Molecular Clock: Testing

Fabia Ursula Battistuzzi, *Arizona State University, Tempe, Arizona, USA*

Alan J Filipski, *Arizona State University, Tempe, Arizona, USA*

Sudhir Kumar, *Arizona State University, Tempe, Arizona, USA*

Based in part on the previous version of this Encyclopedia of Life Sciences (ELS) article, Molecular Clock: Testing by Sudhir Kumar and Alan J Filipski.

Advanced article

Article Contents

Online posting date: 15th July 2011

The molecular clock hypothesis originally rested on the assumption of rate constancy across lineages of a phylogeny, which would produce an approximately steady rate of accumulation of deoxyribonucleic acid or amino acid changes through time. This assumption has been questioned on the basis of increasingly large data sets, which have shown significant variability of rates in evolutionary lineages. To address this issue, tests have been developed to examine whether rates of molecular evolution vary significantly among taxonomic groups or phylogenetic lineages. Two major types of tests exist: those based on comparisons of genetic distances and those based on likelihood ratios. The first ones compare genetic distances between two species (or groups of species) relative to an outgroup; the latter ones compare maximum likelihood values for the same phylogeny calculated with and without the constant rate assumption. In those cases where the rate constancy assumption is violated, modern molecular clocks (relaxed clocks) are now being applied to implement the rate heterogeneity in the time estimation process.

Introduction

In its simplest form, 'Molecular clock' refers to the approximately steady rate of accumulation of changes in deoxyribonucleic acid (DNA) (or protein amino acid) sequences over evolutionary time. A consequence of this rate constancy is a nearly linear relationship between evolutionary distance and time of species divergence. The hypothesis applies only to orthologous sequences of a given gene; evolutionary rates are known to vary widely from gene to gene at the protein sequence level and also in different types of DNA (e.g. mitochondrial versus nuclear). **See also**: Evolution: Tempo and Mode; Molecular Evolution: Introduction; Mutation Rates: Data

eLS subject area: Evolution & Diversity of Life

How to cite:
Battistuzzi, Fabia Ursula; Filipski, Alan J; and Kumar, Sudhir (July 2011) Molecular Clock: Testing. In: eLS. John Wiley & Sons, Ltd: Chichester.
DOI: 10.1002/9780470015902.a0001803.pub2

Margoliash (1963) and Zuckerkandl and Pauling (1965) appear to have been the first to observe that the rates of amino acid substitution in some genes are roughly the same in several lineages. This led to the hypothesis of steady accumulation of amino acid substitutions at the molecular level (see Kumar, 2005 for a historical review). Its utility for estimating times of species divergence was immediately recognised, especially for organisms that left few or no traces in the fossil record. For example, divergence times were estimated between prokaryotes and eukaryotes and among eukaryotic kingdoms using molecular clocks (McLaughlin and Dayhoff, 1970; Dickerson, 1971). **See also**: Molecular Clocks; Molecular Evolution; Semantides and Modern Bacterial Systematics

As more molecular sequence data became available, the initial optimism about the universality of the molecular clock was questioned by reports of significant difference in evolutionary rates among species in some genes and in some lineages. For instance, hominids (humans and their close relatives) seemed to be evolving much more slowly than other mammals, whereas rodents (mouse and rat) were shown to be evolving much faster (reviewed in Nei and Kumar, 2000). Many of these results have been controversial because they were based on species divergence times inferred from the fossil record, which provides only a lower bound on the true time since two species diverged. Easteal *et al*. (1995) demonstrate that the use of the fossil record-based dates may be the reason for the observed deviations from the molecular clock in some of these controversies. Nonetheless, these rate differences can affect the accuracy of the molecular estimates and should be taken into consideration when applying a molecular clock. **See also**: Fossil Record; Fossils in Phylogenetic Reconstruction

Thus, direct comparison of evolutionary rates, without requiring outside knowledge of the evolutionary divergence time, is desirable. The outcome of such comparisons will be the discovery of either a constant or a heterogeneous rate across lineages of a phylogeny. Divergence times for sequences belonging to these two alternative scenarios can then be estimated with different approaches, that is, using a strict clock in the first case and a relaxed clock in the second. We will first address the nature of these rate tests and strategies to maximise the inclusion of rate constant sequences in a phylogeny. Second, we will address rate variability models and tests in the case of analyses of data containing sequences that have evolved with unequal rates.

