Opinion TRENDS in Genetics Vol.20 No.2 February 2004

Reading the entrails of chickens: molecular timescales of evolution and the illusion of precision

Dan Graur[1] and William Martin[2]

[1]Department of Biology and Biochemistry, University of Houston, Houston, TX 77204-5001, USA
[2]Institut für Botanik III, Heinrich-Heine Universität Düsseldorf, Universitätsstraße 1, 40225 Düsseldorf, Germany

For almost a decade now, a team of molecular evolutionists has produced a plethora of seemingly precise molecular clock estimates for divergence events ranging from the speciation of cats and dogs to lineage separations that might have occurred ~4 billion years ago. Because the appearance of accuracy has an irresistible allure, non-specialists frequently treat these estimates as factual. In this article, we show that all of these divergence-time estimates were generated through improper methodology on the basis of a single calibration point that has been unjustly denuded of error. The illusion of precision was achieved mainly through the conversion of statistical estimates (which by definition possess standard errors, ranges and confidence intervals) into errorless numbers. By employing such techniques successively, the time estimates of even the most ancient divergence events were made to look deceptively precise. For example, on the basis of just 15 genes, the arthropod–nematode divergence event was 'calculated' to have occurred 1167 ± 83 million years ago (i.e. within a 95% confidence interval of ~350 million years). Were calibration and derivation uncertainties taken into proper consideration, the 95% confidence interval would have turned out to be at least 40 times larger (~14.2 billion years).

'We demand rigidly defined areas of doubt and uncertainty.' Douglas Adams

People have always been fascinated with dating the past, particularly in the absence of historical records. James Ussher (1581–1656), Archbishop of Armagh and Primate of All Ireland, is considered the first scholar to have employed internal (biblical) and external (astronomical) evidence to date events that were considered undatable by his predecessors. In his *Annales Veteris Testamenti* (*Annals of the Old Testament*) published in 1650, Ussher established the first day of creation as Sunday 'upon the 23 day of the Julian October' 4004 BC. With similar precision, he dated Adam and Eve's expulsion from Paradise, the destruction of Sodom and Gomorra and the landing of Noah's ark on Mount Ararat. Generations of scholars were so captivated by the appearance of precision of these dates, that hardly anyone questioned their veracity.

In a modern rendition of Ussher's feat, a team of molecular evolutionists has inferred ostensibly precise molecular-clock dates for speciation events ranging from the divergence between cats and dogs to the early diversification of prokaryotes [1–12]. The findings were summarized in a *Trends in Genetics* review [13]. With few exceptions [14–24], it has escaped the notice of most readers that all these divergence-time estimates are based on a single calibration point and tenuous methodology. In this article, we document the manner in which a calibration point that is both inaccurate and inexact – and in many instances inapplicable and irrelevant – has been used to produce an exhaustive evolutionary timeline that is enticing but totally imaginary. We will relate a dating saga of ballooning inapplicability and snowballing error through which molecular equivalents of the 23rd October 4004 BC date have been mass-produced in the most prestigious biology journals.

Chapter 1: the origin of the primary 310 ± 0 million-year calibration

The saga starts with 'an accurate calibration point' for obtaining 'reliable estimates of divergence times from molecular data' [1]. From among the many calibration points available in the paleontological literature, 'the relatively well-constrained fossil divergence time between the ancestor of birds (diapsid reptiles) and mammals (synapsid reptiles)' was selected [1]. This divergence time was said to be 310 million years ago (MYA). As a calibrating measurement, the 310-MYA value is treated as extremely accurate and extremely precise. That is, the divergence time estimate between diapsids and synapsids is used as if it had neither *directional* nor *random* errors around the mean (i.e. 310 ± 0 MYA). We note that this number is extremely important because all subsequent divergence-time estimates are based on it.

In an article by Hedges *et al.* [1], the 310 ± 0 MYA date was said to be derived from Volume 2 of *The Fossil Record* edited by M. J. Benton. Because *The Fossil Record* is a one-volume publication, we assume that the authors intended to cite the second edition from 1993 [25]. Because no page numbers were given, we carefully read the relevant portions of the 845-page tome and found no mention of

Corresponding author: William Martin (w.martin@uni-duesseldorf.de).

 0168-9525/$ - see front matter doi:10.1016/j.tig.2003.12.003

'310 MYA'. In fact, in the preface it is written that although the contributing authors 'were invited to use any stratigraphical scheme that they thought appropriate,' such a request 'involved no consideration of the exact ages in million of years.' Moreover, chronological dates were strictly avoided throughout the book, apparently as a matter of principle. This practice is reasonable and commendable because placing a fossil in a stratigraphical context can be done in some cases with relative ease, whereas assigning absolute dates to a stratum can be difficult and controversial.

