + RG + HBT + ALT + LC + HB	15	–320.105	4.239	0.084	1.112	0.334	19.84
~HT + RG + ALT + LC + HB	13	–324.816	4.894	0.060	4.682	0.056	18.66
~RG + HBT + LC + HB	11	–331.020	8.416	0.010	11.616	0.002	17.11

species were classed this time as 'threatened' and *Near Threatened* (formerly, *Lower Risk*) species as 'not threatened' to account for the probability of species not being assessed in the Red List. This left only a total of 310 species of which 268 were 'threatened', making the application of the full model set impractical. Thus, we considered only univariate models and showed that habit accounted for the most deviance in this threat risk response (although overall goodness-of-fit was low; Supplementary Table S7). Removing range-restricted species (Criteria B and D2 species) prevented model convergence due to low sample sizes ($n = 112$ total species). This lack of power suggests caution when interpreting these results, although the similarity in model weights demonstrates the small contributions of each of the factors considered (Supplementary Table S7).

Using *invasive* as the binary response (i.e. 'invasive' or 'not', where the 'not' category represented all 'neutral' and threatened species; $n = 3046$), the saturated model was supported by both *w*BIC and $wAIC_c$ (weights > 0.999; Table 4) and accounted for > 16% of the deviance. The models range + altitude and range + habitat accounted for nearly 9% and 12%, respectively, of the deviance. Even though the saturated model had most support, range, habitat, altitude, life cycle and growth habit (second-ranked model) were the main effects explaining deviance in the response. Although the reduced data set produced equivocal results for the effects of life cycle on threat status (see above Phases 1 and 2; Figs 1 and 2), annual plants tended to be more invasive than perennials (Fig. 3e). In general then, traits that lead to higher threat risk also reduce the likelihood of becoming invasive. Species spanning multiple floristic kingdoms (Fig. 3b), both closed forest and open habitats (Fig. 3c), and both lowland and montane altitudinal ranges (Fig. 3d), were the most likely to be invasive. Likewise, climbers and herbs were generally more likely to be invasive than trees and shrubs (Fig. 3f). Including seed size in the analysis reduced the invasive analysis data set to 1987 species – seed size was including in the most highly BIC-ranked model (saturated; *w*BIC = 0.985). However, the single-term model only accounted for 0.35% of the deviance, suggesting a negligible effect of seed size on invasiveness. We

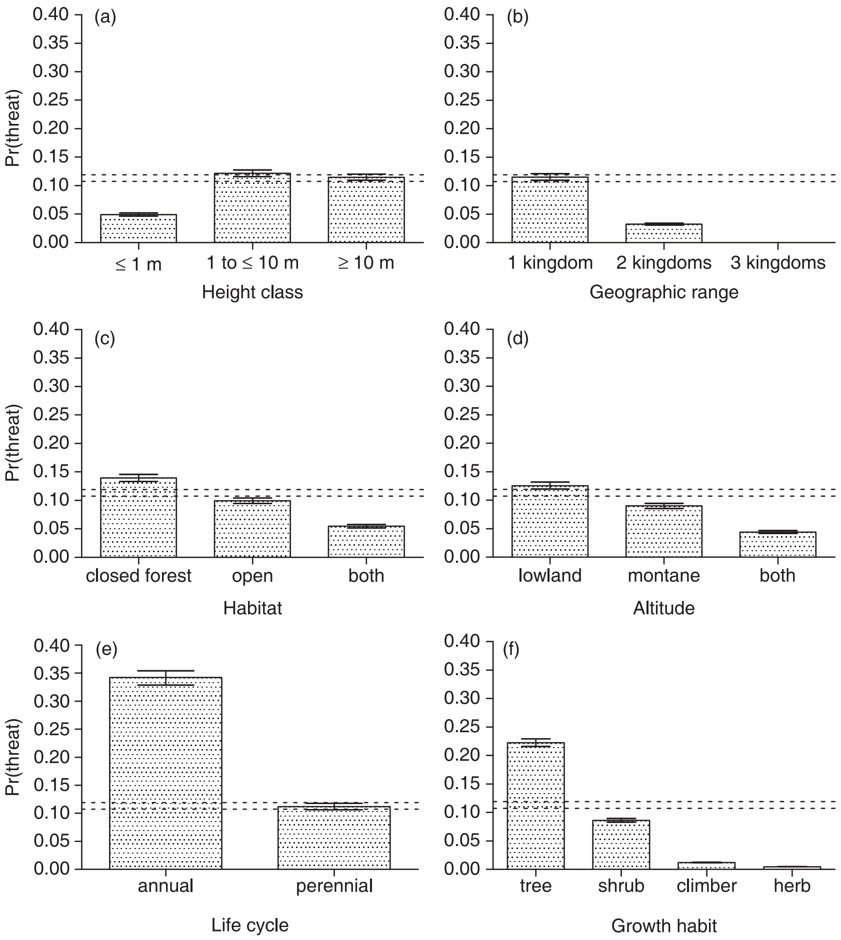

Fig. 1. Probability of a legume species being classed as 'threatened' (Pr[*threat*]) predicted for each category of term included in the most parsimonious generalized linear mixed-effects models (Pr[*threat*] ~ height + range + habitat + altitude + life cycle + habit; Table 2) based on the full data set. Dotted horizontal lines represent the bootstrapped (10 000 iterations) 95% confidence interval limits for Pr[*threat*] derived from the aforementioned model based on the observed data for 2737 species. These predictions of the observed data's threat range are based only on the model above, so they do not describe unexplained variation in the original data set nor do they include data from species missing trait data. Hence, they do not necessarily reflect the proportional results shown in Table 1. Predicted Pr[*threat*] for each term level were calculated by adjusting the original data set so that all species were given the same value for that level (each level in turn), keeping all other terms in the model as in the original data set. Error bars represent the 10 000-iteration bootstrapped upper 95% confidence limits. The percent deviance explained (%DE) in Pr[*threat*] by the univariate models for each term shown are: (a) species ≤ 1 m, 1 to ≤ 10 m or > 10 m in height (%DE = 4.9%), (b) species restricted to either 1, 2 or 3 floristic kingdoms (%DE = 0.8%), (c) species restricted to closed forest, open, or spanning both habitats (%DE = 4.6%), (d) species restricted to lowland, montane, or both altitudinal ranges (%DE = 3.2%), (e) annual or perennial species (%DE = 0.2%), and (f) trees, shrubs, climbers or herbs (%DE = 11.7%). See Appendix S1 for a detailed description of the attributes shown.

also investigated the remaining life-history terms for their contributions given the support for the saturated model (Table 4). Of the six remaining terms, presence/absence of armaments, presence/absence of hairs on vegetative parts, and presence/absence of hooks on fruit were supported. In general, the presence of hooks on fruit, armaments such as spines and prickles, and vegetative hairs increased invasiveness risk (Fig. 4a–c).

The reduced data set (i.e. removing all 'neutral' species and range-restricted 'threatened' species; $n = 606$) using the classification of *threatened* or *invasive* as the binary response variable confirmed the notion that traits increasing the probability of becoming threatened were generally also those that reduced the probability of becoming invasive (Fig. 5). Again, the effect of life cycle was somewhat equivocal (Fig. 5e), although this comparison supported the hypothesis that annual species were less likely to be threatened than perennials. The height + range + habitat + altitude + life cycle + habit model had the strongest support (combined *w*BIC = 0.851) and accounted for *c.* 62% of the deviance (Table 5). Incorporating seed size ($n = 474$ species) provided little support for that term (the saturated model including *seed size* had *w*BIC < 0.02).

