Efficient Methods for Dating Evolutionary Divergences

12

Qiqing Tao, Koichiro Tamura, and Sudhir Kumar

Abstract

Reliable estimates of divergence times are crucial for biological studies to decipher temporal patterns of macro- and microevolution of genes and organisms. Molecular sequences have become the primary source of data for estimating divergence times. The sizes of molecular data sets have grown quickly due to the development of inexpensive sequencing technology. To deal with the increasing volumes of molecular data, many efficient dating methods are being developed. These methods not only relax the molecular clock and offer flexibility to use multiple clock calibrations, but also complete calculations much more quickly than Bayesian approaches. Here, we discuss the theoretical and practical aspects of these non-Bayesian approaches and present a guide to using these methods effectively. We suggest that the computational speed and reliability of non-Bayesian relaxed-clock methods offer opportunities for enhancing scientific rigour and reproducibility in biological research for large and small data sets.

Keywords

Molecular dating · Strict clock · Local clock · Calibration · Maximum likelihood · RelTime

Q. Tao
Institute for Genomics and Evolutionary Medicine, Temple University, Philadelphia, PA, USA

Department of Biology, Temple University, Philadelphia, PA, USA
e-mail: qiqing.tao@temple.edu

K. Tamura
Center for Genomics and Bioinformatics, Tokyo Metropolitan University, Tokyo, Japan

Department of Biological Sciences, Tokyo Metropolitan University, Tokyo, Japan
e-mail: ktamura@tmu.ac.jp

S. Kumar (✉)
Institute for Genomics and Evolutionary Medicine, Temple University, Philadelphia, PA, USA

Department of Biology, Temple University, Philadelphia, PA, USA

Center for Excellence in Genome Medicine and Research, King Abdulaziz University, Jeddah, Saudi Arabia
e-mail: s.kumar@temple.edu

12.1 Introduction

Computational methods to estimate divergence times of genes and species from molecular data have enjoyed a long history of development, spanning more than 50 years (dos Reis et al. 2016; Kumar and Hedges 2016). Divergence times derived by using these methods and molecular data have illuminated the role of geological history in shaping the emergence of species (Hedges et al. 1996; Hedges and Kumar 2009), tempo and mode of speciation (Hedges et al. 2015; Marin et al. 2017), dynamics of genome evolution through gene duplication

S. Y. W. Ho (ed.), *The Molecular Evolutionary Clock*, https://doi.org/10.1007/978-3-030-60181-2_12

(Huerta-Cepas and Gabaldón 2011; Jiao et al. 2011; Yu et al. 2017), and evolution of pathogens (Faria et al. 2014; Worobey et al. 2014; Biek et al. 2015; Metsky et al. 2017). Every year, hundreds of studies report estimates of species divergence times, enabling the assembly of the grand time-tree of life and revealing the fundamental biological processes underlying species diversity (Hedges et al. 2015).

Early statistical methodologies of molecular clock dating (Zuckerkandl and Pauling 1962) were based on the assumption of a constant rate of evolution over time and across lineages (strict clock) and used fossil-age calibrations as point values (Kumar 2005). Over the last two decades, molecular dating methods have become increasingly sophisticated and embrace greater biological realism. They now relax the strict-clock assumption and have the ability to estimate divergence times even when molecules have evolved with vastly different evolutionary rates across loci and lineages (Ho 2014; Ho and Duchêne 2014; Kumar and Hedges 2016). Many modern approaches are also available to incorporate detailed information from the fossil record to generate time-calibrated phylogenies (time-trees).

Statistical development of molecular dating methods remains vibrant even after six decades of development. It is at the centre of systematics, biodiversity, and genome evolution research owing to the ease with which large sequence data sets can now be assembled (Kumar and Hedges 2016). Chronologies of molecular dating methods and their statistical properties have been presented in recent years (Kumar 2005; Ho 2014; Ho and Duchêne 2014; Kumar and Hedges 2016; dos Reis et al. 2016). Therefore, here we focus on a more pragmatic account of molecular dating methods, aimed at assisting researchers to select and utilize available methods.

12.1.1 Non-Bayesian Versus Bayesian Methods

Increased sophistication of molecular dating methods has often been accompanied by increased demand for computational time and memory. There exists a clear dichotomy of molecular dating methods based on their computational resource requirements for large data sets. Bayesian methods are computationally demanding because of their need for extensive sampling from the posterior distribution using the Markov chain Monte Carlo (MCMC) approach (Bromham et al. 2018). The computational burden is usually very high for large data sets and grows with the number of sequences (Crosby and Williams 2017; Tamura et al. 2018). In addition, problems in MCMC mixing can increase the computational time further (Bhatnagar et al. 2011). Sometimes, there is a need to run multiple Bayesian analyses to test different prior assumptions and calibration settings, which might result in the requirement for high-performance computing infrastructure.

In contrast, many non-Bayesian methods tend to have much smaller computational needs, while still allowing rates to vary throughout the tree. For example, both penalized likelihood (Sanderson 2002) and RelTime (Tamura et al. 2012, 2018) are very fast and known to be accurate. Although their computational requirements increase linearly with the number of sequences and sites, they still take orders of magnitude less time than the Bayesian methods (Fig. 12.1). Computational time demands of these non-Bayesian methods are essentially the same as the time taken to estimate branch lengths of a phylogeny, for example by using the maximum-likelihood method. Non-Bayesian methods can also be applied directly to a phylogeny with branch lengths (phylogram), which decreases the computational times further for very large data sets.

In this chapter, our focus is on providing practitioners with a guide to effectively using non-Bayesian methods for molecular dating. We also discuss the advantages and disadvantages of these methods, because the best approach depends on the size of the available data, degree of rate variation among species and loci, nature of clock calibrations, and the availability of computing resources. Table 12.1 shows a summary of different non-Bayesian methods, their statistical properties, and the software packages in which they are implemented.

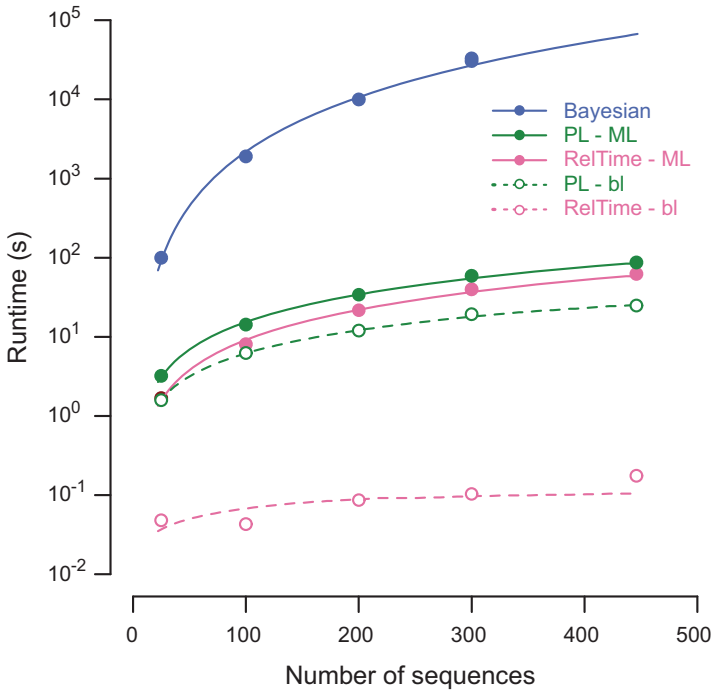

Fig. 12.1 Computational times required by Bayesian, penalized likelihood (PL), and RelTime methods to estimate divergence times for data sets containing an increasing number of sequences (n). Bayesian (blue solid line) is the computational time of the Bayesian method using molecular sequences as input. PL-bl (green dashed line) and RelTime-bl (pink dashed line) are the computational times of PL and RelTime methods using phylogenetic trees with branch lengths as input. PL-ML (green solid line) and RelTime-ML (pink solid line) are the total computational times required by PL and RelTime methods using molecular sequences as input, which are the sum of the computational time of maximum likelihood (ML) inferences of branch lengths and the computational time of PL-bl and RelTime-bl. ML inferences of branch lengths were conducted in MEGA X (Kumar et al. 2018). Bayesian, PL, and RelTime analyses were conducted in MCMCTree (Yang 2007), treePL (Smith and O'Meara 2012), and MEGA X, respectively. All times were estimated on a single-core computer by using an alignment of 4493 sites that was simulated with extensive rate variation (RR50 from Tamura et al. 2012). For this data set, the best-fit exponential equation is $0.06 \times n^{2.28}$, $0.08 \times n^{1.16}$, $0.07 \times n^{0.97}$, $0.03 \times n^{1.27}$, and $0.01 \times n^{0.44}$ for Bayesian, PL-ML, PL-bl, RelTime-ML, and RelTime-bl, respectively

12.2 A Practical Guide to Selecting Non-Bayesian Methods

12.2.1 Using Strict- and Local-Clock Methods

The simplest scenario for molecular dating is when the evolutionary rates are the same (or very similar) across different evolutionary lineages. In this case, methods that assume a strict clock will usually be reliable and produce the most precise time estimates. This assumption was commonly employed in the earliest molecular dating studies that produced many fundamental biological insights, including the finding that humans shared a most recent common ancestor with chimpanzees only five million years (Myr) ago, rather than 15–20 Myr ago based on the classification of *Homo* as a sister group to apes in the early 1960s (Sarich and Wilson 1967, 1973).

Interestingly, methods based on the strict clock continue to be developed and used today. For example, the mean path length (MPL) method (Britton et al. 2002), implemented in the software PATHd8 (Britton et al. 2007), has been used in many recent studies (e.g., Louca et al. 2018; Lu

Table 12.1 A summary of available efficient non-Bayesian dating methods

Software	Statistical basis[a]	Clock type[b]	Calibration type[c]	Confidence interval	References
Lintre	Regression	SC	F	Bootstrap	Takezaki et al. (1995)
PATHd8	MPL	SC	F, B	Bootstrap	Britton et al. (2007)
DAMBE	LS	SC, LC, RC	F, B, S	Bootstrap	Xia and Yang (2011), Xia (2018a)
r8s	LF, NPRS, PL	SC, LC, DC, RC	F, B, S	Bootstrap	Sanderson (1997, 2002, 2003)
treePL	PL	SC, RC	F, B	Likelihood	Smith and O'Meara (2012)
Ape—chronos & chronoMPL	PL, MPL	SC, DC, RC	F, B	Bootstrap	Paradis (2013)
MEGA X—RelTime, RTDT	RRF	SC, RC	F, B, D, R, S	Analytical	Kumar et al. (2018), Tamura et al. (2018), Tao et al. (2019), Miura et al. (2020)
TipDate	Regression	SC	S	Likelihood	Rambaut (2000)
TREBLE	UPGMA	SC	S	Bootstrap	Yang et al. (2007)
Physher	ML	SC, LC, DC	S	Bootstrap	Fourment and Holmes (2014)
LSD	LS	SC, RC	S	Bootstrap	To et al. (2016)
treedater	LS, ML	SC, RC	S	Bootstrap	Volz and Frost (2017)
TreeTime	ML	SC, RC	S	Likelihood	Sagulenko et al. (2018)

[a]*MPL* mean path length, *LS* least squares, *LF* Langley–Fitch method (Langley and Fitch 1974), *NPRS* nonparametric rate-smoothing, *PL* penalized likelihood, *ML* maximum likelihood, *RRF* relative-rate framework
[b]*SC* strict clock, *LC* local multi-rate clock, *DC* discrete multi-rate clock, *RC* relaxed clock
[c]*F* fixed node calibration, *B* node calibration boundary, *D* node calibration density, *R* substitution rate, *S* sampling tip date

et al. 2018). This method assumes that the ratio of ages between two nodes in a phylogeny is proportional to the ratio of their average node-to-tip distances. Therefore, it is only suitable for data sets in which the substitution rates are strictly or nearly constant among lineages throughout the phylogeny (Britton et al. 2002).

