MicroRNAs support a turtle + lizard clade

Tyler R. Lyson[1], Erik A. Sperling[1,†], Alysha M. Heimberg[2], Jacques A. Gauthier[1], Benjamin L. King[3] and Kevin J. Peterson[2,*]

[1]*Department of Geology and Geophysics, Yale University, 210 Whitney Avenue, New Haven, CT 06511, USA*
[2]*Department of Biological Sciences, Dartmouth College, Hanover, NH 03755, USA*
[3]*Mount Desert Island Biological Laboratory, Salisbury Cove, ME 04672, USA*
Author for correspondence (kevin.j.peterson@dartmouth.edu).
[†]*Present address: Department of Earth and Planetary Sciences, Harvard University, Cambridge, MA 02138, USA.*

Despite much interest in amniote systematics, the origin of turtles remains elusive. Traditional morphological phylogenetic analyses place turtles outside Diapsida—amniotes whose ancestor had two fenestrae in the temporal region of the skull (among the living forms the tuatara, lizards, birds and crocodilians)—and allied with some unfenestrate-skulled (anapsid) taxa. Nonetheless, some morphological analyses place turtles within Diapsida, allied with Lepidosauria (tuatara and lizards). Most molecular studies agree that turtles are diapsids, but rather than allying them with lepidosaurs, instead place turtles near or within Archosauria (crocodilians and birds). Thus, three basic phylogenetic positions for turtles with respect to extant Diapsida are currently debated: (i) sister to Diapsida, (ii) sister to Lepidosauria, or (iii) sister to, or within, Archosauria. Interestingly, although these three alternatives are consistent with a single unrooted four-taxon tree for extant reptiles, they differ with respect to the position of the root. Here, we apply a novel molecular dataset, the presence versus absence of specific microRNAs, to the problem of the phylogenetic position of turtles and the root of the reptilian tree, and find that this dataset unambiguously supports a turtle + lepidosaur group. We find that turtles and lizards share four unique miRNA gene families that are not found in any other organisms' genome or small RNA library, and no miRNAs are found in all diapsids but not turtles, or in turtles and archosaurs but not in lizards. The concordance between our result and some morphological analyses suggests that there have been numerous morphological convergences and reversals in reptile phylogeny, including the loss of temporal fenestrae.

Keywords: turtle; microRNA; amniote

1. INTRODUCTION

The phylogenetic position of turtles remains labile, owing in part to the fact that, while many primitive cranial features suggest a basal position outside Diapsida [1–3], almost as many derived post-cranial traits support a turtle + Lepidosauria sister relationship [4–7], which results in two weakly supported morphological signals [8]. In addition, morphological analyses are hampered by the highly modified anatomy exhibited by turtles (e.g. shell) and lepidosaurs (e.g. cranial kinesis). Stem Testudines have recently been described [3,7], but the least controversial of these have already acquired many of the distinctive autapomorphies of the crown. Likewise, the most basal unequivocal stem lepidosaurs are the kuehneosaurs, gliding reptiles from the Late Triassic [9]; comparing reptiles modified for parachute gliding with those designed as armoured tanks has proven difficult. As a result, many evolutionary biologists accept the molecular hypothesis that turtles are sister to (or nested within) the archosaurs ([10–12]; but see [13]), a hypothesis that has little morphological support [14,15].

Despite the conflict between the morphological and molecular datasets [8,12], all three hypotheses actually agree on the topology of the reptile tree, they simply disagree on the position of its root (figure 1). Thus, the problem is not one of deciding inter-relationships *per se*, but one of simply determining polarity. Correct rooting in morphological studies relying on highly divergent outgroups is often hampered by difficulty in establishing polarity for individual characters [16], while rooting in traditional molecular phylogenetic analyses is affected by rate and compositional heterogeneity [17]. Given that there is clear homoplasy in reptile morphology irrespective of the true phylogenetic position of turtles [8], and clear rate heterogeneity in molecular sequences in amniotes [12], an alternative data source is needed to correctly root the reptile tree.

