Rates of evolution in Plio-Pleistocene mammals: six case studies

PHILIP D. GINGERICH

In:

Morphological change in Quaternary mammals of North America

6

Rates of evolution in Plio-Pleistocene mammals: six case studies

PHILIP D. GINGERICH

Evolution, in general, means change over time: The end is different from the beginning. Change is not inevitable, but it seems to be common. And some change can almost always be observed, given enough time. When do we see change? How much time does it take? To study evolution we have to measure change and time, and to understand evolution we have to know how the two are related.

Evolution in biology is often described as a process, but it is really a collection of processes – including the processes of mutation, dispersion, drift, and selection – that together produce whatever change we see. These processes may each be complicated individually, and their interactions produce additional complexity. Evolution is widely acknowledged as a fundamental concept in biology, and yet we have a surprisingly anecdotal and casual knowledge of how evolutionary change takes place. One way to improve this understanding is to quantify change over time in terms of rate. Evolutionary rates are quantitative expressions relating change to time.

Evolution takes place on many scales of time. Field and laboratory experiments can be designed to study change on short time scales, and fossils provide the most direct and best information about evolution on long time scales. The principal problem with the fossil record is that the time scales involved, typically millions of years, are so long that they are difficult to relate to the time scales of our lifetimes (and those of other organisms). Biologists as a group have a surprisingly poor understanding of evolution on a geological scale of time, and paleontologists as a group have a surprisingly poor understanding of evolution on a biological scale of time. One reason for this is that we have almost no

record of changes on intermediate scales of time, scales of hundreds or thousands of years, that would permit evolution on a laboratory scale of time to be related to evolution on a geological scale.

Fortunately biologists can sometimes study evolution on a scale of hundreds or thousands of years by taking advantage of accidents and events that have affected organisms, (e.g., chance colonization events) and have been documented in human history. And in favorable circumstances, paleontologists, too, can resolve time on scales of hundreds or thousands of years. It is generally true that time is easier to resolve the less far back we are in geological time, and the Pleistocene, spanning much of the past 2 million years, is important in this regard. The six Plio-Pleistocene studies discussed here illustrate a paleontological approach to understanding evolution on a wide range of time scales.

Quantification of evolutionary rates

Quantification of evolutionary rates is still in its infancy, and the history of ideas on this can be summarized rather easily. Time is uniquely important in paleontology, and it comes as no surprise that paleontologists figure prominently in the quantification of evolutionary rates. George Gaylord Simpson provided the first substantial treatment of rates in his book *Tempo and Mode in Evolution* (1944). At that time the geological time scale was poorly calibrated numerically, and Simpson was unwilling to trust estimates of the durations of epochs. He published a table of various measures of tooth size in horses and showed graphically that if the length of the ectoloph evolved at a constant rate, then the height of the paracone did not (Simpson, 1944, Figure 2). Simpson also provided a graph of these two characteristics plotted against geological time, showing that in all likelihood neither had evolved at a constant rate (Simpson, 1944, Figure 4). He noted that the slopes of lines in his graphs were proportional to rates of evolution, but did not attempt to quantify these. Simpson's most noteworthy contribution to quantification was recognition from the beginning that proportional change, not absolute change, is the quantity of interest, and consequently that a proportional (logarithmic) scale is the appropriate measurement scale for morphology.

J. B. S. Haldane (1949) quantified evolutionary rates in terms of proportional change and proposed the *darwin*, defined as change by a factor of *e* per million years (where *e* is the base of natural logarithms), as a convenient rate unit. Haldane calculated rates of evolution in horses

for Simpson's two linear measures and for their ratio (Haldane's Table 1), but seems not to have recognized the incomparability of rates quantified for measures of different dimensions. Haldane mentioned the desirability of measuring morphological change in standard deviation units, and he also mentioned the desirability of measuring time in generations, suggestions developed further in this chapter.

Björn Kurtén (1959) measured rates of evolution in various ways in mammals that lived during the pre-Pleistocene (Tertiary), the Pleistocene, and the post-Pleistocene (Holocene), finding that each time period had its own characteristic range. Pre-Pleistocene mammals had rates ranging from 0.003 to 0.2 darwins (mean 0.02 d), Pleistocene mammals had rates ranging from 0.12 to 2.3 d (mean 0.5 d), and post-Pleistocene mammals had rates ranging from 3.7 to 43 d (mean 12.6 d). This led Kurtén to suggest that rates for Quaternary (Pleistocene and post-Pleistocene) mammals may have been higher than those for Tertiary (pre-Pleistocene) mammals, because environmental change in the Quaternary was rapid and revolutionary, whereas environmental change in the Tertiary was slow and gradual. In addition, Kurtén suggested the possibility that slow Tertiary rates might be partially or even wholly spurious because they are based on samples millions of years apart, with intervening histories that may have contained any amount of fluctuation at higher rates. Rates measured over long intervals of time are generally lower than rates measured over shorter intervals, and this is due in part to time-averaging of "fluctuating" higher rates (Gingerich, 1983). This does not mean that any are spurious, but simply that the effect of interval length must be considered when rates are compared.

Log rate–log interval graphs

Evolutionary rates do not exist independently of the interval over which they are measured. This observation is easily accommodated by plotting rates in an interval context. The graphs employed here have interval length plotted on a proportional scale on the abscissa, and the absolute value of evolutionary rate plotted on a proportional scale on the ordinate (log$_{10}$ is used for convenience, because 10 is familiar as the base of our numbering system, but any base could be used). In the following discussion these "log rate–log interval" graphs are referred to as LR-LI graphs. Absolute values of rates are plotted on the assumption that rates of proportional increase are no different than rates of proportional decrease, and assignment of direction is arbitrary in any case.

Processes

Any real process producing systematic change (e.g., natural selection for small body size) does so at some characteristic intrinsic rate that is independent of interval length, meaning that the simple result of a single real process carried out long enough will plot as a horizontal distribution of average slope zero (0.0) on a rate-versus-interval graph (and on an LR-LI graph). The result is a distribution, because intrinsic rates and net rates (discussed later) have variance, and the horizontal line labeled "Process" (for process change) in the inset key in the following LR-LI graphs represents the zero slope of this possibility.

Stasis

Alternatively, a simple process operating at zero rate, or two or more processes that cancel each other out (as well as absence of a process), will produce no systematic change over time (stasis), meaning that differences observed by chance will tend to be constant while interval length increases. Rates calculated from these necessarily will decrease in inverse proportion to interval length, and in this alternative case, if carried out long enough, rates will plot as a distribution with an average slope of negative unity (−1.0) on an LR-LI graph. Again, the result is a distribution, because intrinsic and net rates have variance (lacking variance, stasis at zero rate could not be plotted on an LR-LI graph). The line of slope −1.0 labeled "Stasis" in the inset key of the following LR-LI graphs represents the slope of this possibility.

Pure process and pure stasis are two extreme possibilities in comparing rates measured over a range of time intervals, but neither necessarily happens all the time. Process and stasis can be combined in any proportion, and the slope of the resulting rate distribution on an LR-LI graph will reflect the proportion. The slope will be near zero if process predominates, and near −1.0 if stasis predominates.

