Papers in Natural Resources   Natural Resources, School of

7-1-1989

Testing Macroevolutionary Hypotheses with Cladistic Analysis: Evidence Against Rectangular Evolution

Cliff A. Lemen
University of Nebraska-Lincoln, clemen2@unl.edu

Patricia W. Freeman
University of Nebraska-Lincoln, pfreeman1@unl.edu

TESTING MACROEVOLUTIONARY HYPOTHESES WITH CLADISTIC ANALYSIS: EVIDENCE AGAINST RECTANGULAR EVOLUTION

CLIFF A. LEMEN AND PATRICIA W. FREEMAN
University of Nebraska State Museum, Lincoln, NE 68588-0514

Abstract.—The properties of cladistic data sets from small monophyletic groups (6–12 species) are investigated using computer simulations of macroevolution. Two evolutionary models are simulated: gradualism and the punctuated-equilibrium hypothesis. Under the conditions of our simulations these two models of evolution make consistently different predictions about the distribution of autapomorphies among species. When strict stasis is enforced, the punctuated-equilibrium hypothesis predicts that the most expected number of autapomorphies per species will be zero, no matter how many characters are used in the analysis. As the number of characters used in the analysis increases, the distribution of the number of autapomorphies per species becomes bimodal. Under gradualism, the distribution of autapomorphies remains unimodal under all conditions, but the number of species without autapomorphies can fall to zero. A survey of real cladograms of extant monophyletic groups from a wide range of taxa indicates that the predictions of the punctuated-equilibrium hypothesis about autapomorphies do not hold. This constitutes strong evidence against the punctuated-equilibrium hypothesis.

Macroevolution is a long-term, large-scale process that is emerging as a prominent gap in the understanding of evolution. Macroevolution must consist of speciation, extinction, and genetic change through time. What is uncertain is exactly how these elements interact and combine to produce the evolutionary history of life. Direct observation of macroevolution is difficult, because the relevant evolutionary events may take millennia or longer to unfold. To date, most of the work on macroevolution has used the fossil record (Simpson, 1944; Lazarus, 1983; Malmgren et al., 1983; Gingerich, 1976; Gould et al., 1977). However, others have addressed this question using the morphology of living species (Avise, 1977; Ricklefs, 1980; Lemen and Freeman, 1981, 1984).

We maintain that neontological data sets are at least the equal of the fossil record for testing macroevolutionary hypotheses. We are compelled to this conclusion because of the difference in completeness of these two sources of data. Paleontological data sets suffer from uncertain chronologies, missing species, and geographic variation generated by a temporally and spatially incomplete fossil record and from incomplete data on those species that are found. There must be literally orders of magnitude of differences in the amount of data that could be collected from extant species (hard and soft anatomy,

physiology, and [ultimately] DNA sequences) and the data that are available from even the best-represented fossil species. However, it is not obvious how to use extant species to test macroevolutionary hypotheses, because neontological data sets, for all their virtues, lack the element of time.

Fortunately, the time component is not needed. Neontological data can be used to investigate macroevolution by scrutinizing how species are arranged in phenotypic space. The evolutionary history of a clade can be imagined as a bush that grows taller with time, spreads with phenotypic (and genotypic) change, branches with cladogenesis, and terminates its branches with extinctions. Any instant in time, such as the present, is a cross-section of this bush. How species are arrayed in this cross-section (phenotypic space) is based on micro- and macroevolutionary processes. The problem is to detect what information this cross-section of the evolutionary bush divulges about the underlying macroevolutionary process. The specific phenotypes of the species are unpredictable, because they are the products of the unique history of each species. However, if all clades share a fundamental mode of macroevolution, some repeated patterns dependent on this shared process can be expected. With this premise in mind, we ask the question: do different models of macroevolution make consistently different

predictions about how species will be jux-
taposed in phenotypic space? If so, we can
use these differences in predictions to test
these macroevolutionary hypotheses.

Avise and Ayala (1975) predicted the sta-
tistical properties of genetic distance under
both rectangular (the punctuated-equilib-
rium hypothesis of Eldredge and Gould
[1972]) and gradual modes of evolution.
Avise (1977) concluded that, when evolu-
tionary change is confined to speciation
events, clades with high rates of speciation
should also have high genetic diversity. If
change is gradual, genetic diversity will be
a function of time. Avise (1977) and Doug-
las and Avise (1982) used groups with pre-
sumably similar times of origin but very
different numbers of species (*Notropis* with
about 100 species vs. *Lepomis* with 11
species) to test this prediction. They did not
find greater evolutionary divergence in *No-
tropis* and took this result as counter to the
prediction of the punctuated-equilibrium
hypothesis. This approach to testing modes
of evolution, while interesting, is not with-
out flaws. It requires that *Notropis* and *Le-
pomis* are of about the same age and that
there are not many uncounted speciation
events in *Lepomis* caused by extinctions.
Further, it must be assumed that rectangular
evolution would proceed with similar starts
and stops in these two groups, even given
possible differences such as population
structure, vagility, or generation time.

We previously used a method modified
from Ricklefs (1980) to test gradual and
rectangular evolution (Lemen and Free-
man, 1981). We compared nearest-neigh-
bor distances at four taxonomic levels:
within populations, species, genera, and
families. If two evolutionary mechanisms
are at work (micro- and macroevolutionary
processes) and one is responsible for shifts
in morphology within species and the other
for large shifts among species, then there
might be a significant change between with-
in-species distances and between-species
distances. On the other hand, if changes are
gradual, then these taxonomic levels are
largely arbitrary, and there should be a
smooth transition to successively higher
taxonomic levels. Data from the microchi-
ropteran families Emballonuridae and Mo-
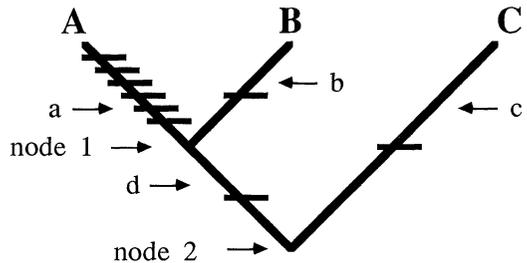lossidae showed a gradual shift in nearest-

FIG. 1. Cladogram of three species to illustrate the
value of cladistic data sets in macroevolutionary stud-
ies and the methods used to quantify cladograms. Hor-
izontal bars are apomorphies; see text for further ex-
planation.

neighbor distances with higher taxonomic
levels, consistent with the gradualistic mode.
Although interesting, the results of Lemen
and Freeman (1981) can be criticized be-
cause of the morphological criterion for as-
sessing nearest neighbors. If there were ma-
jor shifts at some speciation events, these
shifts, by their nature, would initiate new
genera or even families, but such a leap be-
tween the mother and daughter species
would be ignored in a morphological study
because it would not be a nearest-neighbor
distance.

In this paper, we again compare gradual
and rectangular evolution, this time using
cladistic data sets. The cladistic approach is
particularly useful, not only because it is a
different type of data from the nonpolar-
ized-data-set approach used by Avise (1977)
and Lemen and Freeman (1981), but be-
cause cladistic data sets are more directly
applicable to problems in character evolu-
tion and, as a result, solve some of the co-
nundrums encountered in the previous work
employing neontological data. Cladograms
are intended to produce the evolutionary
relationships among species and not just
morphological similarity. This is important
because we will be able to assess distances
among nearest phylogenetic neighbors and
not just distance to the most morphologi-
cally similar species. In Figure 1, species A
and B are sister species, but because B and
C are more similar in overall morphology,
they would have been taken as nearest
neighbors in overall morphology. A second
salient point is that cladistic data sets allow
differences in species to be divided into apo-
morphies (evolution or character-state

transformations) and plesiomorphies (no evolution), allowing a more realistic determination of evolution since speciation. For example, in Figure 1, species A and B differ in character states for seven characters. In an unpolarized data set, this could be represented as a distance of seven. Using the cladistic approach, A has six autapomorphies while B has only one, indicating that A has changed more than B from the ancestral condition for this data set.

