Molecular Phylogenetics

Jordan Lillard

April 28, 2019

**Introduction**

With the earliest known turtle species dating back to over 220 million years ago, they are among the oldest reptiles in the world. Their history is ancient, even more so than both snakes and crocodiles. The purpose of this analysis is to observe how a species with an extensive history, turtles, have evolved through the ages.

Turtles belong to the reptile order *Testudines.* Since they are the only living anapsid amniotes, their evolutionary relationship is essential in determining the basal condition and early transformations of many amniotic characters (Shaffer, 1997). Recent molecular analyses have showed some discrepancies with the previously established morphological hypotheses which led Crawford to perform a genome-scale analysis of turtle phylogeny (Crawford, 2014).

In the research below, three different models - Maximum Parsimony, Distance Analysis, and Bayesian Analysis - are utilized to analyze the entire genome to construct a single resolved phylogeny.

**Materials & Methods**

*Data*

The data was acquired from a Crawford’s 2014 study. The data includes 32 turtle/ingroup taxonomic units (OTUs) and six outgroups (three lepidosaurs [Sphenodon and two squamates], two archosaurs [a crocodilian and a bird], and one mammal [human]; Table 1) for analysis. The 32 ingroup OTUs represent all of the major lineages of turtles: Pleurodira, Trionychia, Testudinoidea, Chelonioidea, Chelydridae, and Kinosternoidea. The data also includes samples from all 14 traditionally accepted families (Turtle Taxonomy Working Group, 2014). Crawford created a database of UCE loci and prepared FASTA files for alignment. Using MAFFT, a monolithic FASTA file was aligned. The resulting alignments were trimmed using phyluce. The generated alignment contained all 36 taxa at every locus. This resulting alignment was used for further analysis.

*Maximum Parsimony*

I performed maximum parsimony analysis using PAUP\*. 100 replications were done using a heuristic search to find the maximum parsimony tree. All trees were saved and, the 50% Majority Rule Consensus tree was found and reconstructed with FigTree (Figure 1). The tree was rooted with the mammals following the approach of a recent analysis that confirms the archosaurs affinities of Testudines (Field, 2014).

*Distance Analysis*

Using PAUP\* again, pairwise distances were used to estimate phylogenetic trees with the neighbor-joining clustering algorithm and the UPGMA clustering algorithm. The distance matrix was calculated from the nucleotides, under Jukes-Cantor model of sequence evolution. The resulting trees were reconstructed with FigTree (Figure 2 and Figure 3).

*Maximum Likelihood*

I referred to the paper’s maximum likelihood analysis of the data. RAxML version 8.0 with the “GTRGAMMA” option and 10,000 bootstrap replicates (Figure 4).

*Bayesian analysis*

I performed Bayesian analysis using MrBayes version 3.2.7a for 500,000 iterations (4 chains; burn-in 25%; thinning: 500. The resulting tree was reconstructed with FigTree (Figure 5).

