

Appendix from L. J. Revell, “On the Analysis of Evolutionary Change along Single Branches in a Phylogeny”

(Am. Nat., vol. 172, no. 1, p. 140)

Details and a Worked Example of the Likelihood Method

The likelihood approach for the analysis of evolutionary changes along isolated branches that I apply in this study is based on the “noncensored” approach of O’Meara et al. (2006; those authors also introduce a “censored” approach, but it is not applicable here.) According to this method, we fit our data to a multivariate normal distribution in which the variances for and covariances among tip values (phenotypic trait values for species) have been rescaled according to a specific hypothesis of rate heterogeneity, and we compare the fit of this model with one in which the evolutionary rate is homogeneous on all branches. In this study, the hypothesis of rate heterogeneity invariably involves heterogeneous rates on isolated branches.

For illustrative purposes, consider the hypothetical (and highly simplified) tree with branch lengths and phenotypic values for species provided in figure A1. Under a hypothesis of rate homogeneity and with Brownian motion as our model for the evolutionary process, the expected variance-covariance matrix for the observations of species phenotype values at the tips of the tree is proportional to the matrix:

$$\mathbf{C} = \mathbf{H} \begin{matrix} & \begin{matrix} A & B & C & D & E & F & G & H & I & J & K & L & M & N & O & P \end{matrix} \\ \begin{matrix} A \\ B \\ C \\ D \\ E \\ F \\ G \\ H \\ I \\ J \\ K \\ L \\ M \\ N \\ O \\ P \end{matrix} & \begin{bmatrix} 4 & 3 & 2 & 2 & 1 & 1 & 1 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 3 & 4 & 2 & 2 & 1 & 1 & 1 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 2 & 2 & 4 & 3 & 1 & 1 & 1 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 2 & 2 & 3 & 4 & 1 & 1 & 1 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 1 & 1 & 1 & 1 & 4 & 3 & 2 & 2 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 1 & 1 & 1 & 1 & 3 & 4 & 2 & 2 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 1 & 1 & 1 & 1 & 2 & 2 & 4 & 3 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 1 & 1 & 1 & 1 & 2 & 2 & 3 & 4 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 4 & 3 & 2 & 2 & 1 & 1 & 1 & 1 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 3 & 4 & 2 & 2 & 1 & 1 & 1 & 1 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 2 & 2 & 4 & 3 & 1 & 1 & 1 & 1 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 2 & 2 & 3 & 4 & 1 & 1 & 1 & 1 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 1 & 1 & 1 & 4 & 3 & 2 & 2 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 1 & 1 & 1 & 3 & 4 & 2 & 2 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 1 & 1 & 1 & 2 & 2 & 4 & 3 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 1 & 1 & 1 & 2 & 2 & 3 & 4 \end{bmatrix} \end{matrix} .$$

This matrix can be computed easily (or, for the highly simplified tree in fig. A1, even generated by hand). This is accomplished by summing the total branch length between the root and the common ancestor of each pair of taxa, i and j , into the matrix position (i, j) .

The basis for so doing lies in the evolutionary process and the manner in which covariance is acquired between species. Under Brownian motion (which is the most common model for the evolution of continuously valued characters and is based on the random diffusion of particles through a fluid), variance among lineages is accumulated in direct proportion to the time since the lineages shared a common ancestor. However, just as variance among taxa is accumulated by separate evolution in proportion to the time that lineages have evolved apart, covariance between species is accumulated in direct proportion to the shared history of taxa.

Thus, from the tree portrayed in figure A1, for example, the element $C(A, B)$ of \mathbf{C} is computed by summing

the three branches each of unit length (see fig. A1 legend) between the root of the tree and the common ancestor of taxa A and B in the matrix position (A, B) to obtain $C(A, B) = 3$. This number is expected to be proportional to the accumulated covariance between taxa A and B. All other elements of \mathbf{C} are computed similarly. Taxa separated by the root of the tree have no expected covariance (they are independent); hence, for the example in figure A1, $C(i, j) = 0.0$ for $i = A-H$ and $j = I-P$ (or vice versa). Zero covariance between taxa separated by the root is expected because such taxa have no shared history (in the context of our phylogenetic tree).

