THE MACROECOLOGY OF CYPRINELLA: CORRELATES OF PHYLOGENY, BODY SIZE, AND GEOGRAPHICAL RANGE

CHRISTOPHER M. TAYLOR* AND NICHOLAS J. GOTELLI†

*Department of Zoology, University of Oklahoma, Norman, Oklahoma 73019;
†Department of Zoology, University of Vermont, Burlington, Vermont 05405

Submitted December 21, 1992; Revised November 2, 1993; Accepted November 9, 1993

Abstract.—Macroecologists interpret correlates of body size and geographical range size in an ecological context, but these patterns may also reflect historical or phylogenetic forces. We examined the relationship between range size and body size for a monophyletic group of 27 North American minnow species. Body size and range size were positively correlated, but both variables were also correlated with latitude. After controlling for effects of latitude, body size and range size were no longer correlated. The basal dichotomy of the cladogram defined an eastern and western clade; they differed in their geological and climatic histories and macroecological patterns. Within the western clade, only Bergmann’s rule was confirmed. Within the eastern clade, both longitude and latitude of geographical range were positively correlated with body size. A simple measure of phylogeny was correlated with range size: species branching near the cladogram root had larger geographical ranges than species branching distally. After statistically removing effects of latitude, longitude, and phylogeny, there was a significant positive correlation between body size and range size. Macroecological patterns are sensitive to phylogeny and speciation history, and they may be most informative for clades that occupy areas with a common climatic history.

Practitioners in the emerging field of macroecology seek to understand the partitioning of physical space and ecological resources by species (Brown and Maurer 1989). In macroecological analyses, individual species function as replicates in searches for correlated patterns of geographical range size, body size, population density, trophic status, and intrinsic rate of increase (Damuth 1981; Brown and Maurer 1987; Gaston and Lawton 1988a, 1988b; Brown and Maurer 1989; Lawton 1990). Analyses are typically carried out on very broad taxonomic and ecological groups, such as on the breeding birds of North America (Brown and Maurer 1987) or the herbivorous insects that feed on bracken (Gaston and Lawton 1988b).

One pattern that frequently emerges from such analyses is a correlation between the mean body size of animal species and the size of their geographical ranges. Brown (1981; Brown and Gibson 1983; Brown and Maurer 1987) predicted a positive relationship between body size and geographical range size. Brown and Maurer (1987) contended that the maximum possible geographical range size of a species is set by spatial constraints (e.g., size of continent for birds), while the minimum geographical range size is set by the minimum viable population size of the species. Thus, the minimum size of a geographical range must in-
crease with body size because large-bodied species have relatively large home-range requirements (McNab 1963; Schoener 1968) and low population densities (Damuth 1981). Consequently, a large-bodied species with a small geographical range would have a small total population size and a high probability of extinction. Although there must exist some upper limit to the density of large-bodied species that is lower than that for small-bodied species, Blackburn et al. (1990) found no relationship between population density and body size for several guilds of taxonomically close bird and beetle species. They argued that body size–population density relationships may often be artifacts due to the relative rareness of species at extreme size ranges.

Gaston (1990) suggested a second mechanism that could lead to a positive correlation between body size and geographical range. If small-bodied species are sensitive to density-independent perturbations such as temperature fluctuations or other climatic disturbances, they may be unable to persist over wide geographical areas (Gotelli and Graves 1990). In this scenario, the differential ability of species to maintain large geographical ranges also leads to a positive correlation between body size and geographical range size.

There is empirical support for this model. The bivariate scatterplot for almost 400 species of North American birds (fig. 3 in Brown and Maurer 1987) indicates a positive relationship between body size and geographical range but does not suggest (to our eyes) a clearly defined minimum, perhaps because it represents a probabilistic boundary (Brown and Maurer 1987). Significant positive correlations between body size and geographical range size have been found for stomatopods (Reaka 1980), fish (McAllister et al. 1986), and mammals (Van Valen 1973; Brown 1981).

A theoretical argument has also been advanced for a negative correlation between body size and geographical range size (Gaston 1988; Gaston and Lawton 1988a, 1988b). The argument is based on the observation that the intrinsic rate of increase, $r$, decreases as a function of increasing body size (Fenchel 1974; Southwood 1981; Gaston 1988). Consequently, population growth of small-bodied species that colonize an empty site will be more rapid than growth of large-bodied species. Rapid population growth will allow small-bodied species to achieve a large population size in a short time, which makes them less vulnerable to stochastic extinction (MacArthur and Wilson 1967; Leigh 1981). Smaller population sizes of large-bodied species will reinforce this pattern (Gaston 1990).

Some empirical studies also support Gaston and Lawton's model. For North American Peromyscus, small-bodied species are more widespread and possess life-history traits (e.g., large litter size, high relative reproductive effort, short life span) that would predispose them to successful colonization (Glazier 1980). Small-bodied mammal species also occupy more sites than large-bodied species on isolated mountain tops in the southern Rocky Mountains (Patterson 1984), although this system is probably driven by extinction rather than colonization (Brown 1971). Finally, body size and site occupancy are negatively correlated for bracken-feeding insects (Gaston and Lawton 1988b). However, the pattern may not hold over their entire geographical ranges, which include other host plant species (Gaston 1990).
The contrasting predictions of the two models depend on whether persistence of a species is controlled by total abundance across the entire geographical range (Brown's model) or abundance and population attributes of local populations (Gaston and Lawton's model). The empirical results are diverse and seem to vary widely among taxa and geographical regions (Gaston 1990).

