Opinion TRENDS in Ecology and Evolution Vol.xxx No.x

Dates from the molecular clock: how wrong can we be?

Mário J.F. Pulquério and Richard A. Nichols

School of Biological and Chemical Sciences, Queen Mary, University of London, London, E1 4NS, UK

Large discrepancies have been found in dates of evolutionary events obtained using the molecular clock. Twofold differences have been reported between the dates estimated from molecular data and those from the fossil record; furthermore, different molecular methods can give dates that differ 20-fold. New software attempts to incorporate appropriate allowances for this uncertainty into the calculation of the accuracy of date estimates. Here, we propose that these innovations represent welcome progress towards obtaining reliable dates from the molecular clock, but warn that they are currently unproven, given that the causes and pattern of the discrepancies are the subject of ongoing research. This research implies that many previous studies, even some of those using recently developed methods, might have placed too much confidence in their date estimates, and their conclusions might need to be revised.

Molecular clocks and substitution rates

This article was motivated by the experience of a colleague who estimated the time since the separation of two taxa from the number of substitutions that had accumulated between their DNA sequences; in other words, he was using the molecular clock. On submitting the work for publication, he was startled to be advised by a referee that his estimate was wrong by a factor of ten. The argument concerned the tick rate of the molecular clock; that is, the rate of accumulation of substitutions per million years. How could the scientific community hold two such contradictory opinions simultaneously?

Our colleague's original calculation was based on a rate estimated from inter-species comparisons, whereas the referee preferred a rate obtained from a pedigree study. Later, we address why such discrepancies exist between estimates of substitution rates. The central lesson for this article, however, is the realization that reasonable scientists working with the molecular clock can be using estimates that are so different. If neither the fast estimate nor the slow estimate were self-evidently wrong, it suggests that it is difficult to validate them using our knowledge of biogeography and the fossil record. Methods are currently being devised that deal with uncertainty about the variation in the rate and about the timing of the calibration points. Here, we consider the prospects of obtaining date estimates that take account of these issues when constructing their standard errors (or analogous measures of uncertainty): is there likely to be so much uncertainty about molecular dating that the estimates are no longer useful? We fear that, for many current studies, the answer is yes. However, it might be possible to gain extra precision using recently developed methods. The degree of improvement depends on the pattern of variation in the rate of molecular evolution and the availability of calibration points. We currently do not know enough to be confident in the prospects of these new methods, and some initial results are discouraging.

Why might the molecular clock have a relatively constant rate?

During the 1960s, Zuckerkandl and Pauling [1] observed that the number of amino acid differences between the haemoglobin of different species had an approximately linear relationship with the time since their common ancestor, as estimated from the fossil record. Kimura [2,3] explained the unexpectedly constant and rapid rate of evolution by assuming that most substitutions have little effect on fitness, contrary to the orthodoxy of the time (Box 1). This 'Neutral Theory' is not a complete explanation, however. For example, it predicts a constant substitution rate per generation, whereas empirical evidence suggests something closer to a constant rate per year [3]. Consider, as an example, two related lineages: the elephants and the elephant shrews, which appear to have diverged ~80 million years ago (Mya) [4]. They have very different generation times: for elephants, it is ~25 years, whereas for elephant shrews, it is approximately two orders of magnitude less. Despite their much shorter generations, the estimated rate of substitution per year is only 2.5 times faster in elephant shrews compared with that in elephants [4].

In fact, Ohta [5] found that generation-time effects are more apparent in DNA-sequence data than in comparisons of amino-acid sequence. She explained this pattern with her Nearly-Neutral extension of Kimura's theory (Box 1), which argues that substitution rates can be elevated in small populations by the fixation of mildly deleterious mutations, and that this effect, among others [6–8], can compensate for longer generation times. Mutations are more likely to be deleterious if they are amino-acid changing (i.e. non-synonymous), hence the compensation between population size and generation time might be more effective for amino-acid sequences than for DNA.