Basic principle of rate testing

Fitch (1976) proposed a test for statistically examining whether the observed difference in evolutionary rates between two sequences is significantly greater than that expected by chance. This test used sequences of a gene from three species. For two of these species (**Figure 1**) the molecular clock is being tested (A and B) and the third is used as an outgroup (C). **See also**: Molecular Evolution: Patterns and Rates; Molecular Evolution: Rates

Let N_{ij} be the number of sequence differences (DNA or protein) between species i and j. In this case, the lengths of the branches ending at A and B (l_A and l_B, respectively) are given by eqns [1] and [2].

$$l_A = \tfrac{1}{2}(N_{AB} + N_{AC} - N_{BC}) \quad [1]$$

$$l_B = \tfrac{1}{2}(N_{AB} + N_{BC} - N_{AC}) \quad [2]$$

Since both A and B evolved from a common ancestor T years ago (**Figure 1**), the time elapsed on each lineage is the same. Therefore, testing the difference between l_A and l_B is equivalent to directly testing the difference in evolutionary rates between lineages A and B. In Fitch (1976) formulation, this difference is tested using a chi-square test with one degree of freedom (eqn [3], where $l_{avg} = \frac{1}{2}(l_A + l_B)$).

$$x^2 = (l_A - l_{avg})^2/l_{avg} + (l_B - l_{avg})^2/l_{avg} \quad [3]$$

This equation can be simplified to eqn [4].

$$x^2 = (l_A - l_B)^2/(l_A + l_B) \quad [4]$$

Although this test illustrates the basic principle behind molecular clock testing, it is no longer used because many more sophisticated and/or powerful relative rate tests have now become available. In the following, we discuss a variety of molecular clock tests, including the relative rate and phylogenetic tests. A detailed account of methods for these and other types of tests can be found elsewhere (e.g. Nei and Kumar, 2000).

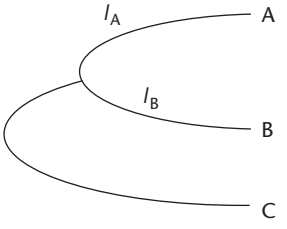

Figure 1 Relative rate test for two species A and B using a third species, C, as an outgroup. Here, l_A and l_B are the branch lengths measured in the number of substitutions or the number of substitutions per site from the common ancestor of A and B.

Relative Rate Tests

Testing rate differences between two species (or sequences)

Under the molecular clock hypothesis for lineages A and B, l_A should be equal to l_B in **Figure 1**. Therefore, the expected value of the quantity $D = (l_A - l_B)$ is zero. Use of the outgroup species C allows us to compute branch lengths l_A and l_B given the evolutionary distances d_{ij}, using eqns [5] and [6].

$$l_A = \tfrac{1}{2}(d_{AB} + d_{AC} - d_{BC}) \quad [5]$$

$$l_B = \tfrac{1}{2}(d_{AB} + d_{BC} - d_{AC}) \quad [6]$$

In general, D is rarely exactly zero because evolutionary processes are stochastic in nature. We therefore need to examine whether the observed D is significantly different from 0. For this purpose, Wu and Li (1985) proposed the use of the Z-test (eqn [7]).