The problem of the equal-rates assumption is illustrated in the analysis of a phylogeny consisting of two clades (X and Y) with an outgroup (Fig. 12.2a). Each clade contains two orthologous DNA sequences of zinc-finger genes *zfx* and *zfy*; this is a simple gene-family tree with two genes that arose from a gene duplication prior to the divergence of human and mouse. Molecular dating methods should produce the same values for t_X and t_Y because they refer to the same evolutionary event: the divergence between human and mouse. Therefore, the expected ratio of t_X and t_Y is 1, which is what a molecular dating method should produce despite the fact that mouse *zfx* gene has evolved four times more quickly than the human *zfx*, and the mouse *zfy* gene has evolved seven times more quickly than the human *zfy*.

Analysis of this data set by the MPL approach in the PATHd8 software produced a $t_X/t_Y = 0.43$, which is much smaller than 1. It estimated that the divergence between human and mouse in clade Y (t_Y) happened much earlier than the same event in clade X (t_X) (Fig. 12.2b). This result is clearly inconsistent with the phylogenetic tree in Fig. 12.2a and shows that strict-clock methods produce biologically incorrect results if they are used for data sets in which evolutionary rates vary extensively among lineages. Smith and O'Meara (2012) have also reported that PATHd8 was not so reliable in analyses of empirical data sets and simulated data sets when evolutionary rates varied. Another least-squares method (Xia and Yang 2011) minimizes the residual sum of squares of patristic distance and distance computed by the rate and time under the global clock. This method, implemented in the DAMBE software (Xia 2018a), also produced an incorrect date ratio of 0.37 (Fig. 12.2c).

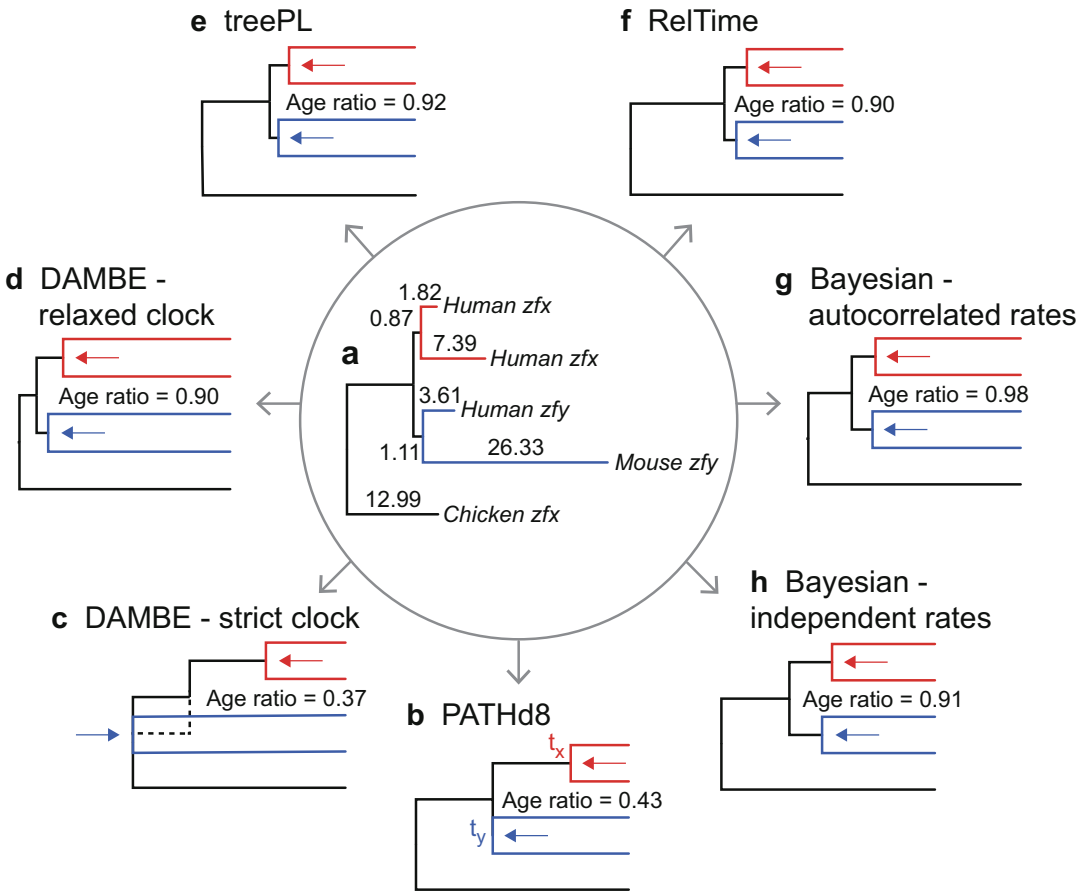

Fig. 12.2 Molecular dating analysis of four DNA sequences. (**a**) An example phylogeny of orthologous DNA sequences of two zinc-finger genes (GenBank accession numbers gi296010876, gi113205066, gi223890138, gi156938288, and gi363728820). The branch lengths are shown in substitutions per 100 base pairs. This is an excellent test case because the expected time for human–mouse species divergence based on gene *zfx* (t_X, red clade) and *zfy* (t_Y, blue clade) should be the same ($t_X/t_Y = 1$), as the gene duplication event occurred prior to the diversification of mammals. Shown are the time-trees produced by (**b**) PATHd8, (**c**) DAMBE with strict clock, (**d**) DAMBE with relaxed clock, (**e**) treePL, (**f**) RelTime, (**g**) MCMCTree (Bayesian) with the autocorrelated branch-rates model, and (**h**) MCMCTree (Bayesian) with the independent branch-rates model. Ratios of node ages for human–mouse divergence based on *zfx* (t_X, red arrow) and *zfy* (t_Y, blue arrow) genes of all resulting time-trees are labelled. One root calibration was used in PATHd8, DAMBE, treePL, and Bayesian analyses. No calibrations were used in the RelTime analysis

Therefore, the use of strict-clock methods is appropriate only if lineages have evolved with a strictly or nearly constant rate. The simplest way to ensure that this condition is valid is to conduct a molecular clock test. An early molecular clock test was proposed by Fitch (1976) for data sets containing two sequences and an outgroup, and was followed by many others (Wu and Li 1985; Muse and Weir 1992; Tajima 1993). For larger data sets, equality of rates on multiple lineages can be evaluated by least squares (Takezaki et al. 1995) and by likelihood-ratio tests (Nei and Kumar 2000). Software packages such as MEGA X (Kumar et al. 2018), LinTre (Takezaki et al. 1995), PAML (Yang 2007), and DAMBE can be used for testing the molecular clock. For example, the difference in log likelihoods with and without assuming the strict clock was

207.33 in PAML for the example data in Fig. 12.2a. The likelihood-ratio test rejects the molecular clock ($P < 10^{-80}$, degrees of freedom = 3) for this data set.

In fact, we expect the hypothesis of the strict molecular clock to be readily rejected for most contemporary data sets, which often consist of many genes and/or genomic segments from many species. Therefore, a practitioner usually needs to use dating methods that do not assume a strict clock. They might choose to apply local clocks that allow different rates in different clades (subtrees) in the phylogeny (Hasegawa et al. 1989; Yoder and Yang 2000). In local-clock methods, a strict clock is assumed to exist within each clade, so one needs to specify clades that show rate homogeneity (clocklike evolution). This is not straightforward to accomplish unless there are clear biological reasons for defining such clades (Sanderson 2002; Ho and Duchêne 2014). Consequently, methods that allow rates to vary throughout the phylogeny are more practical in analyses of real data.

12.2.2 Relaxing the Strict Clock

Relaxed-clock methods allow molecular dating when evolutionary rates vary throughout the tree. We focus on rapid non-Bayesian approaches, as Bayesian approaches have been discussed extensively elsewhere (dos Reis et al. 2016; Nascimento et al. 2017) and in Chaps. 6 and 13. Among the non-Bayesian approaches, penalized likelihood and RelTime are often used. Penalized likelihood estimates divergence times under the statistical criterion that minimizes the squared differences between ancestral and descendent branch rates (Sanderson 1997, 2002). That is, large rate changes are penalized, which is biologically intuitive because an ancestor and its direct descendants are likely to share similar genomic properties, biological attributes, and living environments, and thus will tend to have more similar mutation rates (Gillespie 1984). This property would result in autocorrelation in branch rates (Thorne et al. 1998; Kishino et al. 2001) (Table 12.2).

The penalized-likelihood approach uses a penalty parameter (λ) for penalizing rate changes (Sanderson 2002). A large penalty will favour a strict-clock model, because it will tend to assign very similar rates to ancestor–descendant pairs. Small values of λ will allow rates to vary throughout the tree and will relax the molecular clock. The optimal value of λ depends on the data set being analysed and can be determined by a cross-validation procedure (Sanderson 2002). In this procedure, one terminal branch is removed from the tree at a time, so its immediate ancestral node and other branches are left in place. The rate and node age of the immediate ancestral node is estimated using the remaining branches for a given λ. The optimal value of λ is that which minimizes the difference between the observed substitutions on the ancestral branch and the number of inferred substitutions, which is calculated using the estimated rate and node age. This rate-smoothing approach is effective when applied to the example data in Fig. 12.2a. Penalized likelihood produced an estimate of $t_X/t_Y = 0.92$ (Fig. 12.2e), which is much closer to 1 than that from methods based on a strict clock. The original penalized-likelihood method was implemented in the r8s software (Sanderson 2003) and a faster version is implemented in the treePL software (Smith and O'Meara 2012) and in the R package APE (Paradis 2013). The penalized-likelihood method has also been adopted by Xia and Yang (2011) in their strict-clock method to relax the clock through rate smoothing (Xia 2018a). It produced a time ratio of 0.90 when applied to the example data (Fig. 12.2d).

