

# Diversity, Endemism, and Age Distributions in Macroevolutionary Sources and Sinks

Emma E. Goldberg,<sup>1,\*</sup> Kaustuv Roy,<sup>1,†</sup> Russell Lande,<sup>1,‡</sup> and David Jablonski<sup>2,§</sup>

1. Division of Biological Sciences, University of California, San Diego, La Jolla, California 92093

2. Department of Geophysical Sciences, University of Chicago, Chicago, Illinois 60637

Submitted September 13, 2004; Accepted February 8, 2005;  
Electronically published April 12, 2005

Online enhancement: appendix.

---

**ABSTRACT:** Quantitative tests of historical hypotheses are necessary to advance our understanding of biogeographic patterns of species distributions, but direct tests are often hampered by incomplete fossil or historical records. Here we present an alternative approach in which we develop a dynamic model that allows us to test hypotheses about regional rates of taxon origination, extinction, and dispersal using information on ages and current distributions of taxa. With this model, we test two assumptions traditionally made in the context of identifying regions as “centers of origin”—that regions with high origination rates will have high diversity and high endemism. We find that these assumptions are not necessarily valid. We also develop expressions for the regional age distributions of extant taxa and show that these may yield better insight into regional evolutionary rates. We then apply our model to data on the biogeography and ages of extant genera of marine bivalves and conclude that diversity in polar regions predominantly reflects dispersal of taxa that evolved elsewhere rather than in situ origination-extinction dynamics.

**Keywords:** biogeography, paleontology, macroevolution, biodiversity, endemism, marine bivalve.

---

The processes that produce large-scale spatial patterns of taxonomic diversity remain poorly understood despite the existence of many competing hypotheses. In particular, the role of historical processes in shaping present-day biogeographic

patterns has been a subject of considerable debate (Francis and Currie 1998, 2003; Currie and Francis 2004; Qian and Ricklefs 2004; Ricklefs 2004). Historical hypotheses are considered to be essentially untestable by some (Francis and Currie 1998), while others maintain that past episodes of speciation, extinction, and range expansion or dispersal are major determinants of present-day biogeographic patterns (Qian and Ricklefs 2004; Ricklefs 2004). The fundamental problem here is that direct tests of historical hypotheses require a fossil or historical record with excellent temporal and spatial resolution, information that is not available for most taxa. On the other hand, information about the current spatial distributions of species and higher taxa is readily available for many groups and potentially obtainable for all living taxa. The challenge therefore is to develop a theoretical framework that allows us to use present-day biogeographic data to test hypotheses about historical processes underlying global biodiversity patterns.

The problems inherent in inferring past processes from present observations about the geographic distributions of taxa are perhaps best illustrated in attempts to identify regions as “centers of origin” or “cradles of diversity,” or as their counterparts, “centers of accumulation” or “museums of diversity.” The identification of such regions may aid explanation of spatial diversity patterns and could also guide conservation priorities. The terms “center of origin” and “cradle of diversity” designate regions with a high rate of taxon origination but do not specify relative rates of local extinction, immigration, or emigration (Chown and Gaston 2000; Mora et al. 2003; Briggs 2004). The term “center of origin” has also been used to indicate the region in which a particular taxon first appeared (Darwin [1859] 1975; Ricklefs and Schluter 1993). This may or may not be the same as the region in which it underwent greatest diversification or as a region in which many other taxa originated, as designated by the first meaning of the phrase. A “center of accumulation” is a region that obtains taxa through immigration (Ladd 1960; Palumbi 1996; Mora et al. 2003; Briggs 2004), and a “museum of diversity” is a region with a low rate of local extinction (Stebbins 1974;

\* E-mail: goldberg@biomail.ucsd.edu.

† E-mail: kroy@biomail.ucsd.edu.

‡ E-mail: rlande@biomail.ucsd.edu.

§ E-mail: d-jablonski@uchicago.edu.

Chown and Gaston 2000). Fundamentally, then, these terms are defined by the rates of origination, extinction, and dispersal, but these labels are often assigned to regions on the basis of current taxon richness, the amount of endemism, or other information about extant taxa (reviewed by Ricklefs and Schluter [1993]; Brown and Lomolino [1998]; Briggs [2004]). Many studies have used the qualitative expectations that a region with a higher origination rate should have higher levels of diversity and greater endemism (Willis 1922; Rosenzweig and Sandlin 1997; Mora et al. 2003; reviews by McCoy and Heck [1976] and Ricklefs and Schluter [1993]), but few have explicitly tested this hypothesis (but see Pandolfi 1992). Others have used ages of taxa to distinguish differences in regional rates (e.g., Stehli et al. 1969; Stehli and Wells 1971; Gaston and Blackburn 1996; Briggs 1999). The use of age distributions of taxa living in an area today to infer past diversification rates in that region is also not without problems (Ricklefs and Schluter 1993; Gaston and Blackburn 1996; Chown and Gaston 2000).

The work we present was motivated by a lack of theory on how information about extant taxa can be used to distinguish a “center of origin” from a “center of accumulation.” To draw a sharper distinction between these two types of regions, we use the term “macroevolutionary source” to refer to a region that is a “center of origin” by the first meaning (high rate of origination relative to other regions) and also that does not receive taxa through dispersal (zero immigration). We use the complementary term “macroevolutionary sink” to refer to a region that obtains taxa through immigration, as a “center of accumulation” does, but has no local origination. Under these definitions, the differences between source and sink regions will be maximized, and so the different effects of local origination and dispersal will be highlighted. We address the issue of how a source region can be distinguished from a sink region by considering the more general issue of how distributions of extant taxa are determined by regional rates of origination, extinction, and dispersal. We examine the problem by first developing a theoretical model and then applying the model to empirical data on marine bivalves. We find that high regional diversity and endemism cannot be used alone to infer a high regional rate of origination. We show, however, that age distributions of extant taxa can be used to estimate rates of origination, extinction, and dispersal. In addition, we use this model to show that the origination rate of marine bivalve genera is significantly lower in the polar regions than at lower latitudes and that the rate of movement of genera is greater into the polar regions than out of them, indicating that the poles represent a macroevolutionary sink.