$$Z = |D|/\sqrt{V(D)} \quad [7]$$

where $V(D)$ is the variance of D. The variance of D can be estimated analytically, as was done by Wu and Li, or by using a bootstrap resampling technique. In the analytical method, we need to compute variances and covariances for the distances d_{ij}. On the other hand, bootstrap resampling, whereas computationally intensive, is the simpler of the two methods and has some advantages over the analytical method (see Nei and Kumar, 2000). Suppose that the sequences are n sites long, with x_{ij} representing the ith site of sequence j (**Figure 2**). Each site i may be regarded as a column of values, one from each sequence. A bootstrap replicate in this case refers to a resampled data set consisting of a set of three sequences (each of length n) obtained by sampling n columns with replacement from the original set. For example, if column i is selected as the kth pick of the resampling procedure for a given replicate, then the value of the kth site of each sequence in the replicate set is set to the original x_i value of that sequence. Hundreds of such replicate sets are created and the value of D is computed for each replicate. Let D_b be the value of D for the bth replicate. Then, the bootstrap variance is computed using eqn [8].

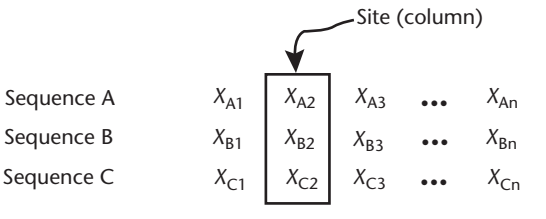

Figure 2 Notations used for explaining the bootstrap resampling of columns for a three-sequence case; n is the number of sites (columns) and x_{ij} is the ith site of the jth sequence.

$$V(D) = \frac{1}{B-1}\sum_{b=1}^{B}(D_b - D)^2 \quad [8]$$

where B is the number of bootstrap replications, D_b is the value of D estimated in the bth bootstrap replication, and is the average of the D_b. **See also**: Molecular Phylogeny Reconstruction

If the Z statistic in eqn [7], computed using the bootstrap or the analytical method, is greater than or equal to 1.96, then the molecular clock hypothesis can be rejected at the 5% significance level. This test is applicable for relative rate analysis using any kind of evolutionary distance measure for DNA or protein sequences. **See also**: DNA Sequence Analysis

Likelihood ratio test

Muse and Weir (1992) constructed a likelihood-ratio test, which is parallel in functionality to the method described above. Given a model of nucleotide sequence evolution and a specific tree (topology), it is possible to compute the maximum likelihood of generating a specific set of sequences at the terminal nodes of the tree. The maximum likelihood value for the tree in **Figure 2** can be computed with and without the condition $l_A = l_B$. The latter corresponds to the molecular clock hypothesis. Muse and Weir used the Hasegawa model of nucleotide substitution for computing the maximum likelihood of obtaining a given data set under these two scenarios (Hasegawa *et al*., 1985). If the maximum likelihood for the molecular clock case is L_c and that for the more general case is L, then the likelihood ratio (LR) test statistic is given by eqn [9].

$$LR = 2|\ln L - \ln L_c| \quad [9]$$

LR is distributed approximately as a chi-square random variable with one degree of freedom. There is only one degree of freedom because the molecular clock assumption imposes only one additional condition ($l_A = l_B$). If LR is 3.84 or higher then the molecular clock hypothesis can be rejected at the 5% level.

Although the Muse–Weir test (Muse and Weir, 1992) was described for Hasegawa's model, their likelihood framework can be used under any model of substitution for DNA and protein sequences.

Nonparametric tests

Both the Wu–Li and Muse–Weir tests are parametric tests because they require information on the model of evolutionary change. In contrast, the Fitch (1976) test described above is nonparametric. Tajima and others have proposed more powerful versions of the Fitch test. Tajima (1993) presents a general framework for testing the molecular clock hypothesis for both DNA and protein sequences. These tests are based on the expected equality of the number of unique mutations that have occurred in the lineages A and B (**Figure 1**). Let us consider a site in sequences A, B and C. This site has undergone a unique mutation in lineage A if the amino acid residue or nucleotide base at that site in sequence A is different from that in B and C, which are identical to each other. Let us denote the number of sites showing a unique mutation in A by m_A. Similarly, we compute the number of sites that have experienced a unique mutation in lineage B by m_B. Under the molecular clock hypothesis, we expect $m_A = m_B$. This expectation can be tested by means of a chi-square test with one degree of freedom.

