Congruence, non-homology, and the phylogeny of basal turtles

Walter G. Joyce[1,2] and Juliana Sterli[3]

[1]Institut für Geowissenschaften, University of Tübingen, 72074 Tübingen, Germany;
[2]Yale Peabody Museum of Natural History, New Haven, Connecticut 06511, USA;
[3]CONICET, Museo Paleontológico Egidio Feruglio, Av. Fontana 140, 9100 Trelew, Chubut, Argentina

Abstract

Joyce, W.G. and Sterli J. 2012. Congruence, non-homology, and the phylogeny of basal turtles.– *Acta Zoologica* (Stockholm) **93**: 149–159.

Modern cladistic analysis is characterized by the assembly of increasingly larger data sets coupled with the use of congruence as the final test of homology. Some critics of this development have recently called for a return to more detailed primary homology analysis while questioning the utility of congruence. This discussion appears to be central to the debate regarding the phylogenetic relationships of basal turtles, as the large data sets developed by us have been criticized recently for utilizing poorly constructed characters and including too many homoplasy-prone characters. Our analysis of this critique reveals that (1) new information regarding poorly understood taxa has a greater impact on the outcome of turtle phylogenies than the characters under dispute; (2) most current turtle phylogenies differ in taxon sampling, not character sampling, and so it appears illogical to condemn a particular analysis for its character sampling; (3) even evolutionary taxonomists should agree that key characters utilized to resolve basal turtle relationships cannot be thought to be 'infallible'; (4) whereas various criteria provide positive evidence for homology, only congruence provides positive evidence for non-homology; and (5) a stalemate between conflicting camps within a congruence frame work is preferable to the *ad hoc* dismissal of data sets, because authoritative statements are untestable.

Walter G. Joyce, Institut für Geowissenschaften, University of Tübingen, 72074 Tübingen, Germany. E-mail: walter.joyce@uni-tuebingen.de

Introduction

The question how to identify characters as synapomorphies is one of the most pervasive issues within systematic biology, because it is only possible to diagnose natural, monophyletic groups of organisms if synapomorphic characters are successfully distinguished from symplesiomorphic and homoplastic ones. Although Hennig (1950, 1966) was the first to develop a cohesive method that demanded the use of synapomorphies only, he still utilized traditional, authoritative criteria to establish which characters he deemed to be synapomorphies. Patterson (1982) synonymized the concept of synapomorphy with the term homology and attempted to establish criteria that can be used to test homology more objectively, such as the 'test of topology' or the 'test of conjunction' (i.e., two structures cannot be homologous if they occur in the same organisms). However, the advance of computer-assisted methods combined with the rationale outlined by 'pattern cladistics'

(e.g., Nelson and Platnik 1981; Rieppel, 1988) led many to conclude that congruence, or the overriding signal found in a data set, should be considered the ultimate test of homology. To avoid confusion, de Pinna (1991) introduced the terms 'primary homology' and 'secondary homology,' to distinguish between the conjectural ideas of homology generated by any given scientist and assessments legitimized through external tests, respectively. Although rarely stated explicitly, the pervasive use of character matrices and computer-generated trees clearly demonstrates that congruence has become the favored method for establishing secondary homologies among morphological systematists. Some objections have been raised over the course of the last decade, in particular that sensible topologies and meaningful secondary homologies cannot be expected to be retrieved from analyses if the primary homologies tabulated in a matrix are poorly constructed (e.g., Rieppel and Kearney 2002, 2007). Wägele (2000) appears to go so far as to reject the use of character matrices completely and calls

for a return to the authoritative, although in principle phylogenetic, taxonomy of Hennig (1950).

This debate on the use of congruence appears to be at the heart of the current dispute regarding the phylogeny of Mesozoic turtles. This debate ultimately originates with Gaffney's (1975) seminal paper in which he outlined a number of characters that can be used to cladistically diagnose the two primary extant clades of turtles, Pleurodira and Cryptodira. These characters were later incorporated into a series of character matrices (e.g., Gaffney *et al.* 1991, 1998, 2007; Gaffney 1996; Hirayama *et al.* 2000), which universally supported their homology within a congruence framework. Key characters that diagnose the stem lineage of crown Pleurodira in these analyses include a fused pelvis and the presence of a trochlea that deflects the adductor jaw musculature around the otic capsule and that is formed by the pterygoid. By contrast, a prefrontal/vomer contact, a vertical pterygoid flange, and the presence of a trochlea that is formed by the otic capsule itself diagnose the stem lineage of crown Cryptodira.

Over the course of the last 15 years, three independent analyses (Rougier *et al.* 1995; Joyce 2007; Tong *et al.* 2009) were published that utilized Gaffney's diagnostic characters but concluded, using congruence, that most should be interpreted as symplesiomorphies or even homoplasies. These data sets were the basis of a number of additional contributions (e.g., Danilov and Parham 2006, 2008; Sukhanov 2006; Sterli 2008; Anquetin *et al.* 2009), all of which arrived at the same conclusion. Although this debate is ultimately focused on character definitions and character observations (Gaffney and Jenkins 2010), the Lower Jurassic turtle *Kayentachelys aprix* has served as the poster child of this debate, as it is thought to be either the world's oldest stem cryptodire (e.g., Gaffney *et al.* 1991, 1998, 2007; Rougier *et al.* 1995; Gaffney 1996) or a stem turtle (e.g., Danilov and Parham 2006, 2008; Sukhanov 2006; Joyce 2007; Sterli 2008; Anquetin *et al.* 2009; Tong *et al.* 2009).

In a series of papers (Joyce 2007; Sterli and Joyce 2007; Sterli 2008), we built explicit data sets that revealed, using congruence, that *K. aprix* may best be viewed as a stem turtle and extensively discussed the implications of this topology regarding character evolution. Gaffney and Jenkins (2010) recently published a response to our papers in which they restate many of Gaffney's original character hypotheses, highlighted a number of objective errors that we had made in our analyses, questioned the use of congruence in establishing homology, and derided out analyses as being 'phenetic' and not 'cladistic,' terminology inspired by the work of Wägele (2000). Whereas we welcome the former two points as part of the scientific debate, we are worried by the implicit return to authoritative taxonomy.

