Molecular clocks: when times are a-changin’

Simon Y.W. Ho and Greger Larson

Department of Zoology, University of Oxford, Oxford, UK, OX1 3PS

The molecular clock has proved to be extremely valuable in placing timescales on evolutionary events that would otherwise be difficult to date. However, debate has arisen about the considerable disparities between molecular and palaeontological or archaeological dates, and about the remarkably high mutation rates inferred in pedigree studies. We argue that these debates can be largely resolved by reference to the ‘time dependency of molecular rates’, a recent hypothesis positing that short-term mutation rates and long-term substitution rates are related by a monotonic decline from the former to the latter. Accordingly, the extrapolation of rates across different timescales will result in invalid date estimates. We examine the impact of this hypothesis with respect to various fields, including human evolution, animal domestication and conservation genetics. We conclude that many studies involving recent divergence events will need to be reconsidered.

The inconstant molecular clock

The molecular clock has become an indispensable tool within evolutionary biology, enabling independent timescales to be placed on evolutionary events. Despite these valuable contributions, date estimates derived from molecular data have not been without controversy. In particular, when molecular clocks have been employed to estimate the timing of recent events already tentatively dated on the basis of palaeontological, archaeological or biogeographic sources, conflicting dates are frequently obtained [1]. Some of the most significant discrepancies have been evident in studies of human evolution, animal domestication and conservation genetics. Previous criticisms of divergence-dating methodology have focused on faulty calibration points [2] and the impact of rate heterogeneity among lineages [3]. Although these points remain valid, many of the disagreements are still unresolved. As we explain in the following sections, we believe that the explanation lies in another major methodological problem that has recently come to light.

It has long been recognized that violations of the strict molecular clock are commonplace [4]. More recently, however, several studies have noted that molecular rates observed on genealogical timescales [<1 million years (My)] are an order of magnitude or more than those measured over geological time (>1 My). In the avian mitochondrial genome, for example, a mutation rate of 95% per My was estimated from the control region of Adélie penguins [5], which is considerably greater than the paradigmatic substitution rate of 1% per My observed among various avian taxa [6–8], even after acknowledging the accepted rate difference between non-coding and protein-coding DNA. Elevated mutation rates (~30% per My) have also been estimated from the control region of recently diverged species within the Bison [9] and Bos [10,11] genera. In the mitochondrial control region of humans, Parsons et al. [12] estimated a mutation rate of 250% per My, whereas a meta-analysis [13] of data pooled from eight pedigree studies yielded an overall rate estimate of 95% per My. Both of these pedigree-based estimates exceed those from phylogenetic studies [14,15] by up to two orders of magnitude.

We argue that these debates actually stem from the same phenomenon: the ‘time dependency of molecular rates’, which has been described in several recent articles [16–18]. The observed rate at which molecular clocks ‘tick’ is not entirely constant over time. Instead, there is a measurable transition from an increased, short-term mutation rate to a low, long-term substitution rate (see Glossary), and this trend can be described mathematically (Box 1). Failure to distinguish between the mutation rate and the substitution rate, and to consider the relationship between the two, is most likely to be the key factor responsible for several prominent controversies in evolutionary biology. We examine some of these in the following sections and explain how these disputes can be resolved by considering the time dependency of molecular rates before discussing the implications for other studies.

Glossary

Mutation rate: the instantaneous rate at which nucleotide changes occur in the genome. Lethal or near-lethal mutations are often ignored in calculations of the mutation rate.

Pedigree rate: an estimate of the mutation rate, assessed by calculating the number of nucleotide changes observed over a known number of reproductive events (based on a known genealogy of individuals and a given or assumed generation time).

Substitution rate: the rate at which mutations are fixed in the population. Because most nucleotide changes (mutations) that appear within a population are eventually eliminated (by purifying or background selection or by drift), there will be fewer observed changes per unit of time. As a result, the substitution rate will always be slower than the mutation rate (except under perfectly neutral conditions).

Phylogenetic rate: an estimate of the substitution rate, calculated by comparing molecular sequence data obtained from different species.
Human migration and origins

Many of the current hypotheses concerning human migration events are based on molecular date estimates. Most of these date estimates rely on two main assumptions: (i) that the human–chimpanzee divergence occurred between four and six million years ago (Mya); and (ii) that the substitution rate inferred from this calibration has remained the same throughout the comparatively recent generation of human genetic diversity. We believe that this second assumption does not hold and that the application of an inappropriately low rate leads to an overestimation of molecular dates.

One prominent example in which this occurs is in studies of the peopling of the Americas, which archaeological evidence places at ~10–12 thousand years ago (kya) with the appearance of the Clovis culture [19]. However, when Ward et al. [20] applied the human–chimpanzee substitution rate to Amerindian sequences, they found that the tribal lineages diverged from each other ~60 kya. On the basis of this evidence, they rejected the notion of a founder effect in the peopling of the Americas. In a similar study, Horai et al. [21] eliminated the possibility of post-migration genetic differentiation and concluded that Native Americans must have entered the New World in several waves.

