

PLANT SPECIES INVASIONS ALONG THE LATITUDINAL GRADIENT IN THE UNITED STATES: COMMENT

Jason D. Fridley,^{1,4} Hong Qian,² Peter S. White,¹
and Michael W. Palmer³

Few biogeographic patterns have been as well documented as the latitudinal gradient in species richness. For North American plants, the decrease in number of species from the subtropics to the arctic is well established (Stevens 1989, Currie 1991, Qian 1999), but becomes less defined as the spatial grain of species richness decreases (e.g., Monk 1967, Glenn-Lewin 1977). There is considerable interest among biogeographers and community ecologists as to whether clear latitudinal patterns are manifest at relatively fine spatial resolutions, where the signature of local processes such as historical disturbances or fine-scale environmental variation could obscure broad-scale geographic drivers of richness (Rahbek and Graves 2001, Whittaker et al. 2001). Because adequate data have generally been lacking, however, there remained until recently no systematic attempt to address the latitudinal gradient of plant richness for the full extent of the United States using relatively fine-scale data.

Stohlgren et al. (2005) examined native and exotic vascular plant species richness at the county level for the 48 conterminous U.S. states in relation to latitude, climate, topographic and biotic factors, and human disturbance. By comparing a variety of multifactor models explaining native and exotic plant richness, Stohlgren et al. (2005) concluded that: (1) there is no relationship between native plant richness and latitude at the county level; (2) the strongest single predictor of exotic species richness is native species richness; and (3) bird species richness, a surrogate for habitat heterogeneity, is the best predictor of native plant richness. If

true, these observations would represent a major advance in the study of plant diversity patterns, as they would suggest scale thresholds where broad-scale geographic patterns are no longer manifest due to the increasing importance of other factors such as local habitat heterogeneity. However, we argue that these conclusions are inaccurate, stemming from uncritical evaluation of completeness in the data set and from artifacts deriving from inappropriate data transformation. Furthermore, we analyze a more appropriate data set to the issue of county-level patterns of native and exotic species richness, and obtain different results than that of Stohlgren et al. (2005).

A major conclusion of Stohlgren et al. (2005:2301) is that “regression analyses showed no relationship between latitude and native plant species density (richness).” We assert that this result is an artifact of two major deficiencies in their study; namely, that (1) their data set of county richness values underestimates true county richness, and (2) richness values were incorrectly transformed to account for variation in survey area. Plant data used by Stohlgren et al. (2005) come from the Biota of North America Program (BONAP), a collection of taxon occurrences from the literature and from herbarium surveys recorded by county (Stohlgren et al. 2003). These data are widely regarded as the “standard plant data set for many government and non-government agencies” (Stohlgren et al. 2005:2299), and we agree that this is the most comprehensive collection of plant occurrence records for North America. However, records for many counties were not themselves collected with the aim of providing complete floristic lists of counties (see Palmer 1995). Thus, although some counties are well collected, others remain less well described for a variety of reasons (such as low survey effort). In addition to adding error to the assessment of county-level richness patterns (thereby undermining the signal of geographic pattern in the data), variation in floristic completeness among counties could cause spurious correlations between native and exotic richness; as well-surveyed counties are more likely to contain more occurrences of both native and exotic species. The use of such data for the analysis of richness patterns is appropriate if researchers make sufficient attempts to correct for, or are at least be able to quantify, the variation in floristic completeness among counties. Stohlgren et al. (2005) give no assessment of data quality in their analysis. Moreover, the only attempt made to reduce the bias associated with including poorly surveyed counties is the removal of 114 counties with fewer than 100 occurrences. This procedure is both insufficient (114 is <4% of the data set) and flawed, in that occurrences are a function of both survey effort and

Manuscript received 3 February 2006; accepted 6 March 2006. Corresponding Editor: H. Hillebrand.

¹ Department of Biology, University of North Carolina, Chapel Hill, North Carolina 27599 USA.

² Research and Collections Center, Illinois State Museum, 1011 East Ash Street, Springfield, Illinois 62703 USA.

³ Botany Department, Oklahoma State University, Stillwater, Oklahoma 74078 USA.

⁴ Present address: Department of Biology, Syracuse University, 130 College Place, Syracuse, New York 13244 USA. E-mail: fridley@syr.edu

actual floristic richness; omission of counties based on low occurrences will spuriously remove legitimate but less rich county richness estimates.

The conclusions of the paper are further eroded by the decision of Stohlgren et al. (2005) to model plant species “density,” rather than raw richness values, by simply dividing county richness by county area (see Gotelli and Colwell 2001). Species richness and area are not linearly related but are most commonly described by the power function $S = cA^z$ (Arrhenius 1921, Rosenzweig 1995). The transformation applied by Stohlgren et al. (2005) essentially assumes that the richness accumulation rate with area (z) is 1, which is a minimum of 4–5 times higher than that commonly reported for county- and state-sized areas (typically between 0.1 and 0.2; Rosenzweig 1995). As a result, larger counties are heavily penalized in their species richness estimates (particularly those in western states and eastern states like Florida and Maine) while those in states of small counties (e.g., Georgia, Indiana, Ohio) get spuriously high values. As there are clear geographic gradients in county size in the United States, this transformation severely inhibits Stohlgren et al. (2005) from uncovering true geographic patterns of county-level richness. It comes as no surprise, then, that this flawed and geographically biased transformation leads Stohlgren et al. (2005) to conclude that there is no latitudinal pattern in richness at the U.S. county level.