The RelTime approach is another relaxed-clock method that minimizes differences between the evolutionary rates of ancestral and descendent lineages (Tamura et al. 2012, 2018). An evolutionary lineage consists of a branch and the descendent clade (including all of the taxa and branches). For example, lineage a contains three branches in Fig. 12.3, so the length of lineage a (L_a) is based on b_1, b_2, and b_5. Tamura et al. (2018) presented a mathematical formulation that produces relative lineage rates purely from the branch lengths in a phylogeny. In their algebraic relative-rate framework, the difference between

Table 12.2 Differences between Bayesian dating methods, penalized likelihood, and RelTime

	Bayesian	Penalized likelihood	RelTime
Framework	Bayesian statistics	Penalized likelihood	Algebra
Rate prior	Independent or autocorrelated rates and probability distributions	Autocorrelated rates and a penalty parameter	Not needed
Tree prior	Birth-death or coalescent process	Not needed	Not needed
Estimate	Node times and branch rates	Node times and branch rates	Node times and lineage rates
Uncertainty	Credibility intervals	Confidence intervals	Confidence intervals
Consider site sampling error	Yes	Yes	Yes
Consider rate variation	Yes	No	Yes
Consider calibrations	Yes; allow the use of boundaries and densities	Yes; allow the use of boundaries	Yes; allow the use of boundaries and densities

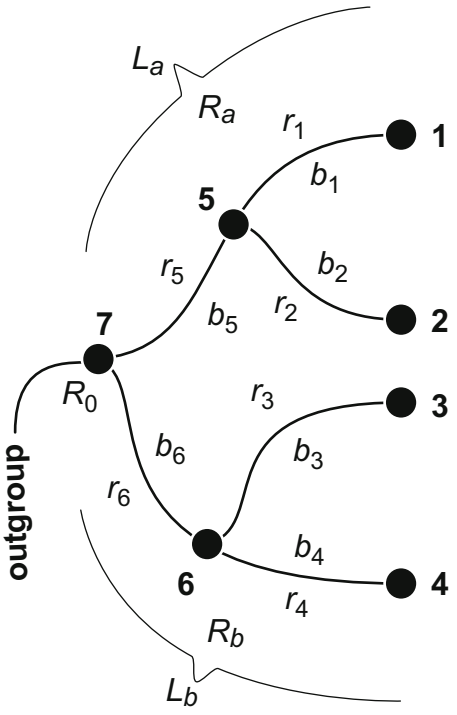

Fig. 12.3 An example phylogeny showing branch lengths (b), branch rates (r), lineage lengths (L), and lineage rates (R). R_a is the rate of the lineage L_a that consists of branches with lengths of b_5, b_1, and b_2, and R_b is the rate of the lineage L_b that consists of branches with lengths of b_6, b_3, and b_4. Lineage rates R_1 to R_4 are the same as branch rates r_1 to r_4, so they are not shown. Relative lineage rates can be computed in MEGA X from branch lengths using Eqs. (6–9, 19–24) for arithmetic means or (28–31, 34–39) for geometric means in Tamura et al. (2018)

rates in ancestral and descendent lineages is minimized and the observed difference in evolutionary rates between sister lineages is accommodated.

The use of lineage rates, rather than the branch rates, is a major difference between RelTime and other relaxed-clock methods (Table 12.2). For example, Bayesian methods use a statistical distribution (e.g., lognormal) as a prior to account for the variation in branch rates across a phylogeny, and the penalized-likelihood method smooths the rate change between ancestral and descendent branch rates using a global penalty function. If we consider node 7 in Fig. 12.3, penalized-likelihood computation will attempt to minimize the difference between branch rates r_5 and r_1 and for other pairs globally. In contrast, RelTime will minimize the difference between lineage rates R_a and R_1 and other pairs individually. Therefore, RelTime does not need to use any penalty functions or distributional priors, which makes it different from penalized-likelihood and Bayesian methods. In the example four-taxon data, RelTime produces a t_X/t_Y ratio of 0.9, which is close to 1.0 (Fig. 12.2f). The RelTime method is available in the MEGA X software. Mello (2018) has provided a detailed protocol for estimating time-trees with RelTime in MEGA X.

Overall, we find that the time ratios produced by non-Bayesian relaxed-clock methods (Fig. 12.2d–f) are similar to the estimate

generated by the Bayesian approach when an independent branch-rate (IBR) model was used (0.91, Fig. 12.2h). The use of an autocorrelated branch-rate (ABR) model produced a time ratio of 0.98, an estimate that is very close to 1 (Fig. 12.2g). The ABR model assumes that the branch-specific molecular rates are autocorrelated, such that closely related branches share similar rates and distantly related branches have more different rates (Thorne et al. 1998; Kishino et al. 2001; Ho and Duchêne 2014). The IBR model assumes molecular rates are independent among branches, such that rates on closely related branches do not need to be similar (Drummond et al. 2006; Ho and Duchêne 2014). Results from Bayesian analyses suggest that the ABR model might fit these data better. In fact, Tao et al. (2019) reported the autocorrelation of branch rates to be the dominant pattern in molecular phylogenies for diverse groups of species in an analysis of DNA and amino acid sequences. Therefore, the assumption of autocorrelation is likely to be valid for this example data set.

12.2.3 Performance of Non-Bayesian Relaxed-Clock Methods

Non-Bayesian relaxed-clock methods have been tested extensively in computer simulations with large data sets. Smith and O'Meara (2012) conducted computer simulations under the ABR model on large phylogenies (100–10,000 species) and reported that the penalized-likelihood method can achieve high accuracy in estimating divergence times (see Fig. 1 in Smith and O'Meara 2012). However, they did not test the performance of penalized likelihood using IBR data sets and did not evaluate the accuracy of divergence-time estimates node-by-node; their investigations conducted using the treePL and r8s software were rather limited in scope and depth. In contrast, RelTime has been extensively tested and has a well-justified mathematical foundation (Tamura et al. 2018).

Tamura et al. (2012) conducted extensive simulations under ABR and IBR scenarios on a master time-tree of 446 taxa. In all scenarios, RelTime produced estimates of node ages that were close to the true values (Fig. 12.4a–c; also see Figs. 3 and 5 in Tamura et al. 2012). RelTime estimates were similar to those from the Bayesian method in the IBR case where rate variation was low (Fig. 12.4a). However, the Bayesian method tended to overestimate divergence times (median deviation = 19%) when the rate variation in IBR was larger (Fig. 12.4b). This pattern might relate to the need to specify a single model of branch rates in Bayesian methods. When the specified rate model is not the correct model for the observed rate variation, biased time estimates might be produced. Model averaging can potentially reduce this bias in Bayesian analysis (Li and Drummond 2012). In contrast, RelTime does not need to model branch rates and it performed much better in this case (Fig. 12.4b, median deviation = −5%). RelTime also performed better (median deviation = −2%) than the Bayesian method (median deviation = 14%) for the ABR data sets (Fig. 12.4c). Mello et al. (2021) also reported RelTime to perform as well as Bayesian methods for dating phylogenies that encompass both species and population divergences using simulated data sets.

Apart from the simulation tests, Chernikova et al. (2011) and Gunter et al. (2016) reported that penalized-likelihood methods produced results consistent with those from Bayesian analyses for some data sets. Mello et al. (2017) and Battistuzzi et al. (2018) have also examined many empirical data sets from different groups across the tree of life and found that RelTime produced time estimates that were very similar to those from Bayesian methods, as long as the equivalent calibration boundaries were used. Tao et al. (2020) developed a method for utilizing calibration densities in RelTime and found that RelTime produced not only time estimates but also the surrounding uncertainties that were comparable to those from Bayesian methods in empirical data analyses.

In fact, some studies have found that non-Bayesian methods performed better than Bayesian methods when some priors (e.g.,

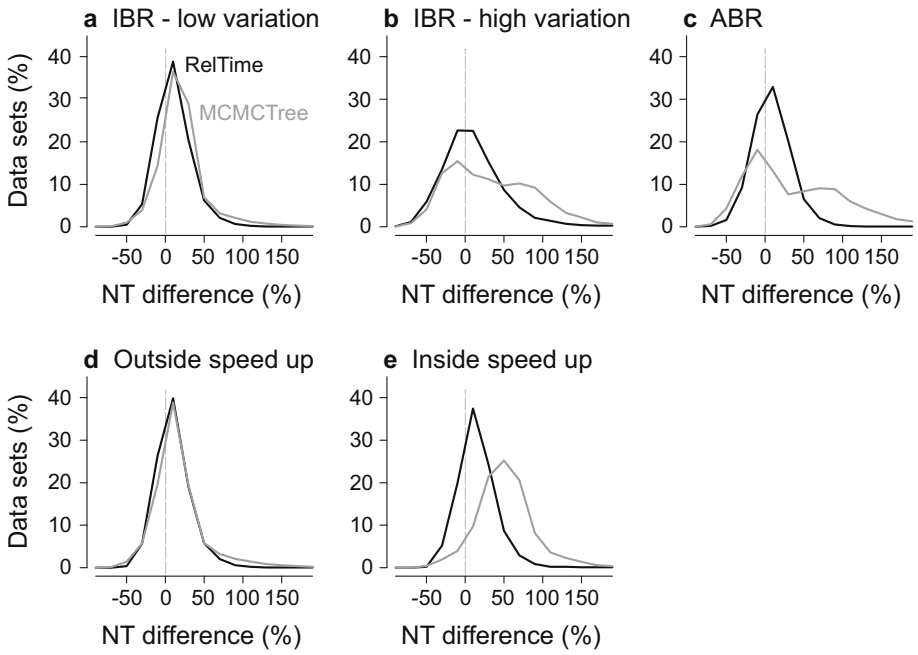

Fig. 12.4 Distributions of the normalized differences between true node times (NT) and estimated times obtained from RelTime and MCMCTree for internal nodes. Comparisons of the performance of RelTime (black curve) and MCMCTree (grey curve) for data sets simulated under (**a**) independent branch-rates (IBR) model with low variation, (**b**) IBR model with high variation, and (**c**) autocorrelated branch-rates (ABR) model. Comparisons of the performance of RelTime (black curve) and MCMCTree (grey curve) for estimating node times (**d**) outside the speed-up clades and (**e**) inside the speed-up clades. Data and results are from Tamura et al. (2012). Dashed grey line indicates the 0% difference in NT

branch-rate model) were incorrectly specified (Tamura et al. 2012, 2018). For example, Tamura et al. (2012) did a simulation test of a clade-specific speed-up, where a random clade of at least 50 taxa was selected to undergo a rate increase while the rest of the branches remained at their original rates simulated under the IBR model. This meant that two different IBR models were applied to the same phylogeny, where one clade had a higher mean rate and the other clade evolved more slowly. The Bayesian method yielded accurate estimates in one clade, but biased estimates in the other clade (Fig. 12.4d–e, also see Fig. 5 in Tamura et al. 2012). This occurred because the single model of branch-rate variation was unable to account for the heterogeneity associated with multiple contrasting clade-specific rate variations. However, RelTime performed well and generated accurate time estimates for both clades (Fig. 12.4d–e), because RelTime does not require the specification of a branch-rate model.

Therefore, the high computational speeds afforded by some of the non-Bayesian dating methods do not come at the expense of accuracy. In fact, whenever possible, it is prudent to analyse data by using methods based on different statistical frameworks to obtain reliable estimates and to assess the potential biases introduced by the assumptions and methods (see Sect. 12.9). However, efficient non-Bayesian methods might be the only feasible option for many users for analysing large data sets containing thousands of genes and species (e.g., Li et al. 2019).

12.2.4 Eliminating Rate Variability Before Molecular Dating

Before proceeding further, let us consider approaches to reduce or eliminate rate variation in data sets containing multiple genes or genomic segments, before applying clock methods. This is important because high rate variation is a key contributor to the uncertainty in time estimates (Zhu et al. 2015; Kumar and Hedges 2016). By reducing the degree of molecular rate variation in a phylogeny, both the accuracy and precision of time estimates might be improved.

