

UC Berkeley

UC Berkeley Electronic Theses and Dissertations

Title

A Hierarchical View of Vertebrate Systematics, with Emphasis on Turtles

Permalink

<https://escholarship.org/uc/item/3k1666zr>

Author

Fong, Jonathan Julio

Publication Date

2011

Peer reviewed|Thesis/dissertation

A Hierarchical View of Vertebrate Systematics, with Emphasis on Turtles

By

Jonathan Julio Fong

A dissertation submitted in partial satisfaction of the
requirements for the degree of

Doctor of Philosophy

in

Integrative Biology

in the

Graduate Division

of the

University of California, Berkeley

Committee in charge:

Professor Jimmy A. McGuire, co-Chair

Professor Jeffrey L. Boore, co-Chair

Professor Richard M. Harland

Spring 2011

A Hierarchical View of Vertebrate Systematics, with Emphasis on Turtles

© 2011

by Jonathan Julio Fong

Abstract

A Hierarchical View of Vertebrate Systematics, with Emphasis on Turtles

by

Jonathan Julio Fong

Doctor of Philosophy in Integrative Biology

University of California, Berkeley

Professor Jimmy A. McGuire, co-Chair

Professor Jeffrey L. Boore, co-Chair

The focus of this research is phylogenetic relationships within vertebrates, with a special emphasis on turtles. Despite a substantial amount of previous research, there are still several outstanding questions regarding relationships within vertebrates. By studying the phylogeny at several hierarchical levels (class, order, family, species), we can begin to understand the processes that produce the biodiversity around us. In addition, turtles provide a good system for phylogenetics, as there is relatively low species diversity allowing for more complete sampling, a rich fossil record to calibrate the phylogeny, and applications to conservation. For this research, I take a genomics/bioinformatics approach to assess vertebrate phylogeny. Using a combination of expressed sequence tags (ESTs) and targeted amplification of cDNA, I developed 75 single-copy, nuclear markers conserved across vertebrates (Chapter 1). I also analyze the use of different data types for higher-level phylogenetics. Comparing NUCL (nucleotides), N12 (1st and 2nd codon positions), DEGEN1 (modified sequences to account for codon degeneracy), and AA (amino acids), I find that the NUCL data-type, due to the high level of phylogenetic signal, performs the best across all divergence times. The remaining three data-types (AA, N12, DEGEN1) are less subject to homoplasy, but have greatly reduced levels of phylogenetic signal relative to NUCL (Chapter 1). I use these molecular markers to build a vertebrate phylogeny to answer questions of relationships between and within major groups. In Chapter 2, I address the phylogenetic position of turtles within the amniote phylogeny. Despite over a century of morphological and molecular research, we still do not know where turtles reside in the vertebrate evolutionary tree. I also analyze different partitioning schemes, the effect of missing data, identifying unstable taxa in a phylogeny (rogue taxa), and the use of different data-types (Chapter 2). For the phylogenetic placement of turtles, different analyses and datasets produce different results. However, after performing topological and statistical tests, the weight of the evidence supports the grouping of turtles with archosaurs (birds and crocodiles), either with turtles being the sister group to Archosauria or Crocodylia.

The focus of Chapter 3 is the phylogenetic relationships within turtles, with divergence dating analyses. Within turtles, a basal Pleurodira-Cryptodira was recovered and within Cryptodira, a basal Trionychia (soft-shell turtles) was recovered, with Cheloniodea next (sea turtles). A novel relationship recovered is the sister relationship between *Platysternon* and Testuguria (Testudinidae and Geoemydidae). Divergence dating analyses using new fossil evidence re-classifying stem cryptodires to be stem turtles find the origin of turtles to be much younger than previously believed (~153mya). For Amphibians, data point towards the diphyletic origin of the group (Chapter 5). Most of the recovered relationships within Squamata are consistent with the currently molecular phylogeny, with my data recovering a basal Dibamidae+Gekkonidae, but these results are in sharp contrast to recent morphological studies (Chapter 5). Mammal relationships in this phylogeny also mirror the current mammal phylogeny, favoring the Theria hypothesis (marsupial-placental sister groups) and a basal Afrotheria group (Chapter 5). For the problematic Scandentia (*Tupaia*) clade, there is phylogenetic signal allying *Tupaia* with Glires and Primates, but the signal with primates is stronger (Chapter 5). The last two groups, Actinopterygii (ray-finned fish) and Aves (birds), had relatively poor internal taxon sampling (Chapter 5). Although my results do not provide any new information, these new markers hold promise in helping to resolve relationships for fish and birds. Lastly, a species-level study was performed on turtles in Taiwan to identify the parental species of hybrid individuals found in the wild (Chapter 4). Through molecular methods, the parental species were identified as *Mauremys sinensis* and *Mauremys reevesii*. Presence of *M. reevesii* alleles on the main island of Taiwan indicates that this species may have been introduced. If so, then *M. reevesii* is non-native and conservation efforts should not be wasted protecting this species.

This is dedicated to my family,
parents Julio and Janet, sisters Juliann and Sherina,
brother-in-law Mark, and nephews Robert and Matthew.
All of this was possible only with all your guidance and support.

TABLE OF CONTENTS

Acknowledgements	iii
------------------------	-----

CHAPTER 1: Evaluating phylogenetic informativeness and data-type usage for new protein-coding genes across Vertebrata	1
--	---

CHAPTER 2: The phylogenetic placement of turtles	38
---	----

CHAPTER 3: Phylogenetic relationships within the Testudines.....	120
---	-----

CHAPTER 4: DNA evidence for the hybridization of wild turtles in Taiwan: possible genetic pollution from trade animals.....	142
--	-----

CHAPTER 5: Phylogenetic relationships within other vertebrate groups (Actinopterygii, Amphibia, Squamata, Aves, Mammalia).....	153
---	-----

ACKNOWLEDGEMENTS

“I have seen further, it is by standing on the shoulders of giants”

~ Sir Isaac Newton (1676)

This quote is often used to pay homage to scientists of the past that paved the way for one's own work. In science, this phrase is a bit cliché and incorrect at the same time. Yes, it is true that we as scientists build on the accomplishments of scientists before us, but this only half of the story. These words overlook the other people around us, the non-scientists that provide emotional and moral support to endure the work we do. This dissertation would not have been possible without the help and support of many people, and it is this motivation that kept me going.

First, I would like to acknowledge and thank my dissertation committee, co-advisors Jeff Boore and Jim McGuire, and outside member Richard Harland. If it wasn't for Jeff, I might not have had the opportunity to pursue my Ph.D. at Berkeley. Working for Jeff at the JGI while applying for Ph.D. programs, Jeff changed the trajectory of my life with a few words, “you know I am an adjunct professor and can accept students at Berkeley.” It has been great working with Jeff, and it was always refreshing to hear stories and advice from a man whose diverse life experiences include being a professor, Lieutenant Colonel in the US Air Force, and the CEO of his own company. Jeff was an amazing advisor in that he was always supportive (even in times of stress such as qualifying exam time) and allowed me independence to pursue the research I was interested in. One of my fondest memories of Jeff will be when we were in China together for the 2004 International Congress of Zoology meeting in Beijing, sharing a beer together on the Great Wall.