Discussion

Our results for legumes support the hypothesis that particular combinations of ecological and life-history attributes

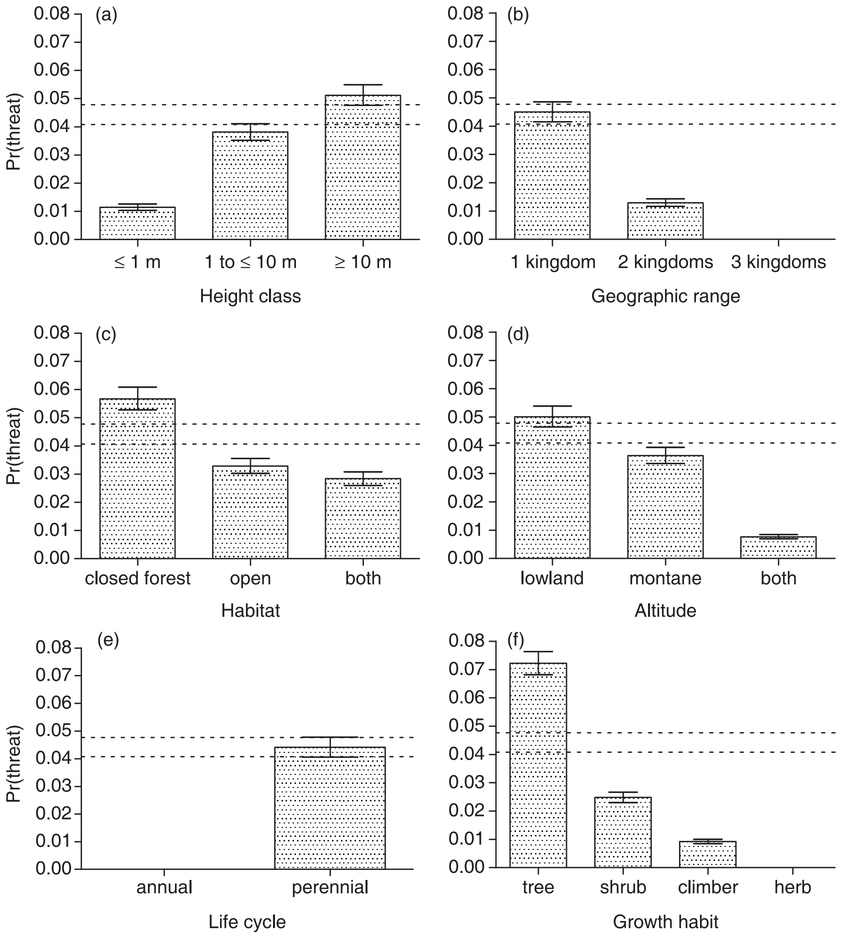

Fig. 2. Probability of a legume species being classed as 'threatened' (Pr[*threat*]) predicted for each category of term included in the most parsimonious generalized linear mixed-effects models (Pr[*threat*] ~ height + range + habitat + altitude + life cycle + habit; Table 3) using the reduced data base (i.e. after removing *threatened* species classed as such due to restricted ranges – see Methods). Dotted horizontal lines represent the bootstrapped (10 000 iterations) 95% confidence interval limits for Pr[*threat*] derived from the aforementioned model based on the observed data for 2426 species. Predicted Pr[*threat*] for each term level were calculated by adjusting the original data set so that all species were given the same value for that level (each level in turn), keeping all other terms in the model as in the original data set. Error bars represent the 10 000-iteration bootstrapped upper 95% confidence limits. The per cent deviance explained (%DE) in Pr[*threat*] by the univariate models for each term shown are: (a) species ≤ 1 m, 1 to ≤ 10 m or > 10 m in height (%DE = 9.4%), (b) species restricted to either 1, 2 or 3 floristic kingdoms (%DE = 0.5%), (c) species restricted to closed forest, open, or spanning both habitats (%DE = 6.0%), (d) species restricted to lowland, montane, or both altitudinal ranges (%DE = 4.4%), (e) annual or perennial species (%DE = 1.1%), and (f) trees, shrubs, climbers or herbs (%DE = 14.6%). See Appendix S1 for a detailed description of the attributes shown.

Table 4. The five most parsimonious generalized linear mixed-effects models investigating the life history correlates of invasiveness for legumes (full data set, *n* = 3046 species, 491 'invasive') according to both (a) the Bayesian information criterion (BIC) and (b) Akaike's information criterion corrected for small sample size (AIC$_c$) (equivalent ranking using either metric). The most highly ranked model (saturated) accounted for > 99% of the information-theoretic weight (*w*BIC and *w*AIC$_c$) of the total of 57 models considered (see Supplementary Table S3). Terms shown are HT = height, RG = range, HBT = habitat, ALT = altitude, LC = life cycle, HB = habit. Also shown are the number of parameters (*k*), the maximum log-likelihood (*LL*), the difference in BIC and AIC$_c$ for each model from the most parsimonious model (ΔBIC and ΔAIC$_c$), AIC$_c$ weight, and the percent deviance explained (%DE) in the response variable by the model under consideration

Model	*k*	*LL*	ΔBIC	*w*BIC	ΔAIC$_c$	*w*AIC$_c$	%DE
Saturated	26	–917.053	0.000	> 0.999	0.000	> 0.999	22.18
~RG + HBT + ALT + LC + HB	14	–968.014	35.527	< 0.001	73.703	< 0.001	17.86
~RG + HBT + ALT	10	–984.298	45.627	< 0.001	97.511	< 0.001	16.48
~HT + RG + HBT + ALT + LC + HB	16	–967.446	45.706	< 0.001	77.040	< 0.001	17.91
~HT + RG + HBT + ALT	12	–983.502	55.369	< 0.001	100.255	< 0.001	16.54

characterize plant species at one or other extreme of a fate continuum between the states of threatened and invasive. We are aware of only two other studies (Sutherland 2004; Cardillo *et al.* 2005) that have compared the attributes of either threatened/non-threatened, or invasive/non-invasive species among a comparable or greater number of species. However, those studies examined only one fate extreme across species from many families and used only univariate (Sutherland 2004) or multiple regression *t*-tests (Cardillo *et al.* 2005) of particular traits, thus potentially confounding perceived relationships by ignoring the simultaneous effects of multiple correlates. Our study is a major advance because (i) we examined the attributes pertaining to a species fate between the two extreme categories (threatened and invasive), which was particularly informative in terms of the models' greater explanatory power (%DE – cf. Tables 1–4 with Table 5) and for confirmation that the traits related to species becoming invasive are generally those that are correlated with lower threat risk; (ii) we compiled information for species from a single (albeit diverse) family, thereby reducing the potentially