## The Model

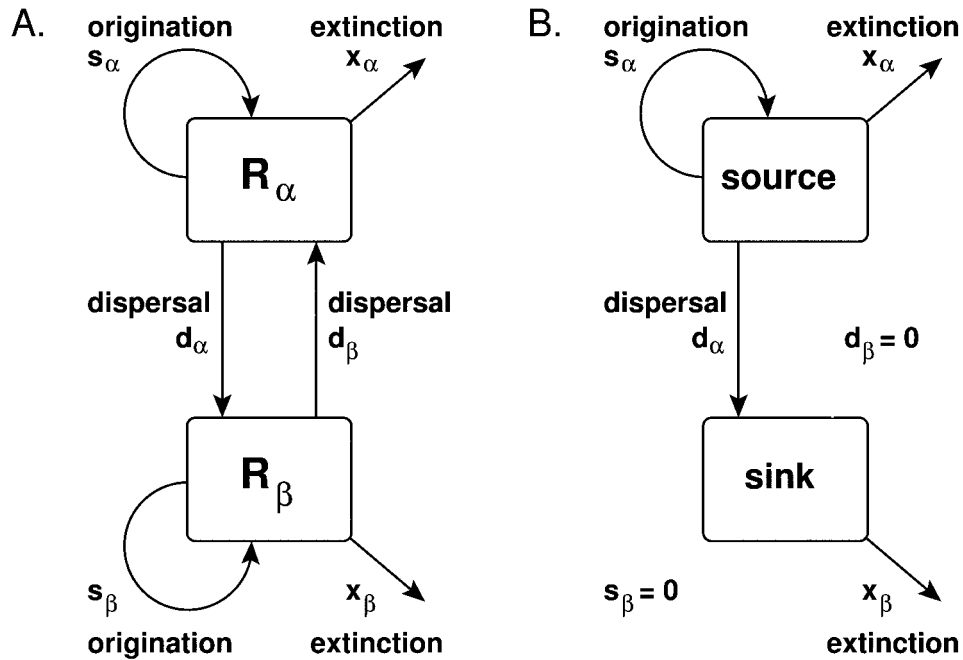
We develop a dynamic model that describes distributions of diversity (taxon richness), endemism, and age in a system consisting of two regions. These regions may differ in their local rates of taxon origination and extinction, and the taxa of the system are able to disperse from each region to the other (i.e., expand their ranges). The two regions are denoted  $R_\alpha$  and  $R_\beta$ . Origination occurs through a branching process at a rate  $s_\alpha$  per taxon in region  $R_\alpha$  and  $s_\beta$  in region  $R_\beta$ . Extinction occurs at per-taxon rates  $x_\alpha$  and  $x_\beta$ . Dispersal occurs at per-taxon rates  $d_\alpha$  from  $R_\alpha$  to  $R_\beta$  and  $d_\beta$  from  $R_\beta$  to  $R_\alpha$ . All rates are nonnegative. We emphasize that extinction of a taxon is independent in the two regions and that some time may pass before a taxon present in one region appears in the other; the expected value of this lag time is the reciprocal of the dispersal rate. A schematic diagram of this system is shown in figure 1A, and a mathematical description is given below.

This general system can be specialized to represent a system consisting of a source region and a sink region by setting  $s_\beta = 0$  and  $d_\beta = 0$  as illustrated in figure 1B. We should clarify that we do not use these terms in the demographic sense; that is, we do not assume that a taxon can persist in a sink region only if it is being continually supplied from the source region. We mean only that all taxa in the system originated in the source region and that taxa do not move from the sink to the source.

We assume that all processes affecting taxa occur independently and that all taxa in a given region are subject to the same rates of origination, extinction, and dispersal, which are constant in time. These are significant assumptions and are addressed in “Discussion.” The model is formulated deterministically in continuous time. Similar results were obtained with discrete time models (not presented here) and stochastic simulations (discussed in the appendix in the online edition of the *American Naturalist*).

### Total Diversity and Endemism

We employ a matrix formulation to express the expected dynamics of the system. Define the column vector  $\mathbf{n}(t) = (n_1, n_2, n_3)^T$  (the superscript T indicates matrix transposition), where  $n_1(t)$  is the expected number of taxa present in both  $R_\alpha$  and  $R_\beta$  at time  $t$ ,  $n_2(t)$  is the expected number of taxa present only in  $R_\alpha$  ( $R_\alpha$  endemics), and  $n_3(t)$  is the expected number of taxa present only in  $R_\beta$  ( $R_\beta$  endemics). The total number of taxa expected in  $R_\alpha$  is thus  $n_1(t) + n_2(t)$ , and the total number of taxa expected in  $R_\beta$  is  $n_1(t) + n_3(t)$ . With this notation, transition rates between these three states by origination, dispersal, and extinction can be written, respectively, as



**Figure 1:** Schematic representation of the model. *A*, In the general case of the model, each region ( $R_\alpha$  and  $R_\beta$ ) has its own rate of taxon origination ( $s_\alpha$  and  $s_\beta$ ) and extinction ( $x_\alpha$  and  $x_\beta$ ), and dispersal (at rates  $d_\alpha$  and  $d_\beta$ ) occurs between the regions. All rates are per taxon and are constant in time and across taxa. *B*, In a special case of the model, one region is a macroevolutionary source and obtains taxa only through local origination, not immigration. The other region is a macroevolutionary sink and obtains taxa only through immigration, not local origination. Our results are for the general case of the model, *A*, except where explicitly stated otherwise. We emphasize that the special case, *B*, is used in the discussion of relative endemism (including eq. [3]; fig. 2) and that the general case, *A*, is used when fitting data.

$$\mathbf{S} = \begin{pmatrix} 0 & 0 & 0 \\ s_\alpha & s_\alpha & 0 \\ s_\beta & 0 & s_\beta \end{pmatrix},$$

$$\mathbf{D} = \begin{pmatrix} 0 & d_\alpha & d_\beta \\ 0 & -d_\alpha & 0 \\ 0 & 0 & -d_\beta \end{pmatrix},$$

$$\mathbf{E} = \begin{pmatrix} -x_\beta & -x_\alpha & 0 & 0 \\ x_\beta & -x_\alpha & 0 & 0 \\ x_\alpha & 0 & -x_\beta & 0 \end{pmatrix},$$

We can now assess the general validity of the claim that a source region will have higher diversity or endemism than a sink region. The equilibrium behavior of equation (2) can be obtained by considering the eigenvector corresponding to the dominant eigenvalue of  $\mathbf{S} + \mathbf{D} + \mathbf{E}$ ; we call this dominant eigenvector  $\mathbf{u} = (u_1, u_2, u_3)^T$ . Since we are interested in the case of a source and a sink, we set  $s_\beta = d_\beta = 0$  to make  $R_\alpha$  a pure source and  $R_\beta$  a pure sink. The proportions of taxa in each region that are endemic there are then

such that the system changes with time according to

$$\frac{d\mathbf{n}(t)}{dt} = (\mathbf{S} + \mathbf{D} + \mathbf{E})\mathbf{n}(t). \quad (1)$$