The fundamental principles associated with the Nearly Neutral Theory mean that the rate of the molecular clock is known to vary between evolutionary lineages, and that it does so in a way that is not precisely predictable,

Corresponding author: Pulquério, M.J.F. (m.pulquerio@qmul.ac.uk).
Available online xxxxxx.

  doi:10.1016/j.tree.2006.11.013

Please cite this article in press as: Pulquério, M.J.F. and Nichols, R.A., Dates from the molecular clock: how wrong can we be?, Trends Ecol. Evol. (2006), doi:10.1016/j.tree.2006.11.013

Box 1. The neutral and nearly neutral rate

The apparently constant rate at which nucleotide substitutions have accumulated among species was initially thought puzzling. This pattern contrasts with the rate of morphological divergence between species, which differs considerably along different lineages [47], probably because of differences in the selection moulding the form and structure of the species. Kimura [3] resolved this problem by proposing that most molecular changes have negligible effects on fitness, and argued that advantageous mutations affecting morphology, and the phenotype in general, are rare. An important feature of his so-called 'Neutral Theory' is that the substitution rate (mutations fixed per generation, as can be inferred by comparing sequences from different species) is equal to the mutation rate, μ_n (mutations per gamete per generation) and is independent of population size. This result can be explained by considering a (diploid) population of size N that is produced by $2N$ gametes, each of which might contain a mutation; thus, $2N\mu_n$ new neutral mutations arrive in a population on average each generation. Because they are neutral, the mutations have just as much chance as any of those in the other $2N$ lineages of eventually becoming fixed, so the probability of fixation is $1/2N$ for each of $2N\mu_n$ mutations per generation, making the overall rate $1/2N \times 2N\mu_n = \mu_n$.

When DNA sequence data became readily available, it was apparent that the rates of nucleotide substitution were more rapid in the evolutionary lineages of organisms with shorter generation times. This trend was not so clearly discerned in the rates of substitution of amino acids. This incongruence was one of the reasons that led Ohta [48] to propose the Nearly Neutral Theory of Evolution, which highlights the effect of population size on the spread of mildly deleterious mutations. Such alleles can increase in frequency by chance, under the action of genetic drift, in a similar manner to genuinely neutral alleles. If the selection acting against a new mutation is sufficiently mild, then the deleterious allele can spread to fixation. Genetic drift becomes sufficiently rapid relative to selection if the population size is smaller than the inverse of the selection coefficient [6]. There does appear to be an inverse relationship between generation time and population size that could obscure the generation time effect on substitution for amino acids (which might be subject to selection), but not on synonymous mutations in the nucleotide sequence.

although broad trends might be anticipated. Indeed, there are many additional attributes of the biology of species that affect the substitution rate [9–13], most of which appear to be associated with differences in the frequency or repair of mutation. The substitution rate variation might not be disastrous for estimating times, because related species will be similar for many attributes, including those that affect the substitution rate. For example, compared with rodents, primates tend to have larger body sizes, longer lives and more stable population sizes, each of which might affect substitution rate. Consequently, primate substitution rates might be relatively constant but differ from those of rodents. Thus, it is possible that large branches of the tree of life have relatively constant rates. However, embarrassingly large discrepancies have been found in some cases when dates were crosschecked by using different molecular analyses and by comparing clock-based dates with independent estimates, including those from the fossil record [14]. For example, twofold differences have been reported for the divergence time between marsupials and eutherians using different methods (104 versus >218 Mya) and for humans and gorillas using different calibration dates (8 versus 18 Mya) [14]. These findings raise the question of how wrong most published dates might be.

How bad is the problem?

Any assessment of the errors associated with molecular clock estimates must face the fundamental problem that there is no definitive yardstick against which they can be checked: we do not know the true dates of the divergences of species. There are three broad approaches to circumvent this issue.

First, the variation in the substitution rate can be detected without knowing any times, because we know logically that, for any two extant species, the time back to their common ancestor is the same for both lineages. The problem with this type of analysis is low power [15]: differences of fourfold can escape detection. Consequently, convincing tests must combine information from many species to characterize the overall degree of rate variation, and our knowledge of a particular evolutionary branch is less precise.

A second, rarely used approach is to check the inferred substitution rate from phylogenetic studies against the mutation rate observed in pedigrees, given that, according to the Neutral Theory, mutation rate and substitution rate should be the same [3]. However, these comparisons suggest large differences: the mutation rate estimated from pedigrees of humans is a hundredfold higher than the substitution rate for the primate mitochondrial DNA control region [16]. One response is to avoid calibrating the molecular clock using pedigree-based mutation rates, but it would be reassuring to understand why these rates differ so much.

The third approach is to compare the clock-based dates with those obtained by other approaches, such as fossil or biogeographical studies. These dates are themselves imprecise: the geological dating has some error; more importantly, the geological events will not always correspond exactly with the branching in the tree of life (such as speciation events). Indeed, times estimated from the fossil record are consistently older, with differences that can reach twofold [17].