There are at least three sources of error that can affect the 310 ± 0 MYA calibration point: (i) errors in the topology of the phylogenetic tree; (ii) errors in the taxonomic identification of the fossil material; and (iii) errors in the chronological assignment of the geological strata [16,24]. The first of these possible errors (Figure 1) will only affect the dating of misassigned branches; the other two can have more profound effects. The 310-MYA value was based on the assumption that *Hylonomus* from the Vereiskian stage of the Carboniferous period is the oldest diapsid, and that *Archaeothyris* from the Myachkovskian stage is the oldest synapsid (S. Blair Hedges, pers. commun.). (The type specimens of *Hylonomus* and *Archaeothyris* were discovered in Joggins, Nova Scotia and Florence, Nova Scotia, respectively.) In the geological literature, the time boundaries of the Vereiskian stage are given as 311.2–309.2 MYA and those for the Myachkovskian stage are 305–303 MYA [26].

Because *Hylonomus* is considered to be the older of the two, only its contribution to the dating of the synapsid–diapsid divergence event needs to be examined. First, it is impossible to state with certainty that *Hylonomus* is a diapsid [16]. The reason is that, although all taxonomically diagnostic characters of synapsids and diapsids are cranial (i.e. detectable mainly in the postorbital and squamosal bones; Figure 1), the vast majority of the relevant fossil specimens are represented by uninformative postcranial fragments [16]. Even 'the best-preserved cranial material' lacks 'the diagnostic squamosal and only possesses an incomplete postorbital" and, in general, 'the manner of preservation at Joggins makes a systematic description of these reptiles difficult' [27,28]. Thus, one should not rely on *Hylonomus* as the sole determinant of the calibration point. The earliest undisputed fossils that are relevant to dating the synapsid–diapsid divergence are from the uppermost Pennsylvanian epoch, ~290 MYA [16,29–31].

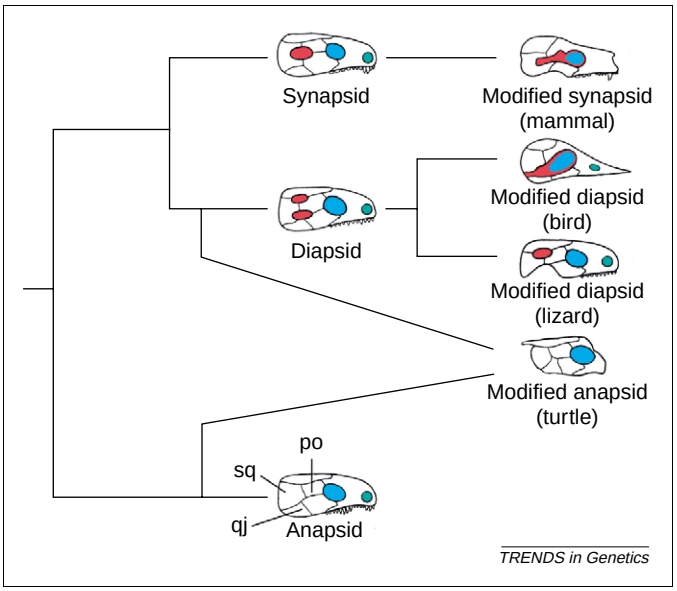

Figure 1. Phylogenetic relationships among amniotes. According to the number of temporal openings or fenestrae (in red), amniotes are traditionally divided into anapsids, diapsids and synapsids [47]. In many extant taxa, this schematic division only holds approximately because openings might have merged, disappeared or closed during evolution. For example, molecular data indicate that turtles might in fact be diapsids [50,51], although morphologically they look undoubtedly anapsid. It is, therefore, possible that synapsids and diapsids are not sister taxa. Owing to the loss of the quadratojugal (qj) bone, lizards have a single temporal opening similar to that of synapsids, although there is little doubt as to their diapsid status. The principal diagnostic characters for classifying amniotes into anapsids, diapsids or synapsids are the postorbital (po) and squamosal (sq) bones. In fossil material, these bones can be absent or incomplete, resulting in uncertain taxonomic assignment.

Let us now deal with the question of stratigraphical dating. Is it possible to date the Joggins formation in Nova Scotia (or any other stratum) with the precision and accuracy implied by the errorless 310 ± 0 MYA value? The answer is negative both in principle and in practice. First, no geological dating is without error [32]. Second, the Joggins tetrapods (and formation) have no absolute age assignment, and the various estimates range from 316.5 to 308.0 MYA, with errors of 4.4–10.5 MYA [33]. Thus, given our present knowledge, a sensible calibration for the synapsid–diapsid divergence would have to be defined as a range, for example from 338 MYA as the earliest estimate [34] to 288 MYA as the most recent estimate [16]. We note that at 313 MYA, the mean remains close to the 310 MYA value. However, using the mean on its own is misleading. At the very least, the '310 ± 0 MYA' is bibliographically misattributed and spuriously stripped of error.