Jim was gracious to co-advise me during my time at Berkeley. I could always count on Jim to help me when there was a problem with my research or to review manuscripts. I was always amazed by how excited Jim would get about talking about research. Especially when I encountered problems with my work, hearing Jim's enthusiasm about my work would reinvigorate me and get me excited again. Just like Jeff, Jim allowed me the freedom to work on research that interested me. This really allowed me to go through the growing pains of research and truly become an independent researcher.

My outside committee member, Richard Harland, provided great guidance and feedback on my research. Despite being a completely different area of research, Dr. Harland took the time to read through and understand my work. Hearing him tell me he found my dissertation “interesting and refreshing” was one of the best compliments he could have given. I would also like to thank Per Palsboll, George Roderick, Craig Moritz, and Kevin Padian for participation in my qualifying exam. It was hard, but the exam prepared for the work ahead of me in my dissertation. Other professors at Berkeley that contributed to my growth as a researcher through discussions or as teaching courses together: Marvalee Wake, David Wake, Rauri Bowie, Eileen Lacey, Jim Patton, Glenys Thomson, Lewis Feldman, Mary Power, Monty Slatkin, and Vince Resh.

Another scientist that I owe my career to is Jim Parham. Jim started off as my GSI for herpetology, and if it hadn't been for his encouragement, I may have never pursued a career in academia. My first trip to the field in China, the reason I began by studying turtles, and meeting my closest friend/collaborator in China (Shi Haitao) was because of Jim. It was amazing to see how selfless Jim was, even as he was busy applying for jobs and grants, he was always available to help me talk through research problems or read over a manuscript/proposal. I not only consider Jim to be a great mentor, but an amazing friend.

Another person from Berkeley that I need to specially thank is Ted Papenfuss. Ted joined forces with Jim Parham to encourage me to pursue a career in academia, and was also there for my first trip to China. If I ever needed to take a break, I could always count on Ted for a chat and some snacks. Ted and his collecting trips are an inspiration to my own fieldwork.

Working overseas, especially in China, the success of your work depends heavily on the people you know, and how much they are willing to support you. I had this in Shi Haitao. I first met Dr. Shi in 2001, when he came to the US to visit for the first time. Despite the language barrier at the time, he became an informal advisor on my master's thesis. Almost ten years later, despite encountering many difficulties, I am still working in China because of Shi. Seeing Dr. Shi ascend in the ranks from being a professor in the department to being the dean of the science department has been amazing. Despite the added responsibilities, Dr. Shi always finds time to help me with my research, introduce me to new scientists, or to blow off steam by playing ping-pong. During my year in China during the Fulbright program, it was truly a gift to be able to spend time with him and his family. Dr. Shi not only accepted me as his student, supporting me every step of the way, but also as a member of his family.

I have to specially thank members of the McGuire lab: Adam Leaché, Matt Brandley, Matt Fujita, Guin Wogan, Shobi Lawalata, Tom Devitt, Brian Lavin, Dan Portik, Sean Reilly, Allison Davis, Rafe Brown, Chris Witt, Bryan Stuart, Becky Chong, Charles, Linkem. They have been the nuclear family within the extended family of the MVZ. Whenever I needed some help with analyses or even time to complain about something, the labmates were there. The moral and scientific support was endless. Of this group I have to recognize Adam Leaché, Matt Fujita, and Brian Stuart. It was that fateful day talking with them that the project for my dissertation was formulated.

Being a member of the MVZ community has been wonderful. As an undergraduate, I would secretly loathe students who were part of the MVZ since they acted like they were the best. But now that I am a part of this community, I know why they acted they way they did (and forgive them). The administration, staff, undergrads, post-docs, and grad students past and present make this place truly special: Robyn Wong, Rayna Bell, Sean Rovito, Juan Parra, Roberta Damasceno, Jose Wellington Alves dos Santos, Charlotte Jennings, Matthew MacManes, Jay McEntee, Chris Nasrallah, Angela Ribeiro, Andrew Rush, Alan Shabel, Sonal Singhal, Elaine Vo, Rachel Walsh, Beth Wommack, Julie Woodruff, Yu Zeng, Jason MacKenzie, Rachel Mueller, Ana Carnaval, Jason Kolbe, Tami Mott, Bill Monahan, Vance

Vredenburg, David Buckley, Rob Bingham, Inigo Martinez-Solano, Ke Bi, Monica Albe, Carla Cicero, Chris Conroy, Karen Klitz, John Wieczorek, Michelle Koo, Stephen Long, Karen Rowe, Kevin Rowe, Lydia Smith, Carol Spencer, Maria Tonione, Ann Trapaga, Bill Lidicker, Jeremy Brown, Bastien Boussau, Billy Dunaway, Anna Ippolito, Jeanne Kassis, Rhanee Guzman-Dungey, and those I forgot to put on the list.

I have also been very fortunate to collaborate with many different scientists. I would like to thank them for the opportunity to work together on various research projects: Yuxiang Liu, Jinzhong Fu, Bob Macey, Si-Min Lin, Bin He, Bob Murphy, Jichao Wang, Meiling Hong, Shiping Gong, Yaping Zhang, Gexia Qiao, Tien-hsi Chen, Xiaolei Huang, Rong-Chien Lin, Jeff Lin, Shou-Hsien Li, Chuan-Chin Huang, Kuang-Yang Lue, Zhongrong Xia, Anchalee Aowphol, Kumthorn Thirakhupt, Bosco Chan, Michael Lau, Nancy Karraker, Chris Shepherd, Doug Hendrie, Feiyan Zhang, KS Cheung, Tim McCormack, Vivian Fu, Wan Pak Ho, Yik Hei Sung, Thomas Diagne, Che Jing, Haijing Wang.

Research would not be possible without funding, and I was fortunate always have enough funding to do my work. I would like to thank the organizations that provided me with funding: Department of Integrative Biology, the Museum of Vertebrate Zoology, the National Science Foundation, The Fulbright Foundation, The Chelonian Research Foundation.

Many people in the half of my life not at a university or in the field need to be thanked. My good friends who always encouraged me and reminded me how great my life was because I get summer and winter vacations: Daniel Tung, Harry Tung, Ryan Ong, Lonnie Huang, Eric Tung, Jordan Chew, Steve Choi, Jon Mao, Jeff Chan, Dave Chang, Amanda Hou, Jason Cheng, Han Fan, Jenny Yu, Jessica Wu, Monica Hong, Min Hsu, Tim Jaeger, and Choong Sohn. A special thanks goes to the woman with many names, 임은혜 Agatha Lim, 林恩惠, my girlfriend, my better half. I would not have survived without her endless support and encouragement.