We can eliminate (or reduce) evolutionary rate variation by excluding species that have evolved significantly more quickly or slowly than the rest in a sequence alignment, or by excluding genes that fail the molecular clock test. For data sets that contain large numbers of genes and genomic segments, this is a viable option for dating species divergences (Hedges et al. 1996; Smith et al. 2018). In the 1990s, Hedges et al. (1996) and Kumar and Hedges (1998) applied this strategy to date major mammalian and vertebrate divergences, respectively, because relaxed-clock methods were not available at that time. In those early multigene studies, genes and species failing the molecular clock test of Tajima (1993) were removed before divergences were dated using a strict clock. These analyses revealed that major orders of placental mammals and of birds were likely to have originated prior to the K-Pg extinction (Hedges et al. 1996), which challenged the hypothesis of adaptive radiation and founded a very active area of biological research (Kumar and Hedges 1998; Eizirik et al. 2001; dos Reis et al. 2014; Phillips 2015; Prum et al. 2015). Takezaki et al. (1995) presented a statistical approach to detect lineages that evolved at rates that were significantly different from the phylogeny-wide average. Using such gene- and species-elimination approaches, evolutionary timescales were assembled from many large data sets, including those for Hawaiian drosophilids (Russo et al. 1995), diatoms (Kooistra and Medlin 1996), metazoans (Wray et al. 1996), and major eukaryote lineages (Doolittle et al. 1996; Feng et al. 1997).

Smith et al. (2018) proposed a ‘gene shopping’ approach that extended the original practice of Hedges et al. (1996) to genes that passed the molecular clock test in large phylogenies. Their strategy also requires that the selected genes have a sufficient number of informative sites and that selected gene trees are highly concordant with the species tree. They reported that the application of strict-clock or relaxed-clock methods on the selected clocklike genes improved the precision of time estimates by more than 50%, as the 95% highest posterior density (HPD) intervals became much narrower. The higher precision is achieved by reducing the rate heterogeneity in the phylogeny, which is a key contributor to wide 95% HPD intervals. Higher precision of estimates enables more powerful tests of biological hypotheses and helps to establish evolutionary and ecological patterns more reliably.

Even after ‘gene shopping’, it is possible that some intrinsic directional rate variation remains in the data set because molecular clock tests are not so powerful when sequences are short or the evolutionary rate is low. This can be remedied by applying a more stringent clock test to exclude genes and species showing even small rate differences (Kumar and Hedges 1998; Hedges and Kumar 2003; Hedges and Shah 2003). We also propose that one should do ‘species shopping’ to remove species that show evolutionary rates significantly different from others before conducting molecular clock dating, to further reduce rate variation and the uncertainty in time estimates (Takezaki et al. 1995; Hedges et al. 1996). In our view, whenever feasible, a combination of gene shopping and species shopping with relaxed-clock methods is the best strategy when many genes and species are available for estimating divergence times.

12.3 Utility of Relative Divergence Times

All of the non-Bayesian methods can generate relative times directly from a phylogeny in which branch lengths are either provided by the user or inferred from the sequence data using a

model of nucleotide or amino acid substitution. The ability to produce relative node ages and rates without using any branch-rate model, speciation model, and even calibration priors can have many benefits (Tamura et al. 2012). First, the relative node ages obtained without any calibrations can be used to identify the calibration constraints or densities that would be expected to have a notable impact on the final time estimation (Marshall 2008). This is because the relative and absolute ('calibrated') node ages should be linearly related when calibration constraints and/or densities do not conflict with the signal from molecular data (Battistuzzi et al. 2015).

Second, the estimates of relative rates can be directly used to identify lineages with significantly lower or higher evolutionary rates, because the standard errors of the relative-rate estimates are available. Those lineages are potentially very interesting because they might indicate the presence of strong selective pressure and other biological factors (Chikina et al. 2016). In addition, the relative rates computed from branch lengths only, without knowing node times, provide insights into evolutionary patterns between the ingroup and outgroup sequences. If the distributions of lineage rates are significantly different, the assumption of the same pattern of rate variation between the ingroup and outgroup taxa might need to be reconsidered.

Third, the relative lineage rates estimated by RelTime can be used for generating new tests of biological hypotheses and for model selection. Tao et al. (2019) used these lineage rates and the machine-learning framework to develop a new statistical test (called CorrTest) that can distinguish between IBR and ABR models, which has been challenging previously (Paradis 2013; Ho et al. 2015a). CorrTest performed better than other methods in detecting the presence of rate autocorrelation in a simulation analysis.

Fourth, the relative divergence times might be useful for detecting clades that have undergone a shift in the rate of diversification, which might indicate the effect of a geological event or the appearance of an ecological niche. Therefore, the knowledge of relative times and rates is useful for discovering exciting biological patterns, developing new methods, and examining the impact of fossil constraints or other prior settings.

12.4 Inferring Absolute Divergence Times

12.4.1 Dating with a Fixed Global Evolutionary Rate

A substantial proportion (12%) of molecular clock studies have been found to use a fixed substitution rate to calibrate the molecular clock (Hipsley and Müller 2014). This is the only choice in cases where no node calibrations are available. An average evolutionary rate from another species group is used to date the divergences in the species group of interest. The estimation of node times is simple in this case: a fixed evolutionary rate is used to convert node heights (in substitutions per site) into divergence times. One just needs to divide all the node heights (in substitutions per site) by the fixed rate of evolution (in substitutions per site per time unit, such as years or million years). Some dating programs (e.g., MEGA X) provide such an option. The use of a fixed rate is only reasonable if there is a good reason to believe that the average evolutionary rates and the biological markers are the same between the species group from which the calibration rate has been derived and the species group to which it is being applied (Wilke et al. 2009). Also, the reliability of the fixed substitution rate depends on the calibrations used in the study from which the rate is obtained (Ho and Phillips 2009).

12.4.2 Dating with a Fixed Node Calibration

A better approach is to derive the clock calibration by using a known divergence time for a node in a phylogeny and then to scale all other node ages in this phylogeny based on this clock calibration. This approach does not require one to assume a molecular clock, because rapid relaxed-clock methods can deal with rate differences

among branches and lineages to generate an ultrametric tree. The clock calibration is the relative node height divided by fixed time in the ultrametric tree, and then this calibration sets the scale to convert relative times into absolute times. In MEGA X and other programs (Sanderson 2003; Britton et al. 2007; Smith and O'Meara 2012; Xia 2018a), this can be easily done by assigning a fixed time to a node, which converts all other node heights into times.

For analyses with fixed node calibrations, calibration times can come from biogeography or from ecological/environmental considerations. In fact, a literature survey of molecular dating studies has shown that 15% used times derived from geological events that were associated with geophysical isolation or the appearance of new habitats (Hipsley and Müller 2014). These calibrations can be derived from vicariance, geodispersal, or biological dispersal (Ho et al. 2015b). The geological event is a good source of calibration especially for the species that were directly affected by that event (see Chap. 9). However, it is not appropriate to use those calibrations if the research goal is to test the impact of those geological events (Kodandaramaiah 2011). Similarly, one can use the fossil record to obtain an estimate of a single divergence time in the tree, which is then used to calibrate the clock. Many early studies used a single calibration point because gene-specific alignments generally contained only a few species (e.g., Hedges et al. 1996).

12.4.3 Dating with Multiple Node Calibrations

The most common approach to calibrate a molecular clock is to use many dates derived from the fossil record (Hipsley and Müller 2014). As expected, this practice is particularly common for fossil-rich groups in the tree of life (Ksepka et al. 2015). In fact, studies have been using increasingly large numbers of calibrations, with some contemporary analyses incorporating many tens of calibrations (e.g., Meredith et al. 2011; dos Reis et al. 2015; Barba-Montoya et al. 2018; Morris et al. 2018).

12.4.3.1 Using Multiple Fixed Calibrations or Calibration Constraints

Efficient non-Bayesian relaxed-clock methods allow the use of multiple point calibrations. For example, RelTime uses a linear regression between the relative node heights in the ultrametric tree and all of the user-supplied fixed calibration points. The resulting scaling factor (f) then converts all of the relative times into absolute divergence times. In practice, however, fossil dates do not correspond directly to actual species divergence times, so they are rarely used as fixed calibration points. Instead, the earliest fossil record usually provides a reliable minimum age constraint on a node in the phylogeny (Hedges and Kumar 2004). In some cases, it is possible to place a maximum age constraint, but these are usually difficult to determine (Marshall 2008; Ho and Duchêne 2014; Bromham et al. 2018; Hedges et al. 2018). In practice, despite these difficulties, many researchers prefer to impose both minimum and maximum constraints on multiple nodes in the phylogeny.

RelTime can use all types of constraints in calibrating the molecular clock. It generates a global time factor (f) that produces time estimates that best satisfy the calibration constraints. If there is a range of f values that do not violate the calibration constraints, then the midpoint of that range becomes the estimate of f. When one or more of the absolute times fall outside the calibration constraints, then f is set so that the deviation from the calibration constraints is minimized. After that, times for calibrated nodes are adjusted to ensure that the calibration constraints are fully respected, such that the estimated times for any offending nodes are between the minimum and maximum constraint times specified by the user. This requires altering local evolutionary rates, which prompts re-optimization of all other node times in the

tree recursively in the RelTime algorithm (Tamura et al. 2013; Tao et al. 2020).

The penalized-likelihood method adds age constraints in the optimization of the penalty functions of rate smoothing, to ensure that the absolute times are within the calibration constraints imposed by the researcher (see the documentation for the r8s software). PATHd8 also smooths rates to resolve the conflicts between estimated ages and calibrations. However, PATHd8 requires the specification of at least one fixed node age as the anchor calibration, which is used to scale relative dates to absolute dates as the first step. Then, the method smooths rates of sister lineages to fit all calibration constraints (Britton et al. 2007). Also, because PATHd8 is fundamentally a strict-clock method, it has limited power in smoothing the rates compared with the relaxed-clock methods (e.g., penalized likelihood and RelTime). The least-squares-based method (DAMBE) utilizes the calibration bounds during the minimization of the residual sum of squares (RSS) of patristic distances and pairwise distances computed based on the evolutionary rate and predicted divergence times (Xia and Yang 2011). In this case, times used to compute the distance are controlled by the calibration constraints imposed in the RSS minimization. To minimize the RSS, the resulting times will be equal to the maximum or minimum bounds in some cases (Xia and Yang 2011).

12.4.3.2 Using Calibration Constraints with Probability Densities

In addition to minimum and/or maximum constraints, it is becoming commonplace to use probability densities that reflect prior belief about the possible location of the true species divergence time relative to the minimum and/or maximum constraints. Early on, Hedges and Kumar (2004) mentioned several possible distributions (triangular, lognormal, and uniform densities) to model such calibration uncertainty. However, they preferred a uniform distribution for their studies due to a lack of additional information about the true density (Meredith et al. 2011; Morris et al. 2018). With the development of Bayesian methods, it became possible to incorporate any desired probability density in molecular dating (Drummond et al. 2006; Yang and Rannala 2006; Barba-Montoya et al. 2017). Indeed, more recent studies use nonuniform distributions (e.g., Cauchy, lognormal, and exponential distributions) in which a stronger constraint is placed on the minimum time. As expected, the quality of the calibrations and the density assumptions have a major impact on divergence-time estimates in Bayesian analyses, even if a huge amount of molecular data is available (Barba-Montoya et al. 2017; Bromham et al. 2018).