A difficult challenge to macroecologists is to sort out the network of highly correlated ecological variables that may be associated with body size and geographical range, including dispersal potential (Glazier 1980), climatic and habitat variation (Karr and James 1975), and longitudinal and latitudinal variation in geographical range boundaries (Brown and Maurer 1989; France 1992). For example, many taxa of both ectotherms and endotherms exhibit an increase in body size at high latitudes (Bergmann's rule; Lindsey 1966). High-latitude species also tend to have greater latitudinal spans in their geographical ranges than more tropical, low-latitude species (Rapoport's rule as designated in Stevens 1989; Rapoport 1982; France 1992). Thus, a positive correlation between geographical range size and body size could result because of underlying latitudinal gradients in each variable (Pagel et al. 1991).

Many macroecological studies have relied on qualitative interpretations of bivariate scatterplots (Brown 1981; Brown and Maurer 1987; Lawton 1990). More quantitative path analyses (e.g., fig. 1 in Gaston and Lawton 1988b) and regression models (Sokal and Rohlf 1981) may reveal more complex ecological relationships involving body size and geographical range.

A more serious concern is that macroecologists have not emphasized phylogenetic and historical processes in their explanations. Phylogenetic processes include correlations and constraints on body size imposed by an evolutionary lineage (Pagel and Harvey 1988). As summarized by Harvey and Pagel (1991), a variety of comparative methods have been developed for analyzing the relationship between two continuous variables (Felsenstein 1985; Huey and Bennett 1987). Analysis of phylogenetic processes requires a cladogram, a branching hierarchical tree that expresses relationships among taxa and is based on shared derived characters (i.e., synapomorphies; Wiley 1981). Phylogenetic effects have been demonstrated for body size (Elgar and Harvey 1987) and other ecological attributes of monophyletic groups (Wanntorp et al. 1990; Brooks and McLennan 1991). Similarly, geographical range size is probably influenced by speciation mode (Lynch 1989) and may even behave as a heritable, species-level trait (Jablonski 1987). The ecological significance of correlations between body size and geographical range cannot be assessed fully unless phylogenetic effects are evaluated simultaneously.

Unfortunately, it has not been easy to study the joint effects of ecology and history in biogeography. Endler (1982, p. 451) suggested one worthwhile avenue: "One possible method would be to explore the ecological factors in sufficient detail so that these factors could be removed, leaving components of the distributions which presumably reflect historical events, but this is a formidable task." Because our goal is an understanding of the relationship between body size and geographical range size, we have taken the opposite approach. Namely, we use phylogeny as a "null hypothesis" (Kochmer and Handel 1986) to try to account
for variation in geographical range size and body size. After controlling for phylogenetic effects, significant correlations between body size and geographical range size are more likely to reflect underlying ecological processes.

An underlying assumption of any comparative study is that taxa represent equivalent and comparable units (Mishler and Donoghue 1982). Mishler and Donoghue argue that this is rarely possible for species taxa because of biological differences among lineages and practicing differences among taxonomists. However, speciation is a central process in evolutionary theory; therefore, species must be real in some evolutionary sense (Brooks and McLennan 1991). We recognize that species boundaries for many groups of organisms may be problematic for comparative biologists. Our analyses are based on a monophyletic subset of the fishes of North America (genus *Cyprinella*), for which we have independent measures of geographical range and body size, as well as a completely resolved phylogeny. Numerous systematic studies exist for *Cyprinella*, and species-level problems have been reasonably well resolved (Mayden 1989). Thus, we feel justified in using the 27 recognized species of *Cyprinella* as phylogenetically comparable units.

**MATERIAL AND METHODS**

We used Mayden's (1989) fully resolved cladogram for the 27 species of *Cyprinella* (fig. 1) as a provisional phylogeny in our analyses. This cladogram is based on 206 osteological and morphological characters. As in all analyses of this sort, the cladogram does not necessarily depict the "true" phylogeny, but it is a parsimonious hypothesis of relationship based on the distribution of shared derived characters of the terminal taxa (Wiley 1981).

We collected most of the data on geographical distribution from Lee et al. (1980), which provides a comprehensive account of the distribution and body size (maximum standard length) of freshwater North American fishes. For most of the Mexican species we relied on other works, including Minckley and Lytle (1969), Lytle (1972), Contreras-Balderas (1975), Hubbs and Miller (1978), Chernoff and Miller (1982), and Mayden and Hillis (1990). Pfieger (1975), Cloutman and Harrell (1987), Robison and Buchanan (1988), and Page and Burr (1991) were used for additional information on body size. There are problems concerning the best measure of adult body size for fishes. Minimum adult reproductive size is unknown for most fish species, and interpopulation size variability is common (Lee et al. 1980). Since "normal" size is subject to so much variation, maximum adult body size may be a better measure for interspecific comparisons (Lee et al. 1980). For *Cyprinella*, maximum standard length was the only measure of body size available for all 27 species.