$$\chi^2 = (m_A - m_B)^2/(m_A + m_B) \quad [10]$$

Note that when m_A and m_B are small, it is better to use Fisher's exact test since the chi-square approximation may not be appropriate. The same nonparametric method can be applied to DNA sequences. The above equation is similar to eqn [4] for the Fitch (1976) test, the only difference being that eqn [4] uses all mutations in each lineage, rather than the unique mutations only. This property makes the Fitch test less powerful in rejecting the null hypothesis when it is false, as the nonunique mutations contribute only to the denominator in eqn [4] making the chi-square value smaller than that computed using eqn [10].

Testing rate differences between two groups of species (or sequences)

When we are interested in comparing evolutionary rates differences between two groups of species, it is desirable to include data available from all species in the relative rate analysis to improve the power of the test. Relative rate tests for two clusters of sequences are essentially identical to those for two sequences, the difference being that A, B and C in **Figure 1** now represent groups containing one or more sequences each. Therefore, we need to extend the two-sequence tests for use in multiple sequence analysis. This generalisation was described by Takezaki *et al*. (1995) for the Wu and Li (1985) method. In this case, we compute $D = l_A - l_B$ by using average distances ($d_{ij,\mathrm{avg}}$) between groups of sequences (eqns [11] and [12]).

$$l_A = \tfrac{1}{2}(d_{AB,\mathrm{avg}} + d_{AC,\mathrm{avg}} - d_{BC,\mathrm{avg}}) \quad [11]$$

$$l_B = \tfrac{1}{2}(d_{AB,\mathrm{avg}} + d_{BC,\mathrm{avg}} - d_{AC,\mathrm{avg}}) \quad [12]$$

$d_{ij,\mathrm{avg}} = \sum d_{ij}/c_i c_j$, c_i is the number of sequences in group i and c_j is the number of sequences in group j. We can then use the Z-test given in eqns [11] and [12] to test the molecular clock hypothesis that the expected value $D = (l_A - l_B) = 0$.

Phylogenetic Tests of the Molecular Clock

The above tests are simple in that they compare rates between two lineages, given an outgroup species. It is also possible to test the molecular clock on multiple lineages

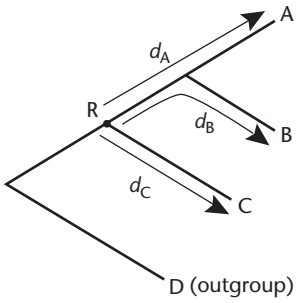

Figure 3 Phylogenetic tree to illustrate the root-to-tip molecular clock tests. The rate along the branch to D is not tested, but serves to locate the root R of the subtree containing species A, B and C for which the molecular clock test is being conducted.

simultaneously in a phylogenetic tree. Some of these tests identify the anomalous groups or lineages, whereas some merely test an entire tree for conformity to the hypothesis. To apply these tests, we need a rooted phylogenetic tree for the given set of sequences (**Figure 3**). The root of the tree is obtained by using the outgroup sequences (D in **Figure 3**). However, the branch leading to the outgroup species is not itself testable. **See also**: Molecular Phylogeny Reconstruction

Least-squares residual sum test

The squared sum of residuals R for a given phylogenetic tree is given by eqn [13].

$$R = \sum_{i<j} (d_{ij} - e_{ij}) \qquad [13]$$

Here, j ranges over the number of sequences, d_{ij} is the observed distance between sequences i and j and e_{ij} is the patristic distance between the same pair. It can be regarded as a measure of the discrepancy between the observed sequence data and the hypothesised phylogenetic tree (Cavalli-Sforza and Edwards, 1967). If we compute this sum without the rate-constancy assumption (R) and then with the molecular clock assumption (R_C), then the following quantity (eqn [14]) follows the F distribution with $m-2$ and $\frac{1}{2}(m-2)(m-3)$ degrees of freedom (Felsenstein, 1988).

$$F = \frac{(R_C - R)/(m-2)}{R/\left[\frac{1}{2}m(m-1) - (2m-3)\right]} = \frac{(m-3)(R_C - R)}{2R} \qquad [14]$$

One should be careful when using this test, however, as it assumes that the distances d_{ij} are independently distributed, which is violated because of the evolutionary history shared by different sequences.