Tao et al. (2020) have developed an approach to incorporate such densities and automatically accommodate the interactions among calibrations in the RelTime method. The new approach resamples calibration constraints from densities many times, to generate a distribution of times for each calibrated node that is analogous to the 'effective prior' in Bayesian approaches, and then derives minimum and maximum bounds (called effective bounds) for use in the RelTime analysis to estimate divergence times and confidence intervals. Confidence intervals produced by this approach overlapped with those reported by the Bayesian analyses and were much narrower than those generated by using the original approach that did not account for interactions among calibrations in RelTime (Tao et al. 2020). The new approach is available in MEGA X for the RelTime method. These effective bounds can also be used in penalized likelihood and other non-Bayesian dating analyses.

12.4.3.3 Using Molecular Dates as Calibrations (Secondary Calibrations)

Many studies use previously published molecular dates to calibrate the clock. These are referred to as secondary calibrations because they are not based on direct fossil or biogeographical data, but rather on inferred molecular dates. A literature

survey found that about 15% of studies have used secondary calibrations (Hipsley and Müller 2014). The use of secondary calibrations traces its origins to Kumar and Hedges (1998). They estimated vertebrate divergence times using a secondary mammalian calibration, which was inferred by using the bird-mammal divergence time from the fossil record. This procedure was needed for inferring intra- and interordinal dates using protein sequence alignments that lacked bird sequences. This approach enabled them to increase the number of genes that could be used to infer divergence times. In fact, Hedges and Kumar (2004) suggested that, in some situations, more accurate time estimates might be obtained by using a secondary calibration from a robust source than by using an unreliable primary fossil calibration.

Secondary calibrations continue to be used for groups that have limited fossil records, such as bacteria (Chriki-Adeeb and Chriki 2016) and fungi (Heckman et al. 2001). They have also been used in several recent studies to increase the total number of available calibrations for dating large phylogenies that contain hundreds of species (e.g., dos Reis et al. 2012, 2018). Ultimately, one must use secondary calibrations judiciously, because this practice might produce results significantly different from those produced by using primary calibrations (Graur and Martin 2004; Sauquet et al. 2012; Schenk 2016). However, Hedges and Kumar (2004) found that the inconsistencies in times estimated using the primary and secondary calibrations reported by Graur and Martin (2004) were caused by an incorrect assumption of Gaussian distribution of multigene times and, thus, an incorrect calculation of the time and confidence intervals. The actual distribution should be very skewed because of the small sample size and a large extrapolation. In fact, Morrison (2008) suggested that a lognormal distribution is the most appropriate to model a secondary calibration. The time estimated using the secondary calibration was consistent with the primary time when a skewed distribution was assumed (Hedges and Kumar 2004). Clearly, further research is needed to inform best practices for using secondary calibrations.

12.5 Molecular Dating with Missing Sequence Information

Modern studies often involve large data sets with hundreds of species and genes, due to the growth of public databases and dramatically decreased sequencing costs. However, a disadvantage of building and using such big data sets is that they might contain a large proportion of missing data. For example, the alignment analysed by Barba-Montoya et al. (2018) had 71.4% missing data. Fortunately, both empirical and simulation studies have found that missing data had little impact on divergence-time estimation by both Bayesian and non-Bayesian dating methods, especially when multiple calibrations were used (Douzery et al. 2004; Filipski et al. 2014; Zheng and Wiens 2015). These results indicate that molecular time estimation is robust even when sequences are missing from the majority of genes for most of the species. However, if the data are highly or systematically sparse, resulting in pairs of species with no common genes, then divergence-time estimation can be seriously misled, especially when only a few or no calibrations are used (Filipski et al. 2014; Zheng and Wiens 2015).

Filipski et al. (2014) showed that time estimates for nodes with zero data coverage (i.e., nodes without any common genes for any pair of species in the immediate descendent clades) were unreliable because there were no data to allow the corresponding branch lengths to be estimated. In general, the accuracy of branch-length estimates is low when the overall number of informative characters is small, which would result in poor time estimates (Wiens and Moen 2008; Wiens and Morrill 2011). Limited numbers of informative sites in sequence alignments can reduce the accuracy and precision of time estimates and, thus, lead to spurious changes in diversification rates (Marin and Hedges 2018) and mislead statistical tests of evolutionary rate correlation (Tao et al. 2019). Therefore, it is important to detect nodes with low or zero data coverage before any dating analysis.

One can use MEGA X to visualize data coverage for each node in a phylogeny (Fig. 12.5). The data coverage for each node in the phylogeny is

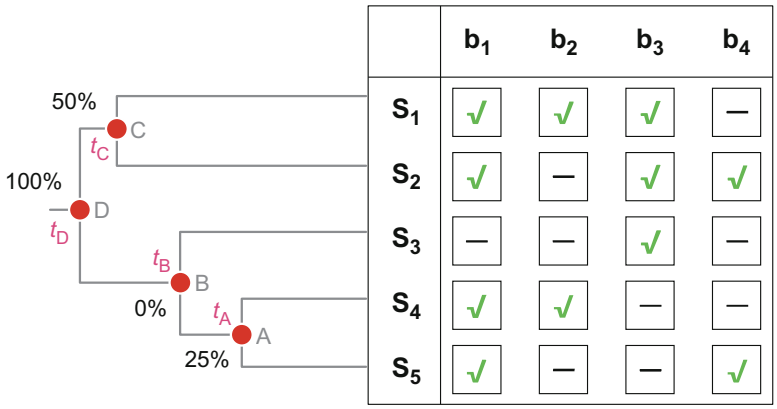

Fig. 12.5 An example of computing node data coverage for a phylogeny containing five species (S) and four nucleotide bases or amino acid residues (b) in the alignment matrix. Node times are given by t_i. Not all bases are available for each species. The available states are designated by check marks and missing ones are indicated by dashes in the matrix. The percentage of data coverage of each internal node is shown

the percentage of positions at which at least one pair of sequences in the descendent clades has a valid nucleotide base or amino acid residue. For example, node A has a data coverage of 25% because only one out of four sites has a valid state between sequences S_4 and S_5. Node B has a data coverage of 0%, because S_3 does not share any positions with a valid state in either S_4 or S_5 (Fig. 12.5). When the data coverage is zero (or low), there is no (or limited) ingroup information to allow the estimation of branch lengths (branch lengths = 0), and RelTime will predict that no time has elapsed on that branch. This results in the age of node B (t_B) becoming the same as the age of node D (t_D) (Filipski et al. 2014). Therefore, dates for nodes with high data coverage are expected to be estimated with higher accuracy.

12.6 Estimation of Confidence Intervals

In Bayesian methods, the credibility intervals or HPD intervals of node ages can be derived from the posterior distributions of times. Although the Bayesian credibility intervals and HPD intervals are not the same as the traditional analytical confidence intervals used in frequentist statistics (Jaynes and Kempthorne 1976), many researchers interpret them in a similar way. However, for non-Bayesian methods, the calculation of confidence intervals is complex. This is because it is difficult to generate analytical equations to account for the variance in node times introduced by the stochastic error in branch-length estimation, the rate heterogeneity among branches, and the uncertainty of calibrations. Therefore, many non-Bayesian methods (e.g., penalized likelihood) compute confidence intervals for divergence times by using the bootstrap approach, in which only sites or genes of molecular sequences are resampled. This leads to overly narrow confidence intervals because the site-bootstrapping approach only captures errors associated with the estimation of branch lengths in the tree. It cannot account for the variance introduced by evolutionary rate differences among lineages, which can have a big impact on the precision of time estimation (Kumar and Hedges 2016) (Table 12.2).

Tamura et al. (2013) suggested a method to generate confidence intervals encompassing the error due to branch-length estimation and rate variation for the RelTime method. Tao et al. (2020) improved this method and presented the analytical equations to compute confidence intervals for RelTime reliably, which is available

in MEGA X. Simulation analyses showed that RelTime performed better than Bayesian methods and produced confidence intervals with high probabilities of containing the true values ($\geq$94%) for both small and large data sets when a minimum number of calibrations was used.

The uncertainty in calibrations is also an important source of estimation error in the inference of divergence times. Therefore, reliable and well-constrained calibrations can be very effective in reducing the widths of confidence intervals. Bayesian methods use different probability densities to accommodate the uncertainty in calibrations and to account automatically for the interaction among calibrations. Tao et al. (2020) have developed a new method for use in the RelTime framework to derive calibration boundaries from probability densities that account for their interactions (mentioned above). The resulting confidence intervals are comparable to the HPD intervals generated from Bayesian methods in empirical analyses (Tao et al. 2020). This method, with modifications, can also be used for other non-Bayesian methods (e.g., penalized likelihood).

12.7 Dating with Non-contemporaneous Molecular Data

In some studies, molecular sequences are obtained from biological samples that have been acquired at different times. This is common in the analysis of DNA and protein sequences from fast-evolving pathogens and those generated from ancient samples (Rambaut 2000; Stadler and Yang 2013; Biek et al. 2015). This makes the tips of the evolutionary tree asynchronous. Several rapid dating methods have been developed for this type of sequence data (see also Chap. 10). As with the evolution of methods for dating analyses of contemporaneous data, the first approaches to be developed were based on a strict clock. In the single-rate dated tips (SRDT) method, the slope of a linear regression between the root-to-tip distances (or pairwise distances from the outgroup sequence) and the sampling dates is used to determine the global rate and the dates for the internal nodes (Li et al. 1988; Bollyky and Holmes 1999; Rambaut 2000). SRDT is a very fast method and has been implemented in the TipDate software (Rambaut 2000). Some UPGMA-like methods, such as serial-sampled UPGMA (Drummond and Rodrigo 2000) and TREBLE (Yang et al. 2007), were also developed under the strict-clock model. The least-squares method of Xia and Yang (2011), implemented in the DAMBE software, can also be modified to analyse non-contemporaneous data to minimize the residual sum of squares under a global clock (Xia 2018b).

Non-Bayesian methods that relax the assumption of rate constancy have also been developed, and they do not require the specification of many priors as in Bayesian approaches (To et al. 2016; Miura et al. 2020). Maximum-likelihood methods have been developed to estimate substitution rates and node dates under local and discrete clocks (Physher; Fourment and Holmes 2014) and under a relaxed clock (TreeTime; Sagulenko et al. 2018). TreeTime uses a normal prior to control the rate variation to be more autocorrelated-like or independent-like. The penalized-likelihood method implemented in r8s can also be used for dating non-contemporaneous data (Sanderson 2003). To et al. (2016) developed a least-squares dating (LSD) method that assumes the noise in molecular rates to be normal-like to account for independent rate variation across branches. Volz and Frost (2017) combined the maximum-likelihood and least-squares criteria to develop treedater. Miura et al. (2020) developed a method based on the RelTime approach, called RelTime with Dated Tips (RTDT), and the method is available in MEGA X.

Many of these non-Bayesian methods have been evaluated using data sets simulated under IBR models. They perform as well as Bayesian methods in estimating substitution rates and the root age (Fourment and Holmes 2014; To et al. 2016; Volz and Frost 2017; Sagulenko et al. 2018). Miura et al. (2020) conducted a benchmark study to assess the performance of various Bayesian and non-Bayesian methods in

estimating divergence times for a large collection of simulated data sets, which were simulated under ABR and IBR models, using different tree shapes, and with strong and weak temporal signals. For data sets with moderate or strong temporal signals, RTDT performed better than other non-Bayesian methods because it produced good node-by-node time estimates and reliable confidence intervals that often contained the true values. Other non-Bayesian methods (e.g., LSD and TreeTime) performed well for IBR data sets, but not for ABR data sets. When there was a weak temporal signal in the data, Bayesian methods provided better estimates than non-Bayesian methods, as long as the correct rate model was specified. Tong et al. (2018) also suggested that non-Bayesian methods produced reliable rate estimates when the evolutionary rate was high, but that Bayesian methods generated slightly better estimates when there was a low evolutionary rate and weak temporal signal.