Establishing the geographical range of a species is not a simple task. Rapoport (1982, p. 1) said that determining the geographical distributions of species is like "measuring, weighing, and studying the behavior of ghosts." Determining the geographical ranges of freshwater fishes is especially challenging, because of the hierarchical spatial pattern of interconnected river basins. Moreover, the distribution of occupied sites for many species is dynamic; local colonizations and extinctions may be common (Starrett 1951; Harrell 1978; Grossman et al.
1982; Matthews 1986; Angermeier and Schlosser 1989). For these reasons, we used Gaston’s (1991) definition of geographical range as the “extent of occurrence,” measured by drawing a smooth border around each species’ entire range (Anderson 1984). We did not include recent introductions or extinctions. Measured in this way, the extent of occurrence does not indicate the area or concentration of occupied sites or the metapopulation structure (Hanski and Gilpin 1991) within the geographical range. Range edges delineated in this way are admittedly crude but may correspond to long-term limits of climatic or environmental tolerance (Gaston 1991). When disjunctions separated by more than 100 km occurred, the disjunct areas were measured separately, then added together. The center of each species’ geographical range was defined as the longitudinal and latitudinal midpoint of the geographical range.

**STATISTICAL ANALYSES**

We used both an absolute and a relative metric of phylogeny in analyses along with other, more traditional macroecological variables. The absolute measure of phylogeny employed for each species was the number of nodes it is separated
from the cladogram root. This number provides a simple index of "primitiveness" (Gotelli and Pyron 1991). In a parsimonious cladogram, species close to the cladogram root will be characterized by many primitive (pleisomorphic) character states, whereas species distant from the root will exhibit more derived (apomorphic) character states, discounting all autapomorphies for species. Norell and Novacek (1992) found a correspondence between the fossil record and this index (equivalent to clade rank in their article) for many vertebrate groups, including teleost fishes. We note that the use of this index does not overcome a fundamental problem in comparative analyses: species are still treated as independent data points (Felsenstein 1985). The primitiveness index also ignores extinction of species and phyletic change in the absence of speciation. Its advantage is that it is a simple cladistic metric that can be introduced into multiple-regression analyses to compete with other ecological predictor variables.

For a relative measure of phylogeny and relatedness, we used the number of nodes separating each pair of terminal (extant) taxa (see Farris 1969). For each ecological variable, such as geographical range center, we also measured the distance (or difference, in kilometers) between the two terminal taxa. These data were organized into two $27 \times 27$ symmetric distance matrices, one for phylogeny and one for geographical distance. The Mantel test measures the degree of association between two such matrices (Mantel 1967; Schnell et al. 1985), providing a probability value for the scatterplot of all possible pairwise distances among the terminal taxa. A significant positive association indicates that closely related species have closer geographical range centers. The Mantel test is nonparametric (Schnell et al. 1985) but is sensitive to skewness and nonlinearity (Cheverud et al. 1989). Differences in geographical range size were extremely skewed, so we a priori applied a log$_{10}$ transformation. Because multiple comparisons were examined within each data set, we used the tablewide sequential Bonferroni adjustment of probability levels (Rice 1989).

We carried out combined analyses for all 27 species. In addition, we assessed separately two major subsets, an eastern *whipplei* clade and a western *lutrensis* clade. These mutually exclusive groups correspond to the branches of the first major dichotomy in the *Cyprinella* cladogram. Intercorrelation was common in all three data sets. In order to evaluate the correlation structure, we used methods from path analysis to estimate magnitudes and directions of interactions among macroecological variables (Sokal and Rohlf 1981). We developed our path models by first using stepwise multiple regression (Wilkinson 1988) to choose sets of predictor variables for body size and geographical range. In the path models, the effects of the predictor variables are shown as single-headed arrows labeled with standardized partial regression coefficients. In the case of a single significant predictor variable, the standardized partial regression coefficient is equal to the correlation coefficient. The unknown variation is also represented by single-headed arrows and represents the explanatory strength of all unaccounted for factors (Sokal and Rohlf 1981). Relationships between body size and range area were identified with double-headed arrows labeled with partial correlation coefficients, obtained by holding constant other predictor variables. We represented the relationship by a double-headed arrow because a cause-effect direction be-
tween body size and geographical range size was not apparent. Before performing the correlation and regression analyses, we checked all variables for departures from normality by examining rankit plots (Sokal and Rohlf 1981, p. 122) prepared for each variable. A log_{10} transformation was necessary to normalize geographical range area. Residuals from path models were examined and indicated that linear statistics were reasonable to describe the relationships among our macroecological variables.