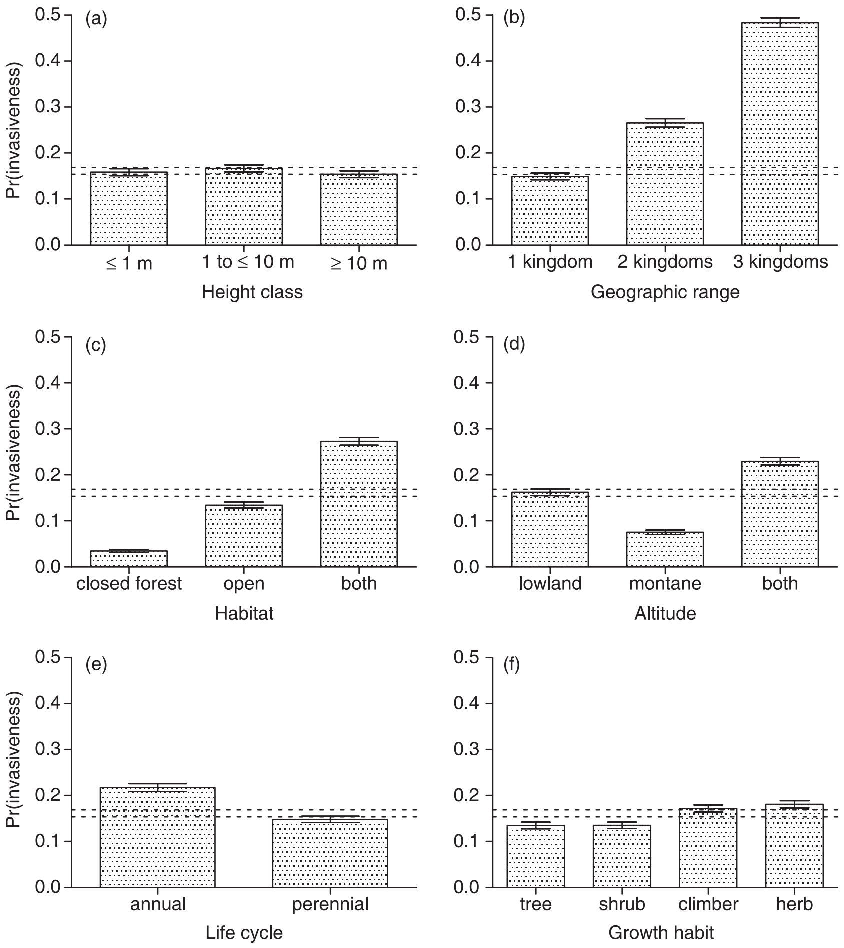

Fig. 3. Probability of a legume species being classed as 'invasive' (Pr[*invasive*]) predicted for each category of term included in the most parsimonious generalized linear mixed-effects models (saturated model: Pr[*invasive*] ~ height + range + habitat + altitude +life cycle + habit + armaments + hairs + floral display + dehiscence + fruit hooks + leaf length; Table 4). Dotted horizontal lines represent the bootstrapped (10 000 iterations) 95% confidence interval limits for Pr[*invasive*] derived from the aforementioned model based on the observed data for 3046 species. Predicted Pr[*invasive*] for each term level were calculated by adjusting the original data set so that all species were given the same value for that level (each level in turn), keeping all other terms in the model as in the original data set. Error bars represent the 10 000-iteration bootstrapped upper 95% confidence limits. The percent deviance explained (%DE) in Pr[*invasive*] by the univariate models for each term shown are: (a) species ≤ 1 m, 1 to ≤ 10 m or > 10 m in height (%DE = 0.1%), (b) species restricted to either 1, 2 or 3 floristic kingdoms (%DE = 3.4%), (c) species restricted to closed forest, open, or spanning both habitats (%DE = 8.7%), (d) species restricted to lowland, montane, or both altitudinal ranges (%DE = 5.8%), (e) annual or perennial species (%DE = 1.9%), and (f) trees, shrubs, climbers or herbs (%DE = 1.2%). See Appendix S1 for a detailed description of the attributes shown.

Table 5. The five most parsimonious generalized linear mixed-effects models investigating the life history correlates of extinction threat (coded 0) versus invasiveness (coded 1) for legumes (full data set, *n* = 606 species due to the removal of 'neutral' and range-restricted 'threatened' species) according to (a) the Bayesian Information Criterion (BIC) and (b) Akaike's Information Criterion corrected for small sample size (AIC_c). The five most highly BIC-ranked models accounted for > 99% of the information-theoretic weight (*w*BIC) of the total of 57 models considered (the top-ranked model accounted for *c.* 1.00 of $wAIC_c$). Terms shown are HT = height, RG = range, HBT = habitat, ALT = altitude, LC = life cycle, HB = habit. Also shown are the number of parameters (*k*), the maximum log-likelihood (*LL*), the difference in BIC and AIC_c for each model from the most parsimonious model (ΔBIC and ΔAIC_c), AIC_c weight, and the percent deviance explained (%DE) in the response variable by the model under consideration.

Model	*k*	*LL*	ΔBIC	*w*BIC	ΔAIC_c	$wAIC_c$	%DE
~RG + HBT + ALT + LC + HB	14	–74.625	0.000	0.851	251.854	< 0.001	61.54
~HT + RG + HBT + ALT + LC + HB	16	–73.612	3.630	0.139	560.491	< 0.001	62.07
Saturated	26	–63.274	9.175	0.009	0.000	> 0.999	67.40
~RG + HBT + LC + HB	12	–83.010	12.424	0.002	183.176	< 0.001	57.23
~HT + RG + HBT + LC + HB	14	–81.634	15.547	< 0.001	219.891	< 0.001	57.93

confounding influence of phylogenetic relatedness (which was controlled for statistically down to the tribal level); (iii) the chosen family spanned the entire globe (except Antarctica) and virtually all terrestrial habitats and (iv) we used linear mixed-effects modelling and multi-model inference to evaluate the relative evidence for an attribute's correlation to a species' status.

We found moderate to good correlations between the attributes supported by model ranking and threat/invasion risk (based on the percent deviance explained in the binary responses, which ranged from 15% to 68%). The upper limit of > 60% likely resulted from our choice to maximize the ecological signal in the reduced threatened/invasive comparison by deliberately contrasting the two fate extremes.