It is apparent that none of these three methods provide a straightforward crosschecking of molecular dates. The twofold errors detected by comparing dates from the fossil record and from the molecular clocks could be the tip of the iceberg. At one extreme, there is evidence of molecular clock rates being 20-fold higher over the short period since human–Neanderthal divergence compared with the longer-term vertebrate rate [18]. Can dating methods reliably accommodate such large differences in clock rates?

Addressing the problem of rate variation among lineages

The most recent statistical methods (e.g. Refs [19–21]) have moved away from the assumption of a constant substitution rate; some even enable each branch of a phylogenetic tree to have its own rate of evolution [22,23]. Until recently, methods that relax the clock have assumed autocorrelation of rates (e.g. Refs [24–26]. In other words, nearby branches on the tree have similar rates, under the assumption that recent shared ancestry implies similar biology, which, in turn, implies similar rates. Does this innovation lead to a rapprochement between dates from the fossils and the molecular clock [17]? In some cases, the

Please cite this article in press as: Pulquério, M.J.F. and Nichols, R.A., Dates from the molecular clock: how wrong can we be?, Trends Ecol. Evol. (2006), doi:10.1016/j.tree.2006.11.013

answer is yes: for example, the improved estimate for the divergence of the Proboscidea is only 1% younger than the time estimated from fossils [17]. However, in other cases, the estimates are much older than those from fossils, for example, 72% older for the origin of the deuterostomes [27], and >200% for Afrosoricida [17].

These estimates depend on the underlying mathematical models of substitution and rate variation, yet a recent study has thrown doubt on the fundamental assumption that the rates are autocorrelated. Drummond *et al.* [21] have developed a method that can be used to investigate the variation in rate between adjacent branches (Box 2), and found no convincing evidence for autocorrelation. Do these patterns suggest a more prominent role for natural selection as suggested by some authors [28,29]? We suspect so, and that as more studies compare the same taxa at multiple loci, evidence of selection will accumulate in the form of rate variation affecting some loci but not others.

Box 2. A new method for relaxing the molecular clock

Drummond *et al.* [21] recently developed a Bayesian method that makes no assumption about the correlation between substitution rates in the tree. At first, it is hard to see how the molecular clock can be used if different branches of the tree can have independent clock rates, which would seem to imply that any branch on the tree could be stretched to fit with an arbitrary belief about the timing. This is not the case, because the rates are not completely unconstrained, but are assumed to be drawn from a statistical distribution. The shape of the distribution (e.g. its mean and variance) can be estimated from the calibration points. For example, if there is a good fossil record, which enables the age of several different parts of the tree to be estimated, then the rate of substitution in each part can be inferred. These separate rates could then be used to estimate both a mean and a variance. In addition, there is a logical constraint: some nodes (i.e. branch points on a tree) are older than others, whereas, for example, nodes on a descendant lineage are younger than the parental node. The user chooses one of the possible statistical distributions (e.g. an exponential distribution or a lognormal distribution), which is then used to specify the probability of a particular rate in a particular branch of the tree.

The Bayesian method estimates the divergence times, the topology of the tree and the rates, all as part of the same calculation. Different combinations of these parameters are chosen by using a Markov Chain Monte Carlo (MCMC) method, an algorithm that samples different parameter combinations using a scheme that is more likely to choose parameters that explain the date well. It therefore builds up a picture of the range of plausible parameter values.

The choice of trees and the dates of branch points are further constrained to reflect what is already known about the species before the genetic data are analysed. The range of plausible evolutionary histories is included in the analysis as prior distributions. For instance, the user can specify that the branching patterns follow a Yule process (a simple uniform probability speciation), in which symmetrical trees are more likely than are asymmetrical trees. Similarly, the dates of the calibration points might be considered equally probable within a range suggested by a palaeontologist, or the prior belief might be specified by a normal distribution. These prior distributions ensure that the analysis appropriately includes the uncertainty associated with the calibration points. The MCMC exploration of parameter space enables the algorithm to estimate a posterior distribution of the parameters: the dates, the rates of substitution, the tree shape, the relative frequency with which the different bases are substituted for each other, and so on. In particular, the estimates of rates along each branch can be used to assess whether they are autocorrelated. The method is implemented in the computer program BEAST [49].

