Darwin’s naturalization conundrum: dissecting taxonomic patterns of species invasions

Jeffrey M. Diez,¹* Jon J. Sullivan,¹ Philip E. Hulme,¹ Grant Edwards² and Richard P. Duncan¹
¹National Centre for Advanced Bio-Protection Technologies, PO Box 84, Lincoln University, Canterbury, New Zealand
²Agriculture and Life Sciences Division, Lincoln University, Canterbury, New Zealand
*Correspondence: E-mail: diezj@lincoln.ac.nz

Abstract
Darwin acknowledged contrasting, plausible arguments for how species invasions are influenced by phylogenetic relatedness to the native community. These contrasting arguments persist today without clear resolution. Using data on the naturalization and abundance of exotic plants in the Auckland region, we show how different expectations can be accommodated through attention to scale, assumptions about niche overlap, and stage of invasion. Probability of naturalization was positively related to the number of native species in a genus but negatively related to native congener abundance, suggesting the importance of both niche availability and biotic resistance. Once naturalized, however, exotic abundance was not related to the number of native congeners, but positively related to native congener abundance. Changing the scale of analysis altered this outcome: within habitats exotic abundance was negatively related to native congener abundance, implying that native and exotic species respond similarly to broad scale environmental variation across habitats, with biotic resistance occurring within habitats.

Keywords
Biotic resistance, invasibility, naturalization, niche overlap, scale, stages of invasion, taxonomy.

INTRODUCTION
It is intuitive that interactions between exotic and resident native species may influence the outcome of species invasions. Theory highlights the potential role of niche overlap in shaping patterns of invasion (Levine et al. 2004; Mitchell et al. 2006) but niche relationships are often difficult to quantify and, as a result, phylogenetic or taxonomic relationships are often used as a metric of niche overlap (e.g. Ricciardi & Atkinson 2004). The idea that the success of exotic species is influenced by their relatedness to natives has a long history in ecology (Darwin 1859), and overlaps conceptually with modern ideas of niche theory (Tilman 2004) and phylogenetic structuring of community organisation (Cavender-Bares et al. 2006; Swenson et al. 2007).

Nevertheless, there are conflicting arguments and some empirical support for both positive and negative associations between relatedness and invasion (Daceyler 2001; Proches et al. 2007). Darwin (1859) recognized that relatedness to the native community should confer an advantage to exotic species due to similar adaptations to the environment. He acknowledged, however, that novel genera are often successful at invading new locations (see also Alphonse de Candolle (1855)), an advantage that could arise because less closely related species are less likely to encounter direct competitors or to share natural enemies with native species. This advantage is now termed ‘Darwin’s naturalization hypothesis’ and has some empirical support (Rejmanek 1999; Ricciardi & Atkinson 2004; Strauss et al. 2006). However, Darwin’s intuition that closely related species may also share traits that pre-adapt them to the new environment, or may increase mutualistic or facilitative interactions (Bruno et al. 2003), is also supported (Duncan & Williams 2002). Still other studies have found no association with natives (e.g. Lambdon & Hulme 2006a; Ricciardi & Mottiari 2006). Thus, even as studies become more numerous and quantitative the expectation for how relatedness influences invasion has not changed considerably since Darwin’s day.

There are several possible reasons why no consensus has emerged on the impact of relatedness. First, relatedness is likely to affect invasion via different but simultaneous mechanisms, such as competition and facilitation, which could alternatively suppress or enhance invasion success. The relative importance of these mechanisms is likely to vary across taxa and ecosystems. Second, the relationship
between relatedness and niche overlap, a key assumption of these studies, is not well understood. For example, if competition among related species is reduced over time through niche divergence, the presence of greater numbers of native relatives may suggest greater potential niche space in the landscape and more opportunities for invasion (Fig. 1a). Alternatively, if the niche space of a group of closely related species is limited, then the presence of greater numbers of native relatives would suggest increased niche overlap and fewer opportunities for invasion if competition limits success (Fig. 1b).

