J Mol Evol (2004) 59:267–276
DOI: 10.1007/s00239-004-2624-9

JOURNAL OF MOLECULAR EVOLUTION



# Error in Estimation of Rate and Time Inferred from the Early Amniote Fossil Record and Avian Molecular Clocks

**Marcel van Tuinen, Elizabeth A. Hadly**

Department of Biological Sciences, Stanford University, Stanford, CA 94305, USA

Received: 2 November 2003 / Accepted: 8 March 2004 [*Reviewing Editor*: Martin Kreitman]

**Abstract.** The best reconstructions of the history of life will use both molecular time estimates and fossil data. Errors in molecular rate estimation typically are unaccounted for and no attempts have been made to quantify this uncertainty comprehensively. Here, focus is primarily on fossil calibration error because this error is least well understood and nearly universally disregarded. Our quantification of errors in the synapsid–diapsid calibration illustrates that although some error can derive from geological dating of sedimentary rocks, the absence of good stem fossils makes phylogenetic error the most critical. We therefore propose the use of calibration ages that are based on the first undisputed synapsid and diapsid. This approach yields minimum age estimates and standard errors of 306.1 ± 8.5 MYR for the divergence leading to birds and mammals. Because this upper bound overlaps with the recent use of 310 MYR, we do not support the notion that several metazoan divergence times are significantly overestimated because of serious miscalibration (*sensu* Lee 1999). However, the propagation of relevant errors reduces the statistical significance of the pre-K–T boundary diversification of many bird lineages despite retaining similar point time estimates. Our results demand renewed investigation into suitable loci and fossil calibrations for constructing evolutionary timescales.

*Correspondence to:* Marcel van Tuinen; *email:* mvtuinen@stanford.edu

**Key words:** Calibration — Molecular clock — Fossil — Birds — Minimum age — Synapsid–diapsid

## Introduction

After 40 years, the molecular clock (Zuckerkandl and Pauling 1965) remains one of the most provocative ideas applicable to evolution, foremost because it extends dating of events in the history of life to organisms without a good fossil record. The molecular clock is contentious because it frequently conflicts with the paleontological timescale in demonstrating older evolutionary lineage times (Easteal et al. 1995; Benton 1999). Furthermore, universal endorsement of the clock is lacking because of its uncertain taxonomic scope and molecular basis and oversimplification of the biases involved (Hillis et al. 1997). Several neutral-theory models have been developed in support of a molecular clock in nonregulatory genome compartments, but the prevalence of the clock (Easteal et al. 1995; Nei and Kumar 2000) and the importance of nearly neutral evolution (Ohta and Gillespie 1996) remain controversial. However, the molecular clock remains the only method for ascertaining evolutionary divergence times where the fossil record is poor (Wilson et al. 1977) and, therefore, will likely prove essential in construction of the timescale of life. For this reason, continued research on the mechanisms of molecular and morphological evolution and on discrepancies between fossil and molecular timescales is desired. Toward this goal, both

**Table 1.** Type of errors in (mostly mtDNA based) molecular clock studies with representative estimates

| Molecular | % error est. | Fossil/calibration | % error est. |
|---|---|---|---|
| Substitution rate variation | 0–100[a] | Fit on tree (max–min/min) | 0–8.6 |
| *Mutation rate variation | 0.8–11.3[b] | Extinction | Unknown |
| *Body size | 0–100[c] | Internal/external | Variable |
| *Generation time | 0–66[d] | Stem/crown/character | Unknown |
| *Other | 0–3[e] | Fossil character diagnosis | Unknown |
| Sequence length | Variable | Fossil age (error/mean) | 1.3–5.2[f] |
| Taxon sampling | Variable | Dating error | 0.5–1.5[g] |
| Gene sampling | 25–35[h] | Stratigraphic error | 0–1.4[i] |
| Model parameters | Unknown | Interpolation | 0–4.0[j] |
| Branch length estimation | 10.7–19.1[k] | Stratigraphic correlation | Unknown |
| Substitution saturation | 0–250[l] | One/multiple/calibrations | Variable |
| Tree building method | Unknown | Regression error | Variable |

[a]Based on assuming a global 2%/MY rate for tetrapod mtDNA (Wilson et al. 1977) and the literature range of substitution for birds (0.62–4%/MYR [Brown et al. 1979; Nunn and Stanley 1998]).
[b]Applies to fourfold sites in mammalian nuclear genes (Kumar and Subramanian 2002).
[c]Applies to reptile cytochrome *b* (Bromham 2002, Fig. 1a).
[d]Applies to mammal cytochrome *b* (Bromham et al. 1996, Fig. 2a).
[e]Significant decrease in hummingbird DNA substitution rate with elevation, independent from body mass and generation time (Bleiweiss 1998).
[f]Minimum and maximum errors are based, respectively, on the minimum age and error of *Hylonomus* and of *Westlothiana* (see Fig. 3 and text for discussion).
[g]Minimum and maximum errors are based, respectively, on the $U^{238}/Pb^{206}$ dating error of the Upper Westphalian B and the $^{40}Ar/^{39}Ar$ dating error of the Upper Stephanian A.
[h]Hedges and Kumar (2003); error may be higher if paralogous gene comparisons are included.
[i]Based on placing *Hylonomus* either in the Westphalian A or Westphalian B (see Fig. 3).
[j]Interpolation between directly dated Namurian A–B and Asbian–Holkerian border (see Fig. 3).
[k]van Tuinen and Hedges (2001); minimum and maximum error from average branch lengths.
[l]Average of 2% MYR is based on uncorrected distance (Brown et al. 1979). A gamma-corrected rate of 5%/MYR in mammals and birds (Arbogast and Slowinski 1998) in return has erroneously been applied to uncorrected distances.

paleontological and molecular timescales are under continuous review (Benton and Ayala 2003; Hedges and Kumar 2003). Fossil discoveries important to the early history of several modern lineages continue to be made, and several important fossils now are placed in a cladistic framework (e.g., Thewissen et al. 2001). The study of molecular clocks has seen major computational and theoretical progress (Nei and Kumar 2000; Arbogast et al. 2002). These advances hold great promise for increasing reliability in divergence time estimation across the entire evolutionary timescale.

In contrast to these developments, errors in genetic data and fossil records have not been addressed adequately. This is surprising given that the errors involved in age estimation are critical to judging the value of a timescale. Here we introduce and discuss the errors associated with a particular molecular-based timescale: that concerning the origin and radiation of modern birds. By using such a prominent group, we hope to raise awareness of the importance of error in molecular clock studies. Our focus is primarily on calibration error because this error is least well understood and nearly universally disregarded. We review the use of the bird–mammal calibration in the molecular literature, but also the primary geologic literature, in order to ascertain the details of the earliest amniote fossils and their corresponding stratigraphic horizons. We then compound fossil and molecular errors and jointly apply this uncertainty to a mitochondrial rRNA data set to assess the impact of molecular clock error on testing of modern bird diversification in the context of the K–T boundary (e.g., van Tuinen and Hedges 2001; Paton et al. 2002). Although our conclusions are limited to a single mitochondrial data set, it illustrates the presentation of confidence intervals in molecular clock-based time estimation based on both fossil and molecular uncertainties.