Random change

One additional possibility, random change, deserves mention. Appearance of randomness usually is taken to mean that real component factors or component processes interact in complex ways that defy separation. For some purposes, these can be adequately modeled together rather than studied individually. Random walks can mimic

process, and they can mimic stasis, but carried out long enough most will plot as a distribution with a slope of −0.5 on an LR-LI graph. The line of slope −0.5 labeled "Random" in the inset key of the following LR-LI graphs represents the slope of this possibility. Random change is an intermediate null model against which both process change and stasis must be compared to evaluate their possible significance (Gingerich, 1992).

Intrinsic rates and net rates

All stepped processes (including random drift) have two rates: an *intrinsic rate* that is the average rate observed at each step of the process (the generation-by-generation rate in evolution), and a *net rate* that is the average rate calculated over one or more steps (generations). The net rate can never exceed the intrinsic rate, although net rates approximate intrinsic rates in any real process. Net rates typically decrease in proportion to the square root of step number for random processes (hence the −0.5 slope on an LR-LI graph), and net rates decrease in proportion to step number in stasis (hence the −1.0 slope on an LR-LI graph).

Empirical limits

Empirically it turns out that rates of evolution have an upper bound and a lower bound, and both decrease systematically with interval length. Log rate of the upper bound and log rate of the lower bound decrease in inverse proportion to log interval, and each plots as a straight line of −1.0 slope. On an LR-LI graph, the upper bound is shown as a solid line because it reflects both a structural limit and an artifice of limited perception, and the lower bound is shown as a dotted line because it is purely an artifact of limited measurement.

The empirical upper and lower bounds of evolutionary rates define the limits of a broad distribution of possible rates having the −1.0 slope characteristic of stasis, which indicates that stasis, rather than process change or drift, predominates in long-term evolution.

Temporal scaling of evolutionary rates

The inverse relationship of evolutionary rates to the interval of time over which they are calculated means that rates can be compared directly

only if they are calculated over intervals of the same length. When a distribution of rates calculated over a range of intervals is available, it is sometimes possible to use the relationship of rate to interval within this distribution to predict rates for some interval outside the original range of intervals. This *temporal scaling* is useful when comparing different distributions, and it is particularly useful for predicting intrinsic rates on a scale of one generation or a scale of 1 year.

Confidence on other time scales

Inference concerning the relationship of evolutionary rates and time intervals is most easily carried out in an LR-LI context. Log rate is a derived variable dependent on log interval; however, ordinary least-squares regression cannot be used for extrapolation and prediction in this context because distributions of derived log rates are usually negatively skewed, which means that medians are better than mean values as measures of location and that squaring deviant outliers would bias regression. The method of slope-and-intercept computation employed here is *robust maximum likelihood* (RML) estimation minimizing absolute deviations. The distribution of residual variation about a line with the computed slope and intercept is shown as an inset histogram in the lower left corner of all LR-LI graphs. Confidence intervals are calculated by bootstrapping, following Efron and Tibshirani (1986). A linear model is used for extrapolation and prediction because of its simplicity, recognizing that any relationship of log rate to log interval may not be linear over its entire range.

Rate units

Haldane (1949) proposed a standard unit for rates of morphological evolution called the *darwin*, quantifying rates of morphological evolution in terms of factors of *e* and time in millions of years:

$$\text{rate } (d) = \frac{\ln x_2 - \ln x_1}{t_2 - t_1},$$ (6.1)

where $\ln x_1$ and $\ln x_2$ are individual natural log (ln) measurements or sample means of ln measurements at times t_1 and t_2, respectively. This is useful when comparing changes in measures of the same dimension (linear, areal, or volumetric) per unit time, but it is not generally useful when comparing changes in length with changes in area, or changes in

length with changes in volume, or changes in length with any compound ratio (shape).

Haldane mentioned the desirability of measuring morphological change in standard deviation units, and our colleague Björn Malmgren has suggested a rate unit called the *simpson*, quantifying rates of morphological evolution in terms of standard deviations per year or per million years. This has the advantage that standard deviations are measured in the original measurement units, making rates in standard deviation units independent of original dimension and thus more widely comparable.

Haldane also mentioned the desirability of measuring time in generations, and another rate unit of general interest is what I call the *haldane* (*h*), quantifying rates of morphological evolution in terms of standard deviations per generation. Morphological change still has to be corrected for proportion, and we are thus interested in change in natural logarithms (ln) of measured quantities, expressed in standard deviations of the natural log measurements. That is,

$$\text{rate } (h) = \frac{(\ln x_2 - \ln x_1)/s_{\ln x}}{t_2 - t_1} = \frac{z_2 - z_1}{t_2 - t_1}, \qquad (6.2)$$

where *x* and *t* are defined as in equation (6.1), and ln x_1 and ln x_2. Natural logarithms are convenient here because the coefficients of variation commonly reported in the literature are a good approximation of the standard deviation of ln measurements (Lewontin, 1966). The term ln(mean *x*) is used to approximate the mean of ln *x* when necessary. The notation h_0 refers to intrinsic evolutionary rate in haldanes estimated at a scale of one generation, that is, standard deviations per generation at a scale of one generation (which is simply the rate intercept where log interval = 0 on any LR-LI graph).

Generation time (*G*) is related to body mass (*M*) in mammals by the power function

$$G = 0.16M^{0.26}, \qquad (6.3)$$

where *G* is measured in years, and *M* is measured in grams. This equation is derived by regression of log *G* on log *M* (*N* = 66, *r* = 0.82), and it can be used to estimate generation time from body mass when body mass is known. Data used in the regression come from Eisenberg (1981), with generation time taken as the sum of age at first mating plus gestation time. Equation (6.3) yields an approximation that should be rounded

downward for "*r*-selected" mammals and upward for "*K*-selected" mammals, especially those living in seasonal climates.

Rates in Plio-Pleistocene mammals: six case studies

Strait Canyon Microtus

The microtine rodent *Microtus pennsylvanicus* (meadow vole) was a common constituent of Pleistocene mammalian faunas of North America (and it is still widely distributed today). A. Barnosky studied samples of *M. pennsylvanicus* from 7 of 46 sampled intervals in Strait Canyon Fissure in Highland County, Virginia (Barnosky, 1990, and personal communication). Two of these sampling intervals 1.35 m apart have radiocarbon ages of 18.420 and 29.870 ky BP. Interpolation and extrapolation suggest that the seven *Microtus*-bearing intervals range in age from about 47.680 ky BP to about 20.960 ky BP. Samples are as few as 1.270 ky or as many as 26.720 ky apart in time. The *M. pennsylvanicus* weigh about 30 g and have a life span of about 1 year. Substitution in equation (6.3) indicates that *Microtus* probably reproduced at a rate of about three generations per year in the Pleistocene.

Rates of change of the length of the last upper molar (M³) in Strait Canyon *Microtus* are shown in the LR-LI graph in Figure 6.1. Twenty-one rates can be calculated for six different intervals of time ranging from 3.810 generations to 80,160 generations. Twenty of the rates are nonzero. Thirteen rates are positive, indicating evolution toward larger size, and seven are negative, indicating evolution toward smaller size. The median rate is 0.000016 standard deviation per generation ($10^{-4.798}$), numbers on a median time scale of about 33,000 generations ($10^{4.519}$), numbers that mean little by themselves. Fortunately, the *M. pennsylvanicus* data contain additional information.

