

## Ecological patterns and biological invasions: using regional species inventories in macroecology

Marc W. Cadotte<sup>1\*</sup>, Brad R. Murray<sup>2</sup> & Jon Lovett-Doust<sup>3</sup>

<sup>1</sup>Complex Systems Group, Ecology and Evolutionary Biology, University of Tennessee, 569, Dabney Hall, Knoxville, TN 37996-0001, USA; <sup>2</sup>Institute for Water and Environmental Resource Management, University of Technology Sydney, Gore Hill, NSW 2065, Australia; <sup>3</sup>Department of Biological Sciences, University of Windsor, Windsor, ON N9B 3P4, Canada; \*Author for correspondence (e-mail: mcadotte@utk.edu)

Received 13 January 2004; accepted in revised form 16 March 2005

**Key words:** comparative studies, database analyses, ecological generalities, macroecology, plant invasions, phylogenetic analyses, predictability

### Abstract

Macroecology depends heavily on a comparative methodology in order to identify large-scale patterns and to test alternative hypotheses that might generate such patterns. With the advent and accessibility of large electronic databases of species and their life history and ecological attributes, ecologists have begun seeking generalities, and examining large-scale ecological hypotheses involving core themes of range, abundance and diversity. For example, combinations of ecological, life history and phylogenetic data have been analysed using large species sets to test hypotheses in invasion biology. Analysis of regional species inventories can contribute cogently to our understanding of invasions. Here we examine several ways in which database analysis is effective. We review 19 studies of comparative invasions biology, each using >100 species of plants in their analyses, and show that invader success is linked to seven correlates: short life cycle, abiotic (mostly wind) dispersal, large native range size, non-random taxonomic patterns (emphasizing certain families or orders), presence of clonal organs, occupying disturbed habitats, and earlier time of introduction. These phylogenetically influenced, comparative analyses using regional species inventories are only just beginning and have much potential.

“Given that we must renounce our quest for case-by-case predictability, what should we do? There is pattern in the data we possess on invasions. The next efforts in the study of invasions should be self-consciously statistical... But for this we will need the raw data on-line in computer databases accessible to all researchers” (Gilpin 1990).

### Introduction

A. S. Watt’s conclusion in his presidential address to the British Ecological Society – ‘Pattern and Process in the Plant Community’ (1947) – continues to

resonate today, in the emergent field of invasions biology and in the larger emergence of macroecology. Watt first showed how a plant community may be described from these two points of view – for diagnosis and classification, and as a working mechanism. For Watt, ‘the ultimate parts of the community are the individual plants, but a description of it in terms of the characters of these units and their spatial relations to each other is impracticable at the individual level. It is, however, feasible in terms of aggregates of individuals and of species which form different kinds of patches’. Later, another B.E.S. president, John Harper (1982) suggested that it was only natural that early stages in the emergence and growth of any science should

consist of the description and ordering of the material for study. The next stage is to search for correlations and possible causes underpinning what has been described.

Harper (1982) showed how historically the field of plant ecology had been dominated by two major themes: (1) description of vegetation (i.e., assemblages of species that are treated as objects for classification or ordination; with an array of environmental features regarded, variously, either as drivers or simple correlates); and (2) autecological single-species descriptions (e.g., *Biological Flora of the British Isles*, in *Journal of Ecology*, or the *Biology of Canadian Weeds* series in *Canadian Journal of Plant Sciences*), intended to produce detailed treatments of individual species – their nomenclature, form, distribution, population ecology, response to environmental factors, etc. Again the emphasis was on the species. As Harper (1982) put it: ‘These two broad categories of description give ecologists the equivalent of the telephone directory and “Yellow Pages”, the one describing who we can find where, and the other describing who (plant or community) does what (and again where they can be found)’.

Both kinds of investigation allow predictions to be made – about where we will find particular vegetation or species, and about what sorts of behaviour, or morphology, or life history we may expect to find among the species in a particular region. Harper (1982) warned that, whereas almost inevitably ecological studies use the species as the basis for description, it remains far from clear ‘that the conservative and stable characters and the breeding isolation that may be used to define such taxa are appropriate to define ecological units – knowing as we do the wide range of ecologically different behaviours that are included within single species’. Harper suggested that questions be formulated in general as: What are the limitations in form and behaviour of organisms that account for present distributions and behaviours? Today this still seems an apt and useful guide.

If as Gilpin (1990) suggests (in the opening quotation), ecologists are unable to predict individual ecological outcomes, then the focus must turn to uncovering broader patterns of ecological generality. Since the advent of the science of ecology (e.g., Clements 1905; Warming 1909) there have been only a limited number of ecological generalities

uncovered. Conspicuous examples include the competitive exclusion principle (e.g., Grinnell 1925; Gause 1934; Hardin 1960), the relative abundance distributions of rare and common species (Preston 1948); species–area and isolation relationships (MacArthur and Wilson 1963); and patterns in neutral macroecology (Bell 2001; Hubbell 2001). The paucity of ecological generalities appears to be inversely related to the general extent of knowledge of any particular phenomenon, and the complex multiplicative and often non-linear interactive nature of many ecological processes. The difficulty is due, in part, to the absence of any standardized method for critically testing and evaluating generalities. For example, does highlighting and noting a limited number of examples of a phenomenon provide enough support for a theoretical generality? Similarly, does determination of particular variants or exceptions necessarily disprove an ecological generality? And of course, even when we think we have identified a strong ecological generality, is causation necessarily implied?

