Evolutionary Change in the Morphological Complexity of the Mammalian Vertebral Column

Daniel W. McShea

Evolution, Volume 47, Issue 3 (Jun., 1993), 730-740.

EVOLUTIONARY CHANGE IN THE MORPHOLOGICAL COMPLEXITY OF THE MAMMALIAN VERTEBRAL COLUMN

DANIEL W. MCSHEA
Museum of Paleontology, University of Michigan,
Ann Arbor, Michigan 48109

Abstract.—The notion that morphological complexity increases in evolution is widely accepted in biology and paleontology. Several possible explanations have been offered for this trend, among them the suggestion that it has an active forcing mechanism, such as natural selection or the second law of thermodynamics. No such mechanism has yet been empirically demonstrated, but testing is possible: if a forcing mechanism has operated, the expectation is that complexity would have increased in evolutionary lineages more frequently than it decreased. However, a quantitative analysis of changes in the complexity of the vertebral column in a random sample of mammalian lineages reveals a nearly equal number of increases and decreases. This finding raises the possibility that no forcing mechanism exists, or at least that it may not be as powerful or pervasive as has been assumed. The finding also highlights the need for more empirical tests.

Key words.—Complexity, entropy, evolutionary trends.

A commonplace in biology is that morphological complexity increases in evolution (e.g., Darwin 1987; Huxley 1953; Stebbins 1969; Saunders and Ho 1976; Bonner 1988; others reviewed in McShea 1991). Some have understood this increase as a passive process. If the first organisms were as simple as possible, then complexity could only have increased in their descendants (Maynard Smith 1970), thus raising the mean and maximum for life as a whole. As diversity increased, the overall mean and maximum would have continued to drift higher, a process that has been described as diffusion or increasing variance in a bounded morphospace (for discussions of the passive mechanism, see Stanley 1973; Fisher 1986; Gould 1988). The critical component needed to make this mechanism work is a simple asymmetry: complexity must have and reach a lower bound (the complexity of the simplest possible organism) but not an upper one.

However, others have understood the increase in complexity to be active or driven and have invoked entropy (Wicken 1979, 1987; Brooks and Wiley 1988) or natural selection (Bonner 1988; Rensch 1960) as forcing mechanisms. The intriguing suggestion in the first case is that evolutionary lineages are dynamic, nonequilibrium systems, and that all such systems contain a randomizing vector (predicted by the second law of thermodynamics), which tends to increase configurational disorder. In organisms this tendency would be manifest as increasing morphological complexity (Wicken 1987). It may be questioned

whether (increasing) entropy is a driving force or merely a description of the overall character of the evolutionary changes in morphology expected on statistical or energetic grounds. But under either interpretation, the prediction is the same: complexity is expected to increase, on the average, in most lineages.

In the second case, the suggestion is that selection favors increases in complexity because complex organisms have more components and the components are more specialized so that the internal division of labor is greater. With greater division of labor, complex organisms will tend to be more efficient (mechanically, if not energetically) or better adapted in some sense (Bonner 1988).

The active and passive mechanisms can be distinguished empirically. If either the entropic or the selective hypothesis is correct, or if any other forcing mechanism operates, then not only should the overall mean and maximum rise, but complexity should tend to increase in most individual lineages. The passive mechanism, on the other hand, predicts only an elevation of mean and maximum, with no directional bias (except among the very simple lineages lying at or near a lower bound of complexity). Here, in the first empirical test for a forcing mechanism, I examine changes in complexity in the vertebral column in a small sample of mammalian lineages that lie well away from a lower bound. In this sample, a nearly equal number of increases and decreases occur, a result that casts some doubt

on the proposed active mechanisms and invites further testing.

MATERIALS AND METHODS

The complexity of a system is generally acknowledged to be some function of the number of different parts it has, or the degree of differentiation among parts, and of the irregularity of their arrangement (Papentin 1980; Hinegardner and Engelberg 1983; Kampis and Csányi 1987; McShea 1991). Thus, heterogeneous, messy, or irregularly configured systems are complex, such as organisms, automobiles, compost heaps, and junk yards.

Order is the opposite of complexity. Ordered systems are homogeneous, redundant, or regular, like picket fences and brick walls. In this definitional scheme, organisms are not especially well ordered, although they are well organized.

Organization has to do with the degree of structuring of a system for some function, independent of the number of parts it has or its configurational heterogeneity. In other words, organization is measured on a scale orthogonal to the complexity-order scale. Thus, for example, a single piece of steel carefully fashioned to open food tins is organized and simple, because it is highly functional and consists of only a single part. An electric can opener with many parts is organized and complex. And an electric can opener that has been smashed to bits is disorganized and complex; it no longer functions, but it still has many parts irregularly configured.

