Contributions to Zoology, 81 (3) 147-158 (2012)

A large phylogeny of turtles (Testudines) using molecular data

Jean-Michel Guillon[1, 2], Loreleï Guéry[1], Vincent Hulin[1], Marc Girondot[1]
[1] *Laboratoire Ecologie, Systématique et Evolution, Université Paris-Sud, UMR 8079 CNRS, AgroParisTech, F-91405 Orsay, France*
[2] *E-mail: Jean-Michel.Guillon@u-psud.fr*

Key words: mtDNA sequence, nuDNA sequence, phylogenetics, Reptilia, taxonomy

Abstract

Turtles (Testudines) form a monophyletic group with a highly distinctive body plan. The taxonomy and phylogeny of turtles are still under discussion, at least for some clades. Whereas in most previous studies, only a few species or genera were considered, we here use an extensive compilation of DNA sequences from nuclear and mitochondrial genes for more than two thirds of the total number of turtle species to infer a large phylogeny for this taxon. Our results enable us to discuss previous hypotheses on species phylogeny or taxonomy. We are thus able to discriminate between competing hypotheses and to suggest taxonomical modifications. Finally, we pinpoint the remaining ambiguities for this phylogeny and the species for which new sequences should be obtained to improve phylogenetic resolution.

Contents

Introduction

Turtles form a monophyletic group with a highly specialized body plan. Their shell makes them easy to identify and there is no confusion with other vertebrates. On the other hand, debate over turtle phylogeny is vigorous. After Gaffney (1984), who used morphological data to provide the first important work on this topic, many authors focused on lower-ranking taxa and proposed various hypotheses for their relationships. Despite this large volume of work, only during the course of the present study has a large phylogenetic analysis been published (Thomson and Shaffer, 2010). Such an extensive work, including species from all main clades, is useful for studying various problems, such as sex determining mechanisms (Janzen and Krenz, 2004), biogeography (Buhlman *et al*., 2009) or for nomenclature (Joyce *et al*., 2004). Until recently, the use of large datasets for phylogeny reconstruction was hampered by computational limitation. Circumventing this problem, the method of 'supertrees' (Sanderson *et al*., 1998) provided a promising approach to obtain large phylogenies from several smaller ones. But some uncertainties remain about the methods and results obtained (Goloboff and Pol, 2002; Bininda-Emonds, 2004). New methods for inferring phylogenies allow the use of extensive datasets and produce outputs in a reasonable time (*e.g.* Goloboff, 1999; Guindon and Gascuel, 2003).

The DNA sequences available in GenBank are either mitochondrial (mtDNA) or nuclear (nuDNA). MtDNA has been and is still very popular in phylogenetic studies. Indeed, among other advantages compared to nuDNA, mitochondrial sequences lack introns and recombinations, making it relatively easy to align. MtDNA has been shown to evolve more rapidly than nuDNA in many eukaryotic animals, and especially in some turtle species (Caccone *et al*., 2004). On the other hand, the microevolutionary rate of turtle mtDNA may be less rapid than first expected (Avise *et al*., 1992). In any case, the fast evolutionary rate of mtDNA may cause higher levels of homoplasy and thus induce errors in phylogenetic reconstructions. However, the impact of high levels of homoplasy in phylogenetic constructions is still unclear. Some authors have even found, in some cases, a positive correlation between the level of homoplasy and the resolution of the phylogenies (Sanderson and Donoghue, 1996;

Kälersjö *et al*., 1999). Engstrom *et al*. (2004) concluded that the use of mtDNA should still be considered, but authors should observe the following precautions: (i) to use 'better data', *i.e*. data from as large a number of species as possible and/or diversified molecular or morphological data, (ii) to use model-based approaches to calculate phylogenetic trees, such as maximum likelihood or Bayesian analyses.

Whenever authors have focused on higher clades of turtles, they have shown that, despite the fact that these clades are commonly recognized and supported, the phylogenetic relationships between some less inclusive clades typically ranked as super-families, families or sub-families are still debatable (Shaffer *et al*., 1997; Fujita *et al*., 2004; Krenz *et al*., 2005). In some cases, previously erected taxa appear to be paraphyletic or polyphyletic. One good example is the Asian big-headed turtle *Platysternon megacephalum*, the sole member of a monotypic family, Platysternidae, which was thought to be closely related to snapping turtles (Chelydridae; Krenz *et al*., 2005; Parham *et al*., 2006). This hypothesis is now rejected by many authors, who consider *P. megacephalum* a member of Testudinoidea (pond turtles and tortoises) based on phylogenetic results. However, its precise position is still uncertain (Parham *et al*., 2006). Besides, the delimitation of some genera is still under discussion (*e.g*. *Emys*; Fritz *et al*., 2011), both because of phylogenetic controversies and because of limitations inherent in rank-based nomenclature (Laurin, 2010). Genera for which monophyly has been questioned include *Elseya* from Chelidae (Seddon *et al*., 1997; Georges *et al*., 1998), *Trachemys* from Emydidae (Stephens and Wiens, 2003; Spinks and Shaffer, 2009), and *Kachuga* from Geoemydidae (Spinks *et al*., 2004; Le *et al*., 2007; Praschag *et al*., 2007b).

To address these questions of taxonomy, we compiled all turtle mtDNA and best represented nuDNA sequences present in GenBank. Contrary to Thomson and Shaffer (2010), we used all information from mtDNA sequences that could be aligned without ambiguity. In order to limit species sampling effect and long-branch attraction that could perturb phylogeny reconstruction, we used all the available species, even those with few sequences. Then, we used maximum

Table 1. Number of species from each taxon included in this study compared to the total number of species in the considered taxon, and number of species for which we have the complete mitochondrial genome.

	Number of species			Complete mtDNA sequence		
Pleurodira	**38**/79			1		
Chelidae			**25**/52			0
Pelomedusoidea		**13**/27			1	
Podocnemididae			**5**/8			0
Pelomedusidae			**8**/19			1
Cryptodira	**192**/238			29		
Chelydridae			**2**/2			2
Platysternidae			**1**/1			1
Chelonioidea		**7**/7			3	
Cheloniidae			**6**/6			3
Dermochelyidae			**1**/1			0
Kinosternoidea		**5**/26			0	
Dermatemydidae			**1**/1			0
Kinosternidae			**4**/25			0
Testudinoidea		**150**/171			19	
Geoemydidae			**68**/72			7
Emydidae			**39**/51			2
Testudinidae			**43**/48			10
Trionychia		**27**/31			4	
Carettochelyidae			**1**/1			0
Trionychidae			**26**/30			4