## Materials and Methods

### *Quantification of Error in Molecular Data*

Molecular errors in the context of the molecular clock derive from various sources, including deviation from rate constancy, and averaging across single gene estimates (Table 1) (Nei and Kumar 2000). Some molecular errors can be estimated by bootstrapping or jackknifing (Nei et al. 2001), while others can be minimized through proper model choice. Departure from substitution rate constancy is a major source of error, especially when a consistent bias exists. Such bias is thought to accrue through clade-specific variation in mutation rate and life history traits such as generation time and body size (Table 1), although the true extent of these and other factors in causing molecular rate differences remains to be

determined. Methods have been developed to test for rate constancy or to utilize multirate data sets (Sanderson 1997; Thorne et al. 1998), but their accuracy is untested. Despite the ability to quantify molecular error to some extent, such error is frequently not measured, and the majority of genetic studies report only point time estimates. Moreover, other sources of molecular error including the effect of character and taxon sampling and choice of branch length (lineage distance) and tree-building models remain poorly studied (Buckley et al. 2001) (Table 1). Here, we do not attempt to resolve each source of molecular error but, instead, illustrate the readily quantifiable error sources in our example of the complete mitochondrial noncoding rRNA genes for birds. We chose this data set because it has been calibrated with the synapsid–diapsid divergence (310 MYA [Kumar and Hedges 1998]) and includes representatives from every modern bird order (van Tuinen and Hedges 2001). Rate heterogeneity is decreased by iteratively testing for rate constancy and then discarding the anomalous taxa by using the branch-length and two-cluster tests (van Tuinen and Hedges 2001). The final branch-length test that yields no additional rate heterogeneous taxa results in a linearized tree and specifies standard error for each branch. We used these error estimates to measure error for each node including that for our selected internal galliform–anseriform (=gamefowl–waterfowl) anchorpoint. The calibration age carries additional molecular (stochasticity in gene sampling) error because it is a multigene average derived from slower-evolving nuclear genes. These genes allow sequence alignment of bird taxa with their mammalian homologues, which is difficult in the mitochondrial ribosomal genes and would result in many indels and loss of phylogenetic resolution within birds. It has been argued elsewhere (van Tuinen and Hedges 2004) that this two-step calibration approach to provide bird divergence times is error-rich but nonetheless preferred over the use of direct fossil-based internal calibrations with unknown error. Instead the galliform–anseriform anchorpoint comes with previously quantified error (van Tuinen and Hedges 2001). However, the published age and error estimate assumed, first, that the initial bird–mammal fossil calibration contains minimal error and, second, that the age of 310 MYA is valid for this divergence. These assumptions are investigated further in this paper.

### *Quantification of Error in Fossil Data*

There are multiple sources of error in dating fossils (Table 1). Because of the increasing reliance on molecular clock studies for many evolutionary questions (e.g., Hedges 2002), we argue here that fossil age errors can and should be documented and actively discussed. Errors involved in assignment of fossil age include error in stratigraphic placement, correlation between fossil-bearing localities and radiometrically dated strata, interpolation between stratigraphic boundaries, and radiometric dating methods. The purpose of this paper is to quantify these errors specifically for the fossils relevant to the stem and crown of the synapsid–diapsid divergence (Table 1). One of the most critical errors in calibration includes placement of the fossil in the proper location on the tree. This phylogenetic uncertainty cannot be quantified readily, thus we use minimum ages obtained from the oldest fossils bearing diagnostic characters. Likely maximum ages are also identified for the synapsid–diapsid divergence and the effect of shifting calibration age from minima to maxima is discussed.

### *Quantifying 95% Confidence Intervals*

Errors are compounded as described previously (van Tuinen and Hedges 2001) and converted to 95% confidence intervals (CI) around each point time estimate across the bird tree (±1.96 SE). The compounded molecular error contains an uncertainty for each node that is fixed because it is derived from the gene sampling (stochasticity) error and the molecular lineage distance estimation error (=branch length error) of the galliform-anseriform calibration (12.4%). It also contains an uncertainty derived from branch length estimation for each node and, hence, is the only variable component. We label this latter error the "time estimation error." The remaining error is in calibration of molecular rate and is labeled here as the "rate estimation error." It is based on the uncertainty in calibration fossil age, and in estimation of the length of the galliform–anseriform branch, and the gene sampling error (12.7%).

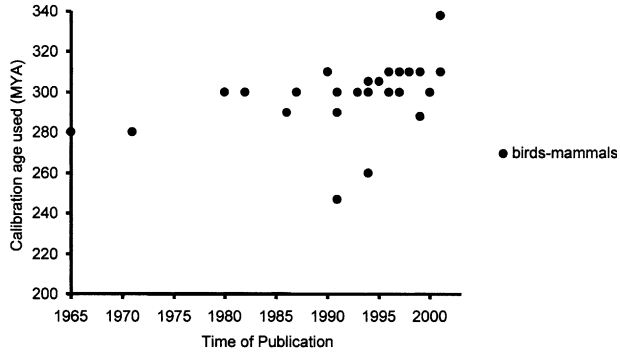


**Fig. 1.** The variety of reported synapsid–diapsid divergence times by time of published studies that use this node for molecular rate estimation.

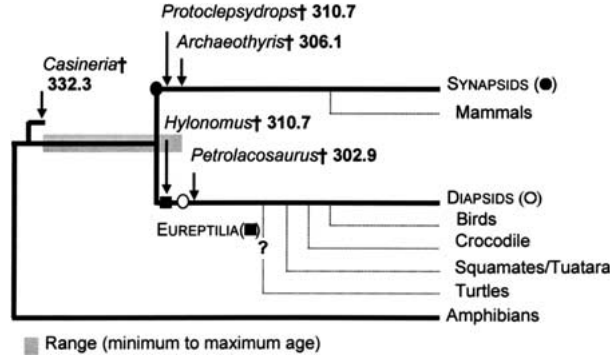


**Fig. 2.** Phylogeny of Synapsida–Diapsida with divergence time constrained by minimum age and maximum age. Minimum age is constrained by the oldest occurrence (*Archaeothyris*) on the line leading to synapsids (306.1 ± 8.5 MYR [Hess and Lippolt 1986; Benton 1990]). The maximum age is 332.3 ± 13.4 MYR based on the stem amniotes *Casineria* and *Westlothiana* (Smithson 1989; Benton 1997; Paton et al. 1999). The oldest fossil relevant to Diapsida is *Petrolacosaurus* from the Lower Carboniferous (302.9 ± 4.6 MYR) of Kansas (Reisz 1972). Thus, the preliminary range based on maximum–minimum ages for the synapsid–diapsid divergence is 332.3–306.1 MYA. The upper bound (306.1 MYA) is probably the more likely estimate. See text for details.