At present, the major mechanism used to test ecological generalities is critical review, and meta-analysis of previous published results. While these provide valuable insight into ecological problems, they are subjected to two potential problems: bias in the selection of the study organism and difficulties due to censored data. The first, bias in study species (see Bonnet et al. 2002), confounds many meta-analyses in ecology. For example, has the ecology of sedges been studied to the same extent as that of lycopod mosses or oak trees, or indeed of salamanders, etc.? Several reasons exist for this, including availability of funding for research on certain (charismatic, relevant, serendipitous) organisms over others, emotional attachment and ease of study. (see Bonnet et al. 2002). Organisms and whole systems may be selected precisely because they may lend themselves to clearly supporting or equally disproving a hypothesis; or because they may easily lend themselves to being manipulated to answering the important questions, or because of an *a priori* belief that these organisms or systems exhibit a certain pattern. An example of this kind of bias is seen among studies examining the effects of forest fragmentation, where there has been a significant bias toward using birds as the organisms of study (Simberloff

2000). While fragmentation studies have led to some interesting and important results (e.g., Bierregaard et al. 1992; Trzcinski et al. 1999), reviews and meta-analyses examining effects of forest fragmentation will be biased towards this unique and exceptionally mobile group of taxa.

The second potential bias, data censoring, is perhaps more insidious and difficult to control. Primary papers – which form the basis for reviews and meta-analyses – are the source of censored data. Journals (like all publications) need to sell themselves and hold a market position (involving scientific standards, reputation, first-publication record for important results, theories, etc.) and are probably biased towards the publishing of substantial results over ‘non-results’ (Murtaugh 2002). Thus papers reporting, say, an absence of any apparent differences between rare and common species, will probably tend not to occupy the pages of top-ranked ecological journals.

If organisms may be selected in a biased manner, and reports preferentially publish positive results, do reviews and meta-analyses have a role in the elucidation of ecological generalities? Of course they do, but as an intermediary step toward developing it. We see reviews and meta-analyses as combining disparate sources of information and results into workable hypotheses, which can then be more thoroughly tested. An overall framework for testing ecological generalities is suggested in Figure 1. We do not argue that the vast array of ecological complexities can necessarily be attributed to a small number of ecological generalities; rather, we see a structure by which ecologists can test the relative importance of proposed generalities, and also the importance of variation or departure from a potential generality. Recently Daehler (2001) rightly noted that a number of floras needed to be analysed in order to adequately test hypotheses of plant naturalization.

With the ability to maintain electronically and transmit large datasets, scientists can analyse complete species datasets [usually maintained by government agencies and conservation organizations (e.g., USGS invasive species, FishBase, IUCN endangered species, etc.)]. Difficulties include differing levels or extents of species inventories, ranging from a few pre-selected species to complete species lists of a region. We see four broad-

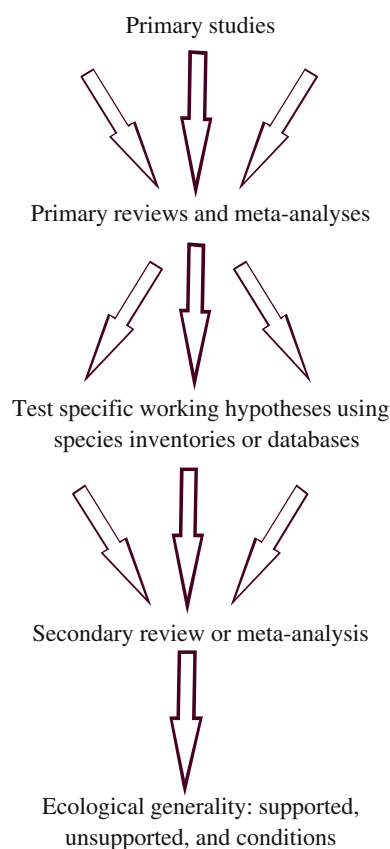


Figure 1. Framework for using primary literature, meta-analyses, and complete species inventories for testing hypotheses of general ecological patterns. Converging arrows represent the consolidation of independent analyses; diverging arrows represent the testing (via multiple independent tests) of the hypotheses, strengthened at the meta-analytic level.

ly different levels of species compilation, available for comparative analyses (Table 1). First is a complete (or near complete) inventory where there has been a concerted effort to record all extant taxa within a large (socio-politically delineated) region or jurisdiction (see also ‘Inventory’). Second is a large species database, which does not attempt to enumerate all the species within an entire region but is, rather, a complete list of species that occur in a large, representative habitat or ecosystem. Third is a species list, which is a limited subset of species, selected on some *a priori* criteria (e.g., sampled from a particular habitat type or biome, or from an individual taxonomic family). The final type is made up of pre-selected small groups of species, such as a

Table 1. Types of comparative analysis and their potential uses for analysing species inventories.

	Type of comparison	Number of species <sup>a</sup>	Potential utility	Example
Inventory	Complete species inventory of a large region	>80% of species in region	Determined in part by history, ecology, and evolution	Cadotte and Lovett-Doust 2002
Database	Large representative group of species	20–80% of species in a region or 100% from sub-regions	Likely to be phylogenetically random	Murray et al. 2002
List	Species list	<20% or 50–200 species	Not phylogenetically random, selected with prior rationale (e.g., community type)	Eriksson and Bremer 1992
Groups	Selected species	Few species, e.g., comparisons among 20 species in single genus	Not random; limited inferential space	Rabinowitz and Rapp 1981

<sup>a</sup>Number is arbitrary.

few pairs of closely-related species. These have received much of the focus thus far, given their utility in considering phylogenetic effects, and the sense of comparative value. In what follows, we emphasize analysis using the first two, less studied levels: species inventories and large species databases.

Given that variables relevant to a hypothesis are available for analysis (e.g., abundance, longevity, body size, fecundity, growth rate, etc.), then a hypothesis (suitably refined by theory and meta-analysis) can be tested, minimizing taxon-specific bias. Such analyses of a complete inventory or large species database may then be combined in order to determine the extent of dependence, the results may have on geography and taxon. The use of large species inventories and databases to test hypotheses explaining large-scale patterns in ecology and as evolutionary biology has led recently to important and diverse conclusions (macroecological studies: Kelly 1996; Kelly and Woodward 1996; Westoby et al. 1996; Pyšek 1998; Hegde and Ellstrand 1999; Cadotte and Lovett-Doust 2001, 2002; Lovett-Doust and Kuntz 2001; Khedr et al. 2002; Murray et al. 2002; Lovett-Doust et al. 2003; evolutionary studies: Eriksson and Bremer 1992; Ricklefs and Renner 1994; Dodd et al. 1999; Gaston and Blackburn 2000; among inventories: Leishman et al. 1995). Most of these examples probably coincide with the recent completion and availability of electronic inventories. Here we wish simply to highlight the importance and value of using large databases for the discovery and validation of ecological generalities. We highlight certain limitations, which bear recognition by ecologists.

