

EVOLUTION

INTERNATIONAL JOURNAL OF ORGANIC EVOLUTION

PUBLISHED BY
THE SOCIETY FOR THE STUDY OF EVOLUTION

Vol. 53

December 1999

No. 6

Evolution, 53(6), 1999, pp. 1637–1653

PERSPECTIVE: THE PACE OF MODERN LIFE: MEASURING RATES OF CONTEMPORARY MICROEVOLUTION

ANDREW P. HENDRY¹ AND MICHAEL T. KINNISON^{2,3}

¹*Department of Zoology, University of British Columbia, 6270 University Boulevard, Vancouver, British Columbia V6T 1Z4, Canada*
E-mail: ahendry@zoology.ubc.ca

²*Department of Biological Sciences, Dartmouth College, Gilman Hall, Hanover, New Hampshire 03755*
E-mail: michael.kinnison@dartmouth.edu

Abstract.—We evaluate methods for measuring and specifying rates of microevolution in the wild, with particular regard to studies of contemporary, often deemed “rapid,” evolution. A considerable amount of ambiguity and inconsistency persists within the field, and we provide a number of suggestions that should improve study design, inference, and clarity of presentation. (1) Some studies measure change over time within a population (allochronic) and others measure the difference between two populations that had a common ancestor in the past (synchronic). Allochronic studies can be used to estimate rates of “evolution,” whereas synchronic studies more appropriately estimate rates of “divergence.” Rates of divergence may range from a small fraction to many times the actual evolutionary rates in the component populations. (2) Some studies measure change using individuals captured from the wild, whereas others measure differences after rearing in a common environment. The first type of study can be used to specify “phenotypic” rates and the later “genetic” rates. (3) The most commonly used evolutionary rate metric, the *darwin*, has a number of theoretical shortcomings. Studies of microevolution would benefit from specifying rates in standard deviations per generation, the *haldane*. (4) Evolutionary rates are typically specified without an indication of their precision. Readily available methods for specifying confidence intervals and statistical significance (regression, bootstrapping, randomization) should be implemented. (5) Microevolutionists should strive to accumulate time series, which can reveal temporal shifts in the rate of evolution and can be used to identify evolutionary patterns. (6) Evolutionary rates provide a convenient way to compare the tempo of evolution across studies, traits, taxa, and time scales, but such comparisons are subject to varying degrees of confidence. Comparisons across different time scales are particularly tenuous. (7) A number of multivariate rate measures exist, but considerable theoretical development is required before their utility can be determined. We encourage the continued investigation of evolutionary rates because the information they provide is relevant to a wide range of theoretical and practical issues.

Key words.—Darwins, evolutionary rates, haldanes, microevolution, population divergence, rapid evolution.

Received October 5, 1998. Accepted June 10, 1999.

How fast, as a matter of fact, do animals evolve in nature?
(Simpson 1944, p. 3).

Human beings are fascinated by extremes. Our imagination, interest, and admiration is inspired by the largest dinosaurs, the smallest computers, the most expansive migrations, and the fastest athletes. The study of evolution is not immune to this bias toward perceived extremes. For instance, a number of recent studies have reported examples of “rapid” evolution and thereby garnered attention from the scientific community and general public alike. Although these studies may make us rethink the rate at which evolution occurs,

claims of rapid evolution mean little without specifying what “rapid” actually means. Indeed, rates of evolution reported in some studies may seem glacial relative to those reported in others. The rate of evolution should therefore be quantified, but this endeavor becomes complicated owing to the varied nature of organismal traits, as well as the imperfect mathematical constructs we use as measures of evolution. Nonetheless, quantitative measures of evolutionary rate are necessary when comparing the tempo of evolution across different studies, traits, taxa, and time frames.

Qualitative rates have provided a basis for evolutionary inference ever since Darwin (1859, p. 84) proclaimed, “We see nothing of these slow changes in progress, until the hand

³ Authors contributed equally; order is alphabetical.

of time has marked the long lapse of ages." Simpson (1944) was the first to provide a detailed comparative analysis of rates of evolution across taxa in his seminal book *Tempo and Mode in Evolution*. Haldane (1949) brought further rigor to the field by developing much needed quantitative measures of evolutionary rate. The consideration of rates has been fundamental to many issues in modern evolutionary biology. In particular, one of Haldane's proposed measures (the *darwin*) has been used by numerous authors to compare rates of evolution across a wide range of taxa, particularly in the paleontological record (e.g., Kurtén 1960; Van Valen 1974) and to argue that microevolution is (or is not) compatible with macroevolution (e.g., Losos et al. 1997; Reznick et al. 1997; Svensson 1997). Estimated rates of evolution have also been compared to random-walk models to evaluate the potential contribution of selection versus random processes in phenotypic evolution (e.g., Lande 1976; Charlesworth 1984; Bookstein 1988; Lynch 1990; Gingerich 1993) and have been combined with information on potential selective factors and genetic covariance among traits to estimate selection differentials and gradients (e.g., Reznick et al. 1997). We predict that a common future application of evolutionary rates will be assessing the potential impact of anthropogenic disturbances (e.g., Singer et al. 1993; Lynch 1996; Stockwell and Weeks 1999). In particular, it is important to determine if populations or species can respond to changing selection pressures rapidly enough to forestall extinction.

The important inferences that can be drawn using evolutionary rates have stimulated a proliferation of studies reporting rates of microevolution in contemporary populations (Stearns 1992; Losos et al. 1997; Reznick et al. 1997; Svensson 1997; Magurran 1998). These studies have often applied the same evolutionary rate measures in the same way as studies of change in fossil lineages. The study of evolution in contemporary populations, however, has its own unique considerations and opportunities that may not be best addressed using traditional approaches. Furthermore, evolutionary biologists in this rapidly growing field may not be familiar with the extensive caveats associated with evolutionary rate calculations (e.g., Gingerich 1983; Fenster et al. 1992; Gingerich 1993). Our goal is to provide an appraisal of methods for quantifying rates of evolution, particularly those relevant to studies of microevolution in contemporary populations, and to make suggestions for improved inference in rate comparisons.

Contemporary Microevolution

We use the term "microevolution" when referring to changes that take place within species or populations. (Dobzhansky was the first to juxtapose the terms "microevolution" and "macroevolution," considering the former to occur "within the span of a human lifetime," and the later to "require time on a geological scale" [1937, p. 12]. Goldschmidt popularized the terms and altered them so that "The latter term will be used here for the evolution of the good species and all the higher taxonomic categories" [1940, p. 8]. Mayr argued that "There is only a difference of degree, not one of kind, between the two classes of phenomena" [1942, p. 291].)

Time frames for microevolutionary change can vary from one generation to many thousands of years, but we have chosen to focus on microevolution occurring in recent times and on short time scales (less than a few centuries). This focus was chosen because studies reporting "rapid" evolution over such time frames are becoming increasingly frequent. We refer to this type of microevolution as "contemporary." Examples include research on Hawaiian mosquitofish (Stearns 1983a), German blackcaps (Berthold et al. 1992), Galápagos finches (Grant and Grant 1995), Trinidadian guppies (Reznick et al. 1997), and Bahamian lizards (Losos et al. 1997). As the foregoing list reveals, most of our examples will be derived from the vertebrate literature, primarily because our experience lies within that arena. Nonetheless, much of our discussion is equally applicable to studies quantifying microevolution in other taxa (e.g., marine invertebrates: Seeley 1986; insects: Carroll et al. 1997; plants: Snaydon and Davies 1972).

The traits we consider typically have a continuous, quantitative genetic basis. A good introduction to estimating rates of molecular evolution can be found in Li (1997), and empirical examples linking molecular and adaptive evolution appear in Givnish and Sytsma (1997). Although our review is aimed at the study of microevolution in the wild, most of the discussion is also relevant to studies undertaken in the laboratory and the fossil record. Readers seeking additional insights directed specifically at quantifying evolutionary rates on paleontological scales can also turn to the work of others (e.g., Charlesworth 1984; Lynch 1990; Fenster et al. 1992; Gingerich 1993).

Estimating a rate of evolution is mechanistically independent of whether that change is caused by selection, phenotypic plasticity, genetic drift, mutation, or gene flow, simply because most rate estimates only require a knowledge of the magnitude of change and the time frame. Thus, evolutionary rates can be calculated and compared without any knowledge of heritabilities, genetic variances and covariances, or population sizes. Such knowledge, however, is instrumental in predicting evolution and in discriminating among different causal mechanisms.

We first consider experimental design and inference in the study of contemporary microevolution. Next, we describe the common measures for quantifying rates of evolution and evaluate their assumptions and utility. We then discuss statistical improvements in the quantification of evolutionary rates, and close with a reconsideration of what constitutes "rapid" evolution.

STUDY DESIGN AND INFERENCE

... the expression "rate of evolution" without further qualification, is extremely ambiguous (Simpson 1953, p. 3).

"Evolution" or "Divergence"?

Microevolutionary studies can be divided into two types, those comparing trait values for the same population at different points in time (referred to hereafter as "allochronic") and those comparing extant populations that had a common

origin at some time in the past ("synchronic"). We use these terms in preference to the analogous "directional" and "non-directional" of Harvey and Purvis (1991) because nondirectional carries little information and can imply a lack of evolution. Allochronic designs also appear in studies of selection, where they are referred to as "longitudinal" (Lande and Arnold 1983). Allochronic study designs are appropriate for inferring rates of "evolution," whereas synchronic study designs allow inference regarding rates of "divergence."

An allochronic design measures the same population at different points in time, such as before and after exposure to environmental change. For almost 25 years, the Grants and their colleagues have monitored body size and beak shape in Darwin's finches (particularly *Geospiza fortis*) in the Galápagos Islands (Grant and Grant 1995). An intense drought between 1976 and 1977 reduced the abundance of small, soft seeds. As a result, selection favored large finches with large beaks, leading to an increase in the average size of these traits (Boag and Grant 1981). In 1983, an El Niño event resulted in 10 times as much rain as the previously recorded maximum, dramatically increasing the abundance of small, soft seeds. Smaller birds with smaller beaks had higher fitness (Gibbs and Grant 1987), and body and beak sizes decreased rapidly to near the pre-1976 levels (Grant and Grant 1995). Because of the allochronic design, rates of evolution for body size and beak shape can be directly estimated from their data (Table 1).

Synchronic designs, where two or more populations of common ancestry are compared at the same time, do not provide direct measures of the rate of evolution. Instead, they measure the rate of divergence among populations. Naturally, such divergence occurs via evolution, but divergence integrates potentially different evolutionary trajectories. As a result, the quantified rate of divergence may represent anywhere from a small fraction of the actual rate of evolution to many times the magnitude of any of the contributing evolutionary trajectories (Fig. 1).

Studies of divergence often consider character states in populations that arose after introduction to a new geographical location. Stearns (1983a,b) compared life-history traits among populations of mosquitofish (*Gambusia affinis*) established by 150 fish introduced to Hawaii in 1905. His study revealed significant genetic differences among the new populations in several traits (Stearns 1983b). Without knowing the character state of the ancestral fish, however, it cannot be ruled out that the Hawaiian populations evolved along a similar trajectory away from the ancestral state, in which case true evolutionary rates would be higher than observed divergence rates. Examples of studies focusing specifically on divergence include those on chinook salmon (*Oncorhynchus tshawytscha*) introduced to New Zealand (Kinnison et al. 1998a,b,c), and sockeye salmon (*Oncorhynchus nerka*) introduced to Lake Washington (Hendry and Quinn 1997; Hendry et al. 1998).