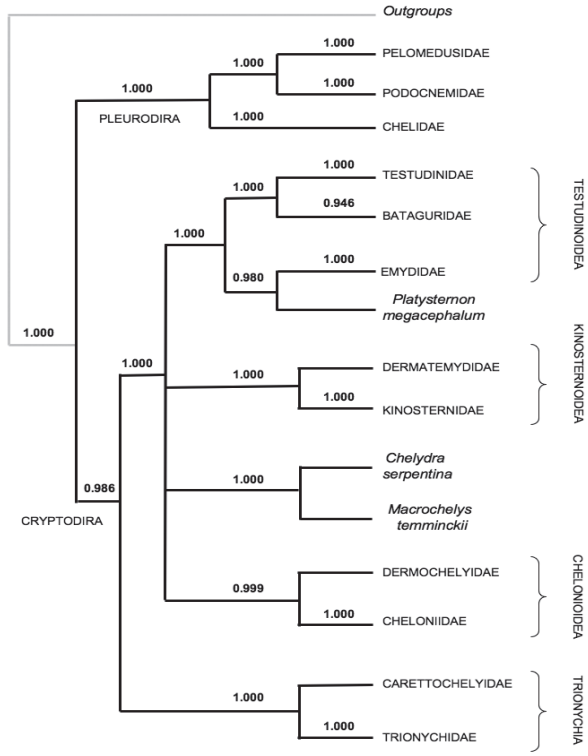

Fig. 1. Phylogenetic relationships between turtle major clades. Confidence values are indicated over each branch. Branches supports are in bold when exceeding 0.9.

likelihood to infer the phylogenetic tree and discuss current debates in turtle evolutionary relationships. The main objectives of this work were multiple: (i) of course, to propose a robust phylogeny for Testudines and help resolving some taxonomic ambiguities, (ii) to favour future work on character evolution for this group, (iii) to pinpoint the remaining ambiguities in the phylogeny to tag the species/groups that need to be sequenced more intensively.

Material and methods

Taxonomic sampling and molecular data

Species sampling was made according to the nomenclature described by Bisby *et al*. (2009). Authorities for each species are indicated in On-line supplementary table S1. We did not include taxa for which there was evidence for a hybrid origin, such as *Mauremys iversoni*, *Mauremys pritchardi*, *Ocadia glyphistoma*, *Ocadia philippeni* and *Sacalia pseudocellata* (Parham *et al*., 2001; Wink *et al*., 2001; Spinks *et al*., 2004; Stuart and Parham, 2007). To obtain a phylogeny as robust as possible, we included both mitochondrial (mtDNA) and nuclear (nuDNA) genes. We compiled all complete mitochondrial sequences available in October 2009 from GenBank (On-line supplementary table S2). We also compiled the sequences of five mitochondrial genes (12S, 16S, COI, NAD4, cytB) and four nuclear genes (R35, c-mos, RAG1 and RAG2). These particular genes were those for which the available number of sequences was highest. We combined the different mtDNA and nuDNA sequences obtained for each species into a single matrix. Only regions of straightforward alignment were taken into account. The length of the final alignment was 20,000 nucleotides (available via: http://purl.org/phylo/treebase/phylows/study/TB2:S12290).

To root the phylogenetic tree, we used total mtDNA for four outgroup species (*Crocodylus porosus* for crocodiles, *Pycnonotus sinensis* for birds, *Lacerta viridis* for squamates, and *Sphenodon punctatus* for Rhynchocephalia). Because nuDNA sequences were not always available for the same taxa, we hereafter

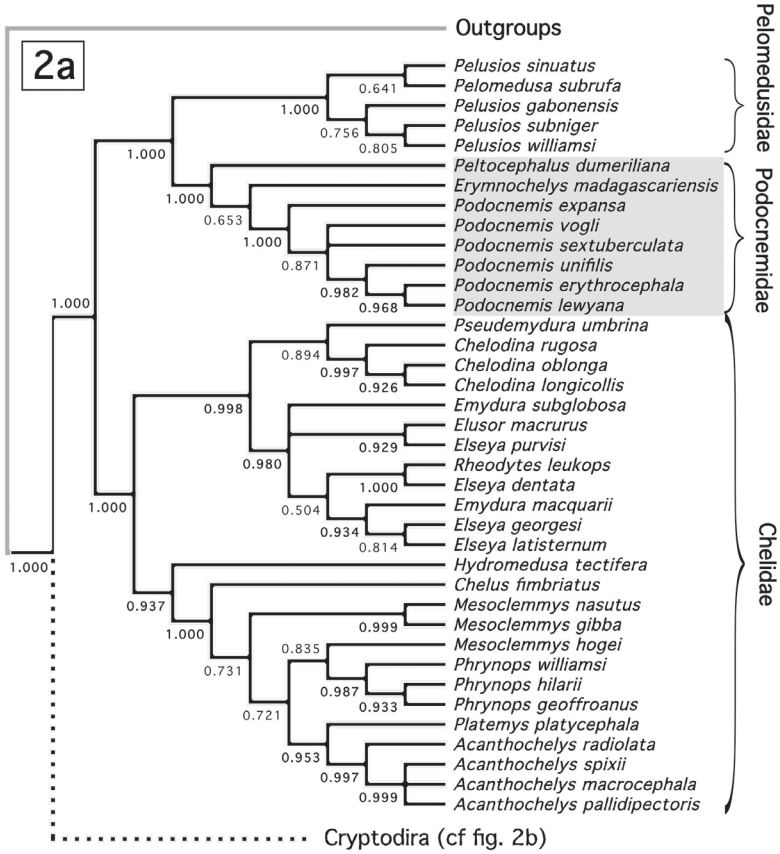

Fig. 2. Maximum likelihood phylogenetic tree obtained using PhyML and the complete DNA dataset (-log(likelihood) = 286215.0). Fitted parameters of the GTR + I + G model are: freqA = 0.36109; freqC = 0.29917; freqG = 0.11484; freqT = 0.22491; A<-> C = 0.91133; A <-> G = 5.21616; A <-> T = 0.94284; C <-> G = 0.50873; C <-> T = 10.42092; G <-> T = 1.00000 ; Proportion of invariant sites = 0.233; Gamma shape parameter = 0.512. Confidence values are indicated under each branch. Branches with confidence value lower than 0.5 have been collapsed. Branch supports are in bold when exceeding 0.9.