### Generalities in invasions biology: much potential for database analyses

By the middle of the past century, invasions by non-native species had begun to be recognized as a significant biological phenomenon and an issue of general concern (Elton 1958; and see reviews of early invasions in Mack et al. 2000; Cadotte, in press). Since then, numerous examples of species thriving outside their native range have been described, with authors inferring 'rules' or generalities of invasiveness (e.g., Baker 1974; Rejmanek 1996). Although these synthetic attempts drew on results or examples from only a limited number of studies and taxa, and were subjected to the potential errors described above, they did shape and direct our understanding of biological invasions. Moreover there is no clear consensus concerning generalities and the capacity of biologists to predict future invasions. Although recent attempts range from relative agnosticism to sanguine confidence (examples from this spectrum, presented in a general range from confident to pessimistic, include: Pimm 1989; Scott and Panetta 1993; Rejmanek and Richardson 1996; Reichard and Hamilton 1997; Goodwin et al. 1999; Pheloung et al. 1999; Williamson 1999; Daehler and Carino 2000; Mack et al. 2000; Kolar and Lodge 2001), there remain relatively few analyses of large databases probing characteristics of invasive species (but see Table 2). In a review of 16 invasion studies, each containing at least 20 species (though definitely not complete species inventories), Kolar and Lodge (2001) showed that generalizations appeared to be emerging, at least among plants and

Table 2. Plant invasion studies using &gt;100 species.

Authors	Dataset	Type of comparison (from Table 1)	Data points	General finding
Anderson 1995	(1) 40 native species of disturbed habitats; (2) 61 aliens of disturbed habitat; and (3) 52 aliens in natural habitats in Denmark.	List	Species	Dispersal strategies seem important (e.g. wind for invaders of natural habitats).
Cadotte and Lovett-Doust 2001	1814 species found in southwestern Ontario, 484 of which are exotics.	Inventory	Species	Taxonomically non-random and exotics were short-lived with longer flowering periods, and small, wind-dispersed fruits.
Crawley et al. 1996	1515 natives and 1169 aliens in the British Isles.	Inventory	Phylogenetic	Aliens were over-represented by heavier seeds, tall height and protracted seed dormancy.
Daehler 1998	1348 'serious' agricultural weeds, 1041 widespread weeds, and 381 natural area invaders.	Inventory	Species	Agricultural weeds were over-represented by rapidly reproducing, abiotically dispersed herbs, while natural area invaders were represented by larger families.
Daehler and Carino 2000	54 invasive and 57 non-invasive non-indigenous species in Hawaii	List	Species	Using Pheloung et al. 1999 methodology they accurately identified >90% of the invasive species.
Goodwin et al. 1999	165 pairs of species native to Europe. First in pair is invasive to New Brunswick, Canada and second is not.	List	Species	Most important trait appears to be geographic range size in Europe.
Lonsdale 1994	463 intentionally introduced species in northern Australia.	Database	Species	<1% were found to be 'useful' without becoming problem species. Also, grasses were more likely than legumes to be weedy.
Pheloung et al. 1999	370 non-native species in Australia and New Zealand.	Database	Species	They created a list of factors (biogeography, ecology, reproduction, dispersal, etc.) which predicted >85% 'weeds' as being of concern.
Prinzinger et al. 2002	183 invaders in Buenos Aires province and 74 in Mendoza province, Argentina.	Database	Species and phylogenetic	Four generalities contributing to invader success: frequency in native range; abiotic niche; native range which covers multiple floristic regions; and ruderal habit.
Pyšek 1997	(1) 457 aliens in Central Europe; (2) World's most aggressive weeds ( $n = 207$ ); (3) Aliens of Auckland region ( $n = 615$ ); and (4) 64 natural area invaders in South Africa.	Multiple	Phylogenetic	Analysis of the influence of clonality; clonals appear more successful in wetter, colder climates. Clonals have dispersal disadvantage but an advantage once established.

Table 2. Continued.

Authors	Dataset	Type of comparison (from Table 1)	Data points	General finding
Pyšek 1998	26 regions around the world, with 57–975 alien species per region.	Multiple	Species <sup>a</sup>	Most effective invaders belong in: Papaveraceae, Chenopodiaceae, Amaranthaceae, Cruciferae, and Polygonaceae.
Pyšek et al. 1995	132 permanently established invasives in Czech Republic compared to native flora.	List	Species	Aliens vs. natives: Taxonomic, life form, life strategies, dispersal, and abiotic requirement differences.
Pyšek et al. 2003	668 species, non-indigenous to the Czech Republic and introduced after 1500.	Inventory	Species	Species adapted to disturbed habitats (e.g., early flowering time, annual life cycle) were more likely to have arrived earlier in history than other species.
Reichard and Hamilton 1997	235 woody plants in North America.	Database	Species	Used discriminant analysis to confirm weedy status. Traits such as species being related to an invader and rapid vegetative reproduction appear important.
Scott and Panetta 1993	242 established South African plants in Australia.	List	Species	These 'weeds' seems to have wide native range size, are weeds elsewhere, have been resident in Australia for long time, and are related to congeneric weeds.
Thébaud and Simberloff 2001	651 species in two datasets: (1) European species in California and <i>vice versa</i> ; and (2) European species in the Carolinas and <i>vice versa</i> .	Database	Species	Species not taller in introduced ranges.
Thompson et al. 1995	211 species that have expanded ranges in England, Scotland, the Republic of Ireland and the Netherlands vs. natives.	List	Species	Invasives are more likely to be clonal and polycarpic perennials.
Williamson and Fitter 1996	Established aliens vs. natives of Britain ( $n = 177$ ).	Inventory	Species	Factors important for invasives: Native range size, size of plant, and propagule pressure.