Last but not least, I would like to thank my family. My parents and sisters were always supportive of my decisions to pursue a career in academia. Because of the sacrifices my parents made and the freedom I was given to follow my dreams, I am where I am today. As other parents “encouraged” their children to become doctors or lawyers, my parents allowed me to study whatever I pleased. I never had to worry about having enough money for school, and was always encouraged to travel for research. It always amazes me how two immigrants to the United States who spoke very little English raised and put three children through school. I am forever indebted to my parents, and if I can be half the parent they were to me for my children, I would consider it a smashing success.

CHAPTER 1

EVALUATING PHYLOGENETIC INFORMATIVENESS AND DATA-TYPE USAGE FOR NEW PROTEIN-CODING GENES ACROSS VERTEBRATA

ABSTRACT

As a resource for vertebrate phylogenetics, I developed 75 new protein-coding genes using a combination of expressed sequence tags (ESTs) available in Genbank, and targeted amplification of complementary DNA (cDNA). In addition, I performed three additional analyses in order to assess the utility of my approach. First, I profiled the phylogenetic informativeness of these new markers using the online program PhyDesign. Next, I compared the utility of four different data-types used in phylogenetics: nucleotides (NUCL), amino acids (AA), 1st and 2nd codon positions only (N12), and modified sequences to account for codon degeneracy (DEGEN1; Regier et al. 2010). Lastly, I use these new markers to construct a vertebrate phylogeny and address the uncertain relationship between higher-level mammal groups: monotremes, marsupials, and placentals. My results show that phylogenetic informativeness of the 75 new markers varies, both in the amount of phylogenetic signal and optimal timescale. When comparing the four data-types, I find that the NUCL data-type, due to the high level of phylogenetic signal, performs the best across all divergence times. The remaining three data-types (AA, N12, DEGEN1) are less subject to homoplasy, but have greatly reduced levels of phylogenetic signal relative to NUCL. My phylogenetic inference supports the Theria hypothesis of mammalian relationships, with marsupials and placentals being sister groups.

INTRODUCTION

Vertebrates, although quite well-studied, still retain numerous groups with uncertain or controversial phylogenetic affinities (Meyer & Zardoya 2003), including 1) monophyly/paraphyly of Agnatha (jawless fishes), 2) primary relationships within Actinopterygii (ray-finned fishes), 3) the branching order of Actinistia (Coelacanth), Dipnoi (Lungfish) and tetrapods, 4) the branching arrangement of the three living orders of amphibians: Gymnophiona (caecilians), Caudata (salamanders), Anura (frogs), 5) placement of Testudines (turtles) within Amniota, 6) basal relationships of Aves (birds), and 7) the primary branching arrangement among the three major mammal clades (monotremes, marsupials, placentals). From the standpoint of molecular phylogenetics, this probably results from an insufficient number of conserved genetic markers across the group, the use of markers optimized for different levels of resolution (i.e. using rapidly evolving markers to infer deeper relationships), the inherent difficulty of the phylogenetic problems (i.e. short internodes), or a combination of all three factors.

To address these issues, I describe 75 new protein-coding markers that have potential primer-binding sites that are conserved across vertebrates. This method uses a combination of expressed sequence tags (ESTs) available in Genbank for marker development and targeted amplification of complementary DNA (cDNA). A significant advantage of this approach is the heavy utilization of publicly available genomic resources (genomes, ESTs, cDNA databases). Since ESTs and cDNAs code for processed transcripts (i.e. without introns), polymerase chain reaction (PCR) amplification of these markers from cDNA preparations avoid the potential complications of long intervening introns that often confound amplifications from genomic DNA, and enjoy the added benefit of highly conserved primer binding sites. The diversity of evolutionary rates within and among these

markers offers an excellent opportunity to examine their utility in resolving difficult and deep phylogenetic relationships.

There are three goals for this study. First, I estimate the phylogenetic utility of each of the new markers. A marker is most useful, or phylogenetically informative, when evolutionary rates are appropriate for divergence times of interest—a marker that is too slow is uninformative, while a marker that is too fast exhibits noise in the form of homoplasy (Graybeal 1994). Two commonly used measures of phylogenetic signal have been pair-wise sequence divergence and saturation curves. However, both have been shown to be poor measures of phylogenetic informativeness (PI), since pair-wise divergence ignores potential saturation, and saturation curves can only inform one end of the PI spectrum—the too-saturated end. In addition, these measures do not consider information on branch lengths, rate heterogeneity, proportion of invariant sites, and character-state space (Fitch 1986a, Lewis et al. 1997, Bjorklund 1999, Simmons et al. 2002), and it has been shown that homoplasious characters at deeper levels can be informative at finer scales (Wenzel & Siddall 1999). A more appropriate measure of PI, described in Townsend (2007) and implemented in the online program PhyDesign, quantifies the signal inherent in a multiple alignment for a gene along a phylogenetic tree by comparing site-wise evolutionary rates against the ideal rate based on tree length. PhyDesign estimates the net PI by measuring the probability that a substitution occurs along an internode and is not later masked by a subsequent substitution (Townsend 2007). The ability of this method to identify those genes that will be most useful to resolve particular nodes in a tree makes it a much more powerful method to assess the utility of molecular markers. Therefore, I profile the PI of each new marker using PhyDesign.

Second, I assess the utility of four different data-types for phylogenetics. Previous work regarding data-types compared the use of nucleotides (NUCL) versus amino acid translations (AA). There seems to be no controversy in using nucleotides to resolve relationships among closely related taxa, but for deeper-level phylogenetics, the answer is not as clear. Traditionally, preference has been given to amino acids over nucleotides, as they evolve more slowly and are not compromised by the more numerous synonymous substitutions that characterize nucleotide data (Meyer 1994, Simon et al. 1994, Blouin et al. 1998). Amino acids also have increased character-state space (20 character states of amino acids versus 4 character states of nucleotides), which reduces the chance of convergence (Albert et al. 1994, Frohlich & Parker 2000, Simmons et al. 2004). However, other studies have argued that nucleotides should be used instead of amino acids, as the larger quantity of phylogenetic signal (due to 3 times as much data from the triplet codon) are much greater than the losses from decreased character-state space (Simmons et al. 2004, Townsend et al. 2008). Two additional data types have also been used for deep-level phylogenetics: N12 (Edwards et al. 1991, Blouin et al. 1998) and DEGEN1 (Regier et al. 2010). N12 refers to a nucleotide dataset that includes 1st and 2nd codon positions only, as it has been argued that 3rd codon positions become saturated and exhibit homoplasy, especially at deeper time scales (Edwards et al. 1991). DEGEN1 is a nucleotide-based dataset that is modified to take codon degeneracy into account by using nucleotide ambiguity codes. This type of dataset may provide advantages by reducing problems dealing with base compositional heterogeneity

among taxa (Regier et al. 2010). I re-assess the arguments of NUC vs. AA, as well as evaluate the performance of these two additional data types (N12, DEGEN1).