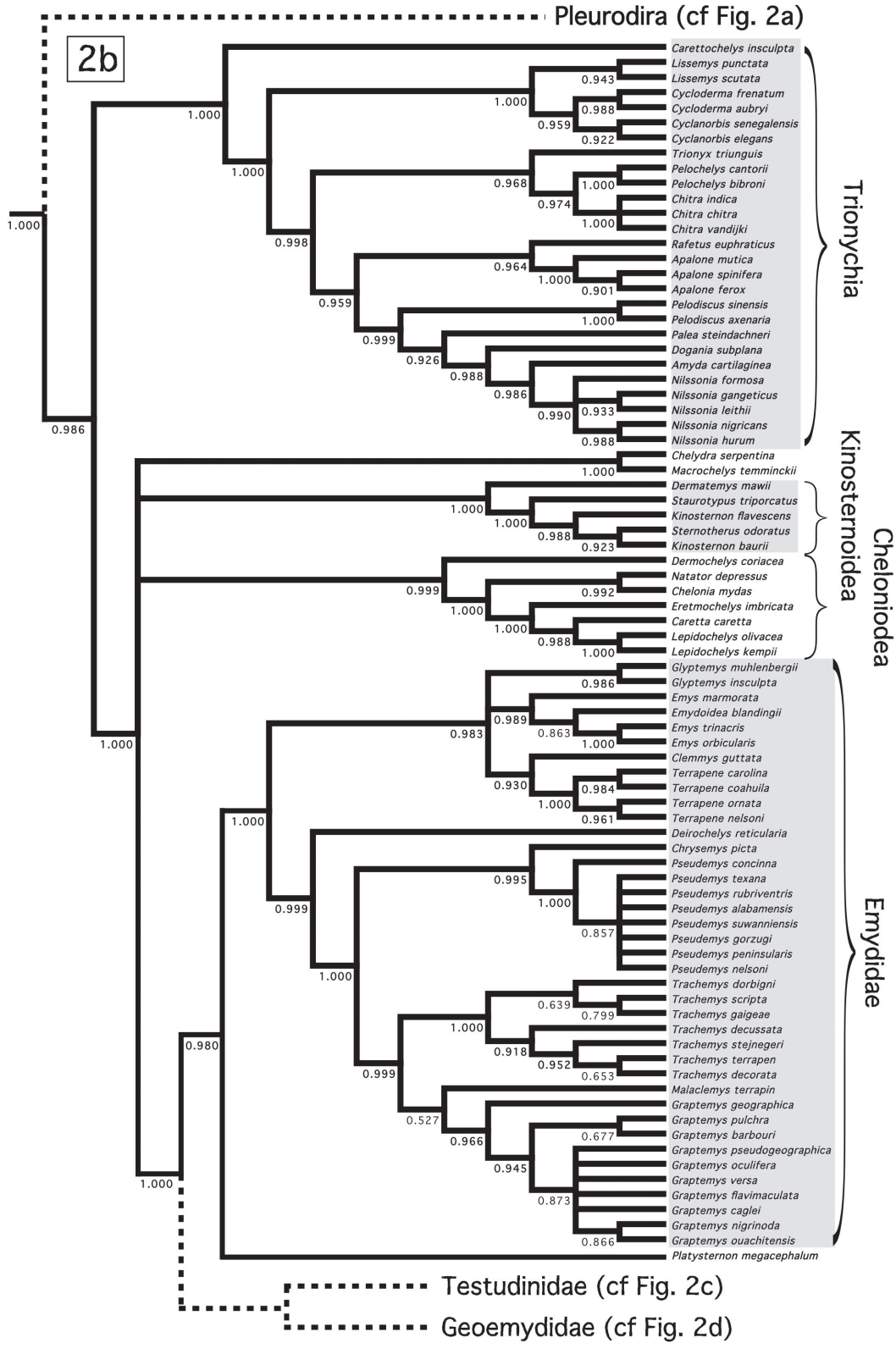

Fig. 2. (continued)

refer to outgroup species as 'crocodiles', 'birds', 'squamates' and 'sphenodons' (references for all outgroup sequences are shown in Table S2). We root the tree either as sister-group of a clade that includes 'birds', 'crocodiles', 'squamates', and 'sphenodons' (Reisz and Laurin, 1991; Laurin and Reisz, 1995; Lee 1997, 2001; Lyson *et al*., 2010) or as sister-group of 'birds' and 'crocodiles' (Zardoya and Meyer 1998, 2001). It does not change the topology within the turtle clade.

Phylogenetic analyses

We used the maximum likelihood algorithm of PhyML (Guindon and Gascuel, 2003) to conduct the phylogenetic analysis, starting with a parsimony tree. Parameterization of PhyML was performed using jModelTest 0.1 (Posada, 2008) to select a model of nucleotide substitution. To quantify branch support, we report confidence values (*cv*) as the result of an approximate likelihood-ratio test performed by PhyML (Anisimova and Gascuel, 2006). Nodes with $cv < 0.5$ have been considered as non-resolved and are polytomized.

Results

Taxonomic sampling and phylogenetic analysis

The complete mitochondrial genome was available in GenBank for 30 turtle species; partial or complete

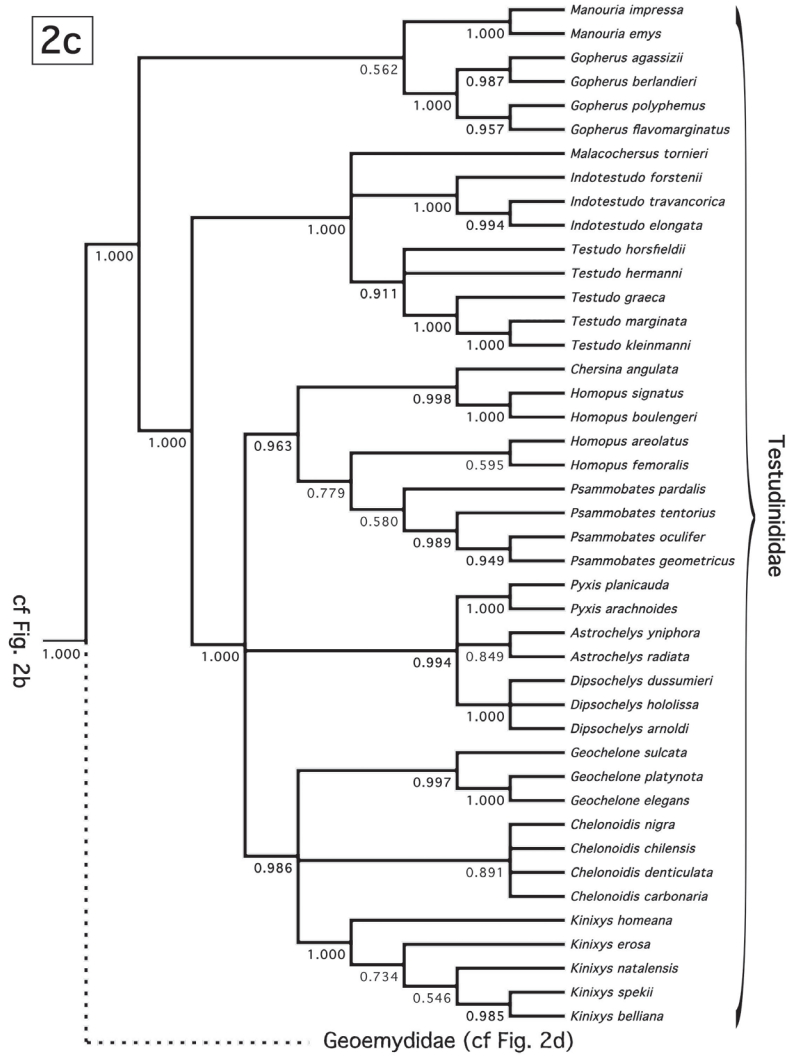

Fig. 2. (continued)

mtDNA sequences were available for 226 species (Table S2). We obtained the sequence of at least one nuDNA gene for 179 species. From a total of 317 extant turtle species (The Reptile Database, 2011), 230 were represented in our phylogeny (Table 1). Amongst the 93 accepted genera, only *Claudius* and *Rhinemys* were not represented.

We ran jModelTest on a restricted alignment, in which only species with complete mtDNA were present. According to the Akaike information criterion, the model of nucleotide substitution selected by jModelTest was a general time reversible model (Lanave *et al*., 1984) allowing for a heterogeneous rate across sites with a 4-categories gamma distribution and for a fitted proportion of sites to be invariable (GTR + I + G). It has been argued that using jModeltest to select the best model of evolution can lead to pitfalls because the program draws on a phenetic BIONJ tree (Marjanović and Laurin, 2007). However, we are rather confident in the model selected, because the second best model was associated with an Akaike weight inferior to 10^{-12}. The numerical outputs of PhyML are presented in the legend to Fig. 2.