Generating predictions from macroevolutionary hypotheses using analytical methods is difficult because of the range of possible conditions and the complex interactions of many factors that must be encompassed in evolutionary models. Therefore, we opted to use computer simulations. These simulations are based on the processes that must exist: extinction, speciation, and morphological change. The simulations differ from one another in the specifics of how these factors are combined. When run, the simulations create artificial clades under different conditions. Each species within the clade has its own history, and the histories of all the species combine to produce the entire evolutionary history of the clade. This history is of little direct value, because we do not have the complete evolutionary history of any real clades with which to compare it. We are interested in how species are arranged at an instant in time in the simulations. It is these relationships that are analogous to and therefore can be compared with neontological data sets. By contrasting the artificial clades generated by the computer, we quantify how different assumptions affect the expected relationships among species. This defines the domain of possible worlds given the range of assumptions that are made. The next step is to compare the results of the models with the results from real data sets to determine whether either model of macroevolution produces data sets that are inconsistent with the real world.

The computer simulations are used to investigate areas we initially predicted should have differences in the predictions of gradual and rectangular evolution. The areas chosen are: 1) the correlation between the total number of apomorphies a species has and the number of branch points in the evolutionary history of that species, 2) the correlation between adjacent branch lengths in cladograms, and 3) the distribution of autapomorphies among species. Results from the computer simulations indicate that only the distributions of autapomorphies are of practical value for testing modes of evolution. Armed with the results of our computer simulations, we looked at a series of published data sets on real species.

The results of this study indicate that patterns predicted by the punctuated-equilibrium hypothesis cannot be found. The telltale fingerprint of stasis, which our simulations indicate must be present, cannot be found in the real data sets surveyed. In the Discussion, we consider how less-extreme interpretations of rectangular evolution may still be compatible with the neontological data. Our results also demonstrate the power of using neontological data sets to address problems in macroevolution.

MATERIALS AND METHODS
Data from Literature

The criteria we used to select studies from the literature were that 1) the group was presumed to be monophyletic, 2) nearly all of the species known for the group were included in the analysis, 3) the study included at least six species, and 4) both synapomorphies and autapomorphies were reported. The studies chosen were *Heterandria* (Osteichthyes; Rosen, 1979), *Rana pipiens* complex (Amphibia; Hillis et al., 1983), *Rana* (Amphibia; Hillis and Davis, 1986), *Rhinoclemmys* (Reptilia; Sites et al., 1981), *Pholidobolus* (Reptilia; Hillis, 1985), *Pappogeomys* (Mammalia; Honeycutt and Williams, 1982), *Atya* (Decapoda; Hobbs and Hart, 1983), *Oreodytes* (*scritulus* [1] and *anguistior* [2] groups) (Coleoptera; Zimmerman, 1985), *Thopeutica* (Coleoptera; Kibby, 1985), and *Styraconyx* (Tardigrada; Kristensen and Higgins, 1985).

From each study, we used information on three relationships. First, we calculated the correlation of the total number of apomorphies a species has and the number of branch points (nodes) between that species and the bottom of the cladogram. In Figure 1, there are two branch points for species A and B

(nodes 1 and 2), while species C has only one node (node 2). In both the real and simulated data sets, there were unresolved trichotomies and higher-level conjunctions. The number of nodes with such unresolved branch points in a species history was calculated as the average of all possible arrangements of unresolved nodes. Second, we determined the correlation of adjacent branch lengths (referred to as branch symmetry) in the cladogram. Branch length was calculated as the total number of character-state transformations that were hypothesized to have occurred between two nodes (synapomorphies) or between a node and a species (autapomorphies). In Figure 1, there are only two pairs of adjacent branches, the pair a and b and the pair c and d. Finally, the distribution of branch lengths for the cladogram was found. We divided the branch lengths into two categories, those that terminate at species, also known as autapomorphies (a, b, and c in Fig. 1), and those that were internal in the cladogram, normally referred to as synapomorphies (d in Fig. 1).

Data from Computer Simulations

The computer simulations used to create the artificial data sets produced characters of known polarity with a 20% reversal rate (one in every five evolutionary events is a reversal). Under the gradual model, change in morphology was a stochastic event that can occur at any time in the history of a species. Under the punctuated-equilibrium model, species changed only at their initial creation. Thus, when a species split into two, one of the new species was left in stasis (mother species), and the other underwent rapid evolutionary change (daughter species).

The data created from these models were analyzed using the PAUP program (PAUP [Phylogenetic Analysis Using Parsimony] is a computer program for phylogenetic analysis written and distributed by David L. Swofford, Illinois Natural History Survey, 607 East Peabody Drive, Champaign, IL 61820). The most-parsimonious trees created were used to find the correlation of total apomorphies and branch nodes, branch symmetry, and the distribution of autapomorphies. If more than one most-parsi-

monious tree existed, the first tree produced was used.

The simulations were designed to contrast the predictions of gradual and rectangular evolution. Because knowledge of macroevolution is limited, simulations were run under a wide variety of conditions to determine whether differences in predictions were robust. This analysis involved the manipulation of four basic factors: clade diversity through time, final number of species in a clade, species definition, and number of characters.

Clade Diversity through Time. — The balance between speciation and extinction determined how the number of species in a clade would wax and wane through time. Simulations were run under three different patterns.

i) Boom clades. — Extinction rates were zero for all species, and the clade monotonically increased in number of species as a function of a stochastically constant chance of speciation for each species in the clade. When the desired number of species was reached (either six or 12), the simulation was stopped (Fig. 2).

ii) Boom/bust clades. — Extinction rates were zero until the clade reached a maximum size (36 species), and then the speciation rate was set to zero and extinction made possible. The clade then decreased in size to a set level (six or 12 species) and the simulation was stopped (Fig. 2).

iii) Equilibrium clades. — The rates of speciation and extinction were functions of number of species in the clade, such that diversity came to a dynamic equilibrium at a set number of species (six or 12). These simulations were run for the largest number of time periods of any of the simulations. An average of 67% of the species created during the run went extinct by the end of the simulation. This indicated that there was considerable species turnover at equilibrium (Fig. 2).

Final Number of Species in the Clade. — Simulations were run with final species counts of six and 12 to determine the effect of clade size. This range bracketed the number of species found in most of the real data sets used.