Global two-cluster test

Takezaki *et al.* (1995) devised a single statistic to test whether all pairs of sequence clusters in a phylogenetic tree satisfy a rate-balance criterion similar to that of the basic two-cluster test described above. In general, a tree with m sequences (not counting the outgroup) has $m-1$ internal nodes, each of which defines a pair of phylogenetic clusters. If A and B are the clusters defined by an internal node i, and C is the outgroup, then rate-constancy would require, as before, that the expected value of $y_i = (l_A - l_B) = 0$. To test all the y_i simultaneously, we first form the row vector $\mathbf{y} = [y_1, y_2, \ldots, y_{m-1}]$ and the variance–covariance matrix $\mathbf{V} = [v_{ij}]$, where v_{ij} is the covariance of y_i with y_j. Then the required statistic is as shown in [15], where $\mathbf{T}$ denotes transpose and -1 denotes the matrix inverse.

$$\mathbf{U} = \mathbf{y}\mathbf{V}^{-1}\mathbf{y}^{\mathbf{T}} \qquad [15]$$

Under the assumption that the y_i are distributed as multivariate normal, $\mathbf{U}$ is distributed as a chi-square random variable with $m-1$ degrees of freedom under the null hypothesis that the expected value of $\mathbf{y}$, $E(\mathbf{y}) = [0, 0, \ldots, 0]$. This provides a test of the hypothesis using standard chi-square tables.

Root-to-tip test

Let d_A, d_B and d_C be the depths to the root node, R, of the tree from the tips A, B and C, respectively; and let d be the average of d_A, d_B and d_C. These d_i are computed by adding the branch lengths from node R to the terminal node corresponding to sequence i. Hence, we refer to them as 'root-to-tip' distances. If the clock hypothesis holds, the expected value of $z_i = (d_i - d)$ is zero. This statistic can be used for two types of tests. First, we can examine species (or sequences) that have evolved with a rate different from the average rate (Uyenoyama, 1995; Takezaki *et al.*, 1995). For instance, for species A we can construct a standard Z-test by computing $z_A = (d_A - d)$ and its variance, $V(z_A)$. It is possible to derive an analytical formula for the variance, but it is easier to use bootstrapping, as mentioned earlier. In the bootstrap method, the variance $V(z_i)$ is given by eqn [16], where B is the number of bootstrap replications, z_{ib} is the value of z_i estimated at the bth bootstrap replication, and z_i is the average of the z_{ib}.

$$(z_i) = \frac{1}{B-1} \sum_{b=1}^{B} (z_{ib} - \bar{z}_i)^2 \qquad [16]$$

This process can be repeated for each sequence and all sequences that are evolving significantly slower or faster can be detected in this way.

Alternatively, we can construct a composite test in which all sequences are simultaneously tested and a single test statistic is obtained for the whole tree. This is accomplished in a way exactly analogous to the global two-cluster test described above, with y_i replaced by z_i.

Global likelihood ratio test

Given a phylogenetic tree of sequences and a model of transition probabilities from one amino acid or nucleotide to another, it is a standard procedure to compute the

likelihood associated with the tree. This is done independently for each site and requires summing over all possible (unknown) ancestral states at internal nodes of the tree. The overall likelihood of a tree is obtained, under the assumption of independent sites, by multiplying the likelihoods (or, equivalently, adding the log-likelihoods) for all sites. We can perform this computation while allowing rates to differ among branches or with the assumption of rate constancy. As in the basic LR test described previously, we may then perform a chi-square test with $m-2$ degrees of freedom on the statistic $LR = 2|\ln L - \ln L_C|$, where L is the unconstrained likelihood value and L_C is the value obtained under the rate-constancy assumption.