Third, I utilize the data collected from these markers to construct a vertebrate phylogeny, focusing on the problem of higher-level relationships within Mammalia. The two alternative hypotheses are the Theria hypothesis (marsupial-placental sister group) and the Marsupionta hypothesis (marsupial-monotreme sister group) (see Grutzner & Graves 2004 for a review). Of fossil and morphological studies, most support the Theria hypothesis, although some studies using controversial interpretations of morphological data support the Marsupionta hypothesis (Gregory 1947, Kuhne 1973). For molecular studies, data from nuclear markers support the Theria hypothesis (Phillips & Penny 2003, Hugall et al. 2007, Luo 2007). Additional studies using DNA-DNA hybridization (Kirsch & Mayer 1998) and mitochondrial genomes (Janke et al. 2002, Zarodoya et al. 2003) support the Marsupionta hypothesis, but these studies have been criticized due to their incomplete taxon sampling or bias due to base compositional heterogeneity (Grutzner & Graves 2004). All three of the major mammal groups are represented by sequenced whole nuclear genomes that are included in this study: *Ornithorhynchus anatinus* (monotreme), *Monodelphis domestica* (marsupial), and *Mus musculus* and *Homo sapiens* (placental).

METHODS

Marker Development

The three steps of the EST/cDNA methodology used in this study for marker development are diagrammed in Figure 1. First, to maximize the use of publicly available data, I began with EST databases (Genbank) for American Alligator (*Alligator mississippiensis*) and Duck-billed Platypus (*Ornithorhynchus anatinus*). The alligator was chosen to expand our taxon sampling, since there are currently no alligator genomes sequenced, and the platypus to supplement the number of ESTs for potential markers. There were a total of 15,130 ESTs available from these two taxa at the outset of this study (5,430 for alligator and 9,700 for platypus). These ESTs were screened for single-copy genes with a Perl script (available upon request) that uses BLAST to compare ESTs to the chicken genome (*Gallus gallus*), since it is the genome most closely related to the alligator that has high coverage and is well annotated. Paralogous genes and genes with high similarity within gene families were identified and removed by looking for multiple, high-scoring BLAST matches. Additional checks of gene copy number were done by identifying gene families on Ensembl (<http://www.ensembl.org>) and Metazome (<http://www.metazome.com>). In the second step, primer sets were developed for potential markers with the aid of a second Perl script (available upon request). This script was used to 1) choose sequences > 500bp in length, 2) run the tBLASTx algorithm to compare the ESTs to a subset of high-coverage, well annotated genomes (*Homo*, *Gallus*, *Xenopus*), and 3) align the sequences via MUSCLE (Edgar 2004). These potential loci were screened by eye and considered good if there were conserved regions for primer sites and few gaps in the alignment. For the third and final step, primers were tested and optimized on a panel of vertebrates (*Ichthyophis bannanicus* [caecilian], *Heteronotia binoei* [gecko], and *Alligator mississippiensis* [alligator]) through PCR. This test panel focuses on reptiles and amphibians, as they are poorly represented in terms of sequenced nuclear genomes. Our PCR procedure is slightly modified from a standard genomic DNA PCR in

that samples are extracted and prepared as cDNA. The methodology for cDNA preparation is described below.

Sample Preparation

All tissue samples used had been appropriately preserved for RNA work, either stored in RNAlater (Ambion) or flash-frozen in liquid nitrogen. These preservation methods protect RNA by minimizing RNase activity and exposure (Sambrook & Russell 2001). Total RNA was extracted using a guanidinium thiocyanate-phenol-chloroform protocol (Chomczynski & Sacchi 1987, Sambrook & Russell 2001). Next, RNA extracts were reverse-transcribed to double-stranded cDNA using a SMART™ technology protocol (Clontech; Zhu et al. 2001). Resulting cDNA preps were used in place of genomic DNA extracts for standard PCR amplification. Information on each primer pair, such as identity (based on HUGO Gene Nomenclature database), annealing temperature, and size can be found in Table 1. PCR products were purified using ExoSAP-IT (USB Corp.), cycle-sequenced using Big Dye Terminator v.3.1, and sequenced using an ABI 3730 automated sequencer.

Datasets

Alignments for each gene consisted of orthologous sequences of single-copy genes from 10 online genomes, and three new samples from the test panel/ESTs. Taxon sampling was selected to span the diversity of vertebrates, and a complete list of taxa and Genbank/Ensembl numbers can be found in Appendix A. Compiling the orthologous genes from the online genomes was a multi-step process. The first step was a “marker development” stage (see above), which included pre-screening the markers based on BLAST searches to increase the probability that the gene is single-copy. Secondly, BLASTN (BLASTv2.2.16) was used to compare the human copy of each gene to cDNA databases from Ensembl (<http://www.ensembl.org>) to identify the orthologous gene of each genome. Ensembl provides two types of cDNA databases: “cdna.all” and “cdna.abinitio”. I used the more inclusive “cdna.all” database, as it includes a set of all transcripts resulting from Ensembl known, novel, and pseudogene predictions. The cdna.abinitio database was not used because its results are based solely on gene prediction algorithms, so not all predictions represent biologically real cDNAs. An E-value of 0.0001 was used as the cut-off value to retain BLAST-hits in the analysis. All the above steps were facilitated by a Perl script (available upon request). The alligator sample in this study was represented by both ESTs and a new cDNA sample. When available for both samples, data were identical; both samples were combined into a single taxon to reduce the amount of missing data.

Newly sequenced samples were assembled and edited using Geneious 5 (Biomatters Ltd.), while pseudogenes and duplicate cDNA sequences (alternative transcripts of the same gene) were removed based on information from cDNA databases. All the remaining sequences were aligned with the MUSCLE plug-in (Edgar 2004) by aligning amino acids back translated to nucleotides, run with a maximum of eight iterations. Alignments were trimmed and exported for the preliminary phylogenetic analyses. Preliminary phylogenetic analyses for individual genes were run using RAxML v7.2.5 (Stamatakis 2008) on ABE (TeraGrid) using the CIPRES portal v2.2 (<http://www.phylo.org/portal2/>; Miller et al., 2009), employing the GTRGAMMA model of evolution for tree inference, and GTRCAT

model of evolution for 1000 bootstrap replicates. These preliminary phylogenetic analyses were the final test of gene orthology. Paralogous gene sequences are a result of gene duplication, and timing of the duplication relative to lineage divergence produces different tree topologies. Duplications post-dating a splitting event results in all copies for a lineage being closely related, and does not affect tree topology. In this case, only the highest matching copy to the human gene (based on E-value) was retained. If the duplication predated a splitting event, the phylogeny shows replicated tree topologies within the larger tree, as is expected due to genome duplications in the ancestral vertebrate (Dehal & Boore 2005). To rectify this situation, the single sub-tree that included sequences from the test panel was retained, while others were deleted.