**Results**

*Maximum Parsimony*

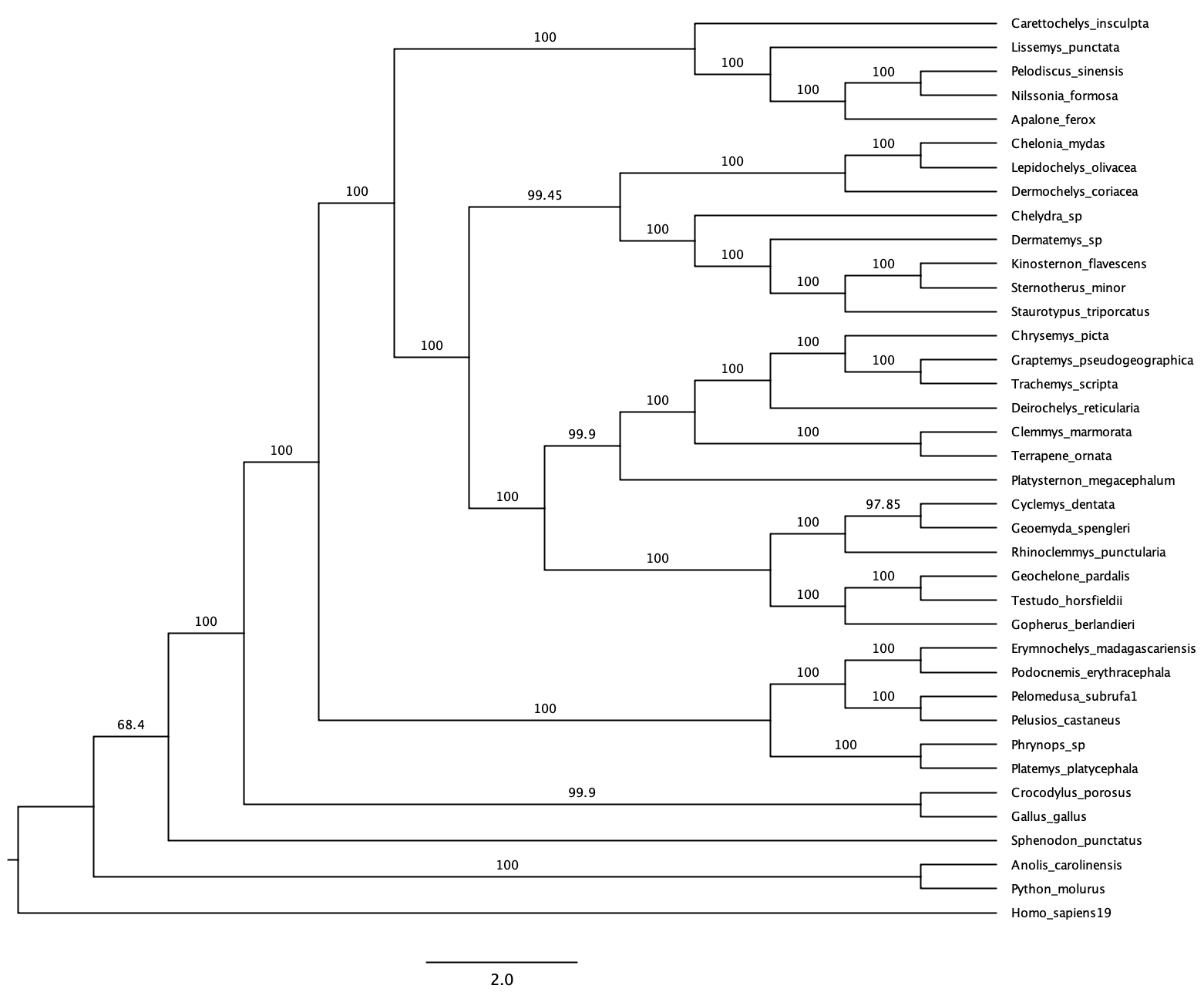
In the analysis, a total of 780,308 rearrangements were tried and one tree was retained with a score of 74,749. Because only one tree was retained, I decided to rerun the heuristic search with 1,000 replicates. The same tree was retained with a maximum parsimony score of 74,749. The bootstrap method with the heuristic search was then done with both 100 and 1,000 replicate and yielded the same tree both times. This tree is consistent with the Maximum Likelihood tree from Crawford’s paper (Figure 4).