All of the *Microtus* rates taken together have an RML slope of -.512, which suggests random change over time. The mean absolute deviation is 0.245. There are few enough data points that a bootstrapped 95% confidence interval on the slope ranges from -0.116 to -1.250, ruling out process change (slope 0.0), but including both randomness (slope -0.5) and stasis (slope -1.0). The RML intercept estimate is h_0 = 0.003 ($10^{-2.487}$), which suggests an intrinsic evolutionary rate of 0.003 standard deviation per generation on a time scale of one generation. Artificial selection experiments achieve intrinsic rates on the order of
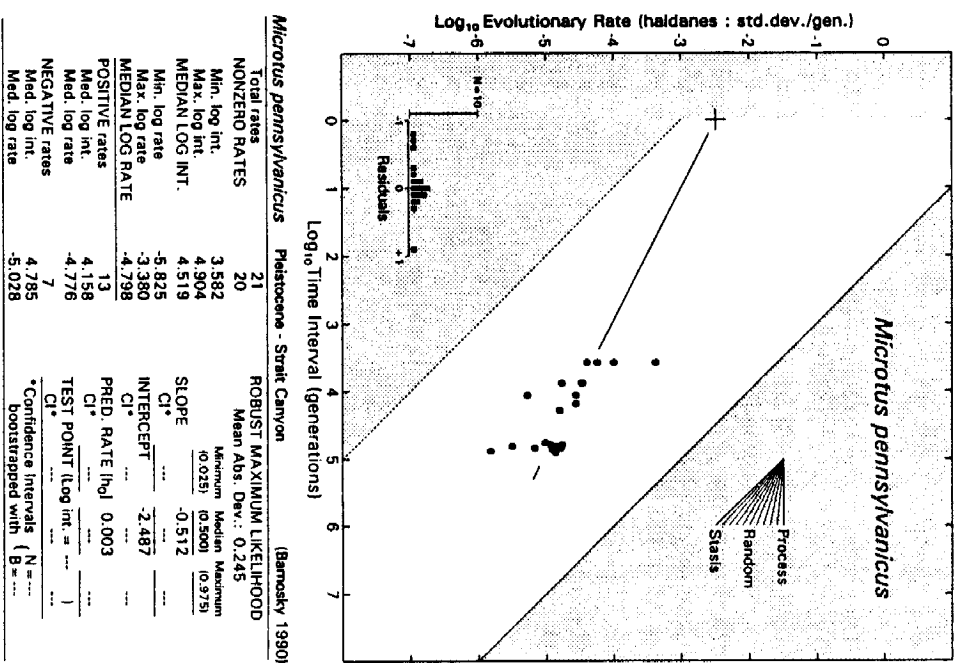
Figure 6.1. LR-LI graph for measurements of M^1 length in seven samples of late Pleistocene *Microtus pennsylvanicus* from Strait Canyon Fissure in Virginia. Solid circles are individual rates. Samples range in age from about 47,680 to 20,960 ky BP. *Microtus* is assumed to have had three generations per year, and rates are calculated over intervals ranging from 3,810 to 80,160 generations. Slope of rate distribution is −0.512. Intercept is −2.487, yielding an intrinsic rate h_0 = 0.003 standard deviation per generation on a time scale of one generation. This sample of rates is small enough that confidence intervals on the slope, extrapolated intercept, and h_0 (not shown) are very broad. Vertical scale bar represents $N = 10$ on inset histogram of residuals (mean absolute deviation is 0.245).

Microtus pennsylvanicus   Pleistocene - Strait Canyon   (Barnosky 1990)

		ROBUST MAXIMUM LIKELIHOOD Mean Abs. Dev.: 0.245		Minimum (0.025)	Median (0.500)	Maximum (0.975)
Total rates	21					
NONZERO RATES	20					
Min. log int.	3.582					
Max. log int.	4.904					
MEDIAN LOG INT.	4.519	SLOPE	CI*	...	-0.512	...
Min. log rate	-5.825	INTERCEPT	CI*	...	-2.487	...
Max. log rate	-3.380	PRED. RATE (h₀)	CI*	...	0.003	...
MEDIAN LOG RATE	-4.798	TEST POINT (Log int. = ...)	CI*		...	
POSITIVE rates	13					
Med. log int.	4.158					
Med. log rate	-4.776					
NEGATIVE rates	7					
Med. log int.	4.785					
Med. log rate	-5.028					

*Confidence Intervals bootstrapped with (N = ... B = ...)

0.200 standard deviation per generation, and this *Microtus* result is about two orders of magnitude less than rates imposed by strong selection. However, here again there are few enough data points in the original rate sample that a bootstrapped 95% confidence interval for the predicted value of h_0 ranges from 0.00005 to 8.300 standard deviations per generation.
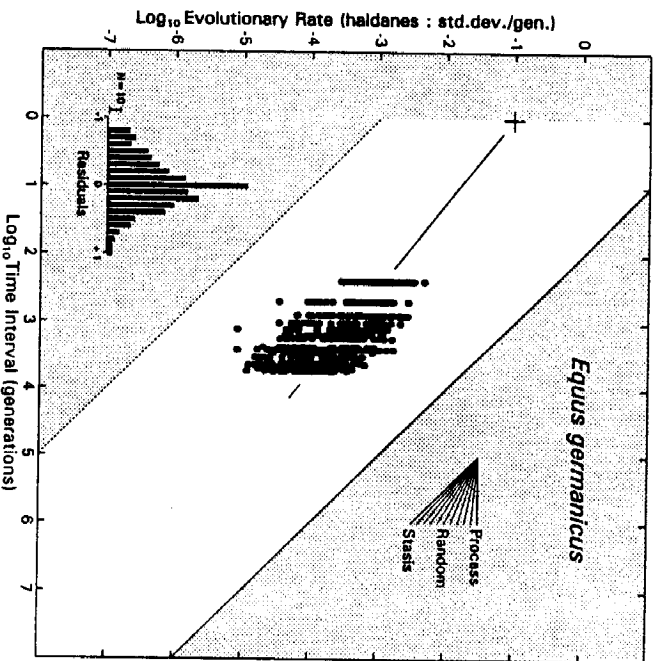
This example illustrates how net rates of evolution calculated on different time scales can be used to predict an intrinsic rate of change on a time scale of one generation, but the result cannot be distinguished from random change over time, and the confidence interval on the prediction is very broad.

Sandalja II Equus

A. Forsten (1990) studied late Pleistocene samples of the horse *Equus germanicus* from 7 of 12 sample intervals in the Sandalja II cave near Pula, in Istria, northwestern Yugoslavia. Two levels in the cave have radiocarbon ages, one near the bottom of the *Equus*-bearing sequence (25,340 ky BP), and one near the top (10,830 ky BP). Extrapolation and interpolation suggest that the seven *E. germanicus*-rich levels have ages of approximately 27,140, 24,740, 19,385, 14,675, 12,320, 11,575, and 10,830 ky BP. Successive samples differ by as few as 745 years and by as many as 5,355 ky, and *Equus* is assumed to have had a generation time of about 3 years in the Pleistocene.