Turtle phylogeny

Phylogenetic relationships between major clades of turtles are indicated in Fig. 1. The complete phyloge-

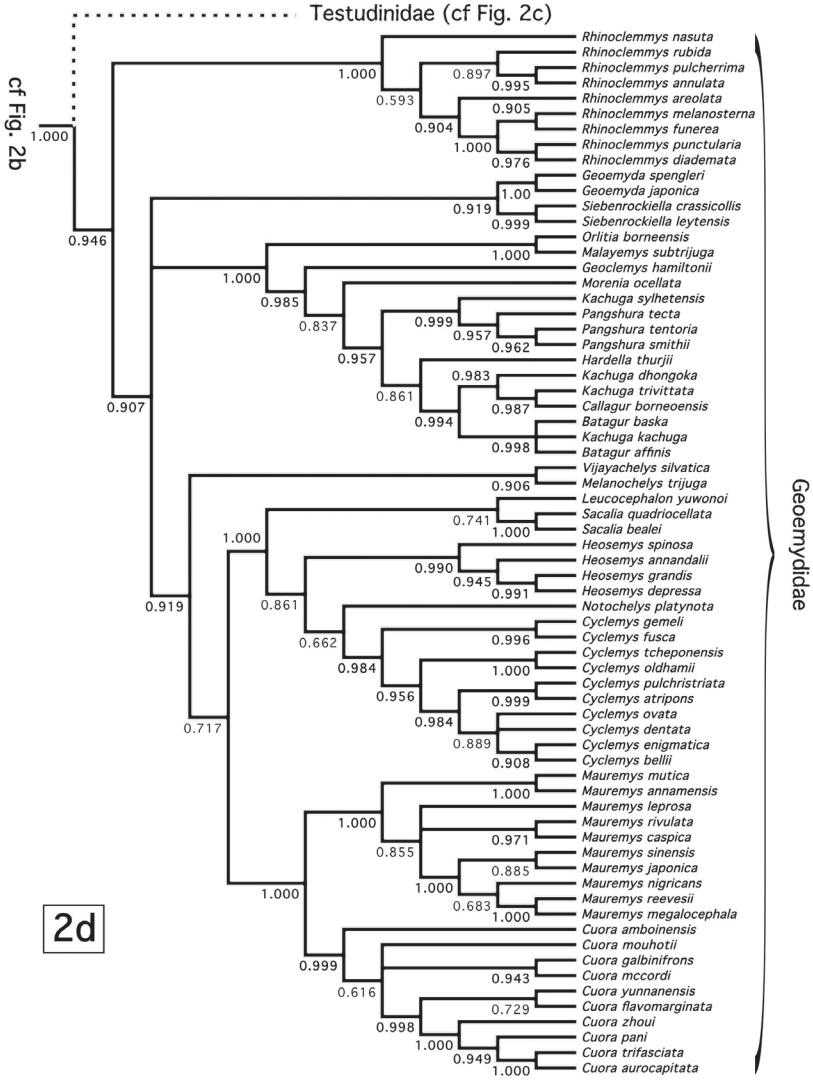

Fig. 2. (continued)

ny with all species is presented in Fig. 2. The distribution of confidence values (*cv*) at nodes is highly skewed in favour of high confidence values (Fig. 3).

Our phylogeny shows a clear separation between monophyletic Pleurodira (*cv* = 1) and Cryptodira (*cv* = 0.986). In Pleurodira (Fig. 2a), Chelidae forms a clade (*cv* = 1), which is the sister-group to the Pelomedusoidea, grouping all Pleurodira except Chelidae (*cv* = 1). The species included in Pelomedusoidea are separated into two monophyletic groups, corresponding to the Podocnemididae (*cv* = 1) and Pelomedusidae (*cv* = 1). We observe a clear separation of Chelidae into three clades, corresponding to Chelidinae (*cv* = 1), Chelodininae (*cv* = 0.998) and Hydromedusinae, although Hydromedusinae is represented in our study by only one species (Table S2). Chelodininae is the sister-group to all other Chelidae, and *Hydromedusa tectifera*, the only species of Hydromedusinae in our phylogeny, is sister to Chelidinae. Four pleurodiran genera are found to be polyphyletic (*Mesoclemmys*, *Elseya*, *Emydura*) or paraphyletic (*Pelusios*). *Mesoclemmys hogei* is grouped with *Phrynops* (*cv* = 0.835), rather than with *M. nasutus* and *M. gibba*. *Elseya dentata* is grouped with *Rheodytes leukops* (*cv* = 1) and *Elseya purvisi* is grouped with *Elusor macrurus* (*cv* = 0.929) rather than with the other sampled species of *Elseya*. *Emydura macquarii* is grouped with *Elseya georgesi* and *Elseya latisternum* (*cv* = 0.934) rather than with *Emydura subglobosa*. *Pelusios sinuatus* is grouped *Pelomedusa subrufa* (*cv* = 0.641) rather than with the other sampled species of *Pelusios*.

Cryptodira is classically organized into five clades (Chelonioidea, Kinosternoidea, Testudinoidea, Trionychia and Chelydroidea, the latter taxon comprising Chelydridae and Platysternidae). Here, Trionychia (*cv* = 1) is sister to the group formed by all other Cryptodira (*cv* = 1). The only species from Carettochelyidae, *Carettochelys insculpta*, is separated from a group including all other Trionychia (*cv* = 1). The monophyly of Trionychinae (*cv* = 0.998) and Cyclanorbinae (*cv* = 1) is also well supported (Table S2, Fig. 2b). All species from the same genus are grouped together.

Trionychia is recovered as the sister group to all remaining clades of cryptodiran turtles which form a tetrapolytomy (Fig. 2b) including: (i) Chelonioidea (*cv* = 0.999), (ii) Chelydridae (*cv* = 1), (iii) Kinosternoidea (*cv* = 1), and (iv) a group formed by Geoemydidae, Testudinidae, Emydidae and *Platysternon megacephalum* (*cv* = 1).

In Chelonioidea, there is a clear separation between Cheloniidae (*cv* = 1) and Dermochelyidae (Table S2, Fig. 2b), as between the two families included in Kinosternoidea, the monotypic Dermatemydidae (*Dermatemys mawii*) and Kinosternidae (*cv* = 1). However, *Kinosternon* appears paraphyletic, as *Kinosternon baurii* is closer to *Sternotherus odoratus* (*cv* = 0.923) than to *K. flavescens*. The fourth clade is composed of a group formed by Testudinidae and Geoemydidae (*cv* = 1), and a group formed by Emydidae and *Platysternon megacephalum* (*cv* = 0.980). According to the usual taxonomy, Testudinidae, Geoemydidae and Emydidae together form the clade Testudinoidea. This clade is then paraphyletic in our phylogeny due to the inclusion of *Platysternon megacephalon* as the sister-group to Emydidae (Fig. 2b).

Within the monophyletic Emydidae (*cv* = 1), Emydinae and Deirochelyinae are both monophyletic (*cv* = 0.983 and *cv* = 0.999, respectively; Table S2, Fig. 2b). All species from the same genus are grouped together except for *Emys*: *Emys orbicularis* and *Emys trinacris* are closer to *Emydoidea blandingii* (*cv* = 0.863) than to *Emys marmorata*.

Within the monophyletic Testudinidae (*cv* = 1), Gopherinae (*cv* = 0.562) and Testudininae (*cv* = 1) are monophyletic (Table S2, Fig. 2c). All species from the same genus are grouped together, except for *Homopus*. *Homopus areolatus* and *Homopus femoralis* are close to *Psammobates* (*cv* = 0.779), whereas *Homopus boulengeri* and *Homopus signatus* are grouped with *Chersina angulata* (*cv* = 0.998).