It is clear that the extrapolation of phylogenetic rates onto population-level data is a recurrent error, but the opposite scenario (the extrapolation of pedigree rates to deeper timescales) also occurs. To place a timescale on the unique divergence between African and non-African humans, Armour et al. [22] applied a relatively rapid pedigree rate to the data and concluded that this event occurred only 15 kya. This date falls well short of the archaeologically defined first appearance of humans in both Europe (40 kya [23]) and Israel (90 kya [24]). The scenarios described here represent only the tip of the iceberg. The timing of many other events in human evolution, such as the origin of HIV and other viruses, the divergence between humans and Neanderthals and the dispersals of other hominids, might also require revision.

One of the most familiar debates in human evolution is the age of the most recent female ancestor of all existing human mitochondrial lineages, commonly referred to as ‘mitochondrial Eve’. Molecular analyses using a human–chimpanzee calibration have placed her age at 150–250 kya [25–27] (Table 1), whereas using a rate derived from pedigree studies yields a date estimate that is patently too recent (e.g. 6 kya [1,28]). In this particular example, the dates estimated on the basis of an intraspecific calibration [29] are probably closest to the truth. The provisional date estimates proposed by Ho et al. [17], which were obtained using a method that accommodates the time dependency of molecular rates, also demonstrate that the actual age of mitochondrial Eve most probably lies between the estimates based on the pedigree (mutation) and phylogenetic (substitution) rates (Box 1).

**Domestication**

Studies of the timing of domestication have persistently generated the greatest inconsistencies between...
archaeological and molecular date estimates. To discriminate between wild and domestic animals, and thus establish when and where domestication began, archaeozoologists have generally relied on the morphological differentiation brought on by the domestication process. This reliance on phenotypic change biases archaeologically based dates towards the later stages of domestication. Because molecular differentiation would have begun as soon as domestic populations were no longer interbreeding with their wild counterparts, the estimates using molecular clocks were expected to pre-date those inferred from archaeological remains. The surprising results revealed that, for virtually every major domestic mammal, molecular date estimates exceeded the accepted archaeological dates by tens or even hundreds of thousands of years (Table 2).

The simplest explanation for this discrepancy is that the molecular dates reflect the splits among multiple wild lineages, all (or most) of which were subsequently domesticated (e.g. Ref. [30]). Most of the studies cited in Table 2 at least acknowledged this possibility, but even those that controlled for this by calculating the age of a single clade containing only domestic animals still produced molecular estimates that significantly contradicted the archaeological record [31,32].

A second general explanation for the vast disparities between molecular and fossil-based date estimates is the problem associated with calibration points. Although this issue applies to all divergence date analyses, it is particularly relevant for domestication studies, given the relatively short time frame of domestication and the large uncertainty surrounding splits between wild progenitors and their sister species. In coyotes and wolves, for example, Kurten [33] places the split at 1.5–4.5 Mya based on the fossil record. Even employing the most recent extreme of the fossil-based ranges (thus producing the fastest rate) for any of the domestic animals in Table 2, however, does not resolve the discrepancy between the molecular and archaeological estimates.

The most likely explanation for the discrepancy between the dates is that substitution rates derived from phylogenetic divergences have been incorrectly used to date splits among conspecific populations. The projection of a rate estimated using palaentologically derived splits between outgroup and wild progenitor taxa (all of which are >1 My) onto sequences obtained from populations of modern domestic animals have consistently produced artificially deep time estimates for the domestication of many animals, including dogs [31], donkeys [34] and sheep [35]. Other domestication studies (e.g. pigs [36] and sheep [37]) that imported general substitution rates derived from mammalian taxa might have also overestimated the timing of the coalescence of domesticated lineages. In turn, these practices might necessitate the postulation of either many independent domestication events or numerous domesticated wild lineages [30].

It remains possible that the earliest phases of domestication began tens of thousands of years ago and that the phenotypic changes associated with the process are a relatively recent phenomenon. What is more likely, however, is that the highly controversial discrepancies are largely the spurious product of invalid extrapolations of molecular rates.

**Implications for other fields**

The inappropriate application of substitution rates to intraspecific evolutionary questions has also occurred in numerous other fields of biology. Because this methodology results in an overestimation of the time required to attain a given level of genetic diversity, the time dependency of molecular rates has a substantial impact on conservation genetics. A study on humpback whales, for example, calculated a molecular rate on the basis of calibration points at 6–25 My. This rate was used to conclude that existing populations of humpback whales underwent ancient divergences [38]. Menotti-Raymond and O’Brien [39] applied a felid substitution rate to cheetah mitochondrial sequences, estimating that a genetic bottleneck occurred at 28–36 kya. More recently, Eizirik et al. [40] used an almost identical approach to date the coalescence of modern jaguar lineages. In both of these felid studies, and in others based on the same methodology, the dates were probably overestimated, which has significant consequences for conservation strategies.