Rates of change of measures of tooth size taken at the base of the crown and at the occlusal surface in Sandalja II *Equus* are shown in the LR-LI graph in Figure 6.2. A total of 714 rates can be calculated for intervals of time ranging from about 248 to 5,432 generations, and 666 of these rates are nonzero. Of the 666, 274 rates are positive, and 392 are negative, indicating a predominance of evolution toward smaller size. The median rate is 0.00021 standard deviation per generation ($10^{-3.686}$), on a median time scale of 2,466 generations ($10^{3.392}$). This median rate is about 13 times that for Strait Canyon *Microtus*, but the *Equus* rates are measured on a shorter median time scale.

All of the *Equus* rates taken together have an RML slope of −0.797, which lies on the stasis side of random. Here there are enough data points to constrain confidence intervals, and bootstrapped confidence intervals for the slope, intercept, and predicted intrinsic rate h_0 are shown in Figure 6.3. The 95% confidence interval on the slope ranges from −0.731 to −0.894. This narrow confidence interval rules out both

Equus germanicus — LR-LI graph (Figure 6.2)

Axis labels: $\log_{10}$ Evolutionary Rate (haldanes : std.dev./gen.) ; $\log_{10}$ Time Interval (generations). Fields labelled Process, Random, Stasis. Inset histogram of Residuals, $N = 10$.

Equus germanicus	Pleistocene - Sandalja II		ROBUST MAXIMUM LIKELIHOOD (Forsten 1990)
			Mean Abs. Dev.: 0.309
Total rates	714		
NONZERO RATES	666		
Min. log int.	2.395		
Max. log int.	3.735		
MEDIAN LOG INT.	3.392		
Min. log rate	-5.081		
Max. log rate	-2.315		
MEDIAN LOG RATE	-3.686		
POSITIVE rates	274		
Med. log int.	3.392		
Med. log rate	-3.716		
NEGATIVE rates	392		
Med. log int.	3.414		
Med. log rate	-3.676		

	Minimum (0.025)	Median (0.500)	Maximum (0.975)
SLOPE CI*		-0.797	
INTERCEPT CI*		-1.015	
PRED. RATE [h_0] CI*		0.097	
TEST POINT (Log int. = ---) CI*			

*Confidence Intervals bootstrapped with (N = --- B = ---)

Figure 6.2. LR-LI graph for measurements of 34 dental characteristics in seven samples of late Pleistocene *Equus germanicus* from Sandalja II cave in Istria (Yugoslavia). Solid circles are individual rates. Samples range in age from about 27.140 to 10.830 ky BP, and rates are calculated over intervals ranging from 248 to 5,432 generations. Slope of rate distribution is −0.797. Intercept is −1.015, yielding rate h_0 = 0.097 standard deviation per generation on a time scale of one generation. Bootstrapped confidence intervals on the slope and intercept are shown in Figure 6.3. Vertical scale bar represents $N = 10$ on inset histogram of residuals (mean absolute deviation is 0.309).

Equus germanicus — LR-LI graph (Figure 6.3)

Axis labels: $\log_{10}$ Evolutionary Rate (haldanes : std.dev./gen.) ; $\log_{10}$ Time Interval (generations). Fields labelled Process, Random, Stasis. Inset histogram of Residuals, $N = 10$.

Equus germanicus	Pleistocene - Sandalja II		ROBUST MAXIMUM LIKELIHOOD (Forsten 1990)
			Mean Abs. Dev.: 0.309
Total rates	714		
NONZERO RATES	666		
Min. log int.	2.395		
Max. log int.	3.735		
MEDIAN LOG INT.	3.392		
Min. log rate	-5.081		
Max. log rate	-2.315		
MEDIAN LOG RATE	-3.686		
POSITIVE rates	274		
Med. log int.	3.392		
Med. log rate	-3.716		
NEGATIVE rates	392		
Med. log int.	3.414		
Med. log rate	-3.676		

	Minimum (0.025)	Median (0.500)	Maximum (0.975)
SLOPE CI*	-0.731	-0.799	-0.894
INTERCEPT CI*	-1.238	-1.015 / -1.007	-0.694
PRED. RATE [h_0] CI*	0.058	0.097 / 0.098	0.202
TEST POINT (Log int. = 3.392) CI*	-3.759	-3.717	-3.681

*Confidence Intervals bootstrapped with (N = 666 B = 1000)

Figure 6.3. LR-LI graph for measurements of 34 dental characteristics in seven samples of late Pleistocene *Equus germanicus* from Sandalja II cave in Istria (Yugoslavia). Same information as in Figure 6.2, with addition of bootstrapped 95% confidence intervals for slope, intercept, and predicted intrinsic rate h_0 based on 1,000 resamplings of original 666 nonzero rates. Confidence intervals on slopes range from −0.731 to −0.894. This range of slopes excludes pure process change (0.0) and pure stasis (−1.0), and it does not conform to expectation for purely random change either (−0.5). Confidence interval on intrinsic rate h_0 ranges from 0.058 to 0.202 standard deviation per generation on a time scale of one generation.

pure process (slope 0.0) and pure stasis (slope −1.0), but does not conform to expectation for purely random change (slope −0.5) either. The best interpretation is probably a combination of real process change and stasis, with the latter predominating.

The RML intercept for the Sandalja II *Equus* is −1.015, yielding a predicted intrinsic rate h_0 = 0.097 standard deviation per generation at a scale of one generation. The 95% confidence interval on h_0 ranges from 0.058 to 0.202.

The intrinsic rate of evolution predicted for Sandalja II *Equus* (0.01) is significantly greater than that calculated for Strait Canyon *Microtus* (0.003), and it is about half the median intrinsic rate characteristic of artificial selection experiments, as discussed later.

Hagerman *Cosomys*

The microtine rodent *Cosomys primus* is common in Blancan Pliocene faunas found near Hagerman, in Idaho. D. Lich (1990) described 10 samples that came from flat-lying sediments at elevations ranging from 3,000 to 3,295 feet above sea level. Radiometric calibration indicates an average rate of sediment accumulation of about 1.8 feet per 1 ky (1.0 foot = 556 years). Successive samples differ by as few as 3.890 ky, and the entire sequence spans about 164.000 ky. *C. primus* is only a little larger than the meadow vole *M. pennsylvanicus*, and it, too, probably produced about three generations per year.

Rates of change of the length and width of the lower first molar (M_1) in Idaho *Cosomys* are shown in the LR-LI graph in Figure 6.4. One hundred thirty-five rates can be calculated for intervals of time ranging from about 11,670 generations to about 492,000 generations. One hundred twenty-nine of the rates are nonzero. Most rates are positive (71), indicating change toward larger size, but there is a substantial number of negative rates (58) as well. The median rate is 0.00002 standard deviation per generation ($10^{-5.68}$), on a median scale of 199,000 generations ($10^{5.299}$). The median rate is substantially higher than that described earlier for *Microtus*, and the median interval is substantially longer as well.

All of the *Cosomys* rates taken together have an RML slope of −0.841, which falls between stasis (−1.0) and random (−0.5), but is well below the slope expected for a real process (0.0). The bootstrapped 95% confidence interval on this slope ranges from −0.630 to −1.196 (Figure 6.4). This interval includes −1.0, but not −0.5, which means

Log₁₀ Evolutionary Rate (haldanes : std.dev./gen.)

