Incorporating phylogenetic uncertainty on phylogeny-based palaeontological dating and the timing of turtle diversification

Juliana Sterli[a,*], Diego Pol[a] and Michel Laurin[b]

[a]*CONICET-Museo Paleontológico Egidio Feruglio, Av. Fontana 140, 9100 Trelew, Chubut Province, Argentina;* [b]*UMR 7207, CNRS/MNHN/UPMC, Muséum National d'Histoire Naturelle, Batiment de Géologie, Case postale 48, 43 rue Buffon, F-75231 Paris Cedex 05, France*

Abstract

Methods improving the performance of molecular dating of divergence time of clades have improved dramatically in recent years. The calibration of molecular dating using the first appearance of a clade in the fossil record is a crucial step towards inferring the minimal diversification time of various groups and the choice of extinct taxa can strongly influence the molecular dates. Here, we evaluate the uncertainty on the phylogenetic position of extinct taxa through non-parametric bootstrapping. The recognition of phylogenetic uncertainty resulted in the definition of the Bootstrap Uncertainty Range (BUR) for the age of first appearance of a given clade. The BUR is calculated as the interval of geological time in which the diversification of a given clade can be inferred to have occurred, based on the temporal information of the fossil record and the topologies of the bootstrap trees. Divergence times based on BUR analyses were calculated for three clades of turtles: Testudines, Pleurodira and Cryptodira. This resulted in extensive uncertainty ranges of topology-dependent minimal divergence dates for these clades.

© The Willi Hennig Society 2012

Dating the tree of life used to be a nearly exclusively palaeontological enterprise, to the point that it was widely accepted that if a taxon lacked a fossil record, its evolutionary radiation could not be dated (Gingerich, 1979; Donoghue et al., 1989). Until the development of molecular dating methods (Zuckerkandl and Pauling, 1962), the only other way to date the appearance of taxa was biogeography combined with geological history, although this criterion is perhaps not as strong as palaeontological or molecular data because the present geographical distribution of taxa may not reflect its past distribution (e.g. Pascual et al., 1992; Gandolfo et al., 2011). Dating the origin of clades based on molecular clock methods assuming a constant evolutionary rate of proteins and DNA has evolved rapidly. In recent years, the development of molecular dating methods that do not require a single, universal rate for the molecular clock (Sanderson, 1997, 2002; Drummond et al., 2006;

Drummond and Rambaut, 2007), along with the notable increase in DNA sequence availability have drastically changed the picture, to the point that in the last decade, work on dating the major evolutionary events in the tree of life has been largely dominated by studies based on molecular data (e.g. Kumar and Hedges, 1998; Bromham et al., 1999; Lee, 1999; van Tuinen and Hedges, 2001; Braun and Kimball, 2002; Tavaré et al., 2002; Douzery et al., 2004; Near et al., 2005; Welch et al., 2005; Donoghue and Benton, 2007; Hugall et al., 2007; Dornburg et al., 2011).

In these methods, however, molecular branch lengths can be converted to time only through calibration criteria, the most frequently used and reliable of which are those based on the fossil record. The calibration of molecular clocks is a critical step in molecular inferences on the time of diversification of clades (Ho and Phillips, 2009), and it has been empirically demonstrated that the choice of calibration constraints has a strong influence on the resulting molecular ages (Brochu, 2004, 2006; Marjanović and Laurin, 2007; Sauquet et al., 2012). Thus, irrespective of methodological concerns on the

*Corresponding author:
E-mail address:* jsterli@mef.org.ar

assumptions of molecular clock methods (Rodríguez-Trelles et al., 2002; Shaul and Graur, 2002; Graur and Martin, 2004; Britton, 2005), progress on dating the tree of life will be achieved if progress in collecting and carefully interpreting palaeontological data relevant to dating the tree of life advances at a pace comparable with gathering molecular data to perform combined analyses.

In the last two decades, compilations and methods were developed to better use the fossil record to date the tree of life. Regarding the treatment of primary data, Parham et al. (2012) developed a specimen-based protocol in five steps to justify the choice of the extinct taxa used for calibration. Benton (1993) edited a large compilation of the fossil record that emphasizes family-level taxa and has been used to select calibration constraints. Several studies have focused on refining the use of stratigraphic information to determine the first appearance of groups in the fossil record. Foote et al. (1999) developed a method based on diversification models and the inferred maximal fossilization rate and suggested that several Cretaceous molecular dates of origin of various high-ranking placental mammal taxa were probably incorrect. Marshall (1990, 1994, 1997) developed phylogeny-free methods that use the stratigraphic distribution of fossils to estimate the confidence interval on the true date of origin or of extinction of taxa. Recently, this was modified to incorporate simple diversification models, as may be required to deal with dates of origin of speciose taxa (Marjanović and Laurin, 2008).

Several authors have noted the importance of considering phylogenetic relationships of extinct taxa for determining the origin of clades and for selecting potential calibration points (e.g. Müller and Reisz, 2005; Benton and Donoghue, 2007). Along these lines, Marjanović and Laurin (2007) and Marshall (2008) also proposed methods that use both phylogenetic and stratigraphic data, in the first case to show that the taxon Lissamphibia probably appeared in the Permian rather than in the Devonian (San Mauro et al., 2005; Roelants et al., 2007) or Early Carboniferous (Zhang et al., 2005), as suggested by various molecular studies. Thus, progress has been made in improving the ways information from the fossil record is used to date the tree of life, although at a slower rate than advances in methods that infer divergence dates based on molecular data.

Incorporation of the uncertainty of the position of fossil taxa in a phylogeny for dating the tree of life is an important (see Ho and Phillips, 2009) but largely overlooked problem and has recently generated other novel approaches to date phylogenetic trees using a combination of molecular, morphological and stratigraphic data (Lee et al., 2009; Pyron, 2011). These analyses require applying model-based (e.g. Bayesian) methods for inferring phylogenies on mixed (molecular and morphological) datasets (Pyron, 2011) but are contingent on the basic assumptions used for analysing

morphological data, such as the adequacy of model-based analysis for morphological data.

Here we propose improvements to the method presented by Marjanović and Laurin (2007) that include the phylogenetic uncertainty associated with the placement of fossil taxa in cladistics analyses and its impact on the inference of divergence dates that can be inferred by integrating the stratigraphic information of extinct taxa in morphology-based phylogenetic analyses. A common fact inherent to fossil specimens is their incompleteness. Usually the hard parts are much more likely to fossilize than the soft tissues. This implies a high percentage of missing data for extinct taxa in a morphological matrix. The presence of high levels of missing data may result in multiple locations of taxa in the most-parsimonious trees (MPTs; Norell and Wheeler, 2003) or may yield alternative positions in marginally suboptimal trees for extinct taxa, reducing the support for their phylogenetic placement in the MPTs. The new method proposed here is a simple and parsimony-based method that incorporates the phylogenetic uncertainty on the position of extinct taxa in morphology-based cladistic analyses. Our approach requires only a phylogenetic dataset (which can be purely morphological, or include both molecular and morphological data) and basic stratigraphic data on the age of the fossils included in the analysis. The data are then subjected to phylogenetic (e.g. parsimony) analysis and bootstrap analyses and the results are integrated with stratigraphic data to infer uncertainty associated with the minimal age implied by the phylogeny for the clades of the tree. We do not tackle the difficult issue of the maximal age of clades, which requires other analytical methods (e.g. Marjanović and Laurin, 2008; Marshall, 2008).

We test this procedure on a new dataset of turtles that has a dense taxonomic sampling of extinct and extant taxa. This provides an ideal test-case for palaeontological dating methods because of the rich fossil record of turtles and because their phylogenetic relationships have been intensively studied during the last decade (e.g. Shaffer et al., 1997; Near et al., 2005; Gaffney et al., 2006, 2007; Joyce, 2007; Sterli, 2008, 2010). Although our approach provides temporal calibration for all nodes of the phylogenetic tree of turtles, we have focused, for the sake of simplicity, on the time of diversification of three major crown clades (Testudines, Pleurodira and Cryptodira) that could be used as dating constraints.

Materials and methods

Phylogenetic analysis

A large-scale morphological data matrix of extant and extinct turtles and several potential sister groups was gathered, expanding the taxon and character sampling

from previous studies (Joyce, 2007; Sterli and de la Fuente, 2011) reaching a total of 100 taxa and 235 characters (Supporting Information, Data S1). The outgroups include four taxa that have been suggested to be close relatives of turtles: *Sphenodon punctatus* (Müller, 2003, 2004; Hill, 2005), *Simosaurus gaillardoti* (Rieppel and deBraga, 1996; Rieppel and Reisz, 1999), *Anthodon serrarius* (Lee, 1995, 2001) and *Owenetta kitchingorum* (Reisz and Laurin, 1991; Laurin and Reisz, 1995). An equally weighted parsimony analysis was performed in TNT (Goloboff et al., 2008a,b) using a heuristic search of 1000 replicates of Wagner trees with random addition of taxa and TBR (saving ten trees per replicate), followed by a second TBR cycle to find all MPTs. Twenty multistate characters were ordered (see supporting Data S1). Nodal support was assessed using Bremer support and Bootstrap analysis and both were calculated in TNT. Bootstrap values were calculated performing 1000 replicates and they are shown as GC frequencies (Group present–Contradicted; Goloboff et al., 2003). The GC frequencies are the difference between the frequency in which a given group is retrieved in the bootstrap replicates and the most frequent incompatible group (Goloboff et al., 2003).