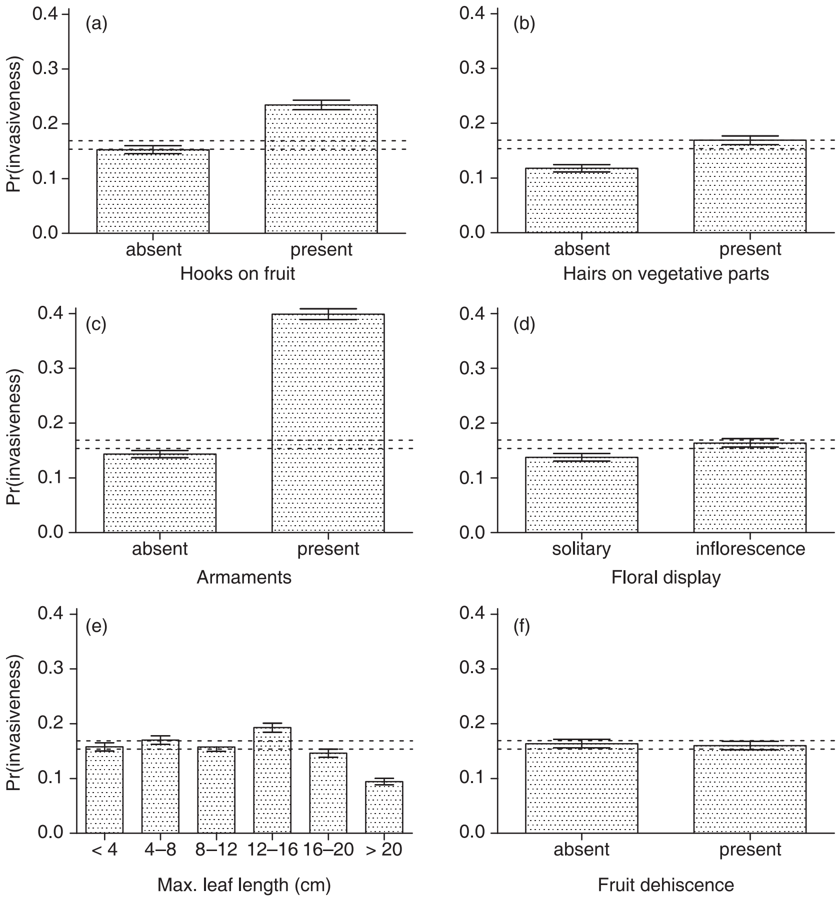

Fig. 4. Probability of a legume species being classed as 'invasive' (Pr[*invasive*]) predicted for each category of the six additional life-history terms included in the saturated model: Pr[*invasive*] ~ height + range + habitat + altitude + life cycle + habit + armaments + hairs + floral display + dehiscence + fruit hooks + leaf length; Table 4). Dotted horizontal lines represent the bootstrapped (10 000 iterations) 95% confidence interval limits for Pr[*invasive*] derived from the aforementioned model based on the observed data for 3046 species. Predicted Pr[*invasive*] for each term level were calculated by adjusting the original data set so that all species were given the same value for that level (each level in turn), keeping all other terms in the model as in the original data set. Error bars represent the 10 000-iteration bootstrapped upper 95% confidence limits. The percent deviance explained (%DE) in Pr[*invasive*] by the univariate models for each term shown are: (a) species with or without hooks on fruit (%DE = 1.6%), (b) species with or without hairs on vegetative parts (%DE = 0.8%), (c) species possessing armaments such as spines and prickles (%DE = 2.8%), (d) species with either solitary flowers or inflorescences (%DE = 0.1%), (e) maximum leaf length category (%DE = 0.4%), and (f) species with or without dehiscing fruit (%DE = 0.03%). See Supplementary Appendix S1 for a detailed description of the attributes shown.

This highlights the importance of contrasting the extremes of species' fates together, especially considering that other studies have generally only found weak correlations between threat/invasion risk and life-history/ecological traits (Jeschke & Strayer 2006; Sodhi *et al.* 2008; see also review by Cadotte *et al.* 2006).

As predicted (Heger & Trepl 2003), the extent of a species' native range, expressed in terms of its geographic, altitudinal and habitat tolerance, explains a large component of the variation in its propensity to become invasive or threatened, even after removing threatened species that had been designated so based on reduced range size (thus removing a large part of any circularity that may have existed in our conclusions). Indeed, geographic range is an important predictor of invasive bird success (Croci *et al.* 2007). This has been rationalized on the basis that widespread species tend to have a higher capacity to tolerate new environments given that they have already encountered a variety of climatic and habitat conditions in their evolutionary history and acquired a relatively higher phenotypic plasticity (Sax & Brown 2000; Croci *et al.* 2007). Widely dispersed species may also harbour greater genotypic variation such that new founder populations may have sufficient genetic variability to adapt to local conditions (Williamson 1996; Tsutsui & Case 2001). Our hypothesis that altitudinally restricted species are more susceptible to becoming threatened and less likely to be invasive was supported by both the single- and two-fate analyses. This corroborates previous work demonstrating that higher altitudes tend to have fewer invasive species (Pauchard & Alaback 2004; Becker *et al.* 2005) and that rare plant species have lower maximum elevations than common species (Hegde & Ellstrand 1999).

Recently, Sodhi *et al.* (2008) demonstrated that angiosperms that have recently gone extinct in tropical Singapore are restricted to inland forest habitats, and Pocock *et al.* (2006) found that the area of occurrence of rare and scarce British plants was predicted by habitat association. These observations agree partially with our results that closed-forest legumes had a higher risk of being classed as threatened (Tables 2 and 3) and a lower likelihood of being classed as invasive (Table 4). This finding also supports the observation of Sutherland (2004) that weeds are more likely to be shade intolerant. Although we found evidence that shorter species had a reduced threat risk (Fig. 2), the influence of plant height on invasiveness was weak (Fig. 3). Taller species of many taxa are known to be more at risk of threat (Duncan & Young 2000; Cardillo *et al.* 2005; Sodhi *et al.* 2008) because smaller species tend to have faster vital rates to offset deterministic drivers of decline, and there is also some evidence that smaller