Third, the processes underlying many ecological patterns may change with scale (Wiens 1989; Levine 2000), and this may change the interpretation of these patterns of invasion (Lambdon & Hulme 2006a; Proches et al. 2007). Competition between species, for example, may be important at local scales (e.g. within habitats) and lead to exclusion of exotics by closely related natives. At larger scales (e.g. among habitats) exotics may be more likely to co-occur with their native relatives because they share traits that pre-adapt them to particular environments and because greater habitat heterogeneity at these broad scales my permit coexistence (Fig. 2). These hypothesized scale relationships parallel those observed more generally with native-exotic richness relationships (Fridley et al. 2007). Finally, invasions can be characterized as a series of stages including introduction, initial establishment and subsequent spread of a species (Williamson & Fitter 1996; Dietz & Edwards 2006; Theoharides & Dukes 2007). Because the processes determining success may differ among stages (Kolar & Lodge 2001; Duncan et al. 2003; Lloret et al. 2005), so too may the direction and importance of relationships between native and exotic relatives. This has not previously been addressed in a single study system.

In this study we examine the association between relatedness and invasion success using a unique long-term dataset that documents changes in the flora of the Auckland region of New Zealand between 1840 and 1985. We show that using both the number and the abundance of native congeners to predict the success of exotic species better distinguishes among competing hypotheses. For example, a positive relationship between invasion success and the number of native congeners may suggest either greater niche opportunities or facilitation. Incorporating abundance can mediate between these possibilities. Whereas a negative relationship with abundance would support the niche opportunity hypothesis, a positive relationship would suggest facilitation (Fig. 3). We further test for scale

Figure 1 Relationships between niche breadth assumptions and invasion hypotheses. Circles represent species, grouped together in genera of different sizes. If greater native species richness implies greater environmental niche breadth (a), a higher naturalization probability could be expected with more native congeners. If generic niche breadth is fixed (b), naturalizing species will likely face greater interaction strength, whether competition (−) and/or facilitation (+).

Figure 2 Scale is hypothesized to influence the interpretation of the relationship between invasion success and the abundance of native relatives. With increased scale comes increased abiotic and biotic heterogeneity and less likelihood of significant direct interactions among species, represented here by decreased shading. The observed relationship, on the x-axis, is likely to suggest different broad classes of processes depending on the scale of study. We suggest some of the processes likely to be dominant at different scales, but they need not be mutually exclusive.
of the region completed by Alan Esler in 1985. Esler (1988) recorded the relative abundance of the native and exotic vascular plant species in a 308 km² area encompassing contemporary urban Auckland. These surveys were performed to complement a similar 1871 survey by Thomas Kirk (Kirk 1871) of the then 174 km² Auckland region; together these surveys document changes in relative abundance of the native and exotic flora over a 114 year period. We supplement these surveys with additional observations of species naturalized in the region prior to 1985, collated by Ewen Cameron (Auckland Herbarium).

To assess the probability of naturalization (defined as the establishment of a self-maintaining population after introduction) we used a national database of 25 049 exotic species known to have been intentionally or accidentally introduced to New Zealand, compiled from a variety of sources and updated from the 24 744 species used in Duncan & Williams (2002). Of these, 932 species naturalized in the Auckland area before 1985. We assume that the national database provides a record of species introduced to the Auckland region because of Auckland’s role as the major hub of transportation and the largest horticultural and trade center of the island (currently home to over a quarter of the country’s population).

The relationships between introduced and native species were assessed using the 399 native species present at the time of Kirk’s 1871 survey (Duncan & Young 2000) plus a further 116 species recorded subsequently, giving a total of 515 native species (Esler 1988). For each introduced species the number of native congeners was used as a measure of relatedness. In addition, we calculated the average abundance of native congeners (in 1985) using the seven relative abundance categories in Esler (1988).

To understand the subsequent spread of naturalized species, we used the three 1985 relative abundance categories that Esler (1988) assigned each naturalized species (updated by Ewen Cameron; unpublished data). Class 1 species were defined as ‘naturalized freely and are common’, class 2 species were naturalized ‘in lower numbers and/or with fairly limited distribution,’ and class 3 species were ‘in very low numbers’ or only marginally naturalized in the region by 1985. To account for the fact that species naturalized at various times in the past, and thus have had different amounts of time to spread, the naturalization date of each species was assigned to one of six periods (pre-1840, 1840–1870, 1870–1900, 1900–1940, 1940–1970, and 1970–1985) using records by a series of botanists between 1870 and 1985, including T. F. Cheeseman (1883, 1906), H. H. Allan (1940), and A. J. Healy (see Esler 1988 for full reference list). Naturalization period was used as a covariate to account for the role of time in explaining greater abundance in 1985. Relatedness to the native community was again measured as the number of congeneric native congeners both within and across habitats, and provide a unique test of how these relationships change across two stages of the invasion process (naturalization and spread). Tying these analyses together is a multilevel framework that accounts for taxonomic dependence, allows incorporation of covariates at multiple levels, and partitions variance to assess how invasion varies across taxonomic levels.

