doi:10.1111/j.1558-5646.2010.01160.x

THE LOCAL-CLOCK PERMUTATION TEST: A SIMPLE TEST TO COMPARE RATES OF MOLECULAR EVOLUTION ON PHYLOGENETIC TREES

Robert Lanfear[1,2]

[1]*Centre for Macroevolution and Macroecology, Ecology Evolution & Genetics, Research School of Biology, Australian National University, Canberra ACT 0200, Australia*

[2]*E-mail: rob.lanfear@anu.edu.au*

Received May 20, 2010
Accepted September 23, 2010

Rates of molecular evolution vary substantially between lineages, and a growing effort is directed at uncovering the causes and consequences of this variation. Comparing local-clocks (rates of molecular evolution estimated from different sets of branches of a phylogenetic tree) is a common tool in this research effort. Here, I show that a commonly used test (the Likelihood Ratio Test, LRT) will not be statistically valid for comparing local-clocks in most cases. Instead, I propose the local-clock permutation test (LCPT), a simple test that can be used to test the significance of differences between local-clocks. The LCPT could also be used to test for differences between any parameter that can be assigned to individual branches on a phylogenetic tree. Using simulated data, I show that the LCPT has good power to detect differences between local-clocks.

KEY WORDS: **Likelihood-ratio test, local-clock, molecular clock, molecular evolution, substitution rate.**

Understanding variation in rates of DNA and protein evolution is of fundamental interest in molecular evolutionary studies, and is important in refining phylogenetic and molecular dating algorithms. It is now widely accepted that DNA sequences rarely evolve in a strictly clock-like manner (Davies et al. 2004; Wright et al. 2006; Duffy et al. 2008; Smith and Donoghue 2008; Thomas et al. 2010). Indeed, some studies have shown that rates can vary dramatically even between closely related lineages (Nabholz et al. 2008, 2009).

One of the most common methods used to understand variation in rates of molecular evolution is the application of local-clocks (e.g., Yoder and Yang 2000; Bromham and Woolfit 2004; Aguileta et al. 2006; Lanfear et al. 2007; Korall et al. 2010; Neiman et al. 2010). A local-clock is a rate of molecular evolution estimated from a particular set of branches on a phylogenetic tree (e.g., the island lineages in Fig. 1). When more than one local-clock is estimated from a single phylogenetic tree, statistical tests can be used to ask whether the substitution rates represented by the local-clocks differ (e.g., Fig. 1). The most common statistical test used to compare local-clocks is the likelihood-ratio test (LRT). The LRT compares the Likelihoods of two nested models of evolution—the "local-clock" model in which separate local-clocks are estimated from sets of branches on a phylogenetic tree, and the "single-clock" model in which a single rate is estimated from the same sets of branches. If the LRT shows that the difference in likelihood between the two models is significant then the single-clock model can be rejected in favor of the local-clock model. A number of authors have interpreted the rejection of a single-clock model as evidence that the underlying mean rates of molecular evolution differ significantly between two sets of branches (see e.g., Bromham and Woolfit 2004; Lumbsch et al. 2008; Zhong et al. 2009; Korall et al. 2010; Neiman et al. 2010).

© 2010 The Author(s). *Evolution* © 2010 The Society for the Study of Evolution.