Concordance between experimental design (synchronic or allochronic) and inference (divergence or evolution) is not always so transparent. Despite first appearances, comparing a derived population to contemporary representatives of their ancestral population does not represent a true allochronic study. Any inference about the rate of evolution would be

reliant on evidence (or the assumption) that the trait has not changed in the ancestral group since the colonizing event. Stability of the ancestral lineage can sometimes be verified by monitoring a "control" population (e.g., Endler 1980; Reznick et al. 1990, 1997). For the purposes of clarity, we suggest that investigators specify whether they used an allochronic or synchronic design and that inferences about rates of evolution or divergence should be made as appropriate. We furnish these distinctions for selected studies of contemporary microevolution in Table 1.

Phenotypic or Genetic?

Most studies measure phenotypic change without documenting the genetic and environmental components of that change. Some authors explicitly acknowledge this ambiguity, whereas others simply assume the change is entirely genetic. For example, Johnston and Selander (1964, p. 541) optimistically asserted "we are safe in assuming that the geographically variable characters of color, pattern, size, and body proportions are in fact genetically controlled." The other main study type employs a common rearing environment (laboratory rearing or reciprocal transplants), increasing the likelihood that only the genetic component of change is measured.

Losos et al. (1997) reported rapid divergence among populations of the lizard *Anolis sagrei*, following their experimental introduction to small islands in 1977 and 1981. In 1991, Losos and colleagues captured lizards from the islands that still harbored populations and from nearby Staniel Cay, the original source for the transplants. Losos et al. (1997) showed that lizards on the different islands had diverged in a manner that was correlated with the extent to which an island's vegetation deviated from that at Staniel Cay and with the perch diameter typically used by lizards on each island. Although their study revealed rapid divergence among populations, the authors acknowledge that they have no direct indication of how much of this divergence can be attributed to genetic change versus phenotypic plasticity (Losos et al. 1997).

The alternative approach, in which genetic change is quantified, is exemplified by the work of Reznick and colleagues (Reznick et al. 1990, 1997). In 1976, guppies (*Poecilia reticulata*) were transferred from a high-predation site (*Crenicichla alta* present) in the Aripo River, Trinidad, to a low-predation site (*C. alta* absent) in the same river (Endler 1980). In 1981, Reznick made a similar transplant in the El Cedro River, Trinidad (Reznick et al. 1990). Subsequent sampling revealed that guppies in the new populations were becoming older and larger at maturity, with smaller brood sizes (Reznick et al. 1990). To measure the genetic component of change in these characters, adults were captured from the introduced and control populations and their offspring were reared in the laboratory for two generations (Reznick et al. 1997). The authors found that significant genetic divergence had occurred in the predicted direction after 11 years in the Aripo River and after 4 years and 7.5 years in the El Cedro River.

Sometimes genetic and phenotypic change are comparable in microevolutionary studies, suggesting that much of the

TABLE 1. Selected studies of microevolution for which rates of evolution or divergence were calculated. Studies are listed in approximate chronological order of their publication (see Appendix for footnotes 1–20, which provide sources of data and notes on calculations). Design indicates whether the study compared extant representatives of populations that had common origin in the past (synchronic, syn) or the same population at different points in time (allochronic, allo). Allochronic studies measure rates of evolution and synchronic studies measure rates of divergence. The time of separation for synchronic studies and the time between collections for allochronic studies are shown in years and generations (Gen). Type indicates whether the study measured genetic differences in a common environment (G) or phenotypic differences in the wild (P). The sign of the rate (+ or –) is provided for allochronic studies, and the range of divergence rates among population pairings is provided for synchronic studies.

Study	Design	Years	Gen	Type	Trait	Darwins ($\times 10^3$)	Haldanes
American house sparrows ¹	syn	111	111	P	body weight	0–1.19	—
				P	wing length	0.16–0.36	0.008–0.024
				P	bill length	0.02–0.38	0–0.010
Trinidadian guppies ²	allo	1.8	3	P	spot length (all)	+205.81	+0.582
				P	spot area	+395.88	+0.688
				P	spot number	+191.99	+0.742
				P	color diversity	+90.29	+0.267
Hawaiian mosquitofish ³	syn	70	140	G	male age	0.13–2.90	0–0.011
				G	male length	0.03–1.62	0–0.010
				G	female age	0.22–3.10	0.001–0.007
				G	female length	0.01–1.13	0–0.006
				G	offspring weight	0.41–3.07	0.001–0.007
New England snails ⁴	allo	17	17	P	spire shape (1898–1915)	–22.35	–0.319
	allo	69	69	P	spire shape (1915–1984)	–4.49	–0.064
Australian rabbits ⁵	syn	125	—	P	body weight	0.02–0.31	—
				P	ear size index	0.19–0.46	—
				P	pes length	0.01–0.07	—
				G	body weight	0.25–0.85	—
				G	ear size index	0.09–0.12	—
				G	pes length	0.04–0.27	—
Trinidadian guppies ⁶	syn	34	59.2	G	male schooling tendency	2.39	0.036
				G	female schooling tendency	2.36	0.029
Norwegian stickleback ⁷	syn	31	31	P	eye diameter	2.26	0.043
				P	spine length	2.41	0.021
Galapagos finches ⁸	allo	2	1	P	body weight (1976–1978)	+32.25	+0.709
				P	bill depth (1976–1978)	+25.92	+0.657
				P	wing length (1976–1978)	+7.24	+0.398
	allo	3	1	P	body weight (1984–1987)	–11.75	–0.376
				P	bill depth (1984–1987)	–8.79	–0.372
				P	wing length (1984–1987)	–5.37	–0.486
Hawaiian honeycreepers ⁹	allo	100	100	P	upper mandible length	–0.18	–0.007
				P	lower mandible length	–0.10	–0.003
Trinidadian guppies ¹⁰	syn	16	27.8	G	male schooling tendency	1.39	0.032
				G	female schooling tendency	0.97	0.002
Columbia River shad ¹¹	allo	56	14	P	migratory timing	—	–0.382
Columbia River sockeye ¹¹	allo	45	11	P	migratory timing	—	–0.066
Bahamian lizards ¹²	syn	10–14	10–14	P	body shape (PC1)	0–1.0	0.004–0.105
				P	body shape (PC2)	0–2.1	0.008–0.081
				P	hindlimb length	0.09–1.2	0.008–0.099
Trinidadian guppies ¹³	syn	4–11	6.9–18.1	G	male age	13.9–45.0	0.061–0.149
				G	male size	5.3–27.1	0.030–0.098
				G	female age	3.7–8.0	0.017–0.036
				G	female size	5.1–13.9	0.014–0.043
Florida soapberry bugs ¹⁴	syn	41	100	P	beak length	3.32–6.01	0.010–0.017
				G	beak length	3.96–6.88	0.015–0.035
Washington sockeye ¹⁵	syn	56	14	P	male body length	0.08–0.86	0.005–0.056
				P	male body depth	0.47–2.30	0.024–0.122
				P	female body length	0.07–1.30	0.005–0.074
Washington sockeye ¹⁶	syn	56	14	G	time to hatch	0.05–0.55	0.009–0.086
				G	time to emerge	0.07–1.01	0.017–0.253
New Zealand chinook ¹⁷	syn	84	26.2	P	male fin size (PC1)	—	0.006–0.010
				P	female fin size (PC2)	—	0.017–0.031
New Zealand chinook ¹⁸	syn	84	26.2	P	GSI	0.80–0.82	0.021–0.026
				P	egg weight	1.26–1.59	0.038–0.048
				G	time to hatch	104	0.020
New Zealand chinook ¹⁹	syn	84	26.2	G	growth rate (stanza 1)	0.79	0.027
				G	growth rate (stanza 2)	0.86	0.032
				G	growth rate (stanza 3)	0.39	0.009
Nevada mosquitofish ²⁰	syn	55	110	G	length at maturity	0.20–1.41	0.001–0.009
				G	fat content (%)	1.21–5.71	0.003–0.014
				G	egg weight	0.54–2.58	0.001–0.005

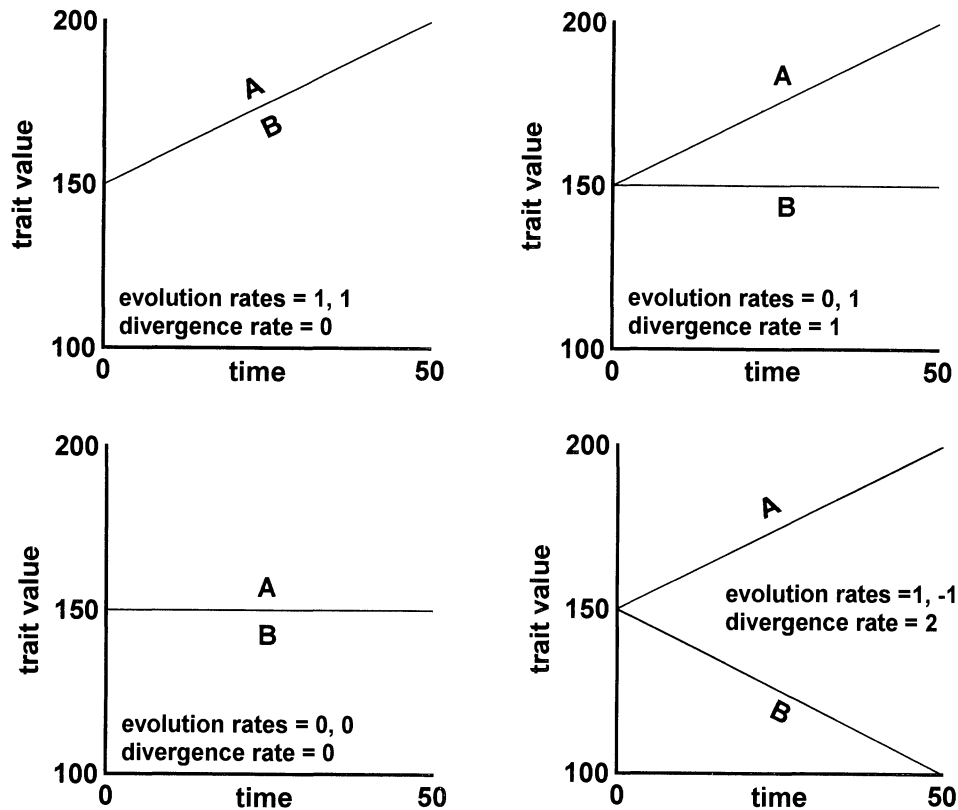


FIG. 1. Potential discrepancies between rates of evolution and rates of divergence in synchronic study designs. A and B depict hypothetical evolutionary trajectories for a trait (the units are arbitrary) in two populations that were established from a common source. If the trait values were only known for the populations at time 50 (years or generations) and not previously, an investigator might draw erroneous conclusions about rates of evolution based on the measured rate of divergence.

observed change in the wild has a genetic basis (e.g., Carroll et al. 1997). Due to genotype-by-environment interactions, however, phenotypic divergence can give a misleading picture of genetic divergence and vice versa. We suggest that rates of change should be qualified whenever stated as “genetic” or “phenotypic.” In this manner, Losos et al. (1997) reported *phenotypic* rates of divergence and Reznick et al. (1997) reported *genetic* rates of divergence (we provide these qualifiers for selected microevolutionary studies in Table 1). Specifying a rate as phenotypic does not imply that the change itself was not genetic, simply that the investigator does not know the relative contributions of genetic and non-genetic effects. Phenotypic and genetic rate estimates provide different but complimentary measures, and the information gained by investigating both is often well worth the effort (e.g., Williams and Moore 1989; Carroll et al. 1997).