METHODS

Data sets

The relatively recent European colonization of New Zealand and records of early botanists enable a unique assessment of a biological invasion process. European settlement of the Auckland region began around 1840, before which the earlier Maori settlers introduced few plants (Leach 2005). Since this time, Auckland has changed from a semi-natural landscape in 1840, to a predominately agricultural landscape with large remaining patches of native vegetation in 1870, to an urbanized landscape with smaller remnant patches of native vegetation by 1985. Most of the data used in this study comes from comprehensive surveys
species, and the abundance of congeneric natives was calculated as the mean abundance of species in the genus.

In order to examine the scale-dependence of the relationship with native congeners, we calculated a second measure of relatedness that included only those native species found in the same habitat as a given naturalized species (either forest, open, wetland, coastal or other). Limiting congeners to those sharing the same habitat reduces spatial scale and environmental heterogeneity and increases the possibility of direct interactions. All congeners were included for 67 naturalized species for which habitat was not recorded. We could not do this within-habitat analysis for the naturalization stage because we lacked data on the habitat associations of many of the 25 049 introduced species. The numbers of native and naturalized species found in the five habitats are summarized in Table 1.

**Statistical models**

The probability of naturalization from the pool of all introduced species (naturalization stage) and their subsequent abundance (spread stage) were modeled within a hierarchical Bayesian framework. The multilevel framework was used to evaluate covariates at the appropriate levels (species and genus) while accounting for non-independence of taxonomically grouped data at the genus and family levels (Harvey & Pagel 1991; Gelman & Hill 2007; McMahon & Diez 2007).

The probability of naturalization was modeled as a Bernoulli process in which each introduced species i has a probability of successful naturalization p_i. This probability is made conditional on the species’ genus through a logit link function and genus level intercept term logit(p_i) = \alpha_g. Relationships with the number and abundance of native congeners are then evaluated at the genus level as:

\[ \alpha_g = \alpha_y + \beta_1(N_{\text{native}_g}) + \beta_2(P_{\text{native}_g}) + \epsilon_g \]

where \( \beta_1 \) describes the relationship between naturalization success and the number of native species in the genus, \( N_{\text{native}_g} \) and \( \beta_2 \) describes the relationship with the average abundance of those native species, \( P_{\text{native}_g} \). These genus-level relationships are further nested within families, which are specified by family-specific intercepts, \( \alpha_f \), drawn from a distribution with overall mean, \( \alpha_0 \), and variance. The residual error term at the genus level, \( \epsilon_g \), is estimated from the data.

The subsequent spread model is similar in structure but uses a latent variable formulation of an ordered, multinomial regression (Gelman & Hill 2007) to model the categorical abundance of naturalized species. The date of naturalization is incorporated as an explanatory variable at the species level in order to account for the effect of time on abundance. At the genus level the effect of both number and abundance of native species are estimated as above, again with family level intercepts to account for taxonomic structure of the data. We ran two models with relatedness to natives calculated in two ways: first, including all native species within the naturalized species’ genus and second, including only those native congeners found in the same habitat as the naturalized species. This second analysis excludes native congeners that are unlikely to interact with the naturalized species because they occur in different habitats.

Both the naturalization and spread models were also fit with no explanatory variables in order to partition the amount of variation explained at the species, genus and family levels (Raudenbush & Bryk 2002; McMahon & Diez 2007). As Bayesian models, all parameters are given prior distributions, which in this case are non-informative to allow data to drive estimation. The overall intercept \( \alpha_0 \) and regression coefficients are assigned Normal prior distributions with mean 0 and variance 1000, and variance terms at each level are given broad uniform priors on the standard deviation following Gelman (2006). Models were fit using OpenBugs v2.10 (Thomas et al. 2006) called using the BRugs package from R 2.4.1 (R Development Core Team 2006) and run for between 70 000 and 100 000 iterations with substantial burn-in periods and thin rates to achieve convergence. Convergence was monitored using the Gelman-Rubin statistic and inspection of sample histories.

**RESULTS**

Variance partitioning revealed that most of the variation in both naturalization and spread was at the species level. For the probability of naturalization, 29, 24, and 47% of the variation was at the family, genus and species levels respectively, and for spread 2, 7, 90% respectively.