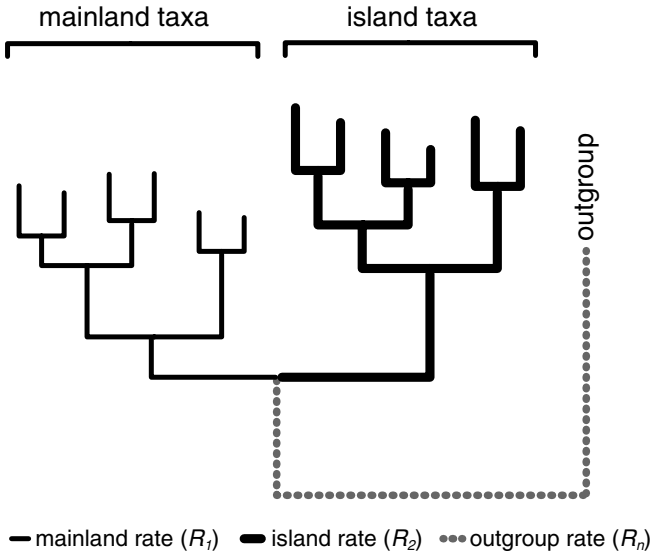

Figure 1. An example of a local-clock hypothesis to compare rates of evolution in mainland and island lineages. Separate local-clocks (i.e., substitution rate parameters) are estimated for the mainland lineages (thin black branches) and island lineages (thick black branches). A third "nuisance" substitution rate parameter is estimated for the outgroup lineage. Statistical tests can be used to compare the fit of a model in which the island and mainland rates are equal, versus one in which they are allowed to differ.

However, this interpretation of the LRT will not usually be statistically valid (although the conclusion that rates differ may not always be incorrect). This is because the single-clock model assumes equal rates in all branches to which all local-clocks are assigned. Therefore, rejection of the single-clock model implies only that there is some variation in rates among branches, but not necessarily that the local-clocks differ significantly. Because of this, a significant result from the LRT should only be interpreted as evidence for a significant difference in local-clocks if there is no variation in rates within either of the local-clocks (i.e., if evolution is locally clock-like). Unfortunately, evolution is unlikely to be locally clock-like in most cases, because of the large number of factors can influence rates of molecular evolution (Bromham 2009), and so the LRT will not usually be an appropriate test with which to compare local-clocks.

In this study, I describe a simple permutation test (the local-clock permutation test, LCPT) that can be used to ask whether substitution rates estimated from different local-clocks differ significantly. The LCPT and the LRT make different assumptions about rates of molecular evolution, and thus test different null hypotheses. The LRT assumes that there is no variation in the rate of molecular evolution within each local-clock, and thus tests the null hypothesis that rates of molecular evolution are identical in all branches to which all local-clocks are assigned. In contrast, the LCPT uses permutations of local-clocks to account for variation in rates within each local clock, and thus tests the null hypothesis that the mean rate of evolution of two local-clocks is equal. That is, the LRT will detect variation in rates both within and between local-clocks, but the LCPT will only detect rate variation between local-clocks. I use simulated data to demonstrate that the LCPT can be used to reliably detect small differences in local-clocks, and to show that the use of the LRT is rarely an appropriate test with which to compare local-clocks.

THE LCPT

The LCPT involves random permutations of the branches to which local-clocks are assigned on phylogenetic trees. For simplicity, I consider here only those cases pertaining to overall substitution rates, but the test would apply equally to analyses involving any parameters that can be applied to individual branches of a phylogenetic tree. As with all local-clock approaches the test applies to those cases in which one has an a priori hypothesis that a given parameter may differ in magnitude among different sets of branches on a phylogenetic tree (Yoder and Yang 2000). The a priori hypothesis is defined by assigning different local-clocks (i.e., substitution rate parameters) to the appropriate branches in the phylogenetic tree under consideration (see e.g., Fig. 1). A third "nuisance" rate parameter should be assigned to branches that have no bearing on the a priori hypothesis (e.g., outgroup lineages), or branches to which neither of the two local-clocks can be assigned with confidence (e.g., Fig. 1).

I use the following notation (following Goldman et al. 2000) to describe the test. T denotes the predefined phylogenetic tree under consideration. The two local-clocks are denoted "R_1" and "R_2," and the nuisance rate parameter assigned to all other branches in the tree "R_n" (e.g., Fig. 1). The vector of assignments of rate parameters (R_1, R_2, and R_n) to the branches of T is denoted θ_{obs}. L_{single} is the maximized log-likelihood of θ_{obs} on T when R_1 are R_2 assumed to be equal (the single-clock hypothesis). L_{local} is the maximized log-likelihood of θ_{obs} on T when R_1 are R_2 are allowed to differ (the local-clock hypothesis). A permutation of θ_{obs} in which the assignments of R_1 and R_2 to the branches of T are randomized is denoted θ_i. L^i_{local} is the maximized log-likelihood of θ_i on T when R_1 and R_2 are allowed to differ.