We have followed the recently proposed phylogenetic definitions of major clades of turtles (Joyce et al., 2004). For the sake of clarity we repeat here the definitions of the three clades of interest. Testudines is defined as the crown clade arising from the last common ancestor of *Chelus* (orig. *Testudo*) *fimbriatus* and *Chelonia mydas*. Pleurodira refers to the crown clade arising from the last common ancestor of *Pelomedusa* (orig. *Testudo*) *subrufa*, *Chelus fimbriatus* and *Podocnemis expansa*. Cryptodira refers to the crown group arising from the last common ancestor of *Chelonia mydas*, *Testudo graeca* and *Trionyx triunguis*.

Phylogenetic inference of divergence times based on the fossil record

As noted above, the method employed for determining the minimal divergence time between turtle clades is based on the calibration of phylogenetic trees with the geological time of first appearance of each extinct taxon included in the phylogenetic analysis. There is an inherent uncertainty in the age of each extinct taxon that should be taken into account, but here we also incorporate the phylogenetic uncertainty on the position of extinct taxa (through the alternative positions these taxa take on bootstrap replicates), which leads to the recognition of a range of uncertainty for the minimal age of divergence for each node of the MPTs.

Geochronological uncertainty in the age of extinct taxa. The temporal information is primarily (but not only) determined by the geological age of the first

appearance datum (FAD) of each extinct taxon included in the phylogenetic analysis. All FADs have some degree of associated uncertainty that stems from the incompleteness of the fossil record, the incompleteness of the known stratigraphic range of taxa influenced by the palaeontological sampling effort, and the fossilization potential of the organisms, which depends partly on the environmental settings in which they lived (Marshall, 1990, 1994, 1997, 2008; Lu et al., 2006; Marjanović and Laurin, 2008; Paul, 2009). Moreover, the age of each fossil also has an uncertainty associated with the dating method used. This age is usually determined by the age of the sediments in which the extinct taxon is recorded. The sediments can sometimes be directly dated, so that the age of the fossil is bracketed by radioisotopic dates obtained from rock samples of the sequence in which it is found. However, more commonly, the age of fossiliferous sediments is indirectly determined by correlation with other rock bodies that have been radiometrically dated, adding to the error associated with the age of the fossil. Therefore, the minimum and maximum possible age of each fossil should be considered when a phylogenetic tree is calibrated using geochronological information of fossil taxa (Pol and Norell, 2006). Several methods have been developed that estimate confidence intervals that extend the stratigraphic range of a fossil taxon beyond the FAD (Strauss and Sadler, 1989; Marshall, 1990, 1994, 1997; Foote and Raup, 1996; Foote, 1997; Solow and Smith, 1997; Wagner, 2000; Solow, 2003; Marjanović and Laurin, 2008). These methods are sometimes difficult to use, especially for extinct taxa known from single specimens (a common case in Mesozoic vertebrates of continental environments). In our approach, we have used the geochronological uncertainty associated with the extinct taxa through the incorporation of a maximum and minimum age for the FAD of each extinct taxon based on the uncertainty on the age in which the fossil was found (supporting Data S2), but the entire procedure could be adapted to incorporate further extensions of the FAD based on any other method. In most cases, this uncertainty represents the duration of one or two geological stages and it expresses uncertainty about the position of the fossil within the relevant stages. For example, if a taxon is found in a formation estimated to be of Aptian–Albian age, we are going to recognize its maximum FAD as being 125 Ma (lower bound of the Aptian following Walker and Geissman, 2009) and its minimum FAD as being 99.6 Ma (upper bound of the Albian following Walker and Geissman, 2009).

Bootstrap Uncertainty Range (BUR). We define the BUR as the range of minimal geological age of a given node, considering the phylogenetic uncertainty as measured by the inclusion/exclusion of extinct taxa in any given clade in the trees resulting from the bootstrap

replicates. The proposed procedure requires a bootstrap analysis of the morphological dataset and the calibration of the trees obtained in each of the bootstrap replicates against geological time (using the maximum and minimum age of FADs). We performed the bootstrap analysis using parsimony, but the method would also work with other resampling techniques (e.g. jackknife) or phylogenetic methods (maximum likelihood, Bayesian, etc.). Each bootstrap replicate creates a slightly modified data matrix by resampling the original characters with replacement (Felsenstein, 1985) and a parsimony heuristic search is conducted on the modified data matrix. The collection of phylogenetic trees derived from bootstrap replicates are saved and calibrated using the temporal information of fossil data. If the inclusion of the oldest fossil of a given clade is poorly supported by the available morphological data, this taxon will be frequently positioned outside that clade in the bootstrap trees (and the minimal age inferred for that node will be younger than in the MPTs). Conversely, if an ancient sister group of a given clade is positioned outside based on scarce morphological data, the ancient fossil will be retrieved within the clade in some of the bootstrap replicates (and the minimal age inferred for that node will be older than in the MPTs).

Rather than relying only on the MPTs, our bootstrap analysis examines the frequency with which various geologically old taxa are positioned within clades of interest defined on the basis of extant specifiers for crown-group clades Testudines, Cryptodira and Pleurodira. This analysis results in a range of the palaeontologically based divergence dates of these clades, which we refer as the BUR. The BUR has a minimum (BUR_{min}) and maximum (BUR_{max}) bound, but both refer to the minimal age that the clade must have.

Treatment of zero-length branches. Phylogenies calibrated against geological time using only palaeontological data (see Pol et al., 2004) usually have multiple zero-length branches. In calibrated phylogenies, branch length represents time (rather than evolutionary transformations or a combination of both), which would imply successive cladogenetic events occurring at the same time. Although multiple cladogenesis can occur (almost) simultaneously, in most cases these result from the incompleteness of the fossil record in capturing successive stages of the evolution of a particular clade. Here we have selected a non-zero minimum branch length for the calibrated trees, following the empirical analyses of Marjanović and Laurin (2007), later further developed for comparative analyses by Laurin et al. (2009). The choice of a minimum branch length depends on the density of the taxonomic sampling, the tree topology, the observed stratigraphic ranges of taxa, and the interaction between these factors that gives indirect evidence about the quality of the fossil

record. Here we explored a minimum branch length of 0.1 to calibrate both the most-parsimonious and the boostrap trees.

Calculating the BUR. BUR_{min} and BUR_{max} are calculated separately, assuming either the youngest or the oldest possible age of each extinct taxon (as determined by the geochronological uncertainty of the strata in which the fossil is found; supporting Data S2), a minimum temporal branch length that extends the temporal branch length of each lineage of the tree, and 95% of the bootstrap trees. To determine BUR_{min} the boostrap trees are calibrated using the most recent possible age of all extinct taxa included in the phylogenetic analysis. The entire procedure is automated in a script for TNT that performs the bootstrap replicates, saves the bootstrap trees, and calibrates them in order to find the frequency of the different divergence ages that are inferred for a given node in the bootstrap replicates (see supporting Data S3). This procedure creates a distribution of inferred ages (for a given node) that varies depending on the frequency at which different extinct taxa are included in or excluded from that particular clade among the bootstrap trees. BUR_{max} is calculated in an identical way, except that we use the oldest possible age of each fossil taxon included in the phylogenetic analysis. Once both procedures are performed, in order to determine the uncertainty age range associated with a given node of the tree (BUR) we discard the 2.5% of the bootstrap trees yielding the youngest ages (from the distribution of ages inferred during the BUR_{min} run) and the oldest ages (from the distribution of the ages inferred during the BUR_{max} run) for the clade being analysed. Thus, the BUR includes only the 95% of the confidence interval of ages of both the BUR_{min} and BUR_{max} procedure, which is akin to computing a 95% confidence interval on the minimal age of the taxon.

The calibration of phylogenetic trees (derived from bootstrap replicates in this case) optimizes a Sankoff character in which the age of each taxon is represented by a character state and the transformation costs are set as irreversible, as detailed by Pol and Norell (2006, 2001). The calibrated trees (with branches representing time) are saved in nexus format and they can be visualized using the Stratigraphic Tools (Josse et al., 2006) for Mesquite (Maddison and Maddison, 2010) and the geological timescale of Walker and Geissman (2009), as well as any other program that displays nexus format trees with branch lengths (e.g. FigTree, Rambaut, 2006–2009).