Uncertainty over the dates of calibration points

Calibration points have often been used without considering their errors, an approach that has recently attracted strong criticism [30,31]. Ignoring this uncertainty in the date attributed to a calibration point at the start of the analysis leads to date estimates with overly optimistically small confidence intervals [32,33], yet many thousands of published dates suffer from this fault.

Assessing the uncertainty is not trivial. In the case of dates inferred from the fossil record, there is an error associated with the process of dating particular fossils [32]. Methods such as radioisotope dating have some intrinsic error and there is additional uncertainty about the correspondence between the dated material and the fossil itself. More fundamentally, the fossil record is incomplete and, accordingly, the most ancient date at which two lineages can be detected represents only the minimum time back to their common ancestor [17]. It is difficult to assess this error, although attempts have been made [34]. This error is worse when the fossil record is patchy. For example, in the different group of bats, the times indicated by the fossils are in general 73% more recent than the times estimated using molecular data [35]. Even identifying the correct position of a fossil in a tree can be difficult [36]. It seems, then, that the very foundation of molecular dating (the calibration of the tree using the fossil record) is prone to substantial error.

There are alternatives to the fossil record for calibrating a tree. Geologists can date events such as the formation of landmasses (e.g. islands) or separation of continents, although again there are errors associated with these timing. For example the dynamics of the formation of an island or the separation of a continent are often difficult to assess [31]. But the main problem is how well the geological date corresponds to the dates at which lineages became separated [31,36]. For example, a species could split into two isolated groups, one of which goes on to colonise a newly appeared landmass much later.

Modern methods enable users to input their assessment of the uncertainty about the date of a calibration point as an upper and lower bound, or as a probability distribution [21,37,38]. In theory, it should be possible to compensate for the uncertainty in individual calibration dates by using several different calibration points, but this strategy has associated problems. Remember that we are essentially using the calibration points to infer the rate of the molecular clock, yet we know that this rate varies across the tree. The use of multiple calibration points will, therefore, only be effective and reliable if we specify correctly how the rate varies across the tree, a topic that is currently unresolved. We can be relatively confident in the dating of clades that have internal calibration points, especially those with calibration points directly above the node (branch point) of interest [33]. Unfortunately, it is rare to have such closely spaced and relevant calibration points. More often, we are attempting to extrapolate from rates estimated from one part of a phylogenetic tree to another, or from one time period to another. In this case, the information from our multiple calibration points might indicate that the rates vary, so that a properly calculated extrapolation would only produce imprecise estimates.

Please cite this article in press as: Pulquério, M.J.F. and Nichols, R.A., Dates from the molecular clock: how wrong can we be?, Trends Ecol. Evol. (2006), doi:10.1016/j.tree.2006.11.013

The illusion of accelerating clock rates

A case-in-point is the problem highlighted by Ho *et al.* [18], who show that the rates of change (substitution or mutation rates) measured over more recent times (<2 Mya) are consistently higher than those measured over longer periods stretching back to more ancient times. Although the trend is species and gene specific, the recent rates can appear over tenfold faster than those measured over longer times. Higher rates still have been estimated from ancient DNA and pedigrees of a few generations.

There are several possible explanations for this inequality. The simplest one would be substantial differences in mutation rate between different sites in the DNA sequence [16]. If some sites mutate particularly quickly, they will tend to be detected in pedigree studies, but will accumulate multiple mutations in phylogenetic studies, which will be misinterpreted as single events. Standard phylogenetic methods attempt to compensate for such rate variation by assuming a gamma distribution [39] in mutation rates, which enables some sites to mutate faster than others. This compensation might be inadequate if more sites have high rates than is suggested by this distribution. Nonetheless, there is only weak evidence for such large differences in mutation rate [18].

A second explanation relates to the action of selection and genetic drift [16,18,40]. Pedigree studies estimate the mutation rate (mutations per gamete per generation), whereas long-term evolutionary comparisons estimate the substitution rate (mutations fixed per generation). Some of the mutations detected in pedigree studies will eventually be fixed, but some will be lost by genetic drift and others will be eliminated by selection. Even comparisons between species can include mutations that have not yet been fixed [40]. If 80% of mutations were deleterious [41], this could account for the discrepancy with rates estimated over periods greater than 2 Mya for protein-encoding regions. However, the most dramatic discrepancy is not in an encoding region, but in the D-loop of mitochondrial human DNA. On the one hand, considering 80% of deleterious mutations would still give a substitution rate ~five-times higher than those obtained by Ho *et al.* [18] for primates. On the other hand, the tenfold difference in the substitution rates for humans obtained for the D-loop in different phylogenetic studies [42] does not appear to have been adequately explained, and needs further analysis.