To compare an alternative hypothesis of rate heterogeneity with a null hypothesis of evolutionary rate homogeneity, we must first find a maximum likelihood estimate for the Brownian motion evolutionary rate, σ^2 . This can be obtained by finding a value for σ^2 that maximizes the equation for the natural logarithm of the likelihood. Based on the multivariate normal distribution, this equation is

$$\ln(L) = -\frac{1}{2}[(\mathbf{x} - \hat{\mathbf{a}}\mathbf{1})'(\sigma^2\mathbf{C})^{-1}(\mathbf{x} - \hat{\mathbf{a}}\mathbf{1})] - \frac{1}{2}\ln|\sigma^2\mathbf{C}| - \frac{1}{2}[n \cdot \ln(2\pi)].$$

In the likelihood equation, \mathbf{x} is an $n \times 1$ vector (for n total taxa in the tree; $n = 16$ in fig. A1) containing the species means for the phenotypic trait for all species; $\hat{\mathbf{a}}$ is the “phylogenetic mean,” whose specific meaning and calculation are described below; and $\mathbf{1}$ is an $n \times 1$ column vector of 1.0s. The phylogenetic mean is the estimated ancestral state at the root node, which in the rate-homogeneous situation can be evaluated as $\hat{\mathbf{a}} = (\mathbf{1}'\mathbf{C}^{-1}\mathbf{1})^{-1}(\mathbf{1}'\mathbf{C}^{-1}\mathbf{x})$ (Rohlf 2001). For the completely balanced tree in figure A1, the phylogenetic mean and the arithmetic mean of \mathbf{x} are identical ($\bar{x} = \hat{\mathbf{a}} = -0.74$); however, this will not generally be the case.

The maximum likelihood estimate of σ^2 can be obtained numerically or by using the analytic equation provided by O’Meara et al. (2006). For the tree and data of figure A1, $\ln(L_1)$ (in which the subscript “1” is used henceforward to indicate that a one-rate model is being fitted) is maximized at $\sigma^2 = 3.3$ and $\ln(L_1) = -39.2$. The interpretation of this finding is that the single-rate Brownian motion process most likely to have produced the data at the tips of the tree in figure A1 has a rate parameter (the rate at which variance is accumulated among species per unit time) of $\sigma^2 = 3.3$.

If an alternative model for the evolutionary history of the group has been proposed—say, a hypothesis involving rate heterogeneity—then we can also investigate the likelihood of that model. Returning to the tree and data of figure A1, we now assume that the phenotypic means of the species at the tips of the tree have evolved by a heterogeneous-rate Brownian motion process. In this case, the heterogeneous process is one in which a single rate prevails on most of the branches of the tree (indicated by solid lines) while a rare second rate is found on several isolated branches (indicated by dashed lines). I refer to the number of branches hypothesized to exhibit a rare second rate as m^* , where $m^* = 5$ for the example in figure A1.

We begin by calculating two matrices. The first, \mathbf{C}_0 , contains elements proportional to the components of variances and covariances for the values of phenotypes at the tips of the tree due to the prevalent rate found on the solid branches of figure A1. The matrix \mathbf{C}_0 is calculated by summing all solid branches into the matrix in a manner equivalent to that used to obtain the \mathbf{C} matrix in the homogeneous-rate case above, the only difference being that dashed branches are not summed into \mathbf{C}_0 . Doing so for the tree in figure A1 results in a matrix \mathbf{C}_0 equal to

App. from L. J. Revell, "Evolution along Single Branches"

	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P
A	3	2	2	2	1	1	1	1	0	0	0	0	0	0	0	0
B	2	3	2	2	1	1	1	1	0	0	0	0	0	0	0	0
C	2	2	4	3	1	1	1	1	0	0	0	0	0	0	0	0
D	2	2	3	4	1	1	1	1	0	0	0	0	0	0	0	0
E	1	1	1	1	4	3	2	2	0	0	0	0	0	0	0	0
F	1	1	1	1	3	4	2	2	0	0	0	0	0	0	0	0
G	1	1	1	1	2	2	3	3	0	0	0	0	0	0	0	0
$C_0 = H$	1	1	1	1	2	2	3	4	0	0	0	0	0	0	0	0
I	0	0	0	0	0	0	0	0	3	2	1	1	1	1	1	1
J	0	0	0	0	0	0	0	0	2	3	1	1	1	1	1	1
K	0	0	0	0	0	0	0	0	1	1	3	2	1	1	1	1
L	0	0	0	0	0	0	0	0	1	1	2	2	1	1	1	1
M	0	0	0	0	0	0	0	0	1	1	1	1	3	2	2	2
N	0	0	0	0	0	0	0	0	1	1	1	1	2	2	2	2
O	0	0	0	0	0	0	0	0	1	1	1	1	2	2	4	3
P	0	0	0	0	0	0	0	0	1	1	1	1	2	2	3	4

In this example, elements of C_0 differing from their counterparts in the homogeneous-rate covariance matrix C are in boldface.