*Distance Analysis*

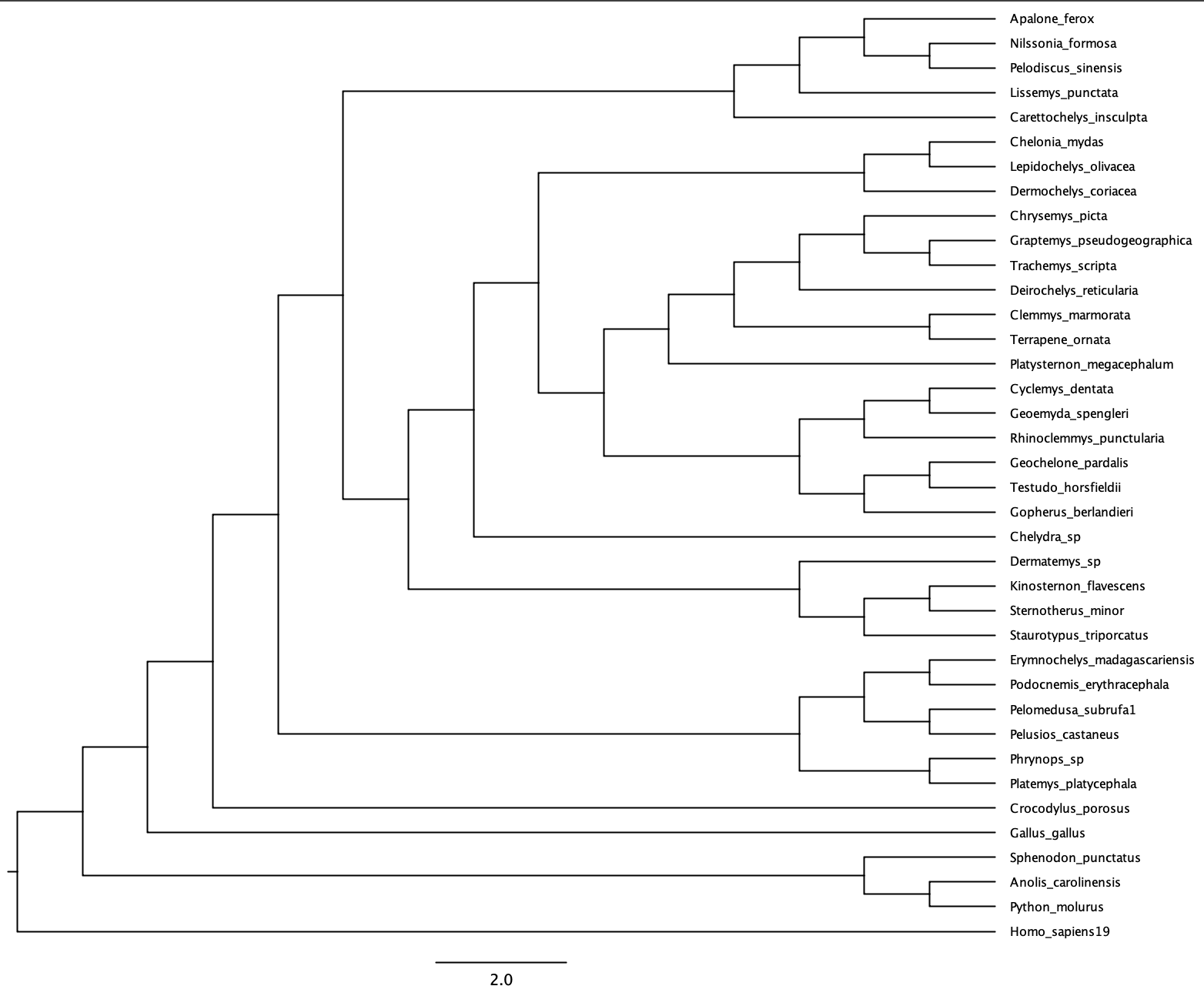
**Discussion**

While all five trees do generally show the same phylogenetic relationships, they are not exactly the same.