## Results

### *Use of the Bird–Mammal Calibration*

The divergence of the lineages leading to birds and to mammals is defined by the synapsid–diapsid split (or synapsid–sauropsid proper) and is one of the most frequently used calibrations for establishing the rate of the molecular clock (e.g., Kimura 1980; Doolittle

et al. 1996; Hedges et al. 1996; Kumar and Hedges 1998). It has been employed as an external calibration in studies of several vertebrate groups and as internal anchorpoint to estimate the time of origin of tetrapods, vertebrates, metazoans, fungi, plants, and life itself (Hedges 2002 and references therein). The bird–mammal calibration is unique in being one of the first examples of applying the fossil record to determine a molecular evolutionary rate for a gene (Margoliash and Smith 1965). First proposed in paleontologic detail by Benton (1990) as 305 MY old, a large range (247–338 MYA) of ages has been found in the molecular clock literature from 1965 to 2002 (Fig. 1) (Shapiro 1991; Bossuyt and Milinkovitch 2001). Interestingly, the use of a Carboniferous age in more recent molecular clock studies (e.g., 310 MYA [Kumar and Hedges 1998]) is not due to discovery of older fossils. Rather, it is a reflection of more detailed interpretation of the original fossils and refinement of the Upper Carboniferous timescale using the $^{40}Ar/^{39}Ar$ radiometric dating method (Menning et al. 2000).

*Definition of Calibration*

The clade defined by the synapsid–diapsid divergence is equivalent to the amniote crown group. The most primitive amniotes are believed to have evolved from anapsid ancestors (Fig. 2) (Carroll 1988; Benton 1997). Mammals are the only living members of a once much larger synapsid group, amniotes that bear one opening behind each orbit. Birds, together with crocodiles, squamates, turtles, and tuatara, evolved separately from mammals on a line characterized by two openings behind each orbit and are referred to as diapsid amniotes. The history and evolutionary origins of diapsids are less well understood because this condition has invariably been lost in some lineages, resulting in a homoplastic anapsid condition (e.g., turtles). Although the origin of amniotes is tied to the first appearance of an extraembryonic egg membrane (Carroll 1988: but see Reisz 1997), membrane tissue is rarely preserved in the fossil record and is thus inferred from crown-group definition. Instead, for the fossil record other amniote-defining morphological characters are crucial for phylogenetic analysis, in particular, the temporal openings in the skull located behind the orbits (Carroll 1988).

*Justification for Calibration Choice*

One of the reasons for the popularity of the synapsid–diapsid evolutionary divergence is that the Carboniferous marks the halfway point in metazoan evolution. Thus this calibration may be equally useful for internal and external calibration of the vertebrate timescale. Another reason is that a large amount of molecular data is available for many genes of both birds and mammals. This calibration choice is widely considered to have support from the relevant fossil record (Margoliash and Smith 1965; Janke and Arnason 1997; Kumar and Hedges 1998). The fossil record that bears on early amniote evolution is remarkable for its unusual preservation of the earliest terrestrial ancestors from the vacuities of tree stumps in otherwise unlikely preservational environments. Of additional importance is that several more derived and easily diagnosable amniote fossils from both synapsid and diapsid clades are known shortly postdating these earliest fossils, thus making it unlikely that the synapsid–diapsid divergence time is greatly overestimated. Minor differences in the skull, combined with otherwise striking morphological similarities between these earliest members on both branches, also suggest that the actual divergence was not significantly earlier. For these reasons, molecular researchers have depended heavily on this particular calibration. Most workers have assumed that the bird–mammal calibration at 310 MYA contains negligible error, partly because there have not been fossil studies that give mean estimate fossil ages with standard errors.

*Age of Synapsid–Diapsid Divergence*

The age of the stem amniote (anapsid) lineage that most closely approaches the time of first appearance of postorbital openings can be used as the lower bound or **maximum** age estimate of the synapsid–diapsid divergence. Proximity of the lower to the upper bound increases our confidence of utilizing a well-constrained fossil lineage. This strategy is especially useful for serial records of fossils with stem and crown taxa and with morphologic characters distinguishing stem from crown (Marshall 1990). The **minimum** age of a calibration must be based on a fossil that displays characters diagnostic to the clade that includes the calibration. In the case of the synapsid–diapsid divergence, it should be based on earliest possession of a diagnosable diapsid or synapsid skull.

Amniotes originated from semiaquatic terrestrial tetrapods, which are common in the Early Carboniferous but nearly absent in the transition to the Late Carboniferous (Carroll 2001). This gap occurs at the transition from large to small tetrapods and may be due to the natural bias of the fossil record against preservation of fully terrestrial and small-sized organisms. At face value, the fossil record suggests that amniotes diversified rapidly from uncertain immediate ancestors. No consensus has yet emerged on the identity of the stem amniote that is closest to the

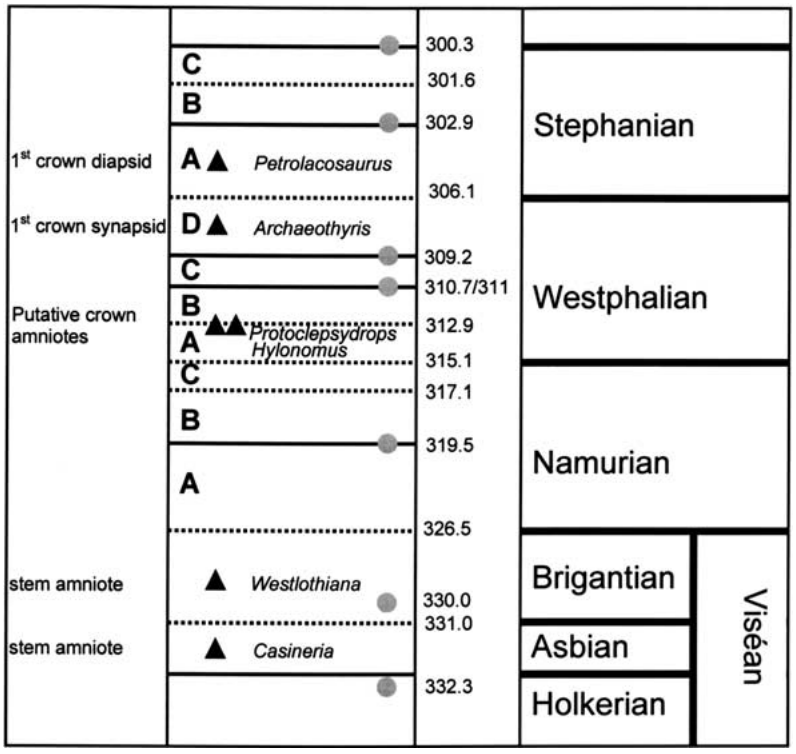


**Fig. 3.** Geological distribution of fossils (black triangles) that are relevant to the divergence of Synapsida and Diapsida. Directly dated geological boundaries are shown with solid lines, and gray circles. Ages for all other geological boundaries involve interpolation given sediment rate estimation based on these directly dated boundaries. Seven geological calibrations are available for the upper and middle Carboniferous timescale (Hess and Lippolt 1986; Menning et al. 2000).