The second matrix, C^* , is calculated by summing all dashed branches of figure A1 into the matrix position $C^*(i, j)$ if the branch is found in the common history of taxa i and j . For the data in figure A1, C^* is calculated as

	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P
A	1	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0
B	1	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0
C	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
D	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
E	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
F	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
G	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0
$C^* = H$	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
I	0	0	0	0	0	0	0	0	1	1	1	1	0	0	0	0
J	0	0	0	0	0	0	0	0	1	1	1	1	0	0	0	0
K	0	0	0	0	0	0	0	0	1	1	1	1	0	0	0	0
L	0	0	0	0	0	0	0	0	1	1	1	2	0	0	0	0
M	0	0	0	0	0	0	0	0	0	0	0	0	1	1	0	0
N	0	0	0	0	0	0	0	0	0	0	0	0	1	1	0	0
O	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
P	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0

In the matrix C^* , elements that are nonzero are in boldface.

Before attempting to find the likelihood of a given hypothesis of evolutionary rate heterogeneity, we must first deal with the complicating factor that the estimated ancestral state at the root node is contingent on our model for evolutionary heterogeneity throughout the tree. This is obvious if one considers the case in which the evolutionary rate to the left of the root is very high and the evolutionary rate to the right of the root is very low. Under these circumstances, clearly the ancestral state at the root will be closer to the values of the tips of the tree in the right half of the tree, where little change from the root would be expected to have accrued, than in the left half of the tree.

The problem is simplified slightly by defining the rate heterogeneity parameter h as $h = \sigma_*^2/\sigma_0^2$, where σ_*^2 and σ_0^2 are the evolutionary rates along exceptional and unexceptional branches, respectively. For a given value of h , the maximum likelihood ancestral state at the root node can then be calculated as

$$\hat{a}_h = [\mathbf{1}'(\mathbf{C}_0 + h\mathbf{C}^*)^{-1}\mathbf{1}]^{-1}[\mathbf{1}'(\mathbf{C}_0 + h\mathbf{C}^*)^{-1}\mathbf{x}].$$

It should be evident that for $h = 1.0$ (i.e., no heterogeneity in the evolutionary rate), the preceding equation reduces to the single-rate equation for \hat{a} .

In the rate-homogeneous case, as noted above, the maximum likelihood evolutionary rate can either be found by maximizing the likelihood numerically or calculated using the analytic solution provided by O'Meara et al. (2006). Similarly, for a given value of h , the maximum likelihood estimates for the two evolutionary rates can be computed analytically as follows:

$$\hat{\sigma}_0^2 = \frac{(\mathbf{x} - \hat{a}_h\mathbf{1})'(\mathbf{C}_0 + h\mathbf{C}^*)^{-1}(\mathbf{x} - \hat{a}_h\mathbf{1})}{n}, \quad (\text{A1})$$

$$\hat{\sigma}_*^2 = h \cdot \hat{\sigma}_0^2. \quad (\text{A2})$$

Thus, the problem is transformed into a univariate optimization in which we must find the value for h (\hat{h} , the maximum likelihood estimate of h) that maximizes the following equation for the likelihood:

$$\ln(L_2) = -\frac{1}{2}\{(\mathbf{x} - \hat{a}_h\mathbf{1})'[\hat{\sigma}_0^2(\mathbf{C}_0 + h\mathbf{C}^*)]^{-1}(\mathbf{x} - \hat{a}_h\mathbf{1})\} - \frac{1}{2}\ln|\hat{\sigma}_0^2(\mathbf{C}_0 + h \cdot \mathbf{C}^*)| - \frac{1}{2}[n \cdot \ln(2\pi)].$$

Of course, it should always be kept in mind that \hat{a}_h , $\hat{\sigma}_0^2$, and $\hat{\sigma}_*^2$ are recalculated for any given value of h whose likelihood must be evaluated.

Figure A2 shows the likelihood surface for $h = \sigma_*^2/\sigma_0^2$ from the data and tree of figure A1. The true value of the rate ratio used to generate the data ($h = 16.0$) and the maximum likelihood estimate of h ($\hat{h} = 19.1$) are indicated by vertical dashed lines (fig. A2). The equation for the likelihood has a maximum value of $\ln(L_2) = -32.9$. The maximum likelihood estimate of the ancestral state at the root node, given $\hat{h} = 19.1$, is $\hat{a}_h = -0.003$, which is much closer to the true ancestral state (0.0) than was the value estimated under the assumption of rate homogeneity ($\hat{a} = -0.74$). Using equations (A1) and (A2), we can also compute the maximum likelihood estimates of the evolutionary rates as $\hat{\sigma}_*^2 = 14.3$ and $\hat{\sigma}_0^2 = 0.75$ (for true values of $\sigma_*^2 = 16.0$ and $\sigma_0^2 = 1.0$, respectively).