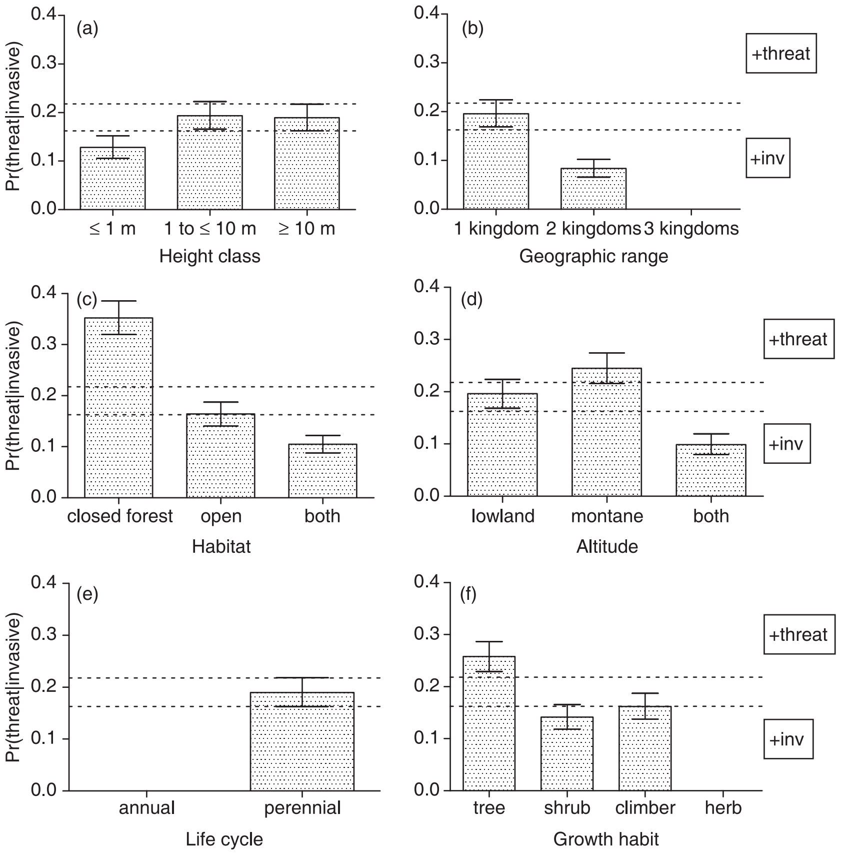

Fig. 5. Probability of a legume species being classed either as 'threatened' or 'invasive' (Pr[*threat*|*invasive*]) predicted for each category of term included in the most parsimonious generalized linear mixed-effects models (Pr[*threat*|*invasive*] ~ *height* +*range* + *habitat* + *altitude* + *life cycle* + *habit*; Table 5) using the reduced data base (i.e. after removing *threatened* species classed as such due to restricted ranges – see Methods). Dotted horizontal lines represent the bootstrapped (10 000 iterations) 95% confidence interval limits for Pr[*threat*|*invasive*] derived from the aforementioned model based on the observed data for 606 species. Predicted Pr[*threat*|*invasive*] for each term level were calculated by adjusting the original data set so that all species were given the same value for that level (each level in turn), keeping all other terms in the model as in the original data set. Error bars represent the 10 000-iteration bootstrapped upper 95% confidence limits. The percent deviance explained (%DE) in Pr[*threat*|*invasive*] by the univariate models for each term shown are: (a) species ≤ 1 m, 1 to ≤ 10 m or > 10 m in height (%DE = 18.6%), (b) species restricted to 1, 2 or 3 floristic kingdoms (%DE = 7.5%), (c) species restricted to closed forest, open, or spanning both habitats (%DE = 30.0%), (d) species restricted to lowland, montane, or both altitudinal ranges (%DE = 12.5%), (e) annual or perennial species (%DE = 5.0%), and (f) trees, shrubs, climbers or herbs (%DE = 32.3%). See Supplementary Appendix S1 for a detailed description of the attributes shown.

species are better invaders (Sutherland 2004; Ruesink 2005), although there are many regional and taxonomic exceptions (Crawley, Harvey & Purvis 1996; Williamson & Fitter 1996). Our analysis of legumes generally supports the trends that larger species are more prone to being threatened (Fig. 1), although it was the coarser size measure, growth habit (trees, shrubs, climbers or herbs), that was more diagnostic overall (after accounting for taxonomy as a random effect). This agrees with the conclusion of Sutherland (2004) that life form rather than generation time (as inferred from plant size) is the more important determinant of a species' proneness to invasion.

The effects of the term life cycle (annual versus perennial) were equivocal partially owing to the absence of threatened annuals in the reduced data set (Figs 2e and 5e) which did not permit useful inference regarding the direction of the effect. Although weedy (invasive) species tend more often to be annuals in the USA (Sutherland 2004), Wu *et al.* (2003) reported that the majority of naturalized legumes in Taiwan were herbaceous perennials. Other studies have failed to find an effect of life cycle on a plant's probability of becoming invasive (Lonsdale 1994; Goodwin *et al.* 1999). The unbalanced samples in our study and the conflicting reports in the literature lead us to hypothesize that the type of life cycle a particular legume species evolves may only affect its fate within particular environmental contexts. Larger samples and evidence from other diverse plant families may eventually provide better insight.

Another attribute commonly associated with invasiveness and plant rarity is seed size (Rejmánek & Richardson 1996; Pocock *et al.* 2006). Species with small seeds may be expected to have better propagule dispersal capacity and so may be better invaders (Rejmánek & Richardson 1996; Pocock *et al.* 2006), whereas larger seeds may require animal dispersal (Pocock *et al.* 2006) such that habitat loss may confer an additional disadvantage to species relying on dispersers that are themselves threatened with extinction (Brook, Sodhi & Bradshaw 2008). However, we found no evidence to support a correlation between this trait and either fate, possibly owing to the fecundity trade-off between seed size and seed number or dispersal mode (Eriksson & Jakobsson 1998; Thompson *et al.* 2002; Pocock *et al.* 2006), and that we only examined a single spatial (global) scale (Hamilton *et al.* 2005). Species invading a wide variety of habitats are potentially drawn from a wide selection of ecological strategies. Given that seed size is a good indicator of plant's ecological strategy, the lack of evidence for a relationship between invasiveness and seed size may result partially from such a large sample of species with different strategies (i.e. regional trends may have been swamped by the large number of strategies sampled – Crawley *et al.* 1996; Pyšek & Richardson 2007). Another potentially confounding influence is the role of vertebrate dispersal of seeds that can mask the effect of small seed size (Richardson *et al.* 2000). Finally, attributes relating to propagule dispersal (e.g. hooks on fruit) and defence from herbivory (e.g. hairs and spines) were also related to an increased probability of being invasive as expected (Sutherland 2004), thus confirming the finding in the opposite direction that species with poor fruit dispersal capacity are more extinction-prone (Duncan & Young 2000).

Many studies have examined which ecological and life-history attributes influence the probability that a plant species becomes extinct, is classed as threatened, or is rare. However, the variance in extinction risk explained by these correlates is typically low, suggesting that innate traits themselves explain the probability of population decline better than those eventually causing extinction given the decoupling of these processes (Caughley 1994; Brook *et al.* 2006; Pocock *et al.* 2006; Brook *et al.* 2008). Rare plants (which may not necessarily have a higher risk of extinction than common species), also tend to demonstrate certain ecological and life-history attributes. Rare mosses and vascular plants tend to depend more on vegetative reproduction, are more likely to be monoecious, and have more bilaterally symmetrical flowers (see review by Kunin & Gaston 1993). Although we were unable to test for the influence of clonality and sexual system on a legume species' probability of being classed as threatened, the accumulated evidence from other taxa suggests that these traits are likely to contribute.

Additional traits such as clonal capacity (Alpert *et al.* 2000; Duncan & Young 2000; Pocock *et al.* 2006; Sodhi *et al.* 2008), growth rate (Grotkopp, Rejmánek & Rost 2002), pollination system (Pocock *et al.* 2006; Sodhi *et al.* 2008), sexual system (Sodhi *et al.* 2008), and relative abundance (Duncan & Young 2000; Pocock *et al.* 2006), are all known or hypothesized to correlate with threat, extinction or invasion risk (Rejmánek 2000; Pyšek & Richardson 2007), but could not be tested here owing to lack of family-wide data (see Supplementary methods in Appendix S1). Similarly, invasiveness and threat risk in some plant groups are influenced by the proportion of the life cycle devoted to reproduction (Brändle *et al.* 2003), fecundity (Burns 2006), duration of the juvenile period (Rejmánek & Richardson 1996), shade tolerance (Sutherland 2004), toxicity (Sutherland 2004), capacity for vegetative reproduction (Burns 2006; Pocock *et al.* 2006), and specific leaf area (Grotkopp *et al.* 2002; Hamilton *et al.* 2005) (a proxy for photosynthetic capacity, leaf nitrogen content, leaf life span and relative growth rate; Reich, Walters & Ellsworth 1997). In legumes, the specificity of rhizobia in nitrogen-fixing species may also be relevant – such species are unlikely to become invasive if their specific rhizobium symbionts are absent during the transport and introduction of seeds (Bala & Giller 2001). Given the unavoidable omission of many traits, and any species-specific or stochastic events, the amount of variation in legume threat and invasion probability described by a few key attributes in our study is remarkably high.