Finally, we examined the frequency of various modes of speciation in Cypri- nella because speciation mode could have a direct influence on range sizes (Lynch 1989). A method developed by Lynch (1989) was used to identify putative cases of sympatric, vicariant, and peripheral-isolate speciation. Sympatric speciation refers to the origin of a new species entirely within the geographical range of its ancestor, while vicariant speciation occurs when an ancestral species is split into two or more relatively large and geographically isolated populations (Mayr 1963). Peripheral-isolate speciation is the origin of a new species from a small, isolated population, often at the periphery of the larger ancestral population or derived from dispersal (Mayr 1963). There are various extrinsic and intrinsic mechanisms that could lead to peripheral-isolate speciation (Wiley and Mayden 1985); if caused by microvicariance, then both allopatric speciation modes can be seen as end points on a continuum of possibilities (Lynch 1989).

Lynch (1989) distinguished among these three modes based on a comparison of geographical ranges of taxa that have been identified in an independent phylogenetic analysis. First, sympatric speciation is invoked if there is substantial overlap of sister taxa in their geographical ranges. Lynch (1989) did not state explicitly how much overlap is necessary before sympatric speciation is invoked; all of his examples show an overlap of 60% or more. We used a more conservative cutoff point of 75% overlap in geographical range for sympatric speciation. However, results were not sensitive to the cutoff point used; we obtained identical results with a 50% cutoff point. Second, vicariant speciation is presumed if sister taxa show little or no overlap in geographical range (< 15%) and have geographical ranges that are comparable in size (i.e., the smaller range is more than 5% of the area of the larger range). Finally, peripheral-isolates speciation is invoked if sister taxa show little overlap in geographical range and if one of the taxa (presumably the isolate) has a range area that is less than 5% of the area of the sister taxon. A further condition is that the two geographical ranges be separated by a distance that would allow for dispersal of a peripheral isolate.

Lynch’s (1989) method is controversial. Geographical ranges can expand or shrink after speciation (an implicit assumption of macroecologists) and may not be indicative of speciation mode. In particular, many authors are unwilling to equate high overlap of the geographical ranges of sister species with sympatric speciation (Wiley 1981). Moreover, it is probably misleading to infer a particular speciation process close to the cladogram root, especially between presumptive taxa that are no longer extant. For this reason, we restricted our analysis to speciation events involving at least one terminal (extant) taxa. For nonterminal sister taxa, we used the combined, nonoverlapping geographical ranges of the component species as the presumed geographical range (Lynch 1989). Finally,
TABLE 1

Correlations between Pairs of Macroecological Variables Based on Cyprinella Species

<table>
<thead>
<tr>
<th>Body Size</th>
<th>Geographical Range</th>
<th>Latitude</th>
<th>Longitude</th>
<th>Phylogeny</th>
</tr>
</thead>
<tbody>
<tr>
<td>Body size</td>
<td>.34 (2.59)*</td>
<td>.30 (2.57)*</td>
<td>.12 (1.61)</td>
<td>.35 (4.87)**</td>
</tr>
<tr>
<td>Geographical range</td>
<td>.51*</td>
<td>.24 (1.99)</td>
<td>-.04 ( -.48)</td>
<td>.08 (1.08)</td>
</tr>
<tr>
<td>Latitude</td>
<td>.64**</td>
<td>.52*</td>
<td>.48 (6.73)**</td>
<td>.32 (4.62)**</td>
</tr>
<tr>
<td>Longitude</td>
<td>-.40</td>
<td>-.37</td>
<td>-.74**</td>
<td>.27 (4.60)**</td>
</tr>
<tr>
<td>Phylogeny</td>
<td>-.18</td>
<td>-.35</td>
<td>-.12</td>
<td>-.39</td>
</tr>
</tbody>
</table>

Note.—N = 27. Product-moment correlations are shown above the diagonal; matrix correlations, below the diagonal. Values in parentheses are Mantel t values. For product-moment correlations, the primitiveness index is the measure for phylogeny. For matrix correlations, the number of nodes in the cladogram separating each species pair is the measure for phylogeny. Asterisks indicate probabilities after sequential Bonferonni adjustments for the tablewide error rate (Rice 1989). See text for details.

* P ≤ .05.

** P ≤ .01.

when using phylogenetic trees to study speciation modes, we must assume that extant ‘‘sister species’’ are each other’s closest relatives (i.e., no extinctions have occurred in the clade; Brooks and McLennan 1991). Hey (1992) found, for small monophyletic groups, that excluding extinction from a null model of speciation yielded predictions closer to observations than did a model including extinction. Thus, in the absence of direct evidence for extinction, it appears useful to think of diversification of recently formed monophyletic groups as a random process without extinction (Hey 1992).

In spite of the difficulties associated with Lynch’s (1989) method (a discussion of which is beyond the scope of our article), it appears to be the only operational protocol for evaluating different speciation mechanisms. To the extent that speciation processes influence geographical range size, we were interested in estimating the frequencies of sympatric, vicariant, and peripheral-isolate speciation events, particularly in the lutrensis and whipplei clades.