A solution to the single-calibration conundrum would be to use multiple primary calibrations because such practices yield better results than those obtained by relying on a single point [35–38]. Indeed, it was stated that 'the use of multiple calibration points from the fossil record would be desirable if they were all close to the actual time of divergence' [4]. However, because no calibrations other than the 310 ± 0 MYA value were ever used in this saga, the authors must have concluded that none exists. This is not true. Moreover, deciding whether a certain fossil is 'close to the actual time of divergence' presupposes a prior knowledge of the time of divergence, which in turn will make the fossil superfluous for dating purposes.

The basic methodology for converting calibration dates and comparative protein data into divergence-time estimates is outlined in Box 1. By using the 310 ± 0 MYA calibration and 19 proteins that were deemed to have evolved at approximately constant rates, an estimate of 97 ± 12 (mean ± standard error) MYA for the divergence of avian orders was obtained [1]. The use of standard errors as error bars is highly misleading [39,40]. A more appropriate practice is to calculate the 95% or 99% confidence intervals, which in this case turn out to be 121–73 MYA and 128–66 MYA, respectively. Even these large time ranges are still much narrower than they should be because the uncertainties surrounding the 310 ± 0 MYA calibration were not taken into account.

Box 1. From calibration dates and comparative protein data to divergence-time estimates

From two homologous protein sequences and by assuming a Poisson process, we can estimate the number of amino acid replacements between the two sequences as:

$$d = -\ln\left(1 - \frac{n}{L}\right) \quad \text{[Eqn 1]}$$

where n is the number of amino acid differences between the aligned sequences and L is the length of the ungapped alignment. The variance of d, $V(d)$, is given by $V(d) = (n/L)/([1 - (n/L)]n)$. It can become large if n/L is large or if L is small. Most molecular clock estimates use the mean d, but $V(d)$ also exists, regardless of how d is calculated [42].

The rate of replacement is:

$$r = \frac{d}{2T} \quad \text{[Eqn 2]}$$

where T, the time of divergence between the two sequences, is usually inferred from paleontological data.

Under the assumption that all lineages in a study evolve at the same rate, and assuming that we know the divergence time between two taxa (T_{cal} = calibration time), we can use the number of amino acid replacements between two sequences from these two taxa (d_{cal}) to calculate a universal rate as:

$$r_{cons} = \frac{d_{cal}}{2T_{cal}} \quad \text{[Eqn 3]}$$

We can, then, take any pair of sequences from any two taxa, estimate d, and calculate the time of divergence as:

$$T = \frac{d}{2r_{const}} \quad \text{[Eqn 4]}$$

Note that the uncertainties associated with T should be large because they are a composite of the errors associated with all the estimates used in the calculation of T. If one turns T into a secondary T_{cal}., then the errors associated with the estimates from the new T_{cal} should be greater than those associated with estimates derived from the primary T_{cal}.

Note also that all estimates of T are estimates for the divergence between the protein sequences used in the calculations. T can be regarded as an estimate of divergence time, only if the genes are orthologous (i.e. related by common ancestry). If the genes are paralogous (i.e. related by gene duplication) and if this duplication occurred in any of the ancestors of the two taxa from which the proteins were derived, then the divergence time will be overestimated, and speciation events will seem more ancient than they actually are [20]. It is usually difficult to ascertain with confidence whether the sequence similarity between two proteins is due to orthology or paralogy.

By considering the uncertainties in the primary calibration (i.e. 338–288 MYA), the diversification of avian orders can be said to have occurred with 95% confidence within the time interval of 132–67 MYA. The divergence times between primates and rodents, primates and artiodactyls and artiodactyls and rodents were estimated originally be 95 ± 7, 90 ± 8, and 113 ± 9 MYA, respectively [1]. Taking the uncertainties in calibration into account would have yielded 95% confidence intervals of 119–74, 117–67 and 145–85 MYA, respectively. These values do not support the conclusion that the ordinal diversification of birds and mammals coincided with the Mesosoic continental breakup [1]. Of course adding other sources of error, such as the rate variation among lineages, the variation in rates with time, the variation in substitution rates among sites, in addition to uncertainties in orthology assignment, would have increased the intervals even further [41–43].

Chapter 2: the origin of the secondary 110 ± 0 million-year calibration

The second chapter purports to estimate 'a molecular timescale for vertebrate evolution' [3], although what are in fact estimated are divergence times between humans and other organisms. Because of a lack of sufficient molecular data from chicken, the primary calibration point can not be used for the vast majority of protein comparisons. Kumar and Hedges [3], thus, opted for a secondary calibration point. (Secondary calibration points are divergence-time estimates that have been derived from one molecular dataset on the basis of a single primary external calibration point, and which are subsequently denuded of their variances and used as if they were independent calibration points on a second dataset [21]).