Here, we apply a novel molecular dataset, the presence versus absence of specific microRNAs (miRNAs)—genes that encode approximately 22 nucleotide non-coding regulatory RNAs—to this problem. miRNA-based phylogenetics have been successfully applied to many metazoan clades, including vertebrates [18], as miRNAs show a number of characteristics that make them ideal phylogenetic characters, including the fact that new miRNAs are continually added to metazoan genomes through time and, once added, are rarely lost in most metazoan taxa [19–21]. In addition, miRNAs show extreme nucleotide conservation of the mature sequence, and structural considerations based on the requirement to fold into the canonical miRNA hairpin structure make convergence highly unlikely, resulting in little homoplasy [19,20]. Because the acquisition of a novel miRNA family represents the gain of a *de novo* trans-acting gene class, where the outgroup state (absence) can be established with certainty, miRNAs are ideal candidates for delineating the position of the root.

2. MATERIAL AND METHODS

RNA was extracted from single individuals of the turtle *Chrysemys picta bellii*, the lizard *Anolis carolinensis* (both purchased from Kingsnake.com) and the alligator *Alligator mississippiensis* (purchased from Watch Them Grow Reptiles) following standard animal care protocols (IACUC number 2009-11302). Total RNA was extracted as described in Wheeler *et al.* [19]. Small RNA libraries were prepared at the Yale University School of Medicine W. M. Keck Facility
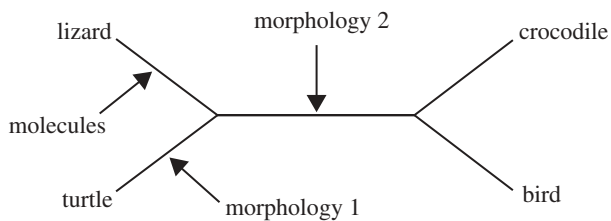
Figure 1. The inter-relationships among the major groups of reptiles. Using an unrooted tree, it is possible to show that although each of the three previous hypotheses concerning turtle relationships—morphology 1 [1−3], morphology 2 [4−7] and the molecular results [10−12]—agree on the topology, they disagree on the position of the root (arrows).

according to manufacturer's instructions, and sequenced on the Illumina Genome Analyzer II platform. The number of reads sequenced per library was 23 765 521 (*C. picta bellii*), 23 488 055 (*A. carolinensis*) and 21 731 314 (*A. mississippiensis*). These reads were analysed using miRMiner [19] to discover previously identified miRNAs (electronic supplementary material, figure S1), and novel miRNA genes specific to *A. carolinensis* were identified with the v. 1 genome assembly using BLASTN from the NCBI BLAST v. 2.2.23+ package (word size = 11, expectation value threshold = 1000).

These were then compared with previously published data in miRBase, coupled with new genomic searches, for the mammals *Ornithorhyncus anatinus* (platypus), *Monodelphis domestica* (opossum) and *Homo sapiens*, the birds *Gallus gallus* (chicken) and *Taeniopygia guttata* (zebrafinch), using *Xenopus laevis* (frog) and the three mammals as the outgroups. A matrix (electronic supplementary material, file S2) of all relevant amniote miRNAs was coded for these taxa, and a maximum parsimony analysis was performed using PAUP 4.0b10 [22] with all characters given equal weight and using the branch and bound search algorithm. Support for each node was measured by calculating the Bremer support [23] using TREEROT v. 3 [24].