RESULTS

For the entire Cyprinella clade, macroecological variables are highly intercorrelated (table 1). Body size is positively associated with geographical range size and with phylogeny: closely related species pairs are more similar in body size than distantly related species pairs. Both body size and geographical range size increase with latitude (Bergmann’s rule and a variant of Rapoport’s rule). Latitude and longitude of species’ geographical ranges also show phylogenetic effects (table 1), so that the more closely related the two species, the less the distance separating their geographical range centers (fig. 2). Phylogenetic effects are not restricted to pairwise comparisons. The deepest branch of the Cyprinella clade (fig. 1) clearly splits the geographical range centers into two groups: a western lutrensis clade and an eastern whipplei clade. The boundary between these two groups (fig. 3) corresponds roughly to the modern Mississippi River basin and
closely to a hypothesized preglacial plains stream system (Metcalf 1966; Cross et al. 1986; Mayden 1988; fig. 4). Path analysis greatly simplifies interpretation of the correlations in table 1. There are positive effects of latitude on both geographical range and body size but no other significant associations. In particular, there is not a significant association between geographical range and body size once latitude is included in the model (fig. 4). For the western lutrensis clade, product-moment correlations and the path analyses gave the same result. Body size increases at higher latitudes, but body size and geographical range are uncorrelated (table 2; fig. 4). In contrast, there are complex relationships among macroecological variables in the eastern whipplei clade (table 3), and these are retained in the path analysis (fig. 4). Longitude is strongly correlated with body size, which increases in a westerly direction. Latitude is not correlated with body size (table 3) but has a significant positive effect after longitude is included in the model (fig. 4). Phylogeny is not associated with body size but is correlated with geographical range. In the eastern whipplei clade, primitive species (i.e., close to the cladogram root) have relatively large geographical ranges (fig. 5). After accounting for these relationships, there is still a significant positive association between body size and geographical range (fig. 4).

Finally, modes of speciation in Cyprinella differ between eastern and western clades. By Lynch’s (1989) criteria, there is evidence for three cases of peripheral-isolate speciation (table 4), all of which occurred in the western lutrensis clade: Cyprinella garmani, Cyprinella xanthiscarca, and Cyprinella bocagrande. In contrast, 10 of 12 speciation events in the eastern whipplei clade fit the vicariant pattern (table 4). The estimated frequencies of vicariant and peripheral-isolate speciation differ significantly between the eastern and western clades (Fisher’s exact test, $P = .022$). These speciation patterns are consistent with the result
Fig. 3.—Location of range centers of 27 species of genus Cyprinella. The dotted line is the demarcation between the western lutrensis clade and the eastern whipplei clade and corresponds to the deepest branching in the cladogram of fig. 1.

A. ALL SPECIES

? ———> BODY SIZE

LATITUDE ——> 0.77

LATITUDE ———> 0.64 **

LATITUDE ———> 0.52 **

LATITUDE ———> 0.85

B. WESTERN LUTRENSIS CLADE

? ———> BODY SIZE

LATITUDE ——> 0.47

LATITUDE ———> 0.88 **

LATITUDE ———> 1.00

C. EASTERN WHIPPLEI CLADE

? ———> BODY SIZE

LATITUDE ——> 0.33 *

LONGITUDE ——> 0.54

LONGITUDE ———> 0.82 **

LONGITUDE ———> 0.67

LONGITUDE ———> 0.52 *

? ———> GEOGRAPHIC RANGE (LOG)

PRIMITIVENESS ——> -0.74 **

GEOGRAPHIC RANGE (LOG) ——> 0.67

GEOGRAPHIC RANGE (LOG) ———> 1

Fig. 4.—Path diagrams for macroecological variables. Only significant partial regression (single-headed arrows) and correlation (double-headed arrows) coefficients are shown. Significance levels are indicated by asterisks (*P ≤ .05, **P ≤ .01).
TABLE 2
CORRELATIONS BETWEEN PAIRS OF MACROECOLOGICAL VARIABLES FOR SPECIES IN THE WESTERN LUTRENSIS CLADE

<table>
<thead>
<tr>
<th></th>
<th>Body Size</th>
<th>Geographical Range</th>
<th>Latitude</th>
<th>Longitude</th>
<th>Phylogeny</th>
</tr>
</thead>
<tbody>
<tr>
<td>Body size</td>
<td>• •</td>
<td>.28 (1.63)</td>
<td>.71 (4.24)**</td>
<td>.22 (1.37)</td>
<td>.14 (.95)</td>
</tr>
<tr>
<td>Geographical range</td>
<td>.29</td>
<td>.48 (2.14)</td>
<td>.38 (1.98)</td>
<td>.10 (.79)</td>
<td></td>
</tr>
<tr>
<td>Latitude</td>
<td>.88**</td>
<td>.18</td>
<td>.25 (1.42)</td>
<td>.17 (1.27)</td>
<td></td>
</tr>
<tr>
<td>Longitude</td>
<td>-.15</td>
<td>-.27</td>
<td>-.33</td>
<td>-.10</td>
<td>-.22</td>
</tr>
<tr>
<td>Phylogeny</td>
<td>-.20</td>
<td>-.45</td>
<td>-.10</td>
<td>-.22</td>
<td>• •</td>
</tr>
</tbody>
</table>

Note.—N = 10. See table 1 for explanation.
** P ≤ .01.