It is at this point in the narrative that two momentous alterations were made. First, for unexplained reasons, the divergence between primates and rodents, which was previously 95 ± 7 MYA, becomes 110 MYA, despite the difference between 95 ± 7 MYA and 110 MYA being statistically significant. (To add to the confusion, this new value is sometimes reported as '100 MYA' [44] or '112 MYA' [45].) The second alteration concerns the disappearance of the uncertainties around the mean of the secondary calibration (including those meager uncertainties stemming from the use of the erroneously errorless 310 ± 0 MYA primary calibration. These two alterations yield a new estimate: 110 ± 0 MYA for the divergence between humans and rodents. We confess that we can not pinpoint exactly the source of the 110 ± 0 MYA value but are certain that it is neither independent nor errorless. As the saga proceeds, fewer and fewer estimates are derived from the original synapsid–diapsid calibration event and – subtly but surely – more and more are based on the secondary tertiary and higher-order derivations.

A study examining the 110 ± 0 MYA point used a 'consistency test' to assess the appropriateness of this calibration [21]. Passing the consistency test entailed meeting two conditions: (i) the estimates for the divergence between birds and mammals had to be larger than the estimates for the divergence between primates and rodents (i.e. time reversibility was not allowed); (ii) the mean inferred time of divergence between birds and mammals should not differ significantly from 310 MYA (i.e. by using the secondary calibration, a divergence time that is close to the primary calibration should be recovered). The results indicated that 25% of the homologous protein sets in birds and mammals failed the first part of the consistency test [21], that is, in one out of four cases the data yielded divergence times between rodents and primates that were older than those obtained for the divergence between synapsids and diapsids. One protein yielded the absurd estimate of 2333 MYA for the

human–chicken divergence event, and as an extreme outlier [46] was discarded. For the remaining proteins, the mean bird–mammalian divergence estimate was 393 MYA with a 95% confidence interval of 471–315 MYA. In other words, the 310 MYA landmark was not recovered. Because neither condition of the consistency test was met, it was concluded that the use of the secondary calibration is unjustified.

In the study by Kumar and Hedges [3], 33 divergence-time estimates were produced. The oldest estimate was 564 ± 74.6 MYA for the divergence between jawless fish (Agnatha) and humans. By using these data and by taking into account the uncertainties associated with the primary and secondary calibrations, we estimate with 95% confidence that the ancestors of humans diverged from the ancestors of agnathans 790–232 MYA (i.e. a period of time spanning the upper Precambrian, Cambrian, Ordovican, Silurian, Devonian, Carboniferous, Permian and lower Triassic). Not surprisingly, the fossil record pinpoints the divergence events between agnathes and humans to a significantly narrower range [47].

Chapter 3: transubstantiation of a secondary calibration into a primary calibration

The reason for the transformation of the secondary calibration date into a primary one, which is equivalent to blood becoming Cabernet Sauvignon, is purportedly based on external evidence: 'Fossil evidence (Archibald 1996) also supports an early divergence time (>90 Ma) for the primate-rodent split' [4]. Unfortunately, the *Science* article by Archibald [48] is entitled *Fossil Evidence for a Late Cretaceous Origin of "Hoofed" Mammals*, and as such deals with neither rodents nor humans. Thus, the mysterious transformation is founded on a numerical interpretation of a schematic figure in Ref. [48], in which rodents and primates are drawn merely as outgroups, although they are neither mentioned in the text nor referenced in the bibliography.

The harvest in this chapter consists of six divergence time estimates that are as ancient as they are ambitious [4]. Two of the six estimates (i.e. for the fungi–plant divergence event and for the divergence between nematodes and arthropods) constitute important precedents in molecular evolution because the lineages being studied do not contain the taxa used for the calibration. (The standard for such practices was most probably set by a seldom-quoted study in which the evolutionary age of Galápagos iguanas was inferred on the basis of a rate calibration involving goats and cattle [49].) The six 'multigene divergence times between animal phyla and among plants, animals and fungi' look extremely precise. For example, on the basis of 15 'constant-rate' genes, the arthropod–nematode divergence event is said to have occurred 1167 ± 83 MYA. This estimate translates into a 95% confidence interval of ~350 million years around the mean. Let us now see what happens if we include the calibration and derivation uncertainties in the calculations. The 95% interval turns out to span a period of ~14.3 billion years (from 14.3 billion to 93 MYA). Of course, to state that the lineages leading to *Drosophila melanogaster* and *Caenorhabditis elegans* diverged from each other subsequent to the formation of the solar system but before the Cretaceous period requires no molecular data.

