

The Shapes of Phylogenetic Trees of Clades, Faunas, and Local Assemblages: Exploring Spatial Pattern in Differential Diversification

Stephen B. Heard* and Graham H. Cox†

Department of Biology, University of New Brunswick, Fredericton, New Brunswick E3B 6E1, Canada

Submitted August 25, 2006; Accepted December 13, 2006;
Electronically published March 1, 2007

Online enhancements: data files.

ABSTRACT: Life on Earth is characterized by strong diversity skewness: related lineages typically show pronounced variation in diversification success, and clades contain hyperdiverse and depauperate subclades. Previous studies have documented diversity skewness only for entire (global) clades. We demonstrate methods for measurement and significance testing of diversity skewness of local assemblages and regional biotas; we illustrate this with an analysis of geographic structure in diversity skewness of primate assemblages. For primates, continental faunas differ in diversity skewness from expectations based on the global phylogeny: South American faunas have significantly low skewness and African faunas have significantly high skewness. However, no local assemblage has diversity skewness different from that expected based on sampling the continental fauna. We also document a latitudinal gradient in diversity skewness for the African assemblages and test for (but do not find) associations of skewness with longitude, local species richness, and net primary productivity. Our data suggest that continental-scale biogeographic events rather than local-scale processes have shaped diversity skewness in modern primate faunas.

Keywords: diversity skewness, tree balance, community structure, primates, phylogeny.

One of the most conspicuous attributes of life on Earth is the astonishing unevenness in biodiversity (“diversity skewness”) among major clades. Every biologist can name clades that seem hyperdiverse (orchids, cichlid fishes) and

clades that seem depauperate (tuataras, ginkgos). Recognition of diversity skewness dates at least to Darwin ([1859] 1968), who commented on the tendency for taxonomic groups to be dominated by a few diverse subtaxa. For instance, within the praying mantises (Insecta: Mantodea), a few large genera make up much of the total diversity, and a frequency plot of generic sizes shows the long-tailed hollow curve (fig. 1) typical of such plots (and which inspires the term “diversity skewness”).

More recently, the modern explosion of phylogenetic data has motivated an explicitly evolutionary perspective in which diversity skewness is quantified using the topology of phylogenetic trees (Heard 1992; Mooers and Heard 1997). When diversity skewness is low, all lineages have had similar diversification success, and the phylogeny is balanced (internal nodes join subclades of similar diversity; fig. 1, *inset, left*). When diversity skewness is high, some lineages have diversified much more than others, and the phylogeny is imbalanced (internal nodes join subclades of very different diversity; fig. 1, *inset, right*). Empirical studies quantifying imbalance in compilations of real phylogenies have led to a strong consensus that high diversity skewness is a fundamental property of Earth’s biota, apparent at all phylogenetic scales and in most studied taxa (for reviews, see Mooers and Heard 1997; Purvis and Agapow 2002; Mooers et al. 2007).

Previous studies have considered diversity skewness of entire (global) clades (Mooers et al. 2007), and any species omitted have been seen only as “incompleteness” and a possible source of tree-shape bias (e.g., Mooers 1995). Such analyses do not consider spatial pattern in skewness at any scale and contrast sharply with the study of species richness, which has long been replete with analyses of geographic structure at multiple spatial scales (e.g., Rosenzweig 1995; Davies et al. 2005). Here, we extend the analysis of diversity skewness from global clades to regional and local scales. We ask whether there are regional or local patterns in diversity skewness, and we test these patterns

* Corresponding author; e-mail: sheard@unb.ca.

† E-mail: m6s78@unb.ca.

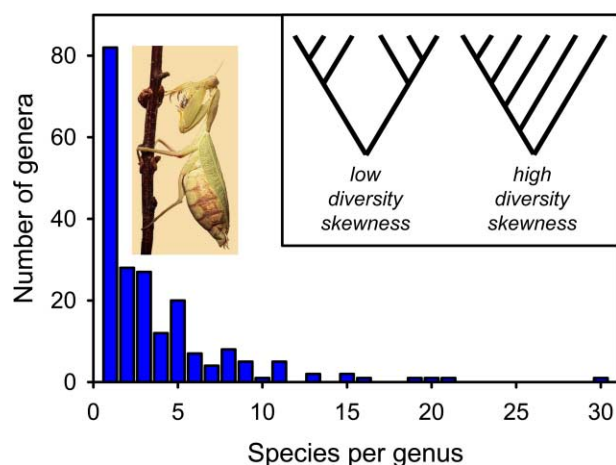


Figure 1: Taxonomic and phylogenetic illustrations of diversity skewness. Diversity skewness in taxonomy of mantids (Mantodea); data from Williams (1964, table 49). *Inset*, diversity skewness in phylogenetic trees. The left-hand tree has low diversity skewness ($I_c = 0.2$), while the right-hand tree has very high diversity skewness ($I_c = 1$). Photo courtesy of Jack Scott/Bio-Ditrl (University of Alberta).

statistically against expectations based on diversity skewness at larger geographic scales.

To introduce the analysis, we begin with a simple example (fig. 2). We know the global phylogeny for a clade of interest, and we have presence/absence data for members of that clade in a number of local assemblages (biotas or communities). For each local assemblage, we can derive from the global phylogeny a smaller phylogenetic tree containing just the members of that assemblage (fig. 2, *right*). We refer to these as local phylogenies, and they have shapes that represent the diversity skewness present in each local assemblage. With an appropriate metric for diversity skewness (see “Methods”), we can now ask two questions.

(1) *Do local assemblages have surprisingly high (or low) diversity skewness?* Local diversity skewness can only be surprisingly high or low by comparison with a null expectation: that is, with the local diversity skewness expected in the absence of any local process that might operate to increase or decrease skewness. There are two potentially interesting nulls in any study of diversity skewness (and we apply both). First, the “phylogenetic null” compares diversity skewness of a set of species with that expected in a monophyletic clade evolving under the so-called equal-rates Markov (ERM) null model. The ERM is a model of macroevolution in which all lineages have equal diversification rates (Moers and Heard 1997), and so a test of a monophyletic clade against the phylogenetic null is a test for differential diversification among lineages. This is the test used in previous studies of (global) diversity skewness (for a review, see Moers and Heard 1997), and

it is appropriately applied to a global clade or to any local or regional assemblage that is also a monophyletic clade. The phylogenetic null, however, does not address questions about local controls on diversity skewness, and it is inappropriate for most assemblages because we can reject it out of hand whenever an assemblage is not a monophyletic clade. The “biogeographic null,” in contrast, compares diversity skewness of a set of species with that expected for random draws from a source pool (larger biota). A test against the biogeographic null asks whether the shape of the local phylogeny is unusual compared with the shapes of communities assembled at random from “available” species in the larger biota. This test takes into account the phylogenetic structure of the global clade, which constrains the ways in which local assemblages can be structured. Most of our tests for significance of diversity skewness are constructed against the biogeographic null, as our principal aim is to test for local or regional processes that might pull diversity skewness of assemblages away from that characterizing the global clade.