3. RESULTS

Two hundred and eighty two miRNA genes were annotated, belonging to 186 miRNA families [20] (electronic supplementary material, file S1), with 77 new miRNA families discovered that appear to be specific to *Anolis* (electronic supplementary material, file S1). As expected, given our understanding of how miRNAs evolve in most taxa [19−21], only two families appear to have been secondarily lost in the lizard: miR-457 and miR-2184. Similarly, in the turtle, reads were not detected for miR-147 and -208, and in the alligator, reads were not detected for miR-726 and -727 (the latter a potential shared loss with birds; electronic supplementary material, file S1). Thus, out of the 100 expected miRNA families in the last common ancestor of Tetrapoda (electronic supplementary material, file S3), each of these species has 98, making the occurrence of secondary loss in all these taxa no more than 2 per cent.

The phylogenetic analysis resulted in a single most parsimonious tree with a tree length of 36 (CI = 0.97; RI 0.99). Each of the reptile clades identified in this study is characterized by at least one miRNA acquisition—the alligator shares one miRNA with the birds (miR-1791), supporting the monophyly of Archosauria, and all reptiles analysed herein share miR-1677 (figure 2a). Importantly, we find that turtles and lizards share four of the 77 unique miRNA gene families identified in *Anolis* that are not found in any

other organisms' genome or small RNA library (figure 2a; electronic supplementary material, file S1). Thus, these sequences are indeed restricted to the reptiles, and a phylogenetic analysis of the presence/absence of miRNAs families in these taxa (electronic supplementary material, file S2) unambiguously supports a turtle + lizard sister group relationship (figure 2b), as no miRNAs were found in all diapsids but not in turtles, or in turtles and archosaurs but not in lizards. Other nodes including Amniota, Mammalia, Theria and Aves are each supported by one or more unique miRNA (figure 2b).

4. DISCUSSION

Turtles as sister group to lizards are recovered in some morphological analyses [4−7] and this is supported by some, mostly postcranial, characters including the fusion of the astragalus and calcaneum in postnatal ontogeny [25]. However, a diapsid affinity also requires several morphological reversals in turtles, including the closure of the temporal fenestrae [25]. In addition, it suggests that molecular analyses that recover an archosaur affinity might be the result of a systematic artefact caused by the attraction of the long-branched lizards towards the outgroup [8].

The consilience between at least some morphological apomorphies and the evolutionary acquisition of miRNAs suggests the validity of the sister group relationship between lepidosaurs and turtles. Hence, we propose 'Ankylopoda' ('fused foot' in reference to the fused ankle bones—astragalus and calcaneum—shared by lepidosaurs and turtles) as the name of the Lepidosauria + Testudines crown clade, which is defined as the last common ancestor of *Chrysemys picta* and *A. carolinensis* and all of its descendants living or extinct (electronic supplementary material, file S4). If Lepidosauria + Testudines is indeed a monophyletic group, then this implies significant convergence in molecular (nuclear and mitochondrial), developmental and morphological characters, including the loss of temporal fenestrae, in the early evolutionary history of turtles. Further, our study suggests that lizards are the appropriate outgroup comparison for understanding the origin of the turtle body plan.

1 Gauthier, J., Kluge, A. G. & Rowe, T. 1988 Amniote phylogeny and the importance of fossils. *Cladistics* **4**, 105−209. (doi:10.1111/j.1096-0031.1988.tb00514.x)

2 Lee, M. S. Y. 1993 The origin of the turtle body plan: bridging a famous morphological gap. *Science* **261**, 1716−1720. (doi:10.1126/science.261.5129.1716)