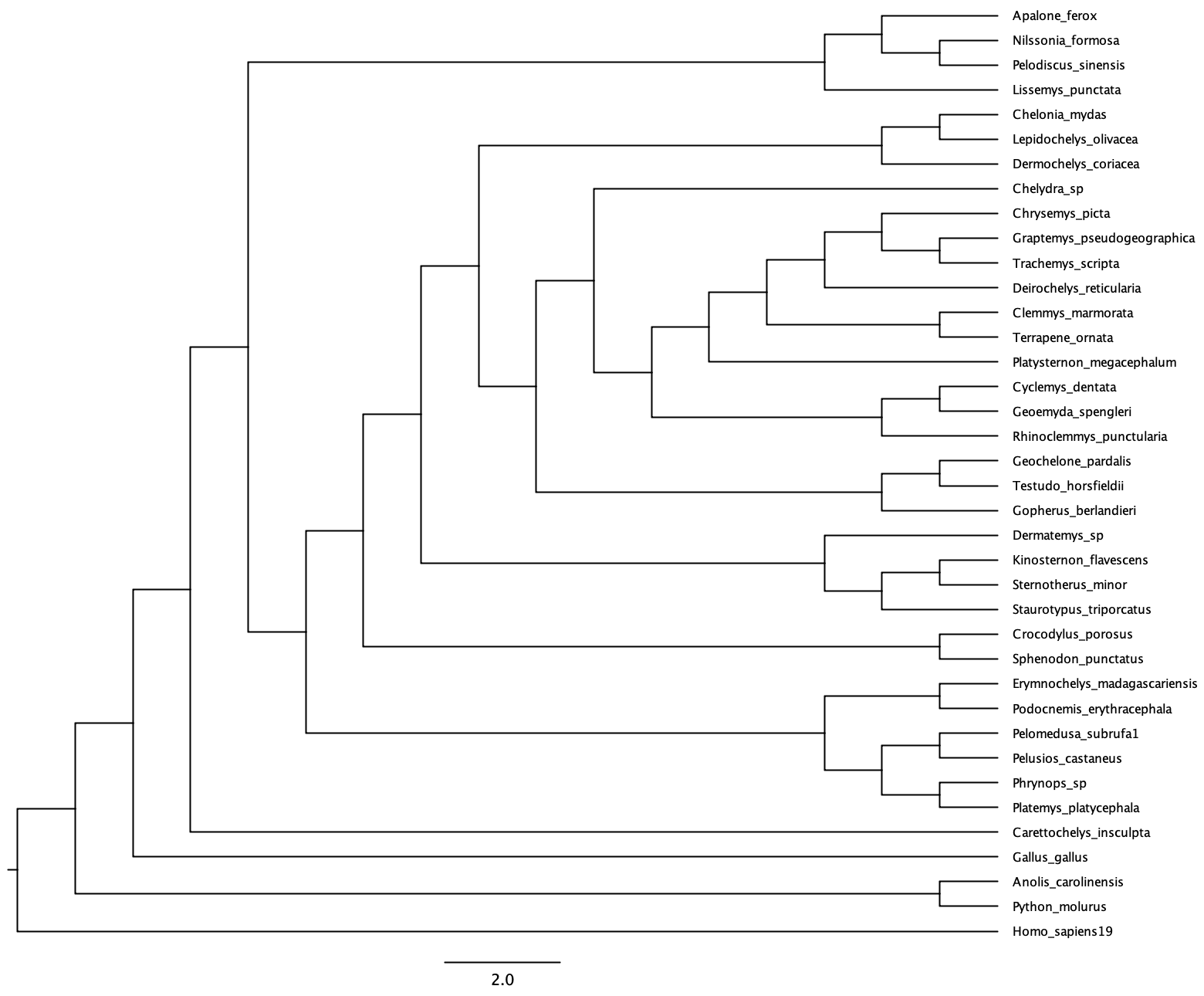
However, to get a full grade, I expect you to ask (and at least attempt to answer) some additional questions based on the phylogenetic trees that you’ve inferred (e.g., correlated evolution, natural selection, testing different topologies, molecular clock analysis, etc.).