Species Definition. — Simulations were run under two assumptions about the recogni-
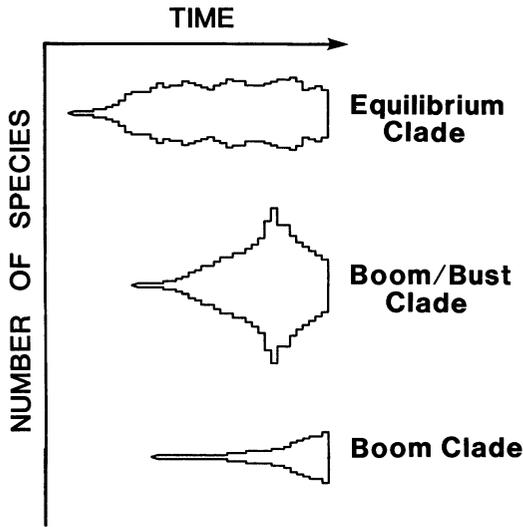
TIME

FIG. 2. These three clade spindles show the number of species extant through time under three different assumptions used in the simulations. All clades begin with one species. The width of the spindle indicates the number of extant species. All clades terminate at the present with 12 species. These spindles are shown to suggest the general nature of the three assumptions; the actual shape of the spindle used in our analysis depends on a pseudorandom number generator. Several thousand different spindles were used in our simulations.

tion of species. To understand how these different methods were applied, it must be remembered that new species were formed when a random-number generator met a criterion. At that point the computer named the new species and gave it a phenotype based on the phenotype of its immediate ancestor and the evolutionary assumptions in place. Because the creation of a new species by the computer was discrete and unambiguous, the number of species could always be known, even if there were no phenotypic differentiation between sibling species. There are cases in which systematists have been able to distinguish species that are phenotypically very similar. To model this situation, we used the assumption that species were recognized as soon as they were created by the computer. However, if a new species has not yet diverged in phenotype from its ancestor, biologists might not recognize it as a new species and would lump it with its sibling into one species. This scenario would be more likely

to occur under the gradualism model, in which gradual divergence is expected. Therefore, we used two methods of defining species. First, we allowed the computer to recognize a new species at the instant of creation, even if that new species had no morphological differences with its sibling species. In the second method, a species was recognized only when it differed from its sibling species by at least 2% of the characters. This method of demanding morphological divergence for species recognition will be referred to as the "distinct-species definition" in the rest of this paper.

Number of Characters. —All characters used in the simulations were binary, and cladograms were created using 20, 40, and 80 characters that were randomly selected (without replacement) from a large data set of, on average, 200 characters that were generated for each species in each artificial clade.

To test fully the interactions of these four factors on the predictions of the gradual and rectangular models, we used a complete factorial design. This produces two modes of evolution, three kinds of clade diversification, two final numbers of species, two species definitions, and three numbers of characters used in the analysis, or 72 different configurations of assumptions. Each of these configurations was run 100 times to obtain an adequate estimate of the results expected under each set of conditions.

RESULTS

*Correlation of Total Apomorphies and
Total Branch Points*

Computer Simulations. —Rectangular evolution normally generated a positive correlation between apomorphies and branch points (solid lines in Fig. 3A). There was considerable variation in these correlations, even when all evolutionary assumptions were held constant, and manipulating the evolutionary variables also affected the distribution of correlations. Analyses based on 12 species produced distributions of correlations with lower variances than analyses based on six species (reduction in the standard deviation averaged 29%; solid lines in Fig. 3B,C). An increase from 20 to 80 in the number of characters used in the analysis resulted in only a 2% decrease in the average standard deviation
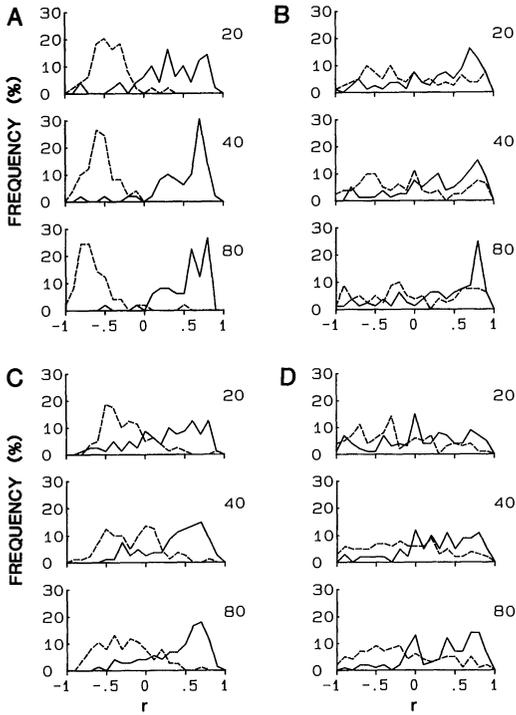
FIG. 3. The distributions of the correlations (r) between total apomorphies and branch nodes (solid lines) and the distributions of the correlations between adjacent branch lengths (broken lines) are shown. These distributions are generated from 100 computer simulations each. The assumptions in force are: A) rectangular evolution, 12 species, distinct species definition, and boom clades; B) rectangular evolution, six species, distinct species definition not used, and equilibrium clades; C) rectangular evolution, 12 species, distinct species definition not used, and equilibrium clades; D) rectangular evolution, six species, distinct species definition, and boom/bust clades. Distributions are shown for 20, 40, and 80 characters used for each set of assumptions.
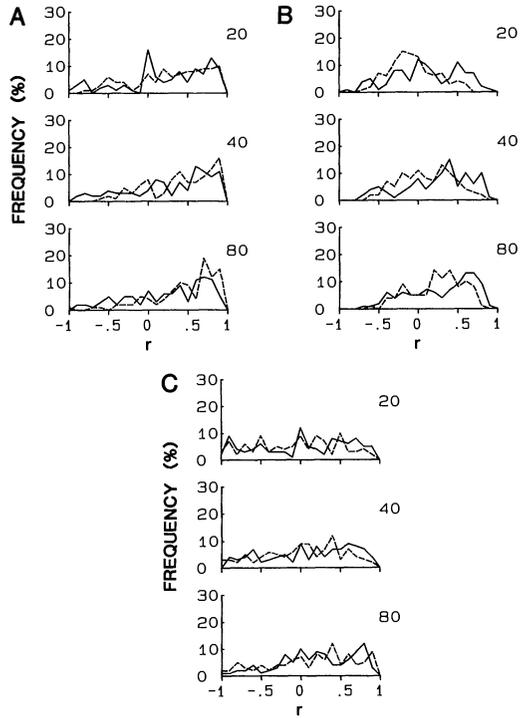
FIG. 4. The distributions of the correlations (r) between total apomorphies and branch nodes (solid lines) and the distributions of the correlations between adjacent branch lengths (broken lines) are shown. These distributions are generated from 100 computer simulations each. The major assumptions in force are: A) gradual evolution, six species, distinct species definition, and equilibrium clades; B) gradual evolution, 12 species, distinct species definition, and equilibrium clades; C) gradual evolution, six species, distinct species definition, and boom/bust clades. Distributions are shown for 20, 40, and 80 characters used for each set of assumptions.

(Fig. 3). The method used to define species also had little impact on this correlation. Simulations with high rates of extinction (equilibrium and boom/bust assumptions) produced higher standard deviations in the correlations than did the boom model with no extinctions (solid lines in Fig. 3C, D vs. those in Fig. 3A). Therefore, the highest variance in correlations was found in the simulation with the fewest species, the fewest characters, and the most extinctions (solid line in Fig. 3D for 20 characters). The result was an almost uniform distribution of correlation coefficients from −1.0 to 1.0. At the other extreme was the simulation

based on many species, many characters and the boom model (solid line in Fig. 3A for 80 characters). In this case, the correlations were highly positively biased, and only 4% of the correlations were negative.

Under gradual evolution, the distribution of correlations was also biased toward positive values, but the pattern was not as pronounced as under rectangular evolution. Lower numbers of species and characters in the analysis tended to increase the variances in distributions. A decrease from 12 to six species resulted in a 27% average increase in standard deviation (Fig. 4A, B). A decrease in number of characters from 80 to 20 resulted in an average increase in the

TABLE 1. The correlation between total apomorphies and branch nodes and the correlation between adjacent branch lengths are shown for the real data sets. N = the number of species in each analysis.

Type of data	Species	N	Between total apomorphies and branch nodes	Adjacent branch lengths
Electrophoretic	Rana (1)	20	0.45	0.36
	Rana (2)[a]	31	0.35	0.29
	Pholidobolus	6	0.74	−0.49
	Rhinoclemmys	6	−0.34	−0.51
	Pappogeomys	7	0.64	0.44
Morphological	Styraconyx	8	0.75	0.17
	Atya	11	0.89	0.40
	Oreodytes (1)	6	0.90	0.58
	Oreodytes (2)	9	0.97	0.41
	Thopeutica	6	0.97	0.00
	Heterandria	9	0.86	0.75

[a] Study used restrictions-fragment analysis of rDNA.

standard deviation of 13% (Fig. 4). The definition of species and assumptions about the rates of speciation and extinction had relatively little impact on the distribution.