Another possible contribution to the apparent acceleration of the clock arises because the time at which the two species split from the ancestral population is different from the time of the common ancestor of the two genes taken from those species (Figure 1). This difference varies among loci, but can amount to several hundred thousand years [43]. The extra divergence occurring during this period inflates the rate estimate, particularly over short periods, when it can make up a substantial proportion of the total divergence [44]. This effect is worse if the fossil evidence provides an underestimate of the time since the species split. These and other possible explanations [42,45] for the bias are not exclusive [46]; they need to be more fully investigated so that appropriate corrections can be made. Otherwise, a date estimate will be biased unless the molecular clock has been calibrated against known dates of a similar age.

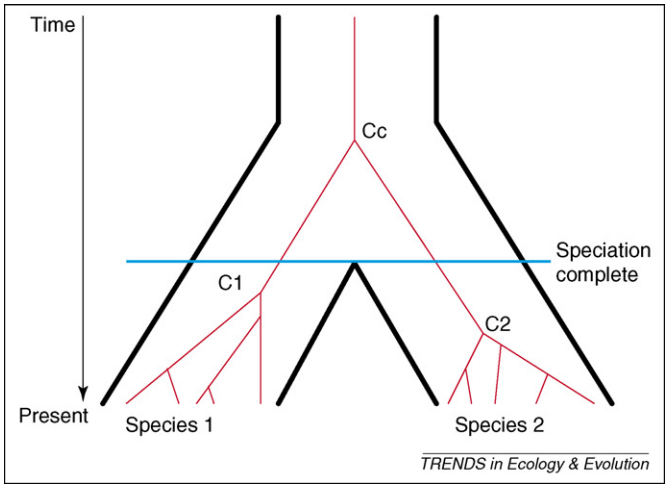

Figure 1. Species and gene trees. The black inverted 'Y' represents the ancestry of two hypothetical species (species 1 and species 2) derived from a single common ancestor. Within each species, the red lines show a gene tree representing the ancestry connecting the DNA sequences sampled from each species. The fossils used to calibrate the single node on this tree will have lived after speciation is complete (below the blue line). For example, to establish that the speciation event has occurred, the morphology of the two species must have diverged, which takes time. In addition an appropriate geological event must occur, leading to the preservation of the fossils. Such events are rare, so additional time will have passed. Conversely, the time at which the common ancestor of the genes (Cc) was alive will be before the blue line. The genes in the ancestral species will have been part of a gene tree that is similar in shape to those of the modern species (the trees descended from C1 and C2). The ancestor Cc could have lived a considerable time before speciation was complete. Although exactly how long before speciation will depend partly on chance (and on the particular history of that gene), it will, on average, be longer in species with larger or more geographically subdivided populations. The genetic differences between the species will reflect the time from Cc to the present. The discrepancy of the two dates (genetic and fossil) will be proportionately larger if the speciation is recent, contributing to the illusion of an accelerating molecular clock.

Conclusions and future directions

The answer to our initial question about dates from the molecular clock ('how wrong could we be?') appears to be that we do not yet know. In most cases, we would currently be dubious about a use for molecular clock, which required a date to be within 30% of the true value and we would be cautious about even bigger differences. The most promising approaches for analysis, such as that of Drummond *et al.* [21], allow for uncertainty in the dates attributed to calibration points and do not impose unproven assumptions about the pattern in clock-rate variation among lineages. Some pressing questions remain, many of which are not theoretical problems but require empirical investigation. If the substitution rates turn out to be autocorrelated after all, or if they are predictable from the biology of the species, then it should prove possible to exploit this knowledge.

For the moment, we need to determine the precision of clock-based date estimates when realistic errors are specified for the dates of calibration points and when appropriate allowance is made for the existence of rate variation. By combining information from many species, the recently developed Bayesian methods enable the extent and pattern of the clock-rate variation to be roughly characterised, and allowed for. When the information is available, the inclusion of additional calibration points to the analysis

Please cite this article in press as: Pulquério, M.J.F. and Nichols, R.A., Dates from the molecular clock: how wrong can we be?, Trends Ecol. Evol. (2006), doi:10.1016/j.tree.2006.11.013

should therefore produce more accurate clock-base dates, but it remains to be seen whether they will be more precise (e.g. have reduced standard errors). The difficulties of verifying clock-based dates mean that some discrepancies might previously have been overlooked or explained away. We await the more rigorous type of assessment with some nervousness, given that we suspect they might reveal that many past studies placed too much confidence in simple molecular clock analyses, and that their conclusions should thus be revisited.