To many, this scheme will seem counterintuitive. Many will argue that the very specific heterogeneity of intact machinery or organisms should not be confounded with the random heterogeneity of junk piles or compost heaps, in other words, that complexity should include highly specified irregularity and exclude random irregularity. Perhaps it should. At least in principle, some word is needed to make that distinction. In practice, however, the distinction often cannot be made. We often do not know the function of a system or how highly specified its parts and configuration are. The smashed electric can opener (or even the compost heap) may be functioning as a work of art, for example, perhaps one requiring a very specific arrangement of parts. Or it may be a functionless, indifferently structured jumble. The point is that pure structure (parts and configuration) can be treated independently of function and specificity, that we need not know anything about function or specificity
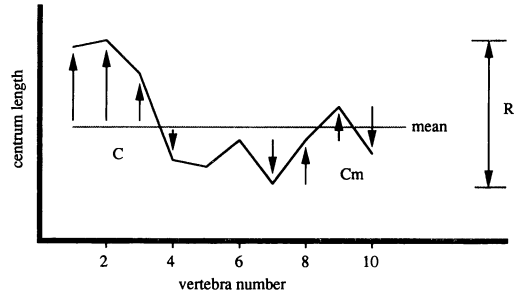

FIG. 1. Plot with artificial data to show how (antilog) values for R, C, and Cm are computed. R is the maximum difference between elements (double-headed arrow, far right). C is twice the average absolute difference between elements and the column mean (shown for four vertebrae, left). (The factor of two was introduced so that the theoretical maximum for C would be R, a nicety that makes no difference here but was relevant in another context [McShea 1992].) Cm is the average absolute difference between adjacent elements (shown for four vertebrae, right). (From McShea 1991.)

to characterize or measure the heterogeneity of structure. To facilitate discussion in this region of ignorance, therefore, some word is needed to characterize the heterogeneity of pure structure, whether functional or not. Here, and increasingly in the technical literature, the word used is "complexity."

Despite the consensus on the abstract meaning of the term, complexity has proven extremely difficult to define operationally, and thus to measure. For whole organisms we lack consistent criteria for identifying and distinguishing parts as well as methods for quantifying their configuration in three dimensions. These problems can be avoided by restricting the investigation to sets of serial structures, such as vertebral columns or arthropod limb series, in which parts are distinct yet comparable and the ordering is linear. Cisne (1974) took this approach in his landmark study of changes in tagmosis (complexity) in aquatic arthropods.

The approach has the virtue that it renders complexity empirically tractable, but has the defect that patterns of change in complexity in a particular organ or substructure may not reflect change in the complexity of the organism as a whole. This defect is no different from that in any sampling procedure, however; a substructure is a kind of sample of a whole organism. And if the sample is unbiased and the sample size large, confidence in the results should be high. In the present study, sample size is small: only one substructure, the vertebral column, was examined.
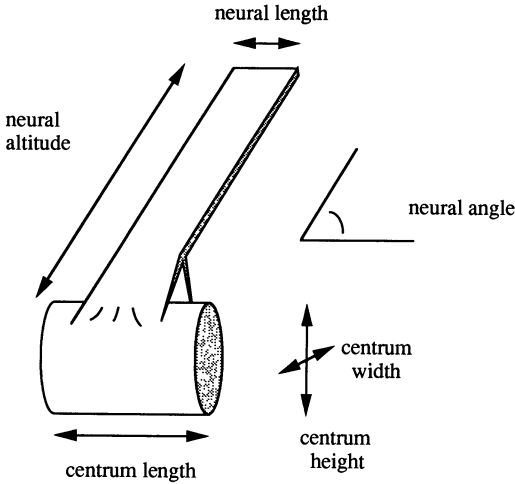
FIG. 2. Schematic of a typical vertebra showing the six vertebral dimensions used in the analysis. (Modified from McShea 1991.)

Thus, further studies measuring change in complexity in other substructures, as well as in other taxa, are needed. However, the sample is probably unbiased, in that we have no a priori or empirical reason to expect that the vertebral column is *not* representative of the organism as a whole, or that its complexity ought to vary independently.

Three Complexity Metrics.—In serial structures, complexity can be understood and measured as the amount of differentiation among the elements along the series. For this study, three metrics were devised to capture three distinct aspects of serial differentiation. The vertebral column was chosen here, but in principle the metrics could be applied to any serial structure.

The metrics are *univariate* and require as raw data a sequence of measurements along a column:

$$R = \log(X_{max} - X_{min})$$

(range)

$$C = \log(2 \Sigma \, |X_i - \bar{x}|/N)$$

(polarization)

$$Cm = \log[\Sigma \, |X_{i+1} - X_i|/(N - 1)]$$

(irregularity)

where X_i is the measurement of ith vertebra, $\bar{x}$ is the mean, and N is the number of vertebrae (fig. 1). R measures complexity as the difference between the elements that are most different from

each other. C is global differentiation, measured as the average absolute difference between the elements and the mean. Cm is local differentiation, or the average absolute difference between adjacent elements. (For a more thorough discussion of the metrics and their properties, see McShea 1992.)

Complexity values were computed for six vertebral dimensions (fig. 2): centrum height, length, and width, and neural altitude, length, and angle. Correlations in complexity values among the dimensions were low (McShea 1991), and for this reason, as well as for simplicity, the six dimensions were treated independently throughout the analysis.

The following measurement protocols were adopted: the atlas was omitted for all taxa, as was the axis for all dimensions except centrum height and width. Only the presacral column was measured. Because fishes lack sacrals, the fish column was truncated at the vertebra just above the pelvic fin, which remains posterior in the species used. Measurements for damaged vertebrae were interpolated using mean values from adjacent vertebrae (see Discussion below).

Size Correction.—The absolute magnitude of the disparities between vertebrae in the columns of large animals tend to be greater than those of small animals, and therefore large animals tend to have systematically greater R, C, and Cm values. Although these first-order differences among taxa could be construed as real differences in morphological complexity, second-order differences or what might be called "shape complexity" is usually of more interest, and some size correction was necessary to measure it with the above metrics. (The correction was not intended to remove all effects of size, or to remove all size-correlated shape variation, but to remove only the first-order component of variation, which is conventionally called size. For this purpose, one of the several standard regression models is appropriate.)