Real Data Sets. — The correlations of total apomorphies and branch points for the real data sets are shown in Table 1. The correlations from electrophoretic studies were significantly lower than results from morphological studies ($t = 27$, $d.f. = 9$, $P < 0.02$).

Branch Symmetry

Computer Simulations. — Under rectangular evolution, these correlations tended to be negative, at least under some assumptions (broken lines in Fig. 3A, C). Simulations with fewer species and, to a lesser extent, fewer characters tended to have increased variances. A decrease from 12 to six species caused an average increase in standard deviation of 82% (broken lines in Fig. 3B, C). A decrease from 80 to 20 characters caused an average increase of 6% in the standard deviation (broken lines in Fig. 3). Species definition had a negligible effect, but assumptions about speciation and extinction patterns had an effect. The boom model had a standard deviation 37% lower than the average for the equilibrium and boom/bust models (broken lines in Fig. 3A vs. those in Fig. 3C, D). The strongest negative correlations of branch symmetry associated with the lowest variances were found in the model with 12 species and a boom clade (broken line in Fig. 3A). At the

other extreme, there was an almost uniform distribution of correlations from −1.0 to 1.0 when only 20 characters were used for only six species (broken lines in Fig. 3B, D).

Under gradual evolution, the correlation of adjacent branches tended to be positive (broken lines in Fig. 4), at least when many characters (40 or 80) were used in the analysis. As with the results from rectangular evolution, the variance in r was often high, even approaching a uniform distribution (Fig. 4C). Fewer species and, to a lesser extent, fewer characters tended to increase variances. A decrease from 12 to six species resulted in a 36% increase in the average standard deviation (Fig. 4A, B). A decrease from 80 to 20 in the number of characters resulted in an 11% increase in the standard deviation (broken lines in Fig. 4). Species definition and the pattern of speciation and extinction did not have strong effects on the distribution of correlations.

Real Data Sets. — The correlations of adjacent branch lengths are shown for both morphological and electrophoretic data sets in Table 1. Correlations in electrophoretic studies were not significantly smaller than those from morphological studies ($t = 1.94$, $d.f. = 9$, $P < 0.1$).

The Distribution of Apomorphies

Computer Simulations. — Under rectangular evolution, branches that terminated at species (autapomorphies) and those within the cladogram (synapomorphies) had similar distributions under all conditions
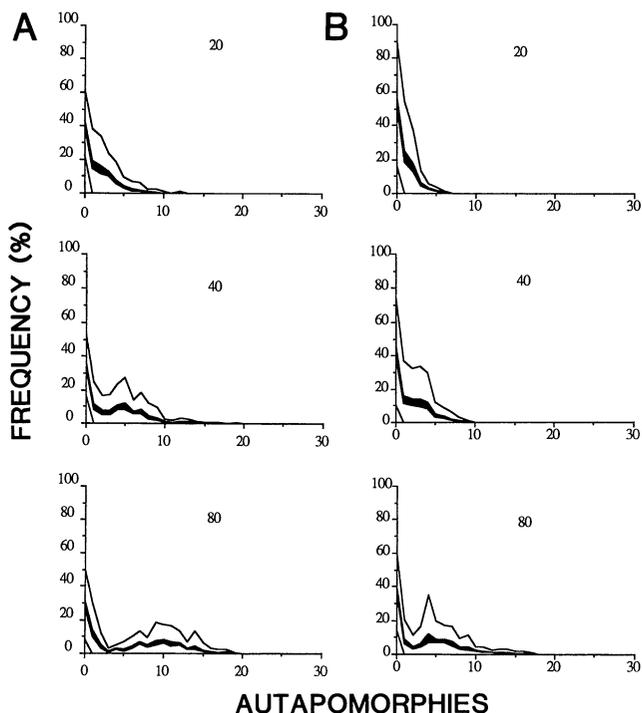
FIG. 5. The distributions of autapomorphies for selected simulations are shown. Each distribution is based on 100 simulations. The light lines delimit two standard deviations, and the blackened areas include two standard errors. Major assumptions for this series of simulations are: A) rectangular evolution, six species, distinct species definition not used, and equilibrium clades; B) rectangular evolution, 12 species, distinct species definition not used, and equilibrium clades. Distributions are shown for 20, 40, and 80 characters used for each set of assumptions.

simulated. When the ratio of characters to species in the analysis increased, the distribution of autapomorphies tended to bimodality (contrast the distributions of autapomorphies for 20 and 80 characters in Fig. 5). In all simulations of rectangular evolution, the single most-expected number of autapomorphies for a species was zero. When we refer to a species as having no autapomorphies, we mean that it has no autapomorphies given the set of characters used in an analysis; this does not necessarily mean that it has no autapomorphies across all characters. The method used to define species and the pattern of speciation and extinction had little effect on the distribution of autapomorphies.

Under gradual evolution, average branch length caused by autapomorphies and synapomorphies increased as the ratio of characters to species in the analysis increased, but the distributions appeared to remain

unimodal (Fig. 6). The method of defining species had a strong impact on the distribution of autapomorphies. When species were defined from the instant they were created by the computer, the most common number of autapomorphies for a species was zero, regardless of how many characters were added to the analysis (Fig. 6B). If distinctness was used to distinguish species, the condition of having no autapomorphies was not the most common condition under all situations (Fig. 6A, C).

The distribution of autapomorphies and synapomorphies were similar under boom and equilibrium models; however, with the boom/bust assumption, there were pronounced differences in the distribution of number of autapomorphies and synapomorphies. This difference in distribution might be interesting for some evolutionary questions, but in our case, it only increased the variance of the distribution of branch
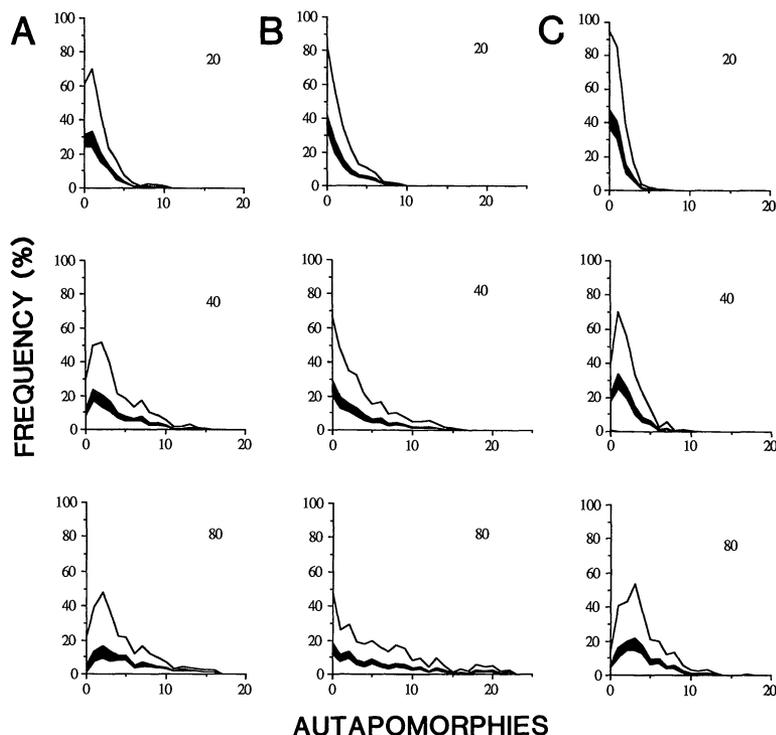
FIG. 6. The distribution of autapomorphies for selected simulations are shown. Each distribution is based on 100 simulations. The light lines delimit two standard deviations, and the blackened areas include two standard errors. Major assumptions for this series of simulations are: A) gradual evolution, six species, distinct species definition, and equilibrium clades; B) gradual evolution, six species, distinct species definition not used, and equilibrium clades; C) gradual evolution, 12 species, distinct species definition, and boom clades. Distributions are shown for 20, 40, and 80 characters used for each set of assumptions.

lengths. For this reason, we concentrated on only autapomorphies in this study.