Alternative model-based approaches for phylogeny calibration using fossils

Model-based methods have been recently proposed for incorporating phylogenetic uncertainty on the

placement of fossils for calibrating phylogenies, namely Bayesian methods that retrieve a posterior distribution for the monophyly and age of a given clade (Lee et al., 2009). Model-based approaches, with greater algorithmic sophistication and the possibility of estimating character evolution models, have been increasingly used during the last decade. However, we propose here a parsimony-based approach to the problem for several reasons (see also Discussion). First, model-based approaches for the analysis of morphological data have not yet gained popularity and the adequacy of these methods for this type of data is not as obvious as for molecular data. In particular, the assumption of a homogeneous Markov model is difficult to accept for morphological datasets and deviations from this assumption can severely affect the performance of model-based methods (Chang, 1996; Kolaczkowski and Thornton, 2004; Thornton and Kolaczkowski, 2005). Second, recent studies have pointed to the existence of problems in the performance of model-based methods (i.e. maximum likelihood, Bayesian) when analysing datasets with large quantities of missing data (Goloboff and Pol, 2005; Simmons, 2011, 2012). In summary, we propose a parsimony-based approach to this issue in accordance with the widespread use of this method for morphological (in particular palaeontolog-ical) datasets, rather than the still poorly understood behaviour of model-based approaches for this type of data.

Results

Phylogenetic analysis

Parsimony analysis of the morphological dataset resulted in 150 MPTs of 892 steps. The general topology of the strict consensus tree (Fig. 1; supporting Data S1 and 4) resembles that of Joyce (2007), Sterli (2008), Anquetin et al. (2009), Danilov and Parham (2008) and Sterli and de la Fuente (2011) to the extent that all the most ancient (Triassic – Middle Jurassic) turtles such as *Proganochelys quenstedti*, *Palaeochersis talampayensis*, *Kayentachelys aprix*, *Condorchelys antiqua* and *Eilean-chelys waldmani* are placed basal to crown-group Testudines (Fig. 1). This contrast with previous hypotheses of turtle evolution that placed some of these taxa as closely related to Pleurodira or Cryptodira (Gaffney et al., 1991; Gaffney, 1996). Testudines is supported by two unambiguous synapomorphies common to all trees (character 132, a single suprapygal; and character 195, first thoracic rib intermediate to short). Cryptodira and Pleurodira are

Fig. 1. Time-calibrated phylogenetic tree of Testudinata assuming minimal internal and terminal branch lengths of 3 Myr (3Myr was chosen to make the tree more legible). The topology represents the strict consensus of the 150 MPTs. Crown-group clades are identified with white circles, and other widely recognized clades are identified with black circles. The calibration was made using the Stratigraphic Tools in Mesquite and the geological timescale of Walker and Geissman (2009). Ages are in million years. For more information on the data matrix and phylogenetic analysis see supporting Data S1. Neo, Neogene; PALEO, Paleogene; PER, Permian.

recovered as monophyletic groups (with their associated stems), forming crown-group Testudines (Fig. 1). Pleurodira is characterized by ten common unambiguous synapomorphies (character 49, cranio-quadrate space floored by quadrate and prootic; character 54, processus trochlearis oticum absent; character 62, pterygoid–basioccipital contact absent; character 79, ventral ridge on opisthotic absent; character 84, basisphenoid pentagonal; character 97, foramen posterius canalis caroticus internus formed by prootic, prootic + basisphenoid or prootic + pterygoid; character 158, xiphiplastron with anal notch; character 162, only one gular; character 174, inframarginals absent; character 216, ischium sutured to plastron by a broad suture) and one synapomorphy present in some trees (character 182, formed cervical articulations). Three unambiguous synapomorphies common to all MPTs characterize Cryptodira (character 1, nasals absent; character 154, axillary buttress on peripheral 4; character 157, inguinal buttress on peripheral 7). Many widely recognized clades are recovered as monophyletic groups. These include Trionychia (Trionychidae + Carettochelyidae), Kinosternoidea (Kinosternidae + Dermatemydidae), Chelonioidea (Cheloniidae + Dermochelyidae), Testudinoidea, Chelidae and Pelomedusoides among extant clades, as well as Adocidae, Baenidae, Pleurosternidae and Meiolaniidae among extinct taxa (Fig. 1).

Divergence times of major turtle clades

The MPTs were calibrated using the maximum and minimum ages for each extinct taxon as determined by the geochronological uncertainty of the strata in which the fossil is found (supporting Data S2). The bootstrap trees were calibrated for determining BUR_{min} and BUR_{max} of the three major clades of turtles of our interest (Testudines, Pleurodira and Cryptodira; see Fig. 2, Table 1; supporting Data S5). Comparison of the ages calculated using the collection of bootstrap trees with those ages calculated using MPTs were conducted to explore how phylogenetic uncertainty on the placement of fossil taxa can affect estimates of divergence times in this morphological phylogenetic study, which includes numerous extinct and extant turtle taxa.
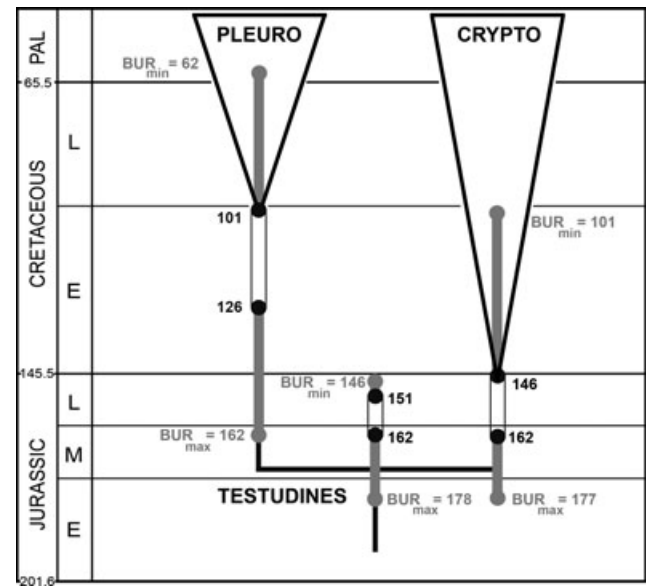
Fig. 2. Schematic tree showing the main clades discussed in the paper (Testudines, Cryptodira, Pleurodira) and the maximum and minimum ages calculated in the MPTs (dark circles) and BUR_{min} and BUR_{max} (grey bars) calculated in the bootstrap trees. Ages are in million years. E, Early; M, Middle; L, Late; PAL, Paleogene.

Age of Testudines. The range of minimum ages inferred for Testudines based only on the MPTs ranges from 162 Ma (Callovian, Middle Jurassic) to 151 Ma (Tithonian, Late Jurassic), given that there are uncertainties in the age of extinct taxa (see supporting Data S2). The minimum age for Testudines of 151 Ma is determined by the presence of three taxa within this clade, all from the Late Jurassic: *Xinjiangchelys latimarginalis*, *Yehguia tatsuensis* and *Caribemys oxfordiensis*. These taxa are deeply nested within Testudines; *C. oxfordiensis* in the stem of Pleurodira and *X. latimarginalis* and *Y. tatsuensis* in the stem of Cryptodira and, therefore, if a non-zero length branch is assumed for the phylogenetic tree, the maximum age would extend the divergence time of Testudines to the early Middle Jurassic (Figs 1 and 2). Irrespective of this point, the major focus of this contribution is how the phylogenetic uncertainty on the position of some fossil taxa affects the inferences of

Table 1
Bootstrap Uncertainty Range (BURs) calculated in this study compared with estimated ages of previous studies

Taxon	BURs			Previous estimates from the literature						
	Max	Min	Span	D11	A09	H07 (nuc., 5 cal.)	H07 (aa, 5 cal.)	N05	D&P 08	J07
Testudines	178	146	32	324–182	225–212	221–193	259–203	210*	176–161	161–145
Pleurodira	162	1	161	163–133	158–17	159–115	220–144	184–168	–	–
Cryptodira	177	101	76	–	212–120	187	–	187–163	151–145	145–140

See supporting Data S5 for the probability distribution of the BURs. A09, Alfaro et al. (2009); D11, Dornburg et al. (2011); H07, Hugall et al. (2007); nuc., nucleotides; cal., calibrations; aa, amino acids; N05, Near et al. (2005); D&P08, Danilov and Parham (2008); J07, Joyce (2007).