Now that we have obtained likelihoods for our one- and two-rate models for the evolutionary process, given the tree and data in figure A1, we can test our hypothesis of rate heterogeneity against the null hypothesis that rates are homogeneous across the branches of the phylogeny. To do so, we first calculate the likelihood ratio statistic as $-2 \cdot \ln(L_1/L_2) = -2 \cdot [\ln(L_1) - \ln(L_2)]$. For the data and tree of figure A1 and the likelihoods calculated above, this statistic evaluates to $-2 \cdot [(-39.2) - (-32.9)] = 12.5$ (obviously with some allowance for rounding). The likelihood ratio is expected to be asymptotically distributed as a χ^2 with degrees of freedom equivalent to the difference between the numbers of parameters estimated in the denominator and numerator models (Chernoff 1954). For two versus one evolutionary rates, $df = 1$. Evaluating the probability of obtaining $\chi^2 \geq 12.5$ ($df = 1$) is a straightforward affair and results in a P value of <0.001 . As mentioned in the main text, and as is more thoroughly discussed in the next section, the null distribution for the likelihood ratio can also be obtained by simulation. When this is performed using 1,000 constant-rate simulations, the resultant P value is identical to that obtained by comparison to the χ^2 ($P < .001$). This should not be surprising because, in this case, the simulated likelihood ratio distribution very closely resembles the χ^2 (fig. A3). Thus, even for quite small trees ($n = 16$, in this case) and few exceptional branches ($m^* = 5$), we can sometimes both reject a hypothesis of rate homogeneity when it is false and obtain reasonable parameter estimates for the evolutionary process.

Number of Exceptional Branches; Number of Taxa

To explore the performance of McPeck's (1995) method and the likelihood approach under broader conditions, I simulated rate-homogeneous and rate-heterogeneous evolution on constant-rate, pure-birth, stochastic phylogenies

both with the same total number of taxa ($n = 100$) and either fewer or more exceptional branches ($m^* = 1, 2, 10, 30, \text{ or } 99$, as opposed to $m^* = 20$) or with fewer total taxa ($n = 10, 20$, as opposed to $n = 100$). I performed the simulations under conditions otherwise identical to those described in "Simulation Tests."

Results from the simulations with fewer or more exceptional branches and fewer taxa are summarized in tables A1 and A2. In general, the likelihood method (table A1) and McPeck's method (table A2) performed similarly to the results for $n = 100$ taxa and $m^* = 20$ exceptional branches, with some caveats.

First, the likelihood method was overly conservative for small m^* and $n = 100$ and had lower power than McPeck's approach (tables A1, A2). This is probably due to the fact that the likelihood ratio test statistic is expected only to asymptotically approximate a χ^2 distribution, and this approximation is poor for small sample sizes. This is evident from a comparison of the likelihood ratio distributions from the simulations in which $m^* = 2$ with those obtained for larger m^* ($m^* = 20$; fig. A4). Results from $m^* = 20$ much better approximated a χ^2 (goodness-of-fit tests: $m^* = 2$, $\chi^2 = 136.1$, $df = 20$, $P < .001$; $m^* = 20$, $\chi^2 = 27.9$, $df = 20$, $P = .11$). This finding suggests that for very small m^* , an approach other than a χ^2 test on the likelihood ratio might be more appropriate. I suggest that a simulation-based hypothesis-testing approach be used for very small m^* because this restores the power of the likelihood method (table A1 for $m^* = 2$). The simulation approach can also be used for larger m^* with no loss of power (table A1 for $m^* = 20$).

Second, the rate-heterogeneity parameter (h in this study) is upwardly biased under several circumstances, in particular for very small and very large m^* . Actually, h is slightly upwardly biased for all values of m^* (table A1). The source of this upward bias is a positive skewness of the distribution of the MLEs for h , which has a very long positive tail for small and large m^* . This positive skewness of the distribution of $MLE(h)$ should not be surprising because $MLE(h)$ is a rate ratio, which is expected to be distributed as a variance ratio, a type of distribution with a very large positive skew (e.g., the F distribution). Log transformation can almost completely remove the bias in h for large m^* (fig. A5 for $h = 16.0$ and $m^* = 99$).

Third, on small trees, the likelihood approach and McPeck's method both suffered reduced power. However, the likelihood approach had power highly comparable to that for McPeck's method with the F -test and higher than that for McPeck's method as typically applied (t - or U -test). As above, when the null distribution for hypothesis testing was generated via numerical simulation, there was no reduction in power (table A1 for $m^* = 4$ and $n = 20$).

Consequences of Misspecifying Exceptional Branches

In this simulation study, the branches with exceptional rate are invariably known without error. However, in reality such information is seldom available in an empirical study. Instead, hypotheses for rate heterogeneity are produced by reconstructing history, a process performed with error.