Chapter 4: tautological comparison of the 310 ± 0 versus 110 ± 0 MYA calibrations

An important detour in the saga of assigning precise timescales to the evolution of everything is found in a paper in which the miracle of transubstantiation attains fulfillment [5]. In this paper, the 310 ± 0 MYA and the 110 ± 0 MYA dates are treated as 'independent' calibrations for purposes of dating avian divergence events. The results based on the 310 ± 0 MYA calibration and the results based on the 110 ± 0 MYA calibration are compared and discussed. Unsurprisingly, because one date was derived from the other, the two values yielded concordant results. Interestingly, other authors have been misled into regarding these two dates as independent calibrations [44].

Chapter 5: errorless molecular estimates substitute for fossil evidence

The fifth chapter in the saga marks the emergence of five tertiary calibration points [6]. That is, five estimates from Wang *et al.* [4] that were derived from the secondary 110 ± 0 MYA calibration, which in turn was derived from the primary 310 ± 0 MYA calibration, are turned into 'errorless' (±0) calibrations, from which further molecular-clock estimates are derived. The tertiary calibration pairs are: plants versus animals, animals versus fungi, plants versus fungi, nematodes versus arthropods and vertebrates, and arthropods versus vertebrates. Interestingly, these calibrations are claimed to be 'derived from an analysis of 75 nuclear proteins calibrated with the vertebrate fossil record' [6]. A careful reading of Wang *et al.* [4], however, reveals that the calibrations are based on 15–39 proteins, not 75.

The elimination of uncertainties in the new calibration points is rationalized by the need 'to reduce extrapolation error' [6]. This statement is exceedingly odd because all molecular-time estimates are based on linear regression analyses, whereby the time of divergence is calculated from graphs in which the X-axis is the molecular divergence between two sequences and the Y-axis is time. The confidence limits of a regression line are known to be shaped as a biconcave belt: the further away we are from the mean, the less reliable our estimates of divergence times will be. In the dating saga before us, the regression line is derived from the calibration point and the intersection of the axes in the graph (i.e. the 'mean' is the primary calibration point). Thus, 'to reduce extrapolation error' essentially violates the mathematical basis of the analysis. The mere fact that the coefficients of variation remain approximately constant or even decrease the further we stray from the primary calibration should have been sufficient to invalidate the entire dating exercise. For example, on the basis of 333 sequences, the divergence time between ferungulates (e.g. cows and horses) and primates is given as 92 ± 1.3 MYA [3] (i.e. the coefficient of variation is 25%). In comparison, on the basis of only two proteins, the divergence time between

Neocallimasticales and all other fungi is said to be 1458 ± 70 MYA [6] (i.e. the coefficient of variation is 7%). These results are simply incompatible with mathematics. The fact that the uncertainties were successively discarded in the process of deriving secondary, tertiary and higher-order calibrations from the primary human–chicken calibration does not mean that the errors are not there. Every estimate, regardless of the method used to derive it, has a mean and a variance [41,42], and the variance neither diminishes nor disappears when one manipulates the mean.

The tertiary derivatives of the chicken–human comparison are used to estimate ten ancient divergence events, such as mosses versus vascular plants, basidiomycetes versus ascomycetes and *Candida albicans* versus *Saccharomyces cerevisiae* [6]. As expected, the iterative use of ± 0 calibration points results in estimates that are accompanied by deceptively low standard errors, leaving the uninitiated reader with a sense of certainty that is as comforting as it is false. Thus, estimates such as 1458 ± 70 MYA for the divergence between Neocallimasticales and all other fungi or 1107 ± 56 MYA for the divergence between Mucorales and Blastocladiales versus Basidiomycota and Ascomycota [6] are imaginary. Indeed, if ancient divergence events can be pinpointed so accurately in a temporal framework, then by using the methodology of Heckman *et al*. [6], we will be able to attain parts-per-billion accuracies in dating less-ancient events. In fact, we might ultimately be able to tell whether the human–chimpanzee divergence occurred on a Monday or not.

Chapter 6: dating *Genesis*

The continuation of the saga is as predictable as it is outlandish. By using tertiary, and possibly quaternary, quinary and senary derivations from the mythical 310 ± 0 chicken–human calibration, five of the most ancient divergence events are dated [2]. The pinnacle is reached with an estimate of 3.97 ± 0.25 billion years ago for the divergence between archaebacteria and eukaryotes. An illustrative example of the extrapolations involved in estimating ancient divergence events is shown in Figure 2.