amniote crown group (Reisz 1997; Carroll 2001), but two recently discovered fossils exhibit striking amniote-like trunk vertebrae characters in combination with a plesiomorphic state in other characters (Carroll 2001). *Westlothiana lizziae* (Smithson 1989) was originally named a reptile (diapsid) based on similarities of the skull, humerus, and vertebrae but comparison to both diapsids and synapsids precluded inclusion in the amniote crown because of the absence of diagnosable characters (Smithson 1989). Despite limited preservation (Reisz 1997), recent investigation has indicated that it belongs on the amniote stem close to perhaps diadectomorphs (Rolfe et al. 1993: Carroll 1995) or lepospondyls (Reisz, personal communication). *Westlothiana* is known from two nearly complete skeletons in the lower East Kirkton Limestone, West Lothian Upper Oil–Shale Formation, Scotland (Rolfe et al. 1993). This formation was a tropical freshwater lake in a volcanic setting and contains a wide variety of fossil plants, fish, arthropods, and amphibians (Rolfe et al. 1993). Most fossil vertebrates represented in this site are terrestrial rather than aquatic, indicating that the preservational environment in the Lower Carboniferous for East Kirkton was unusual. The geological and paleontological succession of the Upper Oil–Shale Formation has been studied in great detail and points to a Brigantian (Visèan) age based on correlation with other similar sites in Scotland, plant biostratigraphy, palynological data, and the presence of the mussel *Posidonia becheri* (Rolfe et al. 1993). The exact position within this Lower Carboniferous stage and the error due to lateral variation in thickness of the fossil-bearing unit are uncertain (Rolfe et al. 1993). The time and span represented by the Brigantian are somewhat unclear because of limited availability of radiometric calibrations, and those calibrations yield inconsistent ages of the Lower Carboniferous (Menning et al. 2000). The relevant radiometric calibrations, based on $U^{238}$–$U^{206}$ and $Pb^{207}$–$Pb^{206}$ zircon dating, indicate that the lower Brigantian is either 319 ($\pm 1$) or 330 ($\pm 2$) MYR old (Menning et al. 2000). The upper Brigantian has not been dated directly but interpolation points to an age of 319 or 326.5 MYA (Menning et al. 2000). Error due to lack of dating and interpolation can be avoided by basing the maximum age of the amniote stem on the maximum age of the lowermost Brigantian, which has been directly dated. This conservative approach yields a maximum age of 330 ($\pm 2$) MYR for *Westlothiana*. Taking into account the uncertainty and range of this age, the quantifiable error can be compounded to 17.1 MYR using the propagated standard error $= \sqrt{(SE_1^2 + SE_2^2 + \ldots SE_n^2)}$.

Parsimony analysis on *Casineria kiddi* (Paton et al. 1999), the first tetrapod with five fingers, suggested a polytomy with *Westlothiana* and Diapsida, but synapsids were not included in those analyses, and the consensus tree was highly instable. Nevertheless, fossil affinities of *Casineria* were deemed as true stem amniote and "reptiliomorph" (Paton et al. 1999). *Casineria* also has been found in Scotland but derives from the slightly older Cheese Bay Shrimp Bed of the

East Lothian Lower Oil Shale. This bed points to freshwater environmental preservation and correlates with the Asbian (Viséan) Wardie Shales Formation based on stratigraphically significant fish faunas (Hesselbo and Trewin 1984) and miospores (Paton et al. 1999). The Asbian also has been poorly dated with radiometric methods but three relevant $U^{238}$–$Pb^{206}$ zircon dates exist: two upper Asbian ages of 325.6 (±2.2) and 326.7 (±2.6) MYR and an uppermost Holkerian age of 332.3 (±1.1) MYR. The maximum age approach would yield an age for *Casineria* and the stem amniote at 332.3 (±13.4) MYR. Based on these analyses the lower bound of the synapsid–diapsid divergence, albeit liberal, can perhaps be set using *Westlothiana* and *Casineria* (Fig. 3). These fossils may not lie on the immediate amniote stem, but until additional discoveries are made, they represent older lineages that temporally and phylogenetically most closely approach the crown node. More detailed analyses and additional fossils will help to further reduce the temporal gap between the lower and the upper bound of the synapsid–diapsid divergence.

Moving toward the crown synapsid–diapsid node, *Hylonomus lyelli*, from the Joggins Formation, Cumberland Group, Nova Scotia (Dawson 1891; Carroll 1964), is widely considered as the oldest fossil relevant to the diapsid line. It has been placed, along with *Paleothyris*, in the family Protorothyrididae (Carroll 1988), a diverse but probably paraphyletic anapsid group from the Westphalian to the Lower Permian (Clark and Carroll 1973). The cranial material of *Hylonomus* also shows an anapsid skull, but its slender limbs and long hand and feet suggest placement within the crown amniote group, as sister to diapsids (Carroll 1991; Benton 1990, 1997). This phylogenetic viewpoint implies that the direct anapsid ancestor of the diapsids, defined as Eureptilia *sensu* Laurin and Reisz 1995, did not give rise to the synapsid lineage (Fig. 2). The problem of the exact position of *Hylonomus* has been avoided in recent literature (Laurin and Reisz 1997; Reisz 1997), possibly because its fragmentary state restricts powerful cladistic analysis. The available material does point unambiguously to an anapsid skull. *Protoclepsydrops haplous*, the oldest fossil pertinent to the synapsid branch has been found in the same formation as *Hylonomus*. Although commonly regarded as the oldest synapsid (e.g., most recently in Carroll [2001]), it is, however, very fragmented and it is unclear if the fossil possesses a synapsid skull (Reisz 1997).