Studies tying evolutionary events to climatic changes in the Pleistocene period are also susceptible to the misuse of long-term substitution rates in the analysis of population-level data. For example, the classic 1% per My avian mitochondrial substitution rate was applied to pairs of songbird subspecies (e.g. Myrtle and Audubon’s warblers) to reject the ‘late Pleistocene origins’ hypothesis, which associates the origins of North American songbirds with Pleistocene glacial events [41]. Correlations of megafaunal-extinction events with either climate change or human activity are highly dependent on accurate estimates of divergence dates, implying that these are particularly sensitive to inappropriately applied substitution rates. By employing a felid substitution rate in the analysis of puma sequences, Culver et al. [42] inferred the coincidence of recolonisation of North America by pumas and a widespread megafaunal-extinction event.

Population genetics and demographics are directly affected by the time dependency of molecular rates because simplifying assumptions about rates must be made in order to separate them from the effects of population size. A recent study [43] on the size of the American founding population extrapolated the human–chimpanzee mutation rate, resulting in an estimated effective size of 80
individuals. If a more appropriate, intraspecific human mutation rate had been used, the effective population size would have been even lower, given the reciprocal relationship between population size and mutation rate (for a given value of the population mutation rate parameter, $\theta$, which is proportional to the product of the mutation rate and the effective population size).

It is possible to list many studies from other fields, such as biogeography [44], in which molecular rates have been inappropriately extrapolated across the population-species boundary. However, the diverse catalogue of examples presented here already provides a cogent indication that the problem is widespread.

**Dating recent events**

It is clear that the direct application of a strict molecular clock across the population-species boundary (in either direction) is unjustifiable, and that the time dependency of molecular rates complicates the process of molecular date estimation. To estimate the timing of recent events in a valid manner, it is paramount to use a method that distinguishes mutation rates from substitution rates. Several different options are apparent, including but not limited to the following:

(i) Use calibration points that are as close as possible to the date being estimated (e.g. intraspecific calibration points for population-level studies [29,45]). Given the time dependency of molecular rates, failure to do so will result in a greater discrepancy between the true rate and the rate that is used in the analysis. This approach could be somewhat circular, however, because it is necessary to have a reasonable idea of what the true date might be to select a suitable calibration;

(ii) If reliable ancient DNA data are available, use radiocarbon-dated sequences [9,46] to decrease the amount of extrapolation required to estimate the age of the root of the tree, as in (i);

(iii) Estimate dates in a relaxed clock framework that permits molecular rates to vary among branches [47], enabling terminal branches to have faster rates (this only works if there is reasonably dense taxon sampling);

(iv) Use an accurately estimated molecular rate curve (Box 1) to derive the rate needed for the timescale in question [17].

Although the last of these methods represents the most straightforward approach, it is not an altogether realistic option owing to the extreme difficulty in obtaining accurate rate curves and because individually generated rate curves cannot (and should not) be generalized across unrelated taxa.

**Concluding remarks**

In the midst of widespread confusion and criticism concerning molecular date estimates of recent evolutionary events, we believe that many outstanding inconsistencies can be explained by the time dependency of molecular rates. Although it would not be reasonable to expect that molecular date estimates should always be identical to those from other sources (considering not only the biased and incomplete nature of the fossil and archaeological records, but also the difference between gene trees and species trees), we should expect that increasingly realistic genetic models will reduce the discrepancy.

The examples provided in this article highlight the difficulty of producing accurate date estimates but suggest that meaningful dates can be obtained from molecular data if the analysis is done correctly. However, the time dependency of rates also provides a strong warning against extrapolating molecular rates across the population-species boundary, unless the transition is well understood and has been quantified. Unfortunately, many current hypotheses about human evolution, domestication, conservation genetics and human demographic history rest on date estimates that have been improperly calculated. Clearly, many previous studies dealing with recent timescales will need to be re-evaluated.

**Acknowledgements**

We thank P. Holland, B. Shapiro, A. Rambaut, and three anonymous reviewers for valuable comments on this article, and A. Cooper, A. J. Drummond, D. Penny and R. Barnett for advice. We also thank the Leverhulme Trust, the Arts and Humanities Research Board, the Commonwealth Scholarship Commission and Linacre College, Oxford, for financial support.

**References**

37 Hiendleder, S. et al. (1998) Analysis of mitochondrial DNA indicates that domestic sheep are derived from two different ancestral maternal sources: No evidence for contributions from urial and argali sheep. J. Hered. 89, 113–120

Elsevier joins major health information initiative
Elsevier has joined with scientific publishers and leading voluntary health organizations to create patientINFORM, a groundbreaking initiative to help patients and caregivers close a crucial information gap. patientINFORM is a free online service dedicated to disseminating medical research and is scheduled to launch in 2005.

Elsevier will provide the voluntary health organizations with increased online access to our peer-reviewed biomedical journals immediately upon publication, together with content from back issues. The voluntary health organizations will integrate the information into materials for patients and link to the full text of selected research articles on their websites.

patientINFORM has been created to allow patients seeking the latest information about treatment options online access to the most up-to-date, reliable research available for specific diseases.

‘Not only will patientINFORM connect patients and their caregivers with the latest research, it will help them to put it into context. By making it easier to understand research findings, patientINFORM will empower patients to have a more productive dialogue with their physicians and make well-informed decisions about care’, said Harmon Eyre, M.D., national chief medical officer of the American Cancer Society.

For more information, visit www.patientinform.org