The procedure for the LCPT is as follows:

- Calculate the test-statistic $\Delta_{obs} = L_{local} - L_{single.}$
- Create i permutations of θ_{obs} such that the R_1 and R_2 assignments are shuffled randomly among the branches to which either was originally assigned in θ_{obs}. The number of branches to which R_1 and R_2 are assigned should be kept constant in every permutation, and R_n should always be assigned to the same branches as it was in θ_{obs}. This step can be conducted using a python script available from the author.

- Calculate L^i_{local} for each θ_i on topology *T*.
- Calculate values of $\Delta_i = L^i_{\text{local}} - L_{\text{single}}$.
- Test whether Δ_{obs} (from the a priori local-clock hypothesis) is a plausible sample of the null distribution given by the Δ_i values by testing whether it falls in the confidence interval given by the upper 5% of the ranked list of Δ_i values. The upper 5% of the ranked list is appropriate (assuming a 5% significance level) because we wish to test whether our a priori hypothesis gives a larger improvement in Likelihood than would be expected by chance alone.

Materials and Methods

SIMULATING RATE VARIATION

The following simulations are designed to capture two important aspects of substitution rate analyses using local-clocks: (1) variation between local-clocks, i.e., variation in rates which is associated with differences in the sets of branches to which the local-clocks are assigned (described below as "between-clock variation"); and (2) variation within local-clocks, i.e., variation in rates which occurs within branches to which a single local-clocks is assigned (described below as "within-clock variation"). For instance, in Figure 1, the between-clock variation is the component of the total variation in rates associated with the difference between island and mainland lineages, and the within-clock variation is the remainder of the variation in rates seen within these lineages. Within-clock variation could result from of the difficulties of accurately measuring substitution rates (Cutler 2000), or from variation in other aspects of species biology which is not captured by the a priori local-clock hypothesis.

Both between- and within-clock rate variation were simulated on phylogenetic trees generated according to the Yule process using BEAST (Drummond and Rambaut 2007). All simulated trees were constrained to produce 16 taxa. The root height of each tree was rescaled to 0.6 substitutions per site in TreeEdit version 1.0a10 (Rambaut and Charleston 2001). This value corresponds to a substitution rate of 1% per million years over a time period of 60 million years, which is roughly equivalent to that expected in broad-scale analyses of mammalian DNA sequences.

Between-clock rate variation was simulated by multiplying either five or 10 randomly chosen branch lengths by a constant rate multiplier. Nine rate multipliers were simulated—from 1.0 (i.e., no between-clock rate variation, local-clocks are equal) to 3.0 (between-clock rate variation in which one local-clock represents a rate three times higher than the other local-clock), in increments of 0.25.

Three levels of within-clock rate variation were simulated by multiplying the branch lengths by random numbers drawn from uniform distributions with means of 1.0 (such that the total tree length remained approximately unchanged) and different ranges. Zero within-clock rate variation was simulated by leaving the branch lengths unchanged. A low level of within-clock rate variation was simulated by multiplying all branch lengths by numbers drawn from $U(0.8, 1.2)$. A high level of within-clock rate variation was simulated by multiplying all branch lengths by numbers drawn from $U(0.5, 1.5)$.

In summary, 54 simulation conditions were considered: between-clock rate variation applied to either five or 10 branches, nine levels of between-clock rate variation, and three levels of within-clock rate variation. All 54 of these conditions were applied to 100 randomly generated Yule trees (see above), giving a total of 5400 simulated phylogenetic trees. For each tree, a DNA dataset of 1000 bp was simulated using the evolver program of the PAML 4.1 package (Yang 2007) under the HKY model of molecular evolution (Hasegawa et al. 1985), with kappa fixed at 2, equal base frequencies, and constant rates of evolution among sites.

