Morgan R. Muell

EEOB 563 Final Project

Dr. Dennis Lavrov

12 April 2018

Phylogenetic relationships among aquatic reptiles (Family: Emydidae): a comparative methodological approach

Introduction

Throughout the wide world of biodiversity, perhaps no organism is more recognizable than the turtle. Turtles have been important symbols in human culture for millenia, and continue to be revered today. However, the picture of turtle evolution is not so clear. A plethora of studies from phylogeneticists have attempted to unveil the evolutionary relationships among various groups of turtles, but not all have been clear in their conclusions. Some find variations in topology when using different genes, while others do not. Furthermore, debate exists on whether or not mitochondrial DNA sequences should continue to be used, or whether a combination of mitochondrial and nuclear DNA would provide the best picture of evolutionary relationships among extant turtles. Much work remains to be done to uncover turtle phylogeny.

It is particularly important to focus on phylogenetic work with turtles because of the perilous conservation status they are currently facing. Due to the onset of global climate warming and given that turtles have temperature-dependent sex determination, turtle populations are facing serious risks for recruitment. In addition, anthropogenic disturbance and increased predator abundance contribute to the declines of many turtle populations, both marine and freshwater. Cryptic species and subspecies that have yet to be unveiled by phylogenetic studies are unable to receive government and conservation protection. In order to advocate for the protection and conservation of turtles in the timely manner required by the changing of the current climate, it is imperative that more phylogenetic research be conducted specifically on vulnerable taxa, including turtles.

The Family Emydidae are a group of freshwater pond turtles that are distributed worldwide. Some work has been done regarding their phylogeny, but many insights remain to be gained within this group. I investigated the evolutionary relationships among representative species in Family Emydidae. I compared the topologies and support for trees produced using Bayesian methods versus maximum likelihood. Within my Bayesian analysis, I compared topologies and support for trees produced using the cytB mitochondrial gene and the ND4 mitochondrial gene. I performed my own Bayesian analysis on a dataset partially collected by Fritz et al. (2011). The cytB gene has been found as a good gene for comparing taxa on higher levels (Lenk et al. 1999). Further, turtles are a lineage historically known for having slowly evolving mitochondrial genes (Lamb et al. 1994), so the use of a rapidly evolving gene such as ND4 is useful for phylogenetic study (Feldman and Parham 2002). I hypothesized that Bayesian methods would produce a different tree topology than found in Fritz et al (2011) through maximum likelihood. I also hypothesized that the tree topology produced from a Bayesian analysis of the ND4 gene will produce a different tree topology than Bayesian analysis on a cytB gene because ND4 is a more rapidly evolving gene.

Methods

*Gene selection and Alignment*

All gene sequences were obtained from GenBank using accession numbers obtained from supplementary information in Fritz et al. (2011). Only one individual was chosen from their dataset to represent each species or subspecies of Emydid in the dataset, and I chose to use the first individual listed in supplementary material containing accession numbers in Fritz et al. (2011), with the exception of *Chrysemys picta dorsalis* (referred to as *Chrysemys picta* in our analyses), in which I used the accession numbers of the second individual. I aligned the sequences using Mafft. and converted the outputted FASTA files to NEXUS files using an online converter.

*Phylogenetic Analyses*

I used MrBayes to carry out all of my Bayesian analyses. I specified a GTR model with a gamma rate of substitution as my prior. I had flat Dirichlet priors on the six exchangeability parameters and the four base frequencies. The shape parameter was set as exponential with a lambda rate equal to 0.05. Multi-chain Monte Carlo analyses were utilized to sample the model distributions. I ran multi-chain Monte Carlo analyses for one million generations, sampling frequencies every 1000 generations. I ran two independent analyses with four chains each. I performed my Bayesian analyses with these exact settings to produce both the cytB gene and the ND4 gene. Tree topologies and support probabilities were visualized and mapped using FigTree version 1.4.3.

Results

Figures 1 and 2 summarize the tree topologies for the cytB and ND4 mitochondrial genes resulting from Bayesian analysis. Strong support exists at every node in the cytB mitochondrial gene tree, with most nodes having probabilities of 1 (Figure 1). The lowest observed probability in the cytB tree topology was found for the sister grouping between the most recent common ancestors of *Trachemys scripta elegans* and *Trachemys gaigeae*, and *Trachemys decussata* *decussata* and *Trachemys decussta angusta*, respectively (probability=.6831; Figure 1).