Once final nucleotide alignments were assembled, data were converted to the three additional datasets (AA, N12, DEGEN1): AA was translated in Geneious, 3rd codon positions were removed for N12 using MacClade v4.08 (Maddison & Maddison 2008), and DEGEN1 was converted using a Perl script (Regier et al. 2010; available at <http://www.phylotools.com>). Concatenation of the individual genes was performed with a Perl script (available upon request).

Phylogenetic Analyses

Final phylogenetic analyses were run using RAxML and MrBayes on data sets representing all four data-types (NUCL, N12, DEGEN1, AA), partitioned by gene. RAxML was run using RAxMLv.7.2.5 (Stamatakis 2008) on ABE (TeraGrid) using the CIPRES portal v2.2 (<http://www.phylo.org/portal2/>; Miller et al. 2009), employing the GTRGAMMA model of evolution for tree inference, and GTRCAT model of evolution for 1000 bootstrap replicates for nucleotide-based analyses and the Protein Gamma model of evolution using empirical frequencies for amino acid-based analyses.

MrBayes was run on BioHPC at the Cornell University Computational Biology Service Unit (<http://cbsuapps.tc.cornell.edu/>) as two independent searches with random starting trees and compared using four chains, 10 million generations, and sampling every 5000th generation. Models of evolution were inferred for nucleotide and amino acid partitions using MrModeltest v2.3 (Nylander 2004) and PROTTEST v2.4 (Abascal et al. 2005), respectively. Burn-in of MCMC chains was evaluated using the online program AWTY by evaluating cumulative plots of posterior probabilities of the 20 most variable splits (Wilgenbusch et al. 2004). It is from these phylogenetic analyses that I examined the two alternative hypotheses of higher-level mammal relationships.

Phylogenetic Informativeness of Markers

The online program PhyDesign (<http://phydesign.townsend.yale.edu/>; Townsend 2007) was used to determine the net PI of each of the new markers. This approach estimates PI by using a model of sequence evolution to measure the probability that a substitution occurring along an internode is not masked by a subsequent substitution (Townsend 2007). In my study, PI is calculated on an uncalibrated tree, so times are relative, with time=0 at the tips and time=1 at the root. A key result from the PI analysis is the time at which PI peaks and reaches its maximum (PI_{max}). It is at this point in time where the marker is most phylogenetically informative, and is directly related to the optimal divergence time a marker

should be used (Figure 2). Although it has already been shown that pair-wise distance (p-dist) is a poor measure of PI, p-dist has never been compared to results from PhyDesign. The relationship between p-dist and PI was explored by examining the correlation between p-dist and PImax; if p-dist is a good measure of PI, then it should be positively correlated with PImax. The p-dist chosen to compare among markers was between mouse (*Mus musculus*) and human (*Homo sapiens*) as these data were available for all but one marker. For this one marker (gene UBE2Q1), the mouse-orangutan (*Pongo pongo*) relationship was used.

PhyDesign requires an ultrametric tree file and an alignment. Similar results were inferred from all analyses (RAxML, MrBayes) for the four data types (NUCL, N12, DEGEN1, AA), which match the general consensus of higher-level vertebrate relationships (Figure 2). The only discrepancy among the analyses was that a paraphyletic Lissamphibia (extant amphibians) was often recovered (AA-Bayes, AA-RAxML, NUCL-Bayes, NUCL-RAxML, N12-Bayes). There is some controversy in the literature on the monophyly of Lissamphibia (Cannatella et al. 2009), with some paleontological studies recovering a diphyletic origin (Carroll 2007, Anderson et al. 2008). However, previous work (Benton 1990, Meyer & Zardoya 2003) supports the monophyly of Lissamphibia, so I use this as the phylogenetic hypothesis. To allow for direct comparisons between the data-types, the tree file from the final N12 RAxML analysis was used for all analyses, which was converted to an ultrametric tree using TreeEdit v1.0a10 (Rambaut & Charleston 2001). PhyDesign was run on concatenated alignments of the four data-types partitioned by gene. The program HyPhy (Pond et al. 2005) using empirical base frequencies and a time-reversible model of substitution (transitions=2, transversions=1) was used to calculate PI of nucleotide-based data types (NUCL, N12, DEGEN1), while rate4site (Mayrose et al. 2004) under a JTT model of evolution and ML inference method was used for the AA dataset. Using this approach I obtained four PI calculations for each gene. Two different comparisons of PI were done, first between genes under each data-type to determine the relative utility of each gene, and next between data-types for each gene to determine the relative utility of each data-type.

RESULTS & DISCUSSION

Marker Development

From the 15,130 alligator and platypus ESTs, more than 5,000 potential markers were identified (Figure 1, step 1). After further screening, ~400 potential primer sets were developed (Figure 1, step 2). Lastly, primer sets were screened on the test panel of cDNA preparations (Figure 1, step 3), and primer sets that amplified specifically were selected. The final result was 75 new nuclear markers conserved across vertebrates. The PCR products of these new markers ranged between 152-789bp in size. Since conservation of primer sites was crucial due to the wide phylogenetic range, it was not always possible to place primers at the ends of a gene, so many markers are ≤500bp despite pre-screening for longer genes. A list of these molecular markers and associated information can be found in Table 1. Combined, these markers comprise 33,843 aligned bp.

Several prior studies have explored the use of ESTs for phylogenetics (Baptiste et al. 2002; Choi et al. 2004; Phillippe et al. 2005; de la Torre et al. 2006; Hughes et al. 2006). These studies depended on the mass sequencing of random clones to obtain homologous

genes. Three potential pitfalls with these studies are missing data (between 35-74%; Rodriguez-Ezpeleta et al. 2005; Hughes et al. 2006), prevalence of mitochondrial transcripts (Hughes et al. 2006), and failure to differentiate between orthologous and paralogous genes. Our approach addresses these issues and mirrors a more recent approach (Regier et al. 2010), which is a targeted EST strategy. For our 75 gene/13 taxon dataset, there was only 17.4% missing data. Such an approach allows the compiling of a more complete dataset and a more efficient use of sequencing funds.

Profiling Markers

It has been argued that p-dist is a poor measure of phylogenetic informativeness (Fitch 1986, Graybeal 1994, Simmons et al. 2002, Townsend 2007, Townsend et al. 2008). Our results support this by showing there is a weak relationship between p-dist and time of PImax ($r^2=0.02$, $\sigma=0.2$; Figure 3).