The general solution to equation (1) is

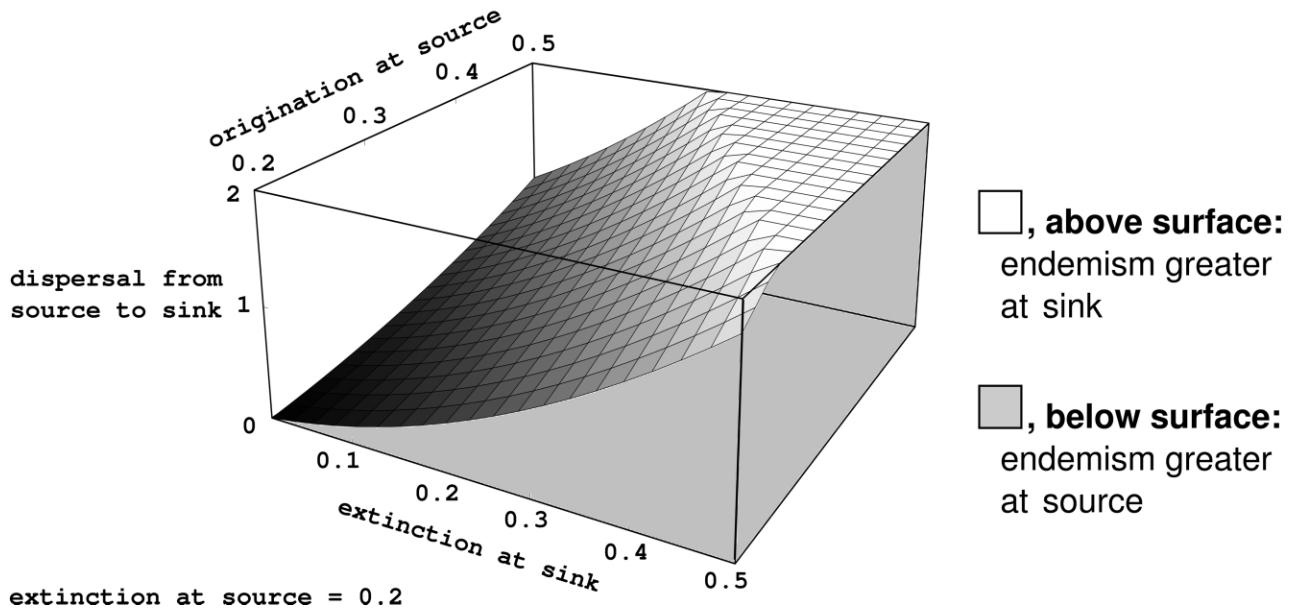
$$\mathbf{n}(t) = e^{(\mathbf{S} + \mathbf{D} + \mathbf{E})(t - t_0)} \mathbf{n}(t_0), \quad (2)$$

where  $t_0$  is a time at which the state of the system is known and the matrix exponential is defined by its Taylor expansion (Apostol 1969).

$$\frac{u_2}{u_1 + u_2} = \frac{s_\alpha + x_\beta}{s_\alpha + d_\alpha + x_\beta} \quad \text{at } R_\alpha, \quad (3a)$$

$$\frac{u_3}{u_1 + u_3} = \frac{x_\alpha}{s_\alpha + x_\beta} \quad \text{at } R_\beta. \quad (3b)$$

Some numerical experimentation reveals that proportional endemism can be greater either at a source or at a sink, depending on the relative values of  $s_\alpha$ ,  $x_\alpha$ ,  $x_\beta$ , and  $d_\alpha$ . An example is shown graphically in figure 2. The number of endemic taxa ( $u_2$  and  $u_3$ ) and the total number of taxa ( $u_1 + u_2$  and  $u_1 + u_3$ ) can also be greater in either of the



**Figure 2:** Relative proportions of endemism. The surface is that of equal proportional endemism at the source and the sink, defined as  $u_2/(u_1 + u_2) = u_3/(u_1 + u_3)$  (see eq. [3]). This illustrates that there is a substantial portion of parameter space in which endemism is greater at the sink than at the source, demonstrating that high levels of endemism in a region do not necessarily indicate high levels of taxon origination there. (The back right corner is not really flat; it has been truncated for the plot.)

two regions. Total diversity and endemism at a single time therefore cannot be used alone to distinguish a source from a sink.

*Age Distributions*

In addition to the expression for the number of taxa in each location as a function of time (eq. [2]), we are interested in the age distribution of taxa in each region at a particular time. Define the vector  $\mathbf{n}'(t, t')$  to describe the number of taxa alive at time  $t$  that survive until a later time  $t'$ . The survival of taxa is described by

$$\frac{d\mathbf{n}'(t, t')}{dt'} = (\mathbf{D} + \mathbf{E})\mathbf{n}'(t, t'), \tag{4}$$

which has the solution

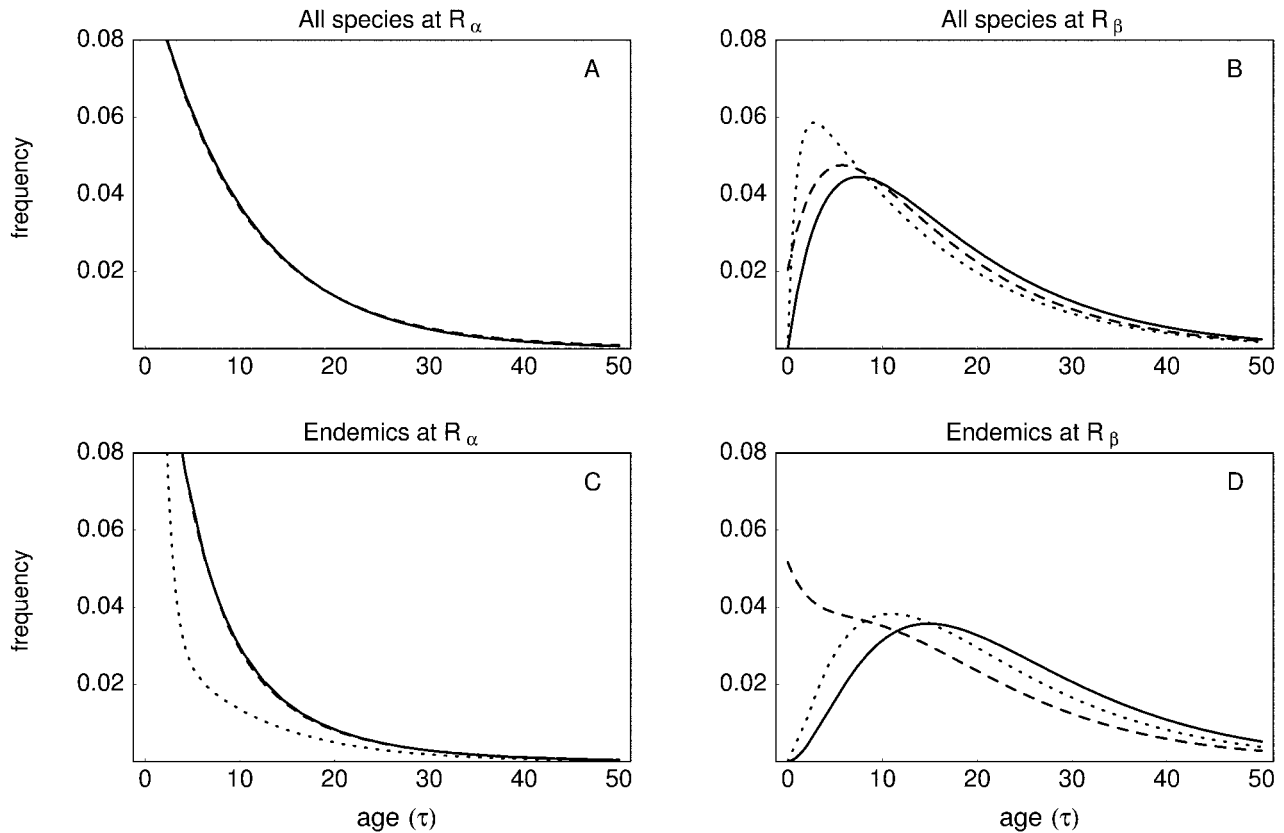
$$\mathbf{n}'(t, t') = e^{(\mathbf{D}+\mathbf{E})(t'-t)}\mathbf{n}'(t, t). \tag{5}$$