The probability of naturalization was positively related to the number of native species in a genus but negatively
related to the abundance of those natives (Fig. 4a), with the 95% credible intervals of both variables excluding 0.

The 1985 abundance of naturalized species was negatively related to the period of naturalization, meaning more recently naturalized species tended to be less abundant. Having controlled for period of naturalization, abundance was not significantly related to the number of native species in a genus but positively related to the average abundance of those natives (94% of the credible interval > 0; Fig. 4b). In contrast, when only congeners sharing the same habitat were included in the analysis, the abundance of naturalized species was negatively related to the abundance of native congeners, while the relationship with the number of congeners remained non-significant (Fig. 4c).

**DISCUSSION**

Phylogenetic and taxonomic relationships are increasingly being used to test longstanding hypotheses about patterns of biological invasions (Daehler 2001; Duncan & Williams 2002; Lambdon & Hulme 2006a; Strauss et al. 2006), but such studies have provided varying results. Our study highlights some of the major conceptual reasons why this might be and provides a unique empirical test across two stages of invasion. We show that patterns can differ with stage of invasion, scale, and both the number and abundance of native relatives. Central to testing these ideas is the availability of a comprehensive dataset documenting a biological invasion coupled with multilevel analysis to simultaneously incorporate covariates at appropriate taxonomic levels.

We found that probability of naturalization for species introduced to the Auckland region was positively associated with the number of native congeners but negatively associated with congeneric abundance. While the positive relationship with the number of species suggests either greater niche opportunity or facilitation (Fig. 1), the negative relationship with abundance points to a combination of native adaption and biotic resistance (Fig. 3). Larger genera may signify greater niche opportunities in the landscape but, having accounted for that, genera with more abundant species are more difficult to invade.

This pattern changed in the second stage of invasion – how widespread a species becomes once naturalized. The regional abundance of naturalized species was not significantly related to the number of native congeners but positively related to their abundance. Again we see the importance of distinguishing the effects of the number of related species from their abundance. While the lack of a relationship with the number of native congeners would suggest that exotic success is unrelated to native relatives, the positive relationship with native abundance suggests two alternative hypotheses: there is facilitation among native and exotic relatives, or they are responding similarly to other factors. Certainly one mechanism behind the positive relationship would be shared traits at the genus level that lead to the abundance of native and exotic species co-varying across the landscape. These could include traits that allow colonization of similar habitats along with traits influencing processes such as dispersal, although such traits are not always obvious (Pysek & Hulme 2005).

One problem with interpreting broad-scale distributional studies is uncertainty as to whether species are actually interacting (Lambdon & Hulme 2006a; Proches et al. 2007). This is problematic when many of the inferred mechanisms depend on these interactions. By analyzing our data both

\[ \begin{align*}
\text{Coefficient Estimate} & \quad \text{Coefficient Estimate} & \quad \text{Coefficient Estimate} \\
-0.8 & & -0.6 \\
-0.6 & & -0.4 \\
-0.4 & & -0.2 \\
-0.2 & & 0.0 \\
0.0 & & 0.2 \\
0.2 & & 0.4 \\
-12 & & -8 \\
-8 & & -4 \\
-4 & & 0 \\
4 & & 8 \\
8 & & 12 \\
\end{align*} \]
among and within-habitats we altered the scale of analysis and attempted to increase the likelihood that related species had opportunities to interact. Doing this altered the outcomes in the way we would predict if species interactions were important: the abundance of exotics was negatively related to the abundance of native relatives within-habitats but positively related among habitats (Fig. 4). This suggests that within habitats, biotic resistance due to competitive interactions or shared natural enemies may be playing a role (Fig. 2). More direct tests are needed however, preferably through experimental manipulation, to assess how relatedness may or may not actually influence mechanisms such as competition (Cahill et al. 2007).

These results are also consistent with studies exploring the relationship between native and exotic species richness. Small-scale experiments most commonly find negative relationships between native diversity and invasibility (Fridley et al. 2007), with this relationship tending to reverse in observational studies at larger scales (Davies et al. 2005; Fridley et al. 2007). At smaller scales, the results more likely reflect the outcome of direct interactions among species. Restricting analysis in this study to within habitats has a similar effect, although the spatial scale of habitats in this study is still larger than the scales at which experiments typically find negative relationships between native and exotic species. We do not have more information about the degree to which related species are actually interacting within habitats, leaving open the possibility that negative relationships result from factors independent of direct interactions.