*Age fixed in previous studies.

divergence times (i.e. BUR). The BUR_{max} of Testudines is 178 Ma (Toarcian, Early Jurassic) (supporting Data S5), as in several bootstrap trees the Middle Jurassic stem Testudines *Chengyunchelys* sp. and *Siamochelys peninsularis* are positioned within Cryptodira (and consequently within Testudines). The bootstrap trees derived from our dataset indicate that there is high support for positioning the Late Triassic *Proterochersis robusta* or the Early Jurassic *Kayentachelys aprix* outside Testudines (contrary to alternative hypotheses of turtle evolution; see Discussion). The BUR_{min} of Testudines is 146 Ma (Tithonian, Late Jurassic) given that in more than 95% of the bootstrap trees (supporting Data S5) the above-mentioned Late Jurassic taxa are placed within Testudines, rejecting a more recent (i.e. Late Cretaceous) origin of this clade. This result suggests that the minimal divergence age of Testudines is reasonably well constrained within the late Early Jurassic to the Late Jurassic by palaeontological data and is not severely affected by phylogenetic uncertainty.

Age of Pleurodira. The range of minimum ages calculated for Pleurodira considering only the MPTs range from 126 Ma (Barremian, Early Cretaceous) to 101 Ma (Albian, Early Cretaceous) and are determined by the inclusion of *Prochelidella cerrobarcinae* and *Araripemys barretoi* within the crown group. The BUR of Pleurodira is much greater than the maximum and minimum calculated ages on the MPTs, indicating a large degree of uncertainty on the phylogenetic position of some extinct taxa. In some bootstrap replicates (21%; supporting Data S5) several stem Pleurodira from the Late Jurassic (i.e. *Notoemys laticentralis*, *Caribemys oxfordiensis*, *Platychelys oberndorferi*) are placed within the crown group, thus increasing the age of the clade by 36 Myr (BUR_{max} = 162 Ma, Callovian, Middle Jurassic). The BUR_{min} is 1 Ma (Quaternary); this decrease in the minimum possible age of Pleurodira occurs when the only two Cretaceous taxa included in the analysis (*Prochelidella cerrobarcinae* and *Araripemys barretoi*) and the Paleocene *Yaminuechelys major* are positioned outside crown-group Pleurodira, although these occur in only 4% of the bootstrap trees (supporting Data S5).

Age of Cryptodira. The range of minimum ages calculated for Cryptodira based only on the MPTs ranges from 162 Ma (Callovian, Middle Jurassic) to 146 Ma (Tithonian, Late Jurassic). In the present phylogeny, the Late Jurassic *Yehguia tatsuensis* is the most ancient cryptodiran turtle and determines the time of diversification of this clade. The minimum age for Cryptodira basically coincides with the minimum possible age of *Yehguia tatsuensis* but when considering non-zero length branches in the calibration of the phylogeny the maximum possible age of Cryptodira extends the

diversification time of this clade back to the Middle Jurassic, given that *Yehguia tatsuensis* is deeply nested within Cryptodira (as closely related to Trionychia). The BUR of Cryptodira extends the ages calculated for the MPTs unevenly towards younger and older dates. The BUR_{min} for Cryptodira is younger than the minimum calculated age by 46 Myr, opening the possibility that the diversification of Cryptodira was about 100 Ma, during the Early Cretaceous (Albian). Such a young age is inferred from bootstrap trees that position *Yehguia tatsuensis* outside Cryptodira (17% of the boostrap trees). This important decrease in the age of diversification of Cryptodira is not only caused by the exclusion of *Yehguia tatsuensis*, but also because of the poor Early Cretaceous fossil record of turtles. The discovery of cryptodiran turtles from the Lower Cretaceous and their inclusion in the dataset would help to reduce the critical role of *Yehguia tatsuensis* and the drastic difference between the minimum age calculated on the MPTs and the BUR_{min} of Cryptodira. The BUR_{max}, instead, extends the possible age of Cryptodira by only 15 Myr (BUR_{max} = 177 Ma, Toarcian, Early Jurassic) as in 20% of the bootstrap trees some basal Middle Jurassic taxa (e.g. *Chengyunchelys* sp., *Siamochelys peninsularis*) are placed within Cryptodira.

Discussion

Comparison between palaeontological and molecular ages of major turtle clades

The approach taken here for obtaining palaeontological dating considering the phylogenetic uncertainty provides uncertainty ranges on the minimal time of diversification of the major clades of turtles. This procedure yields ranges of age estimates that can be substantially broader than the inferred ages when the uncertainty on the placement of extinct taxa is ignored (i.e. considering only the MPTs). Therefore, these ranges are more likely to overlap with molecular dates than the palaeontological dates provided in previous studies. Below, we compare previous palaeontological and molecular dates of origin of three crown clades of turtles analysed in this study: Testudines, Pleurodira and Cryptodira.

Testudines. The diversification time of Testudines (Fig. 3) is one of the most discussed topics in the evolutionary history of turtles (e.g. Gaffney et al., 1987; Near et al., 2005; Danilov and Parham, 2008). According to Gaffney et al.'s (1991) phylogeny, the origin of Testudines dates back to the Late Triassic (Norian, 228–204 Ma). This ancient date reflects the position of *Proterochersis robusta*, the oldest known turtle (Norian), as a stem pleurodiran. The phylogenetic position of
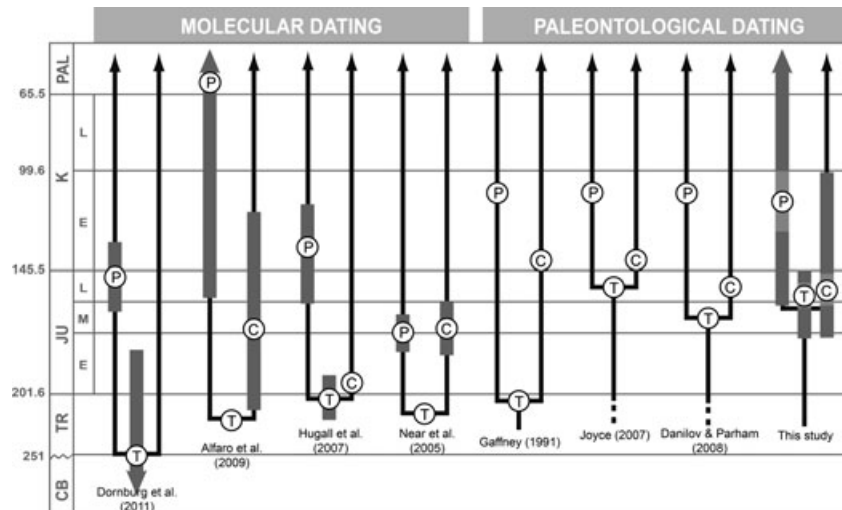
Fig. 3. Comparison of the geological age of Testudines (T), Cryptodira (C) and Pleurodira (P) suggested by various studies. Ages are in million years. PM, Permian (not in scale); E, Early; K, Cretaceous; JU, Jurassic; L, Late; M, Middle; PAL, Paleogene; TR, Triassic. Dark grey bars indicate confidence intervals for node age calculation. In our study, light grey indicates age ranges calculated on the MPTs and dark grey indicates BURs.

Proterochersis robusta (and therefore the ancient age of Testudines) has been recently challenged by several phylogenetic analyses that place this taxon outside crown-group Testudines (e.g. Rougier et al., 1995; Joyce, 2007; Danilov and Parham, 2008; Sterli, 2008). This result is corroborated in our extended analysis, and the oldest member of Testudines is the Late Jurassic cryptodiran *Yehguia tatsuensis*, a result that is consistent (Fig. 3) with the hypothesis of Danilov and Parham (2006, 2008; see below). Thus, except for the phylogenetic analysis of Gaffney et al. (1991), a consensus is emerging from several morphological phylogenies that suggest the oldest crown-group turtle is recorded in the Late Jurassic. In our phylogeny (Fig. 1), *Proterochersis robusta* is not even close to crown-group Testudines (not only in the MPTs but also in the trees from the bootstrap replicates; see above). This leads to a minimal divergence time for Testudines in our phylogenetic analysis that is well constrained to the Middle–Late Jurassic by palaeontological data (range of minimum ages in MPTs = 162–151 Ma; Callovian, Middle Jurassic to Kimmeridgian, Late Jurassic), and the associated phylogenetic uncertainty extends the BUR_{max} only a few million years before, to the late Early Jurassic (178 Ma; Fig. 3). Reinterpretation of the affinities of *Proterochersis robusta* is critical for palaeontological dating of the evolutionary history of turtles (Fig. 3) and also questions the accuracy of molecular dates that used *Proterochersis robusta* as a calibration constraint for the divergence time of crown-group Testudines. For instance, the position of *Proterochersis robusta* retrieved by Gaffney et al. (1991) was used as a calibration constraint in the molecular clock study of Near et al. (2005) and Alfaro et al. (2009).