<sup>a</sup>'Data points' refers to whether studies used species as independent data points or followed some form of explicit phylogenetic analyses.

<sup>a</sup>This study used species to compare taxonomic patterns of invasions.

birds. Kolar and Lodge also showed that this type of investigation (using larger datasets) is only just commencing, with 13 of the studies occurring in the past 10 years.

### The inventory

For purposes of examining issues in invasions biology, an inventory of species may include all of the species, both native and introduced, or just the exotics, depending upon the questions being asked. A bare-bones species inventory is just that – a list of all (or nearly all) species in a given region. How big should the region be? This depends upon the scale of the question, and could limit inferential power. To test potential ecological generalities, the species list should be at a sufficient scale as to transcend the influences of local environments, and individual species' dispersal capabilities (to remove effects that transient species may have on local assemblages), and the vagaries of idiosyncratic and random (including historical) influences. Thus for example, the species inventory for the island Krakatau may be interesting in its own right, but may not be an effective example for studying invasions of intact communities. At the same time, the list cannot be so large as to transcend the regional and geographical forces affecting species distributions and abundances. Global databases (e.g., IUCN red lists) are very important in their own right, and extremely useful for conservation and evolutionary analyses (e.g., Purvis et al. 2000), but may not answer ecological questions without further regional information. For example, constructing a global list of invaders would be difficult, not only because of definitional problems (e.g., Colautti and MacIsaac 2004), but also for historical and ecological ones. Not all species have equal opportunity to expand potentially new habitats, nor are invaders going to be successful in every habitat they come across.

We prefer a large (>100,000 km<sup>2</sup>) political jurisdiction, for several reasons. First, due to the structure of government agencies and their funding lines, political regions will be a logical level of focus for the initial construction of species inventories, especially considering governmental efforts to control invasive species. Second, large

political regions are largely biogeographically arbitrary (excluding small islands). Such arbitrariness means that there is no *a priori* reason for habitat-level structuring to be a factor in the species inventory. Large political regions (e.g., some European and Central American countries, and Canadian, Australian and large US provinces/states) will often contain a number of different geological and biogeographical realms. Third, political regions are repeatable. That is, we can locate arbitrary regions of similar size, climate, species number, etc., for comparative analyses. Finally, data collection at this level will likely contain important 'on the ground' information, such as relative abundances and area of occupation. Analyses of this type of database will compare species that are likely to interact with one another.

A species list on its own is not so very interesting. However once ancillary, correlative evolutionary or ecological information is added, parsing the inventory can become particularly fruitful. Any number of biologically meaningful variables may be added, including but not limited to phylogenetic information, life history attributes, habitat associations, relative abundances, invasion status, and so on. Unfortunately, there is no central repository for databases of this sort, where researchers add variables or fill missing information (but see Peterson and Vieglais 2001), but with the blossoming of bioinformatics as a relatively well-funded area of science, database development and maintenance should continue to grow rapidly. We note the presence of several excellent repositories of large databases – USGS invasive species, FishBase, IUCN endangered species. For example, Ruesink (2003) used FishBase (<http://www.fishbase.org>) to show that the majority of non-indigenous fish had ecological effects on the invaded communities, and that these impacts were more severe in communities with high endemism and longer amount of time since invasion in the community.

### Database analyses

Large databases are amenable to many commonly used statistical analyses (e.g., Zar 1998; Sokal and Rohlf 1994), especially correlation, regression, ANOVA (including MANOVA and ANCOVA)

and their non-parametric equivalents. However, it is crucial that the data points be seen for what they are. One may probe the dataset using individual species as independent datapoints, or by using a phylogenetically-informed analysis. These two approaches can reveal quite different results (Harvey et al. 1995; Westoby et al. 1995) and the researcher should be aware of what their results mean and the potential limitations, or crucial assumptions attached to their analysis.

#### *Incorporating phylogenetic information in database analysis*

Ecological generalities in invasions biology are often sought on the premise that invasive species possess a distinctive suite of biological traits (e.g., Kolar and Lodge 2001). To test this premise, traits of invasive species have been compared with those of native species in invaded habitats (Cadotte and Lovett-Doust 2001), and with those of non-invasive species from the species pool in the source region (Prinzing et al. 2002), as well as among non-native species of differing levels of ecological success (Cadotte, Murray and Lovett-Doust, unpublished manuscript). Such comparisons are possible where trait values for each species contribute single units of information in statistical analyses (Figure 2). These cross-species comparisons are of consequence because they describe present-day relationships among species (Grafen 1989; Price 1997; Ackerly and Donoghue 1998; Westoby et al. 1998; Harvey and Rambaut 2000). Moreover value is given to the present-day functionality of the traits of each species. Cross-species analyses suggest that certain traits, such as clonal growth and hermaphroditism for instance, are functionally important in promoting invasiveness among plants (Thompson et al. 1995; Cadotte and Lovett-Doust 2001).

Cross-species comparisons have traditionally been conducted without explicit consideration of phylogenetic relatedness among species. Coupling information on the traits of extant species with phylogenetic information can detail the extent to which change in one trait (e.g., invasiveness) may be correlated with change in another (e.g., seed mass) throughout a particular phylogeny (or correlated-divergence analysis *sensu* Westoby et al. 1998).