PI of the 75 new markers varies, both in the amount of phylogenetic signal and optimal timescale. Figure 4 displays the PI of all of the genes using the NUCL dataset (PI profiles for N12, DEGEN1, and AA in Figures 5-7). The best way to interpret these graphs is from right to left (tips to root). The PI of the data increases until reaching the PImax, after which the decrease in PI indicates the potential for noise due to homoplasy. To minimize the potential for noise, when choosing a marker, one should choose a marker that peaks deeper in time than the time period of interest (Townsend & Leuenberger in press). In general, shape of PI profiles can be used to estimate overall utility of genes, with low, flat curves indicating less useful genes, curves with a sharp peak characterizing genes most useful for a narrow timescale, and steadily increasing or broad curves representing genes useful over longer time scales.

Based on these data under the NUCL dataset, most of these new markers are informative at relatively narrow, shallow timescales, and will be useful for recovering the phylogenies within the major vertebrate groups. The markers with the deepest timescales reach as far back as the alligator-bird split. This is not to say that these markers will not be useful for resolving deeper relationships, as plots of PI should be used as guidelines rather than definitive indicators of utility, but users should be aware of the increased potential for noise from homoplasy.

Data-type Comparison

A comparison of data-types was performed for each gene, and an exemplar is shown in Figure 2 (other genes found in Figure 8). These results can inform us of the optimal time scales (based on the PImax) and relative utility (comparison of PI curves) of the data-types. Again, PImax is a measure of the optimal divergence time, with tips=0, root=1. When evaluating data-types based on optimal divergence time (PImax), DEGEN1 is most ideal for deeper divergences (average PImax=0.71), with AA (average PImax=0.53) and N12 (average PImax=0.4) for intermediate divergences, and NUCL (average PImax=0.18) for shallow divergences. This is expected, as DEGEN1, AA, and N12 are less likely to exhibit homoplasy (Edwards et al. 1991, Simmons et al. 2004, Regier et al. 2010).

In comparing PI curves, the NUCL data-type was always the most phylogenetically informative, while AA, N12, and DEGEN1 each suggested similarly low levels of PI. In fact,

the PI of NUCL is often upwards of 20-40 times the PI value for the other data-types (Figures 2, 8). The disparity in PI between the NUCL and N12 data-type indicate that the majority of phylogenetic signal is tied to the 3rd codon position. As found in previous studies, the improvement in phylogenetic signal from using nucleotides appears to substantially outweigh the potential benefits of greater character-state space of amino acid conversion (Simmons et al. 2002, Townsend et al. 2008).

In terms of practical recommendations from these results, I suggest that nucleotides (NUCL) should be used at all time scales because of the tremendous levels of phylogenetic signal. However, the impact of phylogenetic noise should be kept in mind, which can be estimated but not explicitly measured in PhyDesign (Townsend & Leuenberger in press). So when the phylogenetic problem deals with deeper divergences, the other three data-types should be tested, as they are less subject to homoplasy. However the relatively low PI levels of AA, N12, and DEGEN1 may nullify these advantages.

Mammal Phylogenetics

All analyses (RAxML, MrBayes) for the four data types (NUCL, N12, DEGEN1, AA) resulted in similar phylogenetic inferences, which match the general consensus of higher-level vertebrate relationships (Figure 2; but see discussion about Lissamphibia above). Of the uncertain relationships within vertebrates outlined in Meyer & Zardoya (2003), taxon sampling allowed me to address the higher-level relationship within mammals. With representatives from all of the major mammal groups, data strongly support the Theria hypothesis placing marsupials and placentals as sister groups (Figure 2). These results are robust with regard to data-type employed, as all analyses recovered the same relationships. However, I caution that this may not be the case with other relationships, and comparisons between data-types should be performed. I am confident that with additional taxon sampling, these new markers will help clarify additional relationships within vertebrates.

CONCLUSION

Herein I have used a novel EST/cDNA marker development protocol to generate 75 new single-copy, protein-coding markers conserved across vertebrates. I also profile the phylogenetic informativeness of these markers and show there is a diversity of levels of informativeness across these genes. Genes should be selected to match the time period of the phylogenetic problem at hand. When comparing different data-types for our taxonomic assemblage, nucleotide data performed best over all time scales, while AA, DEGEN1, and N12 all perform poorly. Lastly, our phylogenetic inference supports the Theria hypothesis of higher-level mammal relationships.

ACKNOWLEDGEMENTS

I would like to thank Drs. Carol Spencer and Jim McGuire (Museum of Vertebrate Zoology, UC Berkeley) for tissue loans. Jeffrey Boore, Richard Harland, Jim McGuire, Guin Wogan, Matt Fujita, and Roberta Damasceno helped in reviewing early drafts of the manuscript. Also, I would like to thank Dr. Jeffrey Townsend (Yale University) for developing and helping us use the PhyDesign program. Funding for this project came from

the National Science Foundation Doctoral Dissertation Improvement Grant (#08-564) and the Museum of Vertebrate Zoology.

CITATIONS

Abascal F, Zardoya R, Posada D (2005) ProtTest: Selection of best-fit models of protein evolution. *Bioinformatics* 21: 2104-2105.

Albert VA, Backlund A, Bremer K, Chase MW, Manhart JR, Mishler BD, Nixon KC (1994) Functional constraints and *rbcL* evidence for land plant phylogeny. *Annals of the Missouri Botanical Garden* 81: 534-567.

Bapteste E, Brinkmann H, Lee JA, Moore DV, Sensen CW, Gordon P, Durufle L, Gaasterland T, Lopez P, Muller M, Philippe H (2002) The analysis of 100 genes supports the grouping of three highly divergent amoebae: Dictyostelium, Entamoeba, and Mastigamoeba. *Proceedings of the National Academy of Sciences* 99: 1414-1419.

Benton MJ (1990) Phylogeny of the major tetrapod groups: morphological data and divergence dates. *Journal of Molecular Evolution* 30: 409-424.

Björklund M (1999) Are third positions really that bad? A test using vertebrate cytochrome b. *Cladistics* 15: 191-197.

Blouin MS, Yowell CA, Courtneely CH, Dame JB (1998) Substitution bias, rapid saturation, and the use of mtDNA for nematode systematics. *Molecular Biology and Evolution* 15: 1719-1727.

Carroll RL (2007) The Palaeozoic ancestry of salamanders, frogs and caecilians. *Zoological Journal of the Linnean Society* 150 (suppl.): 1-140.

Cannatella DC, Vieites DR, Zhang P, Wake MH, Wake DB (2009) Amphibians (Lissamphibia). In *The Timetree of Life*, Hedges, SB, Kumar, S (Eds.) Oxford University Press, 2009. Pp. 353-356.

Choi H-K, Kim D, Uhm T, Limpens E, Lim H, Mun J-H, Kalo P, Penmetsa RV, Seres A, Kulikova O, Roe BA, Bisseling T, Kiss GB, Cook DR (2004) A sequence-based genetic map of *Medicago truncatula* and comparison of marker colinearity with *M. sativa*. *Genetics*: 1463-1502.