All these dating exercises have been summarized as reviews [7,8,13] with attractive figures depicting the age of all vertebrates and model organisms. The appearance of accuracy and the high-quality artwork have resulted in hundreds of citations in which such dates were accepted as factual. Unfortunately, no matter how great our thirst for glimpses of the past might be, mirages contain no water. Trying to estimate the divergence times of fungal, algal or prokaryotic groups on the basis of a partial reptilian fossil

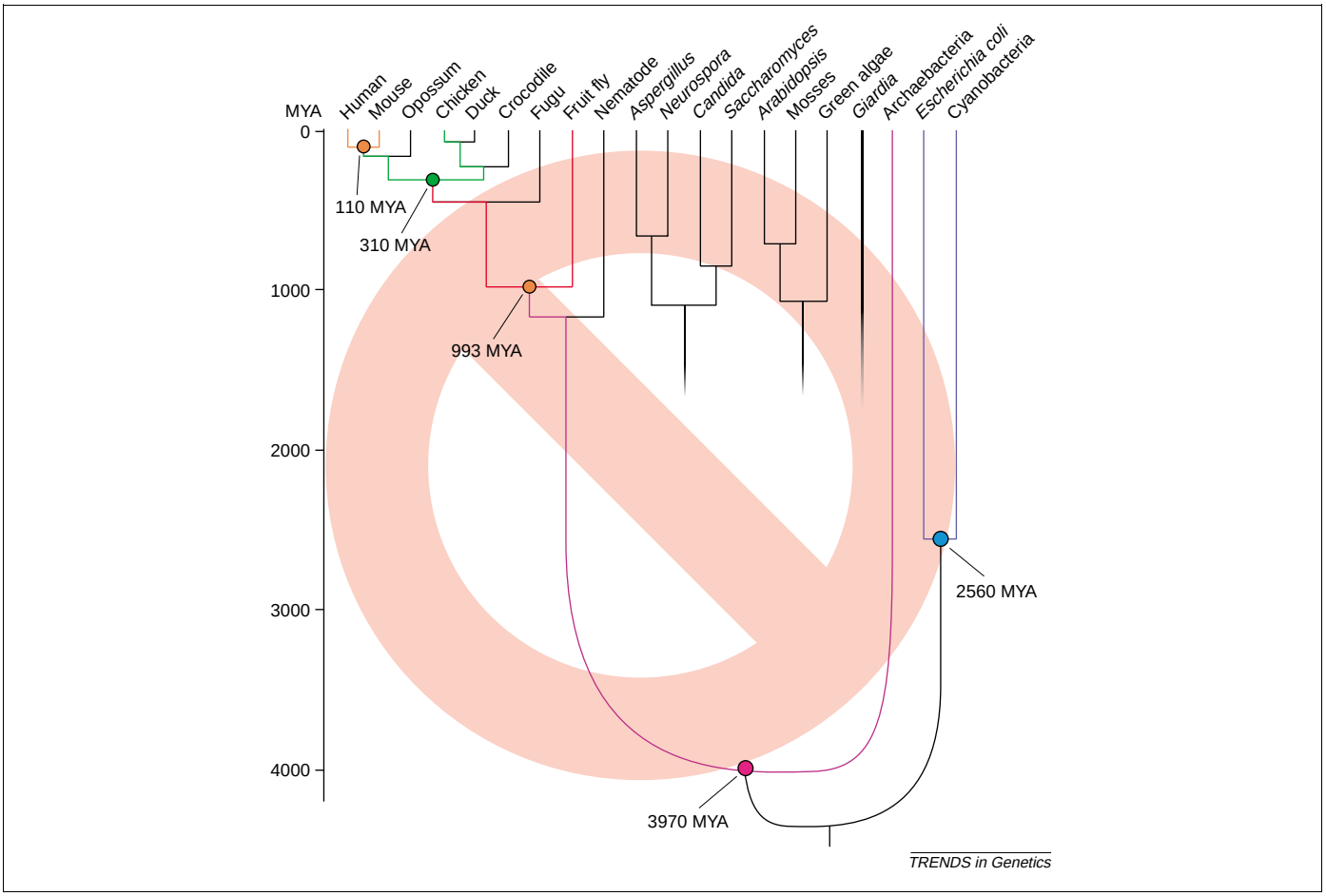

Figure 2. A leap of faith? Estimating the divergence time between *Escherichia coli* and cyanobacteria (blue) in Ref. [2] was accomplished by the following steps: (i) start at the primary human–chicken calibration (green) of 310 ± 0 MYA; (ii) interpolate from 310 ± 0 MYA and modify to 110 ± 0 MYA (orange); (iii) extrapolate from 110 ± 0 MYA and modify to 993 ± 0 MYA (red); (iv) extrapolate from 993 ± 0 MYA and modify to 3970 ± 0 MYA (pink); and (v) bend the corner from 3970 ± 0 MYA and extrapolate to 2560 MYA. The estimation procedure required extrapolations and interpolations over a phylogenetic path equivalent in length to at least 5500 MY of evolution. The total route exceeds the age of the Earth. The faded lines indicate uncertain phylogenetic affiliations.

and protein sequences from mice and humans is like trying to decipher Demotic Egyptian with the help of an odometer and the *Oxford English Dictionary*. This is not to say that molecular estimates of divergence times are flawed *per se*; on the contrary, they are useful when based on solid statistical methodology and multiple fossil calibrations.