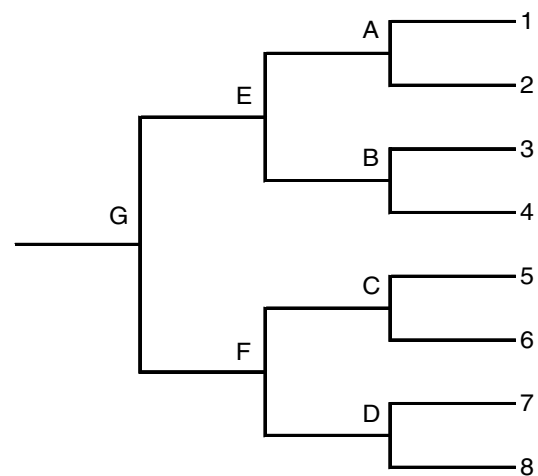


Figure 2. A hypothetical phylogeny linking eight species, numbered 1–8. In cross-species analysis, the trait values of each of the eight species are treated as independent data points ( $n = 8$ ). In a fully resolved phylogeny, there will be  $n - 1$  nodes or radiations; in this example there are seven nodes, labelled A–G. Phylogenetically-structured comparisons use values calculated at nodes as the independent data points, as these values constitute phylogenetically independent contrasts.

The importance of phylogenetically-structured comparisons is that each radiation or node in a phylogeny, rather than each species, contributes a single piece of information in the analysis (Ridley 1983; Felsenstein 1985; Grafen 1989; Harvey and Pagel 1991; see Figure 2). The database of trait values for extant species can be used to calculate values at each node in a phylogeny (a procedure that has been referred to as ‘hanging a variable on a tree’, Grafen 1989). In the simplest case, values at nodes are effectively calculated as differences in trait values between two species diverging from the same node (Felsenstein 1985; see Martins and Hansen 1996 for discussion of more complex situations). The important feature here is that values determined for nodes are independent when they describe separate radiations in a tree. Hence values at nodes can be used in place of trait values of species to perform correlated-divergence analysis in a variety of statistical techniques, including correlation, regression, multiple regression, principal components analysis, and analyses of variance and covariance (Felsenstein 1985; Garland et al. 1999).

The utility of a complementary approach in database analysis, where both cross-species and

correlated-divergence analyses are performed, is that biologically informative patterns may emerge in the comparison of the different analytical methods (Ricklefs and Starck 1996; Price 1997; Brown 1999). For instance, there might be a positive relationship between seed mass and invasiveness within genera or families. Yet this pattern could be obscured in cross-species analysis because of large differences in basic natural history between families or other major clades. The results of cross-species analysis might reveal no relationship or even a negative relationship between invasiveness and seed mass (for a good example of this see Nee et al. 1991 for the relationship between body size and abundance in birds).

A number of techniques are available for formally incorporating phylogenetic information into comparative analyses (Harvey and Pagel 1991; Martins and Hansen 1996). For two programmes that can be used to perform correlated-divergence analysis, packaged, powerful software is available as free downloads from the Internet. CAIC (Comparative Analysis by Independent Contrasts) (Purvis and Rambaut 1995) is available (at <http://www.bio.ic.ac.uk/evolve/software/caic/>), as is the phylogenetic regression (Grafen 1989) (see [users.ox.ac.uk/~grafen/phylo/](http://users.ox.ac.uk/~grafen/phylo/)), as at July 2004. For application of CAIC, see Prinzing et al. (2002), and for application of the phylogenetic regression see Murray et al. (2002). Both these programmes require a working phylogeny of the study species (see Grafen 1989). A good starting point for construction of a working phylogeny for plants is that of the Angiosperm Phylogeny Group (2003). The APG (2003) provides a classification for orders and families of flowering plants. For further phylogenetic resolution, consultation with systematists or additional literature searches are required.

### Database studies of plant invasions

There has been a recent spate of published articles using database analyses to examine patterns and test hypotheses about plant invasions. Table 2 summarizes 19 such studies. Several patterns are evident. First, there is considerable variation in database size and in the objects of comparison, with comparisons of natives vs. exotics (e.g.,

Crawley et al. 1996), both within particular groups of invaders (Daehler 1998), and between invaders and non-invaders (Goodwin et al. 1999).

Relatively few studies have employed phylogenetic information in their analyses, with most using species as independent data points (See Table 2). Many of these studies test very different hypotheses (usually because the databases contain different information). Though the database studies outlined in Table 2 are not readily amenable to examination of how widespread certain generalities are (due to differing sizes and types of information), they do show the great potential for this type of study. From Table 2, it is apparent that there may be widespread patterns emerging, though an adequate test would need more standardized analyses (full species lists and incorporating phylogenetic analyses). Seven possible generalities (patterns occurring in multiple studies) are evident, relating the success of plant invaders to: short life cycle; abiotic (mostly wind) dispersal syndrome; large native range size; non-random taxonomic patterns (emphasizing certain families or orders); presence of clonal organs; occupation of disturbed habitats; and early time since introduction.

### Issues of concern

As long as researchers continue to collect life history and population-level data (about mating systems, genetic variability, competitive interactions, etc.), the information contained in inventories will continue to grow. We believe that ecologists and agencies should strive to develop biotic inventories, which future researchers may build on (see Grime et al. 1988). As long as these datasets remain 'public', that is they are accessible to other researchers to use or add information to, these rich datasets would go a long way toward answering questions about how general and pervasive certain patterns are. Several cautions need particular mention.

#### *Data collection*

Individuals with differing degrees of training do database construction and augmentation. Workshops and meetings devoted to methods of

database construction and analyses (e.g., a special session at the invasive species conference in Beijing, June 2004) attest to the reality that researchers are thinking about ways to make large species databases more accurate. However, the analyst must be cognizant that these databases are not peer-reviewed and that the overall accuracy of the data must be assumed. Some of the aforementioned organized databases (e.g., FishBase) do have rigorous methods of data inclusion. However the methods of construction and inclusion for other databases are less clear. The Ontario plant database used by Cadotte and others (Cadotte and Lovett-Doust 2001, 2002; Cadotte, Murray and Lovett-Doust, unpublished manuscript) was compiled from local botanical lists compiled by regional municipalities and the provincial government, with more limited indication of overall rigour. Additional data (life history attributes, habitat associations, etc.) were compiled by the authors from literature searches. How valid is the assumption that the data are accurate? This can only be answered by systematic review of databases.