As a first step in a test against the biogeographic null, we can compare each local phylogeny with phylogenies of assemblages (having equal species richness) formed by random draws from the global clade. If the real assemblage’s skewness falls outside the 95% envelope for randomly drawn assemblages, then some nonrandom process is inferred to stand between the global clade and its local representation. However, this simple local-versus-global analysis neglects the regional biogeographic context of local assemblages. A more sophisticated approach would recognize regional biotas as source pools for local assemblages, where “regional” biotas might correspond to biogeographic provinces, continental biotas, or biotas of smaller areas such as climate or habitat zones. We could then ask whether local assemblages have surprising diversity skewness compared with random draws from the regional biota, and in turn whether regional biotas have surprising skewness compared with random draws from the global clade. This decomposes question (1) so as to partition diversity skewness among local, regional, and global scales. This decomposition is not limited to three hierarchical levels, since any number of nested biotas could be recognized and tested for significant diversity skewness. Of course (as for any such spatial decomposition), patterns in skewness may occur at more than one spatial scale, and patterns at one scale may be reinforced or countered by patterns at other scales.

(2) *Among local assemblages, can variation in diversity skewness be explained by environmental factors or geographical gradients?* Both geography (e.g., latitude) and environment predict variation in species diversity and other aspects of community structure, and discussion of these patterns and their mechanisms has a very long history in

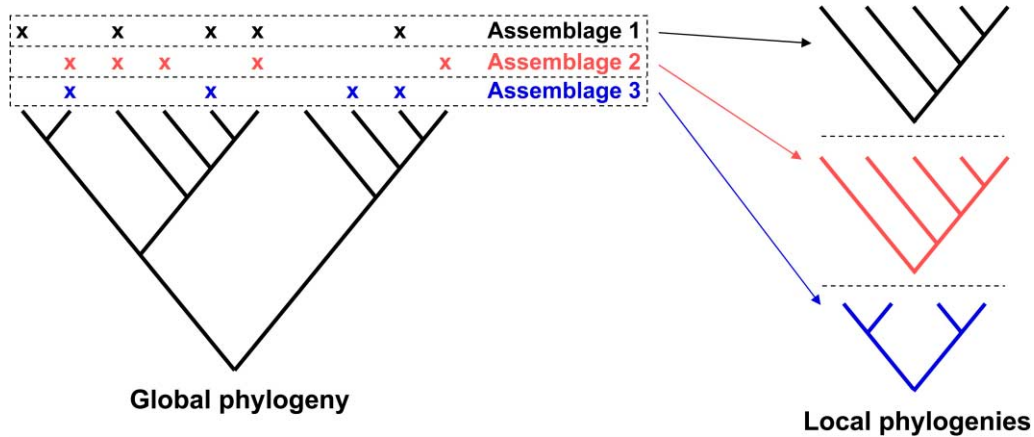


Figure 2: Hypothetical global and local phylogenies. The global clade has 10 species, with three local assemblages having five, five, and four of those species. Each local phylogeny is extracted from the global phylogeny by deleting branches and nodes not locally represented (or tracing paths from the represented species to their union at a single node).

ecology (Rosenzweig 1995; Field et al. 2005; Scheiner and Willig 2005). However, no previous study has considered geographic or environmental patterns in diversity skewness. One could test for latitudinal gradients in diversity skewness or for correlations of diversity skewness with environmental productivity or harshness (and we illustrate both types of analysis).

We illustrate the hierarchical analysis of diversity skewness with an example using primate faunas of large African and South American protected areas (national parks and similar preserves). In this analysis, park faunas are the local assemblages, and we treat continents as regions. We look forward to similar analyses of other phylogenies and other assemblage data, which will build a more general picture of spatial pattern in diversity skewness on Earth.

Methods

Quantifying Diversity Skewness

We quantified diversity skewness using Colless's index of phylogenetic tree imbalance, I_c (Heard 1992), which takes values from 0, for a perfectly balanced phylogeny (low diversity skewness; fig. 1, *inset*), to 1, for a perfectly imbalanced phylogeny (high skewness; fig. 1, *inset*). The I_c is the normalized sum of differences in species richness between the two subclades defined by each internal node:

$$I_c = \frac{2}{(n-1)(n-2)} \sum_{\text{nodes}} |S_R - S_L|, \quad (1)$$

where there are n species in the tree and the right and left

branches at a node define subclades of S_R and S_L species. There are many alternative measures of diversity skewness, but I_c is simple, intuitive, and powerful (Agapow and Purvis 2002; Blum and Francois 2005).

For a monophyletic clade, a test of diversity skewness against the phylogenetic null (equal-rates Markov) is straightforward using the known expectation for I_c (Heard 1992; Rogers 1994; Blum et al. 2006) and confidence intervals (CIs) based on simulations (Heard 1992). (We address a minor complication surrounding resolution of polytomies below, in the context of the biogeographic null.) Unfortunately, tests of single phylogenies against the phylogenetic null tend to be of relatively low power because the distribution of expected tree shapes is flat and, thus, the 95% CIs are broad (Moore and Heard 1997; S. B. Heard, unpublished data). As a result, in a test of a single phylogeny against the phylogenetic null, only a very extreme topology is likely to have significantly high (or low) skewness. For local or regional assemblages, tests of skewness against the biogeographic null are more powerful but are also somewhat more complex.

Assessing Local Diversity Skewness

The Basic Method. Given a global phylogeny and a species list for a local assemblage, measuring local diversity skewness is simple in principle. The local phylogeny is the phylogeny of just the species represented in the local assemblage (fig. 2), and the diversity skewness of the local assemblage is then represented by $I_{c(\text{local})}$. For significance testing against the biogeographic null, $I_{c(\text{local})}$ is compared with a distribution of $I_{c(\text{random})}$ values. For the first test (local assemblage vs. global phylogeny), this distribution

is obtained by calculating I_c for 10,000 assemblages of n_{local} species drawn at random from the global phylogeny. A P value for $I_{c(\text{local})}$ is then determined from the ordered $I_{c(\text{random})}$ values, as is usual for Monte Carlo tests (Manly 1997). The spatial decomposition of this test, yielding comparisons of local versus regional skewness and regional versus global skewness, is straightforward, with each comparison having the smaller-scale I_c tested against random draws from the larger-scale biota.