The fossil record provides little indication of the genetic or environmental components of phenotypic change, and genetic rates of evolution ultimately cannot be specified for paleontological studies. Simpson (1944, p. 3) made this assumption explicit (“phenotypic evolution implies genetic change”), but it has become implicit in many works thereafter. For paleontological studies of microevolution, the assumption that any change must have a genetic basis carries less assurance. Bell et al. (1985) provided measures of morphology and body size at approximately 5000-year intervals over 110,000 years for a Miocene stickleback (*Gasterosteus*

doryssus). It seems likely that some of the observed variation in these characters had an environmental component. Unfortunately, the fossil record will forever remain mute as to the genetic basis for observed change. Thus, if one’s goal is to compare contemporary rates of microevolution to those observed in fossil strata, phenotypic rates may be most appropriate.

Extrapolation and Interpolation

Estimating a rate is mechanistically equivalent to calculating the slope a regression line through trait values over time. Thus, speculation about what might happen if short-term evolutionary rates were sustained over longer time frames is akin to extrapolating a regression line far beyond the range of the data. Additionally, conclusions about the pattern of change between two points that are separated by many generations of evolution is akin to interpolating your inference. Both extrapolation and interpolation have the potential to mislead (Fig. 2). Time series of evolution in Galápagos finches uncovered periods of rapid and reversible evolution caused by specific episodes of directional selection (Grant and Grant 1995). Extrapolation of the rate of evolution observed during a particular selection episode would have yielded incorrect predictions about the future, just as interpolation between the beginning and end values would have erroneously suggested that little change had taken place.

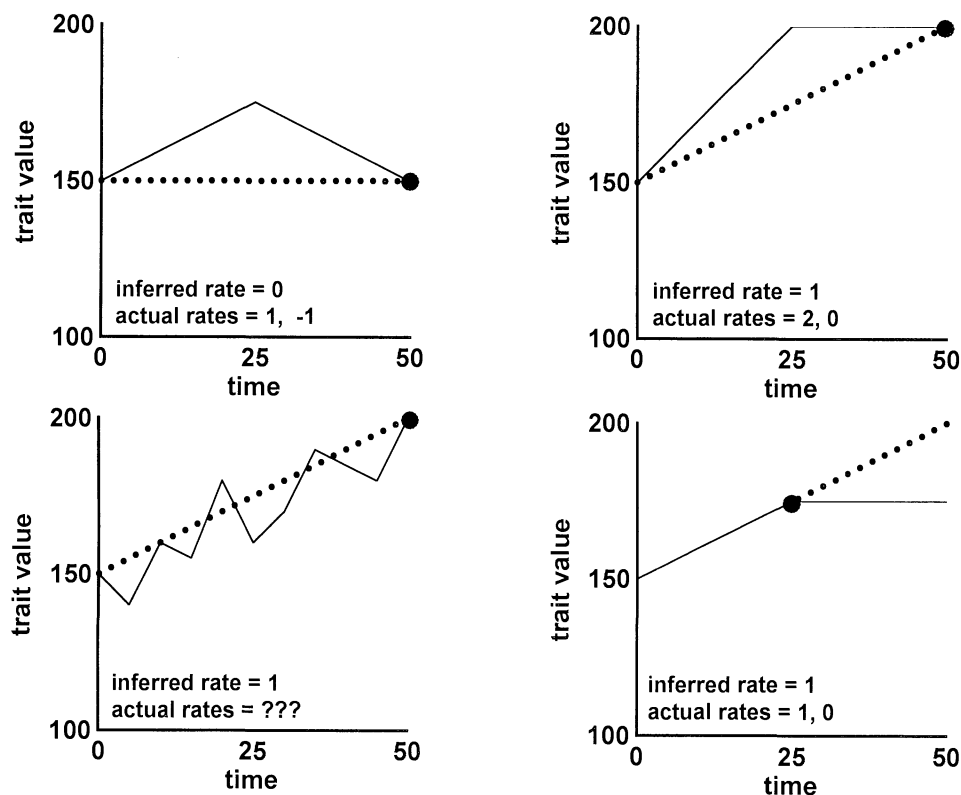


FIG. 2. Potential problems encountered when interpolating an evolutionary rate inference between two endpoints or extrapolating an inference beyond the range of the data. Samples were taken at time 0 and then at either time 25 or time 50 (indicated by the filled circles). The solid line represents the actual evolutionary trajectory of the population and the dotted line represents the inferred rate of evolution.

Informative time series include those for life-history traits in commercially and recreationally important fish species (e.g., 1859–1983 for Atlantic salmon, *Salmo salar*, Bielak and Power 1986). Quinn and Adams (1996) report migratory timing of American shad, *Alosa sapidissima*, from 1938 to 1993 and of sockeye salmon from 1949 to 1993. Using a regression through the entire dataset, the date at which 50% of the fish returned was found to have shifted earlier in the year by about 38 days for shad (0.69 days/year) and six days for sockeye (0.14 days/year). If, however, migratory timing of shad was known for only two of the years, say 1973 and 1984, the rate would appear to be 1.6 days/year in the opposite direction (Hendry and Kinnison 1998). Other random choices of any two years from the dataset would yield spectacularly varying interpretations (see Quinn and Adams 1996, figs. 4, 6). Examples of reversals and variation in evolutionary trajectories are also common in paleontological time series (e.g., Bell et al. 1985).

At the other end of the spectrum, studies of laboratory evolution under artificial selection are particularly well suited to obtaining time series and can often measure evolution over many generations (e.g., 76 in maize experiments by Dudley 1977; 85 in *Drosophila* by Yoo 1980; and 10,000 in *Escherichia coli* by Lenski and Travisano 1994). It is indisputable that time-series data are more informative than simple endpoint comparisons, and investigators should strive to accumulate such series, despite the logistical difficulties encountered when attempting to do so.

QUANTIFYING RATES OF EVOLUTION

It may be found desirable to coin some word, for example a darwin, for a unit of evolutionary rate, such as an increase or decrease of size by a factor of e per million years. . . (Haldane 1949, p. 55).

Darwins and Haldanes

Fifty years ago, Haldane (1949) discussed possible units of measure for quantifying rates of evolution. One of his suggestions has become quite popular, owing in no small part to its ease of application and perhaps also to Haldane's suggestion that it be called the *darwin*. To estimate a rate in darwins, one simply needs to take the natural logarithm (\ln) of a trait value at a particular time (or in a particular population), subtract the natural logarithm of the same trait at some time in the past (or in another population), and divide the resulting value by the length of time in millions of years. Most applications of the darwin have been within the arena for which it was originally designed: morphological traits in the fossil record. However, recently several authors have started reporting rates in darwins for studies of contemporary microevolution (Stearns 1992; Losos et al. 1997; Reznick et al. 1997; Svensson 1997; Magurran 1998; Thompson 1998), probably in large part because the darwin provided the only measure for comparison to previously published studies.

Haldane (1949) presented another potential measure for

quantifying rates of evolution, one that scales the magnitude of change by the amount of variation in the trait (note that we, like Haldane, adopt a broad definition of rate of evolution, rather than reserving the term for a specific rate metric). Lerman (1965, p. 24) further articulated this measure by expressing rates as “differences between the population means in units of standard deviation.” Gingerich (1993) formulated a similar metric and dubbed it the *haldane* (the same measure is designated D'_H by Lynch 1990). A simple formulation of the haldane is:

$$h = \frac{\left(\frac{x_2}{s_p}\right) - \left(\frac{x_1}{s_p}\right)}{g}, \quad (1)$$

where x_2 and x_1 represent mean trait values for each of two populations (sychnronic) or for a single population at two different times (allochronic), s_p is the pooled standard deviation ($[SS_1 + SS_2]/[(n_1 - 1) + (n_2 - 1)]$), and g is the number of generations separating the populations or samples (years divided by generation length). (The equation presented here differs only in form from those in Gingerich [1993] and Clyde and Gingerich [1994]. In comparing the equations, however, one should be aware that those of Gingerich [1993] and Clyde and Gingerich [1994] contain typographical errors [P. Gingerich, pers. comm.].) For many traits, especially morphological characters, standard deviations are expected to increase with the mean (i.e., coefficient of variation remains relatively constant). For such traits, raw data should be transformed to natural logarithms, which will reduce heteroscedasticity (Wright 1968).

For time series, an average rate of evolution in darwins or haldanes can be estimated by plotting $\ln x$ or x/s_p values for each collection against the number of years or generations since the first collection, and then calculating the slope of a regression line through those points. For example, we used the regression approach to calculate rates of evolution for migratory timing of American shad (-0.382 haldanes) and sockeye salmon (-0.058 haldanes) in the Columbia River (Table 1, Appendix). Statistical considerations specifically relevant to the regression approach will be discussed later.

To reduce ambiguity in the interpretation of rates specified in darwins or haldanes, Gingerich (1993) proposed a set of subscripts. We echo this sentiment and suggest an additional subscript distinguishing phenotypic from genetic studies (in the sense described above). In this manner, a rate of 0.05 haldanes for a genetic study (subscript g) over nine generations (\log_{10} generations) would be specified $h_{g(0.95)} = 0.05$. The value inside the parentheses corresponds to that advocated by Gingerich (1993), and our genetic/phenotypic distinction is added outside the parentheses. Similarly, a rate of 50 darwins for a phenotypic study (subscript p) over nine years would be $d_{p(1:0.95,6.0)} = 50$. The subscripts within the parentheses for the darwin also follow Gingerich (1993), such that the first value indicates the dimension of the measured trait (e.g., length = 1, area = 2, and volume = 3), the second indicates the number of years over which the change was measured (in \log_{10} units), and the third value indicates the number of years over which the rate is specified (\log_{10} years). The length of time over which the change was measured

(\log_{10} years or \log_{10} generations) should always be specified to two decimal places to provide enough precision for distinguishing among studies of contemporary microevolution. Note that the subscripts also serve to distinguish an evolutionary rate measured in haldanes from the square root of a heritability estimate.

We calculated rates of evolution and divergence in haldanes for selected microevolutionary studies (Table 1). We used \ln transformations in most cases because standard deviations were expected to be proportional to the mean (raw data were usually not available to examine variance structure). The mean and variance of \ln trait values were not reported for any of the studies, but the haldane has a great degree of computational flexibility that enables approximation of x and s_p for \ln raw values using nontransformed means and standard deviations. First, the mean of \ln values can be approximated as the \ln of the mean value minus half of the square of the coefficient of variation of nontransformed measurements (see Lynch 1990). Second, the s_p of the \ln measurements can be approximated by the coefficient of variation (CV) of the nontransformed measurements (Lynch 1990). We used these two Taylor approximations when calculating rates in haldanes for the studies in Table 1 (see Appendix for exceptions). Interestingly, estimated rates of change calculated using \ln means and standard deviations (Table 1) did not differ appreciably from those calculated using nontransformed means and standard deviations (results not shown). Despite the availability of such approximations, x and s_p values should be calculated from raw data (with \ln transformations when appropriate) whenever possible. When studies reported standard errors or confidence intervals, we first converted these to standard deviations using sample sizes and statistical tables.

Fundamental Differences

The darwin and the haldane differ in two fundamental ways. First, the darwin specifies the rate of proportional change in units of e , whereas the haldane specifies the rate of change in standard deviation units. Second, the time interval is measured in years for the darwin and in generations for the haldane. In principal (but not in practice), these two differences are independent of each other (a proportional change could just as easily be specified per generation or a standardized change per year). If an investigator's goal is to measure a change in some organism that is relevant to time-dependent human interests, then proportionate change and a time unit based on years may be the most desirable characteristics of a rate measure (i.e., darwins). If, however, the goal is to understand how a population responds to environmental change or to estimate the intensity of selection, then standardizing by the trait's variation and using a time scale more relevant to the organism (generations) will provide a more appropriate rate measure (i.e., haldanes).