Further, these databases are delineated by political boundaries, and so likely do not match geographical or biogeographical regions. Again, from the Ontario plant database, many rare Ontario species are relatively common in the USA. Abundance scores recorded in this database are affected by this range truncation. How much this affected the results of Cadotte and Lovett-Doust (2002) are not immediately obvious, though an interesting approach to it would be to compare the edge-of-range rare species with those having their range centres in Ontario.

#### *Generality does not equal predictability*

There have been calls for ecologists to do more to help policy-makers arrive at informed environmental decisions (Shrader-Frechette 2001; Ehrlich 2002). This concern is a prudent response to prevent the invasions of exotic species. In an article highly critical of invasions studies, Shrader-Frechette (2001) argues that the major problem associated with predicting invasive species 'is primarily one of scientific method', and involves the caveat that 'politics and economics play a role' (p. 507). She goes on to say that this scientific inadequacy is three-pronged. First, there is no

firm definition of exotic species; second, current theory has brought no predictive power; and finally, there are no empirical generalizations. Criticism such as this (but see also Ehrlich 2002) contends that ecologists have not met the needs of the conservation community mainly because ecologists have not been able to predict the outcome of certain events, perhaps in a manner parallel to that of classical physics, or have not yet found the mechanistic understanding that, say, geneticists have, with DNA.

True universal predictability in ecology may never be fully attainable, though generalizations may be. Ecology is complex. Many phenomena are time and place-dependent and are likely to have very complex causes, influences and patterns (McMahon and Cadotte 2002). Similarly, the ability to predict precisely which species will have what kind of impact may not be possible (Williamson 1999), due to the complex nature of species interactions with other species and their environments. If we were to select a single strong predictor of invasive plant success, it would probably be showy flowers, since most plant introductions are from horticulture (Mack and Lonsdale 2001; Lovett-Doust 2003). Of course, this would have little biological meaning as to how well a species performs in a new community, competes with other species, or utilizes resources.

Ecology's goal as an applied science need not be to predict the next purple loosestrife or zebra mussel. Rather, we view ecology's applied role as one of examining patterns, discerning generalities and assessing potential consequences. Do we expect climatologists to predict when and where the next hurricane will hit land? No, the meteorological events giving rise to hurricanes are far too complex and stochastic to be precisely predicted. But what these experts can do is to identify high-risk locations and determine the phenomena that may form a hurricane. Much the same should obtain in invasions biology. The fact that ballast water of trans-oceanic ships has been the target of legislation shows that scientists are identifying high risk activities.

Database analyses, by their statistical nature, will never be able to produce a case-by-case predictive application. Rather, the best it can do is to highlight trends and generalities and produce statistical predictions, such as with discriminant

analyses (e.g., Reichard and Hamilton 1997). In understanding statistical patterns in large data sets, ecologists gain a better understanding of how species interact with other species and the abiotic environment.

### Acknowledgements

We thank Jim Carlton, Petr Pyšek, Mark Westoby, Ian Wright and an anonymous reviewer for helpful comments, Sean McMahon for useful discussions and the hurricane analogy, and Jim Drake for suggesting to write this paper.

### References

- Ackerly DD and Donoghue MJ (1998) Leaf size, sapling allometry, and Corner's Rules: phylogeny and correlated evolution in Maples (*Acer*). *American Naturalist* 152: 767–791
- Anderson UV (1995) Comparison of dispersal strategies of alien and native species in the Danish flora. In: Pyšek P, Prach K, Rejmánek M and Wade M (eds) *Plant Invasions: General Aspects and Special Problems*, pp 61–70. Academic Publishing, Amsterdam
- Angiosperm Phylogeny Group (2003) An update of the Angiosperm Phylogeny Group classification for the orders and families of flowering plants: APG II. *Botanical Journal of the Linnean Society* 141: 399–436
- Baker HG (1974) The evolution of weeds. *Annual Review of Ecology and Systematics* 5: 1–24
- Bell G (2001) Neutral macroecology. *Science* 293: 2413–2418
- Bierregaard RO, Lovejoy TE, Kapos V and Dos Santos AA (1992) The biological dynamics of tropical rainforest fragments. *Bioscience* 42: 859–866
- Bonnet X, Besnard M and Van den Driessche J (2002) Taxonomic chauvinism. *Trends in Ecology and Evolution* 17: 1–3
- Brown JH (1999) Macroecology: progress and prospect. *Oikos* 87: 3–14
- Cadotte MW (in press) Darwin to Elton: early ecology and the problem of invasive species. In: Cadotte MW, McMahon SM and Fukami T (eds) *Conceptual Ecology and Invasions Biology: Reciprocal Approaches to Nature*. Kluwer, Dordrecht
- Cadotte MW and Lovett-Doust J (2001) Ecological and taxonomic differences between native and introduced plants of southwestern Ontario. *Ecoscience* 8: 230–238
- Cadotte MW and Lovett-Doust J (2002) Ecological and taxonomic differences between rare and common plants of southwestern Ontario. *Ecoscience* 9: 397–407
- Clements FE (1905) *Research Methods in Ecology*. The University Publishing Company, Lincoln, NB
- Colautti RI and MacIsaac HJ (2004) A neutral terminology to define 'invasive' species. *Diversity and Distributions* 10: 135–141
- Crawley MJ, Harvey PH and Purvis A (1996) Comparative ecology of the native and alien floras of the British Isles. *Philosophical Transactions of the Royal Society, London, Series B* 351: 1251–1259
- Daehler CC (1998) The taxonomic distribution of invasive angiosperm plants: ecological insights and comparison to agricultural weeds. *Biological Conservation* 84: 167–180
- Daehler CC (2001) Darwin's naturalization hypothesis revisited. *American Naturalist* 158: 324–330
- Daehler CC and Carino DA (2000) Predicting invasive plants: prospects for a general screening system based on current regional models. *Biological Invasions* 2: 92–103
- Dodd ME, Silvertown J and Chase MW (1999) Phylogenetic analysis of trait evolution and species diversity variation among angiosperm families. *Evolution* 53: 732–744
- Ehrlich PR (2002) Human natures, nature conservation, and environmental ethics. *Bioscience* 52: 31–43
- Elton CS (1958) *The Ecology of Invasions by Animals and Plants*. Chapman and Hall, New York, NY
- Eriksson O and Bremer B (1992) Pollination systems, dispersal modes, life forms, and diversification rates in angiosperm families. *Evolution* 46: 258–266
- Felsenstein J (1985) Phylogenies and the comparative method. *American Naturalist* 125: 1–15
- Garland T, Midford PE and Ives AR (1999) An introduction to phylogenetically based statistical methods, with a new method for confidence intervals on ancestral values. *American Zoologist* 39: 374–388
- Gaston KJ and Blackburn TM (2000) *Patterns and Process in Macroecology*. Blackwell Science, Oxford, UK
- Gause GF (1934) *The Struggle for Existence*. Williams and Wilkins, Baltimore, MD
- Gilpin M (1990) Ecological prediction. *Science* 248: 88–89
- Goodwin BJ, McAllister AJ and Fahrig L (1999) Predicting invasiveness of plant species based on biological information. *Conservation Biology* 13: 422–426
- Grafen A (1989) The phylogenetic regression. *Philosophical Transactions of the Royal Society of London, Series B* 326: 119–157
- Grime JP, Hodgson JG and Hunt R (1988) *Comparative Plant Ecology*. Unwin Hyman, London, UK
- Grinnell J (1925) Risks incurred in the introduction of alien game birds. *Science* 61: 621–623
- Hardin G (1960) The competitive exclusion principle. *Science* 131: 1292–1297
- Harper JL (1982) After description. In: Newman EI (eds) *The Plant Community as a Working Mechanism*, pp 11–25. Blackwell Scientific Publications, Oxford, UK
- Harvey PH and Pagel MD (1991) *The Comparative Method in Evolutionary Biology*. Oxford University Press, New York, NY
- Harvey PH and Rambaut A (1998) Phylogenetic extinction rates and comparative methodology. *Proceedings of the Royal Society, London, Series B* 265: 1691–1696
- Harvey PH and Rambaut A (2000) Comparative analyses for adaptive radiations. *Philosophical Transactions of the Royal Society, London, Series B* 355: 1599–1605
- Harvey PH, Read AF and Nee S (1995) Why ecologists need to be phylogenetically challenged. *Journal of Ecology* 83: 535–536