The different outcomes we observed at different stages in the invasion process (naturalization vs. spread) could also be due in part to scale-dependent processes. The establishment stage is an inherently more local process than subsequent spread across a landscape (Theoharides & Dukes 2007), and establishment may be analogous to invasion experiments in which propagules are added to local communities (e.g. Levine 2000). While similar biotic and abiotic filters may operate at both the establishment and spread stage, spread can be thought of as repeated establishment and, while able to establish at one location, species may face additional challenges in spreading to other locales. Habitat configuration, dispersal processes and dispersal barriers are likely to become important. Similarly, climate suitability and biotic interactions may be important for both establishment and spread, but to different degrees (Theoharides & Dukes 2007).

**Taxonomy, phylogeny, and measures of relatedness**

Relatedness between exotic species and the native community can be measured in different ways, each carrying assumptions relevant to interpreting patterns of invasion. Most studies of Darwin’s naturalization hypothesis, including our own, have used genus-level taxonomy to measure relatedness. An implicit assumption in these analyses is that the species in each genus are equally related, implying that the degree of relatedness, and hence the strength of biotic resistance, can be conveniently summarized by the number of congeners. However, genera will differ in divergence times, meaning that species in different genera will be related to varying degrees. If there were a systematic relationship between divergence time and genus size, such that species in larger genera were less closely related than those in small genera, then the number of congeneric species will be a less useful measure of relatedness. We are not aware of any study documenting a consistent relationship between genus size and divergence time, and a systematic bias across the wide taxonomic range in this study seems unlikely. Moreover, a recent study of phylogenetic patterns of invasion in Californian grasses obtained comparable results using trees with known and unknown (fixed) branch lengths, implying that adjusting for divergence times had little influence on the outcome (Strauss et al. 2006).

Ultimately, questions about invasion and relatedness hinge on the degree to which relatives overlap ecologically in terms of niche requirements, mutualists and natural enemies. While we may be able to measure relatedness more precisely using well resolved phylogenies, we still need to make assumptions about how the degree of relatedness translates into ecological overlap. Despite this, the findings in this paper are likely to apply regardless of whether we use phylogenetic or taxonomic measures of relatedness: the stage of invasion, the scale of the study, and the performance of relatives in addition to their numbers are likely to mediate associations between invasion outcomes and relatedness.

**Phylogenetic patterns of invasion in context**

Phylogenetic and taxonomic patterns of invasion can inform the growing body of work aimed at understanding the phylogenetic organization of native communities (Webb et al. 2006). Communities can be characterized as phylogenetically under or overdispersed if there is greater or lesser relatedness, respectively, than expected by chance. Such patterns are the net result of opposing forces of phylogenetic attraction (e.g. environmental filtering) and repulsion (e.g. competition), which can act over multiple scales (Cavender-Bares et al. 2006; Kembel & Hubbell 2006; Swenson et al. 2007). For example, Helmus et al. (2007) found evidence for both phylogenetic attraction and repulsion in sunfish communities, yielding a net result of neither under nor overdispersion. More closely related species responded similarly to environmental gradients, but
competition acted to limit coexistence of related species. Our results also suggest that relatedness can influence community composition in multiple ways, with scale important in determining these effects. Examining how the composition of invaded communities change through time is a promising avenue for increasing our understanding of phylogenetic community structure.

Finally, despite the statistically significant taxonomic patterns found in this study and others, it is important to consider their ecological relevance. Nearly half of the variation in naturalization and 90% of the variation in subsequent spread was found at the species level. Thus, although naturalized exotic species are non-randomly clustered at the family level (Dachler 1998; Pysek 1998), our results suggest that species-specific traits or processes are more important drivers of variation in naturalization and particularly subsequent spread. The large amount of variation associated with species-level differences could largely reflect differences in human-mediated processes of introduction effort and transport (Duncan et al. 2003; Lambdon & Hulme 2006b) and/or the biogeography of introduced species (Lloret et al. 2004). Ultimately, we suggest that phylogenetic patterns of invasion may provide interesting insights into the organization of ecological communities, but only a limited basis for predicting invasions.

ACKNOWLEDGEMENTS

This study would not have been possible without the work of several excellent botanists over the years, particularly Thomas Kirk, Alan Esler, and Ewen Cameron. Thanks to Peter Williams and Hazel Gatehouse for help with compiling the databases, and Brian Enquist and three anonymous referees for comments on the manuscript. Funding was provided by the National Centre for Advanced Bio-Protection Technologies, Lincoln University.

REFERENCES