The flawed area transformation of richness values also leads to a spuriously strong correlation between native and exotic richness, and between native plant and bird richness. Because all three variables (native plant richness, exotic plant richness, bird richness) are transformed in the same way, they have a built-in correlation that, as we demonstrate, inflates the strength of each relationship. Moreover, with identical transformation, the correlation of these variables is maintained, while their fit to other potentially strong variables (such as climate factors) is diminished, giving the appearance that climate variables are worse predictors of diversity than these “biotic” variables. In addition, counties that are well surveyed for native species are likely to be well surveyed for exotic species, contributing to an artificially high correlation between these variables. Our revised analysis suggests that native and exotic species richness at the county level are far less correlated than suggested by the analysis of Stohlgren et al. (2005). Indeed, even Stohlgren et al. (2005) assert that native and exotic plant richness patterns are explained by very different models; considering this plus artifacts introduced by data transformation, we assert that native and exotic plant richness at the county level are unlikely to be as tightly coupled ($R^2 = 0.9$) as they report.

To provide a more accurate assessment of plant richness patterns at the county level, we have taken a subset of data from the Floras of North America Project

(Withers et al. 1998, Palmer 2005) to address the county-level latitudinal pattern with more adequate data. The database of the Floras project includes >2700 areas that were surveyed specifically for documenting the vascular flora of a defined location. Although no flora is a perfect representation of the total taxonomic richness of a particular area, taxonomic completeness is usually a goal of a floristic survey (Palmer et al. 1995). Each record summary from the Floras project includes total vascular plant species, exotic species, area, geographic location, and date of publication. From this database we selected floras that are roughly county sized, between 50 km² and 25 000 km² with a mean of 2965 km² (conterminous U.S. county range is from 59 to 52 000 km² with a mean of 2500 km²). We excluded floras that were small islands or peninsulas or did not represent contiguous areas, those that were entirely within wetlands, and those that were published before 1950. Our final set includes 600 floras covering the full extent of North America north of Mexico, including 541 from the conterminous United States (Fig. 1). Note that spatial gaps in Fig. 1 do not necessarily indicate historical gaps of botanical exploration; the floras could have been written before 1950, or for larger (e.g., state) or smaller areas. In addition, although more recent floras should generally be more complete floras, there is no temporal trend in exotic richness in our post-1950 data set ($P > 0.12$ in a linear model of log(exotic richness) over time; see also Palmer [2005]).

To evaluate the latitudinal gradient in native plant species richness we first standardized richness estimates to a common area using an empirically derived species–area relationship. We used the common log transformation of the power law, $\log S = \log c + z(\log A)$, to estimate parameter z via linear regression of the 600 observed area and richness values. Our calculated z (=0.158) was then used to transform all richness estimates to that of the geometric mean area in our data set (equivalent to 1074 km²), using the formula $\log S_c = z(\log A_{\text{mean}} - \log A) + \log S$. We then modeled corrected native richness (log S_c) in response to latitude and longitude, for both the full 600 flora data set and the conterminous U.S. subset of 541 floras corresponding to the spatial extent of the Stohlgren et al. (2005) study (Fig. 2).

For both the full data set and the conterminous U.S. subset, best-fit regression models evaluated by the Akaike Information Criterion include third-order terms for both latitude and longitude (Table 1), and explain 31% and 14% of county-level variance in native richness for North America and the conterminous United States, respectively. Latitude and longitude are significant predictors of native richness, although the nature of their effects depends on which data set is used (Table 1). The latitudinal gradient is much stronger when Canada and Alaska are included (Fig. 2); when confined to the