At this point, all differences in character definitions and character observations have been clearly outlined by both camps, we welcome the community to pass judgment, and we see no need to state our side again. We nevertheless would like to take the opportunity to explore four related issues. We first test the impact of some classic 'cryptodiran characters' on the outcome of a number of analyses that have supported the phylogenetic placement of *K. aprix* as a stem cryptodire relative to minor adjustments in character scoring of this taxon. We then investigate the hypothesis of Gaffney and Jenkins (2010) that our analyses are skewed through the use of irrelevant postcranial characters and by incorrect character definitions. With a certain amount of sympathy for Wägele's (2000) notion that some characters appear to be too unique to have evolved twice, we next discuss whether the classic characters of Gaffney can be equated in their uniqueness to such 'infallible' characters as the vertebrate eye or feathers. Finally, we would like to state our case in support of congruence and argue that scientific impasses over character definitions or character scorings are preferable to any type of authoritative taxonomy, because they can ultimately be resolved through additional data, whereas authoritative statements are untestable and can only be countered by more authoritative statements.

Materials and Methods

To investigate the influence of varying character concepts onto the placement of *K. aprix*, we reran a series of published analyses (i.e., Rougier *et al.* 1995; Gaffney 1996; Hirayama *et al.* 2000; Gaffney *et al.* 2007; and Sterli 2008) but altered the use and/or scoring of two characters that are supposed to be particularly useful for correctly discerning turtle relationships (Gaffney 1975; Gaffney and Jenkins 2010) but for which there is no agreement how they should be scored (Joyce 2007; Sterli and Joyce 2007; Sterli 2008). These characters are the presence of a vertical flange along the lateral edge of the external pterygoid process and the presence of an otic trochlea. Given that the use of these characters is controversial, we informally refer to them as 'problematic characters.' In the vast majority of analyses (e.g., Gaffney *et al.* 1991, 1998, 2007; Rougier *et al.* 1995; Gaffney 1996), the problematic characters are scored as present for *K. aprix*, i.e., this taxon is scored as possessing the derived character state. Joyce (2007) omitted these two key characters from his analysis as he noticed that his observations regarding the distribution of these characters did not overlap with those of previous analyses and he thus felt it better to omit them completely, together with a number of other irreproducible characters. We reiterated these problems in Sterli and Joyce (2007). Sterli (2008) based her analysis on that of Joyce (2007) but decided to include these characters again (contra to the claims of Gaffney and Jenkins 2010), but scored *K. aprix* as lacking a vertical flange on the external pterygoid process and an otic trochlea, i.e., the problematic characters were scored as being plesiomorphic for that taxon. To investigate the influence of these three different approaches to these problematic characters, we reran all analyses while varying their inclusion and/or scoring.

In addition to varying the inclusion and/or scoring of the two 'problem characters' discussed above, we corrected the scoring of *K. aprix* in all data sets based on our observations on all available material. *Kayentachelys aprix* was originally

described by Gaffney *et al.* (1987), but the short format in a highly ranked journal did now allow for the inclusion of a formal description. A number of new specimens were found in the following 20 years that provide novel insights into the morphology of this taxon, but this material has also not been formally described. This lack of descriptions was thoroughly rectified for the cranium, and two descriptions are now available (Sterli and Joyce 2007; Gaffney and Jenkins 2010) that disagree in some minor details, primarily in the amount of confidence with which certain structures are observed. The authors of this paper enjoyed full access to all available *K. aprix* postcranial material, including a number of well-preserved new specimens, and were thus able to address some minor points regarding the postcranial anatomy of this taxon as well. We feel that all modifications are based on sound observations on the available material and are thus uncontroversial. A summary of our changes to each data set is provided in Appendix 1.

Five representative turtle data sets of the last 15 years (i.e., Rougier *et al.* 1995; Gaffney 1996; Hirayama *et al.* 2000; Gaffney *et al.* 2007; Sterli 2008) were selected and manipulated for this study. As described earlier, the corrected data sets were run three times varying in the inclusion and the scoring of 'problem characters' (i.e., the derived presence of a vertical flange in the external process of the pterygoid and the derived presence of an otic trochlea). In the first run, *K. aprix* was scored as plesiomorphic for the problem characters (as done by Sterli 2008). In the second run, *K. aprix* was scored as derived for the problem characters (as done by Gaffney *et al.* 1987, 1991, 2007). In the third run, the problem characters were omitted completely (as was done by Joyce 2007). To make our results more comparable to those of previous authors, we ran their analyses using their analysis criteria (e.g., Rougier *et al.* 1995, characters 9, 15, 45 were run ordered; all character left unordered for all other analyses).

The cladistic analyses were performed using TNT program (Goloboff *et al.* 2008). The data sets of Rougier *et al.* (1995), Gaffney (1996), and Hirayama *et al.* (2000) were run using Implicit Enumeration, while data sets of Gaffney *et al.* (2007) and Sterli (2008) were run using heuristic searches with two cycles of TBR. The first cycle performed 1000 replicates saving 10 trees per replicate, whereas the second cycle used all the trees in the memory to ensure obtaining all possible trees. If more than one most parsimonious tree was retrieved, a strict consensus tree was calculated. The branch support was calculated using bootstrapping, with 1000 replicates. The results of this analysis are presented in Fig. 1.

Discussion

The impact of 'problematic characters'

The phylogenetic placement of various fossil turtles, in particular *K. aprix*, has been controversial for some time (e.g., Gaffney *et al.* 1987; Dryden 1988). A number of reasons exist why

phylogenetic analyses can diverge in their results; the most commonly discussed causes include character and taxon sampling. However, in the case of *K. aprix*, contrasting opinions are perhaps mostly attributable to differences in character interpretations and observations (Gaffney and Jenkins 2010). We here distinguish between two types of observational differences. The first category pertains to conflicting observations that are made as a result of differing quality of studied materials. In the case of *K. aprix*, 20 years of additional collection and preparation have produced a number of specimens that allow objectively correcting observations made based on lesser materials.

The other type of observational difference pertains to characters where different observers cannot agree on the scoring even though the same material is available. In two previous publications (Joyce 2007 and Sterli and Joyce 2007), we argued that we could not objectively replicate observations made regarding the otic trochlear surface or the vertical pterygoid flange of basal turtles in general. However, given that these characters have been argued to be essential for resolving basal turtle relationships, omitting them provides room for criticism (e.g., Gaffney and Jenkins 2010). For simplicity, we refer to these characters as the 'problematic characters.'