TABLE 3
CORRELATIONS BETWEEN PAIRS OF MACROECOLOGICAL VARIABLES FOR SPECIES IN THE EASTERN WHIPPLEI CLADE

<table>
<thead>
<tr>
<th></th>
<th>Body Size</th>
<th>Geographical Range</th>
<th>Latitude</th>
<th>Longitude</th>
<th>Phylogeny</th>
</tr>
</thead>
<tbody>
<tr>
<td>Body size</td>
<td>• •</td>
<td>.47 (2.92)*</td>
<td>-.05 (-.29)</td>
<td>.59 (3.84)**</td>
<td>-.13 (-1.26)</td>
</tr>
<tr>
<td>Geographical range</td>
<td>.61*</td>
<td>.57 (3.16)**</td>
<td>.29 (1.74)</td>
<td>-.21 (-2.09)</td>
<td></td>
</tr>
<tr>
<td>Latitude</td>
<td>.24</td>
<td>.69*</td>
<td>.00 (.02)</td>
<td>-.05 (-.53)</td>
<td></td>
</tr>
<tr>
<td>Longitude</td>
<td>.78**</td>
<td>.24</td>
<td>-.12</td>
<td>-.06 (-.56)</td>
<td></td>
</tr>
<tr>
<td>Phylogeny</td>
<td>-.58*</td>
<td>-.74**</td>
<td>-.78**</td>
<td>-.37</td>
<td>• •</td>
</tr>
</tbody>
</table>

Note.—N = 17. See table 1 for explanation.
* P ≤ .05.
** P ≤ .01.

Fig. 5.—Geographical range vs. primitiveness. Each point represents a species. Geographical range is the area of the range in square kilometers, and primitiveness is the number of nodes separating each species from the cladogram root. Circles represent the western lutreñís clade (r = -.447, P = .195); squares represent the eastern whipplei clade (r = -.737, P = .001).
TABLE 4
ESTIMATED SPECIATION MODES FOR EASTERN AND WESTERN CLADES

<table>
<thead>
<tr>
<th>CLADE</th>
<th>Sympatric</th>
<th>Vicariant</th>
<th>Peripheral Isolate</th>
<th>Questionable</th>
</tr>
</thead>
<tbody>
<tr>
<td>Western</td>
<td>0</td>
<td>2</td>
<td>3</td>
<td>1</td>
</tr>
<tr>
<td>Eastern</td>
<td>1</td>
<td>10</td>
<td>0</td>
<td>1</td>
</tr>
</tbody>
</table>

Note.—Each observation represents a branching in the cladogram that involves at least one terminal (extant) taxon. Classification follows Lynch 1989.

that the average geographical separation of sister taxa in the western *lutrensis* clade (mean = 88.07 km) is greater than in the eastern *whipplei* clade (mean = 20.89 km; Mann-Whitney U = 107.0, P = .038) and that the frequency of overlap among sister groups in the eastern *whipplei* clade (seven of 12) is greater than in the western *lutrensis* clade (one of six), although the difference was not statistically significant (Fisher’s exact test, P = .097).

DISCUSSION

The goal of macroecology is to understand correlations among species in important ecological attributes. Brown (1981) and Gaston and Lawton (1988a, 1988b) have presented reasonable models for explaining associations between geographical range and body size. Although we used linear statistics to describe relationships in our data set, a graphical presentation of geographical range and body size provides additional insight into this relationship and supports Brown’s (1981) model. Figure 6 shows the association between geographical range size and body size and suggests that the significant linear relationship we found (table 1) was primarily due to minimum geographical range size increasing with an increase in body size (lower boundary in fig. 6). Brown and Maurer (1987, 1989) noted (for North American land birds and mammals, respectively) that geographical range size boundaries may be diffuse and suggested (Brown and Maurer 1987, p. 15) that “such boundaries result from the probabilistic processes of origination and extinction, such that the number of species declines gradually across the boundary.”

However, such associations must be viewed cautiously because of the potentially confounding influences of other variables. In particular, associations between geographical range and body size are often confounded by latitudinal gradients in both variables. After statistically controlling for latitude, there was not a correlation between body size and geographical range for the *Cyprinella* clade (fig. 4). Pagel et al. (1991) obtained a similar result for mammals of North America.

Additionally, macroecological analyses may benefit from incorporating phylogenetic information for a historical understanding of continental biotas. Phylogeny may have direct effects on macroecological variables, as shown by the correlation between primitiveness and geographical range (fig. 5). Equally important,
phylogeny appears to control the expression of ecological relationships, which may differ substantially between related monophyletic groups (fig. 4).

In our analyses, both the area and location of geographical ranges were associated with phylogeny, and ecological correlations were different between eastern and western clades. The Rio Grande basin and the southeastern United States also represent centers of endemism for other fish assemblages (McAllister et al. 1986; Mayden 1987a, 1988), and we predict our results may be similar for other aquatic taxa. In the next section, we briefly review the climatic and vicariant histories of the eastern and southwestern United States and indicate how these histories may have contributed to correlations between body size and geographical range.

Speciation in Cyprinella

Since the Pleistocene, the Rio Grande basin has undergone a general progression toward increased aridity (Bryant 1974). Thus, the western lutrensis clade probably has undergone range contractions as formerly widespread river drainages contracted (Smith and Miller 1986). Much of the western lutrensis clade represents remnant fish faunas persisting in contracted relict habitats. In the face of contracting habitats and increasing environmental stress, subtle ecological interactions between body size and geographical range are unlikely to be important. Consequently, these variables are not correlated in the western lutrensis clade.