In accounting for variation, the haldane has a better grounding in the evolutionary process, but this property adds additional complexity when comparing genetic and phenotypic studies. Environmentally-induced variation may be greater in the wild than it is under common conditions. In such cases, genetic haldanes will be higher than phenotypic

haldanes for the same absolute magnitude of change. In contrast, if individuals with extreme phenotypes are selected against in the wild, but not in the laboratory, genetic haldanes may be lower than phenotypic haldanes for the same magnitude of change.

The accuracy of an estimated evolutionary rate depends the accuracy of estimated elapsed time (years for darwins, generations for haldanes). Estimating the number of years that have passed may seem easy, but is not without ambiguity in many instances. Small errors in the estimated year of population founding, for example, can translate into measurable differences in evolutionary rates when the total time interval is relatively short. Obtaining good estimates of the time interval in generations is often more difficult because of added uncertainty in the estimation of generation lengths. As a result, rates estimated in haldanes may be less accurate than those estimated in darwins, regardless of how well each measure reflects the evolutionary process.

For the studies we reviewed (Table 1, Appendix), three techniques for estimating generation length were encountered. First, some studies estimated generation lengths based on age at maturity in the laboratory. Generation lengths calculated thus will tend to be less than those calculated based on age at maturity in the wild, which was the most common technique. In the absence of any better information (see below), we used age at maturity in the wild for our calculations (Appendix). A third technique, the life-table method (Ricklefs 1973), is more accurate than age at maturity but also requires considerably more information. Reznick et al.'s (1997) study was the only one we reviewed that estimated generation lengths using the life-table method. The different techniques can yield very different generation length estimates. For example, Endler (1980) used age-at-maturity to estimate that 15 generations had elapsed in his 1.8-year introduction experiment (8.3 generations/year), whereas Reznick et al.'s (1997) life-table calculation estimated 1.74 generations/year for the same population. The method of generation length estimation should be taken into consideration when comparing rates using haldanes.

It can be difficult to apply the haldane to fossil sequences because doing so requires estimates of phenotypic variance and generation length. However, many paleontological studies of microevolution can generate estimates of variability. For example, Bell et al. (1985) specified sample sizes ($n = 15\text{--}118$), standard deviations, and mean values for six morphological and meristic traits of Miocene stickleback collected at 26 different times. Assuming a generation length of one year, we could calculate average rates of evolution in haldanes for these characters using regression. For example, standard length of the stickleback increased at an average rate of 5.97×10^{-6} haldanes. Examples of the haldane applied to fossil sequences include Eocene horses (Gingerich 1993) and an Eocene adapid primate (Clyde and Gingerich 1994). Generation lengths for fossil organisms are typically approximated by those of their closest extant relatives. If one's confidence in generation length or variance estimates is low, rates in haldanes can be specified as a contour plot of rate at various plausible combinations of generation length and variance.

Other Data Considerations

The dimension in which a trait is measured (e.g., length, area, volume) influences rates of evolution calculated using the darwin (Gingerich 1993). This dimension dependence arises because as length increases, area increases as the square of that change, and volume as the cube of that change (assuming a cubic trait). Endler's (1980) data on guppy color spots allows us to illustrate this dimension dependence. The number, size, and color diversity of spots increased when predation pressure was reduced following an experimental introduction (Endler 1980). Our calculation of darwins for Endler's data suggested that spot area increased about twice as fast as spot length (Table 1). Rates calculated for same data using haldanes, which is not dimension dependent, suggested that spot length and spot area were evolving at a similar rate (Table 1). Rates of evolution will depend on dimension when specified in darwins, but not in haldanes.

Phenotypic data is commonly measured on ratio or interval scales. Data on a ratio scale have a constant interval between adjacent units, and the measurement scale has a precise zero point corresponding to a null quantity. Ratio scale data have the property that doubling a value doubles the actual quantity. The skull of a horse or hominid, to return to Haldane's (1949) examples, has clearly defined zero points from which length can be measured. Similarly, most of the traits shown in Table 1 are measured on a ratio scale, for which rates can be calculated using darwins or haldanes. In contrast, data on an interval scale have a constant interval between adjacent units, but the zero point is arbitrary (e.g., temperature or time of day). The darwin is not appropriate for specifying rates of evolution for interval scale traits.

Behavioral and life-history traits are often measured on an interval scale, such as migratory timing of American shad and sockeye salmon entering the Columbia River (Quinn and Adams 1996). Although rapid change was detected, the darwin cannot be used to estimate rates because the zero point for migration date is not clear. What is day zero? If it is assumed to be the first of the year, the rate of change for shad would be 3878 darwins. If day zero is assumed to be June 1, the rate of change would be 25,205 darwins (for haldanes it remains constant at 0.382 regardless of the assumed zero point). Unlike the darwin, the haldane is suitable for calculating evolutionary rates for migration timing or other traits measured on an interval scale.

Estimating rates for a particular trait assumes that the measurements accurately represent the evolving character and makes no allowances for inclusion or exclusion of additional components. Julian Huxley elaborated at length on such issues in his *Problems of Relative Growth* (1932, pp. 110–118). For example, if an evolutionary change in the beak length of a bird is not proportional along its entire length (e.g., all of the change occurs between the nares and the tip of the beak), a measurement to the anterior edge of the nares would yield a larger evolutionary rate measured in darwins than one taken to the posterior edge of the nares. This discrepancy arises because the same change in only part of the beak will appear smaller when measured from a more distant zero point. The haldane is not as strongly influenced as the darwin by

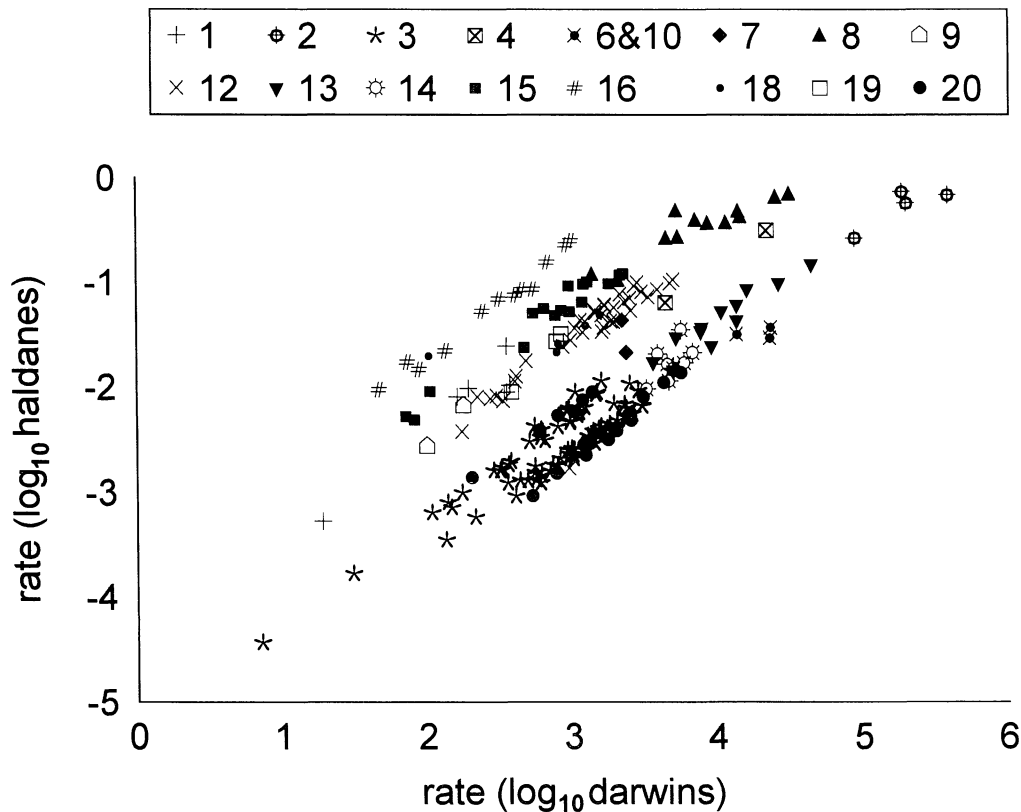


FIG. 3. Comparison of estimated rates of evolution and divergence specified in \log_{10} darwins and \log_{10} haldanes. Each point represents the rate for a single trait and population pairing from a particular study (numbers in the legend refer to the specific studies shown in Table 1 and described in the Appendix). For example, 75 points are plotted for Stearns (1983b) because we used data for 15 pairwise population comparisons (Stearns 1983b) for each of five traits (see Table 1). Negative rates (see Table 1) were expressed as positive rates.

the choice of different landmark points when part of the trait is invariant.

Both the darwin and the haldane are ill-equipped to handle rates of change in nonmetric characters, such as color morphs, presence or absence of a character state, behavioral options (e.g., Berthold et al. 1992), or alternative life histories. Whether a rate measure can be formulated that would allow reliable comparisons between diverse metric and nonmetric traits remains to be seen.

Darwins or Haldanes?

The darwin and the haldane provide different but complementary measures of evolutionary rate. Because both specify a phenotypic difference per unit time, rates calculated using each will be correlated (Fig. 3). However, some studies may have relatively low haldane rates but high darwin rates (e.g., Magurran et al. 1995; Fig. 3), whereas others may have high haldane rates but low darwin rates (e.g., Hendry and Quinn 1997; Fig. 3). The main cause of this difference is variation in estimated generation length (1.74 generations/yr vs. 0.25 generations/yr for the above contrast). Variation in the haldane/darwin relationship also occurs among traits or population pairings within a study (Fig. 3), owing principally to differences in trait variance (e.g., the data from Stearns 1983b). Due to the different conceptual basis of the two mea-

sures and to the different information conveyed by each, it is worth presenting both when possible. It is also important to acknowledge the limitations of each measure as discussed above. However, the specification of a rate, regardless of the measure used, does not give free rein to comparisons across studies, traits, taxa, and time scales (see below).

CONFIDENCE INTERVALS AND RATE COMPARISONS

... the pace of evolution is not alike in all organisms (Dobzhansky 1937, p. 37).

Confidence Intervals

Measures of evolutionary rate are rarely evaluated for statistical significance (with some exceptions for paleontological studies). Instead, studies of contemporary microevolution typically report the statistical significance of observed differences between samples (Stearns 1992; Losos et al. 1997; Reznick et al. 1997; Svensson 1997; Magurran 1998) and then present a rate based on that difference (sometimes excluding comparisons for which the difference was not statistically significant). However, this approach does not indicate how much confidence can be placed in the rate value itself. We believe that in many instances estimated rates of contemporary microevolution are imprecise, but their pre-

TABLE 2. Confidence limits (CL) determined using parametric bootstrapping for divergence rates between two populations of New Zealand chinook salmon (Kinnison et al. 1998a,b,c). Also provided are the results of randomization tests for the probability (P) that a given rate is significantly greater than zero. Superscripts provide references to the Appendix. Conclusions regarding rate significance based on the lower limit of a two-tailed 95% confidence interval do not always correspond to those obtained using randomization tests at $\alpha = 0.05$. Although random chance in both analyses can cause some variation in results, most of this effect is due to the difference between conclusions drawn from a two-tailed test (i.e., does the lower 2.5% confidence limit overlap zero) versus a one-tailed test (the chance of randomly obtaining a rate greater than that observed more than 5% of the time under the null hypothesis that both groups represent samples from the same distribution).