- Hegde SG and Ellstrand NC (1999) Life history differences between rare and common flowering plant species of California and the British Isles. *International Journal of Plant Sciences* 160: 1083–1091
- Hubbell SP (2001) *The Unified Neutral Theory of Biodiversity and Biogeography*. Princeton University Press, Princeton, NJ
- Kelly CK (1996) Identifying functional types using floristic data bases: ecological correlates of plant range size. *Journal of Vegetation Science* 7: 417–424
- Kelly CK and Woodward FI (1996) Ecological correlates of plant range size: taxonomies and phylogenies in the study of plant commonness and rarity in Great Britain. *Philosophical Transactions of the Royal Society, London, Series B* 351: 1261–1269
- Khedr AH, Cadotte MW, El-Keblawy A and Lovett-Doust J (2002) Phylogenetic diversity and ecological features in the Egyptian flora. *Biodiversity and Conservation* 11: 1809–1824
- Kolar CS and Lodge DM (2001) Progress in invasion biology: predicting invaders. *Trends in Ecology and Evolutionary Biology* 16: 199–204
- Leishman MR, Westoby M and Jurado E (1995) Correlates of seed size variation: a comparison among five temperate floras. *Journal of Ecology* 83: 517–530
- Lovett-Doust J (2003) Introduction. In: Day BJ (eds) *The Garden: Myth, Meaning, and Metaphor*. Working Papers in the Humanities, 12, University of Windsor Press, Windsor, ON
- Lovett-Doust J, Biernacki M, Page R, Chan M, Natgunarajah R and Timis G (2003) Effects of land ownership and landscape-level factors on rare-species richness in natural areas of southern Ontario, Canada. *Landscape Ecology* 18: 621–633
- Lovett-Doust J and Kuntz K (2001) Land ownership and other landscape-level effects on biodiversity in southern Ontario's Niagara Escarpment Biosphere Reserve, Canada. *Landscape Ecology* 16: 743–755
- Lonsdale WM (1994) Inviting trouble: introduced pasture species in Northern Australia. *Australian Journal of Ecology* 19: 345–354
- MacArthur RH and Wilson EO (1963) An equilibrium theory of insular zoogeography. *Evolution* 17: 373–387
- Mack RN (2003) Plant naturalization and invasions in the eastern United States: 1634–1860. *Annals of the Missouri Botanical Garden* 90: 77–90
- Mack RN and Lonsdale WM (2001) Humans as global plant dispersers: are we getting more than we bargained for?. *Bioscience* 51: 95–102
- Mack RN, Simberloff D, Lonsdale WM, Evans H, Clout M and Bazzaz FA (2000) Biotic invasions: causes, epidemiology, global consequences, and control. *Ecological Applications* 10: 689–710
- Martins EP and Hansen TF (1996) The statistical analysis of interspecific data: a review and evaluation of phylogenetic comparative methods. In: Martins EP (eds) *Phylogenies and the Comparative Method in Animal Behaviour*, pp 22–75. Oxford University Press, Oxford
- McMahon S and Cadotte M (2002) Complexity and conservation: an ecological perspective. *Bioscience* 52: 460–461
- Murray BR, Thrall PH and Lepschi BJ (2002) Relating species rarity to life history in plants of eastern Australia. *Evolutionary Ecology Research* 4: 937–950
- Murtaugh PA (2002) Journal quality, effect size, and publication bias in meta-analysis. *Ecology* 83: 1162–1166
- Nee S, Read AF, Greenwood JJD and Harvey PH (1991) The relationship between abundance and body size in British birds. *Nature* 351: 312–313
- Peterson AT and Vieglais DA (2001) Predicting species invasions using ecological niche modeling: new approaches from bioinformatics attack a pressing problem. *Bioscience* 51: 363–371
- Pheloung PC, Williams PA and Halloy SR (1999) A weed risk assessment model for use as a biosecurity tool evaluating plant introductions. *Journal of Environmental Management* 57: 239–251
- Pimm SL (1989) Theories of predicting success and impact of introduced species. In: Drake JA, Mooney HA, Castri Fdi, Groves RH, Kruger FJ, Rejmanek M and Williamson M (eds) *Biological Invasions: A Global Perspective*, pp 351–367. John Wiley and Sons, New York, NY
- Preston FW (1948) The commonness and rarity of species. *Ecology* 29: 254–283
- Price T (1997) Correlated evolution and independent contrasts. *Philosophical Transactions of the Royal Society, London, Series B* 352: 519–529
- Prinz A, Durka W, Klotz S and Brandl R (2002) Which species become aliens. *Evolutionary Ecology Research* 4: 385–405
- Purvis A, Agapow PM, Gittleman JL and Mace GM (2000) Nonrandom extinction and the loss of evolutionary history. *Science* 288: 328–330
- Purvis A and Rambaut A (1995) Comparative analysis by independent contrasts (CAIC): an Apple Macintosh application for analysing comparative data. *CABIOS* 11: 247–251
- Pyšek P (1997) Clonality and plant invasions: can a trait make a difference?. In: de Kroon H and van Groenendael J (eds) *The Ecology and Evolution of Clonal Plants*, pp 405–427. Backhuys Publishers, Leiden, The Netherlands
- Pyšek P (1998) Is there a taxonomic pattern to plant invasions?. *Oikos* 82: 282–294
- Pyšek P, Prach K and Šmilauer P (1995) Relating invasion success to plant traits: an analysis of the Czech alien flora. In: Pyšek P, Prach K, Rejmánek M and Wade M (eds) *Plant Invasions: General Aspects and Special Problems*, pp 39–60. Academic Publishing, Amsterdam
- Pyšek P, Sádlo J, Mandák B and Jarošík V (2003) Czech alien flora and the historical pattern of its formation: what came first to Central Europe?. *Oecologia* 135: 122–130
- Rabinowitz D and Rapp JK (1981) Dispersal abilities of seven sparse and common grasses from a Missouri prairie. *American Journal of Botany* 68: 616–624
- Reichard SH and Hamilton CW (1997) Predicting invasions of woody plants introduced into North America. *Conservation Biology* 11: 193–203
- Rejmanek M (1996) A theory of seed plant invasiveness: the first sketch. *Biological Conservation* 78: 171–181
- Rejmanek M and Richardson DM (1996) What attributes make some plant species more invasive?. *Ecology* 77: 1655–1661
- Ricklefs RE and Renner SS (1994) Species richness within families of flowering plants. *Evolution* 48: 1619–1636
- Ricklefs RE and Starck JM (1996) Applications of phylogenetically independent contrasts: a mixed progress report. *Oikos* 77: 167–172