The second lowest observed probability in the cytB tree topology was 0.6831, between *Trachemys venusta panamensis* and *Trachemys venusta grayi*. All other nodes had probabilities above 0.8. Many nodes also showed strong support in the ND4 mitochondrial gene topology (Figure 2). A polytomy exists among the *Pseudemys* and *Graptemys* genera in the ND4 tree. These polytomies were not observed in the cytB mitochondrial gene tree, whose only polytomy occurred between *Trachemys scripta emolli*, *Trachemys venusta panamensis*, and *Trachemys* *venusta grayi*.

Genera were grouped together closely on the tree topology in both the cytB and ND4 mitochondrial gene trees. Species in the *Malaclemys* and *Chrysemys* genera were grouped together in the same way in the cytB mitochondrial gene tree and the ND4 mitochondrial gene tree. Slight variation exists in the grouping of *Trachemys* species and subspecies in topologies, although subtle. Maximum likelihood analysis from Fritz et al. (2011) reveals slightly different topology compared to both the cytB and ND4 mitochondrial gene trees shown in Figure 1 and Figure 2. One striking difference is the grouping of the two species in the *Malaclemys* genus as sister to the *Graptemys* genus, whereas our Bayesian analysis found the *Malaclemys* genus to be sister to the *Trachemys* genus in topologies of both the cytB mitochondrial gene and the ND4 mitochondrial gene. All other aspects of tree topology were very similar between the Bayesian analyses conducted in this study and maximum likelihood analyses conducted by Fritz et al. (2011).

Discussion

*Phylogenetic analysis*

I investigated the comparative methodologies of using two different mitochondrial genes for analysis in a Bayesian framework, and compared the results to a previous study that used maximum likelihood methods on the same dataset. Strong support exists for the monophyly of each of the genera in analyses of both the cytB and ND4 mitochondrial genes, including *Trachemys*, consistent with findings in Fritz et al. (2011). As a whole, few differences existed between the cytB and ND4 mitochondrial gene tree topologies, excluding a few minor polytomies and node supports (Figure 1; Figure 2). Turtles have been characterized as having slowly evolving mitochondrial genes in general, regardless of how fast individual mitochondrial genes may evolve, and this may explain the lack of difference. The biggest difference is the grouping of all Emydid species (excluding the outgroup) as a monophyletic group in the cytB tree topology, whereas the ND4 tree topology fails to resolve a polytomy between the outgroup *Deirochelys reticularia*, a group containing the *Chrysemys* and *Pseudemys* genera, and a group containing the *Malaclemys*, *Graptemys*, and *Trachemys* genera. It is unclear why this grouping difference occurred. Further, few differences existed between maximum likelihood analyses in Fritz et al. (2011) and both Bayesian analyses in this study. Given these results, at present I cannot recommend either gene or methodology above another for use in Emydid phylogenetic analysis. While each method produces a slightly different tree topology, they all produce largely similar results.

*Potential improvements and future directions*

The study could have been improved by using stepping-stone analysis during Bayesian analyses and comparing results with multi-chain Monte Carlo. More models could have also been tested within the Bayesian framework to delineate which model best the data collected by Fritz et al. (2011) of Emydid molecular sequences. Additionally, we could have potentially attempted to replicate the maximum likelihood analysis done in RAxML by Fritz et al. (2011) in our analysis to check for consistencies, and compared our Bayesian results to the replicated work. Many future directions can be taken to build upon this analysis. Many more taxa could be included from within the Emydidae to evaluate how other groups within Emydidae fit within the groups specified in this study. Furthermore, more groups from within a wider range of the turtle lineage could be added into the study to expand the scope. More genes could be compared for usefulness and consistency, specifically nuclear genes. Some scholars argue that using mitochondrial data can lead to misleading results (Wiens et al. 2010), and a comparison between the cytB and ND4 genes and other nuclear genes within the Bayesian framework would provide an adequate means of assessing this argument. Lastly, stepping-stone sampling could be evaluated and compared to multi-chain Monte Carlo sampling in Bayesian analysis to look for potential differences in results.