(a)

```
Aca miR-1677 dG = -31.60

TGG|  A        A     CC-           A   A    GAT
   TATA AGTGTTA TCCT   CCTATTGAA TCA TGG   G
   GTAT TCACAGT AGGA   GGATAACTT AGT ACC   C
GTA^  A        -     TGA           C   C    GTC

Gga-miR-1677 TGACTTCAGTAGGAGCAGGATT
Tgu-miR-1677 TGACTTCAATAGGAGCAGGATT
Ami-miR-1677 TGACTTCACTGAGCTTTGGACA
Aca-miR-1677 TGACTTCAATAGGAGTAGGATG
Cpi-miR-1677 TGACTTCAGTAGAGCCAGGAT-


Aca-miR-5390 Initial dG = -41.30

CTT| C A            C          A   A    TG A
   GT GT CAGGTTGGATGTC CTTATCCG AAT TCTG  G C
   CA CA GTCCGACCTGCAG GAATAGGC TTA AGGT   T G
TTT^ C C            A          C   -    GT A

Aca-miR-5390 GAATTCCGGATAAGAGACGTC
Cpi-miR-5390 GAATTCCGGATTACAGACGT-


Tgu miR-1791 dG = -45.00

AGC|   AC    TG              CCA     TGA
   TGATGC  CATGT  GGCTGCATCAGTCATG  TGTTA   \
   ACTGTG  GTACA  TCGACGTAGTCAGTGT  GCAAT    A
AAC^   AA    GT              AGC     CCA

Gga-miR-1791 TGGGCTGCATCAGTCATGCCATG
Tgu-miR-1791 TGGGCTGCATCAGTCATGCCATG
Ami-miR-1791 TGGGCTGCCTCAGTCATGCCAT-
```
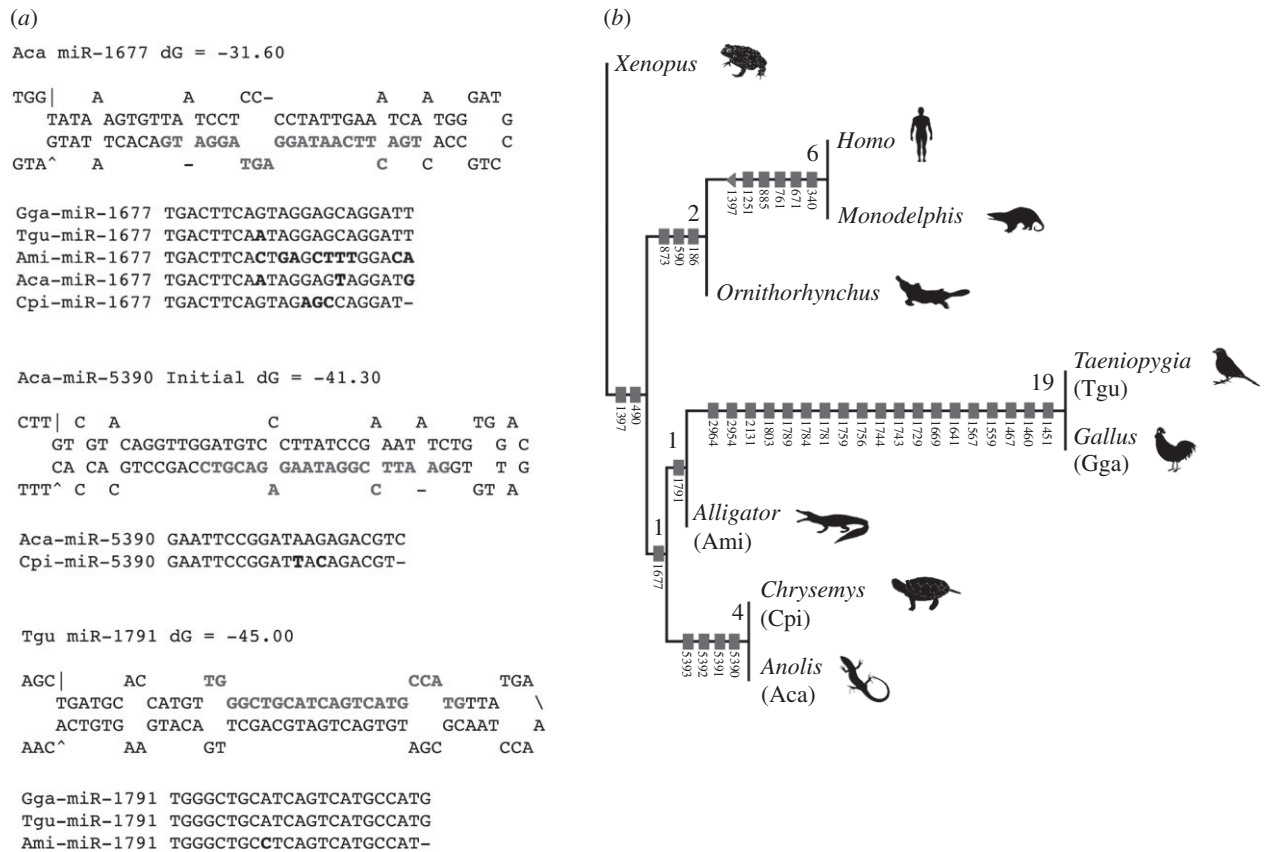
(b)