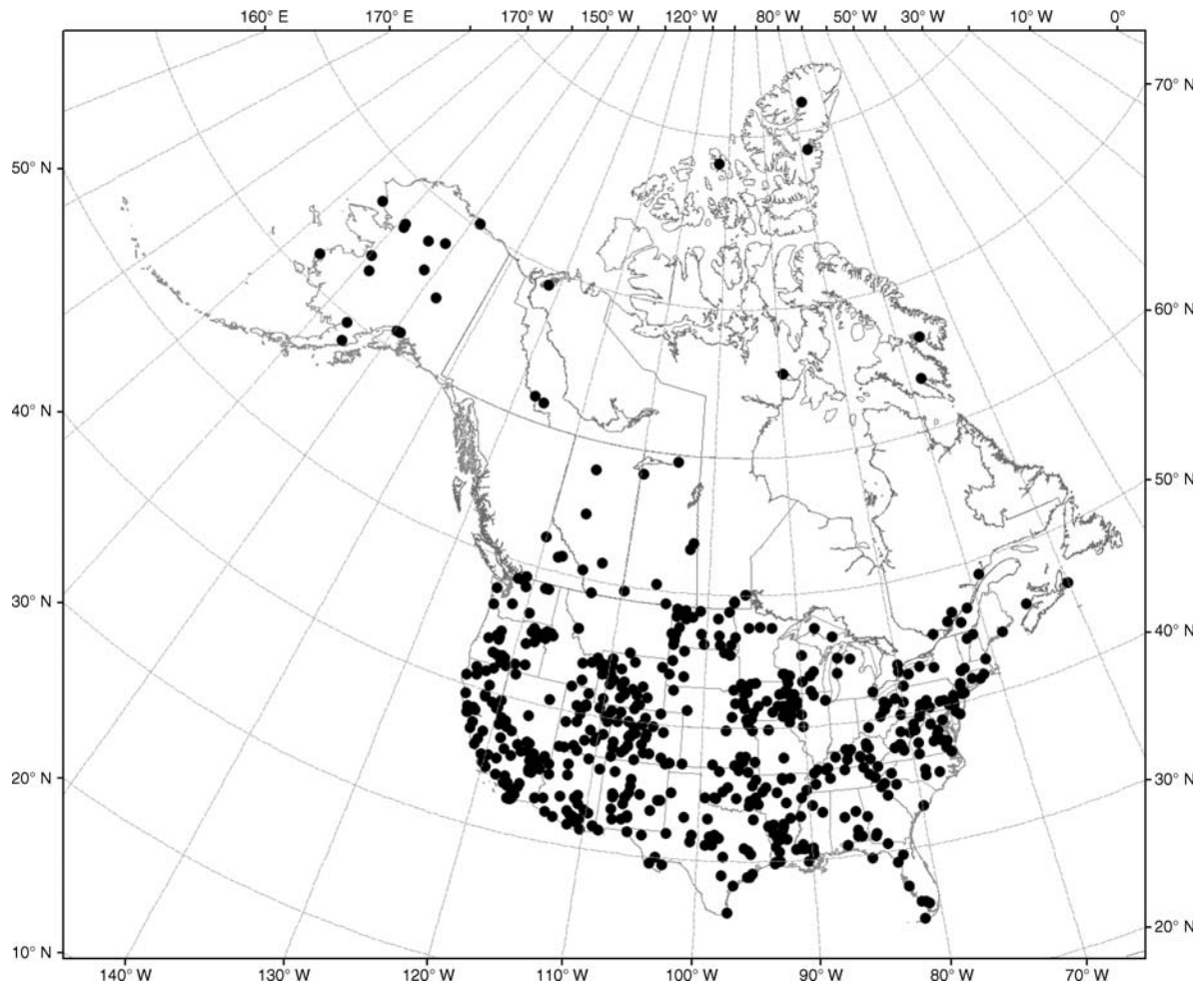


FIG. 1. Locations of 600 county-sized floras from the Floras of North America Project (Withers et al. 1998, Palmer 2005).

latitudes of the conterminous United States, richness decreases weakly and monotonically with latitude, while significant higher order effects emerge with a larger latitudinal range. Significant third-order effects of longitude are detectable in both data sets, where richness is significantly higher along eastern and western coasts. Similar models, but for exotic species richness, contain significant effects of longitude but not latitude.

Although native and exotic species richness are correlated in our data set, the tightness of this relationship is much less than that reported by Stohlgren et al. (2005). When we standardize exotic richness to a constant area (using a fitted z value of 0.212 relating exotic richness and flora area), native richness explains 44% of the variance in area-standardized exotic richness at the county level (Fig. 3), as opposed to the 90% reported by Stohlgren et al. (2005). If instead we use the approach of Stohlgren et al. (2005) of dividing raw native and exotic richness values by flora area, the

variance in exotic richness explained by native richness is artificially enhanced to 74%. Combined with the likely variation in completeness of county lists in the BONAP data, the flawed area transformation of both native and exotic richness by Stohlgren et al. (2005) produces native–exotic richness correlations that are much higher than we believe actually exist in county-level distribution patterns. Although native and exotic plant richness are likely to respond similarly to many environmental factors, we find it likely that other human and historical factors will ultimately prove to be the most powerful predictors of exotic species distributions (Rejmánek 2004, Taylor and Irwin 2004).

We conclude that, contrary to the conclusions of Stohlgren et al. (2005), there is a significant relationship between latitude and native plant richness at the county level for the conterminous United States. The latitudinal gradient is stronger when latitudes above the conterminous United States are considered, which represents a

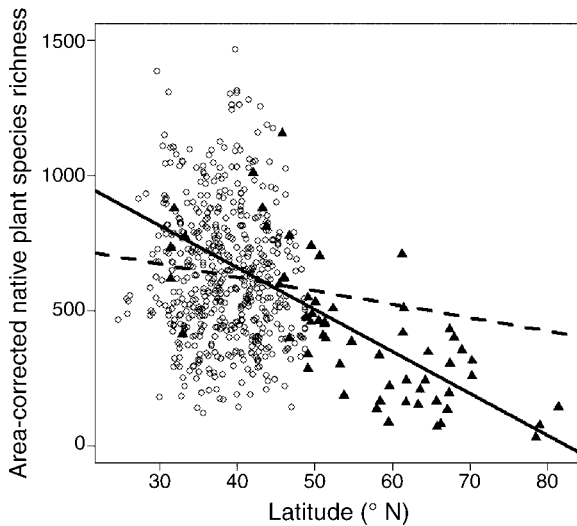


FIG. 2. Relationship between native plant species richness and latitude in 600 county-sized floras for North America (all points, circles and triangles), and for only those floras (541) within the conterminous United States (open circles, dashed regression line). Richness has been corrected to a standard area. Lines represent linear least-squares regressions.

spatial extent not addressed by Stohlgren et al. (2005). Nonetheless, reestablishment of a well-known global pattern at a relatively fine spatial resolution for plants of the United States suggests that broad-scale richness drivers such as climate still play a considerable role in

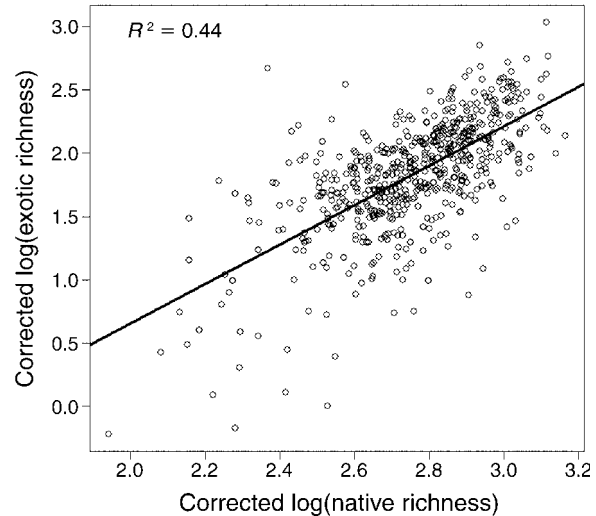


FIG. 3. Relationship between exotic and native plant species richness in 600 county-sized floras from North America, assessed by independently standardizing native and exotic richness to a constant area using an empirically derived species–area relationship. The line represents linear least-squares regression.