Trait	Year	Darwins				Haldanes			
		Lower CL	Mean	Upper CL	P	Lower CL	Mean	Upper CL	P
Male fin size (PC1) ¹⁷	1994	—	—	—	—	-0.013	0.010	0.031	0.190
	1995	—	—	—	—	-0.023	0.006	0.034	0.652
Female fin size (PC2) ¹⁷	1994	—	—	—	—	0.001	0.017	0.035	0.032
	1995	—	—	—	—	0.005	0.031	0.066	0.004
GSI ¹⁸	1994	274	815	1352	0.002	0.010	0.026	0.043	0.001
	1995	-84	797	1681	0.043	-0.002	0.021	0.046	0.042
Egg size ¹⁸	1994	672	1262	1833	<0.001	0.020	0.038	0.061	<0.001
	1995	835	1592	2418	<0.001	0.027	0.048	0.076	<0.001
Time to hatch ¹⁸	1995	-35	104	247	0.097	-0.008	0.020	0.050	0.122
Growth rate (stanza 1) ¹⁹	1995	216	790	1370	0.006	0.007	0.027	0.052	0.009
Growth rate (stanza 2) ¹⁹	1995	333	855	1392	0.002	0.013	0.032	0.053	0.009
Growth rate (stanza 3) ¹⁹	1995	-442	385	1207	0.160	-0.012	0.009	0.031	0.191

cision cannot be evaluated because confidence intervals are not presently reported.

Time series provide an easy way to calculate confidence intervals and perform significance tests because determining an average rate simply involves calculating the slope of a line through the points. We had no difficulty applying the regression approach to the data published by Bell et al. (1985) determining, for example, that the rate of evolution for stickleback body length, although very slow, was nevertheless significantly different from zero ($P = 0.002$, 95% CI = 2.40×10^{-6} – 9.55×10^{-6} haldanes). We also used regressions to calculate 95% confidence intervals for the evolutionary rate of migratory timing in American shad (between -0.271 and -0.493 haldanes) and sockeye salmon (between -0.033 and -0.099 haldanes) in the Columbia River (see Table 1, Appendix). In these two cases, confidence limits did not overlap zero, but were quite large (58% to 120% of the mean), thus reflecting significant rate variation over the interval.

In using the regression approach for time series, the potential influence of autocorrelation among error terms should be evaluated. Simple linear regression assumes that error terms are independent, which will not hold for some time series of evolutionary change (Charlesworth 1984). Violations of this assumption will not bias slope estimates, but will tend to underestimate error variances and the true standard deviation of the slope coefficient (Neter et al. 1989, pp. 484–485). Many statistical packages provide tests for autocorrelation (e.g., Durbin-Watson, DW, test), and offer time series techniques that correct for serial autocorrelation (e.g., Autoregression in SPSS vers. 7.5). Each of the rate estimations that we performed using regression were free of autocorrelation (Bell et al. 1985, standard length DW = 2.11; Quinn and Adams 1996, shad migratory timing DW = 1.98, sockeye migratory timing DW = 1.93).

For synchronic or end-point allochronic comparisons, direct formulae for determining confidence intervals and significance tests have yet to be derived. In the interim, resampling techniques such as bootstrapping and randomization

(Manly 1997) can provide an easy means of estimating confidence limits and testing whether estimated rates are significantly different from zero. We developed S-PLUS routines that perform these calculations (scripts for these routines can be obtained by contacting M. Kinnison). For example, we used resampling to estimate 95% confidence limits and performed significance tests for divergence in New Zealand chinook salmon. As expected, 95% confidence boundaries were quite broad, but rates were still significantly different from zero in many cases (Table 2).

Biological and statistical significance often do not intersect at the same level of difference and the quantification of evolutionary rates should not be limited to instances of rapid, statistically significant change. Statistical significance depends in part on sample size and biologically relevant evolution may occur even without significant differences between population means. An increase in size of 0.001% per generation, which is surely difficult to detect over a short period of time, will nonetheless lead to a 50% increase in body size if sustained for 406 generations. We suggest that evolutionary rates be estimated, along with their confidence intervals, even when statistical significance between group means is not detected. Otherwise, studies of evolutionary rate will be biased toward those that only detect rapid change.

Rate Comparisons

Evolutionary rates are often estimated with the intent of comparison to rates calculated for different traits or species or over different time intervals. Comparisons of rates have a long and distinguished history in evolutionary biology (Simpson 1944; Haldane 1949; Simpson 1953; Gingerich 1983; Stearns 1992), and claims of faster or slower evolution have at times been made without consideration of the probability associated with such inferences. Fortunately, evolutionary rates can be statistically compared using several methods. When raw data are not available, rates can be crudely compared by considering the degree of overlap in confi-

dence intervals. However, this approach is only reliable when the overlap in reported confidence intervals is substantial or if the intervals are far from overlapping. A better method involves testing whether the estimated difference between rates is greater than zero. An S-PLUS script for this type of comparison using bootstrap methods is available from M. Kinnison. For multiple time series of comparable length, ANCOVA can be used to test for heterogeneity of slopes in the relationship between $\ln x$ and elapsed years, or x/s_p , and elapsed generations (with appropriate ANCOVA considerations, Huitema 1980).

Evolutionary rate comparisons at different levels entail varying degrees of confidence. Confidence will be high when comparing rates for the same character, for the same study, over the same length of time. For example, the rate of evolution for Galápagos finch body size could be reliably compared between the two different selection episodes (Table 1). Stepping up a level of complexity, rates can be compared among different traits within the same study, such as beak length and beak width of finches (Table 1), or between the same trait in different populations, such as body size of male guppies from the Aripo and El Cedro Rivers (Reznick et al. 1997). Confidence in such comparisons would remain high because techniques and sample sizes would be similar and because the haldane is robust in its application to different types of traits. When quantitative comparisons are attempted among different studies, confidence may be diminished if techniques vary. In such comparisons, using the haldane instead of the darwin will minimize problems that can arise from variation in how a trait is measured.

A tempting level of inference has frequently been the comparison of microevolutionary rates to macroevolutionary rates (Losos et al. 1997; Reznick et al. 1997; Svensson 1997). Unfortunately this level is also the most tenuous, principally due to the different lengths of time over which micro- and macroevolution are measured. Time, the parameter often of most interest in rate comparisons, is also the most confounding factor in any such inference. Problems with comparisons of rates over different time frames arise if the measure assumes a particular pattern of increase (e.g., the darwin assumes an exponential increase; Gould 1984) or if the pooled standard deviation cannot be made representative of the long-term variation (for the haldane). Even if mechanistic problems were rectified by the invention of a perfect rate measure, comparisons across different time intervals will be confounded by the averaging of disparate rates and trajectories across time.

Temporal Scaling

Gingerich (1983) formalized an objection to inferences about comparisons of evolutionary rates measured over different time intervals. For instance, authors had asserted that mammals evolved more rapidly than invertebrates (Van Valen 1974) and that mammals evolved more rapidly during the Pleistocene than during any of the preceding epochs (Kurtén 1960). Gingerich pointed out that these inferences might not be correct because evolutionary rates are not independent of the measurement interval. To illustrate his point, he provided a simple graph using data from a compilation of studies,

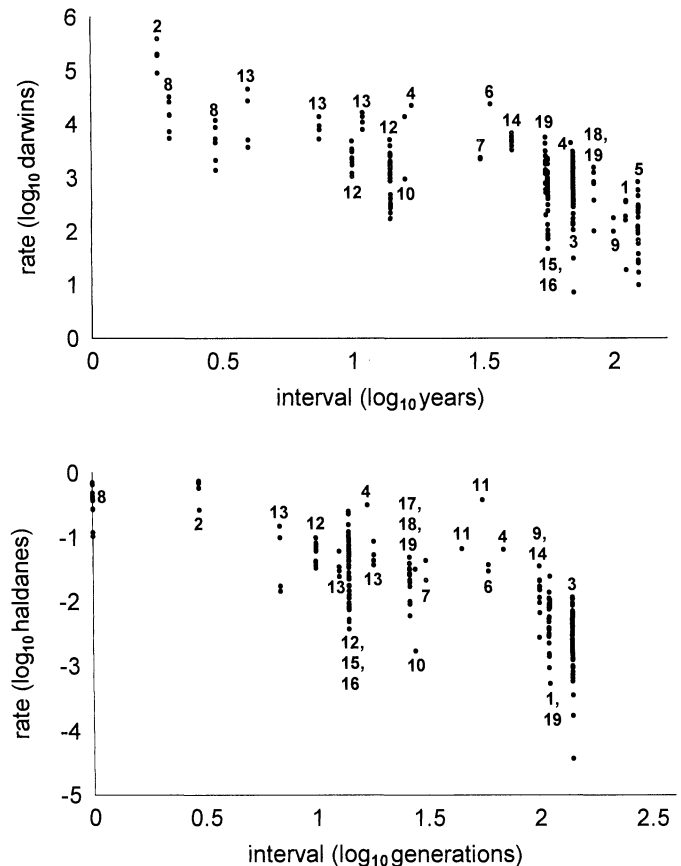


FIG. 4. Distribution of evolutionary rates in darwins (top panel) and haldanes (bottom panel) for the studies summarized in Table 1 (both rates and intervals are expressed as \log_{10} values for presentation purposes and for consistency with Gingerich 1993). Each point represents a single rate (for a single trait or population pairing) and each vertically grouped set of points represents the data for a given study. Negative rates (see Table 1) were expressed as positive rates. The numbers above or below each set of points refer to the specific studies shown in Table 1 and described in the Appendix.

which showed that evolutionary rates in darwins were negatively correlated with the length of time over which they were measured. We found similar associations for both darwins and haldanes in our survey of microevolutionary rates (Fig. 4). The reasons for these correlations are threefold: (1) the darwin assumes an exponential rate of change over time (because calculating a linear rate between two logarithms is equivalent to calculating an exponential rate between two untransformed values; Gould 1984); (2) plotting a rate versus the time interval over which it was calculated produces a negative slope resulting from a mathematical artifact called "spurious self-correlation" (Gould 1984; H. Sheets, pers. comm.); and (3) the longer the time interval, the more likely periods of stasis and evolutionary reversals will be averaged into the final value (Gingerich 1983). The haldane is free of the first constraint (because the mean is divided by the standard deviation, which tends to increase with the mean), but both darwins and haldanes will be sensitive to self-correlation when plotted against time interval. A solution is to plot the absolute amount of evolutionary change (numerator of darwins or haldanes) against time interval. The resulting pattern

in such a plot can be compared to the pattern that would result from constant evolutionary change, and deviations from that line will be reflective of variation in evolutionary trajectories. Performing this analysis on our dataset revealed considerable deviation from the expectation of constancy, indicating that evolutionary change eventually levels off or that disparate trajectories are averaged over longer time periods.

Gingerich (1983) argued that measures of evolutionary rate should be scaled to allow comparisons between sequences spanning different lengths of time. Gingerich (1993) introduced a temporal scaling technique for haldanes that involves examining the slope and intercept of log rate versus log interval (LRI) relationships. This technique can be used to estimate an intrinsic rate in haldanes, representing the mean absolute rate of evolutionary change over one generation (Gingerich 1993). Gingerich used the LRI technique to argue that scaled evolutionary rates were similar for three diverse studies, including Falconer's (1973) laboratory study on mouse body size, Seeley's (1986) study of relative spire height in *Littorina*, and Gingerich's own study of tooth shape in Eocene horses. The LRI technique shows promise, but seems to have limited statistical power when applied over long time frames (bootstrapped confidence intervals on intrinsic rates are large). Under such conditions most rates may appear similar. Nonetheless, it is worth specifying an intrinsic rate of evolution when the data permit.

In closing our consideration of rate comparisons, we caution that quantitative rate values are not sufficient to argue that micro- and macroevolution are (or are not) reflections of the same processes acting over different time scales. Instead, we urge further investigation into using rates of evolution to better understand the mechanisms (selection, mutation, genetic drift, gene flow) and patterns (stabilizing, directional, reversals, asymptotes) of contemporary microevolution. This information will be critical to conserving the biodiversity afforded us by both micro- and macroevolution.