Although it is not the first explicit data set (see Dryden 1988; Gaffney *et al.* 1991), the phylogenetic analysis of Rougier *et al.* (1995) is the oldest one investigated herein. As originally conceived, this analysis supported the placement of *K. aprix* within crown Testudines along the stem of Cryptodira. Interestingly, regardless of the inclusion or scoring of the problematic characters, our revised analyses always place *K. aprix* outside crown Testudines. This indicates that the objective changes that we were able to apply to this data set were sufficient to change the phylogenetic position of *K. aprix* and that the inclusion of the problematic characters has no effect on the outcome of the analysis.

The phylogenetic analysis of Gaffney (1996) was developed independently from the analysis of Rougier *et al.* (1995) and differed primarily by better sampling (i.e., not lumping) the phylogenetic stem and crown Cryptodira. However, in contrast to Rougier *et al.* (1995), this analysis combined the putative stem pleurodire *Proterochersis robusta* and Pleurodira into a single terminal taxon. Our reanalysis of this data set reveals that *K. aprix* is placed within crown Testudines regardless of whether or how the problematic characters are included. This indicates that the analysis of Gaffney (1996) is impervious to changes in the overall understanding of the anatomy of *K. aprix*. Although this could be taken as firm evidence in support of *K. aprix* as a crown turtle, we believe that this may be the result of collapsing all known pleurodires and *P. robusta* into a single terminal taxon. Conversely, whereas we do not doubt that *P. robusta* is more basal than *K. aprix* (e.g., Rougier *et al.* 1995; Sukhanov 2006; Joyce 2007; Sterli 2008), we see little evidence at this point that this taxon is a stem pleurodire. That being said, a significant morphological gap still exists between the most basal pleurodires known from the
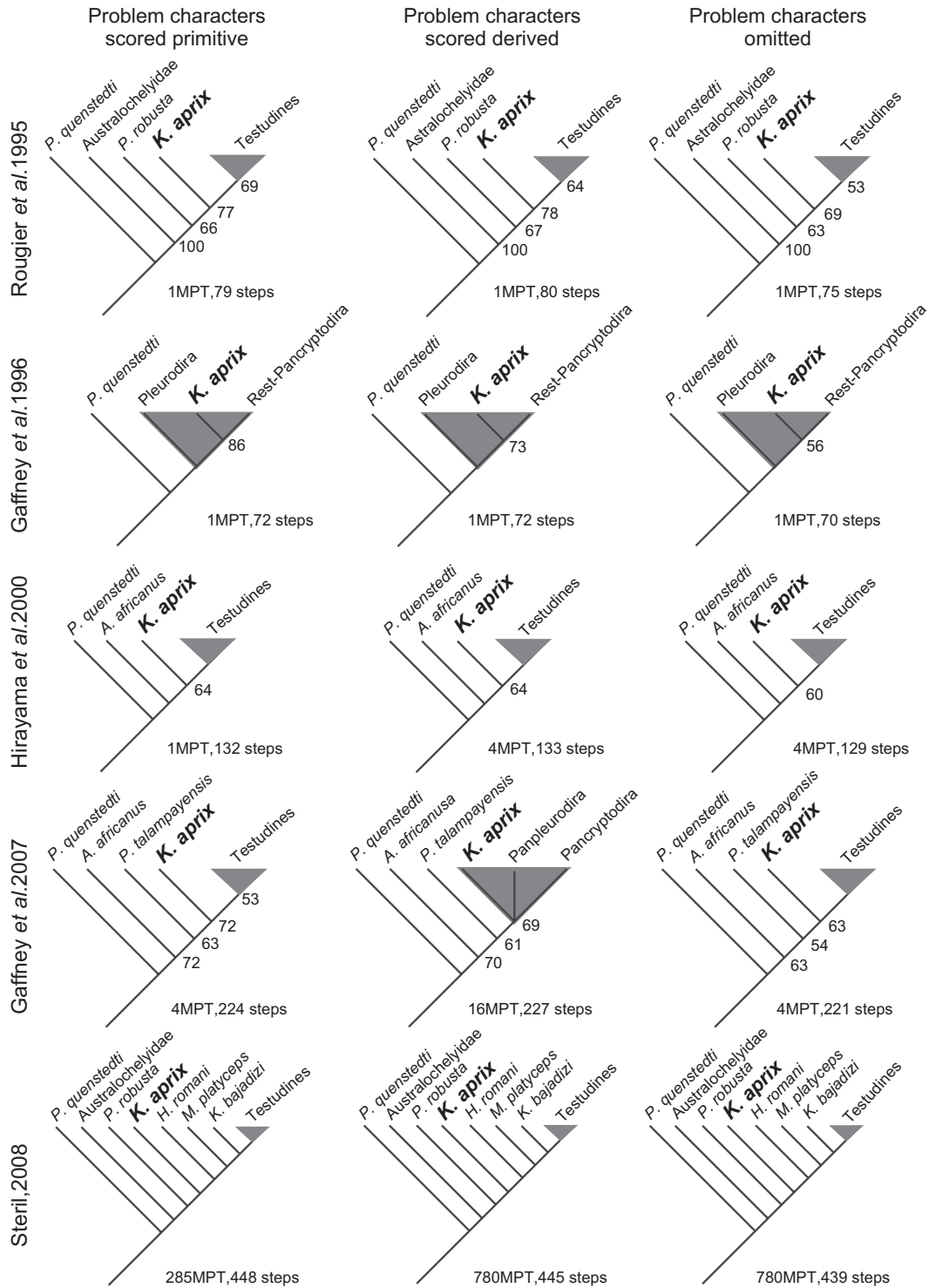
Fig 1—Simplified phylogenetic trees summarizing the results of this study. Numbers next to nodes are bootstrap values. The left column lists the five data sets that were modified slightly by correcting the scoring of *Kayentachelys aprix* based on observations made on better materials discovered over the course of the last 20 years. The second and third columns provide the outcome of the analyses when *K. aprix* was scored as plesiomorphic or derived, respectively, for the two 'problematic characters.' The fourth column displays the results of the analysis when the 'problematic characters' were omitted completely from the analysis. Crown group Testudines is always highlighted as a grey triangle.

Late Jurassic and all remaining turtles (Gaffney and Jenkins 2010). New material from the Lower to Middle Jurassic will thus test our assertion that *P. robusta* is not a stem pleurodire.

The phylogenetic analysis of Hirayama *et al.* (2000) is influenced by the previous two analyses but differs significantly through the inclusion of a large number of additional characters and taxa. As with Rougier *et al.* (1995), this analysis originally supported the placement of *K. aprix* within crown Testudines, but regardless of the inclusion or scoring of the problematic characters, the revised data set supports the placement of that taxon along the phylogenetic stem of crown Testudines.