LIKELIHOOD CALCULATIONS ON SIMULATED DATASETS

All likelihoods were calculated using the baseml program of the PAML version 4.1 package (Yang 2007). The tree topology and model of molecular evolution were set to match the simulation conditions. For each dataset L_{local} was calculated by assigning one rate parameter (R_2) to the branches of the tree that had been simulated to have increased substitution rates (see above), and another rate parameter (R_1) to the same number of randomly chosen branches from the rest of the tree. Note that the decision to assign R_1 and R_2 to an equal number of branches is arbitrary in these simulations—the choice of branches to which each rate parameter is assigned would usually be dictated by the particular a priori local-clock hypothesis under consideration. A third rate parameter, R_n, was assigned to all other branches of the tree. L_{single} was calculated by constraining R_1 and R_2 to be equal.

LRTs were performed by calculating twice the difference in log likelihood between the single- and local-clock models, and comparing this to a chi-squared distribution with one degree of freedom. LCPTs were performed as described above, using 1000 random permutations of local-clocks. All input and output files, and python scripts used to perform the simulations and the LCPT, are available from the author, or at http://datadryad.org/.

Results

WITHIN-CLOCK RATE VARIATION IS DETECTED BY THE LRT BUT NOT THE LCPT

Table 1 shows the proportion of significant results returned by the LRT and LCPT when there was no between-clock rate variation (i.e., the simulated mean rates of molecular evolution in

Table 1. The proportion of significant results returned by the likelihood-ratio test (LRT) and local-clock permutation test (LCPT) when there is no between-clock rate variation, but levels of within-clock rate variation differ. Each proportion is calculated from 100 replicates (see main text). The 95% confidence intervals (shown in brackets) are calculated from the binomial distribution.

Within-clock rate variation	Proportion of significant results (95% CI)			
	Five branches per local-clock		10 branches per local-clock	
	LRT	LCPT	LRT	LCPT
Zero: $U(1.0, 1.0)$	0.13 (0.07–0.20)	0.04 (0.01–0.08)	0.14 (0.08–0.21)	0.04 (0.01–0.08)
Medium: $U(0.8, 1.2)$	0.16 (0.09–0.23)	0.09 (0.04–0.15)	0.20 (0.12–028)	0.04 (0.01–0.08)
High: $U(0.5, 1.5)$	0.44 (0.34–0.54)	0.09 (0.04–0.15)	0.40 (0.31–0.50)	0.04 (0.01–0.08)

the branches to which local-clocks were assigned were identical), but levels of within-clock rate variation differed. The proportion of significant results returned by the LRT increases as the amount of within-clock rate-variation increases, up to about 40% when within-clock rate variation is high, regardless of the number of branches to which each local-clock is assigned (Table 1). In contrast, the proportion of significant results returned by the LCPT does not change as within-clock rate variation increases. Indeed, regardless of the level of within-clock rate variation or the number of branches to which each local-clock is assigned, the 95% confidence intervals of the proportion of significant LCPT results always encompasses 0.05 (Table 1). These results demonstrate that the LCPT is not sensitive to within-clock rate variation, and that it has acceptable false-positive rates when used to compare local-clocks.

Table 1 also shows that the LRT incorrectly identifies variation in substitution rates even when the data were simulated under truly clock-like conditions (i.e., zero between- or within-clock rate variation). In both truly clock-like simulation conditions, the proportion of significant results returned by the LRT was above 0.05, and the 95% confidence intervals did not include 0.05. This is likely to be a result of the difficulty of estimating substitution rates from finite datasets—even if rates are truly clock-like, some variation in rates is likely to be reconstructed as a result of the Poisson nature of the substitution process (Cutler 2000).

THE LCPT CAN DETECT SMALL DIFFERENCES IN LOCAL-CLOCKS

Figure 2 shows the proportion of significant results returned using the LCPT at different levels of within- and between-clock rate variation. Figure 2 shows that the power of the LCPT to detect differences in local-clocks increases as the number of branches to which each local-clock is assigned increases, and as the amount of between-clock variation increases. As expected, the power of the LCPT decreases as the level of within-clock rate variation increases. In the best case (each local-clock assigned to 10 branches, zero within-clock rate variation, Fig. 2B), the LCPT correctly identifies an underlying difference in local-clock substitution rates in the majority of cases when there is at least a 1.25-fold difference in substitution rates (i.e., where $R_2 \geq 1.25 \times R_1$). In the worst case (each local-clock assigned to five branches, high levels of within-clock rate variation, Fig. 2E), the LCPT correctly identifies an underlying difference in local-clock substitution rates in the majority of cases when there is at least a 2.25-fold difference in rates (i.e., where $R_2 \geq 2.25 \times R_1$).