To explore the consequences of occasionally incorrectly specifying the internodes with exceptional evolutionary rate, I stochastically relocated to an adjacent branch the perceived positions of the parts of the tree with exceptional rate. This was accomplished by relocating the perceived position of exceptional branches with a probability 0.2. This is equivalent to the situation in which the data have evolved by heterogeneous rate but the branches of the tree along which the rate was exceptional are known with 20% error, a substantial error rate. Exceptional branches were moved with equal probability to the internode preceding the parent or to either daughter. I performed the perturbation and reanalysis on the trees with $n = 100$ taxa and $m^* = 20$ exceptional branches.

The results of these analyses are summarized in table A3. In general, the consequences of misspecifying exceptional branches were to depress the estimated rate heterogeneity and to decrease the power (cf. table A1). No more-nefarious consequence of exceptional internode misspecification was uncovered by these analyses.

Heterogeneity in the Evolutionary Rate along Exceptional Branches

In this study, rate heterogeneity is assumed to consist of one evolutionary rate for unexceptional branches and a second homogeneous evolutionary rate for exceptional branches in the phylogenetic tree. To determine the effect of heterogeneity in the evolutionary rate among exceptional branches (as opposed to solely between unexceptional and exceptional branches), I performed the following procedure. I simulated heterogeneity of the evolutionary process by generating data by Brownian motion with either the rate σ_0^2 , along unexceptional branches, or the rate $h \cdot \varepsilon \cdot \sigma_0^2$, in which $h = 16.0$ and ε is a random rate-scaling factor drawn from a gamma

distribution with shape parameters $\alpha = 1.0, 2.0,$ or 4.0 and $\beta = 1/\alpha$. Under these conditions, the expected value for $h \cdot \varepsilon$ is h , and the expected variance of $h \cdot \varepsilon$ is h^2/α . Thus, small α corresponds to high heterogeneity in the rate among exceptional branches, whereas $\alpha = \infty$ corresponds to no variation in the rate on exceptional branches in the tree.

I conducted these simulations on 100 trees containing $n = 100$ taxa and $m^* = 10$ exceptional branches. These trees were the first of the 1,000 such trees used in prior simulations (see next section).

I then analyzed the resultant data using two different approaches. First, I fitted the same two-rate model described in "Simulation Tests." I also fitted a multirate model in which 11 rates were estimated (one for each of 10 exceptional branches and the final rate for all unexceptional branches). The multirate model was fitted in the same way as the two-rate model, except that instead of two \mathbf{C} matrices, there were 11 such matrices and a $k \times 1$ vector \mathbf{h} in lieu of the scalar rate-heterogeneity parameter h , in which $k = m^* = 10$ (the number of rates $- 1$). For each scenario, I compared model fit using a model-fitting criterion (the Akaike Information Criterion [AIC]; Akaike 1974).

The results were not as interesting as hoped. Under all simulation conditions, the two-rate model was almost invariably preferred over the multirate model. This is probably due to the large penalty incurred on the AIC for overparameterization. In fact, the only effect of high rate heterogeneity was to slightly increase the frequency with which the single-rate model was chosen. Results are summarized in table A4.

Low Rate on Exceptional Branches

In this study, rare exceptional branches invariably had a higher rate of evolution than common branches. However, detecting rate heterogeneity when the exceptional rate is high is not a limitation of the likelihood method. In order to illustrate this, I simulated rate heterogeneity in which the rare branches had a low rate. For symmetry with the prior simulations, I set $h = 1/16, 1/8, 1/4,$ and $1/2$. I used the same trees and number of exceptional branches ($n = 100$ and $m^* = 20$) as in "Simulation Tests." Results from the likelihood analyses are shown in table A5.

Literature Cited Only in the Appendix

- Akaike, H. 1974. A new look at the statistical model identification. *IEEE Transactions on Automatic Control* 6: 716–723.
- Chernoff, H. 1954. On the distribution of the likelihood ratio. *Annals of Mathematical Statistics* 25:573–578.