Over the past 40 years, there has been considerable modification of the age of the Joggins formation stratum, a view reflected in the molecular clock literature (Fig. 1). This uncertainty highlights one critical source of error: the fossil material is not dated directly. With the first detailed description of *Hylonomus*, the Joggins stratum was referred to the Westphalian B (Carroll 1964), but most recently it has been referred to the older Westphalian A (e.g., Benton 1993; Calder 1994; Reisz, personal communication). Based on the presence–absence pattern of the ancient plants *Calamites*, *Sigillaria*, and *Lepidodendron*, the stratum of the Joggins Formation was estimated to lie between the Lower Westphalian B and the Upper Westphalian A (Bell 1966). A more recent approach is a compromise, which sets the Joggins Formation at the Westphalian A–B border. The age of the North American Early–Mid Pennsylvanian strata is based on correlation with identical strata in Germany and Russia (Milner 1987) and radiometric dating of volcanic tuffs in these strata. The quantifiable error includes the dating errors of these European strata, but it is difficult to quantify the correlation error in this example (Hess and Lippolt 1986; Carroll, personal communication) (Table 1). The original dating of the Westphalian was performed with $^{40}K/^{39}Ar$ decay on two tuff samples and contained standard errors (SE) up to ±9 MYR. Recent advances with $^{40}Ar/^{39}Ar$ dating of these same and additional tuffs have considerably reduced this error (Hess and Lippolt 1986; Menning et al. 2000). Error in the $^{40}Ar/^{39}Ar$ dating of the Westphalian A–B border includes the measurement error of the argon ratios (±1.3 MYR SE), the irradiation standard plus monitor age (±3.9 MYR SE), and lack of calibration at the Westphalian A–B border (Menning et al. 2000). The quantifiable standard error compounds to 4.1 MYR. The age of the A–B border is about halfway between the estimated upper Westphalian B age of 310.7 (±4.1 SE) MYR and the lower Namurian B age of 319.5 (±4.3 SE) MYR (Hess and Lippolt 1986). However, due to extensive sampling gaps for this crucial Carboniferous time period (Menning et al. 2000), it appears best to estimate conservatively (i.e., err on side of more recent) the Joggins Formation and *Hylonomus* and *Protoclepsydrops* at 310.7 (95% CI = 302.5–318.9) MYA. This estimate overlaps with redating the upper Westphalian B at 311 MYA with one of the Hess and Lippolt (1986) samples based on the $U^{238}$–$Pb^{206}$ zircon SHRIMP (sensitive high-resolution ion micro probe) radiometric technique, which yields greater resolution than the argon methods (±1.7 SE) (Claoué-Long et al. 1995; Menning et al. 2000). The recalculated age of the synapsid–diapsid divergence based on the Joggins fauna at 310.7 MYA ($^{40}Ar/^{39}Ar$) or at 311.0 MYA ($U^{238}$–$Pb^{206}$) is close to the recent divergence time (310 MYA) applied to many molecular clock studies (Fig. 1). The use of 310 MYR age often has been justified by referring to Benton (1990). In that influential paper, the 310 MYR age is based on previous minimum age interpretations of *Hylonomus* (305) and *Archaeothyris* (300) and loosely adding 5 MYR to

account for the incompleteness of the amniote fossil record (Benton 1990, pp. 420–421). Hedges et al. (1996) derived their use of 310 MYR from Benton (1993, pp. 683, 686) as a midpoint estimate from placing *Hylonomus* in the lowermost Moscovian stage (309.2–311.3 MYR [Benton 1993]). However, several assumptions were also made to infer the 310.7–311.0 MYR reestimate: that the Joggins Formation lies at the Westphalian A–B boundary, that a judicious minimum estimate and error of this boundary must be based on direct radiometric dating which cannot be performed on fossils themselves, that dating of nearby volcanic tuff approaches the age of the rock, that error does not exist in correlating the Westphalian A and B strata between North America and Europe, and that the biostratigraphic assignment is appropriate. It also assumes that both *Hylonomus* and *Protoclepsydrops*, neither of which can be easily diagnosed as synapsid or diapsid, are crown (or stem) amniotes that closely approach the synapsid–diapsid divergence. The potential phylogenetic assignment bias that accompanies this approach cannot readily be quantified, but can be avoided by conservatively applying a minimum age to the calibration divergence, based on the first undisputed fossil synapsid and diapsid. It is the latter approach that we favor.

If *Hylonomus* is excluded because of lack of the diagnostic diapsid skull (Lee 1999), then *Petrolacosaurus kansensis*, the first true diapsid reptile from the Late Carboniferous (Reisz 1981) of Kansas, becomes the next oldest fossil relevant to the stem diapsid age. Laurin and Reisz (1997) show that several diapsid synapomorphies, including the diapsid skull, reduced ectopterygoid, and complex tibioastragalar joint are present in *Petrolacosaurus* (Laurin and Reisz 1997; Sumida 1997). The diapsid nature of *Petrolacosaurus* is unambiguous and provides an undisputed, albeit conservative synapsid–diapsid age estimate (302.9 ± 4.6 MYR). The slightly older and first undisputed synapsid, *Archaeothyris florensis*, has been uncovered in the Westphalian Florence Formation, Nova Scotia (Reisz 1972; Benton 1990). Despite limited skeletal preservation, *Archaeothyris* shows a true synapsid skull along with unique pelycosaurian (synapsid) characters, e.g., the shape of the pelvic girdle and humerus (Reisz 1972). The placement of *Archaeothyris* within Synapsida is thus unambiguous and also provides an undisputed, albeit conservative, stem synapsid age estimate and thus the upper bound or minimum time estimate (306.1 ± 8.5 MYR) for the synapsid–diapsid divergence.

The earliest diapsid, *Petrolacosaurus*, comes from the Rock Lake Shale of the Stanton Formation, Lansing group in Garnett, Kansas (Reisz 1981; Reisz et al. 1982). Lithologic, petrographic, and biostratigraphic observations on the Rock Lake Shale support an ancient deltaic environment during the upper part of the Missourian stage (Heckel 1975; Reisz et al. 1982; Moussavi–Harami 1990; Kasimovian *sensu* Benton 1993), which corresponds to the upper (Reisz 1977; 1981; Benton 1993), not uppermost (*sensu* Lee 1999), Pennsylvanian series. The North American upper Missourian has not been dated directly but correlates with the lowermost Stephanian of Europe (Hess and Lippolt 1986; Menning et al. 2000). The upper Stephanian A is dated with the $^{40}Ar/^{39}Ar$ method to 302.9 (±4.6 SE) MYR. Interpolation yields a lower Stephanian age of 306.1 MYA (±8.5 SE). Therefore, a conservative approach would be to date *Petrolacosaurus* within the range of 302.9–306.1 MYA, giving a minimum time estimate of the diapsid stem at 302.9 (±4.6) MYR.