the structure of county-sized plant assemblages. This does not preclude a significant role for landscape heterogeneity and environmental favorableness in the fine-scale distribution of species richness: to the contrary, we foresee significant advancement in the under-

TABLE 1. ANOVA results of best-fit models of latitude and longitude predicting area-corrected native species richness in all 600 county-sized floras of (A) North America and (B) a subset of 541 floras from the conterminous United States.

Source of variation	df	ss	ms	F	Coefficient
A) All floras ($N = 600$)					
Latitude	1	4.71	4.71	146.28***	0.0678
Longitude	1	0.64	0.64	19.75***	0.1264
Latitude ²	1	1.41	1.41	43.79***	-0.0024
Longitude ²	1	0.92	0.92	28.49***	8.893×10^{-4}
Longitude \times Latitude ²	1	0.05	0.05	1.68	4.168×10^{-6}
Latitude \times Longitude ²	1	0.03	0.03	0.99	-3.767×10^{-7}
Latitude ³	1	0.16	0.16	5.12*	1.541×10^{-5}
Longitude ³	1	0.33	0.33	10.29**	2.457×10^{-6}
Latitude \times Longitude	1	0.05	0.05	1.63	-7.653×10^{-4}
Residuals	591	19.04	0.03		
B) Conterminous U.S. floras only ($N = 541$)					
Latitude	1	0.14	0.14	4.85*	0.1504
Longitude	1	0.85	0.85	28.63***	-0.4761
Latitude ²	1	0.04	0.04	1.48	-2.963×10^{-4}
Longitude ²	1	0.86	0.86	28.97***	-4.543×10^{-3}
Longitude \times Latitude ²	1	0.06	0.06	2.00	2.312×10^{-5}
Latitude \times Longitude ²	1	0.03	0.03	0.97	2.225×10^{-5}
Latitude ³	1	0.03	0.03	0.91	2.100×10^{-5}
Longitude ³	1	0.53	0.53	17.76***	-1.312×10^{-5}
Latitude \times Longitude	1	0.04	0.04	1.37	2.728×10^{-3}
Residuals	531	15.81	0.03		

Note: Richness was log-transformed.

* $P < 0.05$; ** $P < 0.01$; *** $P < 0.001$.

standing of richness patterns with the increasing availability of geospatial data and a modeling approach that considers the strong scale dependence of richness patterns. Although we commend Stohlgren et al. (2005) for highlighting the dependence of diversity patterns on scale, a key issue in contemporary ecology (Crawley and Harral 2001, Fridley et al. 2005, Palmer, *in press*), we caution that data transformation must be theoretically sound and completeness must be addressed if we are to understand the scale dependence of ecological phenomena.

Acknowledgments

The Floras of North America project is sponsored in part by the National Biodiversity Information Infrastructure, in a contract administered through the Houston Advanced Research Center. J. D. Fridley was supported by the National Parks Ecological Research Fellowship program, a partnership between the National Park Service, the Ecological Society of America, and the National Park Foundation, and funded through a generous grant from the Andrew W. Mellon Foundation.

Literature cited

- Arrhenius, O. 1921. Species and area. *Journal of Ecology* **9**:95–99.
- Crawley, M. J., and J. E. Harral. 2001. Scale dependence in plant biodiversity. *Science* **291**:864–868.
- Currie, D. J. 1991. Energy and large-scale patterns of animal- and plant-species richness. *American Naturalist* **137**:27–49.
- Fridley, J. D., R. K. Peet, P. S. White, and T. R. Wentworth. 2005. Connecting fine- and broad-scale patterns of species diversity: species–area relationships of southeastern U.S. flora. *Ecology* **86**:1172–1177.
- Glenn-Lewin, D. C. 1977. Species diversity in North American temperate forests. *Vegetatio* **33**:153–162.
- Gotelli, N. J., and R. K. Colwell. 2001. Quantifying biodiversity: procedures and pitfalls in the measurement and comparison of species richness. *Ecology Letters* **4**:379–391.
- Monk, C. D. 1967. Tree species diversity in the eastern deciduous forest with particular reference to north central Florida. *American Naturalist* **101**:173–187.
- Palmer, M. W. 1995. How should one count species? *Natural Areas Journal* **15**:124–135.
- Palmer, M. W. 2005. Temporal trends of exotic species richness in North American floras: an overview. *Écoscience* **12**:386–390.
- Palmer, M. W. *In press*. Species–area curves and the geometry of nature. In J. H. Brown, editor. *Scaling biodiversity*. Oxford University Press, Oxford, UK.
- Palmer, M. W., G. L. Wade, and P. R. Neal. 1995. Standards for the writing of floras. *BioScience* **45**:339–345.
- Qian, H. 1999. Spatial pattern of vascular plant diversity in North America north of Mexico and its floristic relationship with Eurasia. *Annals of Botany* **83**:271–283.
- Rahbek, C., and G. R. Graves. 2001. Multiscale assessment of patterns of avian species richness. *Proceedings of the National Academy of Sciences (USA)* **98**:4534–4539.
- Rejmánek, M. 2003. The rich get richer—responses. *Frontiers in Ecology and the Environment* **1**:122–123.
- Rosenzweig, M. L. 1995. *Species diversity in space and time*. Cambridge University Press, Cambridge, UK.
- Stevens, G. C. 1989. The latitudinal gradient in geographic range: how do many species coexist in the tropics? *American Naturalist* **133**:240–256.
- Stohlgren, T. J., D. T. Barnett, C. Flather, J. Kartesz, and B. Peterjohn. 2005. Plant species invasions along the latitudinal gradient in the United States. *Ecology* **86**:2298–2309.
- Stohlgren, T. J., D. T. Barnett, and J. T. Kartesz. 2003. The rich get richer: patterns of plant invasions in the United States. *Frontiers in Ecology and the Environment* **1**:11–14.
- Taylor, B. W., and R. E. Irwin. 2004. Linking economic activities to the distribution of exotic plants. *Proceedings of the National Academy of Sciences (USA)* **101**:17725–17730.
- Whittaker, R. J., K. J. Willis, and R. Field. 2001. Scale and species richness: toward a general, hierarchical theory of species diversity. *Journal of Biogeography* **28**:453–470.
- Withers, M. A., M. W. Palmer, G. L. Wade, P. S. White, and P. R. Neal. 1998. Changing patterns in the number of species in North American floras. Pages 23–32 in T. D. Sisk, editor. *Perspectives on the land-use history of North America: a context for understanding our changing environment*. USGS, Biological Resources Division, BSR/BDR-1998–2003. (<http://biology.usgs.gov/luhna/chap4.html>)