The following procedure was used: for the five linear dimensions (centrum height, length, and width, and neural altitude and length), C values were replaced with their residuals from a reduced major axis computed for C and the log of a "size factor." Likewise R and Cm, plotted on the same scale as C, were replaced with their residuals from the same axis computed for C. Figure 3 shows in principle how this was done; the choice of specimens used to compute regressions in the actual tests is discussed below.

The reduced major axis is the least-biased estimator of the actual relationship between C and the size factor (Rayner 1985; McArdle 1988), but the robustness of the results was tested anyway by repeating the analysis using the major axis and again using a line of slope 1 (i.e., isometry assumed). Because no single size factor captures first-order variation perfectly or completely (Bookstein et al. 1985), the analysis was also repeated using three different size factors (see table 1 caption). Finally, the analysis was conducted using modified metrics in which the data were log-transformed initially and the log operation omitted from the complexity equations. Computed this way, differences become log ratios and no further size correction is necessary.

For neural angle, no size correction was necessary, and the log operation was omitted from the complexity equations.

Removal of Redundancy.—R, C, and Cm are highly correlated with each other, even after size correction, as a consequence of both the mathematical properties of the metrics and the highly coordinated way in which vertebrae in a column vary in development and evolution (McShea 1992). To remove this redundancy, R values (size-corrected) were replaced with their residuals from a reduced major axis computed for R (size-corrected) and C (size-corrected). Figure 4 shows how in principle this was done. These residuals define, in effect, a new kind of complexity, R' or range *relative* to C. Redundancy in Cm was removed in the same way to create Cm' or irregularity relative to C.

Test Cases.—The metrics give intuitively reasonable results for cases in which complexity differences are known, at least qualitatively, in advance. In most fishes and reptiles, vertebrae are quite similar over the whole length of the column, although reptiles show some differentiation in the neck. The typical mammalian column is more complex than both, as reflected in the classical division of the column into distinct cervical, thoracic, and lumbar regions.

Table 1 (cols. 1, 2) shows that these differences are captured by the metrics. In column 1, complexity values for a single fish specimen (*Oncorhynchus*) are compared with mean values for a diverse group of mammals. Regressions for size correction were computed using 31 specimens representing 31 different families and 14 different mammalian orders (see the Appendix). (In computing reduced major axes for R' and Cm', the group of 31 was supplemented with 5 additional
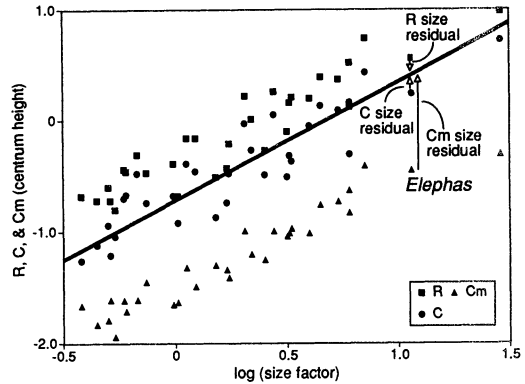

FIG. 3. Plot showing how residuals were computed for size correction. A single regression line was computed for C values and the log of a size factor, and this line used to compute residuals for R (squares), C (circles), and Cm (triangles). Residuals are shown here for only one specimen, an elephant. In this example, the vertebral dimension is centrum height, the regression line is a reduced major axis, the size factor is the log of geometric-mean centrum diameter, and the data are from the 31 mammal specimens (see the text and the Appendix).

families: the Tragulidae, Antilocapridae, Cervidae, Bovidae, and Sciuridae. Because multiple specimens from these groups were measured, mean [size-corrected] values for each family were used in the redundancy-removal calculations.) Similarly, in column 2, mean values for 2 reptile specimens (*Iguana* and *Varanus*) were compared with 31 mammal families (again, 36 for redundancy removal).

Neither the fish nor the reptile specimens were used in computing the regressions, so in effect their complexity values were size- and redundancy-corrected using mammalian data. This protocol is appropriate because the purpose of the comparison is to discover whether the complexity values of these particular fish and reptile specimens could have been drawn from the mammalian population, in other words, whether the columns of these specimens are especially simple *by mammalian standards.* Comparison of mammals with fishes as a group, a different test, would require a common vertebrate size-regression (or independent regressions for fishes and mammals).