Within the monophyletic Geoemydidae (*cv* = 0.946), Batagurinae is monophyletic (*cv* = 0.985) and nested within Geoemydinae, which is thus paraphyletic (Table S2, Fig. 2d). All species from the same genus are grouped together except for *Batagur* and *Kachuga*. *Batagur affinis* and *Batagur baska* are recovered in a polytomy including *Kachuga kachuga* (*cv* = 0.998), whereas *Kachuga trivittata* is grouped with *Callagur borneoensis* (*cv* = 0.987) and *Kachuga sylhetensis* is grouped with *Pangshura* (*cv* = 0.999).

Discussion

As a note of caution, it must be recalled that the present study makes a compilation of GenBank sequences, sequences that may not be devoid of errors. Problems may arise from taxonomic misidentification (Vilgalys, 2003; Stuart and Fritz, 2008; Fritz *et al*., 2010), sequencing errors (Harris, 2003), and pseudogene amplification (Fritz *et al*., 2010). We did not try to remove rogue taxa (Sanderson and Shaffer, 2002) from the

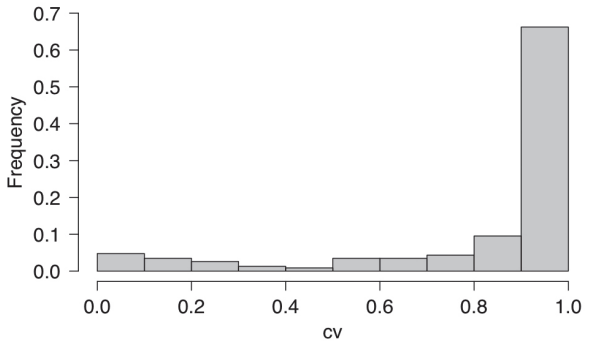

Fig. 3. Distribution of confidence values at nodes.

analysis, because the maximum likelihood method we employed for phylogenetic inference, and the resulting confidence values do not make use of bootstrapping (Guindon and Gascuel, 2003; Anisimova and Gascuel, 2006). Only highly supported branches ($cv > 0.9$) will be discussed here.

Many studies assumed *a priori* the monophyly of Pleurodira and Cryptodira, and rooted the Testudines tree at the branch joining these two groups (Shaffer *et al*., 1997; Fujita *et al*., 2004; Thomson and Shaffer, 2010). Studies that used one or two outgroups (*Gallus* and/or *Alligator*) found ambiguous results concerning the monophyly of Pleurodira and Cryptodira, depending on the method used for phylogenetic inference (Cervelli *et al*., 2003; Krenz *et al*., 2005; Barley *et al*., 2010). Sterli (2010) recently recovered a sister relationship between Pleurodira and Trionychia, based on morphological characters and 12S, 16S, cytb, RAG1 and R35 intron. This unorthodox result may be due to the inclusion of extinct species in the analysis of morphological characters, resulting in the basal position of Chelonioidea as sister to all other extant Testudines lineages. Using outgroups from four clades (Squamata, Rhynchocephalia, Aves and Crocodylidae), we find here good support for a basic divergence of Pleurodira and Cryptodira. The early appearance of Trionychia within Cryptodira is also better supported than in previous molecular studies with outgroups (Krenz *et al*., 2005; Barley *et al*., 2010).

The phylogeny of Cryptodira has been a matter of debate (Shaffer *et al*., 1997; Fujita *et al*., 2004; Krenz *et al*., 2005; Chandler and Janzen, 2009; Thomson and Shaffer, 2010). We find a polytomy composed of four clades: (i) Chelonioidea ($cv = 0.999$), (ii) Chelydridae ($cv = 1$), (iii) Kinosternoidea ($cv = 1$), and (iv) a group formed by Geoemydidae, Testudinidae, Emydidae and *Platysternon megacephalum* ($cv = 1$). Various studies have found different topologies, either grouping Chelonioidea with Kinosternoidea (Fujita *et al*., 2004; using R35 intron), or grouping Chelonioidea with Geoemydidae, Testudinidae, Emydidae and *Platysternon megacephalum* (Parham *et al*., 2006; using complete mtDNA). As the number of species from Kinosternoidea included in phylogenies is usually low (here only five out of 26 species), DNA sequencing of more species from this clade could help clarify these relationships. However, the reasons why certain deep nodes are difficult to resolve are probably twofold: (i) a relatively rapid radiation, and (ii) phylogenetic analyses using genes saturated with mutations. Recently, Barley *et al*. (2010) provided new sequence data for eight nuclear genes, and found good support for a sister relationship of Chelonioidea to a group formed by Chelydridae and Kinosternoidea. We find the same relationships ($cv > 0.9$) after adding the sequences of Barley *et al*. (2010) to our alignment (results not shown).

The phylogenetic position of *Platysternon megacephalum* has long been enigmatic (Parham *et al*., 2006). Indeed, the first analyses, based on morphological characters, have led authors to relate *P. megacephalum* with Chelydridae (*Chelydra serpentina* and *Macrochelys temminckii*). However, some molecular studies found that *P. megacephalum* was grouped with Testudinoidea (Cervelli *et al*., 2003, based on U17 small nucleolar RNA; Krenz *et al*., 2005, based on 12S, cytB and RAG1). In the phylogeny that Parham *et al*. (2006) obtained with complete mitochondrial genomes, *P. megacephalum* was included in Testudinoidea as the sister-species to Emydidae. We here find good support ($cv = 0.980$) for the same hypothesis, as did Thomson and Shaffer (2010).

Within Pleurodira, a problematic taxon is *Elseya*, which appears polyphyletic, with the inclusion of *Rheodytes leukops* as sister to *Elseya dentata* ($cv = 1$), and of *Elusor macrurus* as sister to *Elseya purvisi* ($cv = 0.929$). The same position, albeit less supported, was already found for *Rheodytes leukops* by Seddon *et al*. (1997), based on 12S sequences. Thomson and Georges (2009) recently proposed the erection of a new genus, *Myuchelys*, including species previously named *Elseya latisternum* (type-species of the new genus), *Elseya georgesi*, *Elseya purvisi* and *Elseya novaeguineae*. More sequencing work would be useful to assess the monophyly of the resulting taxon, because it is not monophyletic in our phylogeny. Finally, *Emydura* is problematic, because *Emydura macquarii* is grouped

with *Elseya georgesi* and *Elseya latisternum* (*cv* = 0.934), and not with *Emydura subglobosa*. However, it must be noted that *Emydura subglobosa* is here represented by only one sequence. The topology we find for Podocnemididae is in agreement with the one recently reported by Vargas-Ramirez *et al*. (2008), based on 12 genes (six mitochondrial and six nuclear).