Real Data Sets.—The distributions of branch lengths from real data sets are shown in Figure 7. Note that, in both the electrophoretic and morphological data sets, the condition of having no autapomorphies was never the most common category.

A test to distinguish rectangular and gradual evolution was devised using the distribution of autapomorphies. Each cladistic analysis from the simulated data sets was scored as to whether the lack of autapomorphies was the most common condition for species. Under gradualism, the expected frequency of species lacking autapomorphies was not always the highest expected frequency when the species distinctness rule was used (Fig. 6A). Therefore, low frequencies of species with no autapomorphies can be consistent with gradualism, depending

on the method used to define new species. Alternatively, under rectangular evolution, species without autapomorphies were always the most expected condition, even when the species distinctness rule was enforced. This produced a clear distinction in the predictions of the two models. However, individual runs were highly variable when few species and few characters were used (note standard-deviation ranges in Fig. 6). For example, under the boom/bust assumption with six species, the percentage of the time when species without autapomorphies were the most common condition was only 30%. Even in this simulation, when results were averaged over 100 runs, species without autapomorphies were still the most common condition (note range of standard errors in Fig. 6). The frequency of species without autapomorphies rose quickly as the number of species in the analysis increased.

With 12 species (under the boom/bust assumption), the percentage of the simulations in which the lack of autapomorphies was the most common condition was 81%. Of all our simulations, the boom/bust assumption produced the lowest expected frequencies of zero.

In the real data sets, species without autapomorphies never proved to be the most common condition (Fig. 7). To calculate the probability of this under the rectangular model, we used the probabilities generated from the boom/bust simulations with linear extrapolations for studies of 6–12 species. For studies with more than 12 species, we used the probability associated with 12 species. Probabilities were also adjusted for the number of characters used in the studies. Considering the electrophoretic and morphological data separately, we found a probability of less than 0.02 that these data sets were consistent with the results predicted by our model of rectangular evolution. When all studies were lumped, the combined probability fell to less than 0.00015. This analysis was the least likely to show a significant difference between predicted and actual distributions, because the boom/bust simulations produced the lowest expected frequencies of species that lack autapomorphies. When results from the equilibrium model were used, the total joint probability that the real data sets were consistent with predictions of rectangular evolution fell to 10^{-9}.

To test for possible deviations from unimodality in the distribution of autapomorphies, we needed to find distribution functions that could reasonably approximate the expected distribution of autapomorphies. The distribution function that would best fit the data depended on the assumptions of the simulation (Ricklefs, 1980). Under gradualism and without the species distinctness rule, the distribution often closely approximated an exponential function (Fig. 8A). In the more biologically realistic case with the species distinctness rule, the distribution could often be closely approximated by a Poisson distribution (Fig. 8B).

Under rectangular evolution, as the number of characters increased the distribution of autapomorphies became bimodal. These data often fit a composite distribution made
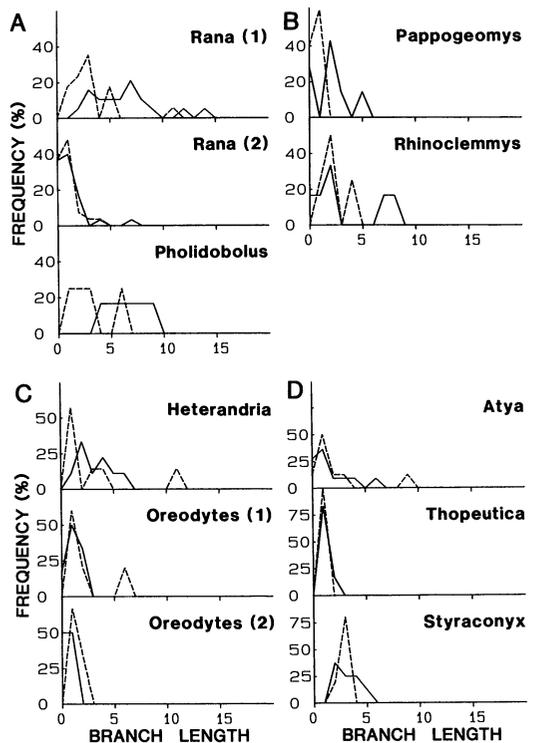

FIG. 7. The distribution of branch lengths for both terminal (solid lines) and internal branches (broken lines) are shown for real data sets. A and B) electrophoretic studies; C and D) morphological studies. See text for references.

up of two Poisson distributions with different means to reflect the difference in change at cladogenesis and anagenesis (Fig. 8C).

Because of the success of the Poisson distribution in fitting the simulated data from gradualistic simulations, we began looking at how well the real data fit Poisson distributions. We quantified deviations from the Poisson distribution using a Kolomogorov-Smirnov test. However, the standard tables for the Kolmogorov-Smirnov statistic, D_{max}, are useless here, because they assume that the true parametric mean of the Poisson distribution is known. We followed the method of Lilliefors (1969) for situations in which the true mean can only be estimated from the data and calculated $D_{critical}$ values with 100 replicate Monte Carlo simulations for each data set. The critical value of D_{max} needed to determine statistical significance had to be calculated by taking into account the fact that 11 D_{max}'s were to be tested (one
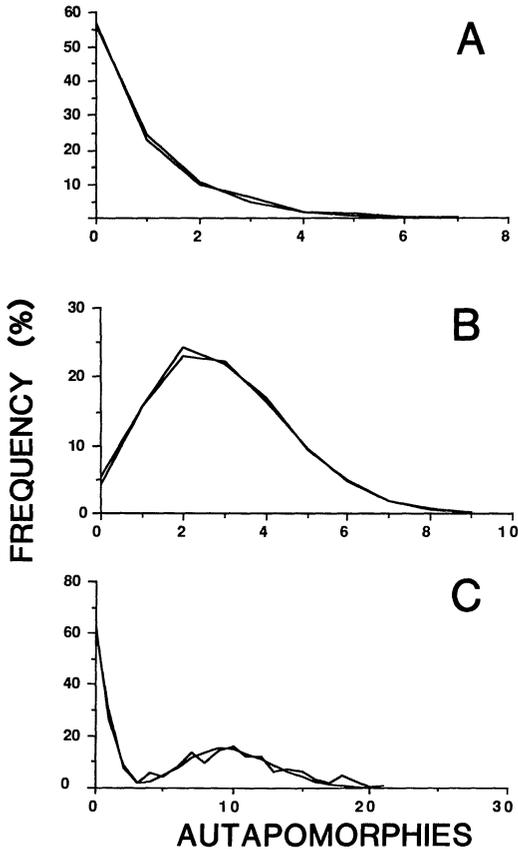
FIG. 8. Graphical representation of the agreement between the distributions of autapomorphies generated from simulations and distribution functions. The results from three different simulation runs are shown with the corresponding distribution functions: A) gradual evolution, 12 species, distinct species definition not used, equilibrium clade, and 80 characters, compared with exponential distribution; B) gradual evolution, six species, distinct species definition used, boom/bust clade, and 80 characters, compared with Poisson distribution; C) rectangular evolution, six species, distinct species definition used, equilibrium clade, and 80 characters, compared with a combination of two Poisson distributions.

for each study). Using the commonly employed $P = 0.05$ level would mean that there would be a 57% chance that one of the D_{max}'s would be found to be significant even if all distributions conformed to the Poisson. The actual level of significance for individual studies was therefore reduced to $P = 0.005$. None of the real data sets was found to deviate significantly from the Poisson, based on this criterion.