Part of the unexplainable variation in risk may also arise because a species' ultimate fate depends on many interacting factors (Lockwood *et al.* 2005; Brook *et al.* 2006). For example, for a species to become a successful invader, it must be transported outside of its native range in sufficient number to avoid chance demographic or genetic failure, become established, and spread (Williamson 1996). As such, an invasive species' immediate history with respect to propagule pressure, its affiliation with humans, its probability of being transported outside its native range, and its phenotypic plasticity (Hamilton *et al.* 2005; Lockwood *et al.* 2005; Jeschke & Strayer 2006) strongly dictate its opportunity to become invasive (Rejmánek 2000; Pyšek & Richardson 2007). Indeed, Chong *et al.* (2006) determined that attributes making species more or less prone to invasion were generally swamped by history and environmental context. Another possible aspect complicating the relationships we found is that particular ecological and life-history attributes may themselves influence the probability of a plant species' fate being modified by humans (Sakai *et al.* 2001). For example, lowland or altitudinal generalist species may be more invasive than montane species not only because the former tend to grow and reproduce more rapidly, but also because they may have a greater probability of being collected and transported to new locations by humans living in similar environments. Thus, higher propagule pressure may account more for the observed higher invasion risk in some species rather than an innate evolved trait *per se*.

One other potential bias that may contribute to the low explanatory power of analyses such as ours is that there is an overrepresentation of woody Fabaceae species in the IUCN Red List (Lughadha *et al.* 2005; Supplementary Fig. S1f). However, this bias is potentially counteracted by the way assessments are prepared; formal assessments are generally made, primarily or exclusively, for species that are considered likely to meet the threatened status criteria, which is reflected by the high proportion of threatened species relative to the total number of species evaluated (Lughadha *et al.* 2005). For instance, 593 out of 770 extant legume species evaluated are categorized in the CR, EN and VU categories. Therefore, the high probability of woody species being assessed is also equally likely to be owing to their higher endangerment status when compared to species of other growth habits. Nonetheless, future assessments of legume species threat status should consider describing more of the smaller Fabaceae to evaluate this potential bias.

Conclusion

Our study is the first to examine explicitly and quantitatively the evolved characteristics that correlate with a higher likelihood of species from a single taxonomic group falling into one of the two fate categories examined: 'threatened' and thus having a higher likelihood of going extinct, or 'invasive' by responding positively to human land-use modifications. Although some biases are plausible and uncertainties remain, we are confident our results are broadly applicable to the Fabaceae family in general. Although our results are not necessarily applicable to all plant taxa (Pyšek & Richardson 2007), the speciose, diverse and important Fabaceae family we investigated certainly provides strong inference (especially where our results corroborate previous findings for other taxa – Pyšek & Richardson 2007).

Conservation managers will therefore benefit from the insight our results provide by being able to rank certain plant species according to their risk of becoming threatened. When land-use changes are imminent, poorly documented species can essentially be ranked according to those ecological and life-history traits that predispose them to respond negatively to habitat modification. Here, species inventories combined with known or expected life-history information (e.g. from congeners) can identify which species may require particular conservation attention such as specific habitat protection measures or artificial rearing. For example, tree and shrub species of the genus *Abarema* endemic to the Neotropics occupy a wide range of habitats including tropical lowland terra firme, montane humid forest and coastal shrublands. Of the 21 *Abarema* species that have been assessed in the IUCN Red List; only one (*A. commutata*) is listed as *Least Concern*. Based on our results, the assessment of tall, lowland, closed-forest *Abarema* tree species (e.g. *A. adenophora*, *A. asplenifolia*, *A. floribunda*, *A. langsdorfii*, *A. macradenia*, *A. mataybifolia*, *A. piresii* and *A. ricoae*) should be prioritized because these species are potentially at higher risk of becoming threatened. The same approach can be used by land managers to rank introduced plant species for their probability of spreading beyond the point of introduction and threatening native ecosystems, and to prioritize management interventions (Goodwin *et al.* 1999; Rejmánek 2000; Kolar & Lodge 2001). Our results are particularly valuable for situations of sustained habitat loss and fragmentation, especially given predictions that global climate change will simultaneously favour the establishment and proliferation of invasive non-indigenous species (Dukes & Mooney 1999) and lead to greater background extinction rates (Pimm 2001; Brook *et al.* 2008).

Acknowledgements

Authors thank L. Prior for helpful suggestions. Three anonymous referees provided helpful comments to improve the manuscript. Funding came from the National University of Singapore (R-154-000–264-112). NSS, HTWT, XG and CJAB conceived the work, XG compiled the data, CJAB did the analysis and took the lead in writing the paper, and all authors contributed to the writing of the paper.

References

Alpert, P., Bone, E. & Holzapfel, C. (2000) Invasiveness, invasibility and the role of environmental stress in the spread of non-native plants. *Perspectives in Plant Ecology, Evolution and Systematics*, **3**, 52–66.

Angiosperm Phylogeny Group (2003) An update of the Angiosperm Phylogeny Group classification of the orders and families of flowering plants: APG II. *Botanical Journal of the Linnean Society*, **141**, 399–436.

Bala, A. & Giller, K.E. (2001) Symbiotic specificity of tropical tree rhizobia for host legumes. *New Phytologist*, **149**, 495–507.

Becker, T., Dietz, H., Billeter, R., Buschmann, H. & Edwards, P.J. (2005) Altitudinal distribution of alien plant species in the Swiss Alps. *Perspectives in Plant Ecology, Evolution and Systematics*, **7**, 173–183.

Blackburn, T.M. & Cassey, P. (2004) Are introduced and re-introduced species comparable? A case study of birds. *Animal Conservation*, **7**, 427–433.

Blackburn, T.M. & Duncan, R.P. (2001a) Determinants of establishment success in introduced birds. *Nature*, **414**, 195–197.

Blackburn, T.M. & Duncan, R.P. (2001b) Establishment patterns of exotic birds are constrained by non-random patterns in introduction. *Journal of Biogeography*, **28**, 927–939.

Brändle, M., Stadler, J., Klotz, S. & Brandl, R. (2003) Distributional range size of weedy plant species is correlated to germination patterns. *Ecology*, **84**, 136–144.

Brook, B.W., Sodhi, N.S. & Bradshaw, C.J.A. (2008) Synergies among extinction drivers under global change. *Trends in Ecology and Evolution*, in press.

Brook, B.W., Traill, L.W. & Bradshaw, C.J.A. (2006) Minimum viable population size and global extinction risk are unrelated. *Ecology Letters*, **9**, 375–382.

Buckley, Y.M., Briese, D.T. & Rees, M. (2003) Demography and management of the invasive plant species *Hypericum perforatum*. II. Construction and use of an individual-based model to predict population dynamics and the effects of management strategies. *Journal of Applied Ecology*, **40**, 494–507.

Burnham, K.P. & Anderson, D.R. (2002) *Model Selection and Multimodal Inference: A Practical Information-Theoretic Approach*, 2nd edn. Springer-Verlag, New York, USA.

Burnham, K.P. & Anderson, D.R. (2004) Understanding AIC and BIC in model selection. *Sociological Methods and Research*, **33**, 261–304.

Burns, J.H. (2006) Relatedness and environment affect traits associated with invasive and noninvasive introduced commelinaceae. *Ecological Applications*, **16**, 1367–1376.