Bootstrap 50% majority rule



NJ

UPGMA

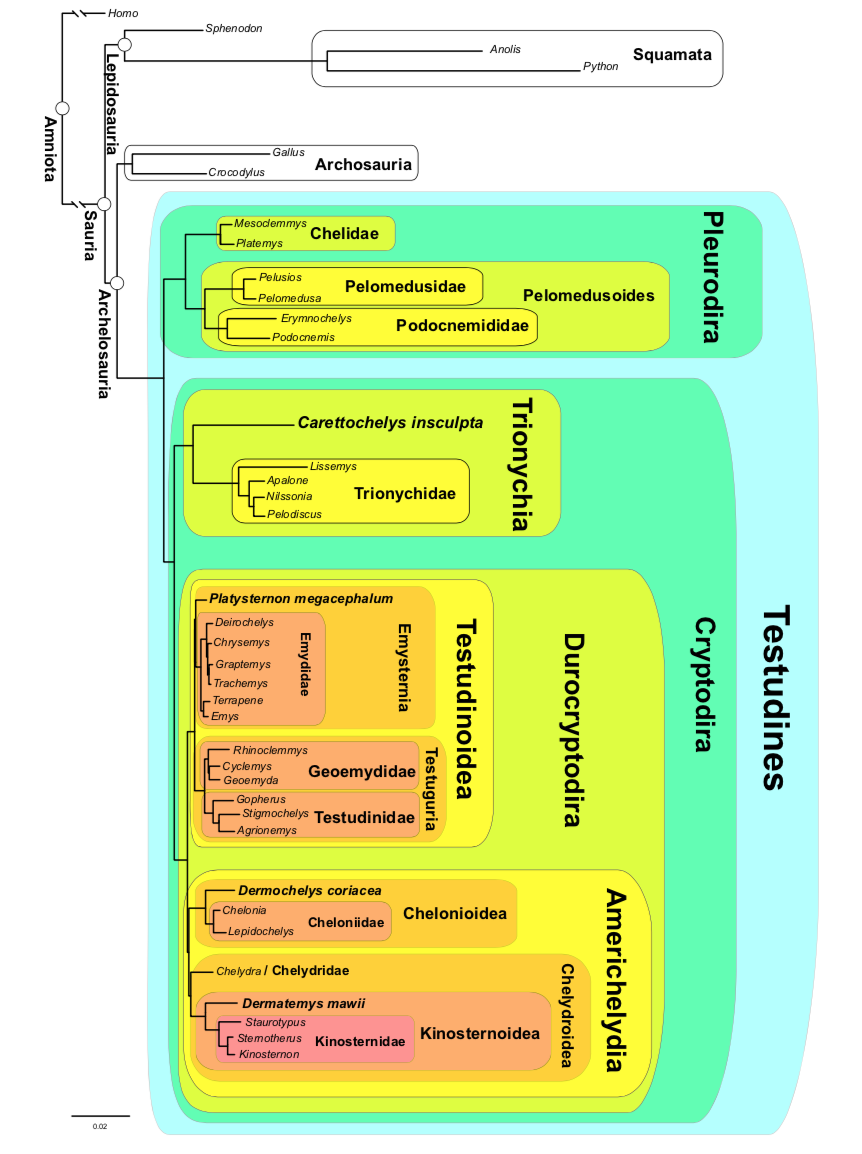
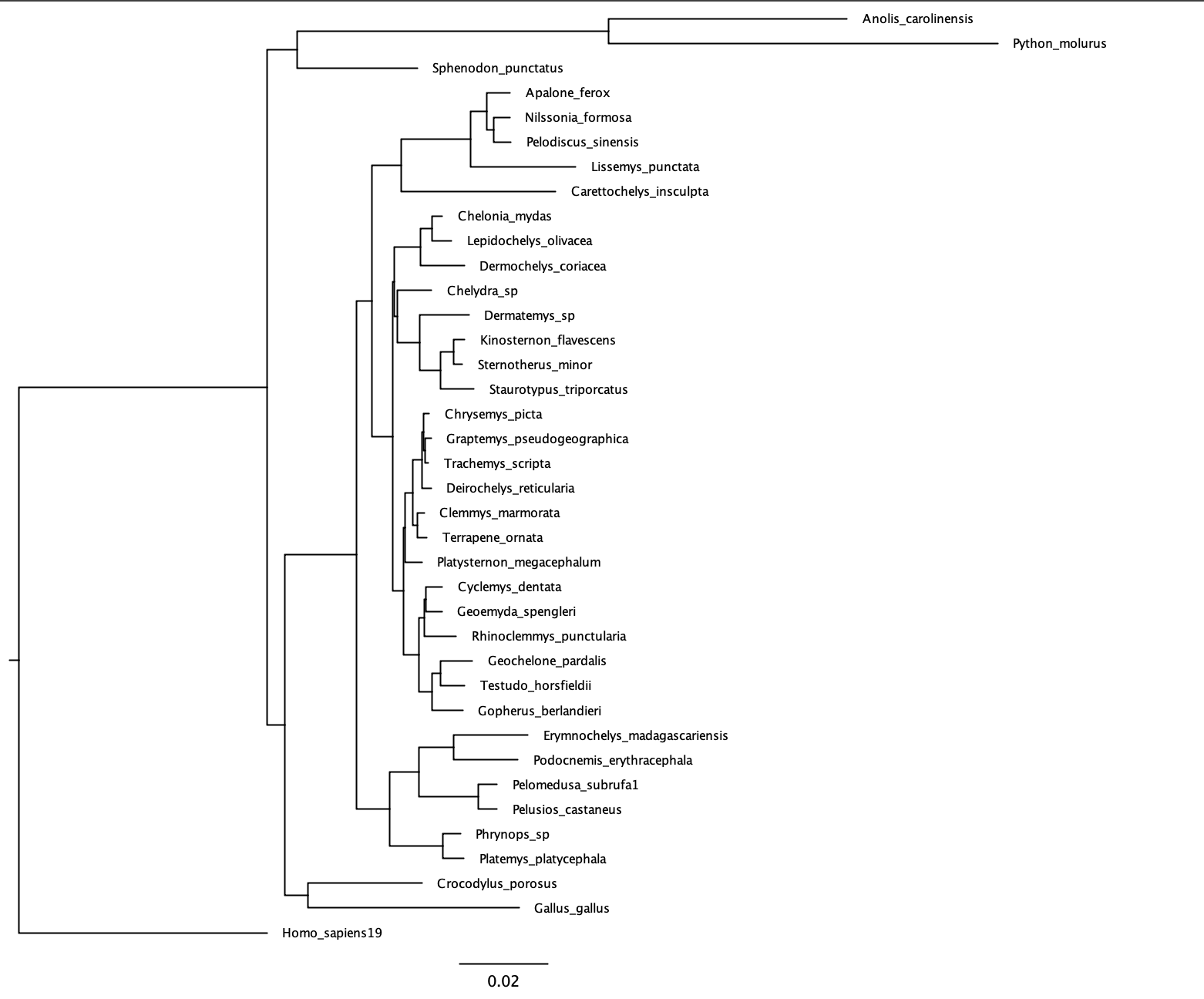


Figure 4. Phylogenetic hypothesis based on RAxML analysis of UCE data showing phylogenetically defined crown clades of turtles (Testudines) (Crawford, 2014).



MrBayes

**References**

**Shaffer, H. B., Meylan, P., & Mcknight, M. L. (1997). Tests of Turtle Phylogeny: Molecular, Morphological, and Paleontological Approaches. *Systematic Biology,* *46*(2), 235. doi:10.2307/2413622**

Rambaut A, Drummond AJ, Xie D, Baele G and Suchard MA (2018) Posterior summarisation in Bayesian phylogenetics using Tracer 1.7. *Systematic Biology*. **syy032**. [doi:10.1093/sysbio/syy032](https://doi.org/10.1093/sysbio/syy032)

Turtle Taxonomy Working Group, 2014. Turtles of the world, 7th edition: annotated checklist of taxonomy, synonymy, distribution with maps, and conservation status. Chelonian Res. Monogr. 5, 329–479.

Field, D.J., Gauthier, J.A., King, B.L., Pisani, D., Lyson, T.R., Peterson, K.J., 2014. Toward consilience in reptile phylogeny: miRNAs support an archosaur, not lepidosaur, affinity for turtles. Evol. Dev.. http://dx.doi.org/10.1111/ede.12081.