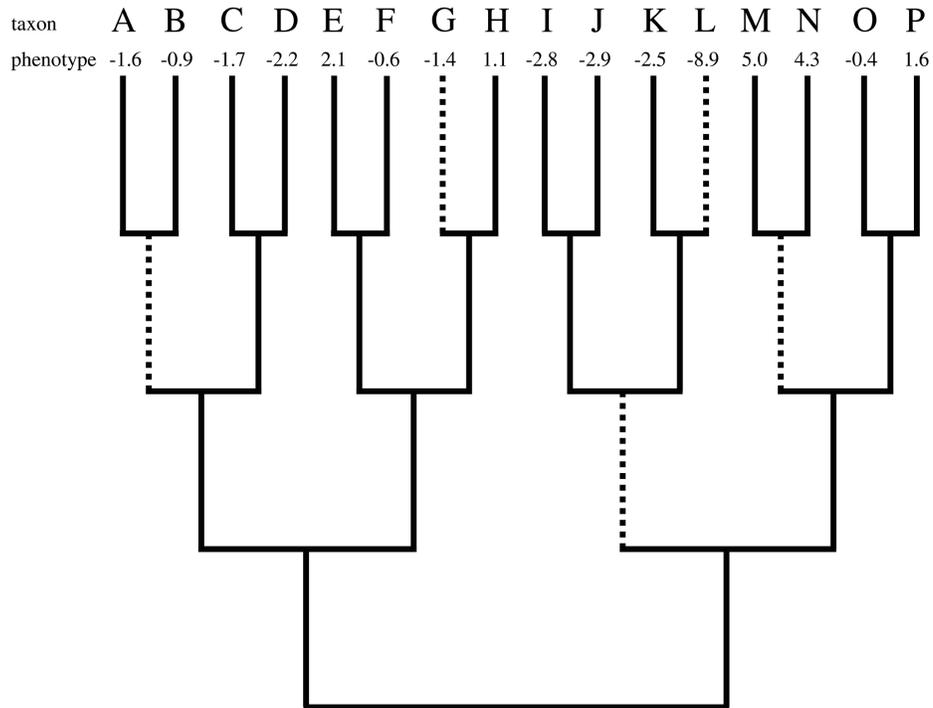


Figure A1: Hypothetical 16-taxon phylogenetic tree, branch lengths, and phenotypic trait values. Data were generated according to a rate-heterogeneous Brownian motion process in which the evolutionary rate was elevated along dashed branches.

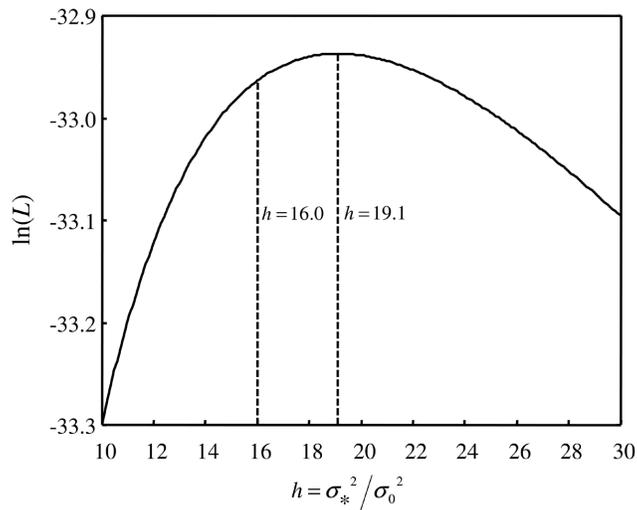


Figure A2: Likelihood surface for the rate heterogeneity parameter, h , defined in the text and in the abscissa label. Likelihoods were calculated for the data and tree in figure A1. The maximum likelihood estimate of h (19.1) and its true value (16.0) are indicated by vertical dashed lines.

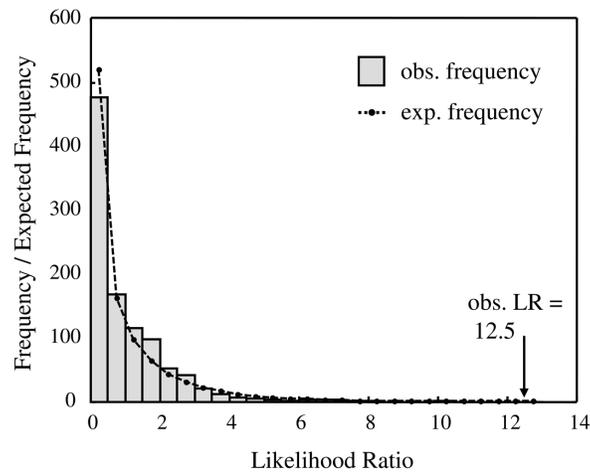


Figure A3: Likelihood ratio frequency distribution ($-2 \ln(L_1/L_2)$) from 1,000 simulations compared with a χ^2 distribution with 1 df.