Non-Bayesian methods also allow the data to have missing sampling dates or to have uncertainties in the sampling dates (Volz and Frost 2017; Sagulenko et al. 2018; Miura et al. 2020). All of these non-Bayesian methods are orders of magnitude faster than Bayesian methods (Volz and Frost 2017; Miura et al. 2020), so they provide the feasibility of dating phylogenies with thousands of tips and sampling dates, which are expected to become increasingly common in molecular epidemiology. Miura et al. (2020) provided brief guidelines for users to select the most appropriate method for tip-dating analysis, based on the characteristics of the data set being analysed.

12.8 Phylogenetic Uncertainty

In the above, we focused on the application of non-Bayesian methods for estimating divergence times and confidence intervals for a given topology, because molecular dating is frequently done after inferring a phylogenetic tree. Ideally, one would obtain a reliable tree topology using maximum likelihood and other methods, and then estimate divergence times and their uncertainties based on this fixed topology. If the inferred tree is inaccurate, divergence times estimated for many of the nodes will be meaningless, because they would not correspond to actual evolutionary divergence events. The placement of calibrations can also become complicated when the phylogenetic tree is not well established. The presence of uncertainty in the tree topology is expected to inflate the uncertainty of divergence-time estimates (Ho 2009).

In some situations, however, one might fix the nodes of interest and allow the rest of the phylogeny to be inferred from the data. In this case, it is possible to apply a chosen non-Bayesian method to each alternative topology and report the mean time estimate, the standard deviation, and a summary confidence interval around the meantime of the node of interest across all of the candidate topologies. For example, it is of great interest to date the origin of a set of pathogenic strains in tip-dating analyses. The accuracy of time estimates for this node has been tested in simulation analyses by using phylogenies inferred from the sequence alignment, rather than fixing the topology (To et al. 2016; Volz and Frost 2017; Sagulenko et al. 2018; Miura et al. 2020). The results of these analyses have been very encouraging, with RTDT and other non-Bayesian methods producing reliable estimates for this node. Similar procedures can be applied to dating species and divergences between duplicated genes by using relaxed non-Bayesian methods.

12.9 Concluding Remarks

We anticipate that RelTime, penalized likelihood, and other non-Bayesian methods will become more widely used for a number of reasons. First, the computational speed and reliable inferences offered by these non-Bayesian methods allow one to use larger data sets for dating the tree of life or for testing biological hypotheses. Because Bayesian methods often demand large amounts of computational time and memory, many researchers adopt a divide-and-conquer approach

by running the Bayesian methods on small partitions and gluing the results together (Misof et al. 2014; Oliveros et al. 2019). Alternatively, researchers might filter the genes until the size of the remaining data set is feasible for Bayesian inference (Hughes et al. 2018). This situation is going to become more acute, because progress in sequencing technology has been a boon for molecular systematics and biodiversity research, leading to a two-dimensional expansion of data sets (sites and species) available for dating studies (e.g., Zeng et al. 2014; Testo and Sundue 2016; Zheng and Wiens 2016; Barba-Montoya et al. 2018; Hughes et al. 2018). For this reason, faster Bayesian implementations are also being developed (Åkerborg et al. 2008; Lartillot et al. 2013).

Second, the use of efficient and reliable methods will enhance scientific rigour by allowing an assessment of the robustness of estimates to the assumptions made in dating analyses. Such analyses might involve studying the effects of using different combinations of genes, species, calibrations, and priors. Owing to computational time requirements, such explorations can be difficult for large data sets. Rapid non-Bayesian methods provide researchers with a toolkit to test the sensitivity of molecular time estimates and to improve downstream investigations of the biological process.

Third, the computational time requirements imposed by Bayesian methods make it challenging to examine the accuracy and precision of their estimates for large data sets, whereas rapid non-Bayesian methods have been tested on data sets with hundreds to thousands of species (Smith and O'Meara 2012; Tamura et al. 2012, 2018). A high computational burden also discourages independent evaluation of Bayesian date estimates by others interested in reproducing the results. Many practitioners are frustrated by the fact that independent attempts to simply reproduce the results of Bayesian dating can take weeks to months, and can only be pursued by research groups with access to extensive computing resources. This delays, and even impedes, scientific discourse and progress. The presence of reliable, efficient non-Bayesian methods is very useful and makes molecular dating accessible to all, including those without ready access to high-performance computing infrastructure.

Admittedly, Bayesian methods are useful when one wishes to incorporate some other information into divergence-time inference (e.g., biogeographic data) or to get a joint inference of some other phylogenetic features (e.g., population dynamics parameters). However, whether the inclusion of additional information or the joint inference will improve the accuracy of divergence-time estimation requires more extensive study, because appropriate settings for priors are usually unknown.

In fact, we suggest that users apply both Bayesian and non-Bayesian methods to obtain estimates of divergence times and their confidence intervals for molecular data sets, where possible. This would allow us to detect potential biases introduced by the assumptions and methods. Nevertheless, it is important to note that concordance between time estimates from Bayesian and non-Bayesian approaches should not be taken to suggest that the estimated times are correct. This is because the estimation of absolute divergence times highly depends on the calibration constraints used, and all methods will be negatively affected if the calibration constraints or densities used are incorrect (Battistuzzi et al. 2015; Hedges et al. 2018). For example, the use of an exponential density indicates a very high probability that the node age is close to the minimum constraint (Hedges and Kumar 2004; Ho and Duchêne 2014). Without proper justification and prior independent data, the choice of calibration density is largely subjective (Heath 2012; Bromham et al. 2018), which can adversely affect molecular date estimates. Different density distributions, even with the same minimum and maximum bounds, can produce different posterior time estimates in Bayesian methods (dos Reis et al. 2015; Barba-Montoya et al. 2017; Warnock et al. 2017; Morris et al. 2018). In addition, there are concerns about the imposition of maximum constraints on node times, because the fossil record only provides reliable minimum constraints (Battistuzzi et al. 2015; Bromham et al. 2018; Hedges et al. 2018). Therefore, one needs to examine the

reliability of calibrations before conducting dating analyses (Andújar et al. 2014; Battistuzzi et al. 2015; Hedges et al. 2018).

In general, we see no reason for avoiding non-Bayesian methods for constructing time-trees, given that they are computationally efficient and produce estimates of divergence times and their surrounding uncertainties that are scientifically rigorous and reproducible. In particular, efficient non-Bayesian methods might be the only feasible option for many users for analysing large data sets containing thousands of genes and species.

References

Åkerborg Ö, Sennblad B, Lagergren J (2008) Birth-death prior on phylogeny and speed dating. BMC Evol Biol 8:77

Andújar C, Soria-Carrasco V, Serrano J, Gómez-Zurita J (2014) Congruence test of molecular clock calibration hypotheses based on Bayes factor comparisons. Methods Ecol Evol 5:226–242

Barba-Montoya J, dos Reis M, Yang Z (2017) Comparison of different strategies for using fossil calibrations to generate the time prior in Bayesian molecular clock dating. Mol Phylogenet Evol 114:386–400

Barba-Montoya J, dos Reis M, Schneider H, Donoghue PCJ, Yang Z (2018) Constraining uncertainty in the timescale of angiosperm evolution and the veracity of a Cretaceous Terrestrial Revolution. New Phytol 218:819–834

Battistuzzi FU, Billing-Ross P, Murillo O, Filipski A, Kumar S (2015) A protocol for diagnosing the effect of calibration priors on posterior time estimates: A case study for the Cambrian explosion of animal phyla. Mol Biol Evol 32:1907–1912

Battistuzzi FU, Tao Q, Jones L, Tamura K, Kumar S (2018) RelTime relaxes the strict molecular clock throughout the phylogeny. Genome Biol Evol 10:1631–1636

Bhatnagar N, Bogdanov A, Mossel E (2011) The computational complexity of estimating MCMC convergence time. In: Goldberg LA, Jansen K, Ravi R, Rolim JDP (eds) Approximation, randomization, and combinatorial optimization. Algorithms and techniques. Springer, Heidelberg, pp 424–435

Biek R, Pybus OG, Lloyd-Smith JO, Didelot X (2015) Measurably evolving pathogens in the genomic era. Trends Ecol Evol 30:306–313

Bollyky PL, Holmes EC (1999) Reconstructing the complex evolutionary history of hepatitis B virus. J Mol Evol 49:130–141

Britton T, Oxelman B, Vinnersten A, Bremer K (2002) Phylogenetic dating with confidence intervals using mean path lengths. Mol Phylogenet Evol 24:58–65

Britton T, Anderson CL, Jacquet D, Lundqvist S, Bremer K (2007) Estimating divergence times in large phylogenetic trees. Syst Biol 56:741–752

Bromham L, Duchêne S, Hua X, Ritchie AM, Duchêne DA, Ho SYW (2018) Bayesian molecular dating: opening up the black box. Biol Rev 93:1165–1191

Chernikova D, Motamedi S, Csürös M, Koonin EV, Rogozin IB (2011) A late origin of the extant eukaryotic diversity: divergence time estimates using rare genomic changes. Biol Direct 6:26

Chikina M, Robinson JD, Clark NL (2016) Hundreds of genes experienced convergent shifts in selective pressure in marine mammals. Mol Biol Evol 33:2182–2192

Chriki-Adeeb R, Chriki A (2016) Estimating divergence times and substitution rates in Rhizobia. Evol Bioinform 12:87–97

Crosby RW, Williams TL (2017) Fast algorithms for computing phylogenetic divergence time. BMC Bioinform 18:514

Doolittle RF, Feng DF, Tsang S, Cho G, Little E (1996) Determining divergence times of the major kingdoms of living organisms with a protein clock. Science 271:470–477

dos Reis M, Inoue J, Hasegawa M, Asher RJ, Donoghue PC, Yang Z (2012) Phylogenomic datasets provide both precision and accuracy in estimating the timescale of placental mammal phylogeny. Proc R Soc B 279:3491–3500

dos Reis M, Donoghue PC, Yang Z (2014) Neither phylogenomic nor palaeontological data support a Palaeogene origin of placental mammals. Biol Lett 10:20131003

dos Reis M, Thawornwattana Y, Angelis K, Telford MJ, Donoghue PC, Yang Z (2015) Uncertainty in the timing of origin of animals and the limits of precision in molecular timescales. Curr Biol 25:1–12

dos Reis M, Donoghue PC, Yang Z (2016) Bayesian molecular clock dating of species divergences in the genomics era. Nat Rev Genet 17:71–80

dos Reis M, Gunnell GF, Barba-Montoya J, Wilkins A, Yang Z, Yoder AD (2018) Using phylogenomic data to explore the effects of relaxed clocks and calibration strategies on divergence time estimation: primates as a test case. Syst Biol 67:594–615

Douzery EJP, Snell EA, Bapteste E, Delsuc F, Philippe H (2004) The timing of eukaryotic evolution: does a relaxed molecular clock reconcile proteins and fossils? Proc Natl Acad Sci USA 101:15386–15391