BEYOND DARWINS AND HALDANES

And it is likely that better indices of evolutionary rate can be made than any which I have suggested (Haldane 1949, p. 55).

Random Walks and Time Series

For the most part we have discussed rates without regard to the evolutionary mechanism that produced them (i.e., selection, drift, mutation, or gene flow). Authors of the papers we reviewed typically assume selection is the reason for any evolution. This inference is often defensible because: (1) the population is very large and the rate of change very high (e.g., Quinn and Adams 1996); (2) the direction of change clearly matches adaptive predictions and is repeated in several parallel situations (e.g., Endler 1980; Williams and Moore 1989; Carroll et al. 1997; Reznick et al. 1997; Losos et al. 1997); or (3) the amount of change matches predictions derived from selection differentials and heritabilities (e.g., Grant and Grant 1995). In some instances, however, observed change may simply conform to a "random walk," owing to

effects of mutation and genetic drift in the absence of sustained selection. Several authors have advocated testing an observed pattern or rate against a null model before inferring selection as the mechanism. Gingerich's (1993) LRI analysis provides one way to test for directional change, random change, or stasis. Other techniques developed for similar purposes have been described and applied to paleontological data by Lande (1976), Charlesworth (1984), Bookstein (1987), Turelli et al. (1988), and Lynch (1990).

The Miocene stickleback data of Bell et al. (1985) was reanalyzed by Bookstein (1988) and tested against the null hypothesis of a random walk. Bookstein concluded that five of the six traits showed patterns consistent with a simple random walk, whereas body length showed evidence of stabilizing selection. Unfortunately, failure to reject a random-walk null hypothesis does not indicate that the change was random, simply that randomness cannot be unequivocally excluded as contributing to the change. The range of evolution/divergence contained within the confidence boundaries of a random walk is clearly obtainable by selection. Moreover, evolutionary series tracking a stochastic selective factor, such as temperature, may undergo many reversals and appear much like a random walk. Random walks are also very sensitive to "noise," making them unable to detect some clear instances of direction selection (H. Sheets, unpubl. data). Finally, random-walk tests are unsuitable for situations of divergence-with-gene-flow because migration can constrain the magnitude of divergence even under diversifying selection. Thus, failure to reject a random-walk model provides limited evidence that selection has not led to the observed pattern. Nevertheless, the *rejection* of a random-walk model can bolster any inference that the observed change is indeed the result of directional or stabilizing selection.

We suggest that time series be empirically examined with regard to shifts in the direction of evolution, especially when changing selective pressures are suspected. A simple approach maybe to use piecewise linear regression, which allows for break points at times that show discontinuities in the rate of change (Neter et al. 1989, pp. 370–374). In this manner, slopes (and therefore rates) can be calculated for parts of time series that show distinct evolutionary trends. More sophisticated techniques also exist for detecting discontinuities in temporal sequences. For example, Bell and Legendre (1987) used the chronological clustering technique of Legendre et al. (1985) to identify several morphological discontinuities in Bell et al.'s (1985) stickleback data. Regardless of the technique employed, it is important to consider (and measure) selective factors that may contribute to evolutionary patterns.

Complex Characters and Character Complexes

Most of our discussion has considered simple measures of single traits evaluated individually. For some evolutionary questions, however, a single measure may not accurately capture a trait's essence (i.e., a complex character). Additionally, several traits may be linked to one another through canalized developmental pathways, allometric constraints, and genetic covariance (i.e., a character complex). For example, Grant and Grant (1995, p. 248) found that "bill length was selected

in the 1984–1986 episode but it did not evolve, partly because the effect was nullified by selection in the opposite direction on positively correlated traits.” As recognition of the complexity and interdependence of traits has increased, a number of multivariate rate measures have been advocated.

An early measure of multivariate evolution was Mahalanobis’s distance (Mahalanobis 1936):

$$D = \sqrt{\sum \sum r^{ij} h_i h_j}, \quad (2)$$

where r^{ij} are the elements of the matrix reciprocal to the matrix of the correlation coefficients between characters (r_{ij}); h_i is the haldane rate of the i th character in two populations; and $i, j = 1, \dots, p$ are the traits, p in number (after Lerman 1965; for a different formulation, see Cherry et al. 1982). D measures the distance between groups as a vector in haldane space, accounting for correlations between characters, and providing a multivariate analog of the haldane. Lerman (1965) shows how Mahalanobis’s distance relates to a set of univariate evolutionary rates specified in darwins.

Another multivariate measure of evolutionary distance is Schluter’s (1984) selection distance:

$$B = \sqrt{\sum \beta_i^2}, \quad (3)$$

where β_i represents the elements of a vector of selection gradients for i traits. Selection gradients are a vector of differences in mean trait values divided by the genetic variance/covariance matrix (Schluter 1984). Selection distance is therefore the length of a vector in Euclidean space, representing the “total net force of selection that has acted to shift mean morphology” (Schluter 1984, p. 922). If differences in trait values are ln transformed and specified in millions of years, then B is somewhat equivalent to a multivariate darwin. If differences in trait values are specified in standard deviations per generation, then B is somewhat equivalent to a multivariate haldane. Unlike D , B uses information on genetic covariances, which are often difficult to obtain but may ultimately provide a better reflection of the process of evolution. Like other measures, D and B rely on specific assumptions that are not always tenable, and both are sensitive to errors in covariance estimates (Cherry et al. 1982; Schluter 1984; Endler 1986, p. 186).

Another approach is to use common techniques for obtaining multivariate descriptors of difference. For example, principal components or discriminant functions can be used to summarize correlated variation in multiple trait measures (Tabachnick and Fidell 1989). Warp scores from thin plate spline applications can be used as descriptors of multidimensional morphological form using bi- or tricoordinate data (Rohlf et al. 1996). Some multivariate descriptors seek to maximize the amount of difference between known groups (e.g., discriminant functions), and therefore may maximize estimates of trait evolutionary rate. Factor scores from such analyses can be used to produce an alternative “trait,” which can be used as one would a single trait to calculate an evolutionary rate in haldanes. The use of factor scores to estimate rates in darwins (e.g., Losos et al. 1997) needs further consideration as many multivariate analyses use scores standardized by subtracting the mean value, which is equivalent

to changing the zero point of the measurement scale, a procedure that may influence rates estimated using the darwin (see above).

A full consideration of the best multivariate measure of evolution remains to be undertaken, and should include an examination of the strengths and weaknesses of each measure, as well as the theoretical and empirical relationships among them. Finally, it must be noted that although multivariate approaches can incorporate covariance between traits or measures, this strength can also be a weakness if applied indiscriminately. It is arguable (e.g., Endler 1986) whether some of these mathematical constructs describe meaningful traits. The reliance of some multivariate measures on phenotypic correlations/covariance may be insensitive to the genetic correlations one would hypothesize underlie true multivariate evolution. Even with acceptable multivariate rate measures, we still face more difficult and beguiling issues surrounding comparisons of multivariate rates. In the absence of a clear consensus on univariate versus multivariate measures and on the best multivariate measure of evolution, we suggest estimating rates using a variety of available techniques.

Estimating Selection

Calculating a rate does not presuppose a particular evolutionary mechanism. However, if selection is the cause and if a trait’s narrow sense heritability can be specified, the selection intensity leading to the observed change (selection differential divided by the phenotypic standard deviation) can be estimated by dividing the haldane rate by the trait’s heritability (derivable from formulae in Endler 1986, p. 175; Falconer 1989, p. 192). The selection differential can be estimated by multiplying the selection intensity by the pooled standard deviation for the trait.

For a suite of potentially correlated traits, the theory of Lande (1979) and Lande and Arnold (1983) can be used to estimate selection from evolutionary rates specified in haldanes or darwins, assuming **P** and **G** (the phenotypic and additive genetic variance/covariance matrices) can be estimated. Differences in trait means can first be converted to an evolutionary response per generation: for haldanes, multiply the rate for each trait by the phenotypic standard deviation for the trait; for darwins, divide the rate for each trait by the number of generations per million years. The resulting evolutionary responses, specified as a vector (**Δz**), can be used to estimate linear selection gradients (**β**) and differentials (**S**), as **β** = **G**⁻¹**Δz** and **S** = **PG**⁻¹**Δz**. Standardized gradients can be obtained by multiplying by the phenotypic standard deviation for each trait, and standardized differentials can be obtained by dividing by the phenotypic standard deviation for the trait (Lande and Arnold 1983). See Reznick et al. (1997) for an example of estimating selection gradients and differentials from evolutionary responses. Grant and Grant (1995) provide an excellent comparison of measured evolutionary responses to those predicted from selection differentials, heritabilities, and genetic covariances.

“RAPID” EVOLUTION RECONSIDERED

The primary purpose of this review and perspective was to consider methods for estimating rates of microevolution.

However, our compilation also enables some preliminary inferences about rates themselves. Considering rates of evolution estimated from the 20 studies we reviewed, the fundamental conclusion that must be drawn is that evolution as hitherto considered "rapid" may often be the norm and not the exception. We suspect that when populations or species are exposed to changing environments, such as in the introduction experiments that constitute a large fraction of microevolution studies, evolution will appear rapid relative to that documented over longer time frames or in undisturbed situations. Thus, claims of rapid microevolution should not necessarily be considered exceptional, and perhaps represent typical rates of microevolution in contemporary populations facing environmental change.

Endler's (1986) review of natural selection in the wild suggested that selection intensities for quantitative characters are often quite high (geometric mean of statistically significant values = 0.59). The haldane is equivalent to the average per generation selection intensity multiplied by the heritability. Thus, even moderate heritabilities (e.g., 0.30) in combination with observed selection intensities should be able to generate evolutionary rates higher than those observed in most microevolutionary studies ($0.3 \times 0.59 = 0.18$). The question may indeed be asked why observed microevolutionary rates are not higher. Of the 224 haldane rates we calculated, 92% did not exceed 0.18 haldanes (Fig. 4B), and 14 of the 18 higher rates were from studies over less than three generations. We suspect the answer lies in the fact that selection (and perhaps additive genetic variance) is not constant or even consistent. Selection in the wild seems sufficiently strong to explain observed microevolutionary rates, but intense selection may only rarely be maintained for more than a few generations (otherwise observed rates would be higher).

What is the critical rate of environmental change that an evolving population can accommodate without mortality exceeding demographic potential? Several theoretical works (Lynch and Lande 1993; Bürger and Lynch 1995; Lynch 1996) suggest that the critical rate is equivalent to the maximum rate of sustainable evolution. For large, sexual populations, this rate is not likely to exceed a few percent of the phenotypic standard deviation per generation, and for small populations would be even less (Bürger and Lynch 1995; Lynch 1996). Of all the haldane rates we calculated, only 27% were greater than 5% of a standard deviation per generation (Fig. 4B), suggesting that evolution over microevolutionary time frames is usually not above the maximum sustainable rate. Of the rates higher than 5% of a standard deviation, 90% were from studies over less than 15 generations. If these high rates were maintained for longer, the probability of extinction might be quite high. For example, the rapid rate of change for Darwin's finches in response to drought conditions was coincident with a decrease in adult population size of 85% in a single generation (Boag and Grant 1981; Grant and Grant 1995). A comparison of maximum sustainable rates of evolution to the strength of selection may ultimately provide a useful tool for predicting population persistence in the face of environmental change.