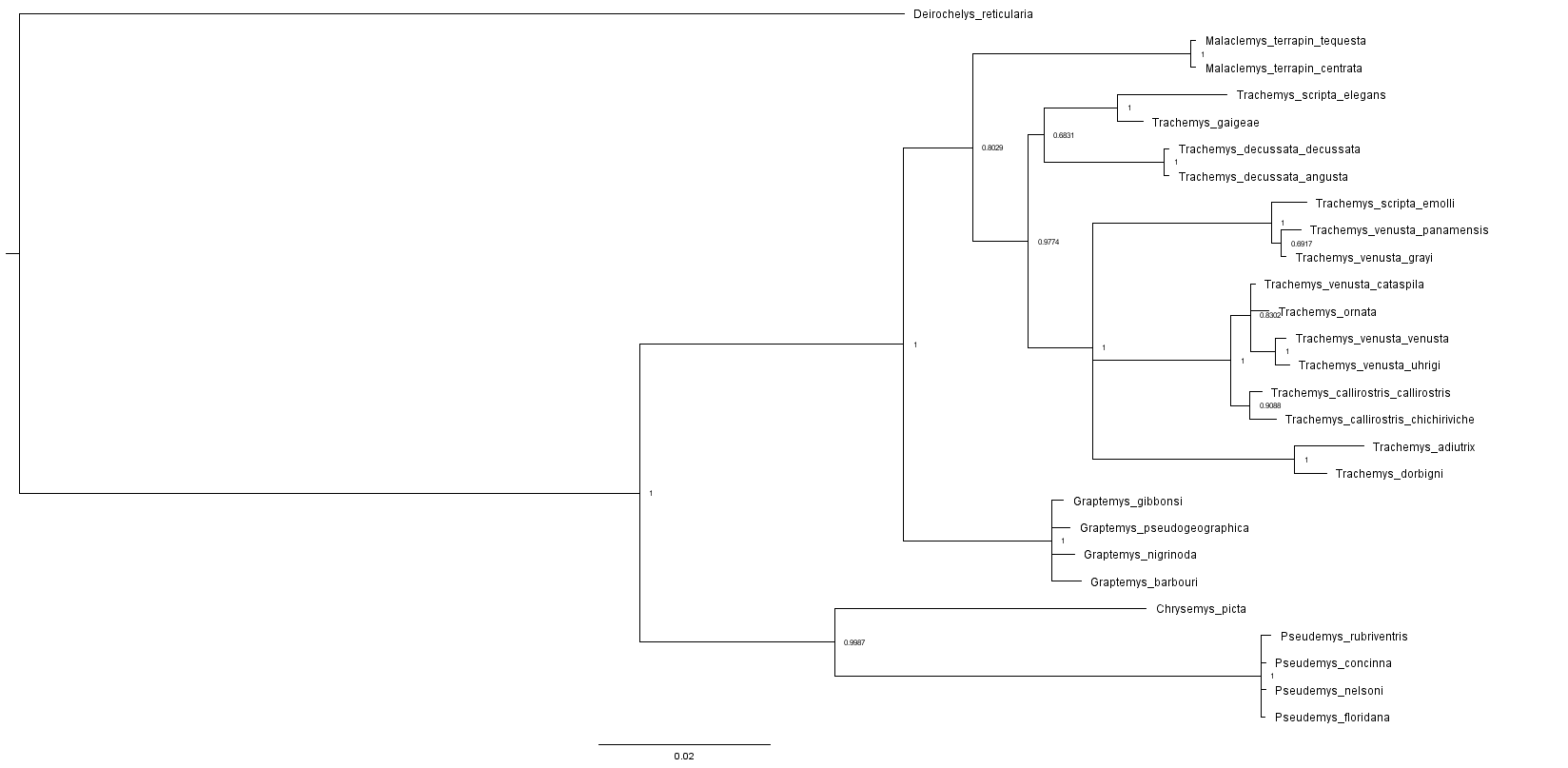


Figure 1. Assembled tree using Bayesian analysis on the cytb mitochondrial gene. Numbers on nodes correspond to probabilities, and the root is indicated by a tick mark on the left.