Local rates in relaxed molecular clocks

In the mid-nineteen eighties, Gillespie (1984) suggested that rates on closely related branches in a phylogeny (e.g. ancestors and descendants) are likely to be more similar to each other than to those on more distant parts of the tree. Following this hypothesis, the autocorrelation model of evolutionary rates was developed and implemented in the first relaxed clock methods (Sanderson, 1997; Thorne *et al.*, 1998). The concept of autocorrelation is motivated by the expectation that closely related lineages will share similar evolutionary and biochemical mechanisms; some of these (e.g. DNA repair mechanisms) directly affect the number of changes that accumulate in the DNA and, therefore, limit the variability of evolutionary rates between lineages. However, more recently, Drummond *et al.* (2006) have questioned the validity of this approach for sequences in phylogenies of both long and short timescales. In fact, when evolutionary distances are large the probability of sister species being exposed to similar selective pressures decreases and, therefore, new biochemical mechanisms will evolve resulting in potentially different evolutionary rates. At short distances, instead, the stochastic nature of the accumulation of differences between sequences plays a stronger role than that of similar inherited factors, once again decreasing the autocorrelation effect. Therefore, uncorrelated changes in evolutionary rates may be better models (Drummond and Rambaut, 2007; Yang, 2007).

Statistically distinguishing between autocorrelation and uncorrelation in evolutionary rates is important because relaxed clock methods assume either one of these models during the time estimation process. Simulation studies have shown that the accuracy of these clocks decreases significantly when a clock assuming an incorrect model of rate evolution is used (Battistuzzi *et al.*, 2010). Therefore, it is desirable to have tests to identify the best-fitting evolutionary rate model to a set of DNA or amino acid sequences. Unfortunately, the available tests are known to be not very powerful or reliable. The most common test, implemented in programs with a Bayesian framework (e.g. PhyloBayes and BEAST; Lartillot and Philippe, 2004, 2006; Drummond and Rambaut, 2007; Lartillot *et al.*, 2007), relies on the calculation of Bayes Factors (BF), which correspond to the ratio of the marginal likelihoods calculated assuming alternatively an autocorrelated or an uncorrelated model of evolutionary rates across lineages (eqn [17]).

$$BF = \frac{\Pr(D|M_1)}{\Pr(D|M_2)} \qquad [17]$$

where D is the data and M is the model of evolutionary rate changes. Increasingly positive BF values would support with greater confidence the use of model 1 (M_1), while increasing negative values the use of model 2 (M_2).

This approach, however, obtains contrasting results based on the taxonomic sampling, the depth of the phylogeny, and the gene selection (Drummond *et al.*, 2006; Lepage *et al.*, 2007; Brown *et al.*, 2008; Ho, 2009). The limitations of this test result in a potential violation of the clock assumption, when a single rate model is chosen. It has been proposed that the inaccuracy produced by such a scenario could be at least partially remediated by using both rate models on the same data set and drawing conclusions on the combination of the results (Battistuzzi *et al.*, 2010).

Molecular clock timing

When the results of the rate tests show constancy across lineages, a strict (also called global) molecular clock may be applied. This strictly uses the assumption of a constant rate across the phylogenetic tree for each gene analysed to estimate divergence times; however, it does not make any assumption on rate constancy across genes (for a review see Hedges and Kumar, 2003).

However, most sequences analysed with the rate tests discussed above will show signs of heterogeneity in their evolutionary rates, limiting the application of the strict molecular clock principle. In these cases there are two possible approaches: one is to purge the violating lineages from the data set so that the remaining ones are rate constant (Takezaki *et al.*, 1995). However, this approach limits the applicability of a molecular clock because lineages with different evolutionary histories will rarely have rates equal to those of their close relatives and, therefore, will not be timed in most data sets. A second approach is to loosen the rate constancy constraint of the original molecular clock methods by using local or relaxed clocks (Uyenoyama, 1995; Takezaki *et al.*, 1995; Sanderson, 1997; Yang, 1997, 2007; Thorne *et al.*, 1998; Drummond and Rambaut, 2007; Yang and Rannala, 2006). Local clocks assume an a priori knowledge of rate differences across lineages or groups and are, therefore, difficult to apply. Relaxed clocks, instead, allow rates to evolve from ancestor to descendant lineages without needing a priori knowledge of rates and are currently the most commonly used methods. **See also**: Molecular Clocks; Molecular Clocks; Molecular Clocks: Determining the Age of the Human–Chimpanzee Divergence; Molecular Evolution: Rates