Dealing with Polytomies. Many available phylogenies contain polytomies, but I_c cannot be calculated for such phylogenies. We believe that few polytomies in published phylogenies truly indicate simultaneous divergence of three or more lineages. Rather, most polytomies reflect our incomplete knowledge of relationships and in principle should be resolved. Some workers have dealt with polytomies by resolving them randomly (e.g., Housworth and Martins 2001; Wiens et al. 2006), but this would be inappropriate here because no simple distribution from which random resolutions could be drawn provides an acceptable representation of diversity skewness in real clades (Mooers and Heard 1997). We therefore took an alternative approach that yields upper and lower bounds for I_c and P values.

For any polytomy in any phylogeny (global or local), we can identify the possible resolution with the highest diversity skewness and the resolution with the lowest diversity skewness. Now we can calculate $I_{c \text{ min}}$, the lower bound on diversity skewness, with all polytomies in their lowest skewness resolution (the desired resolution of each polytomy is independent of any others because the diversity contrasts in I_c are calculated one node at a time). The upper bound on skewness, $I_{c \text{ max}}$, is calculated similarly. For tests against the phylogenetic null, these bounds can be used directly, and interpretation is straightforward if both bounds on I_c are above the null's upper or below its lower 95% confidence limit.

Tests of skewness against the biogeographic null require a slightly more complicated procedure. For a local phylogeny with one or more polytomies, we begin by calculating $I_{c(\text{local}) \text{ min}}$ and $I_{c(\text{local}) \text{ max}}$. Next, for the 10,000 random assemblages, we gather two distributions: that of $I_{c(\text{random}) \text{ min}}$ and that of $I_{c(\text{random}) \text{ max}}$; any random assemblage without a polytomy has its $I_{c(\text{random})}$ included in both. Random assemblages may include polytomies even when no local assemblage does, but this poses no problem. Finally, we combine the two bounds on $I_{c(\text{local})}$ with the two distributions for $I_{c(\text{random})}$. Imagine that $I_{c(\text{local})}$ is larger than the median $I_{c(\text{random})}$. For a one-tailed test of excess diversity skewness, first we compare $I_{c(\text{local}) \text{ min}}$ with the distribution of $I_{c(\text{random}) \text{ max}}$, recording the fraction of randomizations for which $I_{c(\text{random})} > I_{c(\text{local})}$. This is a conservative test, giving

the upper bound on the desired P value. Second, we compare $I_{c(\text{local}) \text{ max}}$ with the distribution of $I_{c(\text{random}) \text{ min}}$, recording the same fraction. This is a liberal test, giving the lower bound on the desired P value. For a two-tailed test, each bound is doubled. When $I_{c(\text{local})}$ is smaller than the median $I_{c(\text{random})}$, the procedure is identical except that the first comparison generates the lower bound and the second generates the upper bound. We report the two bounds on the true P value, for which interpretation will be clear when $P_{\text{min}} > \alpha$ (neither bound makes the test significant) or $P_{\text{max}} < \alpha$ (either bound makes the test significant) but ambiguous otherwise. The procedure for regional phylogenies, of course, is the same.

Software. We conducted our analyses using the software package SkewMatic 2.01, written by S. B. Heard in Microsoft Visual Basic.NET for Windows. SkewMatic 2.01 and instructions for its installation are available in the appendix. We conducted other statistical tests in SAS, version 8.0 (SAS Institute, Cary, NC).

Primate Phylogeny

We illustrate the hierarchical analysis of diversity skewness with African and South American assemblages of the mammalian order Primates. We used a new, nearly complete supertree for the order (Vos 2006) that synthesizes available phylogenetic information (374 source trees) for the primates and is almost fully resolved. It contains 219 primate species, following the nomenclature of Wilson and Reeder (1993). We adopted the same nomenclature, and its relatively conservative species concept, in all cases. Nodes deep in Vos's (2006) tree, to which our index of diversity skewness is most sensitive, have very high support. Most shallower nodes, at the genus level or below, are also well supported (Vos 2006), but there are some exceptions within the langurs and leaf monkeys (*Trachypithecus*, *Presbytis*), none of which occur in the local assemblages we analyze, and the galagos (*Euoticus*, *Galago*, *Galagoides*, *Otolemur*), present in the African assemblages.

As for any phylogenetic estimate, there is some uncertainty in the topology of the primate phylogeny. In particular, biases in tree balance (and hence diversity skewness) may sometimes be introduced during phylogenetic estimation. Such biases are probably mild for most conventional phylogenies (Mooers and Heard 1997) but may be more important for supertrees (Wilkinson et al. 2005). Fortunately, this should not be fatal to our analyses, because in tests against the biogeographic null, any biases should have acted equally on the local phylogenies and the global phylogeny from which they are sampled.

Primate Species Lists

We compiled lists of primate occurrences for large protected areas in Africa and South America. We worked with protected areas recognized by the International Union for Conservation of Nature and Natural Resources (IUCN; <http://www.unep-wcmc.org/wdpa>). These include national parks, nature reserves, forest reserves, biosphere reserves, and a variety of other designations, but all are recognized as dedicated to and managed for biodiversity conservation (IUCN 1994). For simplicity, we refer to all these areas as “parks.” Our data set included all IUCN-recognized terrestrial parks for which protection status and area data were available and that exceeded 900,000 ha.

For each park in our data set, we assembled a list of primate species from the best available distribution maps (Africa [<http://www.gisbau.uniroma1.it/amd/index.htm>]: Instituto Ecologia Applicata 1998; Boitani et al. 1999; South America [<http://www.natureserve.org/getData/mammalMaps.jsp>]: Patterson et al. 2005). We included a primate species in the park’s species list if its geographic range intersected the park (extent-of-occurrence rather than area-of-occupancy data; Gaston 1991). We preferred extent-of-occurrence data because they yielded species lists that were complete, authoritative, and fully comparable among parks. Occupancy data (lists of species actually recorded in the park), in contrast, were available for some parks but not others, and where they were available they varied greatly in apparent quality and authority.

We studied park faunas rather than alternatives (faunas for random points or points on a grid) for two reasons. First, because parks and their faunas are often well studied, we had higher confidence in geographic range data for the vicinity of parks than for elsewhere. Second, many primates are under anthropogenic population pressure in (or have been extirpated from) parts of their ranges. As a result, primate assemblages of protected areas should be more faithful to “real” assemblages as they existed in the predisturbance past.