Acknowledgements

We thank Maria João Cruz for comments on an earlier version of the article and three anonymous referees for helpful comments. M.J.F.P. is supported by the Portuguese Government with a FCT and FSE grant nr. SFRH/BD/17548/2004.

References

1 Zuckerkandl, E. and Pauling, L. (1965) Evolutionary divergence and convergence in proteins. In *Evolving Genes and Proteins* (Bryson, V. and Vogel, H.J., eds), pp. 97–166, Academic Press
2 Kimura, M. (1968) Evolutionary rate at the molecular level. *Nature* 217, 624–626
3 Kimura, M. (1983) *The Neutral Theory of Molecular Evolution,* Cambridge University Press
4 Springer, M.S. *et al.* (2003) Placental mammal diversification and the Cretaceous–Tertiary boundary. *Proc. Natl. Acad. Sci. U. S. A.* 100, 1056–1061
5 Ohta, T. (1993) An examination of the generation-time effect on molecular evolution. *Proc. Natl. Acad. Sci. U. S. A.* 90, 10676–10680
6 Ohta, T. (1992) The Nearly Neutral Theory of molecular evolution. *Annu. Rev. Ecol. Syst.* 23, 263–286
7 Ohta, T. (1998) Evolution by nearly-neutral mutations. *Genetica* 103, 83–90
8 Ohta, T. (2002) Near-neutrality in evolution of genes and gene regulation. *Proc. Natl. Acad. Sci. U. S. A.* 99, 16134–16137
9 Bromham, L. *et al.* (1996) Determinants of rate variation in mammalian DNA sequence evolution. *J. Mol. Evol.* 43, 610–621
10 Bromham, L. (2002) Molecular clocks in reptiles: life history influences rate of molecular evolution. *Mol. Biol. Evol.* 19, 302–309
11 Gillooly, J.F. *et al.* (2005) The rate of DNA evolution: effects of body size and temperature on the molecular clock. *Proc. Natl. Acad. Sci. U. S. A.* 102, 140–145
12 Martin, A.P. and Palumbi, S.R. (1993) Body size, metabolic-rate, generation time, and the molecular clock. *Proc. Natl. Acad. Sci. U. S. A.* 90, 4087–4091
13 Rand, D.M. (1994) Thermal habit, metabolic-rate and the evolution of mitochondrial-DNA. *Trends Ecol. Evol.* 9, 125–131
14 Bromham, L. *et al.* (1999) Growing up with dinosaurs: molecular dates and the mammalian radiation. *Trends Ecol. Evol.* 14, 113–118
15 Bromham, L. *et al.* (2000) The power of relative rates tests depends on the data. *J. Mol. Evol.* 50, 296–301
16 Sigurgardottir, S. *et al.* (2000) The mutation rate in the human mtDNA control region. *Am. J. Hum. Genet.* 66, 1599–1609
17 Benton, M.J. and Ayala, F.J. (2003) Dating the tree of life. *Science* 300, 1698–1700
18 Ho, S.Y.W. *et al.* (2005) Time dependency of molecular rate estimates and systematic overestimation of recent divergence times. *Mol. Biol. Evol.* 22, 1561–1568
19 Yoder, A.D. and Yang, Z.H. (2000) Estimation of primate speciation dates using local molecular clocks. *Mol. Biol. Evol.* 17, 1081–1090
20 Thorne, J.L. and Kishino, H. (2002) Divergence time and evolutionary rate estimation with multilocus data. *Syst. Biol.* 51, 689–702
21 Drummond, A.J. *et al.* (2006) Relaxed phylogenetics and dating with confidence. *PLoS Biol.* 4, 699–710
22 Rutschmann, F. (2006) Molecular dating of phylogenetic trees: a brief review of current methods that estimate divergence times. *Div. Distrib.* 12, 35–48
23 Welch, J.J. and Bromham, L. (2005) Molecular dating when rates vary. *Trends Ecol. Evol.* 20, 320–327
24 Aris-Brosou, S. and Yang, Z.H. (2002) Effects of models of rate evolution on estimation of divergence dates with special reference to the metazoan 18S ribosomal RNA phylogeny. *Syst. Biol.* 51, 703–714
25 Sanderson, M.J. (2002) Estimating absolute rates of molecular evolution and divergence times: a penalized likelihood approach. *Mol. Biol. Evol.* 19, 101–109