The analysis of Gaffney *et al.* (2007) is heavily influenced by the previous three analyses but remains independent by scoring characters differently and sampling other characters and taxa. A basal placement of *K. aprix* is achieved in our revision when the problematic characters are scored as plesiomorphic or omitted, but *K. aprix* is placed in a polytomy with both the pleurodiran and cryptodiran stem lineages in the strict consensus tree of the analysis that includes the problematic characters as derived for *K. aprix*. This indicates that the objective character corrections that we undertook for *K. aprix* in this data set are not sufficient by themselves to achieve a more basal placement for that taxon. Instead, the result of this data set relies on whether the problematic characters are included and what scoring is used.

The final data set to be re-examined is that of Sterli (2008), an updated version of Joyce (2007). This analysis is also heavily influenced by previous analyses, but the attempt was made to explore new, especially postcranial character complexes and include as many potentially basal taxa as possible. Not surprisingly, when omitting the problematic characters or including them with the basal condition, *K. aprix* firmly remains along the phylogenetic stem of Testudines. However, even when *K. aprix* is scored derived for the problematic characters, the hypothesized placement of this taxon does not shift. The overall character support thus dominates the problematic characters in this analysis.

In conclusion, minor, objective changes to the scoring of *K. aprix* have a much greater impact on the phylogenetic placement of this taxon than any particular use or scoring of the 'problematic' characters. This is especially true for the analyses of Rougier *et al.* (1995) and Hirayama *et al.* (2000), which universally retrieve a basal placement of *K. aprix* regardless of character use or character concept. Of course, these results are only of interest within a congruence framework.

The impact of postcranial characters

One of the central points of critique by Gaffney and Jenkins (2010) of the analysis of Joyce (2007), and all analyses that derived from it (Sterli 2008; Anquetin *et al.* 2009), is that it places too much emphasis on postcranial characters. This rationale only makes sense if postcranial characters are presumed *a priori* to be more homoplastic than cranial characters. Although it is true that ecological factors can produce striking convergence among postcranial characters (e.g., the postcranial anatomy of protostegid and chelonioid turtles), the same pressures can also influence the shape of the skull and confound systematic biologists. The bigheaded turtle *Platysternon megacephalum* is a good example within Testudinata. Based on a series of cranial characters, this taxon has been argued to be nested deeply within Chelydridae (e.g., Gaffney 1975; Gaffney and Meylan 1988). However, postcranial characters have always demonstrated an affiliation with Testudinoidea (e.g., Günther 1864; Vaillant 1894; Williams 1950; Romer 1956). Interestingly, well-sampled molecular trees (Krenz *et al.* 2005; Parham *et al.* 2006; Barley *et al.* 2010) and a detailed study on pulmonary morphology (Lambertz *et al.* 2010) have recently come down strongly in favor once again of the testudinoid affiliation of *P. megacephalum*. The combined postcranial and molecular evidence therefore demonstrates that the cranial characters that have been used to unite *P. megacephalum* with Chelydridae are homoplastic. Cranial characters are clearly as prone to homoplasy as postcranial characters. Yet, even though a rationale could now be phrased to justify the exclusion of cranial characters, we would never endorse their omission from a global analysis as they nevertheless carry a true, though apparently homoplastic phylogenetic signal.

Although not explicit, any criticism that postcranial characters skew Joyce's (2007) analysis (Gaffney and Jenkins 2010) implies that analyses that place *K. aprix* along the phylogenetic stem of Cryptodira do not 'suffer' from too many postcranial characters. To further investigate this claim, we calculated the percent of cranial characters utilized in the five analyzed matrices and quantified their average CI relative to that of the postcranial characters (see Data S1 in supporting information for calculated CI values). In the matrix of Rougier *et al.* (1995), 64% of all derived character states are cranial, the highest percentage of all matrices. Yet, although this matrix originally interpreted *K. aprix* as a stem cryptodire, minor adjustments to the scoring of a small number of postcranial characters are sufficient to firmly sway the result in favor of *K. aprix* being a stem turtle. The analyses of Gaffney (1996), Gaffney *et al.* (2007), and Sterli (2008) are intermediate with 45%, 55%, and 42% derived cranial character states, respectively. The analysis of Hirayama *et al.* (2000) includes 41% the smallest amount of derived cranial character states, yet, prior to our analysis, this matrix supported the derived placement of *K. aprix* within crown Testudines. The CI of cranial characters in Rougier *et al.*'s (1995) analysis is approximately 0.89, whereas the CI of postcranial characters is 0.76. The corresponding cranial and postcranial values are 0.59 and 0.58 for Gaffney (1996), 0.75 and 0.56 for Hirayama *et al.* (2000), 0.59 and 0.53 for Gaffney *et al.* (2007), and 0.42 and 0.45 for Sterli (2008), respectively (see Data S1 in supporting information). It is apparent that newer analyses include more homoplasy, and this rise in homoplasy clearly

correlates with the number of terminal taxa used. In general, cranial characters are less homoplastic in most analyses although differences are often negligible.

We cannot discern any correlation between the percentage of cranial character used or homoplasy with any particular topology. Instead, our analyses demonstrate that various published data sets that formerly supported the cryptodiran affinities of *K. aprix* place this taxon along the phylogenetic stem of turtles just by fixing a small number of objective coding errors. Within a congruence framework, this demonstrates that all independently developed analyses of the last 15 years appear to converge upon the same signal (e.g., Rougier *et al.* 1995; Hirayama *et al.* 2000; Gaffney *et al.* 2007; Tong *et al.* 2009) and reveals that the analysis of Joyce (2007) is not an awkward outlier that demands special explanation through the *ad hoc* dismissal of postcranial characters.