Postscript: the 110 ± 0 MYA calibration dies but its descendants survive

In what will surely not be the last chapter in this story, a recent review in *Trends in Genetics* [13] contains four blood-curdling innovations involving statistical methodology, taxonomy, physics of time reversal and logic. The statistical novelty concerns the primate–artiodactyl divergence time, which was 90 ± 8 MYA in Hedges *at al*. [1], whereas in Ref. [13] it is quoted as '90–98 million years ago.' This change turns the mean into the lower end of the range and reduces by 76% the already-too-narrow confidence interval from 34 million years to 8 million years. The taxonomic novelty concerns the 3.97 ± 0.25 billion-years-ago estimate for the divergence between archaebacteria and eukaryotes [2], which in Ref. [13] is not only stripped of error but also assigned to a different evolutionary event (i.e. the 'early divergence among prokaryotes'). The third novelty concerns the claim that two ascomycete fungi (*Saccharomyces* and *Schizosaccharomyces*) diverged from one another (1144 MYA) before their common ancestor diverged from the basidiomycetes (1107 MYA). This claim requires time to run backwards for ~37 million years.

Notwithstanding the concerns discussed, it is the logical innovation in Ref. [13] that is the most extraordinary. In the legend of a figure, Hedges *et al*. state that all 'divergence times involving rodents' were essentially discarded 'because of large differences in published molecular time estimates' [13]. Normally, a house of cards topples if the card at the foundation of the structure is pulled out. Miraculously, however, 36 divergence-time estimates based on the 110 ± 0 MYA calibration are deemed worthy of publication [13], whereas the number on which the calculations were based is not.

Conclusion and recommendation

Despite their allure, we must sadly conclude that all divergence estimates discussed here [1–13] are without merit. Our advice to the reader is: whenever you see a time estimate in the evolutionary literature, demand uncertainty!

Acknowledgements

We thank John H. Calder, Daniel A. Chamovitz, Or M. Graur, S. Blair Hedges, Michael S.-Y. Lee, Michael Ovadia, Robert R. Reisz and Shaul Shaul for information.