To form the age distribution of taxa, first define  $\tau = t' - t$  to be the age of taxa created at time  $t$ , as observed at time  $t'$ . The rate at which new taxa are created at time  $t$  is  $\mathbf{S}\mathbf{n}(t)$ ; equation (2) gives an expression for  $\mathbf{n}(t)$ ; equation (5) describes the survival of these new taxa. The age distribution of extant taxa,  $\mathbf{f}(\tau)$ , can thus be written as

$$\mathbf{f}(\tau) = e^{(\mathbf{D}+\mathbf{E})\tau}\mathbf{S}e^{(\mathbf{S}+\mathbf{D}+\mathbf{E})(t'-\tau-t_0)}\mathbf{n}(t_0). \tag{6}$$

Elements of the vector  $\mathbf{f} = (f_1, f_2, f_3)^T$  are such that  $\int_0^\infty f_i(\tau)d\tau$  equals the total number of taxa in state  $i$  alive at time  $t'$ . The function  $f_i(\tau)$  can therefore be normalized to become a probability density function by dividing by this integral.

Figure 3 shows the behavior of the normalized age distribution function for all taxa and for endemics at  $R_\alpha$  and at  $R_\beta$  for three hypothetical sets of parameter values. It illustrates that the distributions of ages of extant taxa reflect differences in macroevolutionary rates between regions. In particular, there is a clear qualitative difference between the age distributions for source and sink regions (*solid lines*, fig. 3): most taxa in the source region are young, while most taxa in the sink region are of intermediate age. In this case, all taxa originate at the source, and none can disperse from the sink to the source. Young taxa are unlikely to have become extinct at the source and unlikely to have dispersed to the sink. Older taxa are more likely to have become extinct at the source than at the sink because they have been introduced to the source only once (when they were created) but have had many opportunities to disperse to the sink (until they become extinct at the source). Considering only endemic taxa makes



**Figure 3:** Normalized age distribution functions, equation (6). *A*, All taxa in  $R_\alpha$ ,  $f_1(\tau) + f_2(\tau)$ . *B*, All taxa in  $R_\beta$ ,  $f_1(\tau) + f_2(\tau)$ . *C*,  $R_\alpha$  endemics,  $f_2(\tau)$ . *D*,  $R_\beta$  endemics,  $f_3(\tau)$ . Solid lines are for parameter values  $s_\alpha = d_\alpha = 0.1$ ,  $s_\beta = d_\beta = 0$ ,  $x_\alpha = x_\beta = 0.05$ , illustrating the case when  $R_\alpha$  is a pure source and  $R_\beta$  is a pure sink and demonstrating clear differences in the age distributions of the two regions. Dotted lines are for parameter values  $s_\alpha = 0.1$ ,  $d_\alpha = 1$ ,  $s_\beta = d_\beta = 0$ ,  $x_\alpha = x_\beta = 0.05$ , illustrating the case of very high dispersal from the source,  $R_\alpha$ , to the sink,  $R_\beta$ . Dashed lines are for parameter values  $s_\alpha = d_\alpha = 0.1$ ,  $s_\beta = d_\beta = 0.02$ ,  $x_\alpha = x_\beta = 0.05$ , illustrating a situation where the source-sink relationship is relaxed. In all cases the initial condition used is  $\mathbf{n}(-50) = (1, 0, 0)^T$ . With time units of millions of years, these origination and extinction rates are biologically reasonable (Van Valen 1973; Stanley 1985; Sepkoski 1998).

the differences between the source and sink regions more marked.

When there is no dispersal into a region, the shape of the age distribution is exponential with a rate constant equal to the local origination rate, independent of the extinction rate. This can be shown by writing the age distribution for an isolated region,  $f(\tau)$ , following the same reasoning as for the derivation of equation (6). With  $s$  as the origination rate and  $x$  as the extinction rate,  $f(\tau) = e^{-x\tau} s e^{(s-x)(t'-\tau-t_0)} n(t_0) = s e^{(s-x)(t'-t_0)} n(t_0) e^{-s\tau}$ . The only age dependence is in the last factor, so the age distribution normalized as after equation (6) is simply  $se^{-s\tau}$  (see also Pease 1988; Foote 2001).

We also illustrate in figure 3 two situations in which the differences in age distributions between the regions are reduced. First, when dispersal from the source to the sink is very high, the sink region will better mirror the

contents of the source. The peak in the sink age distribution therefore shifts toward the left (*dotted lines*, fig. 3), and with an extremely high dispersal rate, the youngest taxa will dominate the sink age distribution as they do at the source. Note that this effect is less severe when considering only taxa endemic to the sink. Second, when the source-sink relationship is relaxed, the age distributions in the two regions become more similar (*dashed lines*, fig. 3). In particular, origination at the sink increases the proportion of young taxa at the sink, and this is especially so for endemics. The differences in age distributions between source and sink regions do, however, hold over a wide range of parameter values, demonstrating that such age distributions can be a robust means of inferring source or sink properties of a region.

Because our model gives a quantitative description of the expected ages of taxa, we can use it to estimate rates

of origination, extinction, and dispersal from data on taxon ages. Next we discuss the application of this model to biogeographic and paleontologic data on marine bivalves.

### Application to Marine Bivalves

Polar regions of the world's oceans contain significantly fewer species and higher taxa than temperate or tropical areas. While many hypotheses have been proposed to explain why polar regions have so few taxa (Fischer 1960; Connell and Orias 1964; Crame 1992; Rohde 1992, 1999; Rosenzweig 1995; Blackburn and Gaston 1996; Willig et al. 2003; Currie et al. 2004), the evolutionary basis for this pattern remains poorly understood. Wallace (1878) was among the first to argue that the low diversity of the polar regions is largely a reflection of past episodes of glaciations and climatic change that repeatedly drove many high-latitude taxa to extinction, leaving little opportunity for diversity to recover, and this idea has had subsequent proponents (Fischer 1960; Skelton et al. 1990). However, empirical studies provide at best equivocal support for the idea that polar regions are characterized by significantly higher extinction rates compared with temperate or tropical areas (Raup and Jablonski 1993; Crame and Clarke 1997; Crame 2002). An alternative view is that the low diversity of polar regions results from low origination rates there, but again, empirical tests of this idea in the marine realm have produced inconclusive results (Crame and Clarke 1997; Crame 2002).

A central assumption of many previous attempts to explain the differences in diversity between high and low latitudes is that these differences reflect in situ differences in macroevolutionary rates. They either implicitly or explicitly exclude the possibility that such changes in diversity could result from past shifts in the geographic distributions of taxa (Fischer 1960; Stehli et al. 1969; Stenseth 1984; Flessa and Jablonski 1996; Cardillo 1999; Currie et al. 2004; but see Valentine 1968; Hecht and Agan 1972; Gaston and Blackburn 1996; Rosenzweig and Sandlin 1997). Yet there is overwhelming evidence for shifts in geographic distributions of species and higher taxa, not only in response to climate changes (Peters and Lovejoy 1992; Jackson and Overpeck 2000; Roy et al. 2001) but as invaders crossing climatic gradients (Vermeij 1991; Jablonski and Sepkoski 1996), and such shifts over geologic time may be an important determinant of large-scale biodiversity patterns (Wiens and Donoghue 2004). Using our model, we test the relative importances of origination, extinction, and dispersal in determining polar marine bivalve diversity.