Special pleading for special characters

The modern critique of congruence-based methods is focused on the idea that one cannot expect results that reflect the true tree of life if the data set upon which the analysis is based consists of poorly constructed or highly homoplastic characters (e.g., Wägele 2000; Rieppel and Kearney 2002, 2007). We fully agree. It nevertheless appears clear to us that this critique does not apply to the current controversy regarding the phylogeny of Mesozoic turtles. First, as has been demonstrated elsewhere, *a priori* knowledge of homoplasy is not a reason to dismiss a character, because there is no philosophically rigorous basis for the *a priori* dismissal of well-framed characters and because even highly homoplastic characters can exhibit a clear signal (e.g., Poe and Wiens 2000). For instance, it would be unwise to omit the character 'presence of four limbs' from an analysis of vertebrates, just because it is apparent that these structures are lost independently many times, for instance in snakes, amphisbaenids, and various lizards, because the presence of this character supports the monophyly of Tetrapoda and because the loss of this character, at the same time, supports these monophylies of the numerous groups that lost their limbs. Second, even though current fossil turtle data sets are growing in size, they still all consist of less than 200 derived character states and are thus small enough to allow rigorous reanalysis of each character and all scorings. Indeed, although the analyses of Joyce (2007) and Gaffney *et al.* (2007) differ substantially in the choice of terminal taxa (67 versus 27 turtle taxa, respectively), the character sample is nearly identical (169 versus 131 derived character states, respectively), and differences in scoring minimal. Finally, among the large sets of characters being used in such analyses as Joyce (2007) and Gaffney *et al.* (2007), only two are controversial and our analysis demonstrates that these two characters only have a minor impact on a small region of the tree. We thus are confident that no current analysis of turtle relationships suffers from the 'garbage in, garbage out' syndrome.

A more compelling critique of current cladistic methods pertains to weighting (Wägele 2000). In an attempt to minimize *a priori* assumption, most data sets are currently run with all characters weighted equally. However, as Wägele (2000) correctly notes, all character matrices already represent weighting schemes, because it is up to the individual researcher to decide whether a character complex is finely anatomized or lumped, thus adding or reducing the weight of that character complex relative to other characters. As a result, highly complex characters that plausibly could have originated only once during evolution may receive the same weight as poorly constructed or homoplastic ones and may thus have only little impact on the overall analysis. Classic characters that might be construed as a single character, but that might be thought of as too unique to have evolved more than once include the vertebrate eye or feathers. Should a character as menial as the exclusion of the frontal from the orbit be able to contradict the signal preserved in the vertebrate eye? As believers in the power of congruence, we suppose that any character can be tested through the use of 'menial' characters as long as these characters are well defined and correctly scored. With a certain amount of sympathy to the line of reasoning of Wägele (2000), we nevertheless would like to explore whether enough *a priori* knowledge exists that would allow us to speculate that the two characters championed by Gaffney and Jenkins (2010), the vertical flange of the external pterygoid process and the otic trochlear system, should be deemed 'infallible.' Joyce (2007) already outlined much of this debate, but we summarize it herein again, with special respect to this issue.

The two characters in question pertain directly to the trochlear system of turtles. Gaffney (1975) originally explored this character complex and hypothesized that the presence of a vertical flange on the external pterygoid process and an otic trochlea are synapomorphies for Cryptodira and that the pterygoid trochlea is a synapomorphy of Pleurodira. Given that the otic trochlea of cryptodires is mostly formed by the quadrate and prootic and that the pterygoid trochlear of pleurodires is formed by the pterygoid, undoubtedly non-homologous structures, Gaffney (1975) concluded that the trochlear system of both crown groups must have evolved independently from the condition seen in *Proganochelys quenstedti* and cannot have given rise to another.

The topology retrieved from the analysis of Joyce (2007) implies that the pleurodiran trochlear system derived from a cryptodiran-like trochlear system, contrary to Gaffney's (1975) character assessment. Joyce (2007) already noted that this should *not* be misunderstood to mean that the pleurodiran trochlear system derived from the system seen in crown cryptodires in its entire complexity, given that pleurodires are not placed within crown Cryptodira. Instead, the pleurodiran trochlear system likely derived from a rather plesiomorphic trochlear system that only resembled the system of extant cryptodires in that the otic capsule diverted the temporal musculature. In other words, although only few data exist at the

moment that help document the origin of the modern cryptodiran trochlear system, it is reasonable to speculate that it originated in a sequence of steps and that pleurodires split from this evolutionary trend long before its completion. We therefore *must not* attempt to derive a modern pleurodire from a modern cryptodire.

Joyce (2007) took great pain to rigorously assess whether turtles such as *K. aprix* had a trochlear system at all. This taxon clearly lacks a protruding processus trochlearis oticum such as seen in *Chelydra serpentina* (Gaffney 1972), but we agree with Gaffney and Jenkins (2010) that the anterior wall of the otic capsule appears to be thickened. However, does that necessarily imply that a trochlear structure must be present in this taxon? Similarly, although Gaffney and Jenkins (2010) insist that *K. aprix* possessed a vertical flange along the external pterygoid process, we still are not convinced that this taxon is substantially different from *P. quenstedti*, which has never been argued to possess an otic trochlea (Gaffney 1990).

Joyce (2007) suggested that this problem could be addressed by assessing the geometry of the skull of a fossil taxon (Fig. 2). In basal amniotes and basal turtles like *P. quenstedti*, the temporal musculature originated along the coronoid process, inserted at the top of the temporal cavity, and was not intersected by any structures (Fig. 2A). In all living turtles, by contrast, the otic region blocks the direct line that connects the points of origin and insertion of the temporal musculature, necessitating a diversion of the musculature

through the development of a trochlea (Fig. 2B,D). Considering that the skull geometry of *K. aprix* does not appear to necessitate a trochlea, we concluded that this taxon did not possess any type of trochlear system (Sterli and Joyce 2007). It is apparent that we did not convince Gaffney and Jenkins (2010).