References

1 Hedges, S.B. *et al.* (1996) Continental breakup and the ordinal divergence of birds and mammals. *Nature* 381, 226–229
2 Hedges, S.B. *et al.* (2001) A genomic timescale for the origin of eukaryotes. *BMC Evol. Biol.* 1441–2148/1/4 (www.biomedcentral.com)
3 Kumar, S. and Hedges, S.B. (1998) A molecular timescale for vertebrate evolution. *Nature* 392, 917–920
4 Wang, D.Y- *et al.* (1999) Divergence time estimates for the early history of animal phyla and the origin of plants, animals and fungi. *Proc. R. Soc. Lond. B. Biol. Sci.* 266B, 163–171
5 van Tuinen, M. and Hedges, S.B. (2001) Calibration of avian molecular clocks. *Mol. Biol. Evol.* 18, 206–213
6 Heckman, D.S. *et al.* (2001) Molecular evidence for the early colonization of land by fungi and plants. *Science* 293, 1129–1133
7 Hedges, S.B. (2002) The origin and evolution of model organisms. *Nat. Rev. Genet.* 3, 838–849
8 Hedges, S.B. and Kumar, S. (2002) Vertebrate genomes compared. *Science* 297, 1283–1285
9 Hedges, S.B. and Shah, P. (2003) Comparison of mode estimation methods and application in molecular clock analysis. *BMC Bioinformatics* 1471–2105/4/31 (www.biomedcentral.com)
10 Stauffer, R.L. *et al.* (2001) Human and ape molecular clocks and constraints on paleontological hypotheses. *J. Hered.* 92, 469–474
11 Tracy, M.R. and Hedges, S.B. (2000) Evolutionary history of the enolase gene family. *Gene* 259, 129–138
12 Hedges, S.B. and Poling, L.L. (1999) A molecular phylogeny of reptiles. *Science* 283, 998–1001
13 Hedges, S.B. and Kumar, S. (2003) Genomic clocks and evolutionary timescales. *Trends Genet.* 19, 200–206
14 Alroy, J. (1999) The fossil record of North American mammals: evidence for a Paleocene evolutionary radiation. *Syst. Biol.* 48, 107–118
15 Benton, M.J. (1999) Early origins of modern birds and mammals: molecules vs. morphology. *BioEssays* 21, 1043–1051
16 Lee, M.S.Y. (1999) Molecular clock calibrations and metazoan divergence dates. *J. Mol. Evol.* 49, 385–391
17 Bromham, L. *et al.* (2000) The power of relative rates tests depends on the data. *J. Mol. Evol.* 50, 296–301
18 Sanderson, M.J. (2002) Estimating absolute rates of molecular evolution and divergence times: a penalized likelihood approach. *Mol. Biol. Evol.* 19, 101–109
19 Sanderson, M.J. (2003) r8s: inferring absolute rates of molecular evolution and divergence times in the absence of a molecular clock. *Bioinformatics* 19, 301–302
20 Martin, A.P. and Burg, T.M. (2002) Perils of paralogy: Using HSP70 genes for inferring organismal phylogenies. *Syst. Biol.* 51, 570–587
21 Shaul, S. and Graur, D. (2002) Playing chicken (*Gallus gallus*): Methodological inconsistencies of molecular divergence date estimates due to secondary calibration points. *Gene* 300, 59–61
22 Benton, M.J. and Ayala, F.J. (2003) Dating the tree of life. *Science* 300, 1698–1700
23 Bromham, L. and Penny, D. (2003) The modern molecular clock. *Nat. Rev. Genet.* 4, 216–224
24 Rieppel, O. (1999) Turtle origins. *Science* 283, 945–946
25 Benton, M.J. ed. (1993) *The Fossil Record 2*, Chapman and Hall
26 Harland, W.B. *et al.* (1989) *A Geologic Time Scale*, Cambridge University Press
27 Carroll, R.L. (1964) The earliest reptiles. *J. Linnean Soc. Zool.* 45, 61–83
28 Carroll, R.L. (1969) A middle Pennsylvanian captorhinomorph, and the interrelationships of primitive reptiles. *J. Paleontol.* 43, 151–170
29 Peabody, F.E. (1952) *Petrolacosaurus kansensis* Lane, a Pennsylvanian reptile from Kansas. *Paleontol. Contr. Univ. Kansas* 1, 1–41
30 Reisz, R.R. (1986) *Pelycosauria Encyclopedia of Paleoherpetology Pt 17A*, Verlag Dr Friedrich Pfeil
31 Reisz, R.R. (1997) The origin and early evolutionary history of amniotes. *Trends Ecol. Evol.* 12, 218–222
32 Ayala, F.J. *et al.* (1998) Origin of the metazoan phyla: molecular clocks confirm paleontological estimates. *Proc. Natl. Acad. Sci. U. S. A.* 95, 606–611
33 Calder, J.H. (1998) The Carboniferous evolution of Nova Scotia. In *Lyell: The Past is the Key to the Present* (Blundell, D.J. and Scott, A.C., eds), pp. 261–302, Geological Society, London
34 Smithson, T.R. (1989) The earliest known reptile. *Nature* 342, 676–678
35 Sanderson, M.J. (1997) A nonparametric approach to estimating divergence times in the absence of rate constancy. *Mol. Biol. Evol.* 14, 1218–1231
36 Árnason, U. *et al.* (1998) Molecular timing of primate divergences as estimated by two nonprimate calibration points. *J. Mol. Evol.* 47, 718–727
37 Rambaut, A. and Bromham, L. (1998) Estimating divergence dates from molecular sequences. *Mol. Biol. Evol.* 15, 442–448

38 Springer, M.S. *et al.* (2003) Placental mammal diversification and the Cretaceous-Tertiary boundary. *Proc. Natl. Acad. Sci. U. S. A.* 100, 1056–1061
39 Huff, D. (1954) *How to Lie with Statistics*, Norton, New York
40 Topping, J. (1972) *Errors of Observation and their Treatment*, Chapman and Hall
41 Nei, M. (1987) *Molecular Evolutionary Genetics*, Columbia University Press
42 Nei, M. and Kumar, S. (2000) *Molecular Evolution and Phylogenetics*, Oxford University Press
43 Ota, T. and Nei, M. (1994) Estimation of the number of amino acid substitutions per site when the substitution rate varies among sites. *J. Mol. Evol.* 38, 642–643
44 Gu, X. (1998) Early metazoan divergence was about 830 million years ago. *J. Mol. Evol.* 47, 369–371
45 Easteal, S. (1999) Molecular evidence for the early divergence of placental animals. *BioEssays* 21, 1052–1058
46 Barnett, V. and Lewis, T. (1994) *Outliers in Statistical Data*, 3rd edn, Wiley
47 Benton, M.J. (1997) *Vertebrate Paleontology*, 2nd edn, Chapman and Hall
48 Archibald, J.D. (1996) Fossil evidence for a Late Cretaceous origin of 'hoofed' mammals. *Science* 272, 1150–1153
49 Rassmann, K. (1997) Evolutionary age of the Galápagos iguanas predates the age of the present Galápagos Islands. *Mol. Phylog. Evol.* 7, 158–172
50 Cao, Y. *et al.* (2000) Phylogenetic position of turtles among amniotes: evidence from mitochondrial and nuclear genes. *Gene* 259, 139–148
51 Zardoya, R. and Meyer, A. (2001) The evolutionary position of turtles revised. *Naturwissenschaften* 88, 193–200