### The Data

Our analyses are based on 459 genera of marine bivalves living on the continental shelves (depth <200 m). These taxa belong to 14 of the 41 living superfamilies of bivalves and represent about half of the 958 living bivalve genera with a fossil record. We estimated the geological ages of individual genera using an existing database (Jablonski et al. 2003). Geographic distribution of each genus was obtained from an updated version of the data used by Flessa and Jablonski (1996). We then characterized each genus as being present only in the polar regions (defined as poleward of 60° north or south latitude), outside the polar regions, or in both areas. The superfamilies used here are less well represented in the Southern Hemisphere and so our polar data are predominantly from the Northern Hemisphere. Hence, instead of analyzing polar regions of the two hemispheres separately, we combined the data into one polar unit in our analyses. Previous studies have suggested interesting differences in evolutionary dynamics between the northern and southern polar regions (Clarke and Crame 1997, 2003), but the nature of our data prevents us from exploring these differences. We also used an updated version (Jablonski et al. 2003) of data from the Sepkoski (2002) compendium to determine an initial condition for the model as discussed in the next section.

### Model Fit to Data

To estimate rates of dispersal, extinction, and origination of genera, we fit our model to these data using all genera of age 65 million years or less; older genera were not used because the end-Cretaceous mass extinction would severely violate the assumption of time-independent rates. We let  $R_\beta$  refer to the polar region above 60°N latitude and below 60°S latitude, and  $R_\alpha$  refers to the tropical and midlatitude regions between 60°N and 60°S. We emphasize that we fit to the general version of the model (fig. 1A), and so we did not preassign source or sink characteristics to either region.

We used maximum likelihood to estimate the rates. The joint likelihood function and maximization procedure are described in the appendix in the online edition of the *American Naturalist*. We also used the method of least squares to estimate the rates as described in the appendix. Maximum likelihood and least squares emphasize different aspects of the data and make different assumptions, but they yielded nearly identical parameter estimates. We prefer the maximum likelihood approach because it does not require binning the data and therefore takes better advantage of the information available.

For the initial condition,  $\mathbf{n}(t_0)$ , we used data from Jablonski et al. (2003) and Sepkoski (2002) to determine the

number of genera that survived the end-Cretaceous mass extinction and belonged to families in the biogeographic data set. These genera are not included in the data set from which we estimate parameters, even if they are alive today, because they are older than 65 million years. Because biogeographic information for these genera was lacking and because there is evidence that the impact of the extinction on bivalves was globally uniform (Raup and Jablonski 1993), we assumed that they were distributed in the same proportions as present-day diversity. Setting  $t' = 0$  and  $t_0 = -65$  million years, the initial condition was thus  $\mathbf{n}(-65) = (9, 54, 0)$ . Reasonable deviations from this assumption, including the presence of four or five polar endemics (Marincovich 1993), were also considered. These gave parameter estimates within 10% of the nonzero parameter estimates or within the confidence intervals (CIs) of the zero estimates reported in table 1.

A parametric bootstrap was used to calculate a 95% CI for each parameter and to assess the bias and covariance of the parameter estimates (details in the appendix). The maximum likelihood parameter estimates and their CIs are given in table 1. Different widths of the CIs reflect differences in the sensitivity of the model to each parameter.

Significant differences exist between the two regions in the per-genus rates of origination, extinction, and dispersal, as shown in table 1. The rate of origination of new genera is significantly lower in the polar regions than at lower latitudes. The rate at which genera move from polar regions to lower latitudes is significantly lower than the rate of movement in the opposite direction. The rate of extinction of genera also appears higher in the polar regions than at lower latitudes.

To help visualize the fit, figure 4 presents age distributions of the data and of the model with the estimated parameter values. We assembled the data into two age distribution histograms, one for all genera in the polar regions,  $R_\beta$ , and one for all genera at lower latitudes,  $R_\alpha$ . Each of these age distributions is shown with 12 bins of equal width spanning ages from 0 to 65 million years. Using the parameter estimates, we formed age distributions from equation (6). These were in continuous time, so the integral of  $f(\tau)$  over each bin was computed for comparison with the binned data.

In figure 4, the data showed considerable scatter around the model, raising potential concerns about the applicability of this model to these data and especially bringing into question the assumption of constant rates over time. However, the CIs produced by the parametric bootstrap (*dotted lines*, fig. 4; methods in the appendix) show that much of this scatter can be explained by the stochastic nature of the origination-extinction-dispersal process. We cannot, however, entirely rule out the possibility that some

**Table 1:** Parameter estimates

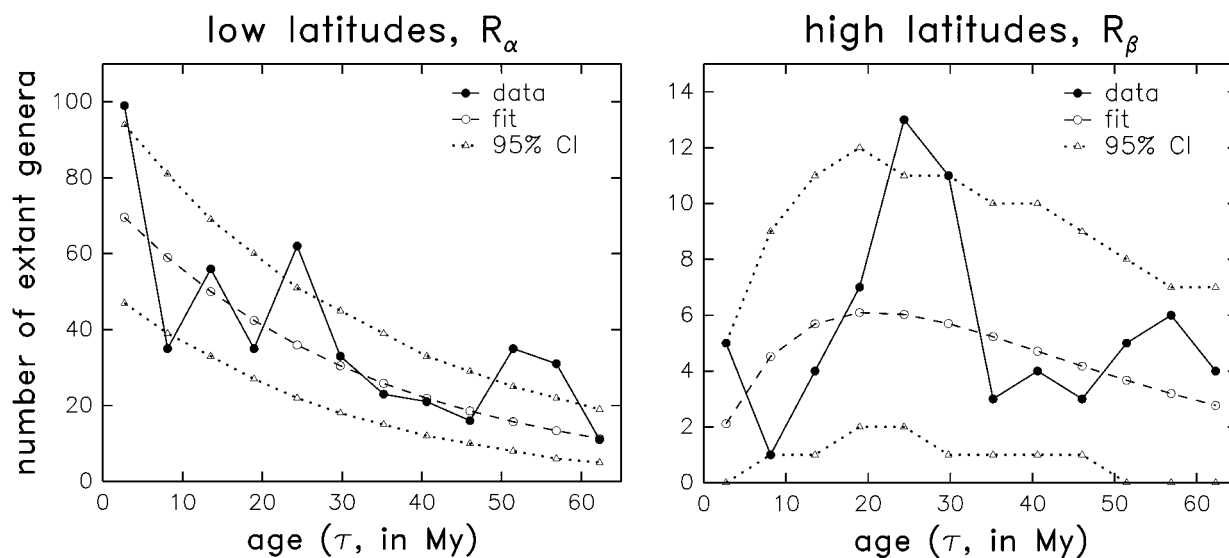
Parameter	ML estimate (genus <sup>-1</sup> Ma <sup>-1</sup> )	95% CI (genus <sup>-1</sup> Ma <sup>-1</sup> )
$s_\alpha$	.03049	(.02886, .03564)
$s_\beta$	.00088	(.00042, .00252)
$x_\alpha$	.00000	(.00000, .00022)
$x_\beta$	.02956	(.00030, .12212)
$d_\alpha$	.01069	(.00633, .02421)
$d_\beta$	.00000	(.00000, .00004)