To a certain degree, the question whether *K. aprix* possessed an otic trochlea is irrelevant to this chapter, because Joyce (2007) ultimately concluded that geometric arguments and morphological correlates indicate that some taxa placed along the phylogenetic stem of Testudines indeed possessed a trochlear system that was redirected by the otic capsule. That implies that a cryptodire-type trochlear system originates before than the pleurodiran system and that the vertical flange of the external pterygoid process originates before than the pterygoid trochlea. To achieve the transition from a cryptodiran-type trochlear system to a pleurodiran trochlear system, Joyce (2007) envisioned a transform of function, where the lateral process of the pterygoid expanded through time and eventually took over the redirecting function from the otic capsule (Fig. 2). If pleurodires were to evolve from a more basal ancestor without a trochlea, we see little difference, as the pterygoid trochlea would still represent a modified lateral pterygoid process and would have had to 'capture' the temporal musculature. Would this evolutionary transition be completely hindered by the presence of a vertical flange of the external pterygoid process? Would the pterygoid be incapable
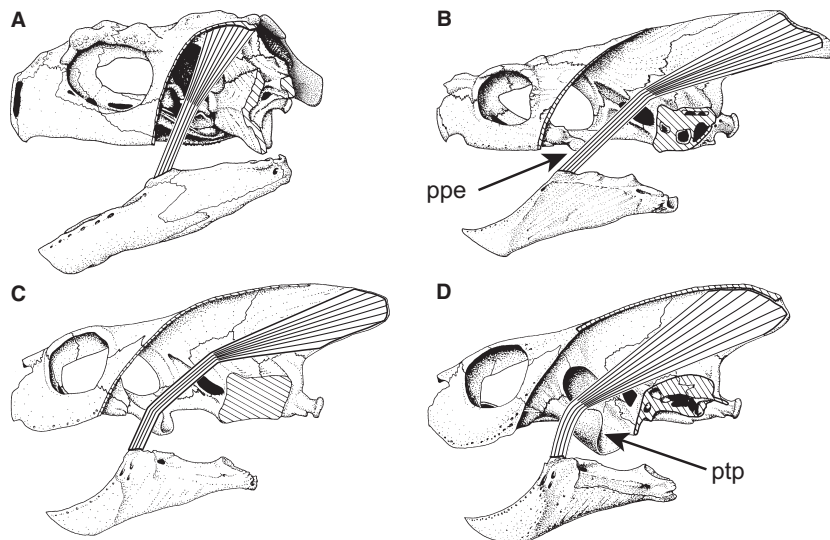

Fig 2—The temporal jaw musculature of turtles. Portions of the dermal roofing and ear region are removed from all illustrations to allow viewing inside the temporal fossa.— **A.** The Late Triassic stem turtle *Proganochelys quenstedti*. The temporal jaw musculature originates at the top of the upper temporal fossa and inserts onto the coronoid process of the mandible. The ear region does not block the direct path of the temporal musculature.— **B.** The extant cryptodiran turtle *Chelydra serpentina*. The temporal jaw musculature originates at the posterior end of the supraoccipital crest and inserts into the coronoid process of the mandible. The otic capsule blocks the direct path of the temporal musculature but forms a trochlear process to redirect the musculature. A small processus pterygoideus externus (ppe) helps guide the lower jaw during adduction.— **C.** A hypothetical intermediate. The processus pterygoideus externus is enlarged and helps redirect the temporal musculature.— **D.** The extant pleurodiran turtle *Elseya dentata*. The processus pterygoideus externus is further enlarged to become the processus trochlearis pterygoidei (ptp), which solely assumed the function of redirecting the temporal musculature.

of redirecting the musculature, just because the otic capsule was already carrying out this function? Considering the complexity of many other evolutionary transitions that most scientists agree to have occurred, can our proposed minor transitions really be deemed categorically impossible? Can Joyce's (2007) proposed hypothetical intermediate (Fig. 2C) simply be disregarded as mechanically implausible (Gaffney and Jenkins 2010) without any supporting arguments? Although our adherence to a congruence framework does not allow us to categorically dismiss the proposed transitions, we feel that even staunch supporters of evolutionary taxonomy would agree with our assessment that the trochlear system of turtles is not philosophically equivalent to feathers in its uniqueness and that this character complex does not deserve any special pleading.

Congruence versus authoritative taxonomy

The identification of synapomorphy is central to systematic biology, as it is only possible to diagnose natural, monophyletic groups of organisms if synapomorphies are distinguished from symplesiomorphies or homoplasies. A number of authors have bemoaned the recent trend toward the development of increasingly larger data matrices, because they feel that less emphasis is placed on developing well-reasoned characters and because they infer that many poorly reasoned characters may swamp the signal of well-reasoned characters with a true phylogenetic signal (e.g., Rieppel and Kearney 2002, 2007). We feel that stating this point is stating the obvious, as any scientist will attest to the importance of using quality data in any type of analysis. At the same time, we feel that this line of critique is highly unconstructive and exhibits hints of authoritative taxonomy, as the *ad hoc* dismissal of a data set does not provide a basis for scientific discourse and asserts scientific superiority without foundation.

As Gaffney and Jenkins (2010) correctly reassert, characters are subjective abstractions created by scientists and must be viewed as testable hypotheses. Characters that *a priori* appear to be non-homologous may actually be homologous, and characters thought to be homologous may actually be non-homologous. The rampant amount of homoplasy that pervades most phylogenetic analyses attests that scientists are not very successful in identifying homology (e.g., Gaffney *et al.* 2007; Joyce 2007). Perhaps Gaffney and Jenkins (2010) are right in that careful analysis of a character using homology criteria such as those of Remane (1952) or Patterson (1982) will allow the identification of characters that are just too unique, too complex, and too unlikely to have evolved twice, characters such as feathers. However, the difference between the analyses of Joyce (2007) and those of Gaffney (e.g., Gaffney *et al.* 1991, 1998, 2007; Gaffney 1996) actually boils down to the opposite as the analysis of Joyce (2007) predicts that structures *not* thought to be homologous (i.e., the trochlear systems of pleurodires and cryptodires) are hypothesized to be homologous after all.

However, whereas it appears plausible to argue that structures could be too similar to *not* be homologous, we cannot think of any criteria that would allow postulating with the same confidence that two structures are too different to be homologous. Whereas there are criteria that allow the positive identification of primary homologies, there are no criteria that positively allow establishing primary non-homology. Yet, despite this lack of objective criteria for non-homology, every systematist continuously has to assert non-homology when defining characters.

The advantage of using congruence as the ultimate test of homology is that it avoids circularity by using external evidence to assess the homology of a character and that it allows for a positive and negative test of homology, thus overcoming the problem discussed earlier in regard to non-homology. Most systematists know that any congruence-based topology remains a hypothesis that is only as good as the primary data. However, as long as all characters are clearly defined, all materials used clearly listed, all observations explicitly coded in a data set, and the search criteria and algorithms outlined, an analysis remains explicit and can be subjected to scientific critique. A scientific impasse may result when two or more groups of systematists cannot agree upon primary homologies, character distributions, or analytical methods, but we feel that such impasses are preferable to authoritative statements. This is because new data may provide resolution to impasses, whereas authoritative statements are immune to new data and remain untestable.

References

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

Barley, A. J., Spinks, P. Q., Thomson, R. C. and Shaffer, H. B. 2010. Fourteen nuclear genes provide phylogenetic resolution for difficult nodes in the turtle tree of life. – *Molecular Phylogenetics and Evolution* **55**: 1189–1194.

Danilov, I. G. and 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. – *Journal of Vertebrate Paleontology* **26**: 573–580.