- Ridley M (1983) *The Explanation of Organic Diversity*. Clarendon Press, Oxford, UK
- Ruesink JL (2003) One fish, two fish, old fish, new fish: which invasions matter?. In: Kareiva P and Levin SA (eds) *The Importance of Species*, pp 161–178. Princeton University Press, Princeton, NJ
- Scott JK and Panetta FD (1993) Predicting the Australian weed status of southern African plants. *Journal of Biogeography* 20: 87–93
- Shrader-Frechette K (2001) Non-indigenous species and ecological explanation. *Biology and Philosophy* 16: 507–519
- Sokal RR and Rohlf FJ (1994) *Biometry: Principles and Practice of Statistics in Biological Research*, 3rd edn. W. H. Freeman and Co., New York, NY
- Simberloff D (2000) What do we really know about habitat fragmentation?. *Texas Journal of Science* 52: 5–22
- Thébaud C and Simberloff D (2001) Are plants really larger in their introduced ranges?. *American Naturalist* 157: 231–236
- Thompson K, Hodgson JG and Rich TCG (1995) Native and alien invasive plants: more of the same?. *Ecography* 18: 390–402
- Trzcinski MK, Fahrig L and Merriam G (1999) Independent effects of forest cover and fragmentation on the distribution of forest breeding birds?. *Ecological Applications* 9: 586–593
- Warming E (1909) *Oecology of Plants: An Introduction to the Study of Plant Communities*. Clarendon Press, Oxford, UK
- Watt AS (1947) Pattern and process in the plant community. *Journal of Ecology* 35: 1–22
- Westoby M, Cunningham SA, Fonseca CR, Overton JMC and Wright I (1998) Phylogeny and variation in light capture area deployed per unit investment in leaves: designs for selecting study species with a view to generalising. In: Lambers H, Poorter H and Van Vuuren MMI (eds) *Inherent Variation in Plant Growth: Physiological Mechanisms and Ecological Consequences*, pp 539–566. Backhuys Publishers, Leiden
- Westoby M, Leishman M and Lord J (1995) On misinterpreting the 'phylogenetic correction'. *Journal of Ecology* 83: 531–534
- Westoby M, Leishman M and Lord J (1996) Comparative ecology of seed size and dispersal. *Philosophical Transactions of the Royal Society, London, Series B* 351: 1309–1318
- Williamson M (1999) Invasions. *Ecography* 22: 5–12
- Williamson MH and Fitter A (1996) The characters of successful invaders. *Biological Conservation* 78: 163–170
- Zar JH (1998) *Biostatistical Analyses*, 4th edn. Prentice Hall, Upper Saddle River, NJ