Note: ML = maximum likelihood; CI = confidence interval. Using the CI of the difference between parameters, we find significant differences between  $s_\alpha$  and  $s_\beta$  ( $P < .005$ ) and between  $d_\alpha$  and  $d_\beta$  ( $P < .001$ ) and a marginally significant difference between  $x_\alpha$  and  $x_\beta$  ( $P = .03$ ). The methods used for calculating CIs and  $P$  values are described in the appendix in the online edition of the *American Naturalist*.

of the variation is caused by rate heterogeneity or possible sampling effects, and future work could explore the effects of such heterogeneities on the model's predictions.

## Discussion

Identifying the role of historical processes in shaping biogeographic patterns of species diversity seen today remains a challenging problem. Direct quantitative tests of historical hypotheses require a complete fossil or historical record that is absent for the vast majority of living taxa. In addition, there is little quantitative theory relating historical processes to present-day biogeographic patterns. The model we present here is an attempt to formulate such a theoretical framework. Our model connects the processes of taxon origination, extinction, and dispersal with present-day regional diversity, endemism, and age distributions. Our results question some widely held assumptions regarding the relationship between endemism, diversity, and origination rates. In particular, we use this model to show that a region that is a macroevolutionary source or "center of origin" need not necessarily have high levels of diversity and endemism, as is often assumed (Willis 1922; Rosenzweig and Sandlin 1997; Mora et al. 2003; reviews by McCoy and Heck [1976]; Ricklefs and Schluter [1993]), but that it must have a high proportion of young taxa. Conversely, a macroevolutionary sink or "center of accumulation" need not necessarily have low levels of diversity and endemism, but its age distribution often will have a single intermediate peak. Our conclusion is therefore that regional measures of diversity and endemism are not sufficient to estimate average regional evolutionary rates but that age distributions often are. By fitting this model to data on extant marine bivalves, we are able to estimate regional rates of origination, extinction, and dispersal, and we show that polar regions have on average a



**Figure 4:** Age distributions of marine bivalve genera and the model fit, using the maximum likelihood parameter estimates given in table 1. Left panel shows all genera in  $R_\alpha$ ; right panel shows all genera in  $R_\beta$ . Filled circles are the data, and open circles are the model, which is  $\int_{\text{bin}} [f_1(\tau) + f_2(\tau)] d\tau$  for  $R_\alpha$  and  $\int_{\text{bin}} [f_1(\tau) + f_3(\tau)] d\tau$  for  $R_\beta$ . The open triangles show approximate 95% confidence intervals on the data determined by the parametric bootstrap, as described in the appendix in the online edition of the *American Naturalist*. The quality of the fit is discussed in the text. The maximum likelihood fitting procedure did not require binning the data; these histograms are used only to display the results.

lower rate of origination, a higher rate of extinction, and a higher rate of immigration of genera than do the lower latitudes.

Our model is in the same spirit as other “neutral” models of biogeography and diversity, notably MacArthur and Wilson’s theory of island biogeography (MacArthur and Wilson 1967) and Hubbell’s unified neutral theory (Hubbell 2001). These share the assumption that the “units” that are considered (species or higher taxonomic units in our case, species in the case of island biogeography, and individuals of all species in the case of unified neutral theory) are all equivalent: there are no intrinsic individual or species differences. Hubbell describes his theory as an extension of MacArthur and Wilson’s theory because he considers not only extinction and immigration of species from a “mainland” pool or “metacommunity” but also speciation in the metacommunity and relative abundances of species. Our model can also be seen as an extension of the theory of island biogeography but in a different direction. Like it, we consider only presence or absence of taxa, not abundances, and like unified neutral theory, we include origination of new taxa. We differ in considering our two regions to be functionally symmetric (though perhaps with different parameter values) rather than designating one area as a mainland pool or metacommunity and one as an island or local community. This enables us to consider relative levels of diversity and endemism be-

tween the regions as functions of regional rates. Unlike previous work, we use the ages of taxa to infer regional rates.

In addition to our “neutral” assumption that all taxa behave equally, our model makes the significant assumption that regional rates of origination, extinction, and dispersal do not change over time. This restriction is not quite as strong as it seems because the requirement is only that the average rates over an entire region remain constant. If, however, the rate parameters were explicit functions of time, equation (1) would still hold but equation (2) would not be its solution, and the subsequent equations for diversity and age distributions would not be valid. To our knowledge, an analytic form of this model cannot be obtained for general time-dependent rates, but specific situations of interest could be investigated numerically or through simulations.

The validity of the assumptions of taxon equivalence and constant rates can be addressed on two levels. The model was developed to identify criteria, based on extant taxa, that can or cannot be used to infer the magnitudes of evolutionary rates. For questions like what the indications are that a region has a high rate of origination, our assumptions are appropriate because the issue is one of average rates over time and over taxa.

It is when we apply our model to data that the validity of the assumptions becomes more important. We took



the obvious precaution of restricting the data set to the last 65 million years to avoid the end-Cretaceous mass extinction, which would be a serious violation of the constant-rates assumption. Inspection of figure 4 indicates that the period 50–60 million years ago may have had higher origination rates, perhaps reflecting a rebound from the end-Cretaceous extinction (Flessa and Jablonski 1996; Jablonski 1998), and 25 million years ago also may have been a time of greater origination. An important observation here is that these anomalies are present in both regions. Given that the average origination rate is two orders of magnitude greater in the lower latitudes (table 1), the anomalous peaks in the polar regions most likely resulted from subsequent dispersal of taxa that originated in the lower latitudes. This highlights the importance of taking into account past dispersal patterns in interpreting present-day regional age distributions: if we assumed that the current distributions of taxa reflected their places of origin, as is commonly done, we would have concluded that both regions had high in situ origination rates during these times.

It is possible that geographic differences in the nature of the fossil record to underestimate taxon ages could introduce a bias into our results. The poorer quality of the fossil record in the tropics (Van Valen 1969; Johnson 2003) could lead to greater underestimation of ages for lower-latitude taxa; this bias could therefore add false support to our conclusion of higher origination rates at lower latitudes. This is unlikely to produce a large effect here, since we defined our “low-latitude” region as both the tropics and also the well-sampled temperate region to take into account such sampling problems. In principle, our approach can be used to compare tropical versus extratropical regions, as many previous studies have done (Stehli et al. 1969; Jablonski 1993; Flessa and Jablonski 1996; Gaston and Blackburn 1996; among many others). However, more work is needed to improve sampling and taxonomic standardization of the tropical fossil record before rigorous analyses are feasible. Similarly, a more complete data set of the ages of taxa living in high-latitude southern oceans would be useful for exploring the differences in the evolutionary dynamics between the two polar oceans (Clarke and Crame 1997, 2003).