Ecology, 87(12), 2006, pp. 3213–3217
© 2006 by the Ecological Society of America

PLANT SPECIES INVASIONS ALONG THE LATITUDINAL GRADIENT IN THE UNITED STATES: REPLY

Curtis H. Flather,¹ Thomas J. Stohlgren,^{2,5}
Catherine Jarnevich,² David Barnett,³ and John Kartesz⁴

We welcome the opportunity to respond to the comments of our colleagues, Fridley et al. (2006), on our recent paper (Stohlgren et al. 2005) regarding plant species invasions along latitudinal gradients. We agree on many aspects of this important line of research. In fact, the two major findings that they report from their analysis of floras are consistent with our main points: (1) that latitude explains very little of variation in native species richness in the conterminous United States; and (2) that native plant species richness is one of the strongest predictors of nonnative species richness. We were not surprised that our conclusions were robust to a different data set and a different data transformation.

Manuscript received 11 April 2006; revised 14 June 2006; accepted 27 June 2006. Corresponding Editor: H. Hillebrand.

¹ USDA Forest Service, Rocky Mountain Research Station, Fort Collins, Colorado 80526 USA.

² National Institute of Invasive Species Science, U.S. Geological Survey, Fort Collins Science Center, Fort Collins, Colorado 80526 USA.

³ Natural Resource Ecology Laboratory, Colorado State University, Fort Collins, Colorado 80523 USA.

⁴ Biota of North America Program, University of North Carolina, Chapel Hill, North Carolina 27599 USA.

⁵ Corresponding author: E-mail: tom_stohlgren@usgs.gov

These acknowledgments notwithstanding, we are concerned with the assertion by Fridley et al. (2006) that our study was characterized by “uncritical evaluation of completeness in the data set” and that our “inappropriate data transformation” (simple species density) resulted in spurious patterns of covariation. We show that these two criticisms do not apply to our data and derive, in part, from their peremptory specification of the species–area functional form. We finish by commenting briefly on their analysis of latitudinal effects, the implications of using different floristic databases, and their overall findings, which largely confirm the most important patterns in our paper.

The issue of incomplete county records

It has long been known that the completeness of biotic surveys will affect how species counts covary with area (Williams 1943, Turner and Tjørve 2005). Therefore, we find ourselves in agreement with Fridley et al. (2006) that survey completeness can affect the inferences drawn from species–area data. However, we disagree with their characterization of the Biota of North America Program (BONAP) database as being composed of records that “... were not themselves collected with the aim of providing complete floristics lists of counties.” Indeed, decades of field research coupled with intense herbarium and museum specimen assessment from collections throughout North America, along with the critical review of vast quantities of published and unpublished monographs, floristic revisions, and county-level summaries of vascular plants were incorporated into BONAP to make our county-level data sets as “comprehensive” as possible. However, as understood by most floristicians, due to the complexity of floristics, even our smallest counties can never be completely inventoried. We freely admitted this in our paper stating: “Due to incomplete data on current species distributions, we assumed that the patterns of past introductions are only a first approximation of current introduction patterns” (Stohlgren et al. 2005:2299). In the absence of a probabilistic sample, it is unlikely that the issue of floristic completeness can be addressed in any kind of rigorous manner (Bunge and Fitzpatrick 1993, Stohlgren et al. 1997).

What we find most disconcerting about their claim that the patterns we report are an artifact of incomplete floristic surveys is that Fridley et al. (2006) offer no evidence that their floras (as compiled by Palmer [2005]) are, by some measure, more complete than those compiled in BONAP. We suspect that their failure to quantify completeness is due, in large part, to the absence of a “... widely accepted method to predict the completeness of a species list” (Palmer et al. 1995:343), at least as derived from typical floristic data. The patterns reported in Stohlgren et al. (2005) were based

upon nearly 2.3×10^6 county-level records of native species among counties with ≥ 100 native plant counts in the database (i.e., an average of 757 native plant species per county). This count appears comparable to the floras compiled by Palmer (2005). Regardless of the record count, unsubstantiated claims by Fridley et al. (2006) of floristic incompleteness (our data) or completeness (their data) do little to advance the difficult task of biodiversity assessments.