The LRT was also performed on all simulated datasets, and uniformly returned a higher proportion of significant results than the LCPT. This difference is an expected consequence of the fact that the LRT and the LCPT test different null hypotheses. The LRT tests the null hypothesis that rates of evolution are identical in all branches to which all local-clocks are assigned, and so will be sensitive to both within- and between-clock variation in substitution rates. The use of the LRT in this study, in which single- and local-clock models are compared, was designed to demonstrate that the LRT will not usually be appropriate for comparing local-clocks. However, the LRT is an appropriate test for detecting deviations from a single-clock (as opposed to comparing rates estimated from local-clocks), and will often be most powerful in this situation when the likelihood of a single-clock model is compared to that of a free-rates model (in which a separate rate parameter is assigned to each branch). In contrast to the LRT, the LCPT tests the null hypothesis that the mean rates of the two local-clocks are equal, and is sensitive only to between-clock variation in substitution rates.

Discussion

As the availability of DNA sequence data increases, and as phylogenetic and molecular dating methods become more advanced, there is a growing interest in uncovering the causes and correlates of rates of molecular evolution (Lanfear et al. 2010). Two common goals in this research effort are to detect variation in rates of molecular evolution, and compare rates of molecular evolution in different clades, or among different sets of branches of a phylogenetic tree (e.g., Fig. 1).