Within Trionychia, our results are comparable with those of Engstrom *et al*. (2004), based on ND4, cytB and R35, except within *Nilssonia*. We find high support for a sister relationship between *Nilssonia hurum* and *Nilssonia nigricans* (*cv* = 0.988), and between *Nilssonia gangeticus* and *Nilssonia leithii* (*cv* = 0.933). This is in agreement with the results reported by Praschag *et al*. (2007a), based on cytB sequences. Within Kinosternoidea, Thomson and Shaffer (2010) found good support for a monophyletic *Kinosternon*. In contrast, we find a sister relationship between *Kinosternon baurii* and *Sternotherus odoratus* (*cv* = 0.923). However, only one DNA sequence is available for *Kinosternon baurii*, so this result is only preliminary and warrants further investigation. The phylogeny we find within Chelonioidea is identical to that already described (Naro-Maciel *et al*., 2008). Within Testudinidae, we find a sister relationship between *Kinixys belliana* and *Kinixys spekii* (*cv* = 0.985), whereas Thomson and Shaffer (2010) found good support for a close relationship between *Kinixys belliana* and *Kinixys natalensis*. The reason for this difference is unclear because gene sampling for *Kinixys* was presumably very similar in both studies. There is also significant disagreement between our phylogeny of Testudininae and the one by Lourenço *et al*. (2012), who found *Indotestudo* and *Malacochersus* nested within *Testudo*. This may result from different species sampling because Lourenço *et al*. (2012) did not include *Testudo hermanni* and *Indotestudo travancorica* in their analysis. This is not the first time that *Homopus* is found to be polyphyletic. Thomson and Shaffer (2010) found the same relationships for *Homopus*, showing *Homopus signatus* and *Homopus boulengeri* grouped with *Chersina angulata* with good support, and *Homopus aerolatus* and *Homopus femoralis* grouped with *Psammobates* with lower support. We suggest that a taxonomic revision could be useful here.

A taxon that has been plagued by nomenclatural problems is Geoemydidae. Based on molecular phylogenies, Spinks *et al*. (2004) made three nomenclatural suggestions. They proposed (i) including all *Chinemys* and *Ocadia* species in *Mauremys*, (ii) re-including *Chelopus annulata* and *Chelopus rubida* in the genus *Rhinoclemmys*, and (iii) classifying *K. tecta*, *K. tentoria* and *K. smithii* as members of a new genus, *Pangshura*. All these propositions are supported by our phylogeny and result in monophyletic genera. However, the taxonomic position of the remaining species of *Kachuga* (*K. dhongoka*, *K. kachuga*, *K. sylhetensis* and *K. trivittata*) is still problematic. As Le *et al*. (2007) and Praschag *et al*. (2007a) already recommended, we propose the inclusion of *Kachuga sylhetensis* in *Pangshura*, with which it is clearly grouped (*cv* = 0.999), and the inclusion of *Callagur borneoensis*, *Kachuga kachuga*, *Kachuga dhongoka*, and *Kachuga trivittata* in *Batagur*. Within Geoemydinae, the taxonomy within *Cuora* and *Cyclemys* has long been uncertain. Here we find *Cuora trifasciata* grouped with *Cuora aurocapitata* (*cv* = 1) and *Cuora galbinifrons* grouped with *Cuora mccordi* (*cv* = 0.943), in contrast with Honda *et al*. (2002), using 12S and 16S sequences, or Stuart and Parham (2004) and He *et al*. (2007), using COI and ND4. The phylogeny we obtain for *Cuora* is consistent with the one obtained by Spinks and Shaffer (2007), using COI and ND4 (but not what they obtained with nuclear DNA). Noticeably, we here used complete mtDNA for *Cuora aurocapitata*, *Cuora flavomarginata* and *Cuora mouhotii*. Uncertainty of specimen identification or hybridization between species, as is known to occur in *Cuora*, may also explain these discrepancies (Spinks and Shaffer, 2007). Within *Cyclemys*, we find close relationships between a group formed by *Cyclemys atripons* and *Cyclemys pulchristriata*, and a group formed by *Cyclemys bellii*, *Cyclemys enigmatica*, *Cyclemys dentata* and *Cyclemys ovata*. This contrasts with the topologies found by Praschag *et al*. (2009) and Fritz *et al*. (2008), using cytB, c-mos, RAG2 and R35 intron. However, these studies obtained different topologies when analyzing separately mitochondrial and nuclear sequences, indicating that mitochondrial introgression may have occurred through hybridization. Finally, we suggest that the Batagurinae and Geoemydinae should be re-delimited, since Geoemydinae is currently not monophyletic (Table S2, Fig. 2d).

Within Emydidae, *Clemmys guttata* has been consistently recovered in two different positions: (i) as sister to *Terrapene* (Feldman and Parham, 2002, based on cytB and ND4; Stephens and Wiens, 2003, based on 16S, ND4 and cytB; Spinks and Shaffer, 2009, based on cytB, R35, RAG1 plus five more nuclear loci) or (ii) as sister to the *Emys* + *Emydoidea* clade (Stephens and Wiens, 2009, based on morphological

data, 16S, cytB, ND4, control region, and R35; Wiens *et al*., 2010, based on cytB, ND4, R35 plus five more nuclear loci). Our results support the first hypothesis (*cv* = 0.930). However, the second hypothesis for *Clemmys guttata* was favored by data sets that were not totally included in ours, so we cannot reject it. Our results on Deirochelyinae are in good agreement with previous topologies (Stephens and Wiens, 2003, 2009; Wiens *et al*., 2010), but differ from those reported by Spinks *et al*. (2009), using seven nuclear loci including R35 and RAG1. We find *Trachemys* monophyletic (*cv* = 1), grouped with *Graptemys* and *Malaclemmys terrapin* (*cv* = 0.999), and *Pseudemys* grouped with *Chrysemys picta* (*cv* = 0.995). Within *Graptemys* and *Pseudemys*, relationships are poorly resolved because mitochondrial DNA seems to exhibit low divergence between species (Wiens *et al*., 2010), and we have excluded fast evolving DNA regions from our 230-species alignment.

A parsimony analysis of the DNA matrix was also performed, and yielded a less resolved tree, with low bootstrap support for some clades that were well resolved in maximum likelihood analysis (results not shown). However, the monophyly of Pleurodira and Cryptodira is supported (bootstrap support = 0.81 and 0.76, respectively), as is the sister relationship of Trionychia (bootstrap support = 0.85) to a group formed by all other Cryptodira (bootstrap support = 0.87). When only bootstrap support > 0.7 is taken into account, the parsimony tree is fully compatible with the one obtained with maximum likelihood.

Our study has provided the largest phylogeny of turtles to date. By using both mtDNA and nuDNA data, we find that most genera are now monophyletic, with strong support, but we suggest some nomenclatural revisions and point at specific taxa that warrant further sequencing work. Polytomies still observed in our phylogeny (*cv* < 0.5) are related with the species with the lowest number of sequences. On the 21 species with only one gene sequenced, one third is directly involved in a polytomy (this proportion is 0.11 when considering all species). New sequences from these seven species (*Cyclemys ovata*, *Emydura subglosa*, *Graptemys oculifera*, *Graptemys versa*, *Pseudemys alabamensis*, *Pseudemys gorzugi*, *Pseudemys suwanniensis*) should be obtained in priority to better resolve the phylogeny. Because all but two turtle genera are represented in our phylogeny, our work provides a solid basis to help in further studies of the evolution of some characters in turtles or the ancient biogeographical distribution of turtles.

Acknowledgements

Authors would like to thank Didier Casane, Marion Hulin, Michel Laurin and Juliana Sterli for valuable comments. This work was financed by regular funding from the Centre National de la Recherche Scientifique, the Université Paris-Sud, and AgroParisTech.