Log₁₀ Time Interval (generations)

Residuals

Cosomys primus

Process
Random
Stasis

N = 10

Cosomys primus — Pliocene - Hagerman — (Lich 1990)

		ROBUST MAXIMUM LIKELIHOOD Mean Abs. Dev.: 0.289		
			Minimum (0.025)	Median (0.500)	Maximum (0.975)

Total rates	135			
NONZERO RATES	129			
Min. log int.	4.067	SLOPE CI*	−0.630	−0.841	−1.196
Max. log int.	5.692	INTERCEPT CI*	−2.268	−1.174	−1.175
MEDIAN LOG INT.	5.299	PRED. RATE [h₀] CI*	0.005	0.067	4.866
Min. log rate	−6.745	TEST POINT (Log int. = 5.299)			
Max. log rate	−4.514	CI*	−5.739	−5.633	−5.549
MEDIAN LOG RATE	−5.648				
POSITIVE rates	71				
Med. log int.	5.413				
Med. log rate	−5.695				
NEGATIVE rates	58				
Med. log int.	5.133				
Med. log rate	−5.587				

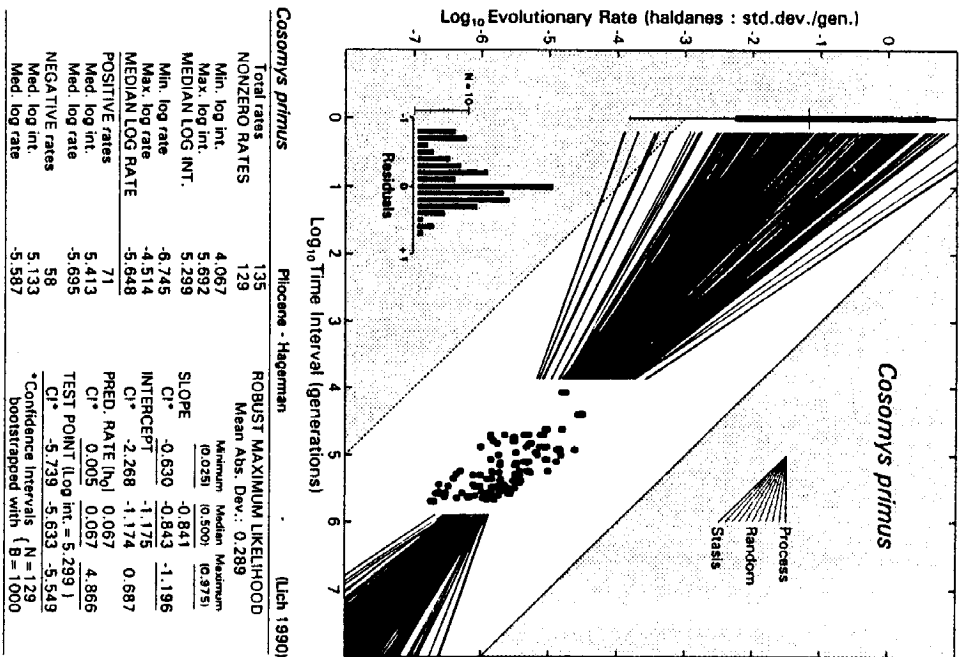*Confidence intervals bootstrapped with { N = 129, B = 1000 }

Figure 6.4. LR-LI graph for length and width measurements of the lower first molar (M_1) in 10 samples of late Pleistocene *Cosomys primus* from Idaho. Solid circles are individual rates. Bootstrapped 95% confidence intervals for slope, intercept, and predicted intrinsic rate h_0 are based on 1,000 resamplings of original 129 nonzero rates. Confidence intervals on slopes range from −0.630 to −1.196, which includes stasis (−1.0) but excludes pure process change (0.0) and random change (−0.5). Predicted intrinsic rate h_0 is 0.067, and the 95% confidence interval on this value ranges from 0.005 to 4.866 standard deviations per generation on a time scale of one generation. Vertical scale bar represents the 95% confidence interval on a time scale of one generation. N = 10 on inset histogram of residuals (mean absolute deviation is 0.289).

that the observed slope of −0.841 cannot be distinguished from that expected of stasis, while both random change and pure process can be ruled out.

The RML intercept estimated for *Cosomys* is −1.175, which yields a predicted intrinsic rate $h_0 = 0.067$ ($10^{-1.175}$) standard deviations per generation at a scale of one generation. This is close to the intrinsic rate calculated for Sandalja II *Equus*, and it is significantly greater than the intrinsic rate of $h_0 = 0.003$ calculated for Strait Canyon *Microtus*. The bootstrapped 95% confidence interval on the *Cosomys* estimate ranges from 0.005 to 4.866 standard deviations per generation. This analysis corroborates Lich's principal conclusion that the *Cosomys* lineage exhibits stasis (Lich, 1990). However, the predicted value $h_0 = 0.067$ indicates that even while in stasis its intrinsic rate of evolution remained relatively high.

North American Mammut

The American mastodon or mastodont made its first appearance in North America about 3.5 my BP and survived until about 10,000 years ago. King and Saunders (1984) described samples of mastodon upper and lower molars from three sites in western Missouri dated at 45.4, 31.4, and 13.3 ky BP. Pleistocene mastodons probably had an average generation time of about 25 years (D. C. Fisher, personal communication).

Rates of change in the lengths and widths of upper and lower second molars (M² and M₂) in Missouri *Mammut* are shown in Figure 6.5. Twelve nonzero rates can be calculated for intervals of time ranging from about 560 generations to about 1,290 generations. All are positive, reflecting a slight increase in molar size over time. The median rate is 0.0007 standard deviation per generation ($10^{-3.145}$), on a median scale of 637 generations ($10^{2.804}$). The median rate is substantially higher than that described earlier for *Microtus*, but here the median interval is shorter.