We encourage the continued examination of evolutionary rates in all taxa and for different traits and time scales. With

standardization and improved statistical inference it will be possible to determine what rates are actually more rapid than others and to quantitatively evaluate patterns and mechanisms of microevolution. Likewise, evolutionary rates may prove instrumental in the conservation of biological diversity in a changing world. Perhaps the greatest contribution that evolutionary rate estimates will ultimately make is an awareness of our own role in the present microevolution of life and a cautious consideration of whether populations and species can adapt rapidly enough to forestall the macroevolutionary endpoint of extinction.

ACKNOWLEDGMENTS

This paper was the result of a number of debates and discussions with our friends and colleagues: S. Hodgson, K. Kay, A. Poon, D. Reznick, J. Rhodes, and members of the Huey/Kingsolver discussion group. The manuscript was improved through comments by M. Bell, S. Carroll, J. Endler, I. Fleming, P. Gingerich, P. Grant, J. Losos, A. Magurran, T. Quinn, D. Reznick, D. Roff, D. Schluter, H. Sheets, S. Stearns, F. Utter, M. Whitlock, and three anonymous reviewers. We hope that our perspective provides somewhat of a consensus on microevolutionary rates, but some of our colleagues undoubtedly disagree with some of our views. Unpublished data were provided by A. Helbig, J. Losos, T. Quinn, K. Warheit, and K. Williams. AH was supported by a Natural Sciences and Engineering Research Council of Canada Postdoctoral Fellowship. MK was supported by a scholarship from the H. Mason Keeler Endowment.

LITERATURE CITED

- Bell, M. A., and P. Legendre. 1987. Multicharacter chronological clustering in a sequence of fossil sticklebacks. *Syst. Zool.* 36: 52–61.
- Bell, M. A., J. V. Baumgartner, and E. C. Olson. 1985. Patterns of temporal change in single morphological characters of a Miocene stickleback fish. *Paleobiology* 11:258–271.
- Berthold, P., A. J. Helbig, G. Mohr, and U. Querner. 1992. Rapid microevolution of migratory behaviour in a wild bird species. *Nature* 360:668–670.
- Bielak, A. T., and G. Power. 1986. Changes in mean weight, sea-age composition, and catch-per-unit-effort of Atlantic salmon (*Salmo salar*) angled in the Godbout River, Quebec, 1859–1983. *Can. J. Fish. Aquat. Sci.* 43:281–287.
- Boag, P. T., and P. R. Grant. 1981. Intense natural selection in a population of Darwin's finches (*Geospizinae*) in the Galápagos. *Science* 214:82–85.
- Bookstein, F. L. 1987. Random walk and the existence of evolutionary rates. *Paleobiology* 13:446–464.
- . 1988. Random walk and the biometrics of morphological characters. *Evol. Biol.* 23:369–398.
- Bürger, R., and M. Lynch. 1995. Evolution and extinction in a changing environment: a quantitative-genetic analysis. *Evolution* 49:151–163.
- Carroll, S. P., H. Dingle, and S. P. Klassen. 1997. Genetic differentiation of fitness-associated traits among rapidly evolving populations of the soapberry bug. *Evolution* 51:1182–1188.
- Charlesworth, B. 1984. Some quantitative methods for studying evolutionary patterns in single characters. *Paleobiology* 10:308–318.
- Cherry, L. M., S. M. Case, J. G. Kunkel, J. S. Wyles, and A. C. Wilson. 1982. Body shape metrics and organismal evolution. *Evolution* 36:914–933.
- Clyde, W. C., and P. D. Gingerich. 1994. Rates of evolution in the

- dentition of early Eocene *Cantius*: comparison of size and shape. *Paleobiology*. 20:506–522.
- Darwin, C. D. 1859. On the origin of species by means of natural selection, or the preservation of favoured races in the struggle for life. J. Murray, London.
- Dobzhansky, T. 1937. Genetics and the origin of species. Columbia Univ. Press, New York.
- Dudley, J. W. 1977. Seventy-six generations of selection for oil and protein percentage in maize. Pp. 459–473 in E. Pollak, O. Kempthorne, and T. B. Bailey, eds. Proceedings of the international conference on quantitative genetics. Iowa State University, Ames, IA.
- Endler, J. A. 1980. Natural selection on color patterns in *Poecilia reticulata*. *Evolution* 34:76–91.
- . 1986. Natural selection in the wild. Princeton Univ. Press, Princeton, NJ.
- Falconer, D. S. 1973. Replicated selection for body weight in mice. *Genet. Res., Camb.* 22:291–321.
- . 1989. Introduction to quantitative genetics. Wiley, New York.
- Fenster, E. J., M. K. Hecht, and U. Sorhannus. 1992. Problems in the measurement of morphological rates of change. *Ann. Zool. Fennici* 28:165–174.
- Gibbs, H. L., and P. R. Grant. 1987. Oscillating selection on Darwin's finches. *Nature* 327:511–513.
- Gingerich, P. D. 1983. Rates of evolution: effects of time and temporal scaling. *Science* 222:159–161.
- . 1993. Quantification and comparison of evolutionary rates. *Am. J. Sci.* 293A:453–478.
- Givnish, T. J., and K. J. Sytsma. 1997. Molecular evolution and adaptive radiation. Cambridge Univ. Press, Cambridge, U.K.
- Goldschmidt, R. 1940. The material basis of evolution. Yale Univ. Press, New Haven, CT.
- Gould, S. J. 1984. Smooth curve of evolutionary rate: a psychological and mathematical artifact. *Science* 226:994–995.
- Grant, P. R., and B. R. Grant. 1995. Predicting microevolutionary responses to directional selection on heritable variation. *Evolution* 49:241–251.
- Haldane, J. B. S. 1949. Suggestions as to quantitative measurement of rates of evolution. *Evolution* 3:51–56.
- Harvey, P. H., and A. Purvis. 1991. Comparative methods for explaining adaptations. *Nature* 351:619–624.
- Hendry, A. P., and M. T. Kinnison. 1998. Taking time with microevolution. *Trends Ecol. Evol.* 13:76–77.
- Hendry, A. P., and T. P. Quinn. 1997. Variation in adult life history and morphology among Lake Washington sockeye salmon (*Oncorhynchus nerka*) populations in relation to habitat features and ancestral affinities. *Can. J. Fish. Aquat. Sci.* 54:75–84.
- Hendry, A. P., J. E. Hensleigh, and R. R. Reisenbichler. 1998. Incubation temperature, developmental biology, and the divergence of sockeye salmon (*Oncorhynchus nerka*) within Lake Washington. *Can. J. Fish. Aquat. Sci.* 55:1387–1394.
- Huitema, B. E. 1980. The analysis of covariance and alternatives. Wiley, New York.
- Huxley, J. S. 1932. Problems of relative growth. The Dial Press, New York.
- Johnston, R. F., and R. K. Selander. 1964. House sparrows: rapid evolution of races in North America. *Science* 144:548–550.
- Kinnison, M., M. Unwin, N. Boustead, and T. Quinn. 1998a. Population-specific variation in body dimensions of adult chinook salmon (*Oncorhynchus tshawytscha*) from New Zealand and their source population, 90 years after introduction. *Can. J. Fish. Aquat. Sci.* 55:554–563.
- Kinnison, M. T., M. J. Unwin, W. K. Hershberger, and T. P. Quinn. 1998b. Egg size, fecundity, and development rate of two introduced New Zealand chinook salmon (*Oncorhynchus tshawytscha*) populations. *Can. J. Fish. Aquat. Sci.* 55:1946–1953.
- Kinnison, M. T., M. J. Unwin, and T. P. Quinn. 1998c. Growth and salinity tolerance of juvenile chinook salmon (*Oncorhynchus tshawytscha*) from two introduced New Zealand populations. *Can. J. Zool.* 76:2219–2226.
- Klepaker, T. 1993. Morphological changes in a marine population of threespined stickleback, *Gasterosteus aculeatus*, recently isolated in fresh water. *Can. J. Zool.* 71:1251–1258.
- Kurtén, B. 1960. Rates of evolution in fossil mammals. Cold Spring Harbor Symp. Quant. Biol. 24:205–214.
- Lande, R. 1976. Natural selection and random genetic drift in phenotypic evolution. *Evolution* 30:314–334.
- . 1979. Quantitative genetic analysis of multivariate evolution, applied to brain:body size allometry. *Evolution* 33:402–416.
- Lande, R., and S. J. Arnold. 1983. The measurement of selection on correlated characters. *Evolution* 37:1210–1226.
- Legendre, P., S. Dallot, and L. Legendre. 1985. Succession of species within a community: chronological clustering, with applications to marine and freshwater zooplankton. *Am. Nat.* 125:257–288.
- Lenski, R. E., and M. Travisano. 1994. Dynamics of adaptation and diversification: a 10,000-generation experiment with bacterial populations. *Proc. Natl. Acad. Sci. USA* 91:6808–6814.
- Lerman, A. 1965. On rates of evolution of unit characters and character complexes. *Evolution* 19:16–25.
- Li, W.-H. 1997. Molecular evolution. Sinauer, Sunderland, MA.
- Losos, J. B., K. I. Warheit, and T. W. Schoener. 1997. Adaptive differentiation following experimental island colonization in *Anolis* lizards. *Nature* 387:70–73.
- Lynch, M. 1990. The rate of morphological evolution in mammals from the standpoint of the neutral expectation. *Am. Nat.* 136:727–741.
- . 1996. A quantitative-genetic perspective on conservation issues. Pp. 471–501 in J. C. Avise and J. L. Hamrick, eds. Conservation genetics: case histories from nature. Chapman and Hall, New York.
- Lynch, M., and R. Lande. 1993. Evolution and extinction in response to environmental change. Pp. 234–250 in P. M. Kareiva, J. G. Kingsolver, and R. B. Huey, eds. Biotic interactions and global change. Sinauer, Sunderland, MA.
- Magurran, A. E. 1998. Population differentiation without speciation. *Phil. Trans. R. Soc. Lond. B Biol. Sci.* 353:275–286.
- Magurran, A. E., B. H. Seghers, G. R. Carvalho, and P. W. Shaw. 1992. Behavioural consequences of an artificial introduction of guppies (*Poecilia reticulata*) in N. Trinidad: evidence for the evolution of anti-predator behaviour in the wild. *Proc. R. Soc. Lond. B Biol. Sci.* 248:117–122.
- Magurran, A. E., B. H. Seghers, P. W. Shaw, and G. R. Carvalho. 1995. The behavioral diversity and evolution of guppy, *Poecilia reticulata*, populations in Trinidad. *Adv. Study Behav.* 24:155–202.
- Mahalanobis, P. C. 1936. On the generalized distance in statistics. *Proc. Nat. Inst. Sci. India* 2:49–55.
- Manly, B. F. J. 1997. Randomization, bootstrap and Monte Carlo methods in biology. Chapman and Hall, New York.
- Mayr, E. 1942. Systematics and the origin of species. Columbia Univ. Press, New York.
- Neter, J., W. Wasserman, and M. H. Kutner. 1989. Applied linear regression models. Irwin, Boston, MA.
- Quinn, T. P., and D. J. Adams. 1996. Environmental changes affecting the migratory timing of American shad and sockeye salmon. *Ecology* 77:1151–1162.
- Reznick, D. N., H. Bryga, and J. A. Endler. 1990. Experimentally induced life-history evolution in a natural population. *Nature* 346:357–359.
- Reznick, D. N., F. H. Shaw, F. H. Rodd, and R. G. Shaw. 1997. Evaluation of the rate of evolution in natural populations of guppies (*Poecilia reticulata*). *Science* 275:1934–1937.
- Ricklefs, R. E. 1973. Ecology. Chiron, Newton, MA.
- Rohlf, F. J., A. Loy, and M. Corti. 1996. Morphometric analysis of Old World Talpidae (Mammalia, Insectivora) using partial-warp scores. *Syst. Biol.* 45:344–362.
- Schluter, D. 1984. Morphological and phylogenetic relations among the Darwin's finches. *Evolution* 38:921–930.
- Seeley, R. H. 1986. Intense natural selection caused a rapid morphological transition in a living marine snail. *Proc. Natl. Acad. Sci. USA* 83:6897–6901.