Our choice of a large park size threshold (900,000 ha) has both benefits and costs. Larger parks are more likely to include all local habitats and so are less likely to be missing habitat specialists that occur in the vicinity but do not find their habitats represented in the park. Furthermore, when parks occur near range boundaries, larger parks are more likely to have broad overlap with the range, minimizing the chance that error in GIS databases or range data will lead to errors in species lists. However, preferring large parks comes at the cost of limiting our inferences about processes operating at the most local scales (truly local communities in which individuals of all species potentially interact). In particular, the species list for a large park can include species that occupy distinct habitats and

never co-occur on the spatial scales that define interspecific interactions. The list for Manu National Park (Peru) is a good example, with the most diverse local assemblages probably including about 10 of the park’s 12 species. The difference arises from species pairs such as *Lagothrix cana* and *Sanguinus imperator*, which are known only from opposite sides of the River Manu (J. Terborgh, personal communication). Our analyses cover a spectrum of areas from global down to the size of our smallest parks ($\approx 10^7$ ha), and we anticipate that future analyses using reliable species lists for very small areas will complement the analyses we report here.

We omitted one primate, *Homo sapiens*, from all species lists. For South America, *Homo* is excluded as a recent (invasive) arrival. For Africa, inclusion or exclusion of *Homo* did not affect any of our results. We also ignored primate populations resulting from human introductions.

Lists of parks and primates in Africa and South America used in our analyses are available as Excel or tab-delimited ASCII files.

Local and Regional Primate Diversity Skewness

We assigned each primate species a continental affiliation (Instituto Ecologia Applicata 1998; Patterson et al. 2005). We considered four continental regions: South America (including Central America), Madagascar, mainland Africa (henceforth just “Africa”), and Asia. Primate faunas of South America and Madagascar represent monophyletic radiations (Vos 2006), whereas Asian and African faunas are paraphyletic. Only one species crosses continental boundaries: the baboon *Papio hamadryas*, which is primarily African but is also found in the Arabian peninsula. We treated this species as African; all other Asian species have ranges much further east.

Using SkewMatic 2.01, we calculated local diversity skewness for each park’s primate assemblage, and we calculated regional skewness for the primates of each continental region. We conducted significance testing of $I_{c(\text{local})}$ against the biogeographic null for the subset of parks that had at least eight primate species (four in Africa, 18 in South America), because individual tests for smaller assemblages have little power (see “Data Requirements and Prospects for Further Analyses”). We first tested $I_{c(\text{local})}$ for each park directly against the global primate phylogeny and then decomposed this test into two hierarchical tests: park versus continental region and continental region versus global phylogeny. We also tested the Asian and Madagascar continental regions versus the global phylogeny, although we had no local assemblage data for those regions. We also tested diversity skewness versus the phylogenetic null for three monophyletic clades of interest: all

primates (global phylogeny) and the South American and Madagascar radiations. All these tests were two tailed.

Because we performed tests for multiple parks in each continent, interpretation of individual P values requires some caution. We had no a priori interest in making tests for any particular park, seeing parks instead as a set of replicate samples. Therefore, rather than apply sequential Bonferroni adjustment, we carried out a single combined test for parks of each continent using Fisher's method for combining probabilities (Sokal and Rohlf 1995). Within each continent there was a consistent direction for the skewness effect, and so we based these combined tests on one-tailed probabilities (but two-tailed tests yielded the same results).

Finally, to illustrate the search for geographic and environmental correlates of diversity skewness, we tested for associations of local skewness (raw I_c score) with net primary productivity (NPP) and for latitudinal and longitudinal gradients in local skewness separately for African and South American parks. We used NPP estimates for the approximate centroid of each park (NASA Earth Observatory, 2002 annual average; <http://earthobservatory.nasa.gov/Observatory/datasets.html>). We quantified associations using Pearson correlation coefficients and tested for latitudinal gradients with raw latitude values and with absolute value of latitude (to test for gradients peaking at the equator). Here we are not concerned with significance testing of I_c for individual parks and so were able to use parks with smaller primate assemblages (five or more species; 11 African and 24 South American parks). All of our results were robust to other choices of species-richness cutoff.

Results

Diversity Skewness of the Primate Clade and of Local Primate Assemblages

The primate supertree has rather low diversity skewness (overall $0.039 < I_c < 0.040$, depending on the resolution of several small polytomies). We cannot reject the phylogenetic null for the primates as a whole: for $n = 219$ species, $E(I_c) = 0.040$, 95% CI = 0.028–0.058.

When compared with expectations based on global phylogeny (using the biogeographic null), local diversity skewness showed significant signal for both African and South American parks. In Africa, skewness was significantly higher than expected for Maiko and suggestively so for Salonga (table 1). Across the four African parks with eight or more primates, skewness was generally higher than expected, and this trend was significant ($\chi^2 = 26.5$, $df = 8$, $P = .0009$). In South America, skewness was significantly lower than expected for Noel Kempff Mercado and

suggestively so for Amazonia and Puiawai (table 1). Across the 18 South American parks with eight or more primates, skewness was generally lower than expected, and this trend was significant ($\chi^2 = 78.3$, $df = 36$, $P < .0001$).

Further analysis, however, revealed that the diversity skewness signals were predominantly continental rather than local. When tested against the appropriate continental fauna, no park had skewness significantly different from what was expected (table 1). In contrast, the continental faunas differed in skewness from expectations based on global phylogeny (table 1): the African fauna had significantly high skewness and the South American fauna had significantly low skewness. Such continental signatures are also apparent for Asia (significantly high skewness) and for Madagascar (suggestively, but not quite significantly, high skewness). When tested instead against the phylogenetic null, neither monophyletic radiation departed significantly from the equal-rates null model, although in both cases the deviations were in the same direction as for the biogeographic null (Madagascar: $I_c = 0.25$, ERM 95% CI 0.097–0.31; South America: $I_c \approx 0.06$, ERM 95% CI = 0.055–0.14).

Geographic Pattern in Local Diversity Skewness

In Africa, local diversity skewness was strongly correlated with latitude (fig. 3; $r = 0.88$, $P = .0004$), with northern parks having higher diversity skewness. Skewness did not peak at the equator: for absolute latitude, $r = 0.14$, $P = .68$. African parks showed no longitudinal pattern ($r = -0.30$, $P = .37$), and South American parks showed no pattern with either latitude ($r = -0.17$, $P = .42$) or longitude ($r = -0.20$, $P = .36$). Local skewness was not significantly correlated with species richness on either continent (Africa: $r = -0.28$, $P = .42$; South America: $r = -0.31$, $P = .14$).

The African latitudinal gradient was partly attributable to variation among parks in the number of species from the basal family Lorisidae (galagos, pottos, and relatives). The northernmost parks had one or two lorises, while the southernmost had two to four (fig. 3, numbers beside data points), and local diversity skewness was negatively correlated with lorisid richness ($r = -0.78$, $P = .004$). However, lorises did not drive the entire gradient: the partial correlation between skewness and latitude, adjusting for lorisid diversity, was weaker than the raw correlation but remained significant ($r = 0.65$, $P = .04$). No other single clade made a major contribution to the skewness gradient. Local diversity skewness was not correlated with net primary productivity on either continent (Africa: $r = -0.48$, $P = .13$; South America: $r = -0.11$, $P = .60$).