The use of such a low P value is statis-

tically correct, but it may obscure a general but individually less dramatic trend throughout the data sets. To test for any such general trend away from unimodality, we computed the expected distribution of the average D_{max} for the five electrophoretic studies. The electrophoretic studies were looked at because we had the most confidence that there was no bias in character selection in these studies (see Discussion). Once again, the Monte Carlo method was used. The mean real D_{max} value fell within one standard deviation of the predicted mean, indicating no significant deviation from the Poisson. The same method was applied to the morphological data sets. In these data sets, there was a significant deviation from the Poisson distribution. The deviation was caused largely by the *Thopeutica* data, for which the deviation from Poisson was not toward bimodality but, rather, toward too low a variance. This is a deviation from the Poisson, but in the opposite direction predicted by rectangular evolution. The low variance may be explained by the researcher's determination to find at least one autapomorphy for each species. Omitting this study from the analysis resulted in an average D_{max} that is within one standard deviation of the predicted D_{max}.

By failing to find evidence of bimodality, we supported the predictions of gradualism. The next statistical problem was to assess the power of this analysis to find bimodality. This question was particularly critical because of the small sample sizes in some of the surveyed data sets. The power of our analysis was estimated by determining the proportion of mean D_{max}'s generated from bimodal distributions that would be larger than 95% of the average D_{max}'s generated from an assumption of a Poisson distribution. Once again, we followed the general method of Lilliefors (1969). The distribution we contrasted with the predictions of the Poisson was based on the sum of two Poisson distributions of different means and equal sample sizes. We used this composite distribution as our alternative because our simulation of rectangular evolution indicated that such a distribution was often produced. We found the expected distribution of average D_{max} (given mean number of au-

tapomorphies and number of species in the real studies) when the mean of the two Poisson distributions differed by factors of 2, 3, 4, 5, and 10. Given the sample sizes of characters, the number of species used in each study, and the total number of studies, we found that the magnitude of the difference between anagenesis and rapid leaps at cladogenesis had to be roughly a factor of five in order to have a 95% chance of rejecting the null hypothesis (unimodal Poisson distribution). The odds of rejecting the null hypothesis fell off quickly at lower differences in mean rates of evolution during leaps and stasis. A ratio of two generated only a 22% chance of rejection.

Reversal Rate

The rate of evolutionary reversal was not varied in our simulations. It was held constant at 20% of the forward rate of evolution. Our impression from looking at a limited number of runs that had no reversals was that they produced similar results to those found with 20% reversals. At higher rates of reversals, it is possible that the patterns we found would not hold true. That question remains to be investigated. We were concerned about whether we could make direct comparisons with our simulated results if reversal rates in the real data sets were too high. We do not have a direct measure of reversal rate in the real data sets. A rough index of the relative rates of reversals in real and simulated data sets can be obtained by comparing consistency indexes of the cladograms. The reported consistency indexes averaged 0.91 (SD = 0.07) for the real data sets and 0.93 (SD = 0.06) in the simulated data sets. While not a direct measure of reversal rate, the high values of consistency indexes and the similarity of values in real and simulated data sets indicate relatively low rates of reversals in the real data sets. With this evidence, we feel that the rate of reversal is not a factor in our analysis of these data sets. However, we recommend that any attempt to use our results for comparison with real data sets should be accompanied by a consideration of reversal rates. Data sets with low consistency indexes or other evidence of high rates of reversal should not be compared directly with our results. High rates of reversal and convergence may obscure macroevolutionary patterns on which our analyses are based.

DISCUSSION

Total Apomorphies and Branch Points

We anticipated that rectangular evolution would produce a strong positive correlation between total apomorphies and total branch points. The computer simulations confirm our intuition. However, there is considerable variation in this relationship over the range of simulated conditions (Fig. 3). The value of this correlation for distinguishing between gradualism and rectangular evolution is further compromised because gradualism also produces highly variable correlations and, in some cases, a tendency toward positive values (Fig. 4).

To our initial surprise, we found that real morphological data sets have much higher correlations of apomorphies and branch points than could be explained by either mode of macroevolution. There are three possible explanations for these high correlations. A biological explanation would be that speciation rate and character evolution are tied closely together. As neither of the modes of evolution simulated in this study includes such an assumption, the real data would cause us correctly to reject both hypotheses as unrealistic. Before such an action can be taken, however, we must consider other, less biologically interesting explanations for these high correlations. First, if a systematic blunder in polarization of characters is made by erroneously defining a species high in the cladogram as the outgroup, a high positive correlation is then expected. Second, bias in character selection can also cause high positive correlations. This would occur if systematists hunt more diligently for synapomorphies to elucidate unresolved portions of cladograms than for redundant synapomorphies and autapomorphies. Such a propensity for character selection is possible because a systematist knows the morphology of the organisms, and that knowledge could consciously or unconsciously affect character selection.

In order to determine which of these three factors is likely to have produced the high correlations between apomorphies and branch points in real data sets, we turned to the electrophoretic data sets. We reason

that electrophoretic data sets are just as likely to be incorrectly polarized as are morphological data sets. Similarly, if there is a strong correlation of evolution and speciation, it might reasonably be expected to affect both morphological and electrophoretic data sets. However, it is improbable that systematists could bias character selection. Character selection is a function of tractability with electrophoretic methodology and contains no preconceived notions of usefulness by systematists. Therefore, the much lower correlations found in electrophoretic data sets suggest that bias in character selection is the likely cause of the high correlations in morphological data sets. In addition, and of greater significance to our central macroevolutionary question, the lower correlations found in the electrophoretic data sets are consistent with the predictions of both modes of evolution modeled here. Therefore, neither model can be rejected on the basis of the correlation of total apomorphies and branch points.

Branch Symmetry

It might be expected that rectangular evolution would produce a negative correlation between adjacent branches in a cladogram if speciation regularly proceeds by the peripheral-isolate model. The negative correlation is produced because one of the species (the one derived from the peripheral isolate) tends to have all the apomorphies associated with the macroevolutionary event at cladogenesis while the other lineage is still maintaining stasis. Our simulations of the punctuated-equilibrium model do tend to produce negative correlations between the lengths of adjacent limbs in a cladogram. The strength of this negative correlation is affected by the number of extinction events in the clade (contrast boom clades with equilibrium and boom/bust clades in Fig. 3 A, C, and D). The prediction of the distribution of correlations under the gradualism model tends to center around zero or even positive values. However, there is a wide variance in the expected distributions, especially with the boom/bust clades (Fig. 4C).

All seven of the correlations between adjacent branches in the morphological data sets are nonnegative (mean $r = 0.36$). The simulations of rectangular evolution show at least a slight bias for negative correlations under all conditions. The probability that all seven correlations under rectangular evolution would be positive due to chance alone is less than 0.01 (0.5^7, using the conservative assumption of equal chance for positive and negative values). This result is counter to the predictions of rectangular evolution. However, as with the correlation of apomorphies and branch points, bias in character selection might artificially inflate these correlations. In the electrophoretic data sets, the correlations are lower, with a mean of nearly zero. Given the range of possible results under the gradual and rectangular models and the relatively small number of electrophoretic studies available, we cannot refute either model with the electrophoretic data alone. We decline to use the morphological data sets as strong evidence to refute rectangular evolution, because of our suspicion of bias in character selection.