- Simpson, G. G. 1944. Tempo and mode in evolution. Columbia Univ. Press, New York.
- . 1953. The major features of evolution. Columbia Univ. Press, New York.
- Singer, M. C., C. D. Thomas, and C. Parmesan. 1993. Rapid human-induced evolution of insect-host associations. *Nature* 366:681–683.
- Smith, T. B., L. A. Freed, J. K. Lepson, and J. H. Carothers. 1995. Evolutionary consequences of extinctions in populations of a Hawaiian honeycreeper. *Conserv. Biol.* 9:107–113.
- Snaydon, R. W., and M. S. Davies. 1972. Rapid population differentiation in a mosaic environment. II. Morphological variation in *Anthoxanthum odoratum*. *Evolution* 26:390–405.
- Stearns, S. C. 1983a. The evolution of life-history traits in mosquitofish since their introduction to Hawaii in 1905: rates of evolution, heritabilities, and developmental plasticity. *Am. Zool.* 23:65–76.
- . 1983b. The genetic basis of differences in life-history traits among six populations of mosquitofish (*Gambusia affinis*) that shared ancestors in 1905. *Evolution* 37:618–627.
- . 1992. The evolution of life histories. Oxford Univ. Press, Oxford, U.K.
- Stockwell, C. A., and S. C. Weeks. 1999. Translocations and rapid evolutionary responses in recently established populations of western mosquitofish (*Gambusia affinis*). *Anim. Conserv.* 2:103–110.
- Svensson, E. 1997. The speed of life-history evolution. *Trends Ecol. Evol.* 12:380–381.
- Tabachnik, B. G., and L. S. Fidell. 1989. Using multivariate statistics. 2d ed. Harper Collins, New York.
- Thompson, J. N. 1998. Rapid evolution as an ecological process. *Trends Ecol. Evol.* 13:329–332.
- Turelli, M., J. H. Gillespie, and R. Lande. 1988. Rate tests for selection on quantitative characters during macroevolution and microevolution. *Evolution* 42:1085–1089.
- Van Valen, L. 1974. Two modes of evolution. *Nature* 252:298–300.
- Williams, C. K., and R. J. Moore. 1989. Phenotypic adaptation and natural selection in the wild rabbit, *Oryctolagus cuniculus*, in Australia. *J. Anim. Ecol.* 58:495–507.
- Wright, S. 1968. Evolution and the genetics of populations. I. Genetic and biometric foundations. Univ. of Chicago Press, Chicago.
- Yoo, B. H. 1980. Long-term selection for a quantitative character in large replicate populations of *Drosophila melanogaster*. *Genet. Res.* 35:1–17.
- 1983b). Note that rates in darwins calculated by Stearns (1992, p. 118) for the same study are not correct (S. Stearns, pers. comm.).
- ⁴ Seeley (1986): Rates were calculated based on changes in relative spire height (\ln shell thickness/ \ln shell width) over two time periods for Nahant samples (Seeley 1986). Standard deviation was only reported for the starting sample (1898). Gingerich (1993) obtained the same rates in haldanes for Seeley's (1986) study. Gingerich (1993) also performed his LRI analysis to show that relative spire height has evolved at an "intrinsic rate" of 1.905 haldanes. Generation length is assumed to be one year.
- ⁵ Williams and Moore (1989): Rates in darwins were calculated using mean values for rabbits from three regional "populations." Phenotypic rates were for rabbits captured in the wild (table 2 in Williams and Moore 1989) and genetic rates were for their progeny raised at 15°C under common conditions (table 4 in Williams and Moore 1989). All traits were standardized to a common body size. Haldane rates could not be calculated because variances and generation lengths were not reported.
- ⁶ Magurran et al. (1992): Rates were calculated using means and standard errors for "Guanapo" versus "Turure CR" estimated from figure 1A in Magurran et al. (1992). Guanapo is considered the most likely source for Haskins's transplant (A. Magurran, pers. comm.) and is used to represent the Turure source population in Magurran (1998). Schooling tendency was the amount of time fish spent within five body lengths of the school (out of a total of 300 sec). These guppies have about 1.74 generations per year (Reznick et al. 1997).
- ⁷ Klepaker (1993): Rates were calculated using means and standard deviations (from table 5 in Klepaker 1993) and represent divergence between marine sticklebacks and the 1991 samples from the freshwater pond. We estimated divergence rates for the two univariate metric characters for which the Klepaker (1993) provided an adaptive explanation, spine length (first dorsal) and longitudinal eye diameter. Generation length was assumed to be one year.
- ⁸ Grant and Grant (1995): Rates were calculated using means, standard errors, and sample sizes for each of two single-generation episodes of selection: 1976–1978 and 1984–1987 (tables 2 and 4 in Grant and Grant 1995). The traits are highly heritable and, although the measured evolution is technically "phenotypic," the underlying basis for the change is genetic. Morphological measurements were not standardized to a common body size.
- ⁹ Smith et al. (1995): Rates were calculated using means, standard errors, and sample sizes for adult males collected prior to 1902 compared to those collected between 1988 and 1991 (table 1 in Smith et al. 1995). Smith et al. (1995) estimate the elapsed years and generations to be approximately 100.
- ¹⁰ Magurran et al. (1995): Rates were calculated using means and confidence intervals for "L Aripo" versus "Aripo (I)" estimated from figure 11 in Magurran et al. (1995). See note for more details.
- ¹¹ Quinn and Adams (1996): Rates in haldanes were calculated from raw data provide by T. Quinn. The migration date of each fish was natural-log transformed and the mean and standard deviation of migration date were calculated in each year. The pooled standard deviation was calculated across all of the years and used to determine x/s_p for each year. Linear regression was used to calculate the slope of the relationship between these points and generations, assuming a generation length of four years for both species (Quinn and Adams 1996). Residuals were not autocorrelated and the 95% confidence interval for the rate of change in migration date was -0.271 to -0.493 ($P < 0.001$) for shad and -0.033 to -0.099 ($P < 0.001$) for salmon. Rates in darwins cannot be calculated because migration date is not on a ratio scale.
- ¹² Losos et al. (1997): Rates in darwins were reported in Losos et al. (1997). Rates in haldanes were calculated using means, standard deviations, and sample sizes provided by K. Warheit. For haldane calculations, two islands ("i3" and "i25") were excluded due to small sample sizes. PC1 and PC2 are multivariate measures of body shape defined using principal components for traits standardized to a common body size. Hindlimb length was also standardized for body size. Generation length was assumed to be one year, but may be 9–18 months (J. Losos, pers. comm.).
- ¹³ Reznick et al. (1997): Rates in darwins were reported in Reznick et al. (1997) and represent the range for three different "ex-

Corresponding Editor: J. Neigel

APPENDIX

Notes on the Calculation of Evolutionary Rates in Table 1

¹ Johnston and Selander (1964): Rates were calculated using means and standard deviations from figures 3, 4, and 5 in Johnston and Selander (1964). Bill lengths for Honolulu and body weights for Mexico City were excluded. Traits were not standardized to a common body size. Rates are lower limits because introduced house sparrows did not reach some sites until well after their introduction in 1852. Generation length is assumed to be one year (Johnston and Selander 1964).

² Endler (1980): Rates were calculated using means and standard errors estimated from figure 4 in Endler (1980). Sample sizes were 100 (J. Endler, pers. comm.). The genetic basis for divergence was not explicitly measured, but male guppy color patterns have a genetic basis. Endler (1980) estimated that 15 generations had passed (about 8.3 generations/yr), but a life-table analysis by Reznick et al. (1997) indicated that guppies in the Aripo River have 1.74 generations per year. We used Reznick's value.

³ Stearns (1983b): Rates were calculated based on means and confidence intervals for all pairwise comparisons among the six Hawaiian populations (from table 1 in Stearns 1983b). Mosquitofish were assumed to have about two generations per year (Stearns

1983b). Note that rates in darwins calculated by Stearns (1992, p. 118) for the same study are not correct (S. Stearns, pers. comm.).

perimental" versus "control" population pairings. Rates in hal-danes were calculated using means, standard errors, sample sizes, and generation times specified for the same pairings (from table 1 in Reznick et al. 1997). Progeny of wild-caught guppies were reared in the laboratory for two generations. Guppies have about 1.74 generations per year (Reznick et al. 1997).

¹⁴ Carroll et al. (1997): Rates were calculated using means and standard errors estimated from figure 2 (females collected from the field) and figure 3 (females reared in the laboratory) in Carroll et al. (1997). Rates represent the range of divergence rates in pairwise comparisons for two ancestral populations (Key Largo and Plantation Key) versus two derived populations (Lake Wales and Leesburg). For the laboratory-reared bugs, only females reared on their "Home" host were used in the rate calculations. The authors estimate that 100 generations have passed for soapberry bugs on their new host plant (Carroll et al. 1997; S. Carroll, pers. comm.).

¹⁵ Hendry and Quinn (1997): Rates were calculated based on means and standard errors for all pairwise comparisons among the three Lake Washington populations of non-native origin (from figures 3, 5, 6 in Hendry and Quinn 1997). Rates were calculated separately for two years (excluding Issaquah female length in 1993). Body length was for fish that matured at four years of age and body depth was standardized to a common body length (Hendry and Quinn 1997). Generation length is about 4 years.

¹⁶ Hendry et al. (1998): Rates were calculated for all pairwise comparisons among groups of sockeye salmon that spawn at different times or places within the Lake Washington watershed (Hendry et al. 1998). Calculations were based on the mean and variance in days from fertilization to hatching (time to hatch) and days from fertilization to emergence (time to emerge) among full-sib families incubated at 9°C (A. Hendry, unpubl. data). Generation length is about four years.

¹⁷ Kinnison et al. (1998a): Rates were calculated using raw data for two years of sampling in New Zealand chinook salmon populations (M. Kinnison, unpubl. data). PC1 and PC2 were multivariate measures of female and male fin shape, respectively, from principal components after log transformation and standardization to a common body length. Generation length (3.21 yr) was determined by averaging age at maturity for the two populations and sexes. Rates were not calculated in darwins because using PC factor scores changes the zero point of the measurement scale.

¹⁸ Kinnison et al. (1998b): Rates were calculated using raw data for divergence between the two New Zealand populations in two different years (M. Kinnison, unpubl. data). Egg weight was standardized to a common body length, GSI was the ratio of gonad weight to somatic weight, and time to hatch was days from fertilization to hatching for embryos reared at 5.9°C (table 1 and text of Kinnison et al. 1998b). Generation length is 3.21 years (as in note 17).

¹⁹ Kinnison et al. (1998c): Rates were calculated using raw data for divergence between the two New Zealand populations (M. Kinnison, unpubl. data). Growth rates were estimated for three different growth stanzas (table 2 and fig. 2 in Kinnison et al. 1998c). Note that growth rates are not log transformed for rate calculations because they already represent \log_e values. Generation length is 3.21 years (as in note 17).

²⁰ Stockwell and Weeks (1999): Rates were calculated using means and standard errors for divergence among four Nevada populations (table 3 in Stockwell and Weeks 1999). Progeny of wild-caught mosquitofish were reared in the laboratory for two generations. Time since colonization was about 55 years (Stockwell and Weeks 1999), and mosquitofish have about two generations per year (after Stearns 1983b).