In contrast to the history of increasing aridity and habitat contraction in the West, the eastern United States remained relatively mesic throughout the Quaternary (Swift et al. 1986; Prentice et al. 1991). Speciation has been due primarily to vicariant events, enhanced by fluctuating sea levels on the coastal plains (Gibbs 1957; Hocutt et al. 1986; Swift et al. 1986; Mayden 1987b).

Compared with the western lutrensis clade, species in the eastern whipplei
clade have had a much greater potential for dispersal and persistence in the mesic eastern United States. Under these circumstances, ecological relationships, such as the correlation between body size and geographical range, are more likely to be expressed. Other correlations with phylogeny, latitude, and longitude are also significant, probably because geographical ranges are not as distorted by habitat contractions.

The contrasting histories of habitat contraction in the arid west and dispersal in the mesic east also are seen in our analyses of geographical separation and speciation mode. Geographical separation of sister taxa in the western lutrensis clade is greater than in the eastern whipplei clade. Furthermore, our speciation analysis found only three likely cases of peripheral-isolate speciation, all in the western clade. Of course, it is difficult to confirm the existence of peripheral-isolate speciation without additional evidence, such as the frequent occurrence of autapomorphies in the peripheral isolate (Lynch 1989). R. L. Mayden (unpublished data) provided us with the number of autapomorphic characters for the three pairs of sister species that fit the peripheral-isolate speciation pattern. Goodness-of-fit G-tests on each species pair indicated that the number of autapomorphic characters did not differ significantly from a 1:1 ratio. Consequently, no evidence suggests that the "peripheral" species has differentiated more than the "ancestral" species, as predicted by the peripheral-isolates model. Given this result, plus the amount of habitat fragmentation and contraction that has occurred in the Rio Grande basin, we suspect these three cases have resulted from microvicariance or range contractions; either mechanism would lead to a very small geographical range for one of the two sister taxa. In contrast, the peripheral-isolates pattern of geographical ranges is never seen in the eastern whipplei clade. The eastern whipplei clade contains the only possible case of sympatric speciation; Cyprinella caerulea and its sister element (containing Cyprinella trichroistea and Cyprinella gibbsii) have geographical ranges that overlap 100%. Whatever the true mode of speciation, our results indicate that the maintenance and possibly the origin of geographical ranges is fundamentally different in eastern and western clades and that these mechanisms are important in understanding the correlation between body size and geographical range.

Geographical Ranges, Dispersal, and Macroecology

Although macroecologists have not emphasized phylogeny, they have conceded that both macroscopic and microscopic processes contribute to local species composition (Ricklefs 1987, 1989; Brown and Maurer 1989). Nevertheless, their emphasis has been on short-term ecological processes: "the origination, spread, and persistence of species in time and space depend on the effect of ecological conditions, dynamics of local populations and the direction and rate of microevolutionary change" (Brown and Maurer 1989, p. 1148). In other words, slower historical processes are thought to set the initial species composition, which is "sorted" by faster, ecological mechanisms. Anderson and Evensen (1978) took a similar perspective in the construction of a null model of dynamic geographical range sizes. In this model, they explicitly assumed that "birth and
death rates” of range size increases occurred faster than vicariant (i.e., historical) events.

An alternative perspective is that historical processes, including vicariance and habitat shifts, have lasting influences (Brooks and McLennan 1991) and that these historical effects on geographical range size are not ameliorated by dispersal on ecological timescales. We expect historical effects to be especially important for freshwater fishes, because their dispersal powers are limited and because relatively recent, post-Pleistocene, vicariant events have influenced their geographical ranges.

Is there any way to test for the relative importance of ecological versus historical effects on geographical range size? Brown and Gibson (1983, p. 529) predicted that “recently derived sister species would have ranges that overlapped less than those species that are products of a more ancient splitting of phyletic lineages.” On the other hand, if historical processes continue to predominate in geographical range size, there would not be a simple relationship between geographical overlap and time of splitting.

Following Lynch (1989), we used speciation level as a measure of the relative age of a speciation event. Speciation level is defined as the number of node levels that must be passed on a cladogram to reach a common ancestor. Thus, a pair of terminal sister species are at speciation level 1. This metric assumes that the time of speciation is correlated with the amount of cladistic change measured in the cladogram and that the analysis is not distorted by a lack of knowledge of “ghost species” (Simberloff et al. 1981), extinct lineages that are not included in the cladogram.

For overlapping sister taxa in the Cyprinella clade, there is no relationship between speciation level and geographical overlap (fig. 7). We obtained similar results for the full data set, not restricting the analysis to sister elements with at least one terminal taxon. This lack of distinctive pattern suggests that historical effects on current geographical ranges are important, even for taxa that speciated in the relatively distant past. Lynch (1989) obtained similar results for several bird, fish, and frog cladograms.