Drummond A, Rodrigo AG (2000) Reconstructing genealogies of serial samples under the assumption of a molecular clock using serial-sample UPGMA. Mol Biol Evol 17:1807–1815

Drummond AJ, Ho SYW, Phillips MJ, Rambaut A (2006) Relaxed phylogenetics and dating with confidence. PLOS Biol 4:e88

Eizirik E, Murphy W, O'Brien S (2001) Molecular dating and biogeography of the early placental mammal radiation. J Hered 92:212–219

Faria NR, Rambaut A, Suchard MA, Baele G, Bedford T, Ward MJ, Tatem AJ, Sousa JD, Arinaminpathy N, Pépin J, Posada D, Peeters M, Pybus OG, Lemey P (2014) The early spread and epidemic ignition of HIV-1 in human populations. Science 346:56–61

Feng DF, Cho G, Doolittle RF (1997) Determining divergence times with a protein clock: update and reevaluation. Proc Natl Acad Sci USA 94:13028–13033

Filipski A, Murillo O, Freydenzon A, Tamura K, Kumar S (2014) Prospects for building large timetrees using molecular data with incomplete gene coverage among species. Mol Biol Evol 31:2542–2550

Fitch WM (1976) Molecular evolutionary clocks. In: Ayala FJ (ed) Molecular evolution. Sinauer, Sunderland, MA, pp 160–178

Fourment M, Holmes EC (2014) Novel non-parametric models to estimate evolutionary rates and divergence times from heterochronous sequence data. BMC Evol Biol 14:163

Gillespie JH (1984) The molecular clock may be an episodic clock. Proc Natl Acad Sci USA 81:8009–8013

Graur D, Martin W (2004) Reading the entrails of chickens: molecular timescales of evolution and the illusion of precision. Trends Genet 20:80–86

Gunter NL, Weir TA, Slipinksi A, Bocak L, Cameron SL (2016) If dung beetles (Scarabaeidae: Scarabaeinae) arose in association with dinosaurs, did they also suffer a mass co-extinction at the K-Pg boundary? PLOS ONE 11:e0153570

Hasegawa M, Kishino H, Yano T (1989) Estimation of branching dates among primates by molecular clocks of nuclear DNA which slowed down in Hominoidea. J Hum Evol 18:461–476

Heath TA (2012) A hierarchical Bayesian model for calibrating estimates of species divergence times. Syst Biol 61:793–809

Heckman DS, Geiser DM, Eidell BR, Stauffer RL, Kardos NL, Hedges SB (2001) Molecular evidence for the early colonization of land by fungi and plants. Science 293:1129–1133

Hedges SB, Kumar S (2003) Genomic clocks and evolutionary timescales. Trends Genet 19:200–206

Hedges SB, Kumar S (2004) Precision of molecular time estimates. Trends Genet 20:242–247

Hedges SB, Kumar S (2009) The timetree of life. Oxford University Press, Oxford, UK

Hedges SB, Shah P (2003) Comparison of mode estimation methods and application in molecular clock analysis. BMC Bioinform 4:31

Hedges SB, Parker PH, Sibley CG, Kumar S (1996) Continental breakup and the ordinal diversification of birds and mammals. Nature 381:226–229

Hedges SB, Marin J, Suleski M, Paymer M, Kumar S (2015) Tree of life reveals clock-like speciation and diversification. Mol Biol Evol 32:835–845

Hedges SB, Tao Q, Walker M, Kumar S (2018) Accurate timetrees require accurate calibrations. Proc Natl Acad Sci USA 115:E9510–E9511

Hipsley CA, Müller J (2014) Beyond fossil calibrations: realities of molecular clock practices in evolutionary biology. Front Genet 5:138

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

Ho SYW (2014) The changing face of the molecular evolutionary clock. Trends Ecol Evol 29:496–503

Ho SYW, Duchêne S (2014) Molecular-clock methods for estimating evolutionary rates and timescales. Mol Ecol 23:5947–5965

Ho SYW, Phillips MJ (2009) Accounting for calibration uncertainty in phylogenetic estimation of evolutionary divergence times. Syst Biol 58:367–380

Ho SYW, Duchêne S, Duchêne D (2015a) Simulating and detecting autocorrelation of molecular evolutionary rates among lineages. Mol Ecol Resour 15:688–696

Ho SYW, Tong KJ, Foster CS, Ritchie AM, Lo N, Crisp MD (2015b) Biogeographic calibrations for the molecular clock. Biol Lett 11:20150194

Huerta-Cepas J, Gabaldón T (2011) Assigning duplication events to relative temporal scales in genome-wide studies. Bioinformatics 27:38–45

Hughes LC, Ortí G, Huang Y, Sun Y, Baldwin CC, Thompson AW, Arcila D, Betancur-R R, Li C, Becker L, Bellora N, Zhao X, Li X, Wang M, Fang C, Xie B, Zhou Z, Huang H, Chen S, Venkatesh B, Shi Q (2018) Comprehensive phylogeny of ray-finned fishes (Actinopterygii) based on transcriptomic and genomic data. Proc Natl Acad Sci USA 115:6249–6254

Jaynes ET, Kempthorne O (1976) Confidence intervals vs Bayesian intervals. In: Harper WL, Hooker CA (eds) Foundations of probability theory, statistical inference, and statistical theories of science. Springer, Dordrecht, pp 175–257

Jiao Y, Wickett NJ, Ayyampalayam S, Chanderbali AS, Landherr L, Ralph PE, Tomsho LP, Hu Y, Liang H, Soltis PS, Soltis DE, Clifton SW, Schlarbaum SE, Schuster SC, Ma H, Leebens-Mack J, dePamphilis CW (2011) Ancestral polyploidy in seed plants and angiosperms. Nature 473:97–100

Kishino H, Thorne JL, Bruno WJ (2001) Performance of a divergence time estimation method under a probabilistic model of rate evolution. Mol Biol Evol 18:352–361

Kodandaramaiah U (2011) Tectonic calibrations in molecular dating. Curr Zool 57:116–124

Kooistra WH, Medlin LK (1996) Evolution of the diatoms (Bacillariophyta): IV. A reconstruction of their age from small subunit rRNA coding regions and the fossil record. Mol Phylogenet Evol 6:391–407

Ksepka DT, Parham JF, Allman JF, Benton MJ, Carrano MT, Cranston KA, Donoghue PC, Head JJ, Hermsen EJ, Irmis RB, Joyce WG, Kohli M, Lamm KD, Leehr D, Patané JL, Polly D, Phillips MJ, Smith NA,

Smith ND, Van Tuinen M, Ware JL, Warnock RCM (2015) The Fossil Calibration Database, a new resource for divergence dating. Syst Biol 64:853–859

Kumar S (2005) Molecular clocks: four decades of evolution. Nat Rev Genet 6:654–662

Kumar S, Hedges SB (1998) A molecular timescale for vertebrate evolution. Nature 392:917–920

Kumar S, Hedges SB (2016) Advances in time estimation methods for molecular data. Mol Biol Evol 33:863–869

Kumar S, Stecher G, Li M, Knyaz C, Tamura K (2018) MEGA X: molecular evolutionary genetics analysis across computing platforms. Mol Biol Evol 35:1547–1549

Langley CH, Fitch WM (1974) An examination of the constancy of the rate of molecular evolution. J Mol Evol 3:161–177

Lartillot N, Rodrigue N, Stubbs D, Richer J (2013) PhyloBayes MPI: phylogenetic reconstruction with infinite mixtures of profiles in a parallel environment. Syst Biol 62:611–615

Li WLS, Drummond AJ (2012) Model averaging and Bayes factor calculation of relaxed molecular clocks in Bayesian phylogenetics. Mol Biol Evol 29:751–761

Li WH, Tanimura M, Sharp PM (1988) Rates and dates of divergence between AIDS virus nucleotide sequences. Mol Biol Evol 5:313–330

Li H-T, Yi T-S, Gao L-M, Ma P-F, Zhang T, Yang J-B, Gitzendanner MA, Fritsch PW, Cai J, Luo Y, Wang H, van der Bank M, Zhang S-D, Wang Q-F, Wang J, Zhang Z-R, Fu C-N, Yang J, Hollingsworth PMN, Chase MW, Soltis DE, Soltis PS, Li D-Z (2019) Origin of angiosperms and the puzzle of the Jurassic gap. Nat Plants 5:461–470

Louca S, Shih PM, Pennell MW, Fischer WW, Parfrey LW, Doebeli M (2018) Bacterial diversification through geological time. Nat Ecol Evol 2:1458–1467

Lu L-M, Mao L-F, Yang T, Ye J-F, Liu B, Li H-L, Sun M, Miller JT, Mathews S, Hu H-H, Niu Y-T, Peng D-X, Chen Y-H, Smith SA, Chen M, Xiang K-L, Le C-T, Dang V-C, Soltis PS, Soltis DE, Li J-H, Chen Z-D (2018) Evolutionary history of the angiosperm flora of China. Nature 554:234–238

Marin J, Hedges SB (2018) Undersampling genomes has biased time and rate estimates throughout the tree of life. Mol Biol Evol 35:2077–2084

Marin J, Battistuzzi FU, Brown AC, Hedges SB (2017) The timetree of prokaryotes: new insights into their evolution and speciation. Mol Biol Evol 34:437–446

Marshall CR (2008) A simple method for bracketing absolute divergence times on molecular phylogenies using multiple fossil calibration points. Am Nat 171:726–742

Mello B (2018) Estimating timetrees with MEGA and the TimeTree resource. Mol Biol Evol 35:2334–2342

Mello B, Tao Q, Tamura K, Kumar S (2017) Fast and accurate estimates of divergence times from big data. Mol Biol Evol 34:45–50

Mello B, Tao Q, Kumar S (2021) Molecular dating for phylogenies containing a mix of populations and species by using Bayesian and RelTime approaches. Mol Ecol Resour (in press)

Meredith RW, Janečka JE, Gatesy J, Ryder OA, Fisher CA, Teeling EC, Goodbla A, Eizirik E, Simão TL, Stadler T, Rabosky DL, Honeycutt RL, Flynn JJ, Ingram CM, Steiner C, Williams TL, Robinson TJ, Burk-Herrick A, Westerman M, Ayoub NA, Springer MS, Murphy WJ (2011) Impacts of the Cretaceous Terrestrial Revolution and KPg extinction on mammal diversification. Science 334:521–524

Metsky HC, Matranga CB, Wohl S, Schaffner SF, Freije CA, Winnicki SM, West K, Qu J, Baniecki ML, Gladden-Young A, Lin AE, Tomkins-Tinch CH, Ye SH, Park DJ, Luo CY, Barnes KG, Shah RR, Chak B, Barbosa-Lima G, Delatorre E, Vieira YR, Paul LM, Tan AL, Barcellona CM, Porcelli MC, Vasquez C, Cannons AC, Cone MR, Hogan KN, Kopp EW, Anzinger JJ, Garcia KF, Parham LA, Ramírez RMG, Montoya MCM, Rojas DP, Brown CM, Hennigan S, Sabina B, Scotland S, Gangavarapu K, Grubaugh ND, Oliveira G, Robles-Sikisaka R, Rambaut A, Gehrke L, Smole S, Halloran ME, Villar L, Mattar S, Lorenzana I, Cerbino-Neto J, Valim C, Degrave W, Bozza PT, Gnirke A, Andersen KG, Isern S, Michael SF, Bozza FA, Souza TML, Bosch I, Yozwiak NL, MacInnis BL, Sabeti PC (2017) Zika virus evolution and spread in the Americas. Nature 546:411–415