The earliest synapsid, *Archaeothyris*, like the Joggins fauna, was also preserved in stumps of upright lycopod trees but is derived from the younger Florence locality, Morien Group, Nova Scotia. Biostratigraphic analysis shows that its age is the same as the coal swamp deposits in Linton, Ohio, and Nyrany, Czech Republic (Bell 1966), which are of Westphalian D age (Benton 1990). The bottom of this stratigraphic unit has been directly dated from European sanidine tuffs with the $^{40}Ar/^{39}Ar$ method at 309.2 MYA (±6.8 SE). However, a more conservative approach would be to characterize the Florence locality at the upper Westphalian D (i.e., Westphalian–Stephanian) border. The Westphalian ends at 306.1 MYA (±8.5 SE), assuming equal duration of Stephanian A and Westphalian D (Hess and Lippolt 1986; Menning et al. 2000). Thus, the age of *Archaeothyris* likely lies within the range 306.1–309.2 MYA, yielding a minimum time estimate of synapsid stem and synapsid–diapsid crown of 306.1 (±8.5) MYR. The compounded error derives from error around the $^{40}Ar/^{39}Ar$ dating of the nearest directly dated boundaries above (Stephanian A–B) and below (Westphalian D–C), including measurement error of the argon ratios and of the irradiation standard plus monitor age (respectively, 4.6 MY/302.9 MYA and 6.8 MYR SE/309.2 MYA [Hess and Lippolt 1986]), as well as interpolation error (2.27 MY/306.1 MYA) caused by lack of direct radiometric calibration at the Westphalian B–Stephanian A boundary (Menning et al. 2000).

### *Differing Calibration Age and Clock Error and the Avian Evolutionary Timescale*

Previous molecular point time estimates (without error) have placed several major bird divergences in the Cretaceous. These divergences include all modern superordinal diversification events, the origin of most orders, and the diversification within some orders (Passeriformes, Charadriiformes, Galliformes, ratites,

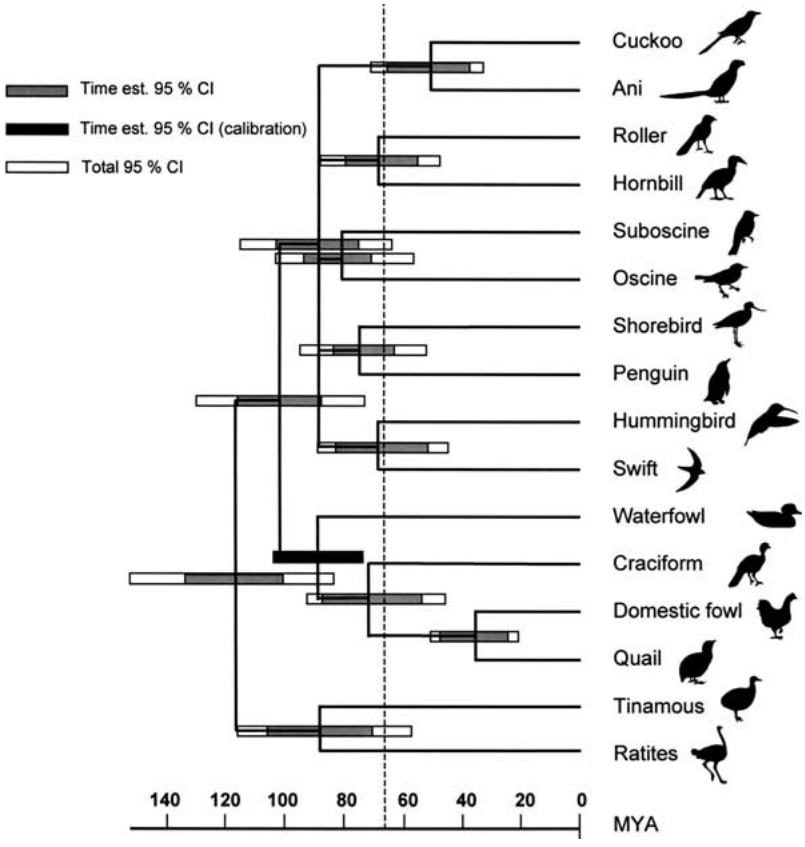


**Fig. 4.** Effect of molecular clock error in interpretation of the mtRNA early-bird timescale. Error is depicted here with 95% confidence intervals (CI) of age based on error in time estimation only and based on (total) error also accounting for genetic and fossil uncertainty in the calibration age. The dotted line shows the position of the K–T boundary.

and tinamous [Cooper and Penny 1997; van Tuinen and Hedges 2001; Paton et al. 2002, 2003]). Several of these results are robust to different data sets, models and calibrations, while taking into account time-estimation error (Rambaut and Bromham 1998; Paton et al. 2002, 2003). Yet these results are unsubstantiated by the fossil record. Here we determine which major lineage ages statistically remain within the Cretaceous when also accounting for error in molecular rate estimation. The upper 95% CI range based on time-estimation error alone statistically remains in the Cretaceous for all ordinal diversification except for within Galliformes, ratites, and tinamous (Fig. 4) (Paton et al. 2002, 2003). However, when also considering rate-estimation error, the number of Cretaceous lineages dwindles to only those encompassing all of Neornithes and all of Neognathae. Interestingly, the contribution of total error in reducing statistic resolution is large for the mtRNA data set, and even a shift to a 332-MYR age for the bird–mammal calibration maintains overlap with the K–T boundary for ordinal lineages (Cretaceous divergence found only for superordinal divergences). Thus, our conservative approach, though not arguing against Cretaceous (or K–T) ordinal diversification, suggests that the mtDNA data set alone cannot accurately resolve the time of ordinal diversification at the 95% confidence level. To further reduce error, nuclear sequencing efforts will be needed as well as similar scrutiny of other available data sets (complete mitochondrial genomes, nuclear c-mos) and calibrations. Locating solid internal calibration fossils and more rate-homogeneous loci will likely reduce future error in branch length estimation and in calibration age.