In columns 1 and 2, each entry (each string of letters) shows results for various combinations of size factors and size-correction models, including the logarithmic metric (see table 1 caption). Significance of differences between means (or, in the fish comparison, between the single

TABLE 1. Differences in mean complexity between taxa (columns 1 and 2: G, complexity of mammals is greater than fish/reptiles; S, complexity of mammals is smaller than fish/reptiles; g and s, nonsignificant difference) and changes within "lineages" (columns 3–7: I, increase; D, decrease; i and d, nonsignificant change). Entries in the table show results for various combinations of size-correction model and size factor, plus the logarithmic metric (LM, see the text). Models are reduced major axis (RMA), major axis (MA), and isometry (ISOM). Size factors are geometric-mean diameter (MD), or the square root of mean centrum height times mean width; mean centrum length (LENG); and generalized hypotenuse (GEN), or the sum of squared means for the five linear dimensions. For columns 1–4, the combinations are (from left to right within columns): (1) RMA, MD; (2) RMA, LENG; (3) RMA, GEN; (4) MA, MD; (5) ISOM, MD; and (6) LM. For columns 5–7, comparisons were between single specimens, so the mammalian data (31 specimens) were used for computation of regressions, and distributions for bootstrapping were imported from other taxa, specifically *Cephalophus* (CEPH) or *Sciurus* (SCI) (see fig. 5 caption). For column 5: (1) RMA, MD, CEPH; (2) RMA, LENG, CEPH; (3) ISOM, LENG, CEPH; (4) RMA, MD, SCI; (5) RMA, LENG, SCI; and (6) ISOM, LENG, SCI. For columns 6 and 7: (1) RMA, LENG, CEPH; (2) ISOM, LENG, CEPH; (3) RMA, LENG, SCI; (4) ISOM, LENG, SCI. Gaps occur in the table: because neural process dimensions were not measured in the fish, size correction using GEN was not possible; because the original neural length data contain zero values, the log model could not be applied to this dimension; no size correction was necessary for neural angle (and the log operation was omitted from the metrics). I excluded the lumbar vertebrae from the analysis for *Camelus* because they were absent from the *Poebrotherium* specimen. Data for centrum width in *Aetiocetus* were taken from Emlong (1966); measurements for some cervical vertebrae were missing, thus for centrum width all cervicals were excluded for both *Aetiocetus* and *Balaenoptera*.

	1	2	3	4	5	6	7
	Fish/Mam	Rept/Mam	Hyem/Rum	Sciu/Squ	Pat/Man	Aet/Bal	Poe/Cam
Centrum height							
C	Gs GGG	GGGGGG	dDDDii	idiiii	IiiIII		
R'	GG GGG	gggggg	iiiiid	didddd	ddddDD		
Cm'	GG GGG	sggssg	dDdddd	dddddd	dDDddd		
Centrum length							
C	GG GGG	SGsSSG	ididII	idiiiI	IIIIII	DdDD	DDDD
R'	Gg GGg	SgsSSs	IIIIII	DDDDDD	DDDDDD	IIII	DDDD
Cm'	GG GGG	sgsssg	IiIIII	dddddd	DDDDDD	DDDd	DDDd
Centrum width							
C	GG GGG	GGGGGG	idiiii	iddiii	DDDDDD	DDDD	
R'	GG GGG	SSSSSS	iiiiii	dddddd	IIIIII	iiii	
Cm'	GG GGG	SgsSSs	diidii	diddid	iddidd	IIii	
Neural altitude							
C		GGGGGG	IiIIII	iddidi			
R'		ssssss	idiiid	iiiiii			
Cm'		sssssg	dDDdDD	dddddd			
Neural length							
C		GGGGG	iiiii	iddii			
R'		SSSSS	iddid	DDDDD			
Cm'		SSSSS	ddddd	ddddd			
Neural angle							
C		G	d	D			
R'		G	I	d			
Cm'		G	i	d			

values for the fish and the means for mammals) was assessed with a bootstrap (two-tail test, $P \leq 0.05$, 500 resamplings with replacement). In the table, G indicates that the mean complexity of the mammals was significantly greater than the test case, S indicates that the mammalian mean was significantly smaller, and g and s indicate a nonsignificant difference. Note that for most entries in columns 1 and 2 (and throughout table 1), different size factors and models gave concordant results.

The metrics agree with the qualitative assessment. The mammal columns are significantly more complex than the fish column in all three dimensions measured, for all three complexity measures, and for almost all size-correction schemes. Comparing the mammals with two reptiles, mammals have significantly greater *C* for

all dimensions except centrum length. In the relative measures, R' and Cm', the reptiles are more complex in several dimensions, a result that may be surprising at first. The result is reasonable, however: because *absolute* complexity values (C) for reptiles are extremely low, the small differentiation in the neck region is quite significant by *relative* standards (R' and Cm').

A Test for a Forcing Mechanism.—The test consists of a series of comparisons within "lineages," that is, between certain mammalian taxa and their relatively underived sister taxa, which here serve as ancestor surrogates. If a forcing mechanism operates, then complexity should increase among these lineages more frequently than it decreases.

The selection of lineages was based only on the availability of suitable underived sister taxa, and was therefore random with respect to complexity. The following taxa were used:

1. The extant chevrotain (*Hyemoschus*) as an ancestor surrogate for all extant ruminants (Janis 1984; Janis and Scott 1987).
2. The extant tree squirrel (*Sciurus*) for the rest of the extant Sciuridae; Emry and Thorington (1982, 1984) considered *Protosciurus* ancestral to the Sciuridae and found few differences between it and *Sciurus,* none of them in the vertebral column. The propriety of placing *Protosciurus* in the Sciuridae has been challenged (Wood 1980; Vianey-Liaud 1984), but its primitiveness has not been questioned.
3. The extinct *Patriomanis* for the extant pangolin (*Manis*); Emry (1970) noted that the Oligocene *Patriomanis* is the first known member of the Manidae and described it as intermediate in many respects between the probable (more distant) manid ancestor *Palaenodon* and the modern genus. [This conclusion is disputed by Patterson (1978).]
4. The extinct *Aetiocetus* for the extant finback whale (*Balaenoptera*). Emlong (1966) described *Aetiocetus* as a close antecedent of the modern Mysticeti, and Mchedlidze (1984) described it as intermediate between them and the Archaeoceti.
5. The extinct *Poebrotherium* for the modern camel (*Camelus*) (Scott 1891; Herre 1982).