Misof B, Liu S, Meusemann K, Peters RS, Donath A, Mayer C, Frandsen PB, Ware J, Flouri T, Beutel RG, Niehuis O, Petersen M, Izquierdo-Carrasco F, Wappler T, Rust J, Aberer AJ, Aspöck U, Aspöck H, Bartel D, Blanke A, Berger S, Böhm A, Buckley TR, Calcott B, Chen J, Friedrich F, Fukui M, Fujita M, Greve C, Grobe P, Gu S, Huang Y, Jermiin LS, Kawahara AY, Krogmann L, Kubiak M, Lanfear R, Letsch H, Li Y, Li Z, Li J, Lu H, Machida R, Mashimo Y, Kapli P, McKenna DD, Meng G, Nakagaki Y, Navarrete-Heredia JL, Ott M, Ou Y, Pass G, Podsiadlowski L, Pohl H, von Reumont BM, Schütte K, Sekiya K, Shimizu S, Slipinski A, Stamatakis A, Song W, Su X, Szucsich NU, Tan M, Tan X, Tang M, Tang J, Timelthaler G, Tomizuka S, Trautwein M, Tong X, Uchifune T, Walzl MG, Wiegmann BM, Wilbrandt J, Wipfler B, Wong TK, Wu Q, Wu G, Xie Y, Yang S, Yang Q, Yeates DK, Yoshizawa K, Zhang Q, Zhang R, Zhang W, Zhang Y, Zhao J, Zhou C, Zhou L, Ziesmann T, Zou S, Li Y, Xu X, Zhang Y, Yang H, Wang J, Wang J, Kjer KM, Zhou X (2014) Phylogenomics resolves the timing and pattern of insect evolution. Science 346:763–767

Miura S, Tamura K, Tao Q, Huuki LA, Pond SLK, Priest J, Deng J, Kumar S (2020) A new method for inferring timetrees from temporally sampled molecular sequences. PLOS Comput Biol 16:e1007046

Morris JL, Puttick MN, Clark JW, Edwards D, Kenrick P, Pressel S, Wellman CH, Yang Z, Schneider H, Donoghue PC (2018) The timescale of early land

plant evolution. Proc Natl Acad Sci USA 115:E2274–E2283
Morrison DA (2008) How to summarize estimates of ancestral divergence times. Evol Bioinform 4:75–95
Muse SV, Weir BS (1992) Testing for equality of evolutionary rates. Genetics 132:269–276
Nascimento FF, dos Reis M, Yang Z (2017) A biologist's guide to Bayesian phylogenetic analysis. Nat Ecol Evol 1:1446–1454
Nei M, Kumar S (2000) Molecular evolution and phylogenetics. Oxford University Press, Oxford, UK
Oliveros CH, Field DJ, Ksepka DT, Barker FK, Aleixo A, Andersen MJ, Alström P, Benz BW, Braun EL, Braun MJ, Bravo GA, Brumfield RT, Chesser RT, Claramunt S, Cracraft J, Cuervo AM, Derryberry EP, Glenn TC, Harvey MG, Hosner PA, Joseph L, Kimball RT, Mack AL, Miskelly CM, Peterson AT, Robbins MB, Sheldon FH, Silveira LF, Smith BT, White ND, Moyle RG, Faircloth BC (2019) Earth history and the passerine superradiation. Proc Natl Acad Sci USA 116:7916–7925
Paradis E (2013) Molecular dating of phylogenies by likelihood methods: a comparison of models and a new information criterion. Mol Phylogenet Evol 67:436–444
Phillips MJ (2015) Geomolecular dating and the origin of placental mammals. Syst Biol 65:546–557
Prum RO, Berv JS, Dornburg A, Field DJ, Townsend JP, Lemmon EM, Lemmon AR (2015) A comprehensive phylogeny of birds (Aves) using targeted next-generation DNA sequencing. Nature 526:569–578
Rambaut A (2000) Estimating the rate of molecular evolution: incorporating non-contemporaneous sequences into maximum likelihood phylogenies. Bioinformatics 16:395–399
Russo C, Takezaki N, Nei M (1995) Molecular phylogeny and divergence times of drosophilid species. Mol Biol Evol 12:391–404
Sagulenko P, Puller V, Neher RA (2018) TreeTime: maximum-likelihood phylodynamic analysis. Virus Evol 4:vex042
Sanderson MJ (1997) A nonparametric approach to estimating divergence times in the absence of rate constancy. Mol Biol Evol 14:1218–1231
Sanderson MJ (2002) Estimating absolute rates of molecular evolution and divergence times: a penalized likelihood approach. Mol Biol Evol 19:101–109
Sanderson MJ (2003) r8s: inferring absolute rates of molecular evolution and divergence times in the absence of a molecular clock. Bioinformatics 19:301–302
Sarich VM, Wilson AC (1967) Immunological time scale for hominid evolution. Science 158:1200–1203
Sarich VM, Wilson AC (1973) Generation time and genomic evolution in primates. Science 179:1144–1147
Sauquet H, Ho SYW, Gandolfo MA, Jordan GJ, Wilf P, Cantrill DJ, Bayly MJ, Bromham L, Brown GK, Carpenter RJ, Lee DM, Murphy DJ, Sniderman JM, Udovicic F (2012) Testing the impact of calibration on molecular divergence times using a fossil-rich group: the case of *Nothofagus* (Fagales). Syst Biol 61:289–313
Schenk JJ (2016) Consequences of secondary calibrations on divergence time estimates. PLOS ONE 11: e0148228
Smith SA, O'Meara BC (2012) treePL: divergence time estimation using penalized likelihood for large phylogenies. Bioinformatics 28:2689–2690
Smith SA, Brown JW, Walker JF (2018) So many genes, so little time: a practical approach to divergence-time estimation in the genomic era. PLOS ONE 13: e0197433
Stadler T, Yang Z (2013) Dating phylogenies with sequentially sampled tips. Syst Biol 62:674–688
Tajima F (1993) Simple methods for testing the molecular evolutionary clock hypothesis. Genetics 135:599–607
Takezaki N, Rzhetsky A, Nei M (1995) Phylogenetic test of the molecular clock and linearized trees. Mol Biol Evol 12:823–833
Tamura K, Battistuzzi FU, Billing-Ross P, Murillo O, Filipski A, Kumar S (2012) Estimating divergence times in large molecular phylogenies. Proc Natl Acad Sci USA 109:19333–19338
Tamura K, Stecher G, Peterson D, Filipski A, Kumar S (2013) MEGA6: molecular evolutionary genetics analysis version 6.0. Mol Biol Evol 30:2725–2729
Tamura K, Tao Q, Kumar S (2018) Theoretical foundation of the RelTime method for estimating divergence times from variable evolutionary rates. Mol Biol Evol 35:1170–1182
Tao Q, Tamura K, Battistuzzi FU, Kumar S (2019) A machine learning method for detecting autocorrelation of evolutionary rates in large phylogenies. Mol Biol Evol 36:811–824
Tao Q, Tamura K, Mello B, Kumar S (2020) Reliable confidence intervals for RelTime estimates of evolutionary divergence times. Mol Biol Evol 37:280–290
Testo W, Sundue M (2016) A 4000-species dataset provides new insight into the evolution of ferns. Mol Phylogenet Evol 105:200–211
Thorne JL, Kishino H, Painter IS (1998) Estimating the rate of evolution of the rate of molecular evolution. Mol Biol Evol 15:1647–1657
To T-H, Jung M, Lycett S, Gascuel O (2016) Fast dating using least-squares criteria and algorithms. Syst Biol 65:82–97
Tong KJ, Duchêne DA, Duchêne S, Geoghegan JL, Ho SYW (2018) A comparison of methods for estimating substitution rates from ancient DNA sequence data. BMC Evol Biol 18:70
Volz E, Frost S (2017) Scalable relaxed clock phylogenetic dating. Virus Evol 3:vex025
Warnock RCM, Yang Z, Donoghue PCJ (2017) Testing the molecular clock using mechanistic models of fossil preservation and molecular evolution. Proc R Soc B 284:20170227
Wiens JJ, Moen DS (2008) Missing data and the accuracy of Bayesian phylogenetics. J Syst Evol 46:307–314

Wiens JJ, Morrill MC (2011) Missing data in phylogenetic analysis: reconciling results from simulations and empirical data. Syst Biol 60:719–731

Wilke T, Schultheiß R, Albrecht C (2009) As time goes by: a simple fool's guide to molecular clock approaches in invertebrates. Am Malacol Bull 27:25–45

Worobey M, Han G-Z, Rambaut A (2014) A synchronized global sweep of the internal genes of modern avian influenza virus. Nature 508:254–257

Wray GA, Levinton JS, Shapiro LH (1996) Molecular evidence for deep Precambrian divergences among metazoan phyla. Science 274:568–573

Wu C-I, Li W-H (1985) Evidence for higher rates of nucleotide substitution in rodents than in man. Proc Natl Acad Sci USA 82:1741–1745

Xia X (2018a) DAMBE7: New and improved tools for data analysis in molecular biology and evolution. Mol Biol Evol 35:1550–1552

Xia X (2018b) Bioinformatics and the cell: modern computational approaches in genomics, proteomics and transcriptomics. Springer International, New York

Xia X, Yang Q (2011) A distance-based least-square method for dating speciation events. Mol Phylogenet Evol 59:342–353

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

Yang Z, Rannala B (2006) Bayesian estimation of species divergence times under a molecular clock using multiple fossil calibrations with soft bounds. Mol Biol Evol 23:212–226

Yang Z, O'Brien JD, Zheng X, Zhu H-Q, She Z-S (2007) Tree and rate estimation by local evaluation of heterochronous nucleotide data. Bioinformatics 23:169–176

Yoder AD, Yang Z (2000) Estimation of primate speciation dates using local molecular clocks. Mol Biol Evol 17:1081–1090

Yu Y, Xiang Q, Manos PS, Soltis DE, Soltis PS, Song B-H, Cheng S, Liu X, Wong G (2017) Whole-genome duplication and molecular evolution in *Cornus* L. (Cornaceae) – Insights from transcriptome sequences. PLOS ONE 12:e0171361

Zeng L, Zhang Q, Sun R, Kong H, Zhang N, Ma H (2014) Resolution of deep angiosperm phylogeny using conserved nuclear genes and estimates of early divergence times. Nat Commun 5:4956

Zheng Y, Wiens JJ (2015) Do missing data influence the accuracy of divergence-time estimation with BEAST? Mol Phylogenet Evol 85:41–49

Zheng Y, Wiens JJ (2016) Combining phylogenomic and supermatrix approaches, and a time-calibrated phylogeny for squamate reptiles (lizards and snakes) based on 52 genes and 4162 species. Mol Phylogenet Evol 94:537–547

Zhu T, dos Reis M, Yang Z (2015) Characterization of the uncertainty of divergence time estimation under relaxed molecular clock models using multiple loci. Syst Biol 64:267–280

Zuckerkandl E, Pauling L (1962) Molecular disease, evolution, and genic heterogeneity. In: Kasha M, Pullman B (eds) Horizons in biochemistry. Academic, New York, pp 189–225