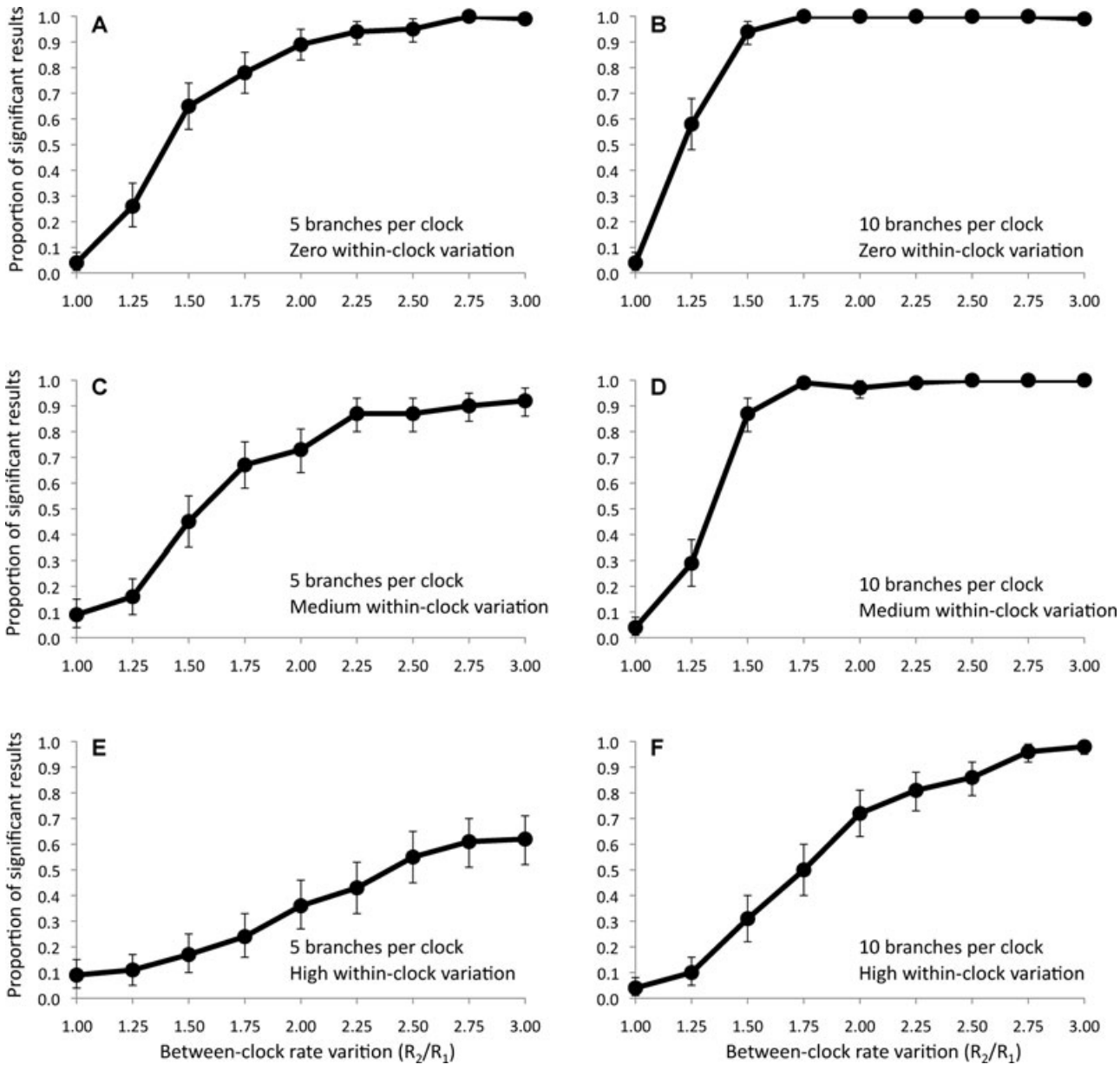

Figure 2. **The ability of the LCPT to correctly determine differences in substitution rates at different levels of between-clock and within-clock rate variation. The *y*-axis shows the proportion of significant results calculated from 100 randomly generated replicates datasets (see main text). The *x*-axis shows the degree of between-clock rate variation, calculated as the ratio of the faster rate to the slower rate. Three levels of within-clock rate variation were tested: zero variation (A, B), medium variation (C, D), and high variation (E, F) (see main text for details). Between-clock rate variation was applied to either five (A, C, E) or 10 (B, D, F) branches of each tree. Error bars represent 95% confidence intervals, calculated from the binomial distribution.**

The results of this study clarify that while the LRT is useful and powerful test with which to detect variation in rates of molecular evolution (see e.g., Table 1), it is not usually an appropriate test with which to compare substitution rates estimated from local-clocks. This is because the LRT will often (and correctly) return significant results even when there is no difference in mean rate between local-clocks (Table 1). Put another way, the use of the LRT to compare local-clocks will only be valid when rate variation occurs only between, but not within, local-clocks. Given the increasing evidence that variation in rates of molecular evolution is ubiquitous (Davies et al. 2004; Wright et al. 2006; Duffy et al. 2008; Smith and Donoghue 2008; Thomas et al. 2010), it seems that the LRT will rarely be an appropriate test with which to compare substitution rates estimated from local-clocks.

In this study, I introduce a new test, the LCPT that can be used to compare substitution rates estimated from local-clocks. Simulations show that in contrast to the LRT, the LCPT is not sensitive to variation in substitution rates within local-clocks (Table 1),

and can detect relatively small differences in substitution rates between local-clocks (Fig. 2). Because of this, the LCPT can used to test whether the mean substitution rates of local-clocks differ significantly. The LCPT is most powerful when each local-clock is assigned to a large number of branches, and when those local-clocks explain a high proportion of the total variation in substitution rates in the dataset (Fig. 2). Given the ever-increasing availability of molecular sequence data, the increasing sophistication of models of molecular evolution, and the high observed variation in substitution rates, the LCPT should be a useful test with which to detect differences in substitution rates using local-clocks.

ACKNOWLEDGMENTS

I would like to thank two anonymous reviewers for improving the quality of the manuscript, and to S. Ho and M. Phillips for commenting on earlier drafts. RL is funded by the Australian Research Council.

LITERATURE CITED

Aguileta, G., J. P. Bielawski, and Z. Yang. 2006. Evolutionary rate variation among vertebrate beta globin genes: implications for dating gene family duplication events. Gene 380:21–29.