26 Thorne, J.L. *et al.* (1998) Estimating the rate of evolution of the rate of molecular evolution. *Mol. Biol. Evol.* 15, 1647–1657
27 Blair, J.E. and Hedges, S.B. (2005) Molecular phylogeny and divergence times of deuterostome animals. *Mol. Biol. Evol.* 22, 2275–2284
28 Rodriguez-Trelles, F. *et al.* (2001) Erratic overdispersion of three molecular clocks: GPDH, SOD, and XDH. *Proc. Natl. Acad. Sci. U. S. A.* 98, 11405–11410
29 Ayala, F.J. (1999) Molecular clock mirages. *Bioessays* 21, 71–75
30 Graur, D. and Martin, W. (2004) Reading the entrails of chickens: molecular timescales of evolution and the illusion of precision. *Trends Genet.* 20, 80–86
31 Heads, M. (2005) Dating nodes on molecular phylogenies: a critique of molecular biogeography. *Cladistics* 21, 62–78
32 van Tuinen, M. and Hadly, E.A. (2004) Error in estimation of rate and time inferred from the early amniote fossil record and avian molecular clocks. *J. Mol. Evol.* 59, 267–276
33 van Tuinen, M. and Hadly, E.A. (2004) Calibration and error in placental molecular clocks: a conservative approach using the cetartiodactyl fossil record. *J. Hered.* 95, 200–208
34 Marshall, C.R. (1990) Confidence-intervals on stratigraphic ranges. *Paleobiology* 16, 1–10
35 Teeling, E.C. *et al.* (2005) A molecular phylogeny for bats illuminates biogeography and the fossil record. *Science* 307, 580–584
36 Magallon, S.A. (2004) Dating lineages: molecular and paleontological approaches to the temporal framework of clades. *Int. J. Plant Sci.* 165, S7–S21
37 Kishino, H. *et al.* (2001) Performance of a divergence time estimation method under a probabilistic model of rate evolution. *Mol. Biol. Evol.* 18, 352–361
38 Yang, Z. and 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
39 Yang, Z. (1994) Maximum-likelihood phylogenetic estimation from DNA-sequences with variable rates over sites – approximate methods. *J. Mol. Evol.* 39, 306–314
40 Penny, D. (2005) Evolutionary biology – relativity for molecular clocks. *Nature* 436, 183–184
41 Fay, J.C. *et al.* (2001) Positive and negative selection on the human genome. *Genetics* 158, 1227–1234
42 Santos, C. *et al.* (2005) Understanding differences between phylogenetic and pedigree-derived mtDNA mutation rate: a model using families from the Azores Islands (Portugal). *Mol. Biol. Evol.* 22, 1490–1505
43 Nichols, R. (2001) Gene trees and species trees are not the same. *Trends Ecol. Evol.* 16, 358–364
44 Edwards, S.V. and Beerli, P. (2000) Perspective: gene divergence, population divergence, and the variance in coalescence time in phylogeographic studies. *Evolution* 54, 1839–1854
45 Galtier, N. *et al.* (2006) Mutation hot spots in mammalian mitochondrial DNA. *Genome Res.* 16, 215–222
46 Howell, N. *et al.* (2003) The pedigree rate of sequence divergence in the human mitochondrial genome: there is a difference between phylogenetic and pedigree rates. *Am. J. Hum. Genet.* 72, 659–670
47 Cherty, L.M. *et al.* (1978) Frog perspective on the morphological difference between humans and chimpanzees. *Science* 200, 209–211
48 Ohta, T. (2003) Origin of the neutral and nearly neutral theories of evolution. *J. Biosci.* 28, 371–377
49 Drummond, A.J. and Rambaut, A. (2003) *BEAST v1.0*, available from http://evolve.zoo.ox.ac.uk/beast/

Please cite this article in press as: Pulquério, M.J.F. and Nichols, R.A., Dates from the molecular clock: how wrong can we be?, Trends Ecol. Evol. (2006), doi:10.1016/j.tree.2006.11.013