References

Anisimova M, Gascuel O. 2006. Approximate likelihood-ratio test for branches: a fast, accurate, and powerful alternative. *Systematic Biology* 55: 539-552.

Avise JC, Bowen BW, Lamb T, Meylan AB, Bermingham E. 1992. Mitochondrial DNA evolution at a turtle's pace: evidence for low genetic variability and reduced microevolutionary rate in the Testudines. *Molecular Biology and Evolution* 9: 457-473.

Barley AJ, Spinks PQ, Thomson RC, Schaffer HB. 2010. Fourteen nuclear genes provide phylogenetic resolution for difficult nodes in the turtle tree of life. *Molecular Phylogenetics and Evolution* 55: 1189-1194.

Bininda-Emonds ORP. 2004. The evolution of supertrees. *Trends in Ecology and Evolution* 19: 315-322.

Bisby FA, Roskov YR, Orrell TM, Nicolson D, Paglinawan LE, Bailly N, Kirk, PM, Bourgoin T, Baillargeon G. 2009. Species 2000 & ITIS Catalogue of Life: 2009 Annual Checklist. Available via www.catalogueoflife.org/annual-checklist/2009/ [visited on 15 Oct 2009].

Buhlmann KA, Akre TSB, Iverson JB, Karapatakis D, Mittermeier RA, Georges A, Rhodin AGJ, van Dijk PP, Gibbons JW. 2009. A global analysis of tortoise and freshwater turtle distributions and identification of priority conservation areas. *Chelonian Consevation Biology* 8: 116-149.

Caccone A, Gentile G, Burns CE, Sezzi E, Bergman W, Ruelle M, Saltonstall K, Powell JR. 2004. Extreme difference in rate of mitochondrial and nuclear DNA evolution in a large ectotherm, Galapagos tortoises. *Molecular Phylogenetics and Evolution* 31: 794-798.

Cervelli M, Oliverio M, Bellini A, Bologna M, Cecconi F, Mariottini P. 2003. Structural and sequence evolution of U17 small nucleolar RNA (snoRNA) and its phylogenetic congruence in chelonians. *Journal of Molecular Evolution* 57: 73-84.

Chandler CH, Janzen FJ. 2009. The phylogenetic position of the snapping turtles (Chelydridae) based on nucleotide sequence data. *Copeia* 2009: 209-213.

Engstrom TN, Shaffer HB, McCord WP. 2004. Multiple data sets, high homoplasy, and the phylogeny of softshell turtles (*Testudines : Trionychidae*). *Systematic Biology* 53: 693-710.

Feldman CR, Parham JF. 2002. Molecular phylogenetics of emydine turtles: taxonomic revision and the evolution of shell kinesis. *Molecular Phylogenetics and Evolution* 22: 388-398.

Fritz U, Guicking D, Auer M, Sommer RS, Wink M, Hundsdörfer, AK. 2008. Diversity of the Southeast Asian leaf turtle genus *Cyclemys*: how many leaves on its tree of life? *Zoologica Scripta* 37: 367-390.

Fritz U, Daniels SR, Hofmeyr MD, González J, Barrio-Amorós CL, Siroky P, Hundsdörfer AK, Stuckas H. 2010.

Mitochondrial phylogeography and subspecies of the wide-ranging sub-Saharan leopard tortoise *Stigmochelys pardalis* (Testudines: Testudinidae) - a case study for the pitfalls of pseudogenes and GenBank sequences. *Journal of Zoological Systematics and Evolutionary Research* 48: 348-359.

Fritz U, Schmidt C, Ernst CH. 2011. Competing generic concepts for Blanding's, Pacific and European pond turtles (*Emydoidea*, *Actinemys* and *Emys*) - Which is best? *Zootaxa* 2791: 41-53.

Fujita MK, Engstrom TN, Starkey DE, Shaffer HB. 2004. Turtle phylogeny: insights from a novel nuclear intron. *Molecular Phylogenetics and Evolution* 31: 1031-1040.

Gaffney ES. 1984. Historical analysis of theories of chelonian relationship. *Systematic Zoology* 33: 283-301.

Georges A, Birrell J, Saint KM, McCord WP, Donnellan S. 1998. A phylogeny for side-necked turtles (Chelonia: Pleurodira) based on mitochondrial and nuclear gene sequence variation. *Biological Journal of the Linnean Society* 67: 213-246.

Goloboff PA. 1999. Analysing large data sets in reasonable times: solutions for composite optima. *Cladistics* 15, 415-428.

Goloboff PA, Pol D. 2002. Semi-strict supertrees. *Cladistics* 18: 514-525.

Guindon S, Gascuel O. 2003. A simple, fast and accurate algorithm to estimate large phylogenies by maximum likelihood. *Systematic Biology* 52: 696-704.

Harris DJ. 2003 Can you bank on GenBank? *Trends in Ecology and Evolution* 18: 317-319.

He J, Zhou T, Rao D, Zhang Y. 2007. Molecular identification and phylogenetic position of *Cuora yunnanensis*. *Chinese Science Bulletin* 52: 3305-3309.

Honda M, Yasukawa Y, Hirayama R, Ota H. 2002. Phylogenetic relationships of the Asian box turtles of the genus *Cuora* sensu lato (Reptilia: Bataguridae) inferred from mitochondrial DNA sequences. *Zoological Science* 17: 1305-1312.

Janzen FJ, Krenz JG. 2004. Phylogenetics: which was first, TSD or GSD? Pp. 121-130 in: Valenzuela N, Lance VA, eds, Temperature-dependent sex determination in vertebrates. Washington: Smithsonian Books.

Joyce WG, Parham JF, Gauthier JA. 2004. Developing a protocol for the conversion of rank-based taxon names to phylogenetically defined clade names, as exemplified by turtles. *Journal of Paleontology* 78: 989-1013.

Kälersjö M, Albert VA, Farris JS. 1999. Homoplasy increases phylogenetic structure. *Cladistics* 15: 91-93.

Krenz JG, Naylor GJP, Shaffer HB, Janzen FJ. 2005. Molecular phylogenetics and evolution of turtles. *Molecular Phylogenetics and Evolution* 37: 178-191.

Lanave C, Preparata G, Saccone C, Serio G. 1984. A new method for calculating evolutionary substitution rates. *Journal of Molecular Evolution* 20: 86-93.

Laurin M. 2010. The subjective nature of Linnean categories and its impact in evolutionary biology and biodiversity studies. *Contributions to Zoology* 79: 131-146.

Laurin M, Reisz RR. 1995. A reevaluation of early amniote phylogeny. *Zoological Journal of the Linnean Society* 113: 165-223.

Le M, McCord WP, Iverson JB. 2007. On the paraphyly of the genus *Kachuga* (Testudines: Geoemydidae). *Molecular Phylogenetics and Evolution* 45: 398-404.

Lee MSY. 1997. Reptile relationships turn turtles. *Nature* 389: 245-246.

Lee MSY. 2001. Molecules, morphology, and the monophyly of diapsid reptiles. *Contributions to Zoology* 70: 1-22.

Lourenço JM, Claude J, Galtier N, Chiari Y. 2012. Dating cryptodiran nodes: origin and diversification of the turtle superfamily Testudinoidea. *Molecular Phylogenetics and Evolution* 62: 496-507.

Lyson TR, Bever GS, Bhullar B-AS, Joyce WG, Gauthier JA. 2010. Transitional fossils and the origin of turtles. *Biology Letters* 6: 830–833.