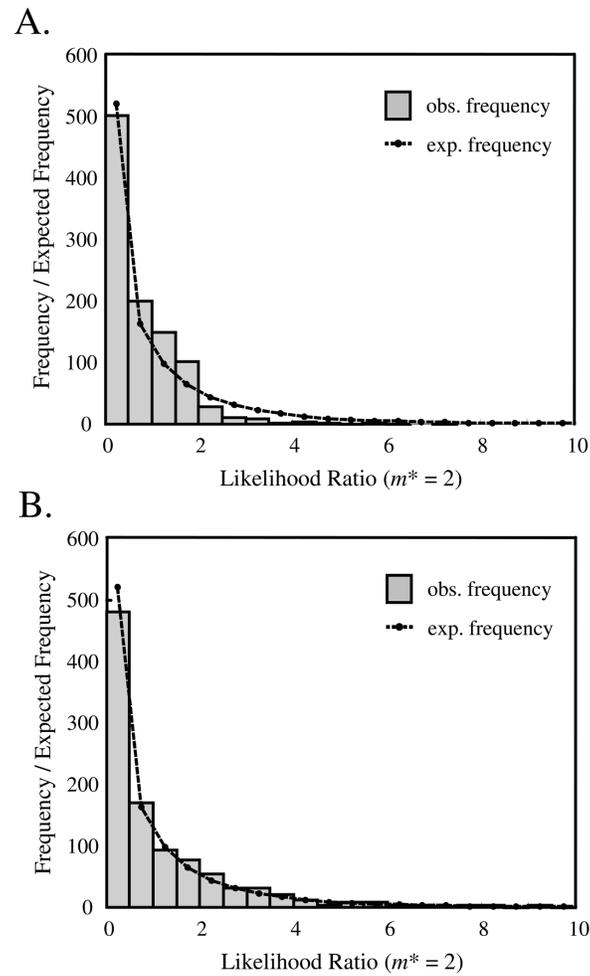


Figure A4: Likelihood ratio frequency distributions from simulations with $m^* = 2$ exceptional branches (A) and $m^* = 20$ exceptional branches (B). Both are compared to a χ^2 distribution with 1 df.

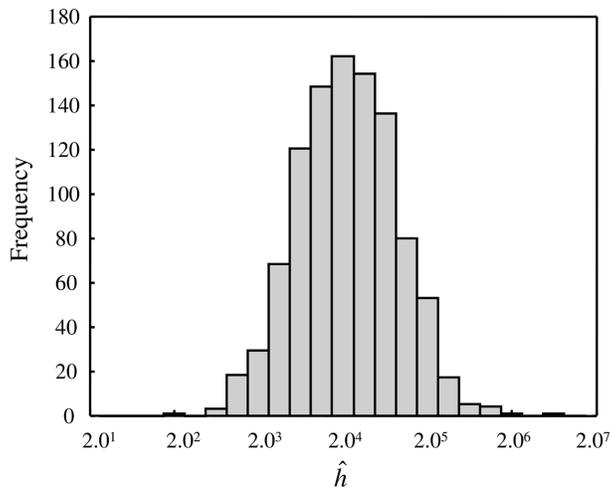


Figure A5: Frequency distribution for the maximum likelihood estimates of the evolutionary rate heterogeneity parameter, h . The true value of h is 16.0, around which the \log_2 distribution is centered. For these simulations, $n = 100$ and $m^* = 99$, as discussed in "Number of Exceptional Branches; Number of Taxa."

Table A1

Summary of results for analyses using the likelihood method

h	$\overline{\text{MLE}(h)}$	Type I error / power	Type I error / power (simulation) ^a
$n = 100, m^* = 1:$			
1.00	3.49	.005	...
2.00	3.28	.035	...
4.00	4.89	.079	...
8.00	8.38	.174	...
16.00	17.40	.344	...
$n = 100, m^* = 2:$			
1.00	1.20	.006	.08 ^b
2.00	2.38	.048	.12 ^b
4.00	4.93	.169	.33 ^b
8.00	8.44	.370	.46 ^b
16.00	16.55	.574	.71 ^b
$n = 100, m^* = 10:$			
1.00	1.11	.065	...
2.00	2.04	.150	...
4.00	4.15	.531	...
8.00	8.07	.874	...
16.00	16.50	.985	...
$n = 100, m^* = 20:$			
1.00	1.08	.049	.04 ^b
2.00	2.07	.262	.31 ^b
4.00	4.24	.787	.67 ^b
8.00	8.14	.984	.97 ^b
16.00	16.95	1.000	1.00 ^b
$n = 100, m^* = 30:$			
1.00	1.07	.068	...
2.00	2.08	.323	...
4.00	4.20	.893	...

Table A1 (Continued)

h	$\overline{\text{MLE}}(\bar{h})$	Type I error / power	Type I error / power (simulation) ^a
8.00	8.43	.997	...
16.00	16.70	1.000	...
$n = 100, m^* = 99:$			
1.00	1.09	.053	...
2.00	2.14	.454	...
4.00	4.51	.939	...
8.00	9.03	1.000	...
16.00	18.01	1.000	...
$n = 10, m^* = 2:$			
1.00	2.75	.030	...
2.00	3.79	.071	...
4.00	7.32	.163	...
8.00	13.96	.336	...
16.00	26.16	.509	...
$n = 20, m^* = 4:$			
1.00	1.45	.021	.08 ^b
2.00	2.93	.110	.14 ^b
4.00	5.09	.267	.38 ^b
8.00	10.17	.535	.51 ^b
16.00	19.97	.762	.86 ^b

Note: All results from 1,000 trees and data sets unless otherwise indicated. $\overline{\text{MLE}}(\bar{h})$ is the mean maximum likelihood estimate of the rate ratio.