Figure 2. microRNAs support a turtle–lizard relationship. (a) Structures and alignments of the mature sequences for three of the 35 families analysed phylogenetically, the reptile-specific miR-1677, the archosaur-specific miR-1791 and miR-5390, a novel miRNA shared between *Anolis carolinensis* and *C. pictus*. Mature sequences within the context of the miRNA hairpin are shown in grey; changes in the mature sequence with respect to the reference sequence, either chicken or lizard, are shown in bold. (b) To arbitrate among these competing hypotheses (see figure 1), eight amniote taxa were scored for the presence/absence of 35 miRNA families with the frog *Xenopus laevis* as the outgroup. Using a combination of small RNA library sequencing coupled with genomic searches (electronic supplementary material, file S1), we find that the turtle *C. picta* shares four miRNA families with the lizard *A. carolinensis* that are not found elsewhere in the animal kingdom, supporting the rooting position of 'morphology 2' in figure 1 [4–7]. Bremer support indexes are indicated at each node. Character changes on branch: grey boxes, acquisition of miRNA family; filled triangle, loss of miRNA family.

3 Lyson, T. R., Bever, G. S., Bhullar, B.-A. S., Joyce, W. G. & Gauthier, J. A. 2010 Transitional fossils and the origin of turtles. *Biol. Lett.* **6**, 830–833. (doi:10.1098/rsbl.2010.0371)

4 Rieppel, O. & deBraga, M. 1996 Turtles as diapsid reptiles. *Nature* **384**, 453–455. (doi:10.1038/384453a0)

5 Müller, J. 2004 The relationships among diapsid reptiles and the influence of taxon selection. In *Recent advances in the origin and early radiation of vertebrates* (eds M. V. H. Arratia & R. Cloutier), pp. 379–408. München, Germany: Verlag Dr. Friedrich Pfeil.

6 Hill, R. V. 2005 Integration of morphological data sets for phylogenetic analysis of Amniota: the importance of integumentary characters and increased taxonomic sampling. *Syst. Biol.* **54**, 530–547. (doi:10.1080/10635 150590950326)

7 Li, C., Wu, X.-C., Rieppel, O., Wang, L.-T. & Zhao, J. 2008 An ancestral turtle from the Late Triassic of southwestern China. *Nature* **456**, 497–501. (doi:10.1038/nature07533)

8 Lee, M. S. Y., Reeder, T. W., Slowinski, J. B. & Lawson, R. 2004 Resolving reptile relationships: molecular and morphological markers. In *Assembling the tree of life* (eds J. Cracraft & M. J. Donoghue), pp. 451–467. New York, NY: Oxford University Press.

9 Evans, S. E. & Jones, M. E. H. 2010 The origin, early history and diversification of Lepidosauromorph reptiles. In *New aspects of mesozoic biodiversity* (ed. S. Bandyopadhyay), pp. 27–44. Berlin, Heidelberg: Springer.