## Discussion

Our investigation of the primary literature indicates that the most conservative, or minimum, estimate for the bird–mammal calibration should be based on the stem age of synapsids (*Archaeothyris*: 306.1 ± 8.5 MYR). The older *Hylonomus* may lie on the stem to diapsids within Eureptilia, following the synapsid–diapsid divergence (Carroll 1988; Laurin and Reisz 1995). However, its fragmentary state, lack of diagnostic diapsid skull, and apparent lack of cladistic analyses make it ambiguous with respect to the origin of diapsids. *Protoclepsydrops* is an uncertain synapsid at best. Nevertheless, new findings or analyses could lock these two fossils in their alleged places and shift the minimum age for the bird–mammal calibration to 311 MYA (±1.7 SE). The nature of the immediate stem to living amniotes is currently unclear, and the fossil record therefore provides a liberal, maximum estimate for the synapsid–diapsid split (332.3 ± 13.4 MYR). The upper bound limit (306.1 MYA) is perhaps more likely because the earliest diapsids and synapsids

are morphologically similar. We note that a minimum age approach also was championed by Lee (1999), but his significantly younger age estimate (288 MYR) is not supported here. First, Lee oddly discounted the use of *Archaeothyris* because it exhibits only a single synapsid character even though that character (synapsid skull) defines Synapsida. Second, his divergence time of 288 MYR was based on placing *Petrolacosaurus* in the uppermost Pennsylvanian. This stratigraphic assignment conflicts with the first-hand literature of *Petrolacosaurus*, which describes a Missourian age (302.9–306.1 MYA). Furthermore, the Pennsylvanian–Permian boundary is currently set at 296 MYA using $^{40}Ar/^{39}Ar$ dating (Menning et al. 2000). Thus, use of a conservative minimum estimate supported by primary literature appears to be the most judicious approach, with the realization that the synapsid–diapsid split could, in the future, be pushed somewhat deeper into the Carboniferous but not likely beyond 332 MYA. Such revision could become necessary with uncovering new fossils possessing diagnostic characters, more local and comprehensive radiometric dating, but also with a shifting geologic timescale and increased resolution for the Carboniferous (Menning et al. 2000). Implementation of our new minimum age and error to the avian mtDNA data set shows decreased statistical resolution for extensive taxonomic-scale diversification occurring in the Cretaceous, as opposed to near the K–T boundary.

Besides implementing error in molecular clock studies, more study on genetic and fossil error is needed. Several major questions remain to be researched in detail, including (a) whether one should take the mean (or mode) of single gene estimates or use a concatenated gene approach (Nei et al. 2001), (b) whether the distance and tree models that are phylogenetically optimal also work best for molecular time estimation, and (c) whether reduction in error from better modeling of molecular evolution (Sanderson 1997; Thorne et al. 1998) is overwhelmed by additional error due to parameter increase and deviation from the appointed rates. Due to the often-observed complex nature of molecular substitution pattern, parameter-rich models are frequently preferred for best resolution of phylogenetic results, but they also carry more variance (Nei and Kumar 2000). Another question to be examined is (d) whether assuming insignificant error in molecular rate estimation is warranted in existing and future molecular clock studies. Our synapsid–diapsid example shows that error in rate estimation exceeds error in time estimation and that propagation of these errors has significant statistical ramification. A great deal of work remains to assess the validity of available methods and fossils and their utility in minimizing molecular clock error and maximizing precision and accuracy in dating the diversification of life.

*Acknowledgments.* For valuable comments we thank Robert Reisz, Hans Thewissen, Blair Hedges, Susan Holmes, Stephen Porder, three anonymous reviewers, and members of the Hadly lab.

## References

Arbogast BS, Slowinski JB (1998) Pleistocene speciation and the mitochondrial clock. Science 282:1955a

Arbogast BS, Edwards SV, Wakeley J, Beerli P, Slowinski JB (2002) Estimating divergence times from molecular data on phylogenetic and population genetic timescales. Annu Rev Ecol Syst 33:707–740

Bell WA (1966) Carboniferous rocks and fossil floras of northern Nova Scotia. Can Geol Surv Memoir 238:1–277

Benton MJ (1990) Phylogeny of the major tetrapod groups: morphological data and divergence dates. J Mol Evol 30:409–424

Benton MJ (1993) The fossil record 2. Chapman & Hall, London

Benton MJ (1997) Vertebrate paleontology, 2nd ed. Chapman & Hall, London

Benton MJ (1999) Early origins of modern birds and mammals: Molecules vs. morphology. Bioessays 21:1043–1051

Benton MJ, Ayala F (2003) Dating the tree of life. Science 5626:1698–1700

Bleiweiss R (1998) Slow rate of molecular evolution in high-elevation hummingbirds. Proc Natl Acad Sci USA 95:612–616

Bromham L (2002) Molecular clocks in reptiles: Life history influences rate of molecular evolution. Mol Biol Evol 19:302–309

Bromham L, Rambaut A, Harvey PH (1996) Determinants of rate variation in mammalian DNA sequence evolution. J Mol Evol 43:610–621

Brown WM, George M Jr, Wilson AC (1979) Rapid evolution of animal mitochondrial DNA. Proc Natl Acad Sci USA 67:1967–1971

Bossuyt F, Milinkovitch MC (2001) Amphibians as indicators of early Tertiary "out-of-India" dispersal of vertebrates. Science 292:93–95

Buckley TR, Simon C, Chambers GK (2001) Exploring among-site rate variation models in a maximum likelihood framework using empirical data: Effects of model assumptions on estimates of topology, branch lengths, and bootstrap support. Syst Biol 50:67–86

Calder JH (1994) The impact of climate change, tectonism and hydrology on the formation of Carboniferous tropical intermontane mires: The Springhill coalfield, Cumberland Basin, Nova Scotia. Palaeogeogr Palaeoclimat Palaeoecol 106:323–351

Carroll RL (1964) The earliest reptiles. J Linn Soc Zool 45:61–83

Carroll RL (1970) The ancestry of reptiles. Philos Trans R Soc Lond B 257:267–308

Carroll RL (1988) Vertebrate paleontology and evolution. Freeman, New York

Carroll RL (1991) The origin of reptiles. In: Schultze HP, Trueb L (eds) Origins of the higher groups of tetrapods: Controversy and consensus. Cornell University Press, Ithaca, NY, pp 331–354

Carroll RL (1995) Phylogenetic analysis of Paleozoic choanates. Bull Mus Nat Hist Nat Paris 17:389–445

Carroll RL (2001) The origin and early radiation of terrestrial vertebrates. J Paleo 75:1202–1213

Claoué-Long JC, Compston W, Roberts J, Fanning CM (1995) Two Carboniferous ages: a comparison of SHRIMP zircon dating with conventional zircon ages and $^{40}Ar/^{39}Ar$ analyses. In: Berggren WA, Kent DV, Aubry M-P, Hardenbol JA (eds) Geochronology, time scales, and global stratigraphic correlations: A unified temporal framework for an historical geology. Soc Econ Geol Mineral, Tulsa, OK, 54, pp 3–21

Clark J, Carroll RL (1973) Romeriid reptiles from the Lower Permian. Bull Mus Comp Zool Harvard Univ 144:353–407

Cooper A, Penny DP (1997) Mass survival of birds across the Cretaceous–Tertiary: Molecular evidence. Science 275:1109–1113

Dawson Sir JW (1891) Acadian geology, 4th ed. Macmillan, London

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

Easteal S, Collet C, Betty D (1995) The mammalian molecular clock. Springer Verlag, Heidelberg, Germany