The excellent fossil record for marine bivalves makes it possible to determine regional macroevolutionary rates and range shifts explicitly, and some of this has been done (Vermeij 2001). Such analyses, however, are not possible for many other taxa, and we hope that the approach taken by our model may be useful for cases in which only more limited information about extant taxa is available. In particular, it would be quite useful if this method could be applied to the rapidly increasing number of taxa for which phylogenetic trees and estimates of branching times are

available. There is a large body of work (Nee et al. 1994; Pybus and Harvey 2000; among others) on estimating rates of origination and extinction from branching times, but this does not allow consideration of differences in rates between regions. Our model does not require an explicit phylogeny, but we assume that an origination event creates one daughter taxon and leaves the age of the parent taxon unaffected, as is the convention with phylogenies determined from the fossil record. In phylogenies determined from molecular data, taxa do not have absolute ages and the most recent branching time of a lineage therefore depends on the survival of potential sister taxa (Gaston and Blackburn 1996). Because of this difference, modification of our model would be necessary in order to apply it to lineage ages from this second kind of phylogeny, but simulations (not shown) do suggest that similar patterns in age distributions will hold.

The spatial extent of our system was the entire globe and our data were at the level of genera, but the model presented here could be applied to closed systems of two regions on any spatial or taxonomic scale. This model could also be extended in a straightforward manner to multiple regions. This could allow, for example, quantitative description of expected age distributions in a region where diversity is elevated by the overlap of biogeographic provinces, or it could lead to a model for expected distributions of range sizes.

In general, findings for marine bivalve genera clearly show that shifts in geographic ranges can play an important role in determining global patterns of biodiversity. Future attempts to estimate regional origination and extinction rates for any taxon therefore should not be based explicitly on the assumption of in situ origination, and the possible effects of dispersal between regions should be evaluated. Spatially explicit models, in which in situ processes interact with biotic interchanges, should prove important for our understanding of past and future dynamics of biological diversity.

#### Acknowledgments

We thank T. J. Case, P. Fenberg, G. Hunt, S. B. Menke, and B. Pister for helpful discussions; W. F. Briskin and J. P. Huelsenbeck for statistical advice; and F. He and an anonymous reviewer for comments on the manuscript. This work was supported by a National Science Foundation (NSF) Graduate Research Fellowship to E.E.G. and NSF grants to R.L., K.R., and D.J.

#### Literature Cited

- Apostol, T. M. 1969. *Calculus*. Vol. 2. Wiley, New York.  
Blackburn, T. M., and K. J. Gaston. 1996. A sideways look at patterns

- in species richness, or why there are so few species outside the tropics. *Biodiversity Letters* 3:44–53.
- Briggs, J. C. 1999. Coincident biogeographic patterns: Indo-West Pacific Ocean. *Evolution* 53:326–335.
- . 2004. A marine center of origin: reality and conservation. Pages 255–269 in M. V. Lomolino and L. R. Heaney, eds. *Frontiers of biogeography: new directions in the geography of nature*. Sinauer, Sunderland, MA.
- Brown, J. H., and M. V. Lomolino. 1998. *Biogeography*. Sinauer, Sunderland, MA.
- Cardillo, M. 1999. Latitude and rates of diversification in birds and butterflies. *Proceedings of the Royal Society of London B* 1425: 1221–1225.
- Chown, S. L., and K. J. Gaston. 2000. Areas, cradles and museums: the latitudinal gradient in species richness. *Trends in Ecology & Evolution* 15:311–315.
- Clarke, A., and J. A. Crame. 1997. Diversity, latitude and time: patterns in the shallow sea. Pages 122–147 in R. F. G. Ormond, J. D. Gage, and M. V. Angel, eds. *Marine biodiversity: patterns and processes*. Cambridge University Press, New York.
- . 2003. The importance of historical processes in global patterns of diversity. Pages 130–151 in T. M. Blackburn and K. J. Gaston, eds. *Macroecology: concepts and consequences*. Blackwell, Malden, MA.
- Connell, J. H. and E. Orias. 1964. The ecological regulation of species diversity. *American Naturalist* 98:399–414.
- Crame, J. A. 1992. Evolutionary history of the polar regions. *Historical Biology* 6:37–60.
- . 2002. Evolution of taxonomic diversity gradients in the marine realm: a comparison of late Jurassic and Recent bivalve faunas. *Paleobiology* 28:184–207.
- Crame, J. A., and A. Clarke. 1997. The historical component of marine taxonomic diversity gradients. Pages 258–273 in R. F. G. Ormond and J. A. Gage, eds. *Marine biodiversity patterns and processes*. Cambridge University Press, New York.
- Currie, D. J., and A. P. Francis. 2004. Regional versus climate effect on taxon richness in angiosperms: reply to Qian and Ricklefs. *American Naturalist* 163:780–785.
- Currie, D. J., G. Mittelbach, H. V. Cornell, R. Field, J.-F. Guégan, B. A. Hawkins, D. M. Kaufman, et al. 2004. Predictions and tests of climate-based hypotheses of broad-scale variation in taxonomic richness. *Ecology Letters* 7:1121–1134.
- Darwin, C. (1859) 1975. *On the origin of species by means of natural selection*. Harvard University Press, Cambridge, MA. Original edition, J. Murray, London.
- Efron, B., and R. Tibshirani. 1986. Bootstrap methods for standard errors, confidence intervals, and other measures of statistical accuracy. *Statistical Science* 1:54–75.
- Fischer, A. G. 1960. Latitudinal variations in organic diversity. *Evolution* 14:64–81.
- Flessa, K. W., and D. Jablonski. 1996. The geography of evolutionary turnover: a global analysis of extant bivalves. Pages 376–397 in D. Jablonski, D. H. Erwin, and J. H. Lipps, eds. *Evolutionary paleobiology*. University of Chicago Press, Chicago.
- Foote, M. 2001. Evolutionary rates and the age distributions of living and extinct taxa. Pages 245–294 in J. B. C. Jackson, S. Lidgard, and F. K. McKinney, eds. *Evolutionary patterns: growth, form, and tempo in the fossil record*. University of Chicago Press, Chicago.
- Forney, K. A., and J. Barlow. 1998. Seasonal patterns in the abundance and distribution of California cetaceans, 1991–1992. *Marine Mammal Science* 14:460–489.
- Francis, A. P., and D. J. Currie. 1998. Global patterns of tree species richness in moist forests: another look. *Oikos* 81:598–602.
- . 2003. A globally consistent richness-climate relationship for angiosperms. *American Naturalist* 161:523–536.
- Gaston, K. J., and T. M. Blackburn. 1996. The tropics as a museum of biological diversity: an analysis of the New World avifauna. *Proceedings of the Royal Society of London B* 263:63–68.
- Hecht, A. D., and B. Agan. 1972. Diversity and age relationships in Recent and Miocene bivalves. *Systematic Zoology* 21:308–312.
- Hubbell, S. P. 2001. *The unified neutral theory of biodiversity and biogeography*. Princeton University Press, Princeton, NJ.
- Jablonski, D. 1993. The tropics as a source of evolutionary novelty through geologic time. *Nature* 364:142–144.
- . 1998. Geographic variation in the molluscan recovery from the end-Cretaceous extinction. *Science* 279:1327–1330.
- Jablonski, D., and J. J. Sepkoski Jr. 1996. Paleobiology, community ecology, and scales of ecological pattern. *Ecology* 77:1367–1378.
- Jablonski, D., K. Roy, J. W. Valentine, R. M. Price, and P. S. Anderson. 2003. The impact of the pull of the Recent on the history of marine diversity. *Science* 300:1133–1135.
- Jackson, S. T., and J. T. Overpeck. 2000. Responses of plant populations and communities to environmental changes of the late Quaternary. *Paleobiology* 26(suppl.):194–220.
- Johnson, K. G. 2003. New data for old questions. *Paleobiology* 29: 19–21.
- Ladd, H. S. 1960. Origin of the Pacific island molluscan fauna. *American Journal of Science* 258-A:137–150.
- Lo, N. C. H. 1994. Level of significance and power of two commonly used procedures for comparing mean values based on confidence intervals. *California Cooperative Oceanic Fisheries Investigations Reports* 35:246–253.
- MacArthur, R. H., and E. O. Wilson. 1967. *The theory of island biogeography*. Princeton University Press, Princeton, NJ.
- Marincovich, L., Jr. 1993. Danian mollusks from the Prince Creek formation, northern Alaska, and implications for Arctic Ocean paleogeography. *Journal of Paleontology* 67(suppl.):1–35.
- McCoy, D. E., and K. L. Heck Jr. 1976. Biogeography of corals, seagrasses and mangroves: an alternative to the center of origin concept. *Systematic Zoology* 25:201–210.
- Mora, C., P. M. Chittaro, P. F. Sale, J. P. Kritzer, and S. A. Ludsin. 2003. Patterns and processes in reef fish diversity. *Nature* 421:933–936.
- Nee, S., R. M. May, and P. H. Harvey. 1994. The reconstructed evolutionary process. *Philosophical Transactions of the Royal Society of London B* 344:305–311.
- Nelder, J. A., and R. Mead. 1965. A simplex method for function minimization. *Computer Journal* 7:308–313.
- Palumbi, S. R. 1996. What can molecular genetics contribute to marine biogeography? an urchin's tale. *Journal of Experimental Marine Biology and Ecology* 203:75–92.
- Pandolfi, J. M. 1992. Successive isolation rather than evolutionary centres for the origination of Indo-Pacific reef corals. *Journal of Biogeography* 19:593–609.
- Pease, C. M. 1988. Biases in the survivorship curves of fossil taxa. *Journal of Theoretical Biology* 130:31–48.
- Peters, R. L., and T. E. Lovejoy. 1992. *Global warming and biological diversity*. Yale University Press, New Haven, CT.