Two criteria were essential in establishing an ancestral or near-ancestral relationship: first the ancestor had to be a sister taxon to the descendant group, and second, it had to be primitive in all or most respects. That is, it had to have
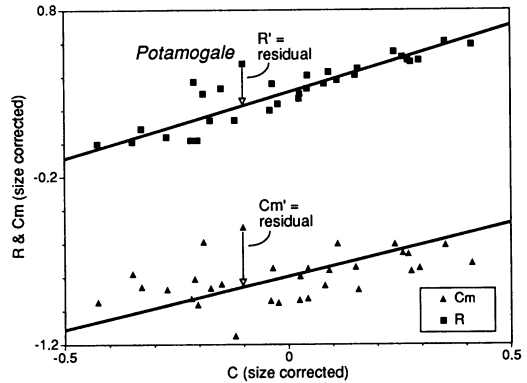

FIG. 4. Plot showing how redundancy was removed from R (squares) and Cm (triangles). R' values are the residuals of R from a reduced major axis for R and C. Cm' values are the residuals of Cm from a reduced major axis for Cm and C. Residuals are shown here for a single specimen, an insectivoran. As in figure 3, the vertebral dimension is centrum height and the data are from the 31 mammal specimens (see the text and the Appendix).

few or no unique specializations (autapomorphies) of its own. Both of these conditions are met for the ancestor surrogates used (see references above). There are exceptions, of course: for example, *Hyemoschus* (like other Tragulids) has some unique derived features in the areas of the postorbital bar and the auditory bullae (Janis 1984). But exceptions are few, and in no case were any significant autapomorphic features noted in the vertebral column.

In the phylogenetic work relied on here, I judged the primitiveness of the character states of the ancestor surrogates using a combination of outgroup comparisons and fossils. In all cases the focus was on cranial (including dental) characters, not on the vertebral column, but there is no reason to think the characters used are not a representative sample of the morphology of the whole organism and therefore a good indication of its overall primitiveness.

RESULTS

Table 1 (cols. 3–7) shows the direction of change in complexity in the lineages and indicates which changes are statistically significant (two-tailed test, $P \leq 0.05$; 500 resamplings with replacement; D, significant decrease; I, significant increase; d and i, nonsignificant change). In the comparison of *Hyemoschus* with other ruminants (col. 3), mean complexity values for 5 *Hyemoschus* specimens were compared with means for 18 ruminant

Cephalophus distribution

TABLE 2. Summary of significant complexity changes. To be counted, a change must have registered as significant in at least half of the various size-correction schemes (if there was more than one) in table 1, and the direction of change must have been the same in all schemes. The question marks are explained in the text.

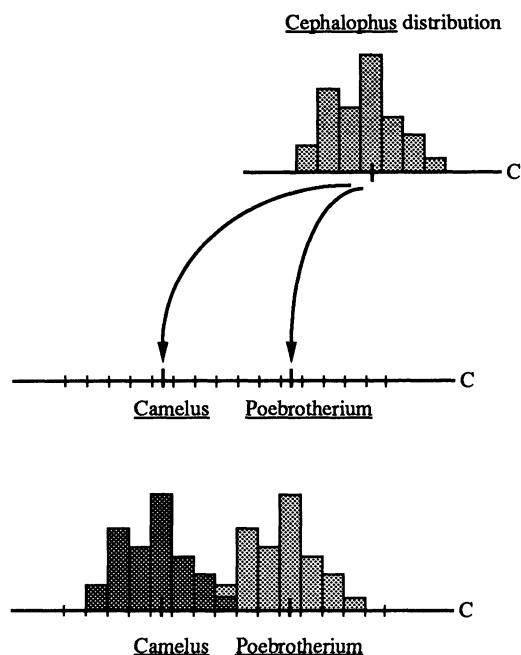|	Increases	Decreases
Ruminants and sciurids		
C	1	1
R'	2	2
Cm'	1	1
Fossils		
C	2	4
R'	2	2
Cm'	1	3 ??

FIG. 5. Schematic showing how the significance of changes was evaluated in the fossil comparisons (columns 5–7 in table 1). Comparisons were between single specimens, so distributions for bootstrapping were imported from other taxa. In the example shown, the (size-corrected) C value for *Poebrotherium* is compared to that of its descendant, *Camelus*. A distribution of C values from another genus, in this case *Cephalophus*, is centered on the C values of the single specimens, and the difference between the two C values evaluated with a bootstrap (two-tailed test, $P \leq 0.05$, 500 resamplings with replacement). The assumption is that distributions of complexity values are comparable among genera. The distribution shown is not an actual one. The actual distributions used consist, in one set of tests (see table 1 caption), of values for seven specimens from seven different species of *Cephalophus*. In the other set, they consist of seven specimens from seven different species of *Sciurus*.

specimens from 18 different ruminant genera (see the Appendix). Size-correction and redundancy-removal regressions were based on these same 18 specimens. For the comparison of *Sciurus* with other sciurids (col. 4), mean values for 7 specimens from 7 different *Sciurus* species were compared with 12 specimens from 12 different sciurid genera (see the Appendix). Size-correction and redundancy-removal regressions were also based on these 12 specimens.