All of the *Mammut* rates taken together have an RML slope of −0.004, which is almost exactly zero. The RML intercept estimated for *Mammut* is −3.133. This yields a predicted intrinsic rate $h_0 = 0.001$ ($10^{-3.133}$), which is at (or below) the lower limit to be expected for intrinsic rates. This example suggests that change in *Mammut* was due to a real process carried out over many generations and operating at a very low rate. However, both the number of rates and the interval range they
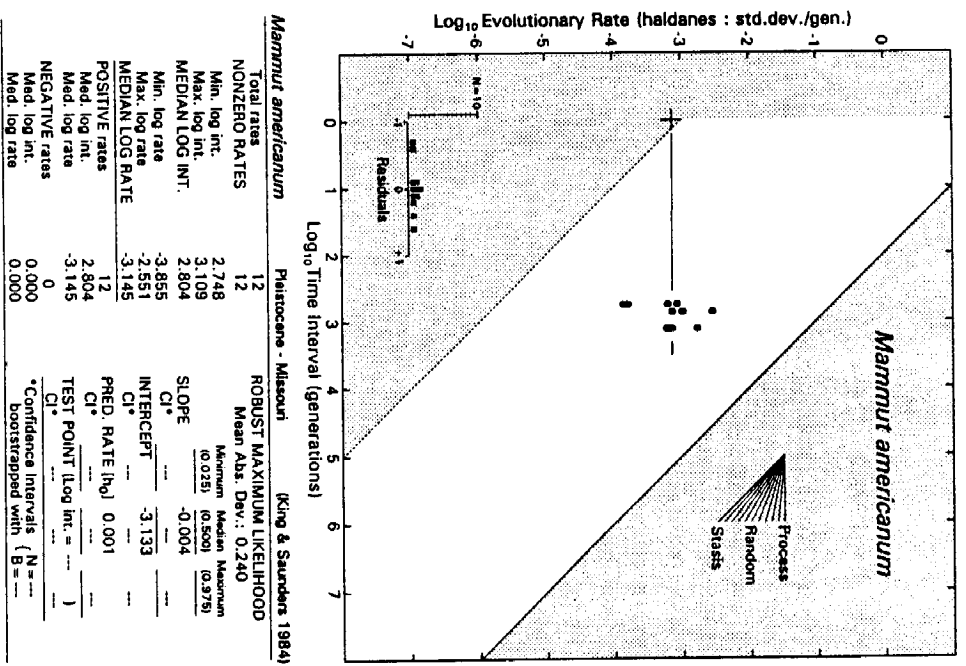
Mammut americanum	Pleistocene - Missouri	ROBUST MAXIMUM LIKELIHOOD (King & Saunders 1984)		
		Mean Abs. Dev.: 0.240		
			Minimum (0.025)	Median (0.500)	Maximum (0.975)
Total rates	12			
NONZERO RATES	12			
Min. log int.	2.748	SLOPE	−0.004	
Max. log int.	3.109	CI*	---	
MEDIAN LOG INT.	2.804	INTERCEPT	−3.133	
Min. log rate	−3.855	CI*	---	
Max. log rate	−2.551	PRED. RATE (h₀)	0.001	
MEDIAN LOG RATE	−3.145	CI*	---	
POSITIVE rates	12	TEST POINT (Log int. = ---)		
Med. log int.	2.804	CI*	---	
Med. log rate	−3.145			
NEGATIVE rates	0	*Confidence intervals (N = ---		
Med. log int.	0.000	bootstrapped with (B = ---)		
Med. log rate	0.000			

Figure 6.5. LR-LI graph for measurements of upper and lower molar (M² and M₂) lengths and widths in three samples of late Pleistocene *Mammut americanum* from western Missouri [sample statistics given by King and Saunders (1984)]. Solid circles are individual rates. Quantification of change yields 12 rates, all positive, ranging from 0.0001 to 0.003 haldane on a time scale of about 637 generations. This can be explained by process evolution sustained for the interval order of 0.001 standard deviation per generation at rates on the sampled, shown with the solid horizontal line. However, this sample of rates is too small and too closely spaced to rule out other interpretations, including random change or stasis at higher intrinsic rates.

span are small: These are not adequate to rule out random change or even stasis.

North American Bison

Bison entered North America from Eurasia in the early Pleistocene, and two species, *Bison latifrons* and *B. antiquus*, were present though much of the Pleistocene (McDonald, 1981). The transition from the ancestral species *B. antiquus antiquus* to modern *B. bison bison* took place rapidly between about 5,000 and 4,000 ky BP. J. McDonald (1981) measured skulls and limbs in large samples of ancestral *B. a. antiquus* and descendant *B. b. bison*, which evidently changed little before and after this speciation event. According to McDonald, *B. an- tiquus occidentalis* is a possible intermediate present from about 10,000 to about 4,000 ky BP, but I have used measurements for *B. a. antiquus* (McDonald's Tables 21 and 22) and *B. b. bison* (McDonald's Tables 29 and 30), together with the minimum 1,000-year duration of the transition, to maximize calculated rates. The generation time of Holocene *Bison* probably averaged about 3 years.

Rates of change in skulls and limbs of North American *Bison* are shown in the LR-LI graph in Figure 6.6. Sixty-six rates can be calculated for the 333-generation interval. All of these are negative, indicating decreases in size of all measured characteristics. Rates range from about 0.001 to about 0.022, with a median rate of 0.009 standard deviation per generation on a time scale of 333 generations. Multiplying 0.009 by 333 indicates that the species *B. a. antiquus* and *B. b. bison* differed in size, on average, by about 3.0 standard deviations.

McDonald (1981, p. 52) regarded speciation in *Bison* as conforming to the "punctuated equilibrium" concept of Eldredge and Gould (1972), or as one step in a presumed staircase pattern. Quantification of the best-studied and most rapid "punctuation" in *Bison* evolution, that be- tween *B. a. antiquus* and *B. b. bison*, suggests that staircase evolution involves intrinsic rates ranging from as low as 0.001 to as high as 0.022 haldane, sustained over intervals ranging from 1 to about 333 generations (horizontal dashed line in Figure 6.6). The intrinsic rate may have been much higher if this distribution represents a lineage in stasis (oblique dashed line in Figure 6.6).

Isle of Jersey Cervus

The red deer *Cervus elaphus* has a 400,000-year history, and it is still widely distributed in Europe today. During the last interglacial, some
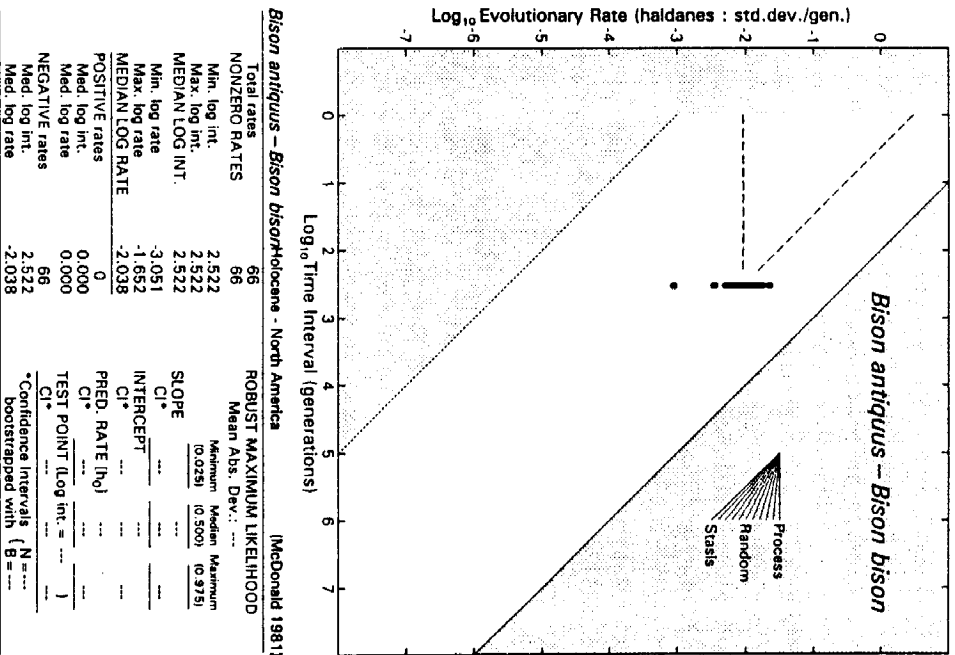
Bison antiquus – Bison bison/Holocene - North America		ROBUST MAXIMUM LIKELIHOOD		(McDonald 1981)
		Mean Abs. Dev.		
		Minimum (0.025)	Median (0.500)	Maximum (0.975)
NONZERO RATES	66			
Total rates	66			
Min. log int.	2.522			
Max. log int.	2.522			
MEDIAN LOG INT.	2.522	SLOPE CI*	---	---	---
Min. log rate	-3.051			
Max. log rate	-1.652			
MEDIAN LOG RATE	-2.038	INTERCEPT CI*	---	---	---
POSITIVE rates	0			
Med. log int.	0.000	PRED. RATE [b_0] CI*	---	---	---
Med. log rate	0.000			
NEGATIVE rates	66	TEST POINT (Log int. = ---)		
Med. log int.	2.522	CI*	---	---	---
Med. log rate	-2.038			