Other recent studies obtained ages of Testudines that largely exceed our palaeontological-based estimates (even considering the BUR), even though they relied on other calibration points. Alfaro et al. (2009) calculated an age for Testudines ranging from 225 to 212 Ma (Norian, Triassic). On the other hand, Hugall et al. (2007) inferred the age for Testudines from 259 to 193 Ma (Wuchiapingian, Late Permian to Sinemurian, Early Jurassic) (Table 1, Fig. 3) by using relaxed molecular clocks calibrated with other turtle clades based on the phylogeny of Gaffney et al. (1991). In the latest molecular clock estimates for turtles, Dornburg et al. (2011) calibrated the molecular tree using *Santanachelys gaffneyi* for Cryptodira and *Araripemys barretoi* for Pleurodira (both from the Lower Cretaceous). In their analysis, the divergence time estimated for Testudines ranged from 324.5 to 182.2 Ma (Serpukhovian, Early Carboniferous to Toarcian, Early Jurassic) using lognormal priors and from 314.4 to 169.2 Ma (Bashkirian, Late Carboniferous to Bajocian, Middle Jurassic) using exponential priors.

Irrespective of the selected calibration constraint and the molecular dating method, all published molecular dates are older than minimal palaeontological ages (Fig. 3) when the mean molecular dates and the palaeontological dates from the MPTs are considered, ranging from moderate to large differences (62 Myr; Alfaro et al., 2009; 87 Myr; Dornburg et al., 2011), but is interesting to note that in the latest molecular clock study (Dornburg et al., 2011) the youngest end of the credibility interval for the age of Testudines closely approximates the most ancient end of the BUR of our analysis. Many long-standing discrepancies between molecular clock estimates and palaeontological dates

are dwindling with the addition of more data and methods that incorporate the inherent uncertainties from both sources of information.

Nonetheless, the age of diversification of crown-group Testudines seems to be the most controversial point in the evolutionary history of turtles. Previous disparate results on this topic were influenced by alternative phylogenetic hypotheses stemming from different palae-ontological studies (e.g. the debated position of *Proter-ochersis robusta*), the use of different molecular clock methods, and the selection of different calibration constraints (Fig. 3). Future venues of research to solve this topic should include improving the relatively poor palaeontological sampling of Early to Middle Jurassic turtles, further anatomical studies to robustly determine the controversial position of some critical fossils, as well as increasing the taxon and gene sampling in molecular clock studies. If the time of diversification of crown-group Testudines is to be resolved, it will certainly need progress and new data, both palaeontological and molecular.

Pleurodira. The origin of Pleurodira has not been as intensively discussed in the palaeontological literature as the origin of Cryptodira or Testudines. Previous palae-ontological studies (Gaffney et al., 1991; Joyce, 2007; Danilov and Parham, 2008) suggested Pleurodira orig-inated around 110 Ma (Albian, Early Cretaceous) given the inclusion in this clade of the Early Cretaceous *Araripemys barretoi*. In our analysis the estimated age of Pleurodira of the MPTs ranges between 126 and 101 Ma (Barremian–Albian, Early Cretaceous; Figs 2 and 3, Table 1), but the phylogenetic uncertainty on the position of several fossil taxa greatly extends this range (BUR = 162–1 Ma) due to the instability of several fossil taxa among the bootstrap replicates. In the case of Pleurodira, the long BURs obtained in this study suggest that a more exhaustive sampling of extinct pleurodiran taxa in phylogenetic analyses will be critical to thoroughly test the uncertainties on the time of origin of this clade.

Molecular dates for the time of diversification of Pleurodira have different degrees of congruence with the BUR obtained in this study (Fig. 3). The palaeontolog-ical dates obtained in this study are in agreement with those calculated by Alfaro et al. (2009) and they partially overlap with those calculated by Hugall et al. (2007) and Dornburg et al. (2011) (Fig. 3). Only Near et al. (2005) obtained molecular dates much older than our palaeontological dates (Fig. 3). It must again be underlined that our palaeontological-based minimal estimates are congruent with molecular clock estimates only when comparing the range of ages derived from the phylogenetic uncertainty on the position of fossil taxa (using bootstrap trees to estimate the BUR) and the credibility range of molecular clock estimates (rather

than the mean age estimates; Fig. 3). This suggests, again, that the discrepancies between molecular and palaeontological estimates of divergence dates can be overestimated when the inherent uncertainties of both sources of data are ignored.

Cryptodira. The oldest cryptodiran turtle included in the phylogenies of Gaffney et al. (1991) and Joyce (2007) are from the Early Cretaceous, so their phylogenetic studies suggested Cryptodira originated at least by that time (Fig. 3). Danilov and Parham (2008) were the first to include *Yehguia tatsuensis* in a cladistic analysis, extending the first appearance of this clade at least to the Late Jurassic (Fig. 3). In the present study, the range of minimum ages calculated in MPTs for Cryptodira almost occupy the entire Late Jurassic (from 162 to 146 Ma; latest Callovian, Middle Jurassic to Tithonian, Late Jurassic; Fig. 2, Table 1), but the associated phylogenetic uncertainty on the position of some Juras-sic turtles created a much more extensive BUR (177–101 Ma; Fig. 3), ranging from the late Early Jurassic to the latest Early Cretaceous.

The time of origin of Cryptodira based on molecular data was calculated only by Near et al. (2005) and Alfaro et al. (2009). The range calculated by Alfaro et al. (2009) spans 92 Myr between 212 Ma (Norian, Upper Triassic) and 120 Ma (Aptian, Lower Creta-ceous), partially overlapping with the BUR inferred in this study. The ages calculated by Near et al. (2005) also partially overlap with the BUR (Fig. 3). As in the case of Pleurodira, the palaeontological dates obtained in this study and the available molecular clock estimates on the diversification of Cryptodira are largely congruent but the partial agreement only arises when the uncer-tainty ranges from both sources of data are compared (i.e. credibility intervals of molecular clocks and BUR estimates from palaeontological data); the mean age estimates of both methods still differ by several million years, as expected given that our palaeontological dates are minimal divergence time estimates.

The position of Trionychia and the origin of Cryptodira

A remarkable difference between the molecular and morphological phylogenies is the position of the cryp-todiran clade Trionychia. In molecular phylogenies Trionychia is recovered as the sister clade of remaining cryptodiran turtles (including Kinosternoidea) (e.g. Shaffer et al., 1997; Fujita et al., 2004; Near et al., 2005) whereas in morphological phylogenies, Trionychia is usually recovered as the sister group of Kinosternoi-dea and both are deeply nested within Cryptodira (e.g. Joyce, 2007; Sterli and de la Fuente, 2011). Interestingly, the oldest known cryptodiran turtle is a stem Trionychia (*Yehguia tatsuensis*) from the Late Jurassic of China. If we consider the topology of cryptodiran turtles obtained

from molecular data (Trionychia as the sister group of remaining Cryptodira) to calculate the age of the node Cryptodira, the resulting age is identical to the age calculated in the present analysis because *Yehguia tatsuensis* is the oldest fossil cryptodiran and determines the minimum age of diversification of Cryptodira irrespective of the internal topology of this clade. By contrast, the age of remaining cryptodiran clades (Chelonioidea, Chelydridae, Testudinoidea and Kinosternoidea) is affected if Trionychia is retrieved as the sister clade of remaining groups because in the morphological topology (Fig. 1) there are long ghost lineages (ca. 60 Ma) subtending each of the four mentioned clades. Calibrating the molecular topology with fossils, these four ghost lineages do not exist and there is a single long lineage leading to the clade formed by the four remaining cryptodiran clades (Chelonioidea, Chelydridae, Testudinoidea and Kinosternoidea). We focused this contribution only on three major clades of turtles rather than in the internal relationships of Cryptodira, but the phylogenetic position of Trionychia is an unresolved issue in turtle phylogenetics that should be further studied as it has important evolutionary implications and is also influential in determining the time of origin of the different clades of cryptodiran turtles.

Uncertainties associated with the fossil record

The criteria for selecting calibration constraints and the associated uncertainties are often not discussed at length in molecular dating studies, even though several authors have discussed the risks of overlooking this issue (Conroy and van Tuinen, 2003; Dyke and van Tuinen, 2004; van Tuinen and Hedges, 2004; Reisz and Müller, 2004; Müller and Reisz, 2005; Benton and Donoghue, 2007; Rutschmann et al., 2007; Gandolfo et al., 2008; Ho and Phillips, 2009; Lee et al., 2009; Parham et al., 2012). In particular, the use of artificially precise dates to calibrate molecular dating studies (e.g. Kumar and Hedges, 1998), ignoring the stratigraphic and phylogenetic uncertainties inherent to extinct taxa, can seriously bias the resulting estimates (Shaul and Graur, 2002; Graur and Martin, 2004).

The temporal information of the fossil record is critical for inferring the timing of the evolutionary history of a given group both for palaeontological dating methods and for molecular dating studies. However, the fossil record has an inherently associated uncertainty related to the geochronological uncertainty of FADs and the possibility that the FAD is recovered much later than the origin of the group. Many molecular clock studies, especially older ones, used punctual dates (e.g. temporal midpoint of a stratigraphic stage) from the fossil record instead of using the uncertainty ranges to calibrate the tree of life (e.g. Kumar and

Hedges, 1998; Near et al., 2005; Pyron, 2011). The broad ages obtained in our study for selected nodes of turtle phylogeny highlight the need for considering phylogenetic uncertainty about extinct taxa for obtaining palaeontological dates, which subsequently can be used as calibration constraints for molecular clock estimations.