The criticisms of Fridley et al. (2006) would have benefited from evidence showing: (1) a systematic bias in the BONAP data set due to incomplete sampling of some counties; and (2) that the systematic bias in completeness significantly influenced our model results. If the claims of Fridley et al. (2006) were accurate, then one might predict that the patterns of covariation that we observed would change under a reanalysis focused on the most complete county floras. We have recently reanalyzed our data in such a fashion (Jarnevich et al. 2006) and found no evidence of bias attributable to survey incompleteness in BONAP. Jarnevich et al. (2006) relied on 177 counties with the most complete floras (307 162 occurrence records; an average of 1735 species per county). The 177 counties were selected as the four greatest county floras in 44 of the 48 conterminous states (Georgia, Maryland, and Mississippi had less complete statewide data at the time of analysis and were dropped from the data set, and we included one of the three counties that comprise the state of Delaware), and with the same simple density transformation the following corroborative results were found: (1) the same general patterns of covariation between native and nonnative species densities emerged, as we previously reported, with nearly identical models and predictor variables; and (2) the strongest predictor of county-level nonnative species density was native species density (fit with the same polynomial model as in Stohlgren et al. [2005]; $R^2 = 0.92$; $P < 0.0001$). No other factor or suite of factors came close to predicting nonnative species densities.

Thus, the modeling results presented in Stohlgren et al. (2005) were largely unchanged whether we used a small sample ($n = 177$) of well-studied counties or a large sample ($n = 3004$) of well-studied to moderately sampled counties.

Effect of area and area-corrected transformations

Fridley et al. (2006) correctly assert that species richness and area are often not linearly related but are commonly described by the power function $S = cA^z$ (Arrhenius 1921, Rosenzweig 1995). However, Fridley et al. (2006) fail to acknowledge that species–area relationship research has been largely phenomenological, and that assuming that the power function is the appropriate specification for all species–area data without testing other functional forms is perhaps

TABLE 1. Relationships of native-species richness per county (S) to area (A ; km²) using various common models.

Data set	Model	Specification	Fit statistics			Normality (ϵ)			
			R^2	F	P	A^2	P	Skewness (γ_1)	P
Full county (Stohlgren et al. 2005) ($n = 3004$)	untransformed	$S = c + zA$	0.161	577.2	<0.001	11.20	<0.005	0.460	<0.001
	semilog	$S = c + z \log(A)$	0.111	374.1	<0.001	9.73	<0.005	0.466	<0.001
	log-log	$\log(S) = c + z \log(A)$	0.128	439.1	<0.001	40.55	<0.005	-0.769	<0.001
Well-sampled county (Jarnevich et al. 2006) ($n = 177$)	untransformed	$S = c + zA$	0.395	114.2	<0.001	0.78	0.043	0.023	0.451
	semilog	$S = c + z \log(A)$	0.343	91.4	<0.001	1.29	<0.005	-0.124	0.324
	log-log	$\log(S) = c + z \log(A)$	0.356	96.8	<0.001	2.79	<0.005	-1.083	<0.001

Notes: Normality of errors (ϵ) was evaluated with the Anderson-Darling (A^2) test. Skewness (γ_1) is reported as an index of asymmetry and $H_0: \gamma_1 = 0$ (i.e., the distribution of ϵ is symmetric) is tested using the procedure defined by Sokal and Rohlf (1981:174).

symptomatic of “unwarranted confidence” (Connor and McCoy 1979:795). Past attempts to specify the form of the species–area relationships based on theory (Fisher et al. 1943, Preston 1962, MacArthur and Wilson 1963, Hubbell 2001) have, as yet, failed to offer a general explanation. Thus, the search for a canonical model specification has largely eluded ecology (Lomolino 2000, Tjørve 2003, Whittaker 2006), and it has been difficult to specify the functional form a priori (Williams 1964, Palmer and White 1994). In a reexamination of 100 species–area data sets, Connor and McCoy (1979:795) found that the species–area function was best fit by the untransformed model (i.e., species counts were linearly related to area) in 35 cases. Furthermore, only 36 of the remaining 65 data sets were best fit by the log–log approximation of the power function.

Given the uncertainty surrounding the functional form of the species–area relationship, we took an empirical approach to model specification rather than assuming a priori that any one functional form took precedence. In a paper completed prior to the publication of Stohlgren et al. (2005), we examined the effect of area on species richness (see Stohlgren et al. 2006) and evaluated a number of competing models. Here, we tested three competing linear specifications of the species–area relationship: untransformed, semilog, and log–log (Table 1). We used two general criteria to evaluate these specifications. First, we looked at indicators of model fit. Because the dependent variables across these three specifications change scale (arithmetic scale for the untransformed and semilog specification; log scale for the log–log specification), fit statistics are not directly comparable (Scott and Wild 1991, Anderson-Sprecher 1994). Comparability was achieved by backtransforming the predictions from the log–log model and estimating R^2 (using $R^2 = 1 - \frac{\sum_i^n [y_i - \hat{y}]^2}{\sum_i^n [y_i - \bar{y}]^2}$ as recommended by Kvålseth [1985]) and the F statistic in the original arithmetic units using the method of Duan (1983) to correct for backtransformation bias. We could have reported Akaike’s information criterion for each model using the equation given in Burnham and Anderson (2002:63) for the least-squares

case, but since it is derived from the residuals and the number of estimated parameters for each specification is the same, it resulted in identical model rankings.