*Confidence Intervals   {N = ---}
bootstrapped with   {B = ---}

Figure 6.6. LR-LI graph for measurements of 66 dental, cranial, and postcranial characteristics in two samples of Pleistocene-Holocene *Bison antiquus antiquus* and Holocene *Bison bison bison* drawn from many sites in North America [sample statistics given by McDonald (1981)]. Solid circles are individual rates. Transition from one species to the other took place between 5,000 and 4,000 ky BP. Quantification of change across this 1,000-year interval yields 66 rates, all negative, ranging from 0.001 to 0.022 haldane on a time scale of about 333 generations. McDonald regards this transition as a "punctuation" event. It can be explained by process evolution sustained for 333 generations, shown with the hori- zontal dashed line. Samples of rates measured over a single interval are insuf- ficient to permit calculation of slopes or projected intercepts, and this rate distribution could possibly represent stasis at a much higher intrinsic rate, as shown with the oblique dashed line.

120.000 ky BP, the rising sea level isolated a population of *C. elaphus* on the English Channel island of Jersey, and this population dwarfed rapidly. A. Lister (1989) described the geological context, showing that the maximum duration of isolation lasted 5.800 ky. Half this duration, or 2.900 ky, is probably a good estimate of the time necessary to produce this dwarfed form. Lister measured *C. elaphus* teeth and limb bones from Belle Hougue Cave on Jersey and from contemporary sites in England. Ten measurements were made on samples large enough to quantify variability. These suggest a reduction in mean body mass of male *C. elaphus* from about 200 kg to about 36 kg and permit calculation of rates of dwarfing. The generation time of mainland *Cervus* probably was about 3 years, which probably decreased as dwarfing proceeded. An average generation time of 2.5 years has been used here in calculating evolutionary rates.

Rates of change in teeth and limbs of Jersey *Cervus* are shown in the LR-LI graph in Figure 6.7. Ten rates can be calculated for the 1,160-generation interval. All of these are negative, indicating decreases in size of all measured characteristics. Rates range from about 0.003 to about 0.017, with a median rate of 0.007 standard deviation per generation on a time scale of 1,160 generations. Multiplying 0.007 by 1,160 indicates that ancestral and dwarfed *C. elaphus* differed in size, on average, by about 8.0 standard deviations, which is much greater than the average difference between species of *Bison*.

Lister (1989) regarded dwarfing *Cervus* as an example of rapid evolution in an allopatric isolate. Quantification indicates that "rapid" evolution in *Cervus* probably involved intrinsic evolutionary rates ranging from as low as 0.003 to as high as 0.017 haldane, sustained over intervals ranging from 1 to about 1,160 generations (horizontal dashed line in Figure 6.7). The intrinsic rate may have been much higher if this sample represents a lineage in stasis (oblique dashed line in Figure 6.7).

Discussion

Six case studies are too few to justify broad conclusions concerning evolutionary tempo and mode, especially when each of the individual studies has as many assumptions underlying temporal calibration as these do. However, each case analyzed here is an independent empirical study of population variation in samples of fossils in known stratigraphic context carried out with great effort and published to further understanding
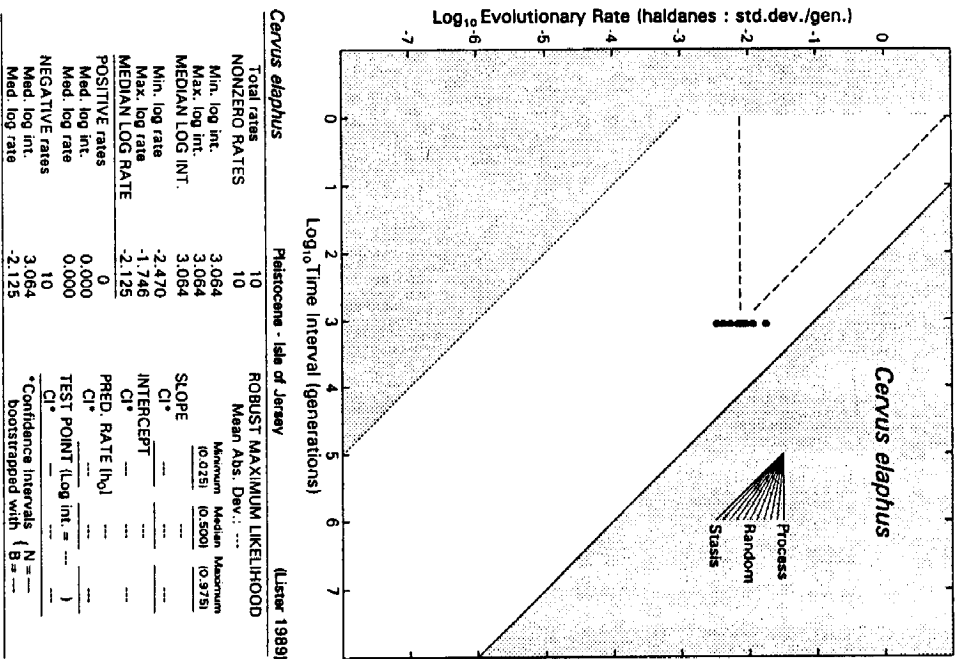
Cervus elaphus	Pleistocene - Isle of Jersey				(Lister 1989)
			ROBUST MAXIMUM LIKELIHOOD		
			Mean Abs. Dev:		
			Minimum (0.025)	Median (0.500)	Maximum (0.975)
Total rates	10				
NONZERO RATES	10				
Min. log int.	3.064	SLOPE	CI*	---	---
Max. log int.	3.064				
MEDIAN LOG INT.	3.064	INTERCEPT	CI*	---	---
MEDIAN LOG RATE	-2.125				
Min. log rate	-2.470	PRED. RATE [h₀]	CI*	---	---
Max. log rate	-1.746				
POSITIVE rates	0	TEST POINT (Log int.)	CI*	---	---
Med. log int.	0.000	CI*	---		)
Med. log rate	0.000				
NEGATIVE rates	10	*Confidence intervals	N = ---		
Med. log int.	3.064	bootstrapped with	B = ---		
Med. log rate	-2.125				

Figure 6.7. LR-LI graph for measurements of 10 dental and postcranial characteristics in two samples of late Pleistocene *Cervus elaphus*, one a British last-interglacial sample representing an ancestral stock, and the other a dwarfed descendant stock isolated on the Isle of Jersey in the English Channel [sample statistics given by Lister (1989)]. Solid circles are individual rates. Transition from one species to the other took place in a 2,900-year interval about 120.000 ky BP. Quantification of change across this interval yields 10 rates, all negative, ranging from 0.003 to 0.017 haldane on a time scale of about 1,160 generations. Lister regards this as an example of "rapid evolution in an allopatric isolate." It can be explained by process evolution at rates on the order of 0.003 to 0.017 standard deviation per generation sustained for 1,160 generations, shown with the horizontal dashed line. Samples of rates measured over a single interval are insufficient to permit calculation of slopes or projected intercepts, and this rate distribution could possibly represent stasis at a much higher intrinsic rate, as shown with the oblique dashed line.

of evolution. It would be less than scientific to ignore the empirical record, and quantitative comparison of the six case studies does provide some insight.