- Press, W. H. 1992. Numerical recipes in C. Cambridge University Press, New York.
- Pybus, O. G., and P. H. Harvey. 2000. Testing macro-evolutionary models using incomplete molecular phylogenies. *Proceedings of the Royal Society of London B* 267:2267–2272.
- Qian, H., and R. E. Ricklefs. 2004. Taxon richness and climate in angiosperms: is there a globally consistent relationship that precludes region effects? *American Naturalist* 163:773–779.
- Raup, D. M., and D. Jablonski. 1993. Geography of end-Cretaceous marine bivalve extinctions. *Science* 260:971–973.
- Ricklefs, R. E. 2004. A comprehensive framework for global patterns in biodiversity. *Ecology Letters* 7:1–15.
- Ricklefs, R. E., and D. Schluter. 1993. Species diversity: regional and historical influences. Pages 350–363 *in* R. E. Ricklefs and D. Schluter, eds. *Species diversity in ecological communities*. University of Chicago Press, Chicago.
- Rohde, K. 1992. Latitudinal gradients in species diversity: the search for the primary cause. *Oikos* 65:514–527.
- . 1999. Latitudinal gradients in species diversity and Rapoport's rule revisited: a review of recent work and what can parasites teach us about the causes of the gradients? *Ecography* 22:593–613.
- Rosenzweig, M. L. 1995. *Species diversity in space and time*. Cambridge University Press, Cambridge.
- Rosenzweig, M. L., and E. A. Sandlin. 1997. Species diversity and latitudes: listening to area's signal. *Oikos* 80:172–176.
- Roy, K., D. J. Jablonski, and J. W. Valentine. 2001. Climate change, species range limits and body size in marine bivalves. *Ecology Letters* 4:366–370.
- Sepkoski, J. J., Jr. 1998. Rates of speciation in the fossil record. *Philosophical Transactions of the Royal Society of London B* 353:315–326.
- . 2002. A compendium of fossil marine animal genera. *Bulletin of American Paleontology* 363:1–560.
- Skelton, P. W., J. A. Crame, N. J. Morris, and E. M. Harper. 1990. Adaptive divergence and taxonomic radiation in post-Palaeozoic bivalves. Pages 91–117 *in* P. D. Taylor and G. P. Larwood, eds. *Major evolutionary radiations*. University of Chicago Press, Chicago.
- Stanley, S. 1985. Rates of evolution. *Paleobiology* 11:13–26.
- Stebbins, G. L. 1974. *Flowering plants: evolution above the species level*. Harvard University Press, Cambridge, MA.
- Stehli, F. G., and J. W. Wells. 1971. Diversity and age patterns in hermatypic corals. *Systematic Zoology* 20:115–126.
- Stehli, F. G., R. G. Douglas, and N. D. Newell. 1969. Generation and maintenance of gradients in taxonomic diversity. *Science* 164:947–949.
- Stenseth, N. C. 1984. The tropics: cradle or museum? *Oikos* 43:417–420.
- Valentine, J. W. 1968. Climatic regulation of species diversification and extinction. *Geological Society of America Bulletin* 79:273–276.
- Van Valen, L. 1969. Climate and evolutionary rate. *Science* 166:1656–1658.
- . 1973. A new evolutionary law. *Evolutionary Theory* 1:1–30.
- Vermeij, G. J. 1991. When biotas meet: understanding biotic interchange. *Science* 253:1099–1104.
- . 2001. Community assembly in the sea: geologic history of the living shore biota. Pages 39–60 *in* M. D. Bertness, S. D. Gaines, and M. E. Hay, eds. *Marine community ecology*. Sinauer, Sunderland, MA.
- Wallace, A. R. 1878. *Tropical nature and other essays*. Macmillan, London.
- Wiens, J. J., and M. J. Donoghue. 2004. Historical biogeography, ecology and species richness. *Trends in Ecology & Evolution* 19:639–644.
- Willig, M. R., D. M. Kaufman, and R. D. Stevens. 2003. Latitudinal gradients of biodiversity: pattern, process, scale, and synthesis. *Annual Review of Ecology, Evolution, and Systematics* 34:273–309.
- Willis, J. C. 1922. *Age and area*. Cambridge University Press, London.

Editor: Jonathan B. Losos  
Associate Editor: Steven L. Chown