Cadotte, M.W., Murray, B.R. & Lovett-Doust, J. (2006) Ecological patterns and biological invasions: using regional species inventories in macroecology. *Biological Invasions*, **8**, 809–821.

Cardillo, M., Mace, G.M., Jones, K.E., Bielby, J., Bininda-Emonds, O.R.P., Sechrest, W. & Orme, C.D.L. (2005) Multiple causes of high extinction risk in large mammal species. *Science*, **309**, 1239–1241.

Caughley, G. (1994) Directions in conservation biology. *Journal of Animal Ecology*, **63**, 215–244.

Chong, G.W., Otsuki, Y., Stohlgren, T.J., Guenther, D., Evangelista, P., Villa, C. & Waters, A. (2006) Evaluating plant invasions from both habitat and species perspectives. *Western North American Naturalist*, **66**, 92–105.

Crawley, M.J., Harvey, P.H. & Purvis, A. (1996) Comparative ecology of the native and alien floras of the British Isles. *Philosophical Transactions of the Royal Society B*, **351**, 1251–1259.

Croci, S., Le Quilliec, P. & Clergeau, P. (2007) Geographical range as predictor of spatial expansion of invading birds. *Biodiversity and Conservation*, **16**, 511–524.

Dukes, J.S. & Mooney, H.A. (1999) Does global change increase the success of biological invaders? *Trends in Ecology and Evolution*, **14**, 135–139.

Duncan, R.P. & Young, J.R. (2000) Determinants of plant extinction and rarity 145 years after European settlement of Auckland, New Zealand. *Ecology*, **81**, 3048–3061.

Eriksson, O. & Jakobsson, A. (1998) Abundance, distribution and life histories of grassland plants: a comparative study of 81 species. *Journal of Ecology*, **86**, 922–933.

Felsenstein, J. (1985) Phylogenies and the comparative method. *American Naturalist*, **70**, 115.

Gerlach, J.D. & Rice, K.J. (2003) Testing life history correlates of invasiveness using congeneric plant species. *Ecological Applications*, **13**, 167–179.

Goodwin, B.J., McAllister, A.J. & Fahrig, L. (1999) Predicting invasiveness of plant species based on biological information. *Conservation Biology*, **13**, 422–426.

Grotkopp, E., Rejmánek, M. & Rost, T.L. (2002) Toward a causal explanation of plant invasiveness: seedling growth and life-history strategies of 29 pine (*Pinus*) species. *American Naturalist*, **159**, 396–419.

Hamilton, M.A., Murray, B.R., Cadotte, M.W., Hose, G.C., Baker, A.C., Harris, C.J. & Licari, D. (2005) Life-history correlates of plant invasiveness at regional and continental scales. *Ecology Letters*, **8**, 1066–1074.

Harborne, J.B. (1994) Phytochemistry of the Leguminosae. In *Phytochemical Dictionary of the Leguminosae* Vol. 1 (ed. F.A. Bisby, J. Buckingham& J.B. Harborne), pp. xx–xxiii. Chapman and Hall, London.

Hegde, S.G. & Ellstrand, N.C. (1999) Life history differences between rare and common flowering plant species in California and the British Isles. *International Journal of Plant Sciences*, **160**, 1083–1091.

Heger, T. & Trepl, L. (2003) Predicting biological invasions. *Biological Invasions*, **5**, 313–321.

Jeschke, J.M. & Strayer, D.L. (2006) Determinants of vertebrate invasion success in Europe and North America. *Global Change Biology*, **12**, 1608–1619.

Kolar, C.S. & Lodge, D.M. (2001) Progress in invasion biology: predicting invaders. *Trends in Ecology and Evolution*, **16**, 199–204.

Kunin, W.E. & Gaston, K.J. (1993) The biology of rarity: patterns, causes and consequences. *Trends in Ecology and Evolution*, **8**, 298–301.

Kunin, W.E. & Gaston, K.J., eds (1997) *The Biology of Rarity: Causes and Consequences of Rare-Common Differences*. Chapman and Hall, London.

Lake, J.C. & Leishman, M.R. (2004) Invasion success of exotic in natural ecosystems: the role of disturbance, plant attributes and freedom from herbivores. *Biological Conservation*, **117**, 215–226.

Lewis, G., Schrire, B., Mackinder, B. & Lock, M. (2005) *Legumes of the World*. Royal Botanic Gardens, Kew, United Kingdom.

Link, W.A. & Barker, R.J. (2006) Model weights and the foundations of multimodel inference. *Ecology*, **87**, 2626–2635.

Lloret, F., Medail, F., Brundu, G., Camarda, I., Moragues, E., Rita, J., Lambdon, P. & Hulme, P.E. (2005) Species attributes and invasion success by alien plants on Mediterranean islands. *Journal of Ecology*, **93**, 512–520.

Lockwood, J.L., Cassey, P. & Blackburn, T.M. (2005) The role of propagule pressure in explaining species invasions. *Trends in Ecology and Evolution*, **20**, 223–228.

Lonsdale, W.M. (1994) Inviting trouble: introduced pasture species in northern Australia. *Australian Journal of Ecology*, **19**, 345–354.

Lughadha, E.N., Baillie, J., Barthlott, W., Brummitt, N.A., Cheek, M.R., Farjon, A. *et al*. (2005) Measuring the fate of plant diversity: towards a foundation for future monitoring and opportunities for urgent action. *Philosophical Transactions of the Royal Society B*, **360**, 359–372.

McIntyre, S., Lavorel, S. & Tremont, R.M. (1995) Plant life-history attributes: their relationship to disturbance response in herbaceous vegetation. *Journal of Ecology*, **83**, 31–44.

Pauchard, A. & Alaback, P.B. (2004) Influence of elevation, land use, and landscape context on patterns of alien plant invasions along roadsides in protected areas of south-central Chile. *Conservation Biology*, **18**, 238–248.

Pimm, S. (2001) *The World According To Pimm: A Scientist Audits the Earth*. McGraw-Hill, New York.

Pimm, S., Raven, P., Peterson, A., Sekercioglu, C.H. & Ehrlich, P.R. (2006) Human impacts on the rates of recent, present, and future bird extinctions. *Proceedings of the National Academy of Sciences of the USA*, **103**, 10941–10946.

Pocock, M.J.O., Hartley, S., Telfer, M.G., Preston, C.D. & Kunin, W.E. (2006) Ecological correlates of range structure in rare and scarce British Plants. *Journal of Ecology*, **94**, 581–596.

Purvis, A., Gittleman, J.L., Cowlishaw, G. & Mace, G.M. (2000) Predicting extinction risk in declining species. *Proceedings of the Royal Society B*, **267**, 1947–1952.

Pyšek, P. & Richardson, D.M. (2007) Traits associated with invasiveness in alien plants: where do we stand? *Biological Invasions. Ecological Studies 193* (ed. W. Nentwig), pp. 97–126. Springer-Verlag, Berlin.

Pyšek, P., Prach, K. & Smilauer, P. (1995) Relating invasion success to plant traits: an analysis of the Czech alien flora. *Plant Invasions – General Aspects and Special Problems* (eds P. Pyšek, K. Prach, M. Rejmánek & M. Wade), pp. 39–60. SPB Academic Publishing, Amsterdam.