On the other hand, the pattern in figure 5 provides some support for Brown and Gibson’s (1983) dispersal hypothesis. For the eastern whipplei clade, species close to the cladogram root (which we presume to be relatively old) have larger geographical ranges than species distant from the root. This pattern does not hold for the western lutrensis clade, again suggesting that historical range contractions have been more important in the arid west than in the mesic east.

CONCLUSIONS

Our analyses of the Cyprinella clade illustrate the interplay of historical and ecological processes in producing relationships between body size and geographical range. Based on our results, we suggest four caveats for further studies in macroecology.

First, patterns of covariation should be explored statistically. All macroecologi-
Fig. 7.—Geographical overlap vs. speciation level \( (r = 0.067, P = .887) \). Each point represents a pair of sister taxa, at least one of which is an extant species. Geographical overlap is the percentage overlap of the smaller of two species ranges onto the larger. Counting from terminal taxa, speciation level is the number of nodes in the cladogram that must be crossed to reach the split between two taxa. For sister taxa that comprise more than one extant species, geographical range is defined as the nonoverlapping sum of component ranges. Results are similar if the sympatric outlier is excluded. Symbols are as in fig. 5.

cal data sets evaluated to date, including ours, are characterized by substantial variability and strong intercorrelation. Under these circumstances, qualitative graphical interpretations (see, e.g., Brown and Maurer 1987) may be confounded by other variables. Although path analyses and multiple regression cannot help distinguish correlation from causation (James and McCulloch 1990), they are an important step in analyzing macroecological data and eliminating redundant or spurious correlations.

Second, phylogenetic information is fundamental to comparative biology (Felsenstein 1985) and should be incorporated into analyses when possible. Unfortunately, cladograms are often not available; macroecologists have attempted to control for phylogenetic effects by restricting their analyses to large monophyletic groups, such as the class Aves. However, even within a single fish genus, we detected important historical effects. Within any monophyletic group, macroecological patterns may vary among different clades, as we found for the eastern *whipplei* and western *lutrensis* clades. The primitiveness index and the number of nodes separating a pair of taxa are two simple phylogenetic metrics that can be extracted from a cladogram and incorporated into macroecological studies.

We note in passing that a more detailed analysis of the relationship among geographical range size, body size, and phylogeny is possible. We could have reconstructed body sizes and geographical ranges for interior nodes of the cladogram and then tested for correlations amongst the extant taxa (A. J. Letcher and P. H. Harvey, unpublished manuscript). However, such an analysis assumes a simple underlying model of character change. At least for geographical range
size, our analysis suggests fundamentally different mechanisms of range size evolution in the eastern and western lineages of *Cyprinella*.

Third, speciation mechanisms and climatic histories of a region should be explored. Attempts to explain the size and location of a species' geographical range should consider the speciation process by which the geographical range originated and the climatic history, which sets the potential for range expansion or contraction. Because historical ecology (sensu Brooks 1985) is largely a narrative process, there is a danger of involving post hoc explanations to account for observed patterns. Nevertheless, a historical perspective is important for interpreting statistical analyses of macroecological data (see, e.g., Brooks and McLennan 1991).

Finally, macroecological analyses may be problematic when carried out at the level of an entire continent. The spatial scale of an entire continent may be too large for effective evaluation of macroecological patterns because a continental land mass is composed of several geographical subunits, each of which has a separate history (Noonan 1988). If a phylogenetic approach is taken, small monophyletic groups are likely to map onto geographical regions with a homogeneous history (Cracraft 1982), especially for organisms with a limited ability to disperse. Analyses of independent cladograms may provide additional confirmation of the integrity of these biogeographical subunits (Rosen 1978).

Although we have advocated a phylogenetic approach to macroecology, our methods suffer from some limitations. First, there are relatively few data sets available that include both a well-resolved cladogram and quantitative information on macroecology. Second, the cladogram itself is a hypothesis of relationship, so the analysis will depend on the quality of the phylogenetic information. Finally, it may be inappropriate to attempt to clearly partition macroecological variation into ecological and phylogenetic components. Our procedures will not reveal ecological effects that are correlated with phylogeny, in the same way that environmental and genetic effects are difficult to partition in studies of IQ. However, by removing phylogenetic effects first, we are taking a conservative approach that makes it more likely that the remaining patterns have ecological significance.

Our analysis provides support for predictions of Brown's (1981) model. Even after statistically controlling for effects of latitude, longitude, and phylogeny, body size and geographical range size were positively correlated for species of the eastern United States. For the western clade, no such correlation emerged, but there are compelling historical reasons for the lack of association of geographical range size and body size. These insights would not be possible without a phylogenetic and historical perspective on macroecology.

ACKNOWLEDGMENTS

We thank J. H. Brown, A. A. Echelle, G. R. Graves, P. H. Harvey, M. V. Lomolino, J. D. Lynch, W. J. Matthews, B. A. Maurer, B. D. Mishler, G. D. Schnell, and C. C. Vaughn for comments on the manuscript. We thank R. L. Mayden for providing us with unpublished character data for *Cyprinella*. C. McCallister helped draft figure 3.
MACROECOLOGY OF CYPRINELLA


MACROECOLOGY OF CYPRINELLA


Associate Editor: Brent D. Mishler