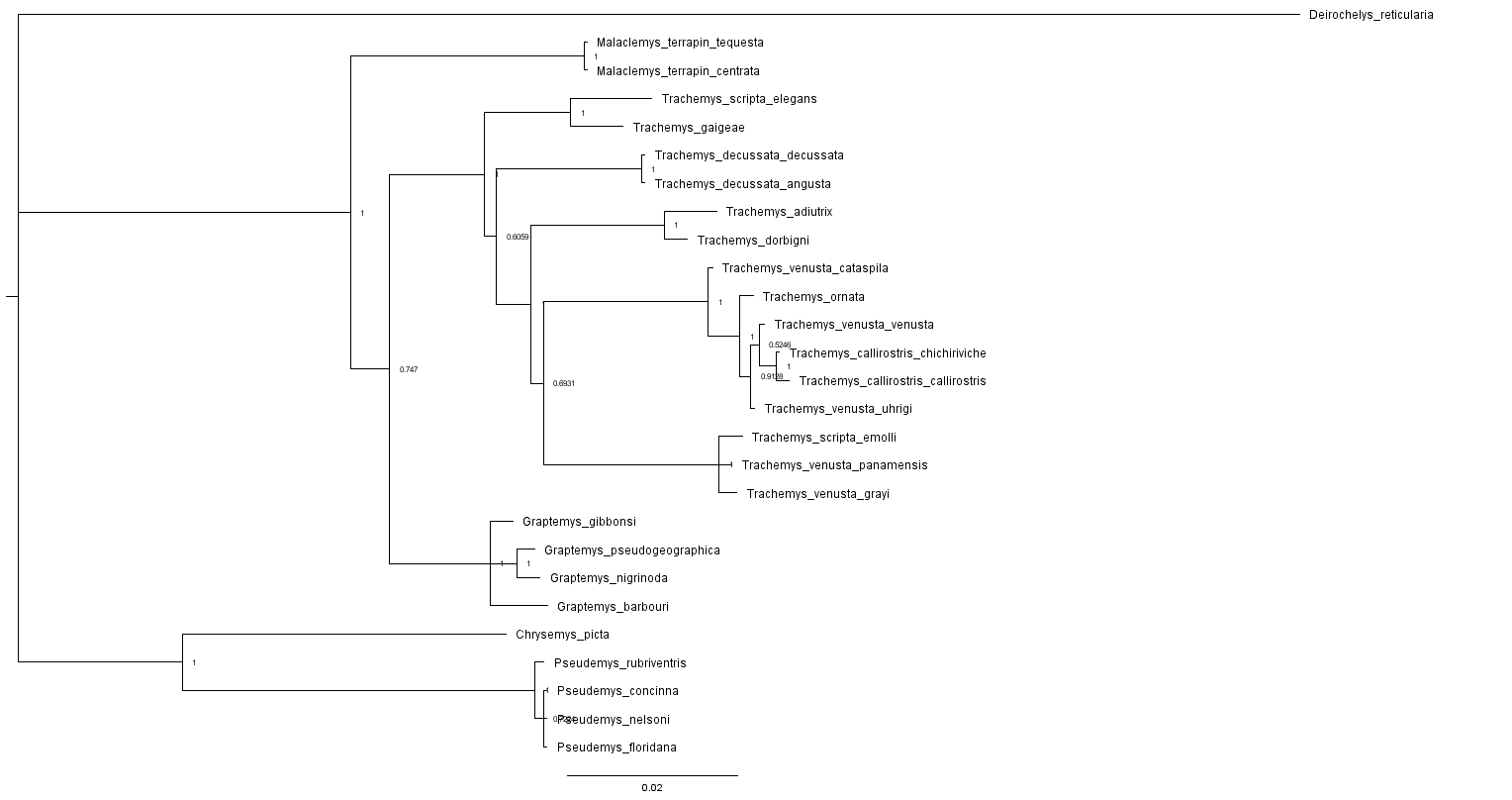


Figure 2. Assembled tree using Bayesian analysis on the ND4 mitochondrial gene. Numbers on nodes correspond to probabilities, and the root is indicated by a tick mark on the left.

References

Bickham, J. W., Lamb, T., Minx, P., & Patton, J. C. (1996). Molecular systematics of the genus Clemmys and the intergeneric relationships of emydid turtles. *Herpetologica*, 89-97.

Feldman, C. R., & Parham, J. F. (2002). Molecular phylogenetics of emydine turtles: taxonomic revision and the evolution of shell kinesis. *Molecular phylogenetics and evolution*, *22*(3), 388-398.

Lamb, T., Lydeard, C., Walker, R. B., & Gibbons, J. W. (1994). Molecular systematics of map turtles (Graptemys): a comparison of mitochondrial restriction site versus sequence data. *Systematic Biology*, *43*(4), 543-559.

Lenk, P., Fritz, U., Joger, U., & Wink, M. (1999). Mitochondrial phylogeography of the European pond turtle, Emys orbicularis (Linnaeus 1758). *Molecular ecology*, *8*(11), 1911-1922.

Wiens, J. J., Kuczynski, C. A., & Stephens, P. R. (2010). Discordant mitochondrial and nuclear gene phylogenies in emydid turtles: implications for speciation and conservation. *Biological Journal of the Linnean Society*, *99*(2), 445-461.