Second, we examined the residuals for departure from the assumed normal distribution based on the Anderson-Darling test where larger values of the test statistic (A^2) indicate a greater departure from normality (Stephens 1974). Because least-squares regression is robust to deviations from normality as long as the errors are not severely asymmetric (Vasu 1979), we also looked at skewness (γ_1) to quantify the asymmetry in the error distribution (skewness = 0 for any symmetric distribution). These comparisons were carried out on the “full county” ($n = 3004$) and “well-sampled county” ($n = 177$) data sets using native species richness as the response variable.

We found the untransformed model to fit the county richness data relatively well, while the other two specifications resulted in poorer fits (Table 1). For native plants, area accounted for 16% and 40% of the variation in species counts for the full and well-sampled county data sets, respectively. For the semilog specification, area accounted for 11% and 34% of the variation in species counts; for the log–log specification, area accounted for 13% and 36% of the variation in species counts for the full and well-sampled county data sets, respectively.

Based on the Anderson-Darling test (A^2), there was evidence that error distributions were nonnormal for all model specifications, with the log–log specification deviating from normality to a greater degree than either the semilog or untransformed models (Table 1). Furthermore, the log–log specification had the greatest degree of asymmetry for both the full and well-sampled data sets. For the well-sampled data set, it is notable that the error distributions for the untransformed and semilog specifications did not deviate from a symmetric distribution ($P > 0.32$), while the errors under the log–log specification were asymmetrical ($P < 0.001$). Similar patterns were observed for exotic species, but exotic richness was only weakly correlated with area (full data set: $R^2 < 0.007$; well-sampled data set: $R^2 < 0.14$). We

concluded that: (1) an untransformed linear function between area and species richness fit the data better than either of the transformed specifications; and (2) the use of a simple density transformation (richness per area) was justified to remove the slight “area effect” we saw in the full county data set.

Latitudinal effects

We reanalyzed simple correlations between latitude and native plant richness using the transformation of Fridley et al. (2006) and our simple density transformation for 3004 counties. We restricted our analysis to the main effect of latitude and chose not to replicate the more complicated formulation in Fridley et al. (2006) involving third-order polynomial and interaction terms of both latitude and longitude (a formulation that is addressing much more than the simple latitudinal gradient that we addressed in our paper, and a formulation that is atypical of most latitudinal gradient analyses [see Rosenzweig 1995:25]). Latitude was significant ($P < 0.05$) but weak in all tests, explaining $\leq 2\%$ of the variation in native species regardless of the transformation used. Fridley et al. should not be surprised by weak positive or negative relationships between species richness and latitude given the synthesis by Hillebrand (2004) of nearly 600 latitude studies, especially considering the latitude range of 20° – 30° of our study (cf. Hillebrand 2004:203, Fig. 6). We were not surprised that by adding points in Canada and Alaska, Fridley et al. (2006) detected a stronger latitudinal gradient than we reported. However, this pattern (see Fig. 2 in Fridley et al.) is largely driven by the points above 48° latitude, which is an interesting but irrelevant “comment” on our analysis.

The contrasting data set of floras

Apart from survey completeness and data transformation, there are at least three other factors that may explain why the two analyses differ: (1) differences in the basic sample units; (2) differences in sample size (and the environmental gradients covered); and (3) the level of standardization of taxonomy. In Stohlgren et al. (2005) the observation unit is a county, whereas in Fridley et al. (2006) the observation unit is a vaguely defined county-sized unit. The sample sets likely contain differences within and across units in land use, land cover, physiography, climate, and other factors. The much smaller sample size of Fridley et al. (541 floras compared to 3004 counties) may not adequately represent the complex vegetation patterns in the United States that are affected by biotic, environmental/topographic, and human variables (where “area” plays a relatively minor role [Stohlgren et al. 2005, 2006]).

A potentially important factor accounting for the different findings is the underlying taxonomy used to generate species counts. Fridley et al. (2006) recognize

that “no flora is a perfect representation of the total taxonomic richness of a particular area” (also see Palmer et al. 1995). However, they fail to mention that all floras can be greatly affected by the taxonomic expertise of the field crews (usually over decades), the specific area surveyed or not surveyed, the level of intensity of the survey, the ability to detect the species in the area, and the taxonomic standards employed. We note that Palmer (2005:387) did not standardize his floras to a common nomenclature so that species names used in one flora, based on local taxonomic keys, may not be similar or consistent with the same taxon in another flora. Differences in taxonomic classification systems alone can alter species counts of a particular area by scores, if not hundreds, of entries. BONAP, on the other hand, has established and published its international taxonomic standard, which accounts for virtually all scientific names in use for the past century. A fundamental feature of our research is the ability to cross spatial scales (i.e., plot scales to subcontinental scales) using BONAP as a consistent taxonomic standard. Without a consistent taxonomic standard, accurate species richness comparisons between any two geographic areas are problematic.

A more definitive treatment of why the analyses presented in Stohlgren et al. (2005) and Fridley et al. (2006) differ would require a detailed comparison of the underlying criteria for accepting floras into the respective data sets, an analysis of the potential effects of using varying nomenclature, and subjecting the data set of Fridley et al. to the same competing species–area formulations as we have done.

In summary, the selection of an appropriate species–area curve is largely an empirical exercise. Theory does not unambiguously support any one species–area functional form. Therefore, an empirical examination of the data seems the most appropriate way to select a “best” species–area relationship. Based on our data, species richness appeared to covary with county area in a linear fashion, and area accounted for a small proportion of the total variation in plant species counts at the full county data set. Fridley et al. (2006) do not offer the fit statistics associated with competing species–area formulations based on their floras. In the absence of that comparison it is difficult for us to argue that their transformation is incorrect for their data set, but they are incorrect in asserting that their log–log transformation is appropriate for our data (Table 1).