Chomczynski P, Sacchi N (1987) Single-step method of RNA isolation by acid guanadinium thiocyanate-phenol-chloroform extraction. *Analytical Biochemistry* 162: 156-159.

Dehal P, Boore JL (2005) Two rounds of whole genome duplication in the ancestral vertebrate. *PLoS Biology* 3: e314.

de la Torre JEB, Egan MG, Katari MS, Brenner ED, Stevenson DW, Coruzzi GM, DeSalle R (2006) ESTimating plant phylogeny: lessons from partitioning. *BMC Evolutionary Biology* 6: 48.

Edgar RC (2004) MUSCLE: multiple sequence alignment with high accuracy and high throughput. *Nucleic Acids Research* 32: 1792-1797.

Edwards SV, Arctander P, Wilson AC (1991) Mitochondrial resolution of a deep branch in the genealogical tree for perching birds. *Proceedings: Biological Sciences* 243: 99-107.

Fitch WM (1986) The estimate of total nucleotide substitutions from pairwise differences is biased. *Philosophical Transactions of the Royal Society of London B* 312: 317-324.

Frolich MW, Parker DS (2000) The mostly male theory of flower evolutionary origins: from genes to fossils. *Systematic Botany* 25: 155-170.

Graybeal A (1994) Evaluating the phylogenetic utility of genes: a search for genes informative about deep divergences among vertebrates. *Systematic Biology*: 174-193.

Gregory WK (1947) The monotremes and the palimpsest theory. *Bulletin of the American Museum of Natural History* 88: 1-52.

Grutzner F, Graves JAM (2004) A platypus' eye view of the mammalian genome. *Current Opinion in Genetics and Development* 14: 642-649.

Hugall AF, Foster R, Lee MSY (2007) Calibration choice, rate smoothing, and the pattern of tetrapod diversification according to the long nuclear gene RAG-1. *Systematic Biology* 56: 543-563.

Hughes J, Longhorn SJ, Papadopoulou A, Theodorides K, de Riva A, Meja-Chang M, Foster PG, Vogler AF (2006) Dense taxonomic EST sampling and its applications for molecular systematics of the Coleoptera (beetles). *Molecular Biology and Evolution* 23: 268-278.

Janke A, Magenell O, Wiczorek G, Westerman M, Arnason U (2002) Phylogenetic analysis of 18S rRNA and the mitochondrial genomes of the wombat, *Vombatus ursinus*, and the spiny anteater, *Tachyglossus aculeatus*: increased support for the Marsupionta hypothesis. *Journal of Molecular Evolution* 54: 71-80.

Kuhne WG (1973) The systematic position of monotremes reconsidered (Mammalia). *Zoomorphology* 75: 59-64.

- Lewis LA, Mishler BD, Vilgalys R (1997) Phylogenetic relationships of liverworts (Hepaticae), a basal embryophyte lineage, inferred from nucleotide sequence data of the chloroplast gene *rbcL*. *Molecular Phylogenetics and Evolution* 7: 377-393.
- Luo Z (2007) Transformation and diversification in early mammal evolution. *Nature* 450: 1011-1019.
- Maddison WP, Maddison DR (2008) *MacClade: analysis of phylogeny and character evolution*. Sinauer, Sunderland, Mass.
- Meyer A (1994) Shortcomings of the cytochrome b gene as a molecular marker. *Trends in Ecology and Evolution* 9: 278-280.
- Meyer A, Zardoya R (2003) Recent advances in the (molecular) phylogeny of vertebrates. *Annual Review of Ecology, Evolution, and Systematics* 34: 311-338.
- Miller MA, Holder MT, Vos R, Midford PE, Liebowitz T, Chan L, Hoover P, Warnow T, (2009) The CIPRES portals. CIPRES. 2009-08-04. URP: http://www.phylo.org/sub_sections/portal. Accessed 2009-08-04. (Archived by WebCite(r) at <http://www.webcitation.org/5imQ1JeQa>).
- Nylander JAA (2004) *MrModeltest v2*. Program distributed by the author. Evolutionary Biology Centre, Uppsala University.
- Philippe H, Lartillot N, Brinkmann H (2005) Multigene analyses of bilaterian animals corroborates the monophyly of Ecdysozoa, Lophotrochozoa, and Protosomia. *Molecular Biology and Evolution* 22: 1246-1253.
- Philips MJ, Penny D (2003) The root of the mammalian tree inferred from whole mitochondrial genomes. *Molecular Phylogenetics and Evolution* 28: 171-185.
- Rambaut A, Charleston M (2001) see <http://tree.bio.ed.ac.uk/software/treeddit/>
- Rodriguez-Ezpeleta N, Brinkmann H, Burey SC, Roure B, Burger G, Loffelhardt W, Bohnert HJ, Philippe H, Lang BF (2005) Monophyly of primary photosynthetic eukaryotes: green plants, red algae, and glaucophytes. *Current Biology* 15: 1325-1330.
- Regier JC, Shultz JW, Zwick A, Hussey A, Ball B, Wetzer R, Martin JW, Cunningham CW (2010) Arthropod relationships revealed by phylogenomic analysis of nuclear protein-coding sequences. *Nature* 463: 1079-1083.
- Sambrook J, Russell DW (2001) *Molecular cloning: a laboratory manual*. Cold Springs Harbor Laboratory Press. Cold Springs Harbor, New York.

Simmons MP, Ochoterena H, Freudenstein JV (2002) Amino acids vs. nucleotide characters: challenging preconceived notions. *Molecular Phylogenetics and Evolution* 24: 78-90.

Simmons MP, Reeves A, Davis JI (2004) Character-state space versus rate of evolution in phylogenetic inference. *Cladistics* 20: 191-204

Simon C, Frati F, Beckenbach A, Crespi B, Liu H, Flook P (1994) Evolution, weighting, and phylogenetic utility of mitochondrial gene sequences and a compilation of conserved polymerase chain reaction primers. *Annals of the Entomological Society of America* 87: 651-701.

Stamatakis A, Hoover P, Rougemont J (2008) A rapid bootstrap algorithm for the RAxML web servers. *Systematic Biology* 57: 758-771.

Townsend JP (2007) Profiling phylogenetic informativeness. *Systematic Biology* 56: 222-231.

Townsend JP, Lopez-Giraldez F, Friedman R (2008) The phylogenetic informativeness of nucleotide and amino acid sequences for reconstructing the vertebrate tree. *Journal of Molecular Evolution* 67: 437-447.

Townsend JP, Leuenberger C (in press) Point of View: taxon sampling and the optimal rates of evolution for phylogenetic inference. *Systematic Biology*.

Wenzel JW, Siddall ME (1999) Noise. *Cladistics* 15: 51-64.