^a An ellipse indicates that simulation results are not available.

^b Results from 100 simulations on each of the first 100 trees.

Table A2

Summary of results for analyses using McPeck's method

h	Type I error / power (t -test) ^a	Type I error / power (U -test) ^b	\bar{F}	Type I error / power (F -test)
$n = 100, m^* = 1:$				
1.00	.047 ^c	.022 ^d	.966	.048
2.00	.094 ^c	.046 ^d	1.429	.096
4.00	.181 ^c	.079 ^d	2.269	.186
8.00	.304 ^c	.161 ^d	3.783	.311
16.00	.453 ^c	.335 ^d	8.329	.458
$n = 100, m^* = 2:$				
1.00	.173	.055 ^d	.996	.044
2.00	.135	.067 ^d	1.471	.130
4.00	.114	.149 ^d	2.574	.286
8.00	.094	.244 ^d	4.526	.497
16.00	.102	.381 ^d	8.397	.671
$n = 100, m^* = 10:$				
1.00	.062	.065 ^d	1.057	.068
2.00	.066	.111 ^d	1.511	.240
4.00	.185	.315 ^d	2.517	.620
8.00	.400	.559 ^d	4.438	.913
16.00	.655	.790 ^d	8.456	.991
$n = 100, m^* = 20:$				
1.00	.047	.051	1.066	.064
2.00	.143	.166	1.572	.371
4.00	.480	.486	2.619	.840

Table A2 (Continued)

h	Type I error / power (t -test) ^a	Type I error / power (U -test) ^b	\bar{F}	Type I error / power (F -test)
8.00	.868	.822	4.554	.990
16.00	.980	.970	8.874	1.000
$n = 100, m^* = 30:$				
1.00	.065	.038	1.079	.080
2.00	.199	.197	1.599	.415
4.00	.715	.621	2.696	.935
8.00	.968	.924	4.879	.999
16.00	.999	.998	9.061	1.000
$n = 100, m^* = 99:$ ^c				
1.00
2.00
4.00
8.00
16.00
$n = 10, m^* = 2:$				
1.00	.092	. . . ^f	1.436	.047
2.00	.088	. . . ^f	2.057	.101
4.00	.091	. . . ^f	3.632	.227
8.00	.103	. . . ^f	6.847	.399
16.00	.133	. . . ^f	12.38	.569
$n = 20, m^* = 4:$				
1.00	.064	.039 ^d	1.170	.044
2.00	.062	.079 ^d	1.834	.167
4.00	.072	.147 ^d	2.923	.347
8.00	.090	.286 ^d	5.402	.608
16.00	.160	.450 ^d	10.07	.815

Note: All results from 1,000 trees and data sets. $\overline{MLE}(h)$ is the mean maximum likelihood estimate of the rate ratio. An ellipse indicates that results are not available or are not defined because of constraints on the test.

^a Assuming unequal variances unless otherwise indicated.

^b From normal approximation unless otherwise indicated.

^c Assuming equal variances.

^d From exact U -test.

^e McPeck's approach is not possible for $m^* = 99$ because of inevitable adjacency of exceptional branches.

^f Significant U -test is impossible for $n_1 = m^* = 2$ and $n_2 = n - m^* - 1 = 7$.

Table A3

Summary of results for analyses in which exceptional branches were misspecified with a probability of 0.2

h	$\overline{MLE}(h)$	Type I error / power
1.0	1.08	.051
2.0	1.86	.184
4.0	3.51	.638
8.0	6.32	.919
16.0	11.58	.989

Note: $\overline{MLE}(h)$ is the mean maximum likelihood estimate of the rate ratio.

Table A4

Consequence of rate heterogeneity among exceptional branches on the probability of accepting alternative models for the evolutionary process using the Akaike Information Criterion (AIC) to measure model fit

Rate heterogeneity α	Frequency of accepted model (using AIC)		
	One-rate	Two-rate	11-rate
1.0	.05	.95	.00
2.0	.01	.98	.01
4.0	.00	1.00	.00

Table A5

Results from analyses in which exceptional branches had a decreased, as opposed to an increased, rate

h	$\overline{\text{MLE}(h)}$	Power
1/16	.13	.648
1/8	.17	.539
1/4	.28	.384
1/2	.53	.158

Note: $\overline{\text{MLE}(h)}$ is the mean maximum likelihood estimate of the rate ratio.