Table 1: Diversity skewness of primate assemblages for African and South American parks and primate faunas of four continents

	Richness	I_c	P (vs. world)	Skewness direction ^a	P (vs. continent)
Park assemblages:					
Africa:					
Bamingui	6	1.0			
Comoe	8	.67	.21	High	.7
Kemeia	6	.20			
Maiko	20	.40	.004	High	.4
Manova	6	1.0			
Ruaha	5	.33			
<i>Salonga</i>	14	.46	<i>.064–.066^b</i>	High	.64
Serengeti	6	.50			
Southern	5	1.0			
Tsavo East	5	.33			
Upemba	12	.31	1		.41
South America:					
Alto Orinoco Casiquiare	10	.22	.31		.93
<i>Amazonia</i>	10	.11	<i>.072</i>	Low	.24
Bahuaja Sonene	9	.21	.35		.84
Canaima	6				
Central Suriname	8	.14	.1	Low	.32
Chiribiquete	9	.21	.35		.84
El Caura	7				
Imataca	6				
Isiboro Secure	7				
Jau	9	.21	.35		.84
Madidi	9	.21	.35		.84
Manu	12	.24	.59		.97
Manuripi Heath	14	.19	.3		1
Noel Kempff Mercado	12	.07	.012	Low	.068
Otuquis	7				
Pacaya Samiria	11	.24	.52		1
Parima Tapirapeco	10	.22	.31		.93
Pico de Neblina	9	.21	.35		.84
<i>Puinawai</i>	10	.11	<i>.072</i>	Low	.24
Rio Pure	12	.16	.12	Low	.56
San Matias	6				
Serrania de la Neblina	9	.21	.35		.84
Sipapo	8	.43	1		.64
Yasuni	14	.17	.18	Low	.76
Continental faunas:					
Africa	52	.18	<.0002	High	
South America	80	.06	.04–.06^b	Low	
Asia	57	.15–.16 ^b	<.0002	High	
Madagascar	28	.25	<i>.05–.052^b</i>	High	

Note: Parks with four or fewer primate species (three South American, six African) are not shown; parks with five to seven species are shown but without individual tests of significance (see “Methods”). P values are two-tailed tests. Boldface indicates $P < .05$; italic indicates $.05 < P < .10$.

^a Skewness of local phylogeny compared with random-sampling expectation. Direction is not shown for parks with very weak skewness patterns ($P > .25$).

^b Ranges arise because of polytomies in the supertree (see “Dealing with Polytomies”).

Discussion

Our methods allow the decomposition of diversity skewness into global, regional, and local components, with significance testing against the biogeographic null possible at each spatial scale. Our analyses complement other recent investigations of phylogenetic aspects of community structure (Webb et al. 2002; Anderson et al. 2004; Cavender-Bares et al. 2004; Stephens and Wiens 2004; Silvertown et al. 2006).

There are two conspicuous features of our primate diversity skewness data. First, significant skewness (deviation from the biogeographic null) exists at the continental scale but not the local scale; no park in our data set has a fauna with skewness either higher or lower than expected based on random draws from the appropriate continental assemblage. This does not mean that local primate faunas are actually assembled randomly and does not preclude local pattern in other aspects of primate community struc-

ture. It does mean, however, that there is no evidence for local processes affecting the diversity skewness of park faunas. In contrast, all four continental regions deviate in skewness from the biogeographic null: Africa, Asia, and Madagascar have significantly (or nearly so) high skewness and South America has significantly low skewness. Second, despite the lack of significant diversity skewness in any individual assemblage, we were able to detect among-assemblage pattern in the form of an African latitudinal gradient.

Interpreting Spatial Scale in Diversity Skewness Patterns

What might the continental scale of signal in diversity skewness tell us about how clades evolve and how local faunas are assembled? At a global scale, diversity skewness arises from among-lineage variation in speciation and/or extinction rates, but at more local scales, skewness can also arise from among-lineage variation in local representation: that is, from a tendency for local assemblages to include disproportionately many (or few) species drawn from particular lineages. Among processes that can produce diversity skewness, we can recognize some that should act primarily at global scales and others that should act primarily at regional or local scales.

Global Skewness. Several potential drivers of global-scale diversity skewness are known from modeling efforts (for a review, see Mooers et al. 2007), including elevated speciation rates in new lineages and speciation- or extinction-rate variation arising because those rates are linked to heritable (ancestor to descendent) traits of species. However, in the case of the primates, we see no evidence for such processes: overall primate I_c is exactly that expected under the phylogenetic null. Importantly, our analyses establish that a lack of excess diversity skewness in a global phylogeny can conceal strong diversity skewness at smaller geographic or taxonomic scales.

Regional Skewness. Regional patterns in diversity skewness can arise either through diversification or through faunal assembly. Geography can shape skewness through diversification in at least two different ways. First, as a result of phylogenetically conserved traits, different subclades in a global phylogeny might achieve diversification success in different areas. For instance, plant lineages that have evolved C_4 photosynthesis tend to radiate in the tropics but have few boreal representatives, while their C_3 relatives often show the opposite pattern (Kellogg 1999; Sage et al. 1999). This could lead to both boreal and tropical floras with much higher diversity skewness than the global phylogeny formed by their union, and cases such as this may be quite common. Second, differential rates of diversifi-

cation among subclades might arise from geographic controls on speciation or extinction rates. Strong diversity skewness might arise, for instance, from allopatric speciation along linear features such as island chains if repeated speciation events occur in connection with stepwise dispersal (e.g., Mendelson and Shaw 2005). Geographic controls could also lead to regional faunas with very low skewness; if vicariance events simultaneously divide the ranges of many subclades, leading to speciation in each, then diversity will increase similarly across subclades and diversity skewness will be low.

Regional diversity skewness may also reflect processes constraining the assembly of faunas. This is most obvious when major barriers to dispersal mean that regional faunas are assembled without representation of lineages evolving on the other side of those barriers, making regional skewness quite different from that expected based on geographically naive sampling from the global clade. More generally, regional patterns in skewness can arise whenever radiations are geographically limited and especially where some members of a radiation extend their ranges further than others (see “The African Latitudinal Gradient in Diversity Skewness”).