Because fossil comparisons (cols. 5–7) were between single specimens, no entirely appropriate distributions were available for computing re-

gressions or assessing significance of differences. Thus, the mammalian regressions were used for size correction and redundancy removal, and distributions for bootstrap tests were imported from other genera (fig. 5). As with the fish and reptile test cases, each entry in the table shows results for various combinations of size factor and size-correction model (see table 1 caption).

Table 2 summarizes the results: no tendency to increase is evident, either in the ruminant and sciurid lineages or in the fossil lineages.

The reason the two sets of comparisons are tallied separately is that the difference in robustness of the results may be substantial. The fossil comparisons are less reliable, because they are based on single specimens and because the significance of changes was evaluated using distributions imported from other taxa. The use of alien distributions is not clearly justified, but ignoring the statistical significance of results in columns 5–7 would not change the interpretation anyway. Only three additions would have to be made to table 2: one decrease in Cm', one increase in R', and one decrease in R'.

The fossil comparisons are also less reliable because fossil specimens tend to be somewhat deformed and to have rougher surfaces. As a result, measurements from fossils contain more noise than those from the modern columns, which artificially increases fossil complexity values, especially in Cm'. The effect of noise on complexity can be demonstrated with simulations in which another dose of measurement error is added to that already present. Most R', C, and Cm' values increase, with a maximum increase of only 5% for R' and C but as much as 28% for Cm'. Thus, as indicated by question marks, the decreases in

Cm' in the comparisons between fossil and modern columns may not be robust.

An opposite bias is introduced by the interpolation of measurements for missing vertebrae. This protocol ought to have little or no effect on *R'* and *C,* but it tends to decrease *Cm'.* The bias should not affect the comparisons involving modern specimens, because the number of interpolations is small (less than 0.5% of all measurements were interpolated), and because the decrease in *Cm'* should be the same, on the average, in both ancestors and descendants. The effect in the fossil comparisons, however, should be to produce spurious increases in complexity, because the frequency of interpolations in the fossil specimens is much greater (about 10%) than in the modern specimens to which they are compared. Such a bias is acceptable in the present context, because it renders the test more conservative. It also works in the opposite direction from the noise bias and may offset it somewhat, although the relative magnitudes of these two biases are unknown.

DISCUSSION

The preceding discussion takes a trend for granted in order to inquire into its mechanism. That is, it accepts the evolutionary conventional wisdom that a complexity trend occurs at some large scale, one encompassing many or most higher taxa and some large portion of the history of life. Actually, with the exception of Cisne's (1974) study, almost no trend data are available at any taxonomic or temporal scale (McShea 1991), but many people share the subjective impression of increase and this gestalt may have a foundation in reality. In any case, if the absence of hard evidence is overlooked, and if a large-scale pattern of increase in mean and maximum complexity is granted, then either a passive or an active mechanism (or both) is needed to explain it. (For further discussion of these alternative mechanisms, see Gould 1988.)

The Passive Mechanism.—It seems inescapable that complexity must have increased passively, at least to some extent. Consider the largest scale, life as a whole over its entire history. Suppose that, in the absence of a forcing mechanism, complexity increases and decreases are about equally likely among all lineages. Then, if diversity increases (on average), if a complexity lower bound (the complexity of the simplest possible organism) exists and is reached, and finally if no *upper* bound is reached—all reasonable assumptions—mean complexity will tend to rise. Among the most complex lineages, increases will occur occasionally, perhaps with the advent of new designs, and the complexity maximum will rise. (For example, this mechanism may account for the increase in complexity in the Precambrian transitions from unicellular to multicellular organisms). A passive mechanism may also operate sometimes at smaller scales, to raise mean and maximum complexity within individual higher taxa or in structures within organisms (such as the vertebral column).

Importantly, the passive mechanism does not guarantee a long-term, monotonic, increasing trend. Mean and maximum complexity must passively decrease sometimes also. For example, maximum complexity is expected to decrease when diversity decreases, perhaps during mass extinctions.

Also, the claim that certain complexity increases were passive does not deny that the specific morphological changes involved were adaptive or that they were driven by natural selection. The implication of the passive mechanism is only that complexity increases are *not* the result of selection favoring complexity itself, or some close correlate of complexity—perhaps specialization or mechanical efficiency.

For example, a passive mechanism may account for the increase in vertebral column complexity in the evolutionary transition from fish to reptile. The lower bound for the vertebral column is roughly the fish condition, in which the vertebrae are more or less the same all along the column. The fish species used here to demonstrate the metrics is highly derived, but the actual ancestor probably had an equally undifferentiated column, and column complexity could only have increased in the course of the modifications leading to reptiles. The particular vertebral modifications that occurred in this transition would have been mainly adaptations to allow locomotion on legs and life on land, and would therefore have been driven by natural selection. If, however, this transition involved no increase in specialization, mechanical efficiency, or any other property thought to be connected with complexity (and I see no reason to think it did), then the increase in complexity itself would have been merely an adjunct effect and therefore, in an important sense, passive.