Extinct taxa also have a variable degree of phylogenetic uncertainty that is in many cases influenced by the limited amount of morphological information preserved (incompleteness of the fossils) and the absence of molecular data on all but the most recent fossils. These uncertainties should not be ignored, as alternative positions of extinct taxa can severely alter the palaeontological dating of nodes of interest (Marjanović and Laurin, 2007; Lee et al., 2009; Pyron, 2011). Our study reveals that the uncertainty on the inferred age of the clades can increase substantially when trees from bootstrap replicates are considered in comparison with the ranges of minimal ages of clades inferred considering the uncertainties in the age of FADs for calibrating MPTs (Pol and Norell, 2006), the possible resolutions of polytomies in supertrees (Marjanović and Laurin, 2007), consensus trees (Boyd et al., 2011), or the use of alternative values of a minimum temporal branch length (Marjanović and Laurin, 2007).

Despite the amount of missing data, the inclusion of all taxa used for calibrating a phylogenetic tree is an essential requirement of the methodology we propose here. It is clear to us that all taxa used for establishing the palaeontological date of a node must be included in a phylogenetic analysis (see also Parham et al., 2012). Whether the oldest taxon of a clade (e.g. oldest chelydrid) is fragmentary is not consider to be problematic. If there are sufficient informative and apomorphic characters that support their affinities, the analysis will result in a narrow uncertainty interval for the palaeontological dating of this clade. If there are not sufficient characters, the hypothesis supporting the chelydrid affinities of the fragmentary taxon will not be particularly strong, and our method should retrieve a larger uncertainty range on the palaeontological dating of this clade. It is obviously recommended to add all relevant characters to the analysis so that the analyses performed actually evaluate all relevant data.

As noted above, recently published papers (Lee et al., 2009; Pyron, 2011) have applied a Bayesian approach to calculate molecular divergence times while considering the phylogenetic uncertainty of extinct taxa, representing an alternative way to deal with this problem. These model-based methods require the integration of molecular and morphological data into a combined dataset. Such an approach should also include a good sample of the extinct taxonomic diversity and the relevant character data, which result in a set of conditions that are rarely met by empirical datasets. Potential problems

discussed above (see Materials and methods) with these approaches include the adequacy of model-based methods for morphological data (in particular for fossils with copious numbers of missing entries).

Conclusions

Here we explore the effect of incorporating phylogenetic uncertainty associated with fossil data for determining the time of diversification of several clades of interest in the evolutionary history of turtles based on a large matrix of morphological data scored across extinct and extant turtles. In particular, we introduce a procedure to incorporate the effect of two types of uncertainties on the estimation of the age of origin of three clades. These are the uncertainty on the age of the FAD of taxa in the fossil record and the uncertainty on the phylogenetic position of extinct and extant taxa based on bootstrap replicates. The inferences on the age of a given node made on bootstrap trees result in BURs, which are often a far more extensive range of ages than those calculated using the FAD on the MPTs. The BURs of major clades of turtles reveal a large degree of previously hidden uncertainty on palaeontological dating of the time of diversification of the major clades of Testudinata, some of which reduce the disagreement with molecular dates, especially when the BURs are compared with the credibility intervals generated by relaxed molecular clocks based on Bayesian analysis. Yet, some disagreements exists, such as the age of crown-group Testudines, and further research and new molecular and palaeontological data will certainly clarify some outstanding issues for understanding the evolutionary history of turtles.

These results emphasize the need to consider the uncertainties associated with the age and phylogenetic position of extinct taxa, not only for deriving palaeontological dates, but also for their use as potential calibration constraints for molecular dating, rather than using arbitrarily chosen point ages (a practice that is fortunately being abandoned) and ignoring phylogenetic uncertainty of fossil taxa (which is still too frequent). The long uncertainty ranges of palaeontological dating obtained in this study (BURs) indicate that the remarkable developments of methods that allow us to incorporate associated uncertainties on multiple parameters of molecular models and dating (e.g. Drummond et al., 2006; Drummond and Rambaut, 2007) should be accompanied by cautious use of the temporal and phylogenetic information from the fossil record to date the many branches of the tree of life, as recently seen in some studies (Lee et al., 2009; Pyron, 2011). We hope that this modest methodological progress will find other applications in the thriving field of time-tree construction (e.g. Hedges and Kumar, 2009).

References

Alfaro, M.E., Santini, F., Brock, C., Alamillo, H., Dornburg, A., Rabosky, D.L., Carnevale, G., Harmong, L.J., 2009. Nine exceptional radiations plus high turnover explain species diversity in jawed vertebrates. Proc. Natl. Acad. Sci. U S A 106, 13410–13414.

Anquetin, J., Barrett, P.M., Jones, M.E.H., Moore-Fay, S., Evans, S.E., 2009. A new stem turtle from the Middle Jurassic of Scotland: new insights into the evolution and palaeoecology of basal turtles. Proc. R. Soc. Edin. B 276, 879–886.

Benton, M.J., 1993. Fossil Record 2. Chapman & Hall, London.

Benton, M.J., Donoghue, P.C.J., 2007. Paleontological evidence to date the tree of life. Mol. Biol. Evol. 24, 26–53.

Boyd, C.A., Cleland, T.P., Marrero, N.L., Clarke, J.A., 2011. Exploring the effects of phylogenetic uncertainty and consensus trees on stratigraphic consistency scores: a new program and a standardized method. Cladistics 27, 52–60.

Braun, E.L., Kimball, R.T., 2002. Examining basal avian divergences with mitochondrial sequences: model complexity taxon sampling, and sequence length. Syst. Biol. 51, 614–624.

Britton, T., 2005. Estimating divergence times in phylogenetic trees without a molecular clock. Syst. Biol. 54, 500–507.

Brochu, C.A., 2004. Patterns of calibration age sensitivity with quartet dating methods. J. Paleontol. 78, 7–30.

Brochu, C.A., 2006. Empirical exploration of calibration sensitivity in molecular dating methods – maximising the role of palaeontology in molecular evolution. In: Yang, Q., Wang, Y., Weldon, E.A.

(Eds.), Ancient Life and Modern Approaches. Abstracts of the Second International Palaeontological Congress. University of Science and Technology of China Press, Beijing, pp. 138.

Bromham, L., Phillips, M.J., Penny, D., 1999. Growing up with dinosaurs: molecular dates and the mammalian radiation. Trends Ecol. Evol. 14, 113–118.

Chang, J.T., 1996. Inconsistency of evolutionary tree topology reconstruction methods when substitution rates vary across characters. Math. Biosci. 134, 189–215.

Conroy, C.J., van Tuinen, M., 2003. Extracting time from phylogenies: positive interplay between fossil and genetic data. J. Mammal. 84, 444–455.

Danilov, I.G., Parham, J.F., 2006. A redescription of 'Plesiochelys' tatsuensis from the Late Jurassic of China, with comments on the antiquity of the crown clade Cryptodira. J. Vertebr. Paleontol. 26, 573–580.

Danilov, I., Parham, J.F., 2008. A reassessment of some poorly known turtles from the Middle Jurassic of China, with comments on the antiquity of extant turtles. J. Vertebr. Paleontol. 28, 306–318.

Donoghue, P.C.J., Benton, M.J., 2007. Rocks and clocks: calibrating the Tree of Life using fossils and molecules. Trends Ecol. Evol. 22, 424–431.

Donoghue, M.J., Doyle, J.A., Gauthier, J.A., Kluge, A.G., Rowe, T., 1989. The importance of fossils in phylogeny reconstruction. Annu. Rev. Ecol. Syst. 20, 431–460.

Dornburg, A., Beaulieu, J.M., Oliver, J.C., Near, T.J., 2011. Integrating fossil preservation biases in the selection of calibrations for molecular divergence time estimation. Syst. Biol. 60, 519–527.

Douzery, E.J.P., Snell, E.A., Bapteste, E., Delsuc, F., Philippe, H., 2004. The timing of eukaryotic evolution: does a relaxed molecular clock reconcile proteins and fossils? Proc. Natl. Acad. Sci. U S A 101, 15386–15391.

Drummond, A.J., Rambaut, A., 2007. BEAST: Bayesian evolutionary analysis by sampling trees. BMC Evol. Biol. 7, 1–8.

Drummond, A.J., Ho, S.Y.W., Phillips, M.J., Rambaut, A., 2006. Relaxed phylogenetics and dating with confidence. PLoS Biol. 4, 699–710.

Dyke, G.J., van Tuinen, M., 2004. The evolutionary radiation of modern birds (Neornithes): reconciling molecules, morphology and the fossil record. Zool. J. Linn. Soc. 141, 153–177.