For the primates, we were able to demonstrate significant continental pattern in diversity skewness (against the biogeographic null): African, Asian, and Madagascar faunas have high skewness, while the South American fauna has a strikingly low-skewness topology (Vos 2006) very different from that for the rest of the primates. (Our failure

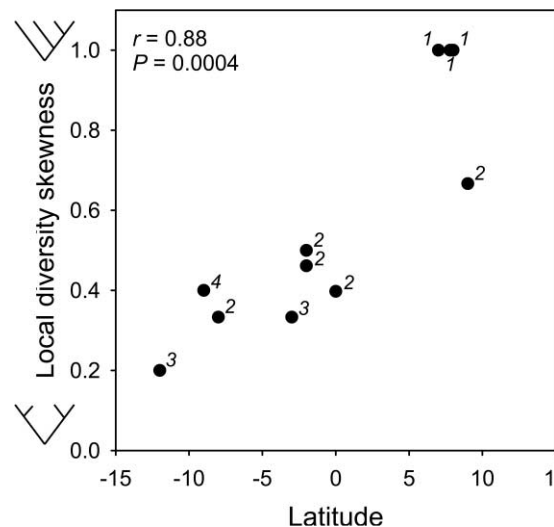


Figure 3: Latitudinal gradient in local diversity skewness for African parks (11 parks with five or more primate species). Beside each data point is the number of lorises (galagos, pottos, etc.) included in that local assemblage.

to detect departure from the phylogenetic null for either the Madagascar or the South American faunas most likely reflects the low power of tests against that null.) The contrast among continents indicates very different histories of diversification and/or faunal assembly on the two sides of the Atlantic and suggests that something in the geographic context of New World and Old World primate radiations may have shaped their diversification differently. We do not know what the critical differences might have been, and attempts to speculate involve post hoc reasoning. For example, repeated glacial fragmentation of South American habitats (Haffer 1997) might have driven the kind of simultaneous across-lineage vicariance that can produce low diversity skewness, but we do not know whether similar fragmentation occurred elsewhere, and the role of climate cycling and glacial refugia in controlling diversification remains controversial (Cheviron et al. 2005; Rull 2005; Weir 2006). On the extinction side, the Madagascar fauna has experienced severe and size-selective Holocene extinction (Godfrey et al. 1997; Burney 1999), and nonrandom extinction can increase diversity skewness (Heard and Mooers 2002). We do not know, however, whether African and Asian faunas (also high skewness) experienced similar nonrandom extinctions, and we cannot reject the occurrence of such extinctions in South America (low skewness). Many other hypotheses are conceivable, and hypothesis tests will be difficult, but the first step is obvious: we need to know whether the New World/Old World contrast in diversity skewness is unique to the primates or is a shared feature of other South American radiations. This will require analyses like ours for a number of other clades.

Local Skewness. Local drivers of diversity skewness would be distinguished from regional ones by their dependence on interactions between co-occurring species. Notably, if niche breadth is phylogenetically conservative, then in competitively structured communities, species packing during community assembly could favor high skewness. In this case, lineages with narrow niches could be tightly packed and heavily represented in local communities, while lineages with broad niches would contribute fewer representatives. Phylogenetic conservatism of niche breadth has been demonstrated for at least some clades (e.g., passerines; Brandle et al. 2002), but other studies show high enough rates of niche breadth evolution to suggest that phylogenetic signal will be far from absolute (e.g., Morse and Farrell 2005; Nosil and Mooers 2005; Sipes and Tepedino 2005). Therefore, we cannot yet assess the likelihood that niche breadth effects lead to local diversity skewness in real clades.

Because no primate assemblage showed significant local diversity skewness (vs. continental source pools), we have

no evidence for any such community-assembly effects for primates. However, our use of only large parks may have limited our ability to detect truly local effects, and of course such effects may exist for other clades even if they do not for primates.

The African Latitudinal Gradient in Diversity Skewness: The “Possum Effect”

The lack of local signal in primate diversity skewness does not preclude the existence of among-assemblage patterns, and indeed for Africa we found a strong regional pattern: local diversity skewness shows a strong gradient from low skewness in the north to high skewness in the south. Unlike more familiar latitudinal gradients in species richness (Rosenzweig 1995), the skewness gradient does not peak at the equator. While other latitudinal gradients may arise because latitude is a proxy for local environmental variables such as productivity or harshness (Rosenzweig 1995), we do not believe such explanations are relevant here (and skewness had no correlation with NPP).

Instead, the diversity-skewness gradient appears to reflect biogeographic pattern in African primate radiations. In particular, a strong role is played by the distribution of the family Lorisidae, which has a number of African representatives (pottos, galagos, bushbabies). The lorisesids have a complex biogeographic distribution (Instituto Ecología Aplicada 1998), but with respect to our result, the most important feature is that lorisesids tend to be more species rich in southern parks (fig. 3). Because the lorisesids are sister to all other African primates, their local species richness strongly influences diversity skewness (nodes near the phylogeny's root are weighted heavily by I_c because the largest diversity contrasts are possible there). The inclusion of just one loriseid in a local assemblage makes that assemblage highly diversity skewed, while the inclusion of just a few more greatly reduces diversity skewness. Clades like the lorisesids have similarly large effects on calculations of “evolutionary heritage” and distinctiveness-weighted conservation value (e.g., Mooers et al. 2004; Soutullo et al. 2005).

A familiar version of this effect occurs in North American mammal assemblages, where the Virginia opossum (*Didelphis virginiana*) ranges far north of any other member of its basal (marsupial) clade. Therefore, along a north-to-south transect, North American mammal assemblages should have low diversity skewness north of the range of *D. virginiana*, dramatically higher skewness across the zone where *D. virginiana* is the only marsupial, and then lower skewness in Central America, where *D. virginiana* is joined by several other opossums. This “possum effect” raises local diversity skewness in temperate North America be-

cause *D. virginiana* represents the outer envelope of a geographically limited radiation.

Inspection of range maps in almost any field guide makes it clear that the kind of staggered boundaries that drive the possum effect are very common. Possum effects should, therefore, be quite common in biogeographic studies of diversity skewness. Our lorised result highlights the importance of understanding the ecology and evolution of range boundaries (e.g., Keitt et al. 2001; Gaston 2003; Swenson and Howard 2005) and thus the biogeographic structure of adaptive radiations (Schluter 2000; Böhm and Mayhew 2005).

Data Requirements and Prospects for Further Analyses

Our primate analysis is only the first step in the study of geographic pattern in diversity skewness. As for any other macroecological pattern, we anticipate learning much more from parallels and differences among clades in the structure of diversity skewness than we can learn from analysis of any single group. However, this raises the issue of whether analyses such as ours are currently feasible for other taxa. The analysis of local diversity skewness requires a phylogeny and a set of local species lists for the clade of interest. There are fairly stringent requirements for each, and so we suspect there are as yet relatively few clade/assemblage combinations for which analysis is possible.

With respect to phylogeny, our analysis requires a phylogenetic estimate that is complete at the species level and that is reasonably well resolved. Lack of phylogenetic resolution complicates significance testing, with inconclusive tests most likely when polytomies are numerous, large, or deep in the tree. Branch length data are not necessary.