Pyšek, P., Richardson, D.M. & Williamson, M. (2004) Predicting and explaining plant invasions through analysis of source area floras: some critical considerations. *Diversity and Distributions*, **10**, 179–187.

R Development Core Team (2004) *R: A language and environment for statistical computing*. R Foundation for Statistical Computing, Vienna, Austria.

Randall, R.P. (2002) *A Global Compendium of Weeds*. RG and FJ Richardson, Melbourne, Australia.

Reich, P.B., Walters, M.B. & Ellsworth, D.S. (1997) From tropics to tundra: global convergence in plant functioning. *Proceedings of the National Academy of Sciences of the USA*, **94**, 13730–13734.

Rejmánek, M. (2000) Invasive plants: approaches and predictions. *Austral Ecology*, **25**, 497–506.

Rejmánek, M. & Richardson, D.M. (1996) What attributes make some plant species more invasive? *Ecology*, **77**, 1655–1661.

Richardson, D.M., Allsop, N., D'Antonio, C.M., Milton, S.J. & Rejmánek, M. (2000) Plant invasions – the role of mutualisms. *Biological Reviews*, **75**, 65–93.

Ruesink, J.L. (2005) Global analysis of factors affecting the outcome of freshwater fish introductions. *Conservation Biology*, **19**, 1883–1893.

Sakai, A.K., Allendorf, F.W., Holt, J.S., Lodge, D.M., Molofsky, J., With, K.A. *et al*. (2001) The population biology of invasive species. *Annual Review of Ecology and Systematics*, **32**, 305–332.

Sax, D.F. & Brown, J.H. (2000) The paradox of invasion. *Global Ecology and Biogeography*, **9**, 363–371.

Sodhi, N.S., Koh, L.P., Peh, K.S.-H., Tan, H.T.W., Chazdon, R.L., Corlett, R.T., Lee, T.M., Colwell, R.K., Brook, B.W., Sekercioglu, C.H. & Bradshaw, C.J.A. (2008) Correlates of extinction proneness in tropical angiosperms. *Diversity and Distributions*, **14**, 1–10.

Stohlgren, T.J., Binkley, D., Chong, G.W., Kalkhan, M.A., Schell, L.D., Bull, K.A., Otsuki, Y., Newman, G., Bashkin, M. & Son, Y. (1999) Exotic plant species invade hot spots of native plant diversity. *Ecological Monographs*, **69**, 25–46.

Stohlgren, T.J., Chong, G.W., Schell, L.D., Rimar, K.A., Otsuki, Y., Lee, M., Kalkhan, M.A. & Villa, C.A. (2002) Assessing vulnerability to invasion by nonnative plant species at multiple spatial scales. *Environmental Management*, **29**, 566–577.

Sutherland, S. (2004) What makes a weed a weed: life history traits of native and exotic plants in the USA. *Oecologia*, **141**, 24–39.

Takhtajan, A. (1986) *Floristic Regions of The World*. University of California Press, Berkeley.

Thompson, K. (1994) Predicting the fate of temperate species in response to human disturbance and global change. *Biodiversity, Temperate Ecosystems, and Global Change* (eds T.J.B. Boyle & C.E.B. Boyole), pp. 61–76. Springer-Verlag, Berlin.

Thompson, K., Rickard, L.C., Hodkinson, D.J. & Rees, M. (2002) Seed dispersal: the search for trade-offs. *Dispersal Ecology* (eds J.M. Bullock, R.E. Kenward & R.S. Hails), pp. 152–172. Blackwell Science, Oxford.

Tsutsui, N.D. & Case, T.J. (2001) Population genetics and colony structure of the Argentine ant (*Linepithema humile*) in its native and introduced ranges. *Evolution*, **55**, 976–985.

Turner, I.M., Tan, H.T.W., Wee, Y.C., Ibrahim, A.B., Chew, P.T. & Corlett, R.T. (1994) A study of plant species – extinction in Singapore – lessons for the conservation of tropical biodiversity. *Conservation Biology*, **8**, 705–712.

Walker, K. & Preston, C. (2006) Ecological predictors of extinction risk in the flora of lowland England, UK. *Biodiversity and Conservation*, **15**, 1913–1942.

Weber, E. (2003) *Invasive Plant Species of the World: A Reference Guide to Environmental Weeds*. CABI Publishing, Wallingford, UK.

Westoby, M. (1998) A leaf-height-seed (LHS) plant ecology strategy scheme. *Plant and Soil*, **199**, 213–227.

Williamson, M. (1996) *Biological Invasions*. Chapman & Hall, London.

Williamson, M.H. (1999) Invasions. *Ecography*, **22**, 5–12.

Williamson, M.H. & Fitter, A. (1996) The characters of successful invaders. *Biological Conservation*, **78**, 163–170.

Wu, S.-H., Chaw, S.-M. & Rejmánek, M. (2003) Naturalized Fabaceae (Leguminosae) species in Taiwan: the first approximation. *Botanical Bulletin of Academia Sinica*, **44**, 59–66.

Received 3 October 2007; accepted 30 April 2008
Handling Editor: Peter Bellingham

Supplementary material

The following supplementary material is available for this article:

Appendix S1. Species filtering methods and details of the ecological and life-history traits used to predict threat and invasion risk among legume species.

Table S1 List of subspecies removed and retained in the data set.

Table S2 Spearman's ρ and Kendall's τ rank correlation coefficients among species traits (factors) considered.

Table S3 A set of generalized linear mixed-effects models used to examine the correlation between a legume species' ecological and life-history attributes and its threat proneness or invasiveness.

Table S4 Results for the generalized linear mixed-effects models used to examine the correlation between a legume species' ecological and life-history attributes and its threat proneness as assessed under the NatureServe data base of rare and endangered species for Canada and the USA.

Table S5 Results for the generalized linear mixed-effects models used to examine the correlation between a legume species' ecological and life-history traits and its threat proneness with the traits *height* and *maximum length of leaf lamina* treated as continuous covariates.

Table S6 Results for the generalized linear mixed-effects models used to examine the correlation between a legume species' ecological and life-history attributes and its threat proneness with the taxonomic control of the random effect set as Tribe nested within nominal phylogenetic group (NPG – see Table 1 and Supplementary Appendix S1, Fig. S3)

Table S7 Results for the univariate generalized linear mixed-effects models examining the probability of a legume species being classed as either 'threatened' (*Critically Endangered*, *Endangered* or *Vulnerable*) or *Near Threatened* (IUCN Red List categories)

Figure S1 Proportional distribution of 8906 species sampled compared to the global distribution of described Fabaceae species. The proportional distributions between described and IUCN Red-Listed species are also shown.

Figure S2 Proportion of threatened, neutral and invasive Fabaceae species per major trait level and among six floristic kingdoms.

Figure S3 Phylogenetic tree of subfamily and tribal classifications of sampled and catalogued species in Family Fabaceae.

Appendix S2 Lists of Floras, Manuals and publications consulted.

Appendix S3 List of all Fabaceae species considered in the study.

This material is available as part of the online article from:
http://www.blackwell-synergy.com/doi/full/10.1111/j.1365-2745.2008.01408.x
(This link will take you to the article abstract).