Distribution of Autapomorphies

Rectangular evolution will produce many branches of zero length in cladograms, because mother and daughter species only differ by the autapomorphies of the daughter species. This produces a long branch to the daughter species and a zero-length branch to the mother species. Similarly, if a mother species gives rise to several daughter species and then goes extinct, branches of zero length will still be generated, because all these daughter species evolved from the same ancestral morphology.

The commonness of the branches of zero length under gradualism depends on how species are defined. If we were omnipotent (and in the case of the simulated data we are), the instant of speciation is definable, and zero branch lengths will be the most common class for autapomorphies (Fig. 6B). New species are generated and then run the risk of extinction; therefore, the most common condition of a species is to be new (analogous to newborns being the most common age class in a survivorship curve if we assume that births and deaths occur randomly through time). These new species will tend to occur in pairs, with initially similar character states. If they survive long enough, they will diverge. These common

new-species pairs play a large role in the production of the zero-length category of branches in Figure 3B.

The idea that the first instant of a speciation event is real and definable is logically tied to the view that speciation is a discrete event (as in rectangular evolution). In a gradualistic model, the process is gradual, and small differences accumulate. Over time, differences are seen to be typical of populations, then subspecies, and finally species. The exact placement of a speciation event by systematists is somewhat arbitrary, and biologists cannot distinguish current differences in populations that will become species differences in time. As a practical matter, species are recognized after evolution has produced significant morphological or reproductive differences between populations. To simulate this effect in the gradualistic model, we changed the criteria for recognition of a new species to include the idea that species have to be morphologically distinct from sister species. This methodology radically changed the predicted distribution of autapomorphies (Fig. 6A, C). When distinctness of species is demanded, the lack of autapomorphies may not be the most expected condition. This will not happen with rectangular evolution, even if the distinctness criterion is used on these data sets, because mother and daughter species diverge and become distinct instantaneously at cladogenesis.

In the real data sets, the lack of autapomorphies is never the most common condition in either the morphological or electrophoretic studies (Fig. 7). This pattern provides statistically significant evidence against rectangular evolution. The pattern also provides support for the species-distinctness definition.

Possible Problems in Our Analysis

Flaws in Real Cladograms. — Cladograms of extant species offer evidence against the punctuated-equilibrium hypothesis; however, cladograms are often referred to as "least-refuted hypotheses" and one hypothesis cannot be used to test another. Fortunately, a cladogram is more than a simple hypothesis, because there is an underlying data set. Certain assumptions are made, and the patterns in the data set produce a particular set of relationships or cladogram. However, it might be prudent to investigate how systematists might have gone wrong in the production of these cladograms and how their mistakes might affect this test of gradualism and rectangular evolution. First, there is the problem of homoplasies when a systematist incorrectly codes two different conditions as the same state. To simulate this effect, we dictated that 20% of all evolutionary events would be reversals. Even with this high level of reversal our results produce clear distinctions in the predictions of rectangular and gradual evolution. Lower levels of reversals would not affect our results. We have not simulated any higher reversal rates. A second problem, bias in the selection of characters, appears to be more severe and forces us to be at least a little skeptical about the use of morphological data sets. However, bias is unlikely in biochemical data sets. Third, there is the problem of assigning polarity to character-state transformation. As already noted, a consistent error in polarity can profoundly affect the correlation of total apomorphies and branch points. For this and other reasons, that correlation is of little use in separating rectangular and gradual evolution. More relevant in this context is the effect that errors in polarity have on the distribution of autapomorphies. We simulated errors in polarity in two ways. First, we assumed that the polarities of all characters are incorrect. Second, we simulated the condition in which half of the characters are incorrectly polarized. Neither of these errors altered the basic conclusion that a lack of autapomorphies will be the most expected condition for species under rectangular evolution.

Based on our analysis, we conclude that our results are robust over a variety of evolutionary conditions and a wide range of errors that might be made by systematists in constructing cladograms. Put another way, the impact of rectangular evolution on neontological data sets is so dramatic that, if it does exist, it is almost impossible to miss.

Modifications to the Punctuated-Equilibrium Hypothesis

One way to modify the punctuated-equilibrium hypothesis to make it consistent with

our results is to maintain that this hypothesis only applies some of the time or only to some species and that we simply did not find any examples of rectangular evolution. As both gradually changing and static species have been reported in the fossil record, it appears to some (Gould and Eldredge, 1986) that the resolution of this controversy hinges on determining the relative frequency of rectangular evolution. This point of view abandons the original promise of the punctuated-equilibrium hypothesis as strongly nomothetic:

> "We believe that punctuated change dominates the history of life. . . . phyletic gradualism is very rare and too slow." (Gould and Eldredge, 1977 p. 115).

If rectangular evolution occurs only sometimes and in some groups, and if the times and groups cannot be specified but only inferred after the fact, the hypothesis becomes highly idiographic and untestable. Rejection of the punctuated-equilibrium hypothesis does not necessitate a return to gradualism. If elements of the punctuated-equilibrium hypothesis are to be retained, they need to be incorporated into a more specific hypothesis capable of specifying when and where rectangular evolution will apply.

We have a further problem with the approach of determining the relative frequency of rectangular evolution by scoring the presence or absence of apparent stasis in the fossil record (Barnosky, 1987). The problem with this approach is that it assumes that apparent stasis verifies the presence of rectangular evolution. It was the apparent pattern of stasis that caused Eldredge and Gould (1972) to propose the punctuated-equilibrium hypothesis in the first place. Thus, it would be circular to confirm this hypothesis with the presence of a pattern that was used in the construction of the model. Demonstrating that this pattern exists, even sometimes in some organisms, is a vindication of the initial premise that went into the construction of the punctuated-equilibrium hypothesis, not a confirmation of the hypothesis itself. What is needed is an independent test, which is why we have pursued our approach of using neontological data from diverse taxonomic groups. The patterns predicted by the punctuated-equilibrium hypothesis cannot be found in these data sets.

The absence of evidence for stasis in the neontological record is strong evidence against the punctuated-equilibrium hypothesis. Strict stasis can be rejected; but, what if stasis is not absolute but only relative? Gould and Eldredge (1977) noted that the punctuated-equilibrium hypothesis should be allowed to have some anagenesis, as long as larger shifts evolve during speciation. A great difference in the magnitude of change at cladogenesis and anagenesis will leave our conclusions unaltered. A lack of autapomorphies will still be the most expected condition when there is no bias in character selection. The anagenetically produced autapomorphies are uncommon enough that they will only rarely be sampled in cladistic studies with few characters (the normal condition). Higher rates of anagenesis or larger numbers of characters in the analysis can increase the number of anagenetic autapomorphies found until a lack of autapomorphies is not the most expected condition. If a large difference in anagenesis and change at speciation remains (rectangular evolution is not allowed to converge too closely to gradual evolution), then a bimodal distribution of autapomorphies reflecting this difference is expected. There is no evidence present in the surveyed data sets to indicate any deviation from unimodal distributions. An estimate of the statistical power of our test of unimodality indicates that there would have to be at least a five-fold difference in anagenesis and change at cladogenesis to have a 95% chance of rejecting the null hypothesis of unimodality. Larger data sets will be needed to investigate the problem of stasis and rapid change differing by less than a factor of five.

Another point to consider before discarding rectangular evolution is the problem of scale. The morphological groups analyzed here are congeners and are assumed to be closely related. It could be argued that the differences among congeners are relatively small and may have evolved through gradualism, but that this is not the stuff of major evolutionary change. In this scenario, similar in some ways to Simpson's megaevolution (Simpson, 1944), the shifts to form

new genera, families, orders, and other higher taxonomic categories account for most change and are the leaps of interest. Indeed, Lemen and Freeman (1984) suggested that the morphological shifts within and between genera are qualitatively different evolutionary events.