Danilov, I. G. and 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. – *Journal of Vertebrate Paleontology* **28**: 306–318.

Dryden, L. S. 1988. *Paraphyly of the Cryptodira and Phylogenetic Systematics of Turtles*. Unpublished thesis for MA in University of Kansas Library, 142 pp.

Gaffney, E. S. 1972. An illustrated glossary of turtle skull nomenclature. – *American Museum Novitates* **2486**: 1–33.

Gaffney, E. S. 1975. A phylogeny and classification of the higher categories of turtles. – *Bulletin of the American Museum of Natural History* **155**: 389–436.

Gaffney, E. S. 1990. The comparative osteology of the Triassic turtle *Proganochelys*. – *Bulletin of the American Museum of Natural History* **194**: 1–263.

Gaffney, E. S. 1996. The postcranial morphology of Meiolania platyceps and a review of the Meiolaniidae. – *Bulletin of the American Museum of Natural History* **229**: 1–166.

Gaffney, E. S. and Jenkins, F. A. Jr 2010. The cranial morphology of *Kayentachelys*, an Early Jurassic cryptodire, and the early history of turtles. – *Acta Zoologica (Stockholm)* **91**: 335–368.

Gaffney, E. S. and Meylan, P. A. 1988. A phylogeny of turtles. In: Benton, M. J. (Ed.): *The Phylogeny and Classification of the Tetrapods*, Vol. 1, Amphibians, Reptiles, Birds. – *Systematics Association Special Volume* **35A**: 157–219.

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

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

Gaffney, E. S., Kool, L., Brinkman, D., Rich, T. H. and Vickers-Rich, P. 1998. *Ottwayemys*, a new cryptodiran turtle from the Early Cretaceous of Australia. – *American Museum Novitates* **3233**: 1–28.

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

Goloboff, P. A., Farris, J. S. and Nixon, K. 2008. A free program for phylogenetic analysis. – *Cladistics* **24**: 774–786.

Günther, A. C. L. G. 1864. *The Reptiles of British India*, pp 452. R. Hardwicke, London.

Hennig, W. 1950. *Grundzüge einer Theorie der phylogenetischen Systematik*. Deutscher Zentralverlag, Berlin.

Hennig, W. 1966. *Phylogenetic Systematics*, pp 263. University of Illinois Press, Urbana, IL.

Hirayama, R., Brinkman, D. B. and Danilov, I. G. 2000. Distribution and biogeography of non-marine Cretaceous turtles. – *Russian Journal of Herpetology* **7**: 181–198.

Joyce, W. G. 2007. Phylogenetic relationships of Mesozoic turtles. – *Bulletin of the Yale Peabody Museum* **48**: 3–102.

Krenz, J. G., Naylor, G. J. P., Shaffer, H. B. and Janzen, F. J. 2005. Molecular phylogenetics and evolution of turtles. – *Molecular Phylogenetics and Evolution* **37**: 178–191.

Lambertz, M., Böhme, W. and Perry, S. F. 2010. The anatomy of the respiratory system in *Platysternon megacephalum* Gray, 1831 (Testudines: Cryptodira) and related species, and its phylogenetic implications. – *Comparative Biochemistry and Physiology, Part A* **156**: 330–336.

Nelson, G. and Platnik, N. 1981. *Systematics and Biogeography*. Columbia University Press, New York.

Parham, J. F., Feldman, C. R. and Boore, J. L. 2006. The complete mitochondrial genomen 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, DOI: http://www.biomedcentral.com/1471-2148/6/11.

Patterson, C. 1982. Morphological characters and homology, In: Joysey, K. A. and Friday, A. E. (Eds): *Problems in Phylogenetic Reconstruction*, pp. 21–74. Systematics Association Special Volume 21. Academic Press, London.

de Pinna, M. G. G. 1991. Concepts and tests of homology in the cladistic paradigm. – *Cladistics* **7**: 367–394.

Poe, S. and Wiens, J. J. 2000. Character selection and the methodology of morphological phylogenetics. In: Wiens, J. J. (Ed.): *Phylogenetic Analysis of Morphological Data*, pp. 20–36. Smithsonian Institution Press, Washington, DC.

Remane, A. 1952. *Die Grundlagen des nat[aaa]lichen Systems Der Vergleichenden Anatomie und der Phylogenetik*. Akademische Verlagsgesellschaft Geest & Portig K.-G, Leipzig.

Rieppel, O. 1988. *Fundamentals of Comparative Biology*. Birkhauser Verlag, Basel.

Rieppel, O. and Kearney, M. 2002. Similarity. – *Biological Journal of the Linnean Society* **75**: 59–82.

Rieppel, O. and Kearney, M. 2007. The poverty of taxonomic characters. – *Biology and philosophy* **22**: 95–113.

Romer, A. S. 1956. *Osteology of the Reptiles*, pp 772. University of Chicago Press, Chicago.

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

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

Sterli, J. and Joyce, W. G. 2007. The cranial anatomy of the lower Jurassic turtle *Kayentachelys aprix*. – *Acta Palaeontologica Polonica* **52**: 675–694.

Sukhanov, V. B. 2006. An archaic turtle, *Heckerochelys romani* gen. et sp. nov., from the Middle Jurassic of Moscow region, Russia, In: Danilov, I. G. and Parham, J. F. (Eds): *Fossil Turtle Research Volume 1*, pp. 104–111. Zoological Institute, Russian Academy of Sciences, St. Petersburg.

Tong, H., Claude, J., Naksri, W., Suteethorn, V. and Buffetaut, E. 2009. *Basilochelys macrobios* n. gen. and n. sp., a large cryptodiran turtle from the Phu Kradung Formation (latest Jurassic-earliest Cretaceous) of the Khorat Plateau, NE Thailand. – *Geological Society, London, Special Publications* **315**: 153–173.

Vaillant, L. 1894. Essai sur la classification générale des chéloniens. – *Annales des Sciences Naturelles* **16**: 331–345.

Wägele, J. W. 2000. *Grundlagen Der Phylogenetischen Systematik*, pp 315. Verlag Friedrich Pfeil, München.

Williams, E. E. 1950. Variation and selection in the cervical central articulations of living turtles. – *Bulletin of the American Museum of Natural History* **94**: 511–561.

Supporting Information

Additional supporting information may be found in the online version of this article.

Data S1. A summary of changes to the five data sets used in the phylogenetic analyses.

Please note: Blackwell Publishing are not responsible for the content or functionality of any supporting information

supplied by the authors. Any queries (other than missing material) should be directed to the corresponding author for the article.