Conclusions

The idea of a clock-like substitution fits well with the neutral theory of molecular evolution. However, rather than expecting broad acceptance or rejection of the principle, it seems better to treat the molecular clock as a hypothesis to be tested for each gene and lineage. For this purpose, several statistical tests for rate constancy among two or more sequences or groups of sequences have been developed, some of which are presented above. If a particular gene satisfies the strict molecular clock tests over a wide number of lineages, then it may be relied upon to help establish dates for phylogenetic events using the simple constant-rate-of-evolution model. On the other hand, if some lineages show high variation, this fact is interesting in itself and may point the way to a need for biological explanations. Before one can apply a strict clock to this type of data set, one must either remove those genes and/or lineages that violate the rate constancy assumption or use a relaxed clock model to accommodate the variation. In this latter case, however, caution should be used in choosing the model (autocorrelation or uncorrelation) that best represents the empirical changes in evolutionary rates among lineages. It must also be borne in mind that more complex models with more parameters inevitably lead to greater variance of estimates. As more and more sequence data becomes available, careful application of the molecular clock principle should allow us to make best use of a limited amount of palaeontological data to provide dates for many important events in the history of life (see a review in Hedges and Kumar, 2009). **See also**: Molecular Clocks; Molecular Evolution: Neutral Theory; Neutrality and Selection in Molecular Evolution: Statistical Tests

References

Battistuzzi FU, Filipski A, Hedges SB and Kumar S (2010) Performance of relaxed-clock methods in estimating evolutionary divergence times and their credibility intervals. *Molecular Biology and Evolution* **27**: 1289–1300.

Brown JW, Rest JS, Garcia-Moreno J, Sorenson MD and Mindell DP (2008) Strong mitochondrial DNA support for a cretaceous origin of modern avian lineages. *BMC Evolutionary Biology* **6**: 6.

Cavalli-Sforza L and Edwards A (1967) Phylogenetic analysis: models and estimation procedures. *American Journal of Human Genetics* **19**: 233–257.

Dickerson R (1971) The structure of cytochrome *c* and the rates of molecular evolution. *Journal of Molecular Evolution* **1**: 26–45.

Drummond AJ, Ho SY, Philips MJ and Rambaut A (2006) Relaxed phylogenetics and dating with confidence. *PloS Biology* **4**: e88.

Drummond AJ and Rambaut A (2007) BEAST: Bayesian evolutionary analysis by sampling trees. *BMC Evolutionary Biology* **7**: 214.

Easteal S, Collet C and Betty D (1995) *The Mammalian Molecular Clock*. Austin, TX: RG Landes.

Felsenstein J (1988) Phylogenies from molecular sequences: inference and reliability. *Annual Review of Genetics* **22**: 521–565.

Fitch W (1976) Molecular evolutionary clocks. In: Ayala FJ (ed.) *Molecular Evolution*, pp. 160–178. Sunderland, MA: Sinauer Associates.

Gillespie JH (1984) The molecular clock may be an episodic clock. *Proceedings of the National Academy of Sciences of the USA* **81**: 8009–8013.

Hasegawa M, Kishino H and Yano T (1985) Dating the human–ape splitting by a molecular clock of mitochondrial DNA. *Journal of Molecular Evolution* **22**: 160–174.

Hedges SB and Kumar S (2003) Genomic clocks and evolutionary timescales. *Trends in Genetics* **19**: 200–206.

Hedges SB and Kumar S (2009) *The Timetree of Life*. New York: Oxford University Press.

Ho SYW (2009) An examination of phylogenetic models of substitution rate variation among lineages. *Biology Letters* **5**: 421–424.