Felsenstein, J., 1985. Confidence limits on phylogenies: an approach using the bootstrap. Evolution 39, 783–791.

Foote, M., 1997. Estimating taxonomic durations and preservation probability. Paleobiology 23, 278–300.

Foote, M., Raup, D.M., 1996. Fossil preservation and the stratigraphic ranges of taxa. Paleobiology 22, 121–140.

Foote, M., Hunter, J.P., Janis, C.M., Sepkoski, J.J. Jr, 1999. Evolutionary and preservational constraints on origins of biologic groups: divergence times of eutherian mammals. Science 283, 1310–1314.

Fujita, M.K., Engstrom, T.N., Starkey, D.E., Bradley, H., 2004. Turtle phylogeny: insights from a novel nuclear intron. Mol. Phylogenet. Evol. 31, 1031–1040.

Gaffney, E.S., 1996. The postcranial morphology of Meiolania platyceps and a review of the Meiolaniidae. Bull. Am. Mus. Nat. Hist. 229, 1–165.

Gaffney, E.S., Hutchinson, J.H., Jenkins, F.A.J., Meeker, L.J., 1987. Modern turtle origins: the oldest known cryptodire. Science 237, 289–291.

Gaffney, E.S., Meylan, P.A., Wyss, A.R., 1991. A computer assisted analysis of the relationships of the higher categories of turtles. Cladistics 7, 313–335.

Gaffney, E.S., Tong, H., Meylan, P.A., 2006. Evolution of the side-necked turtles: the families Bothremydidae, Euraxemydidae, and Araripemydidae. Bull. Am. Mus. Nat. Hist. 300, 1–698.

Gaffney, E.S., Rich, T.H., Vickers-Rich, P., Constantine, A., Vacca, R., Kool, L., 2007. Chubutemys, a new eucryptodiran turtle from the Early Cretaceous of Argentina, and the relationships of the Meiolaniidae. Am. Mus. Novit. 3599, 1–35.

Gandolfo, M.A., Nixon, K.C., Crepet, W.L., 2008. Selection of fossils for calibration of molecular dating models. Ann. Mo. Bot. Gard. 95, 34–42.

Gandolfo, M.A., Hermsen, E.J., Zamaloa, M.C., Nixon, K.C., González, C.C., Wilf, P., Cúneo, N.R., Johnson, K.R., 2011. Oldest known Eucalyptus macrofossils are from South America. PlosOne 6, e21084. doi:10.1371/journal.pone.0021084.

Gingerich, P.D., 1979. Paleontology, phylogeny, and classification: an example from the mammalian fossil record. Syst. Zool. 28, 451–464.

Goloboff, P.A., Pol, D., 2005. Parsimony and Bayesian phylogenetics. In: Albert, V. (Ed.), Parsimony, Phylogeny, and Genomics. Oxford University Press, Oxford, UK, pp. 148–159.

Goloboff, P.A., Farris, J.S., Källersjö, M., Oxelman, B., Ramírez, M.ı.J., Szumik, C.A., 2003. Improvements to resampling measures of group support. Cladistics 19, 324–332.

Goloboff, P.A., Mattoni, C.I., Quinteros, A.S., 2008a. TNT, a free program for phylogenetic analysis. Cladistics 24, 774–786.

Goloboff, P., Farris, J., Nixon, K., 2008b. TNT: tree search using new technology, vers. 1.1 (Willy Hennig Society Edition). Program and documentation available at http://www.zmuc.dk/public/phylogeny/tnt.

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

Hedges, S.B., Kumar, S., 2009. Discovering the timetree of life. In: Hedges, S.B., Kumar, S. (Eds.), The Timetree of Life. Oxford University Press, New York, pp. 3–18.

Hill, R.V., 2005. Integration of morphological data sets for phylogenetic analysis of Amniota: the importance of integumentary characters and increased taxonomic sampling. Syst. Biol. 54, 530–547.

Ho, S.Y.W., Phillips, M.J., 2009. Accounting for calibration uncertainty in phylogenetic estimation of evolutionary divergence times. Syst. Biol. 58, 367–380.

Hugall, A.F., Foster, R., Lee, M.S.Y., 2007. Calibration choice, rate smoothing, and the pattern of tetrapod diversification according to the long nuclear gene RAG-1. Syst. Biol. 56, 543–563.

Josse, S., Moreau, T., Laurin, M., 2006. Stratigraphic tools for Mesquite, version 1.0. http://mesquiteproject.org/packages/stratigraphicTools/

Joyce, W.G., 2007. Phylogenetic relationships of Mesozoic turtles. Bull. Peabody Mus. Nat. Hist. 48, 3–102.

Joyce, W.G., Parham, J.F., Gauthier, J.A., 2004. Developing a protocol for the conversion of rank-based taxon names to phylogenetically defined clade names, as exemplified by turtles. J. Paleontol. 78, 989–1013.

Kolaczkowski, B., Thornton, J.W., 2004. Performance of maximum parsimony and likelihood phylogenetics when evolution is heterogeneous. Nature 431, 980–984.

Kumar, S., Hedges, B.S., 1998. A molecular timescale for vertebrate evolution. Nature 392, 917–920.

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

Laurin, M., Canoville, A., Quilhac, A., 2009. Use of paleontological and molecular data in supertrees for comparative studies: the example of lissamphibian femoral microanatomy. J. Anat. 215, 110–123.

Lee, M.S.Y., 1995. Historical burden in systematics and the interrelationships of 'parareptiles'. Biol. Rev. 70, 459–547.

Lee, M.S.Y., 1999. Molecular clock calibrations and metazoan divergence dates. J. Mol. Evol. 49, 385–391.

Lee, M.S.Y., 2001. Molecules, morphology, and the monophyly of diapsid reptiles. Contrib. Zool. 70, 1–18.

Lee, M.S.Y., Oliver, P.M., Hutchinson, M.N., 2009. Phylogenetic uncertainty and molecular clock calibrations: a case study of legless

lizards (Pygopodidae, Gekkota). Mol. Phylogenet. Evol. 50, 661–666.

Lu, P.J., Yogo, M., Marshall, C.R., 2006. Phanerozoic marine biodiversity dynamics in light of the incompleteness of the fossil record. Proc. Natl. Acad. Sci. U S A 103, 2736–2739.

Maddison, W.P., Maddison, D.R., 2010. Mesquite: a modular system for evolutionary analysis. Version 2.74. http://mesquiteproject.org.

Marjanović, D., Laurin, M., 2007. Fossils, molecules, divergence times, and the origin of lissamphibians. Syst. Biol. 56, 369–388.

Marjanović, D., Laurin, M., 2008. Assessing confidence intervals for stratigraphic ranges of higher taxa: the case of Lissamphibia. Acta Palaeontol. Pol. 53, 413–432.

Marshall, C.R., 1990. Confidence intervals on stratigraphic ranges. Paleobiology 16, 1–10.

Marshall, C.R., 1994. Confidence intervals on stratigraphic ranges: partial relaxation of the assumption of randomly distributed fossil horizons. Paleobiology 20, 459–469.

Marshall, C.R., 1997. Confidence intervals on stratigraphic ranges with nonrandom distributions of fossil horizons. Paleobiology 23, 165–173.

Marshall, C.R., 2008. A simple method for bracketing absolute divergence times on molecular phylogenies using multiple fossil calibration points. Am. Nat. 171, 726–742.

Müller, J., 2003. Early loss and multiple return of the lower temporal arcade in diapsid reptiles. Naturwissenschaften 90, 473–476.

Müller, J., 2004. The relationships among diapsid reptiles and the influence of taxon selection. In: Arratia, G., Wilson, M.V.H., Cloutier, R. (Eds), Recent Advances in the Origin and Early Radiations of Vertebrates. Dr. Friedrich Pfeil, München, pp. 379–408.

Müller, J., Reisz, R.R., 2005. Four well-constrained calibration points from the vertebrate fossil record for molecular clock estimates. BioEssays 27, 1069–1075.

Near, T.J., Meylan, P.A., Shaffer, H.B., 2005. Assessing concordance of fossil calibration points in molecular clock studies: an example using turtles. Am. Nat. 165, 137–146.

Norell, M.A., Wheeler, W., 2003. Missing entry replacement data analysis: a replacement approach to dealing with missing data in paleontological and total evidence data sets. J. Vertebr. Paleontol. 23, 275–283.

Parham, J.F., Donoghue, P.C.J., Bell, C.J., Calway, T.D., Head, J.J., Holroyd, P.A., Inoue, J.G., Irmis, R., Joyce, W.G., Ksepka, D.T., Patané, J.S.L., Smith, N.D., Tarver, J.E., van Tuinen, M., Yang, Z., Angielczyk, K.D., Greenwood, J.M., Hipsley, C.A., Jacobs, L., Makovicky, P.J., Müller, J., Smith, K.T., Theodor, J.M., Warnock, R.C.M., Benton, M.J., 2012. Best practices for justifying fossil calibrations. Syst. Biol. 61, 346–359.