Marjanović D, Laurin, M. 2007. Fossils, molecules, divergence times, and the origin of Lissamphibians. *Systematic Biology* 56: 369-388.

Naro-Maciel E, Le M, FizSimmons NN, Amato G. 2008. Evolutionary relationships of marine turtles: a molecular phylogeny based on nuclear and mitochondrial genes. *Molecular Phylogenetics and Evolution* 49: 659-662.

Parham JF, Simison BW, Kozak KH, Feldman CR, Shi H. 2001. New Chinese turtles: endangered or invalid. A reassessment of two species using mitochondrial DNA, allozyme electrophoresis and known-locality specimens. *Animal Conservation* 4: 357-367.

Parham JF, Feldman CR, Boore, JL. 2006. The complete mitochondrial genome of the enigmatic bigheaded turtle (*Platysternon*): description of unusual genomic features and the reconciliation of phylogenetic hypotheses based on mitochondrial and nuclear DNA. *BMC Evolutionary Biology* 6: 11.

Posada D. 2008. jModelTest: Phylogenetic model averaging. *Molecular Biology and Evolution* 25: 1253-1256.

Praschag P, Hundsdörfer AK, Fritz U. 2007a. Phylogeny and taxonomy of endangered South and South-east Asian freshwater turtles elucidated by mtDNA sequence variation (Testudines: Geoemydidae: *Batagur*, *Callagur*, *Hardella*, *Kachuga*, *Pangshura*). *Zoologica Scripta* 36: 429-442.

Praschag P, Hundsdörfer AK, Reza AHM, Fritz U. 2007b. Genetic evidence for wild-living *Aspideretes nigricans* and a molecular phylogeny of South Asian softshell turtles (Reptilia: Trionychidae: *Aspideretes*, *Nilssonia*). *Zoologica Scripta* 36: 301-310.

Praschag P, Hundsdörfer AK, Fritz U. 2009. Further specimens and phylogenetic position of the recently described leaf turtle species *Cyclemys gemeli* (Testudines: Geoemydidae). *Zootaxa* 2008: 29-37.

Reisz RR, Laurin M. 1991. *Owenetta* and the origin of turtles. *Nature* 349: 324-326.

Sanderson MJ, Donoghue MJ. 1996. The relationship between homoplasy and confidence in a phylogenetic tree. Pp. 67-89 in: Sanderson MJ, Hufford L, eds, Homoplasy: The recurrence of similarity in evolution. San Diego: Academic Press.

Sanderson MJ, Shaffer BH. 2002. Troubleshooting molecular phylogenetics analyses. *Annual Review of Ecology and Systematics* 33: 49-72.

Sanderson MJ, Purvis A, Henze C. 1998. Phylogenetic supertrees: assembling the trees of life. *Trends in Ecology and Evolution* 13: 105-109.

Seddon JM, Georges A, Baverstock PR, McCord WP. 1997. Phylogenetic relationships of chelid turtles (Pleurodira : Chelidae) based on mitochondrial 12S rRNA gene sequence variation. *Molecular Phylogenetics and Evolution* 7: 55-61.

Shaffer HB, Meylan PA, McKnight ML. 1997. Tests of turtle phylogeny: molecular, morphological, and paleontological approaches. *Systematic Biology* 46: 235-268.

Spinks PQ, Shaffer HB. 2007. Conservation phylogenetics of the Asian box turtles (Geoemydidae, *Cuora*): mitochondrial introgressions, numts, and inferences from multiple nuclear loci. *Conservation Genetics* 8: 641-657.

Spinks PQ, Shaffer HB. 2009. Conflicting mitochondrial and nuclear gene phylogenies for the widely disjunct *Emys* (Testudines: Emydidae) species complex, and what they tell us about biogeography and hybridisation. *Systematic Biology* 58: 1-20.

Spinks PQ, Shaffer HB, Iverson JB, McCord WP. 2004. Phylogenetic hypotheses for the turtle family Geoemydidae. *Molecular Phylogenetics and Evolution* 32: 164-182.

Spinks PQ, Thomson RC, Lovely GA, Shaffer HB. 2009. Assessing what is needed to resolve a molecular phylogeny: simulations and empirical data from emydid turtles. *BMC Evolutionary Ecology* 9: 56.

Stephens PR, Wiens JJ. 2003. Ecological diversification and phylogeny of emydid turtles. *Biological Journal of the Linnean Society* 79: 577-610.

Stephens PR, Wiens JJ. 2009. Evolution of sexual size dimorphisms in emydid turtles; ecological dimorphism, Rensch's rule, and sympatric divergence. *Evolution* 63: 910-925.

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. *Contributions to Zoology* 79: 93-106.

Stuart BL, Parham JF. 2004. Molecular phylogeny of the critically endangered Indochinese box turtle (*Cuora galbinifrons*). *Molecular Phylogenetics and Evolution* 31: 164-177.

Stuart BL, Parham JF. 2007. Recent hybrid origin of three rare Chinese turtles. *Conservation Genetics* 8: 169-175.

Stuart BL, Fritz U. 2008. Historical DNA from museum type specimens clarifies diversity of Asian leaf turtles (*Cyclemys*). *Biological Journal of the Linnean Society* 94: 131-141.

The Reptile Database. 2011. Available at http://www.reptile-database.org/ [visited on 15 Oct 2009].

Thomson RC, Shaffer HB. 2010. Sparse supermatrices for phylogenetic inference: taxonomy, alignment, rogue taxa, and the phylogeny of living turtles. *Systematic Biology* 59: 42-58.

Thomson S, Georges A. 2009. *Myuchelys* gen. nov. - a new genus for *Elseya latisternum* and related forms of Australian freshwater turtle (Testudines: Pleurodira: Chelidae). *Zootaxa* 2053: 32-42.

Vargas-Ramirez M, Castano-Mora OV, Fritz U. 2008. Molecular phylogeny and divergence times of ancient South American and Malagasy river turtles (Testudines: Pleurodira: Podocnemididae). *Organisms, Diversity & Evolution* 8: 388-398.

Vilgalys R. 2003. Taxonomic misidentification in public DNA databases. *New Phytologist* 160: 4-5.

Wiens JJ, Kuczynski CA, Stephens PR. 2010. Discordant mitochondrial and nuclear genes phylogenies in emydid turtles: implications for speciation and conservation. *Biological Journal of the Linnean Society* 99: 445-461.

Wink M, Guiking D, Fritz U. 2001. Molecular evidence for hybrid origins of *Mauremys iversoni* Pritchard and McCoy, 1991, and *Mauremys pritchardi* McCord, 1997 (Reptilia: Testudines: Bataguridae). *Zoologische Abahndlungen (Dresden)* 52: 41-49.

Zardoya R, Meyer A. 1998. Complete mitochondrial genome suggests diapsid affinities of turtles. *Proceedings of the National Academy of Sciences of the USA* 95: 14226-14231.

Zardoya R, Meyer A. 2001. The evolutionary position of turtles revised. *Naturwissenschaften* 88: 193-200.

Received: 23 September 2011
Revised and accepted: 7 March 2012
Published online: 4 July 2012
Editor: J.W. Arntzen

On-line supplementary material

S1. Authorities for species used in this study.
S2. GenBank accession numbers of the sequences used in this study.