Another example makes the same point for passive change with no increase in complexity (i.e., in the absence of a lower bound). The in-
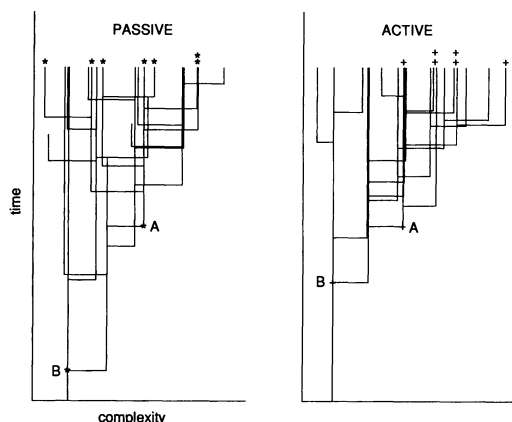
FIG. 6. Distinguishing active (or driven) and passive mechanisms. Using B as an ancestor, the two mechanisms would be indistinguishable, because in both systems, most of B's descendants are more complex than it. However, using A as an ancestor, starting closer to the middle of the complexity range, the results are different. In the driven system, most of A's descendants (crosses) are more complex than it, while in the passive system, the same number (asterisks) of increases and decreases occur.

crease in neural altitude C in the ruminants (table 1) may have been the direct result of adaptive changes to accommodate increased size, specifically, disproportionate (in relation to size) lengthening of the neural processes in the neck for support of a disproportionately heavier head. And the decrease in centrum length and centrum width C in the whale lineage (table 1) may reflect a change toward a more uniform column to accommodate the more uniform distribution of stresses associated with life in the water. The point is that the various adaptive modifications of the vertebral column in evolution will sometimes entail complexity increases and sometimes decreases, but the important selective factors may well be related to column function, not complexity itself.

A passive process such as this probably accounts for the major empirical finding of this study, the roughly equal number of increases and decreases in table 2.

A Methodological Aside. — These examples also make a methodological point that will be important in any future empirical tests for active and passive mechanisms. In the fish-reptile transition, an increase would be expected whether or not selection favors complexity itself (or its correlates). Near a lower bound, the active and passive mechanisms make the same prediction.

However, because the vertebral columns of the lineages used in this study lie far from the lower bound, complexity can either increase or decrease. Increases will occur more frequently if selection favors complexity itself, but in the absence of any such selection, the expectation is that increases and decreases will occur in about equal numbers (as was found). The important methodological point is that to distinguish the active and passive patterns, the lineages examined must lie far from a lower bound.

Figure 6 shows how this works in principle. The system on the left is purely passive. Complexity starts low, and increases and decreases occur with equal probability at each branching event, except that decreases below a fixed lower bound are not allowed. (In this case, the lower bound is the y-axis.) The system on the right is active or driven. The starting point and lower bound are the same as in the passive system, but increases are significantly more likely than decreases. (In this particular case, which is decidedly unrealistic but illustrates the point, only two decreases occurred.)

In both systems, B is an ancestor lying near the lower bound of complexity, and almost all of its descendants (the terminal lineages) are more complex than it. Thus, if B were the ancestor used in the comparison, the active and passive mechanisms would be indistinguishable. However, using A as the ancestor, starting closer to the middle of the complexity range, the results are different. In the driven system, all but one of A's descendants (marked with crosses) are more complex than it, whereas in the passive system, the same number (asterisks) of increases and decreases occur. (Note that the experimental setup for the test in this study was different in detail but applied the same logic. Instead of counting the number of increases and decreases from a single ancestor, multiple ancestor-descendant pairs were used, all from somewhere around the middle of the complexity range.)

The Active Mechanism. — The existence of an active forcing mechanism remains uncertain. In this first attempt to treat the problem empirically, none was detected. However, the point has been made that the comparisons presented here cover a rather short time span (about 30 million years) and were constrained by the limitations on vertebral column variability within fairly specialized locomotory modes. In some cases, the point is well taken: if complexity did not increase, on the average, in the ruminants, it could

be that the range of ways in which ruminants run, walk, support their weight, and so forth has expanded little throughout the history of the group, so little opportunity for modification in the vertebral column has occurred. (In sciurids, the point is less apt.)

Thus, no general conclusion should be drawn from this single, limited finding. Forcing mechanisms may yet be found, perhaps on longer time scales and in comparisons spanning a wider range of locomotory modes or adaptive zones. They may also be found in other characters, in larger samples, or in other taxa. These are precisely the areas in which further empirical tests are needed, but this study does raise for serious consideration the possibility that no forcing mechanism exists for complexity.

LITERATURE CITED

Bonner, J. T. 1988. The evolution of complexity. Princeton University Press, Princeton, N.J.

Bookstein, F., B. Chernoff, R. Elder, J. Humphries, G. Smith, and R. E. Strauss. 1985. Morphometrics in evolutionary biology. Academy of Natural Sciences of Philadelphia, Special Publication no. 15, Philadelphia, Pa.

Brooks, D. R., and E. O. Wiley. 1988. Evolution as entropy, 2d ed. University of Chicago Press, Chicago.

Cisne, J. L. 1974. Evolution of the world fauna of aquatic free-living arthropods. Evolution 28:337–366.

Darwin, C. 1987. Notebook E. P. 42 in P. H. Barrett et al., eds. Charles Darwin's notebooks. Cornell University Press, Ithaca, N.Y.

Emlong, D. 1966. A new archaic cetacean from the Oligocene of northwest Oregon. Bulletin, University of Oregon Museum of Natural History, 3.