Pascual, R., Archer, M., Jaureguizar, E.O., Prado, J.L., Godthelp, H., Hand, S.J., 1992. First discovery of monotremes in South America. Nature 356, 704–706.

Paul, C.R.C., 2009. The fidelity of the fossil record: the improbability of preservation. Palaeontology 52, 485–489.

Pol, D., Norell, M.A., 2001. Comments on the Manhattan stratigraphic measure. Cladistics 17, 285–289.

Pol, D., Norell, M.A., 2006. Uncertainty in the age of fossils and the stratigraphic fit to phylogenies. Syst. Biol. 55, 512–521.

Pol, D., Norell, M.A., Siddall, M.E., 2004. Measures of stratigraphic fit to phylogeny and their sensitivity to tree size, tree shape, and scale. Cladistics 20, 64–75.

Pyron, R.A., 2011. Divergence-time estimation using fossils as terminal taxa and the origins of Lissamphibia. Syst. Biol. 60, 466–481.

Rambaut, A., 2006–2009. FigTree v1.3.1. Free software available at http://tree.bio.ed.ac.uk/software/figtree/.

Reisz, R.R., Laurin, M., 1991. *Owenetta* and the origin of turtles. Nature 349, 324–326.

Reisz, R.R., Müller, J., 2004. Molecular timescales and the fossil record: a paleontological perspective. Trends Genet. 20, 237–241.

Rieppel, O., deBraga, M., 1996. Turtles as diapsid reptiles. Nature 384, 453–455.

Rieppel, O., Reisz, R.R., 1999. The origin and early evolution of turtles. Annu. Rev. Ecol. Syst. 30, 1–22.

Rodríguez-Trelles, F., Tarrío, R., Ayala, F.J., 2002. A methodological bias toward overestimation of molecular evolutionary time scales. Proc. Natl. Acad. Sci. U S A 99, 8112–8115.

Roelants, K., Gower, D.J., Wilkinson, M., Loader, S.P., Biju, S.D., Guillaume, K., Moriau, L., Bossuyt, F., 2007. Global patterns of diversification in the history of modern amphibians. Proc. Natl. Acad. Sci. U S A 104, 887–892.

Rougier, G.W., de la Fuente, M.S., Arcucci, A., 1995. Late Triassic turtles from South America. Science 268, 855–858.

Rutschmann, F., Eriksson, T., Salim, K.A., Conti, E., 2007. Assessing calibration uncertainty in molecular dating: the assignment of fossils to alternative calibration points. Syst. Biol. 56, 591–608.

San Mauro, D., Vences, M., Alcobendas, M., Zardoya, R., Meyer, A., 2005. Initial diversification of living amphibians predated the breakup of Pangaea. Am. Nat. 165, 590–599.

Sanderson, M.J., 1997. A nonparametric approach to estimating divergence times in the absence of rate constancy. Mol. Biol. Evol. 14, 1218–1231.

Sanderson, M.J., 2002. Estimating absolute rates of molecular evolution and divergence times: a penalized likelihood approach. Mol. Biol. Evol. 19, 101–109.

Sauquet, H., Ho, S.Y.W., Gandolfo, M.A., Jordan, G.J., Wilf, P., Cantrill, D.J., Bayly, M.J., Bromham, L., Brown, G.K., Carpenter, R.J., Lee, D.M., Murphy, D.J., Sniderman, J.M.K., Udovicic, F., 2012. Testing the impact of calibration on molecular divergence times using a fossil-rich group: the case of *Nothofagus* (Fagales). Syst. Biol. 61, 289–313.

Shaffer, H.B., Meylan, P., McKnight, M.L., 1997. Tests of turtle phylogeny: molecular, morphological, and paleontological approaches. Syst. Biol. 46, 235–268.

Shaul, S., Graur, D., 2002. Playing chicken (*Gallus gallus*): methodological inconsistencies of molecular divergence date estimates due to secondary calibration points. Gene 300, 59–61.

Simmons, M.P., 2011. Misleading results of likelihood-based phylogenetic analyses in the presence of missing data. Cladistics 27, 1–15.

Simmons, M.P., 2012. Radical instability and spurious branch support by likelihood when applied to matrices with non-random distributions of missing data. Mol. Phylogenet. Evol. 62, 472–484.

Solow, A.R., 2003. Estimation of stratigraphic ranges when fossil finds are not randomly distributed. Paleobiology 29, 181–185.

Solow, A.R., Smith, W.K., 1997. On fossil preservation and the stratigraphic ranges of taxa. Paleobiology 23, 271–277.

Sterli, J., 2008. A new, nearly complete stem turtle from the Jurassic of South America with implications for turtle evolution. Biol. Lett. 4, 286–289.

Sterli, J., 2010. Phylogenetic relationships among extinct and extant turtles: the position of Pleurodira and the effects of the fossils on rooting crown-group turtles. Contrib. Zool. 79, 93–106.

Sterli, J., de la Fuente, M.S., 2011. A new turtle from the La Colonia Formation (Campanian–Maastrichtian), Patagonia, Argentina, with remarks on the evolution of the vertebral column in turtles. Palaeontology 54, 63–78.

Strauss, D., Sadler, P.M., 1989. Classical confidence intervals and Bayesian probability estimates for ends of local taxon ranges. Math. Geol. 21, 411–427.

Tavaré, S., Marshall, C.R., Will, O., Soligo, C., Martin, R.D., 2002. Using the fossil record to estimate the age of the last common ancestor of extant primates. Nature 416, 726–729.

Thornton, J.W., Kolaczkowski, B., 2005. No magic pills for phylogenetic error. Trends Genet. 21, 310–311.

van Tuinen, M., Hedges, S.B., 2001. Calibration of avian molecular clocks. Mol. Biol. Evol. 18, 206–213.

van Tuinen, M., Hedges, B.S., 2004. The effect of external and internal fossil calibrations on the avian evolutionary timescale. J. Paleontol. 78, 45–50.

Wagner, P.J., 2000. Likelihood tests of hypothesized durations: determining and accommodating biasing factors. Paleobiology 26, 431–449.

Walker, J.D., Geissman, J.W. (Eds), 2009. Geologic Time Scale. Geological Society of America, Denver, CO.

Welch, J.J., Fontanillas, E., Bromham, L.D., 2005. Molecular dates for the Cambrian explosion: the influence of prior assumptions. Syst. Biol. 54, 672–677.

Zhang, P., Zhou, H., Chen, Y.-Q., Liu, Y.-F., Qu, L.-H., 2005. Mitogenomic perspectives on the origin and phylogeny of living amphibians. Syst. Biol. 54, 391–400.

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

Supporting Information

Additional Supporting Information may be found in the online version of this article:

Data S1. Two files containing: 1, data matrix in Nexus format with time-calibrated consensus trees using 0.1 and 3 Myr as minimum branch lengths; and one block of 150 MPTs calibrated; 2, stratigraphic data extracted from Walker and Geissman (2009).

Data S2. Table with taxon ages (of observed stratigraphic range), references, and the assigned states for DataMSM.tnt. "Stratigraphic distribution" denotes the geological stages where a taxon was found. The lower (defining the maximum age) and the upper (defining the minimum age) bounds of the stratigraphic distribution of a taxon were recovered from Walker and Geissman (2009). "DataMSM age" denotes the age ranges created for tree calibration. The ranges were defined considering the maximum and minimum ages of taxa. For example, all the taxa sharing state G are from the Aptian/Albian (99.6–125 Ma) and all taxa sharing state H are from the Albian (99.6–112 Ma).

Data S3. TNT scripts and MSM data matrix: 1, dataMSM.tnt (matrix with the age states and values for each taxa)—dataMSM.tnt must be modified according to the lower and upper bounds of the stratigraphic distribution of each taxon to calculate the maximum and minimum ages for each node in bootstrap trees; 2, BOOTAGES.RUN (calculates the bootstrap trees and the temporal distribution of inferred ages for selected nodes); 3, MSM.RUN (calculates the stratigraphic fit of phylogenetic trees—MPTs or bootstrap trees) and outputs calibrated phylogenies in which branch length represents time, including a user-defined minimum branch length).

Data S4. File containing consensus tree of 150 MPTs showing Bremer support and bootstrap values (1000 replicates) of clades and list of synapomorphies. Bremer support/bootstrap values shown as percentages. Bootstrap values are shown as GC frequencies. Numbers in bold represent node names.

Data S5. Bootstrap replicates. Left column: BURmax (assuming the maximal age of each fossil, within its uncertainty range); right column: BUR_{min} (assuming the minimal age of each fossil, within its uncertainty range).

Please note: Wiley-Blackwell is not responsible for the content or functionality of any supplementary materials supplied by the authors. Any queries (other than missing material) should be directed to the corresponding author for the article.