A more interesting and urgent problem in ecology is quantifying patterns of nonnative species richness and abundance. To that end, we are finding much stronger evidence that simple native species richness and density are strong predictors of nonnative species richness and density at plot (Stohlgren et al. 2003), county (Stohlgren et al. 2005, Jarnevich et al. 2006), state (Stohlgren et al. 2006), and regional scales (Stohlgren et al. 2005).

Literature cited

- Anderson-Sprecher, R. 1994. Model comparisons and R^2 . *American Statistician* **48**:113–117.
- Arrhenius, O. 1921. Species and area. *Journal of Ecology* **9**:95–99.
- Bunge, J., and M. Fitzpatrick. 1993. Estimating the number of species: a review. *Journal of the American Statistical Association* **88**:364–373.
- Burnham, K. P., and D. R. Anderson. 2002. Model selection and multimodel inference: a practical information–theoretic approach. Second Edition. Springer, New York, New York, USA.
- Connor, E. F., and E. D. McCoy. 1979. The statistics and biology of the species–area relationship. *American Naturalist* **113**:791–833.
- Duan, N. 1983. Smearing estimate: a nonparametric retransformation method. *Journal of the American Statistical Association* **78**:605–610.
- Fisher, R. A., A. S. Corbet, and C. B. Williams. 1943. The relation between the number of species and the number of individuals in a random sample of an animal population. *Journal of Animal Ecology* **12**:42–58.
- Fridley, J. D., H. Qian, P. S. White, and M. W. Palmer. 2006. Plant species invasions along the latitudinal gradient in the United States: Comment. *Ecology* **87**:3209–3213.
- Hillebrand, H. 2004. On the generality of the latitude diversity gradient. *American Naturalist* **163**:192–211.
- Hubbell, S. P. 2001. The unified neutral theory of biodiversity and biogeography. Princeton University Press, Princeton, New Jersey, USA.
- Jarnevich, C. S., T. J. Stohlgren, D. Barnett, and J. Kartesz. 2006. Filling in the gaps: modelling native species richness and invasions using spatially incomplete data. *Diversity and Distributions* **12**:511–520.
- Kvålseth, T. O. 1985. Cautionary note about R^2 . *American Statistician* **39**:279–285.
- Lomolino, M. V. 2000. Ecology's most general, yet protean pattern: the species–area relationship. *Journal of Biogeography* **27**:17–26.
- MacArthur, R. H., and E. O. Wilson. 1963. An equilibrium theory of insular zoogeography. *Evolution* **17**:373–387.
- Palmer, M. W. 2005. Temporal trends of exotic species richness in North American floras: an overview. *Écoscience* **12**:386–390.
- Palmer, M. W., G. L. Wade, and P. R. Neal. 1995. Standards for the writing of floras. *BioScience* **45**:339–345.
- Palmer, M. W., and P. S. White. 1994. Scale dependence and the species–area relationship. *American Naturalist* **144**:717–740.
- Preston, F. W. 1962. The canonical distribution of commonness and rarity: Part I. *Ecology* **43**:185–215.
- Rosenzweig, M. L. 1995. Species diversity in space and time. Cambridge University Press, Cambridge, UK.
- Scott, A., and C. Wild. 1991. Transformations and R^2 . *American Statistician* **45**:127–129.
- Sokal, R. R., and F. J. Rohlf. 1981. Biometry: the principles and practice of statistics in biological research. Second edition. W. H. Freeman, San Francisco, California, USA.
- Stephens, M. A. 1974. EDF statistics for goodness of fit and some comparisons. *Journal of the American Statistical Association* **69**:730–737.
- Stohlgren, T. J., D. Barnett, C. Flather, P. Fuller, B. Peterjohn, J. Kartesz, and L. L. Master. 2006. Species richness and patterns of invasion in plants, birds, and fishes in the United States. *Biological Invasions* **8**:427–447.
- Stohlgren, T. J., D. Barnett, C. Flather, J. Kartesz, and B. Peterjohn. 2005. Plant species invasions along the latitudinal gradient in the United States. *Ecology* **86**:2298–2309.
- Stohlgren, T. J., D. Barnett, and J. Kartesz. 2003. The rich get richer: patterns of plant invasions in the United States. *Frontiers in Ecology and the Environment* **1**:11–14.
- Stohlgren, T. J., G. W. Chong, M. A. Kalkhan, and L. D. Schell. 1997. Rapid assessment of plant diversity patterns: a methodology for landscapes. *Ecological Monitoring and Assessment* **48**:25–43.
- Tjørve, E. 2003. Shapes and functions of species–area curves: a review of possible models. *Journal of Biogeography* **30**:827–835.
- Turner, W. R., and E. Tjørve. 2005. Scale-dependence in species–area relationships. *Ecography* **28**:721–730.
- Vasu, E. S. 1979. Non-normality in regression analysis: Monte Carlo investigation under the condition of multicollinearity. Working papers in methodology. Number 9. Institute for Research in Social Science, University of North Carolina. Chapel Hill, North Carolina, USA.
- Whittaker, R. J. 2006. Island species–energy theory. *Journal of Biogeography* **33**:11–12.
- Williams, C. B. 1943. Area and number of species. *Nature* **152**:264–267.
- Williams, C. B. 1964. Patterns in the balance of nature: and related problems in quantitative ecology. Academic Press, New York, New York, USA.