10 Hedges, S. B. & Poling, L. L. 1999 A molecular phylogeny of reptiles. *Science* **283**, 998–1001. (doi:10.1126/science.283.5404.998)

11 Rest, J. S., Ast, J. C., Austin, C. C., Waddell, P. J., Tibbetts, E. A., Hay, J. M. & Mindell, D. P. 2003 Molecular systematics of primary reptilian lineages and the tuatara mitochondrial genome. *Mol. Phylogenet. Evol.* **29**, 289–297. (doi:10.1016/S1055-7903(03)00108-8)

12 Hugall, A. F., Foster, R. & Lee, M. S. Y. 2007 Calibration choice, rate smoothing, and the pattern of tetrapod diversification according to the long nuclear gene RAG-1. *Syst. Biol.* **56**, 543–563. (doi:10.1080/10635150701477825)

13 Becker, R. E., Valverde, R. A. & Crother, B. I. 2011 Proopiomelanocortin (POMC) and testing the phylogenetic position of turtles (Testudines). *J. Zool. Syst. Evol. Res.* **49**, 148–159. (doi:10.1111/j.1439-0469.2010.00589.x)

14 Rieppel, O. 2000 Turtles as diapsid reptiles. *Zool. Scr.* **29**, 199–212. (doi:10.1046/j.1463-6409.2000.00039.x)

15 Lee, M. S. Y. 2001 Molecules, morphology, and the monophyly of diapsid reptiles. *Contrib. Zool.* **70**, 121–138.

16 Rota-Stabelli, O., Campbell, L., Brinkmann, H., Edgecombe, G. D., Longhorn, S. J., Peterson, K. J.,

Pisani, D., Philippe, H. & Telford, M. J. 2010 A congruent solution to arthropod phylogeny: phylogenomics, microRNAs and morphology support monophyletic Mandibulata. *Proc. R. Soc. B* **277**, 298–306. (doi:10.1098/rspb.2010.0590)

17 Bergsten, J. 2005 A review of long-branch attraction. *Cladistics* **21**, 163–193. (doi:10.1111/j.1096-0031.2005.00059.x)

18 Heimberg, A. M., Cowper-Sal Lari, R., Sémon, M., Donoghue, P. C. & Peterson, K. J. 2010 microRNAs reveal the interrelationships of hagfish, lampreys, and gnathostomes and the nature of the ancestral vertebrate. *Proc. Natl Acad. Sci. USA* **107**, 19 379–19 383. (doi:10.1073/pnas.1010350107)

19 Wheeler, B. M., Heimberg, A. M., Moy, V. N., Sperling, E. A., Holstein, T. W., Heber, S. & Peterson, K. J. 2009 The deep evolution of metazoan microRNAs. *Evol. Dev.* **11**, 50–68. (doi:10.1111/j.1525-142X.2008.00302.x)

20 Sperling, E. & Peterson, K. 2009 microRNAs and metazoan phylogeny: big trees from little genes. In *Animal evolution—genomes, trees and fossils* (eds M. J. Telford & D. T. J. Littlewood), pp. 157–170. Oxford, UK: Oxford University Press.

21 Campo-Paysaa, F., Sémon, M., Cameron, R. A., Peterson, K. J. & Schubert, M. 2011 miRNA complements in deuterostomes: origin and evolution of miRNAs. *Evol. Dev.* **13**, 15–27. (doi:10.1111/j.1525-142X.2010.00452.x)

22 Swofford, D. L. 2002 *PAUP** Phylogenetic analysis using parsimony (* and other methods); v.4.0b10 for Macintosh. Sunderland, MA: Sinauer Associates.

23 Bremer, K. 1994 Branch support and tree stability. *Cladistics* **10**, 295–304. (doi:10.1111/j.1096-0031.1994.tb00179.x)

24 Sorenson, M. D. & Franzosa, E. A. 2007 *TREEROT v3.0*. Boston, MA: Boston University.

25 deBraga, M. & Rieppel, O. 1997 Reptile phylogeny and the interrelationships of turtles. *Zool. J. Linn. Soc.* **120**, 281–354. (doi:10.1111/j.1096-3642.1997.tb01280.x)