Emry, R. J. 1970. A North American Oligocene pangolin and other additions to the Pholidota. Bulletin of the American Museum of Natural History 142:456–510.

Emry, R. J., and R. W. Thorington, Jr. 1982. Descriptive and comparative osteology of the oldest fossil squirrel, Protosciurus (Rodentia, Sciuridae). Smithsonian Contributions to Paleobiology, 47.

———. 1984. The tree squirrel Sciurus (Sciuridae, Rodentia) as a living fossil. Pp. 23–31 in N. El-

dredge and S. M. Stanley, eds. Living fossils. Springer, New York.

Fisher, D. C. 1986. Progress in organismal design. Pp. 99–117 in D. M. Raup and D. Jablonski, eds. Patterns and processes in the history of life. Springer, Berlin.

Gould, S. J. 1988. Trends as changes in variance: A new slant on progress and directionality in evolution. Journal of Paleontology 62:319–329.

Herre, W. 1982. Zur Stammesgeschichte der Tylopoden. Verhandlungen der Deutschen Zoologischen Gesellschaft 1982:159–171.

Hinegardner, R., and J. Engelberg. 1983. Biological complexity. Journal of Theoretical Biology 104:7–20.

Huxley, J. S. 1953. Evolution in action. Harper and Brothers, New York.

Janis, C. M. 1984. Tragulids as living fossils. Pp. 87–94 in N. Eldredge and S. M. Stanley, eds. Living fossils. Springer, New York.

Janis, C. M., and K. M. Scott. 1987. The interrelationships of higher ruminant families with special emphasis on the members of the Cervoidea. American Museum Novitates, 2893.

Kampis, G., and V. Csányi. 1987. Notes on order and complexity. Journal of Theoretical Biology 124:111–121.

Maynard Smith, J. 1970. Time in the evolutionary process. Studium Generale 23:266–272.

McArdle, B. H. 1988. The structural relationship: Regression in biology. Canadian Journal of Zoology 66:2329–2339.

Mchedlidze, G. A. 1984. General features of the paleobiological evolution of the Cetacea. Oxonian Press, New Delhi.

McShea, D. W. 1991. Complexity and evolution: What everybody knows. Biology and Philosophy 6:303–324.

———. 1992. A metric for the study of evolutionary trends in the complexity of serial structures. Biological Journal of the Linnean Society 45:39–55.

Papentin, F. 1980. On order and complexity. I. General considerations. Journal of Theoretical Biology 87:421–456.

Patterson, B. 1978. Pholidota and Tubulidentata. Pp. 268–278 in V.J. Maglio and H.B.S. Cooke, eds. Evolution of African mammals. Harvard University Press, Cambridge, Mass.

Rayner, J.M.V. 1985. Linear relations in biomechanics: The statistics of scaling functions. Journal of Zoology 206:415–439.

Rensch, B. 1960. Evolution above the species level. Columbia University Press, New York.

Saunders, P. T., and M. W. Ho. 1976. On the increase in complexity in evolution. Journal of Theoretical Biology 63:375–384.

Scott, W. B. 1891. On the osteology of Poebrotherium: A contribution to the phylogeny of the Tylopoda. Journal of Morphology 5:1–79.

Stanley, S. M. 1973. An explanation of Cope's Rule. Evolution 27:1–26.

Stebbins, G. L. 1969. The basis of progressive evolution. University of North Carolina Press, Chapel Hill.

Vianey-Liaud, M. 1984. Possible evolutionary relationships among Eocene and Lower Oligocene ro-

740 DANIEL W. McSHEA

dents of Asia, Europe, and North America. Pp. 277–309 *in* W. P. Luckett and J.-L. Hartenberger, eds. Evolutionary relationships among rodents. Plenum, New York.

Wicken, J. S. 1979. The generation of complexity in evolution: A thermodynamic and information-theoretical discussion. Journal of Theoretical Biology 77:349–365.

———. 1987. Evolution, thermodynamics, and information. Oxford University Press, Oxford.

Wood, A. E. 1980. The Oligocene rodents of North America. Transactions of the American Philosophical Society 70:1–68.

Corresponding Editor: C. Hickman

APPENDIX

Genera Names of 31 Mammal Specimens	Genera Names of 18 Ruminant Specimens
Ornithorhynchus	*Antilocapra*
Tachyglossus	*Odocoileus*
Didelphis	*Cervus*
Macropus	*Mazama*
Dasyurus	*Rangifer*
Hydrochoerus	*Hydropotes*
Dasyprocta	*Hemitragus*
Pedetes	*Bison*
Lagostomus	*Tetracerus*
Canis	*Oreotragus*
Acinonyx	*Gazella*
Ursus	*Antilope*
Lutra	*Ammodorcas*
Phoca	*Cephalophus*
Odobenus	*Capra*
Arctocephalus	*Ammotragus*
Macaca	*Damaliscus*
Homo	*Kobus*
Potamogale	**Genera Names of 12 Sciurid Specimens**
Solenodon	
Elephas	*Cynomys*
Phacochoerus	*Callosciurus*
Giraffa	*Marmota*
Camelus	*Ratufa*
Tapirus	*Tamiasciurus*
Equus	*Xerus*
Rhinoceros	*Sundasciurus*
Trichechus	*Spermophilus*
Manis	*Dremomys*
Lepus	*Petaurista*
Balaenoptera	*Aeromys*
	Hylopetes