Bromham, L. 2009. Why do species vary in their rate of molecular evolution? Biol. Lett. 5:401–404.

Bromham, L., and M. Woolfit. 2004. Explosive radiations and the reliability of molecular clocks: island endemic radiations as a test case. Syst. Biol. 53:758–766.

Cutler, D. J. 2000. Understanding the overdispersed molecular clock. Genetics 154:1403–1417.

Davies, T. J., V. Savolainen, M. W. Chase, J. Moat, and T. G. Barraclough. 2004. Environmental energy and evolutionary rates in flowering plants. Proc. Biol. Sci. 271:2195–2200.

Drummond, A. J., and A. Rambaut. 2007. BEAST: Bayesian evolutionary analysis by sampling trees. BMC Evol. Biol. 7:214.

Duffy, S., L. A. Shackelton, and E. C. Holmes. 2008. Rates of evolutionary change in viruses: patterns and determinants. Nat. Rev. Genet. 9:267–276.

Goldman, N., J. P. Anderson, and A. G. Rodrigo. 2000. Likelihood-based tests of topologies in phylogenetics. Syst. Biol. 49:652–670.

Hasegawa, M., H. Kishino, and T. Yano. 1985. Dating of the human-ape splitting by a molecular clock of mitochondrial DNA. J. Mol. Evol. 22:160–174.

Korall, P., E. Schuettpelz, and K. M. Pryer. 2010. Abrupt deceleration of molecular evolution linked to the origin of arborescence in ferns. Evolution 64:2786–2792.

Lanfear, R., J. A. Thomas, J. J. Welch, T. Brey, and L. Bromham. 2007. Metabolic rate does not calibrate the molecular clock. Proc. Natl. Acad. Sci. USA 104:15388–15393.

Lanfear, R., J. J. Welch, and L. Bromham. 2010. Watching the clock: studying variation in rates of molecular evolution between species. Trends Ecol. Evol. 25:495–503.

Lumbsch, H. T., A. L. Hipp, P. K. Divakar, O. Blanco, and A. Crespo. 2008. Accelerated evolutionary rates in tropical and oceanic parmelioid lichens (Ascomycota). BMC Evol. Biol. 8:257.

Nabholz, B., S. Glemin, and N. Galtier. 2008. Strong variations of mitochondrial mutation rate across mammals—-the longevity hypothesis. Mol. Biol. Evol. 25:120–130.

———. 2009. The erratic mitochondrial clock: variations of mutation rate, not population size, affect mtDNA diversity across birds and mammals. BMC Evol. Biol. 9:54.

Neiman, M., G. Hehman, J. T. Miller, J. M. Logsdon, Jr., and D. R. Taylor. 2010. Accelerated mutation accumulation in asexual lineages of a freshwater snail. Mol. Biol. Evol. 27:954–963.

Rambaut, A., and M. Charleston. 2001. TreeEditv1.0a10 Carbon, Oxford, Evolutionary Biology Group, University of Oxford.

Smith, S. A., and M. J. Donoghue. 2008. Rates of molecular evolution are linked to life history in flowering plants. Science 322:86–89.

Thomas, J. A., J. J. Welch, R. Lanfear, and L. Bromham. 2010. A generation time effect on the rate of molecular evolution in invertebrates. Mol. Biol. Evol. 27:1173–1180.

Wright, S., J. Keeling, and L. Gillman. 2006. The road from Santa Rosalia: a faster tempo of evolution in tropical climates. Proc. Natl. Acad. Sci. USA 103:7718–7722.

Yang, Z. 2007. PAML 4: phylogenetic analysis by maximum likelihood. Mol. Biol. Evol. 24:1586–1591.

Yoder, A. D., and Z. Yang. 2000. Estimation of primate speciation dates using local molecular clocks. Mol. Biol. Evol. 17:1081–1090.

Zhong, B., T. Yonezawa, Y. Zhong, and M. Hasegawa. 2009. Episodic evolution and adaptation of chloroplast genomes in ancestral grasses. PloS ONE 4:e5297.

Associate Editor: D. Posada