Appendix 1

Character corrections for Rougier et al. (1995)

Objective changes. Character 25, basisphenoid/basioccipital medial process unpaired (0), paired (1), absent (2): Currently, this character is scored as absent (2), but our review of the *K. aprix* material (see Sterli and Joyce 2007) indicates that these processes are paired (1).

Character 26, recessus scale tympani and perilymphatic fenestra defined by bone: Currently, these structures are scored as being defined by bone (1), but our observations (see Sterli and Joyce 2007) indicate that they are unossified (0).

Character 27, jugular posterior foramina defined by bone: As a consequence of our changes to the scoring of character 26, the posterior jugular foramina must be scored as absent (0), not present (1).

Character 28, small anterior jugular foramina: Similar to character 27, the anterior jugular foramina is actually poorly defined in *K. aprix* (see Sterli and Joyce 2007) and thus must be scored as basal (0), not derived (1).

Character 30, acquisition of a supraoccipital crest: The available material of *K. aprix* clearly reveals that this taxon did not possess an extended supraoccipital crest and should thus be scored as plesiomorphic (0), not derived (1).

Character 40, eleven peripheral bones: Prior to this contribution, it was universally accepted that *K. aprix* possessed 11 pairs of peripherals (1). However, our reanalysis of the available material indicates that this taxon should be scored polymorphic, as at least one individual exhibits 12 peripherals (0/1).

Character 49, first dorsal rib smaller than second: Our analysis of the available *K. aprix* clearly reveals that this taxon, like many other basal turtles, possessed a large first thoracic vertebra and thus should be scored plesiomorphic (0), not derived.

Character 51, acromial process triangular plate: This character is somewhat difficult to interpret. However, our review of the postcranial anatomy of *K. aprix* revels that this taxon did not have a rod-like acromial process, as seen in all more derived turtles, but rather one that was tri-radiate, the basal condition seen in *Proganochelys quenstedti*. We thus score this character as plesiomorphic (0), not uncertain (?).

Problematic characters. Character 12: This character documents morphological variation regarding the acquired presence of the processus trochlearis oticum and is considered 'problematic' herein (see text for discussion).

Character 13: This character documents morphological variation of the vertical pterygoid flange and is considered 'problematic' herein (see text for discussion).

Character corrections for Gaffney (1996)

Objective changes. Characters 22, 24-27, 29–30, 33: These eight characters pertain to the morphology of the vertebral column. Unfortunately, although isolated vertebra are known for *K. aprix* (see description presented herein), most of these characters cannot be scored objectively and should thus be scored as unknown. However, given that these characters pertain to significantly more derived morphologies, all scorings are reasonable and thus left untouched.

Problematic characters. Character 4, pterygoid flange: see Rougier *et al.* 1995

Character 7, processus trochlearis oticum: see Rougier *et al.* 1995

Character corrections for Hirayama et al. (2000)

Objective changes. Character 7, processus inferior parietalis: The descending process of the parietal of *K. aprix* generally resembles that of *P. quenstedti* and should thus be scored as reduced (0), not present (1) (see Sterli and Joyce 2007).

Character 16, palatal teeth: Palatal teeth are clearly absent (1), not present (0) in all available material of *K. aprix* (see Sterli and Joyce 2007).

Character 25, pterygoid–basioccipital contact: A pterygoid–basioccipital contact is clearly absent (0), not present (1) in all available material of *K. aprix* (see Sterli and Joyce 2007).

Characters 35, 36–44, 47, 49–50: This series of characters pertain to the vertebral morphology of turtles. Although they cannot be scored objectively, all are left as they stand (see critique of Gaffney 1996 for rationale)

Character 45, first thoracic vertebra: This character should be scored 0, not '?' (see character 49 of Rougier *et al.* 1995).

Character 52, acromion process: This character should be scored 0, not '?' (see character 51 of Rougier *et al.* 1995).

Character 56, short digits: As it stands, *K. aprix* is scored as possessing elongate hands (0). However, given the lack of quality material that would clarify this observation, we think it better to score this character as '?.'

Problematic characters. Character 18, pterygoid flange: see critique of Rougier *et al.* 1995

Character 21, processus trochlearis oticum: see critique of Rougier *et al.* 1995

Character corrections for Gaffney et al. (2007)

Objective changes. Character 10, size of processus inferior parietalis: This character should be scored 0, not 1 (see character 7, Hirayama *et al.* 2000).

Character 19, length of the postorbital: Precisely quantifying the size of the postorbital is difficult. However, we feel that the relative size of the postorbital does not differ in any significant way from the plesiomorphic condition seen in

P. quenstedti, and we thus score *K. aprix* as exhibiting the plesiomorphic condition (see Sterli and Joyce 2007) 0, not 1.

Character 26, palatal teeth: This character should be scored 1, not ? (see character 16, Hirayama *et al.* 2000).

Character 37, interpterygoid space: In all well-preserved *K. aprix* specimens, the interpterygoid space is well developed, and we thus score this taxon as exhibiting the plesiomorphic (0) and not the derived (1) condition (see Sterli and Joyce 2007).

Character 55, recessus scalae tympani: This character should be scored 0, not 1 (see character 26, Rougier *et al.* 1995).

Character 61: number of basisphenoid/basioccipital tubercles: This character should be scored 1, not 2 (see character 25, Rougier *et al.* 1995).

Character 63, splenial: The splenial bone is clearly present in a number of specimens of *K. aprix* specimens (Sterli and Joyce 2007) and thus should be scored 0, not ?.

Characters 66–81, 91: This series of characters pertain to the vertebral morphology of turtles. Although they cannot be scored objectively, all are left as they stand (see critique of Gaffney 1996 for rationale)

Character 92, ninth pair of costals: Gaffney *et al.* (1987) originally reported the presence of a ninth pair of costals in *K. aprix*, but we were not able to confirm this in our systematic analysis of all available material. We thus score *K. aprix* as exhibiting the derived (1) and not the plesiomorphic (0) condition.

Character 109, humeropectoral sulcus: The humeropectoral sulcus in all available *K. aprix* material never laps onto the epiplastra, and we thus confidently score thus character as plesiomorphic (0), not '?.'

Problematic characters. Character 28, pterygoid flange: see critique of Rougier *et al.* 1995

Character 50, processus trochlearis oticum: see critique of Rougier *et al.* 1995

Character 51, thickening to the anterior region of the ear: see critique of Rougier *et al.* 1995