The intrinsic rate of change (h_0) in Strait Canyon *Microtus* is estimated at about 0.003 haldane, and this is interpreted as random change because the slope of the empirical LR–;I distribution is close to −0.5. The intrinsic rate of change in Sandalja II *Equus* is estimated at about 0.098 haldane, with a 95% confidence interval ranging from 0.058 to 0.202, and this is interpreted as a combination of stasis and process change because the slope of the empirical LR-LI distribution is −0.80, differing significantly from those expected of pure stasis, purely random change, and pure process change. The intrinsic rate of change in Hagerman *Cosomys* is estimated at about 0.067 haldane, and this is interpreted as evolutionary stasis because the slope of the empirical LR-LI distribution differs significantly from those expected of purely random change and pure process change, but not from that expected of stasis. The intrinsic rate of evolution in North American *Mammut* may have been as low as about 0.001 haldane. The intrinsic rate of change during speciation in North American *Bison* is estimated to have been between 0.001 and 0.022 (median 0.009) haldane, and that during dwarfing in Jersey *Cervus* is estimated to have ber   etween 0.003 and 0.017 (median 0.007) haldane.

The maximum rates calculated for *Bison* and *Cervus* undergoing rapid "punctuation" change, 0.022 and 0.017, respectively, lie at the lower end of the 95% confidence interval of intrinsic rates calculated for *Cosomys* in stasis, and these lie well below the lower end of the 95% confidence interval of intrinsic rates calculated for *Equus*. This suggests that speciation in *Bison* and dwarfing in *Cervus* may have taken place at rates higher than those calculated here, which may mean that the observed changes took place in less time than either McDonald (1981) or Lister (1989) considered in writing about them, but higher rates involve higher selection intensities (Gingerich, 1992), and species cannot survive when too small a proportion of individuals are selected in each generation.

Observation that rates of change during "punctuation" are no higher than intrinsic rates associated with stasis may also help to clarify the difference between "punctuation" and "stasis" in the punctuated equilibrium theory of Eldredge and Gould (1972). Eldredge and Gould's title proclaimed punctuated equilibrium as an alternative to phyletic

gradualism, but sexual reproduction takes place generation by generation, step by step, which is why rates are appropriately calculated on a generational time scale. There is no alternative to gradualism because there is no alternative to generation-by-generation inheritance.

Process change ("punctuated" or not) and stasis differ in net rate but probably not in intrinsic rate, and thus the punctuation-or-stasis fate of lineages cannot, as Eldredge and Gould argued, be intrinsically determined (because what is intrinsic is observed not to be different). Further, when rates are examined in the appropriate quantitative LR-LI context, process rates (including punctuation in *Bison* and *Cervus*) and stasis rates (in *Cosomys*) are seen as the two extremes of an empirically fan-shaped spectrum of possible net rate distributions all rooted in the same narrow range of intrinsic rates (symbolized by the inset fan in the upper right quadrant of the LR-LI graphs). This fan-shaped spectrum includes random change in *Microtus* and a similar complex combination of process and stasis in *Equus*. Evolution and evolutionary change are manifestly more complicated than a simplistic punctuation-stasis dichotomy would indicate, and their full complexity can be characterized only by quantification.

The estimates of intrinsic rate (h_0) calculated here, spanning the two orders of magnitude from about 0.001 to about 0.100 haldane, appear to lie within the range of intrinsic rates a geneticist might suggest for change measured in phenotypic standard deviations per generation. My own analyses of intrinsic rates observed in artificial selection experiments indicate that these range from about 0.010 to 1.000 haldane, with a median h_0 of about 0.200 haldane. Economy and simplicity of design dictate that artificial selection experiments run at maximum rate, and it is reasonable that the range of rates calculated for the natural cases analyzed here is consistently about one order of magnitude less than the range observed experimentally.

In conclusion, stratigraphic sequences of fossils contain much information about the tempo of evolutionary change that can be extracted only by quantitative study of population variation and numerical calibration. When population variation and age are known to reasonable approximation, it is possible to characterize and compare changes in terms of rates. The six case studies reviewed here are exemplary in this regard and indicate that it is possible to relate change on a geological scale of time to change on a generational scale.

References

Barnosky, A. D. (1990). Evolution of dental traits since latest Pleistocene in meadow voles (*Microtus pennsylvanicus*) from Virginia. *Paleobiology*, 16:370–83.

Efron, B., and R. Tibshirani (1986). Bootstrap methods for standard errors, confidence intervals, and other measures of statistical accuracy. *Statistical Science*, 1:54–77.

Eisenberg, J. E. (1981). *The Mammalian Radiations: An Analysis of Trends in Evolution, Adaptation, and Behavior*. Chicago: University of Chicago Press.

Eldredge, N., and S. J. Gould (1972). Punctuated equilibria: an alternative to phyletic gradualism. In T. J. M. Schoph (ed.), *Models in Paleobiology* (pp. 82–115). San Francisco: Freeman.

Forsten, A. (1990). Dental size trends in an equid sample from the Sandalja II cave of northwestern Yugoslavia. *Paläontologische Zeitschrift*, 64:153–60.

Gingerich, P. D. (1983). Rates of evolution: effects of time and temporal scaling. *Science*, 222:159–61.

——— (1992). Quantification and comparison of evolutionary rates. Unpublished manuscript.

Haldane, J. B. S. (1949). Suggestions as to quantitative measurement of rates of evolution. *Evolution*, 3:51–6.

King, J. E., and J. J. Saunders (1984). Environmental insularity and the extinction of the American mastodont. In P. S. Martin and R. G. Klein, (eds.), *Quaternary Extinctions: A Prehistoric Revolution* (pp. 315–39). Tucson: University of Arizona Press.

Kurtén, B. (1959). Rates of evolution in fossil mammals. *Cold Spring Harbor Symposia on Quantitative Biology*, 24:205–15.

Lewontin, R. C. (1966). On the measurement of relative variability. *Systematic Zoology*, 15:141–2.

Lich, D. K. (1990). *Cosomys primus*: a case for stasis. *Paleobiology*, 16:384–95.

Lister, A. M. (1989). Rapid dwarfing of red deer on Jersey in the last interglacial. *Nature*, 342:539–42.

McDonald, J. N. (1981). *North American Bison: Their Classification and Evolution*. Berkeley: University of California Press.

Simpson, G. G. (1944). *Tempo and Mode in Evolution*. New York: Columbia University Press.