Whatever the truth about these large-scale shifts, we note that as originally formulated the punctuated-equilibrium hypothesis was conceived to function at the level of speciation events and the genetics of small populations. The hypothesis was created to explain the suggested pattern of stasis and sudden appearance of new species without intermediate forms in the fossil record. Eldredge and Gould (1972) offered a provocative explanation of stasis by suggesting that stasis is inevitable for widespread species and the special times of change occur when small populations are in the process of forming new species. As we have demonstrated, there is strong evidence for significant anagenesis in the genera reviewed here. If true, this finding destroys the pattern on which the punctuated-equilibrium hypothesis is founded. There may or may not be a pattern of abrupt large changes (megaevolutionary changes generating the differences between genera or higher taxonomic levels), but without species-level stasis there is no longer a compelling reason to prefer the punctuated-equilibrium hypothesis. This opens the field to competing hypotheses of megaevolution. It is first necessary to quantify the patterns of evolutionary changes above the genus level (Lemen and Freeman, 1984) before proliferating hypotheses of megaevolution. We believe that neontological data can play a critical role in this investigation.

Eldredge and Gould (1972) contended that there was a pattern of morphological stasis in the fossil record and hypothesized a process to explain this pattern. Because of the nature of the punctuated-equilibrium hypothesis, we felt it should apply to electrophoretic data sets as well. Gould and Eldredge (1977) did not agree with this view after electrophoretic data sets used by Avise (1977) and coworkers yielded results contrary to the punctuated-equilibrium hypothesis. They noted that changes in electromorphs and morphology may be ruled by separate evolutionary forces. This is an unresolved problem in evolutionary biology. We note that while we worry about bias in character selection in morphological data sets, the distributions of autapomorphies in morphological data sets confirm the finding in the electrophoretic data sets. They are inconsistent with the predictions of the punctuated-equilibrium hypothesis.

Conclusion

Our results, using cladistic data sets, are consistent with those of Avise (1977) and Lemen and Freeman (1981). Some criticisms can be made of each of these studies individually (and of studies using the fossil record as well). However, the three different approaches used in these studies all fail to find evidence of stasis punctuated with abrupt change as envisioned by Eldredge and Gould (1972). Taken together, these studies provide strong evidence against the punctuated-equilibrium hypothesis. The legacy of Eldredge and Gould's (1972) hypothesis may not be its validity, but its role in rekindling curiosity in large-scale, long-term evolutionary problems.

LITERATURE CITED

AVISE, J. C. 1977. Is evolution gradual or rectangular? Evidence from living fishes. Proc. Nat. Acad. Sci. USA 74:5083–5087.

AVISE, J. C., AND F. J. AYALA. 1975. Genetic change and rates of cladogenesis. Genetics 81:757–773.

BARNOSKY, A. D. 1987. Punctuated equilibrium and phyletic gradualism: Some facts from the Quaternary mammalian record. pp. 109–147. In H. H. Genoways (ed.), Current Mammalogy, Vol. 1. Plenum, N.Y.

DOUGLAS, M. E., AND J. C. AVISE. 1982. Speciation rates and morphological divergence in fishes: Tests of gradual versus rectangular modes of evolutionary change. Evolution 36:224–232.

ELDREDGE, N., AND S. J. GOULD. 1972. Punctuated equilibria: An alternative to phyletic gradualism, pp. 82–115. In T. J. M. Schopf (ed.), Models in Paleobiology. Freeman, San Francisco, CA.

GINGERICH, P. D. 1976. Paleontology and phylogeny: Patterns of evolution at the species level in early Tertiary mammals. Amer. J. Sci. 276:1–28.

GOULD, S. J., AND N. ELDREDGE. 1977. Punctuated equilibria: The tempo and mode of evolution reconsidered. Paleobiology 3:115–151.

———. 1986. Punctuated equilibrium at the third stage. Syst. Zool. 35:143–148.

GOULD, S. J., D. M. RAUP, J. SEPKOSKI, T. J. M. SCHOPF, AND D. S. SIMBERLOFF. 1977. The shape of evolution: A comparison of real and random clades. Paleobiology 3:23–40.

HILLIS, D. M. 1985. Evolutionary genetics of the An-
dean lizard genus *Pholidobolus* (Sauria: Gymnoph-
thalmidae): Phylogeny, biogeography, and a com-
parison of tree construction techniques. Syst. Zool.
34:109–123.

HILLIS, D. M., AND S. K. DAVIS. 1986. Evolution of
ribosomal DNA: Fifty million years of recorded
history in the frog genus *Rana*. Evolution 40:1275–
1288.

HILLIS, D. M., J. S. FROST, AND D. A. WRIGHT. 1983.
Phylogeny and biogeography of the *Rana pipiens*
complex: A biochemical evaluation. Syst. Zool. 32:
132–143.

HOBBS, H. H., AND C. W. HART. 1983. The shrimp
genus *Atya* (Decapoda:Atyidae). Smithson. Contr.
Zool. 364:1–143.

HONEYCUTT, R. L., AND S. L. WILLIAMS. 1982. Genic
differentiation in pocket gophers of the genus *Pap-
pogeomys*, with comments on intergeneric relation-
ships in the subfamily Geomyinae. J. Mammal. 63:
208–217.

KIBBY, G. 1985. A review of *Thopeutica*, a subgenus
of *Cicindela*, (Coleoptera: Cicindelidae), with a key
to species and description of two new taxa. J. Nat.
Hist. 19:21–36.

KRISTENSEN, R. M., AND R. P. HIGGINS. 1985. Re-
vision of *Styraconyx* (Tardigrada: Halechinisci-
dae), with description of two new species from Dis-
ko Bay, West Greenland. Smithson. Contr. Zool.
391:1–40.

LAZARUS, D. 1983. Speciation in pelagic Protista and
its study in the planktonic microfossil record: A
review. Paleobiology 9:327–340.

LEMEN, C. A., AND P. W. FREEMAN. 1981. A test of
macroevolutionary problems with neontological
data. Paleobiology 7:311–315.

———. 1984. The genus: A macroevolutionary prob-
lem. Evolution 38:1219–1237.

LILLIEFORS, H. W. 1969. On the Kolmogorov-Smir-
nov tests for the exponential distribution with mean
unknown. J. Amer. Stat. Assoc. 64:387–389.

MALMGREN, B. A., W. A. BERGGREN, AND G. P.
LOHMANN. 1983. Evidence for punctuated grad-
ualism in the Late Neogene *Globorotalia tumida*
lineage of planktonic foraminifera. Paleobiology 9:
377–389.

RICKLEFS, R. E. 1980. Phyletic gradualism vs. punc-
tuated equilibrium: Applicability of neontological
data. Paleobiology 6:271–275.

ROSEN, D. E. 1979. Fishes from the uplands and
intermontane basins of Guatemala: Revisionary
studies and comparative biogeography. Bull. Amer.
Mus. Nat. Hist. 162:267–376.

SIMPSON, G. G. 1944. Tempo and Mode in Evolution.
Columbia Univ. Press, N.Y.

SITES, J. W., I. F. GREENBAUM, AND J. W. BICKHAM.
1981. Biochemical systematics of neotropical tur-
tles of the genus *Rhinoclemmys* (Emydidae: Bata-
gurine). Herpetologica 34:256–264.

ZIMMERMAN, J. R. 1985. A revision of the genus
Oreodytes in North America (Coleoptera: Dytisci-
dae). Proc. Acad. Nat. Sci. Philadelphia 137:99–
127.

Corresponding Editor: J. A. Endler