With respect to species lists, our analysis requires complete lists for a set of local assemblages (and for analysis including tests against regional source pools, it requires regional species lists as well). As for any biogeographic analysis, local assemblage data must be comparable across sites, which can be surprisingly difficult to guarantee over large geographic scales. More restrictive is a requirement for fairly high species richness of the chosen clade in each local assemblage, arising because analysis of diversity skewness for a single assemblage is likely to be inconclusive unless that assemblage includes at least about eight species. Smaller phylogenies have few possible topologies, and this leads to low statistical power to distinguish any topology from a random expectation (although smaller assemblages can still contribute to broader-scale patterns such as the latitudinal pattern in fig. 3). To meet the requirement for a diverse local assemblage, one can choose locations with high α diversity or one can define the clade of interest broadly (e.g., passerines rather than warblers). Each of these has its disadvantages: restricting analysis to locations

with high α diversity might exclude much of the globe, but clades defined very broadly are likely to lack complete, well-resolved phylogenies and may include members unlikely to interact in local communities.

While data sets meeting the stringent requirements of our analysis are currently uncommon, we do not wish to seem unduly pessimistic. The availability of phylogenetic estimates and compilations of range data is improving constantly, and so data sets amenable to analysis should become rapidly more common. Our primate data, of course, give us only the first glimpse of the biogeographic structure of diversity skewness. It will require similar analyses for many clades in many sets of places before we can know whether our results (significant continental and regional patterns but no evidence for a local contribution to diversity skewness) are typical or unusual. Diversity skewness is a fundamental property of life on Earth (Mooers and Heard 1997), but surprisingly, its study is still in its infancy.

Acknowledgments

We thank A. Mooers and R. Vos for early access to their primate supertree and J. Terborgh for discussion of Manu primates. O. Bininda-Emonds, F. Bokma, J. Davies, R. Grenier, S. Hendrix, A. Mooers, S. Renner, and the Heard and Quiring labs provided valuable comments on our ideas. We also thank the researchers who compiled the primate range maps. For South America, maps were compiled by NatureServe with G. Ceballos, B. Patterson, W. Sechrest, M. Tognelli, the Nature Conservancy Migratory Bird Program, Conservation International, World Wildlife Fund (United States), and Environment Canada WILDSPACE. For Africa, maps were from the African Mammals Databank, a project of the Instituto Ecologia Aplicada (led by L. Boitani). Our own work was funded by the National Science Foundation and the Canadian Natural Sciences and Engineering Research Council.

APPENDIX

SkewMatic 2.01 Installation Instructions

SkewMatic 2.01 is released for the Windows XP operating system (although it should run on Windows versions as old as Windows 98). However, it requires a Windows component called the ".NET Framework". If you are not sure whether or not you have the .NET Framework, proceed with SkewMatic installation; if the .NET Framework is not available on your computer, the installer will advise you of this. If this happens, visit the Microsoft Download Center (<http://www.microsoft.com/downloads/details.aspx?FamilyID=262d25e3-f589-4842-8157-034d1e7cf3a3&>

displaylang=en). Run the .NET Framework installer provided there, and then retry SkewMatic installation.

Download and unzip the SkewMatic installer to any location on your machine. Five files will appear; locate and click on "Setup.exe". This will launch an installation wizard that will guide you through the remainder of the installation. Instructions for using SkewMatic ("Instructions.pdf") are located in the "SkewMatic\SkewMatic 201\Samples" folder.

Literature Cited

- Agapow, P. M., and A. Purvis. 2002. Power of eight tree-shape statistics to detect nonrandom diversification: a comparison by simulation of two models of cladogenesis. *Systematic Biology* 51:866–872.
- Anderson, T. M., M. A. Lachance, and W. T. Starmer. 2004. The relationship of phylogeny to community structure: the cactus yeast community. *American Naturalist* 164:709–721.
- Blum, M. G. B., and O. Francois. 2005. On statistical tests of phylogenetic tree imbalance: the Sackin and other indices. *Mathematical Biosciences* 195:141–153.
- Blum, M. G. B., O. Francois, and S. Janson. 2006. The mean, variance and joint distribution of two statistics sensitive to phylogenetic tree balance. *Annals of Applied Probability* 16:2195–2214.
- Böhm, M., and P. J. Mayhew. 2005. Historical biogeography and the evolution of the latitudinal gradient of species richness in the Papionini (Primate: Cercopithecidae). *Biological Journal of the Linnean Society* 85:235–246.
- Boitani, L., F. Corsi, G. Reggiani, J. Sinibaldi, and P. Trapanese. 1999. A databank for the conservation and management of African mammals. Instituto di Ecologia Applicata, Rome.
- Brandle, M., A. Prinzing, R. Pfeifer, and R. Brandl. 2002. Dietary niche breadth for central European birds: correlations with species-specific traits. *Evolutionary Ecology Research* 4:643–657.
- Burney, D. A. 1999. Rates, patterns, and the processes of landscape transformation and extinction in Madagascar. Pages 145–164 in R. D. E. MacPhee, ed. *Extinctions in near time: causes, contexts, and consequences*. Plenum, New York.
- Cavender-Bares, J., D. D. Ackerley, D. A. Baum, and F. A. Bazzaz. 2004. Phylogenetic overdispersion in Floridian oak communities. *American Naturalist* 163:823–843.
- Cheviron, Z. A., S. J. Hackett, and A. P. Capparella. 2005. Complex evolutionary history of a Neotropical lowland forest bird (*Lepidothrix coronata*) and its implications for historical hypotheses of the origin of Neotropical avian diversity. *Molecular Phylogenetics and Evolution* 36:338–357.
- Darwin, C. (1859) 1968. *The origin of species*. Penguin, London.
- Davies, T. J., V. Savolainen, M. W. Chase, P. Goldblatt, and T. G. Barraclough. 2005. Environment, area, and diversification in the species-rich flowering plant family Iridaceae. *American Naturalist* 166:418–425.
- Field, R., E. M. O'Brien, and R. J. Whittaker. 2005. Global models for predicting woody plant richness from climate: development and evaluation. *Ecology* 86:2263–2277.
- Gaston, K. J. 1991. How large is a species' geographic range? *Oikos* 61:434–438.
- . 2003. *The structure and dynamics of geographic ranges*. Oxford University Press, Oxford.
- Godfrey, L. R., W. L. Jungers, K. E. Reed, E. L. Simons, and P. S. Chatrath. 1997. Subfossil lemurs: inferences about past and present primate communities in Madagascar. Pages 218–256 in S. M. Goodman and B. D. Patterson, eds. *Natural change and human impact in Madagascar*. Smithsonian Institution, Washington, DC.
- Haffer, J. 1997. Alternative models of vertebrate speciation in Amazonia: an overview. *Biodiversity and Conservation* 6:451–476.
- Heard, S. B. 1992. Patterns in tree balance among cladistic, phenetic, and randomly generated phylogenetic trees. *Evolution* 46:1818–1826.
- Heard, S. B., and A. O. Mooers. 2002. Signatures of random and selective mass extinctions in phylogenetic tree balance. *Systematic Biology* 51:889–897.
- Housworth, E. A., and E. P. Martins. 2001. Random sampling of constrained phylogenies: conducting phylogenetic analyses when the phylogeny is partially known. *Systematic Biology* 50:628–639.
- Instituto di Ecologia Applicata. 1998. African mammals databank: a databank for the conservation and management of the African mammals. Report to the Director-General for Development (DGVIII/A/1) of the European Commission, project no. B7-6200/94-15/VIII/ENV. European Commission, Brussels.
- IUCN. 1994. Guidelines for protected areas management categories. International Union for Conservation of Nature, Cambridge.
- Keitt, T. H., M. A. Lewis, and R. D. Holt. 2001. Allee effects, invasion pinning, and species' borders. *American Naturalist* 157:203–216.
- Kellogg, E. A. 1999. Phylogenetic aspects of the evolution of C₄ photosynthesis. Pages 411–444 in R. F. Sage and R. K. Monson, eds. *C₄ plant biology*. Academic Press, San Diego, CA.
- Manly, B. F. J. 1997. *Randomization, bootstrap, and Monte Carlo methods in biology*. 2nd ed. Chapman & Hall, London.
- Mendelson, T. C., and K. L. Shaw. 2005. Sexual behaviour: rapid speciation in an arthropod. *Nature* 433:375–376.
- Mooers, A. O. 1995. Tree balance and tree completeness. *Evolution* 49:379–384.
- Mooers, A. O., and S. B. Heard. 1997. Inferring evolutionary process from phylogenetic tree shape. *Quarterly Review of Biology* 72:31–54.
- Mooers, A. O., S. B. Heard, and E. Chrostowski. 2004. Evolutionary heritage as a metric for conservation. Pages 120–138 in A. Purvis, J. L. Gittleman, and T. M. Brooks, eds. *Phylogeny and conservation*. Cambridge University Press, Cambridge.
- Mooers, A. O., L. J. Harmon, D. H. J. Wong, and S. B. Heard. 2007. Some models of phylogenetic tree shape. Pages 147–168 in O. Gascuel and M. Steel, eds. *New mathematical models for evolution*. Oxford University Press, Oxford.
- Morse, G. E., and B. D. Farrell. 2005. Ecological and evolutionary diversification of the seed beetle genus *Stator* (Coleoptera: Chrysomelidae: Bruchinae). *Evolution* 59:1315–1333.
- Nosil, P., and A. O. Mooers. 2005. Testing hypotheses about ecological specialization using phylogenetic trees. *Evolution* 59:2256–2263.
- Patterson, B. D., G. Ceballos, W. Sechrest, M. F. Tognelli, T. Brooks, L. Luna, P. Ortega, I. Salazar, and B. E. Young. 2005. Digital distribution maps of the mammals of the Western Hemisphere. Version 2.0. NatureServe, Arlington, VA.
- Purvis, A., and P. M. Agapow. 2002. Phylogeny imbalance: taxonomic level matters. *Systematic Biology* 51:844–854.
- Rogers, J. S. 1994. Central moments and probability distribution of Colless's coefficient of tree imbalance. *Evolution* 48:2026–2036.
- Rosenzweig, M. L. 1995. *Species diversity in space and time*. Cambridge University Press, Cambridge.