Heckel PH (1975) Stratigraphy and depositional framework of the Stanton Formation in Southeastern Kansas. Kans Geol Bull 210:1–45

Hedges SB (2002) The origin of evolution of model organisms. Nat Rev Gen 3:838–849

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

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

Hess JC, Lippolt HJ (1986) $^{40}Ar/^{39}Ar$ ages of tonstein and tuff sandines: new calibration points for the improvement of the Upper Carboniferous time scale. Chem Geol 59:143–154

Hesselbo SP, Trewin NH (1984) Deposition, diagenesis and structures of the Cheese Bay Shrimp Bed, Lower Carboniferous, East Lothian. Scott J Geol 20:281–296

Hillis DM, Mable BK, Moritz C (1997) Applications of molecular systematics. In: Hillis DM, et al. (eds) Molecular systematics. Sinauer Associates: Sunderland, MA, pp 515–544

Janke A, Arnason U (1997) The complete mitochondrial genome of *Alligator mississippiensis* and the separation between recent archosauria (birds and crocodiles). Mol Biol Evol 14:1266–1272

Kimura M (1980) A simple method for estimating evolutionary rates of base substitutions through comparative studies of nucleotide sequences. J Mol Evol 16:111–120

Kumar S, Hedges SB (1998) A molecular timescale for vertebrate evolution. Nature 392:917–919

Kumar S, Subramanian S (2002) Mutation rates in mammalian genomes. Proc Natl Acad Sci USA 99:803–808

Laurin M, Reisz RR (1995) A reevaluation of early amniote phylogeny. Zool J Linn Soc 113:165–223

Laurin M, Reisz RR (1997) A new perspective on tetrapod phylogeny. In: Sumida SS, Martin KLM (eds) Amniote origins: Completing the transition to land. Academic Press, San Diego, pp 9–59

Lee MSY (1999) Molecular clock calibrations and metazoan divergence dates. J Mol Evol 49:385–391

Marshall CR (1990) The fossil record and estimating divergence times between lineages: Maximum divergence times and the importance of reliable phylogenies. J Mol Evol 30:400–408

Margoliash E, Smith E (1965) Structural and functional aspects of cytochrome *c* in relation to evolution. In: Bryson V, Vogel HJ (eds) Evolving genes and proteins. Academic Press, New York, pp 221–242

Menning M, Weyer D, Drozdzewski G, van Amerom HWJ, Wendt I (2000) A Carboniferous timescale 2000: Discussion and use of geological parameters as time indicators from Central and (Western Europe). Geol Jahrb A 156:3–44

Milner AR (1987) The Westphalian tetrapod fauna: some aspects of its geography and ecology. J Geol Soc Lond 144:495–506

Moussavi-Harami R (1990) Stratigraphy, petrology, and depositional environment of sandstones in the Rock Lake Shale member of the Stanton Limestone (Missourian stage, Upper Pennsylvanian) in Southeastern Kansas. Kans Geol Surv 5:1–44

Nei M, Kumar S (2000) Molecular clocks and linearized trees. In: Nei M, Kumar S (eds) Molecular evolution and phylogenetics. Oxford University Press, New York, pp 187–206

Nei M, Xu P, Glazko G (2001) Estimation of divergence times from multiprotein sequences for a few mammalian species and several distantly related organisms. Proc Natl Acad Sci USA 98:2497–2502

Nunn GB, Stanley SE (1998) Body size effects and rates of cytochrome b evolution in tube-nosed seabirds. Mol Biol Evol 15:1360–1371

Ohta T, Gillespie JH (1996) Development of neutral and nearly neutral theories. Theor Pop Biol 49:128–142

Paton RL, Smithson TR, Clack JA (1999) An amniote-like skeleton from the Early Carboniferous of Scotland. Nature 398:508–513

Paton T, Haddrath O, Baker AJ (2001) Complete mitochondrial DNA genome sequences show that modern birds are not descended from transitional shorebirds. Proc R Soc Lond B 269:839–846

Paton TA, Baker AJ, Groth JG, Barrowclough GF (2003) RAG-1 sequences resolve the phylogenetic relationships within Charadriiform birds. Mol Phylogenet Evol 29:268–278

Rambaut A, Bromham L (1998) Estimating divergence dates from molecular sequences. Mol Biol Evol 15:442–448

Reisz RR (1972) Pelycosaurian reptiles from the Middle Pennsylvanian of North America. Bull Mus Comp Zool Harvard Univ 144:27–62

Reisz RR (1977) *Petrolacosaurus*, the oldest known diapsid reptile. Science 196:1091–1093

Reisz RR (1981) A diapsid reptile from the Pennsylvanian of Kansas. Special Publ Mus Nat Hist Univ Kans 7:1–74

Reisz RR (1997) The origin and early evolutionary history of amniotes. TREE 12:218–222

Reisz RR, Heaton MJ, Pynn B (1982) Vertebrate fauna of late Pennsylvanian rock lake shale near Garnett, Kansas: Pelycosauria. J Paleo 56:741–750

Rolfe WDI, Clarkson ENK, Panchen AL (eds) (1993) Volcanism and early terrestrial biotas, Trans R Soc Edinburgh Earth Sci 84, parts 3 and 4

Sanderson MJ (1997) A nonparametric approach to estimating divergence times in the absence of rate constancy. Mol Biol Evol 14:1218–1231

Shapiro SG (1991) Uniformity in the nonsynonymous substitution rates of embryonic beta-globin of several vertebrate species. J Mol Evol 32:122–127

Smithson TR (1989) The earliest known reptile. Nature 342:676–678

Smithson TR, Carroll RL, Panchen AL, Andrews SM (1993) *Westlothiana lizziae* from the Viséan of East Kirkton, West Lothian, Scotland, and the amniote stem. Trans R Soc Edinburgh Earth Sci 84:383–412

Sumida SS (1997) Locomotor features of taxa spanning the origin of amniotes. In: Sumida SS, Martin KLM (eds) Amniote origins: Completing the transition to land. Academic Press, San Diego, pp 353–398

Thewissen JGM, Williams EM, Roe LJ, Hussain ST (2001) Skeletons of terrestrial Cetaceans and the relationship of whales to artiodactyls. Nature 413:277–281

Thorne JL, Kishino H, Painter IS (1998) Estimating the rate of evolution of the rate of molecular evolution. Mol Biol Evol 15:1647–1657

van Tuinen M, Hedges SB (2001) Calibration of avian molecular clocks. Mol Biol Evol 18:206–213

van Tuinen M, Hedges SB (2004) The effect of external and internal fossil calibrations on the avian evolutionary timescale. J Paleo 78:45–50

Wilson AC, Carlson SS, White TJ (1977) Biochemical evolution. Annu Rev Biochem 46:573–639

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