Kumar S (2005) Molecular clocks: four decades of evolution. *Nature Reviews Genetics* **6**: 654–662.

Lartillot N, Brinkmann H and Philippe H (2007) Suppression of long-branch attraction artefacts in the animal phylogeny using a site-heterogeneous model. *BMC Evolutionary Biology* **7**: S4.

Lartillot N and Philippe H (2004) A Bayesian mixture model for across-site heterogeneities in the amino-acid replacement process. *Molecular Biology and Evolution* **21**: 1095–1109.

Lartillot N and Philippe H (2006) Computing Bayes factors using thermodynamic integration. *Systematic Biology* **55**: 195–207.

Lepage T, Bryant D, Philippe H and Lartillot N (2007) A general comparison of relaxed molecular clock methods. *Molecular Biology and Evolution* **24**: 2669–2680.

Margoliash E (1963) Primary structure and evolution of cytochrome *c*. *Proceedings of the National Academy of Sciences of the USA* **50**: 672–679.

McLaughlin P and Dayhoff M (1970) Eukaryotes versus prokaryotes: an estimate of evolutionary distance. *Science* **168**: 1469–1471.

Muse S and Weir B (1992) Testing for equality of evolutionary rates. *Genetics* **132**: 269–276.

Nei M and Kumar S (2000) *Molecular Evolution and Phylogenetics*. New York: Oxford University Press.

Sanderson M (1997) A nonparametric approach to estimating the divergence times in the absence of rate constancy. *Molecular Biology and Evolution* **14**: 1218–1231.

Tajima F (1993) Simple methods for testing molecular clock hypothesis. *Genetics* **135**: 599–607.

Takezaki N, Rzhetsky A and Nei M (1995) Phylogenetic test of the molecular clock and linearized tree. *Molecular Biology and Evolution* **12**: 823–833.

Thorne JL, Kishino H and Painter IS (1998) Estimating the rate of evolution of the rate of molecular evolution. *Molecular Biology and Evolution* **15**: 1647–1657.

Uyenoyama M (1995) A generalized least squares estimate of the origin of sporophytic self-incompatibility. *Genetics* **139**: 975–992.

Wu C-I and Li W-H (1985) Evidence for higher rates of nucleotide substitution in rodents than in man. *Proceedings of the National Academy of Sciences of the USA* **82**: 1741–1745.

Yang ZH (1997) PAML: a program package for phylogenetic analysis by maximum likelihood. *Computer Applications in the Biosciences* **13**: 555–556.

Yang ZH (2007) PAML 4: phylogenetic analysis by maximum likelihood. *Molecular Biology and Evolution* **24**: 1586–1591.

Yang ZH and Rannala B (2006) Bayesian estimation of species divergence times under a molecular clock using multiple fossil calibrations with soft bounds. *Molecular Biology and Evolution* **23**: 212–226.

Zuckerkandl E and Pauling L (1965) Evolutionary divergence and convergence in proteins. In: Bryson V and Vogel H (eds) *Evolving Genes and Proteins*, pp. 97–166. New York: Academic Press.

Further Reading

Drummond AJ and Suchard MA (2010) Bayesian random local clocks, or one rate to rule them all. *BMC Biology* **8**.

Ho SYW (2009) An examination of phylogenetic models of substitution rate variation among lineages. *Biology Letter* **5**: 421–424.

Koonin EV (2009) Darwinian evolution in the light of genomics. *Nucleic Acids Research* **37**: 1011–1034.

Nei M, Suzuki Y and Nozawa M (2010) The neutral theory of molecular evolution in the genomic era. *Annual Review of Genomics and Human Genetics* **11**: 265–289.

Pagel M and Meade A (2008) Modelling heterotachy in phylogenetic inference by reversible-jump Markov chain Monte Carlo. *Philosophical Transactions of the Royal Society B* **364**: 3955–3964.

Yang ZH (2006) *Computational Molecular Evolution*. Oxford: Oxford University Press.