- Rull, V. 2005. Biotic diversification in the Guayana Highlands: a proposal. *Journal of Biogeography* 32:921–927.
- Sage, R. F., D. A. Wedin, and M. Li. 1999. The biogeography of C₄ photosynthesis: patterns and controlling factors. Pages 313–376 in R. F. Sage and R. K. Monson, eds. *C₄ plant biology*. Academic Press, San Diego, CA.
- Scheiner, S. M., and M. R. Willig. 2005. Developing unified theories in ecology as exemplified with diversity gradients. *American Naturalist* 166:458–469.
- Schluter, D. 2000. *The ecology of adaptive radiation*. Oxford University Press, Oxford.
- Silvertown, J., K. McConway, D. Gowing, M. Dodd, M. F. Fay, J. A. Joseph, and K. Dolphin. 2006. Absence of phylogenetic signal in the niche structure of meadow plant communities. *Proceedings of the Royal Society B: Biological Sciences* 273:39–44.
- Sipes, S. D., and V. J. Tepedino. 2005. Pollen-host specificity and evolutionary patterns of host switching in a clade of specialist bees (Apoidea: *Diadasia*). *Biological Journal of the Linnean Society* 86: 487–505.
- Sokal, R. R., and F. J. Rohlf. 1995. *Biometry: the principles and practice of statistics in biological research*. W. H. Freeman, New York.
- Soutullo, A., S. Dodsworth, S. B. Heard, and A. O. Mooers. 2005. Distribution and correlates of carnivore phylogenetic diversity across the Americas. *Animal Conservation* 8:249–258.
- Stephens, P. R., and J. J. Wiens. 2004. Convergence, divergence, and homogenization in the ecological structure of emydid turtle communities: the effects of phylogeny and dispersal. *American Naturalist* 164:244–254.
- Swenson, N. G., and D. J. Howard. 2005. Clustering of contact zones, hybrid zones, and phylogeographic breaks in North America. *American Naturalist* 165:581–591.
- Vos, R. A. 2006. *Inferring large phylogenies: the big tree problem*. PhD thesis. Simon Fraser University, Burnaby.
- Webb, C. O., D. D. Ackerley, M. A. McPeck, and M. J. Donoghue. 2002. Phylogenies and community ecology. *Annual Review of Ecology and Systematics* 33:475–505.
- Weir, J. T. 2006. Divergent timing and patterns of species accumulation in lowland and highland Neotropical birds. *Evolution* 60: 842–855.
- Wiens, J. J., M. C. Brandley, and T. W. Reeder. 2006. Why does a trait evolve multiple times within a clade? repeated evolution of snakelike body form in squamate reptiles. *Evolution* 60:123–141.
- Wilkinson, M., J. A. Cotton, C. Creevey, O. Eulenstein, S. R. Harris, F. J. Lapointe, C. Levasseur, J. O. McInerney, D. Pisani, and J. L. Thorley. 2005. The shape of supertrees to come: tree shape related properties of fourteen supertree methods. *Systematic Biology* 54: 419–431.
- Williams, C. B. 1964. *Patterns in the balance of nature*. Academic Press, London.
- Wilson, D. E., and D. M. Reeder, eds. 1993. *Mammal species of the world*. Smithsonian Institution, Washington, DC.

Associate Editor: Susanne S. Renner

Editor: Michael C. Whitlock