Wilgenbusch JC, Warren DL, Swofford DL (2004) AWTY: a system for graphical exploration of MCMC convergence in Bayesian phylogenetic inference. <http://ceb.csit.fsu.edu/awty>.

Zardoya R, Malaga-Trillo E, Veith M, Meyer A (2003) Complete nucleotide sequences of the mitochondrial genome of a salamander, *Mertensiella luschni*. *Gene* 317: 17-27.

Zhu YY, Machleder EM, Chenchik A, Li R, Siebert PD (2001) Reverse transcriptase template switching: a SMART approach to full-length cDNA library construction. *BioTechniques* 30: 892-897.

Table 1: The 75 new protein-coding markers developed and used in this study, along with associated information.

Name	Full Gene Name (HUGO Gene Nomenclature Committee)	Primer 1 (forward)	Primer 2 (reverse)	Annealing Temp (°C)	marker size (bp)
ABCE1	ATP-binding cassette, sub-family E (OABP), member 1	AAACTTCACAGGTTGCCT	AARATGTTTATGCCTTCTCT	60	587
AMD1	adenosylmethionine decarboxylase 1	AATGACAGCTGACAGCA	TCTGAAGGTTTGAAATCAAT	60	423
ARCN1	archain 1	ATGGAGAARCTSTAYATGGT	TTYARYTTGTCCACAAAGTT	52	492
C22orf28	chromosome 22 open reading frame 28	GATCMRSAAGGGCTTCGT	ATGGARGARCGGTTGAC	55	725
C7orf28B	chromosome 7 open reading frame 28B	TTYTGGATGGTCATGGT	CACTCATGGCCTTRTAAAC	60	620
CBX3	chromobox homolog 3	GCAGAGCCTGAAGARTTT	GGRCAWGAATGCCAAGTTA	60	406
CPSF3	cleavage and polyadenylation specific factor 3, 73kDa	GARAGYATGGACAARATGAAAC	AWGAAATGTAATCHACAGACAT	52	676
CSNK2A2	casein kinase 2, alpha prime polypeptide	CGCGASTACTGGGACTA	GTCAGYCTCTGYTGATGGTC	60	695
CTNNB1	catenin (cadherin-associated protein), beta 1, 88kDa	CARGGDTTYTCTCAGTCCTT	CYARAGCTTGCAATCCACC	60	656
DSCR3	Down syndrome critical region gene 3	TVCARCTHAGTGCSAAAAGTGT	GCTCTCTCTTTTRACRTTCTG	53	276
EED	embryonic ectoderm development	TAYTCYTTCAARTGTGTAAATAG	ATKGCAITTTTCAACAGAC	52	513
EIF3B	eukaryotic translation initiation factor 3, subunit B	TATGAAACWCCTTCYATGGG	CCRITGTTTTDCACTGGTA	50	404
EIF3M	eukaryotic translation initiation factor 3, subunit M	TCTAAAGGAGCKGAGATYTC	TCCAGTCAGAAAATCCATT	55	333
EXOC2	exocyst complex component 2	GGGTGKAGKAAATGAAYTAT	CCTTCYACTTTTTCYGTTC	60	323
GAPDH	glyceraldehyde-3-phosphate dehydrogenase	AARCAYAACTGYATGGC	TCYTCAATWATCCAGAYGA	55	242
GDE1	glycerophosphodiester phosphodiesterase 1	TGAYGTCAAAGGCCATG	TCYASCADGCTGTCGTGAT	56	369
GNAS	GNAS complex locus	GTSAARGCATGCTTYGAGAG	GTGCATYCKTGRATGATGTC	60	445
GSPT1	G1 to S phase transition 1	RYSATGATGCCAAACAAGC	TTYTCTGGAACAGTTTCA	60	413
H2AFY2	H2A histone family, member Y2	AGSGTMCCTGCCMAGAATTC	ATGTGRCAGTGGATGACAAA	60	495
HIAT1	hippocampus abundant transcript 1	GAAAGAAGYATGGCHTATGG	GCCACATCATCCAA	59	484
HPD	4-hydroxyphenylpyruvate dioxygenase	GAMTWCMTYGTCAGAAAAGC	TGTCATCACSGACCAGAA	55	327
IFT57	intraflagellar transport 57 homolog (Chlamydomonas)	GCTGCTTGGYTKATHARTAAAG	TGATTGTTGATGTAATTTTC	52	615
KDM6A	lysine (K)-specific demethylase 6A	GACACTATKTGTGATAGGCAATA	GTKGAAATGGCTGAAGATG	60	645
KPNA4	karyopherin alpha 4 (importin alpha 3)	CARAARGAAGCTGCTTGG	GGTACATTGGCAGATGAAT	60	265
MAT2B	methionine adenosyltransferase II, beta	GDCTTCTTGGYMGAGCTGT	GRAGGTGAAKGCATCTGC	59	633
NFIA-005	nuclear factor I/A, transcript variant 5	TTTGTCACATCAGGTGTTTT	CTTGCCTTGGCTGCT	60	662
NFIA-007	nuclear factor I/A, transcript variant 7	TTCATYGAAGCACTTCTKC	CCAARTATAAATCGAGTTCCCT	60	378
NFIB	nuclear factor I/B	AGGCACCTTCTCCRCA	AGARGACATATCTTGATCTC	58	789
NFYB	nuclear transcription factor Y, beta	ATAATGAAAAAYGCCATACC	ATAAGAACTGTCTACAAATCTC	59	386
NR2C2	nuclear receptor subfamily 2, group C, member 2	ATGGGCATGAAAATGGA	TACTCTGGCATTGGACTDG	59	601
NTS	neurotensin	GAAGAGGAAATGAAAGC	CTYTTAGATATGTADGGTCTTC	55	345
OAT	ornithine aminotransferase	GCWGTGAYYATGAAAATGT	CKRAKYTCRTCYTCCTTRAT	55	380
OAZ1	ornithine decarboxylase antizyme 1	ACCCACCCCTGAAGAT	ACAATCTCAAAGCCCAA	60	318
OLA1	Obg-like ATPase 1	CAAGCAAAARTCCWGCC	TTTTTATCAACCATTGTGTTTTTC	52	431
PABPC1	poly(A) binding protein, cytoplasmic 1	CATYATGTGGTTCYACGG	ACCCTGGTATCTGGTGAT	60	538
PAPOLA	poly(A) polymerase alpha	CARTCTGTAAATHGAAAATGT	AGTTTGGACCACTCTGC	52	525
PDCD10	programmed cell death 10	AAGGCWGAAGAAAAGAAATCC	CCACAGTTTTGAAGGTCTG	60	432
POLR2B	polymerase (RNA) II (DNA directed) polypeptide B, 140kDa	GATGTCTGTDCARAGAAATGT	ATTTCCATCATYTCCTGG	52	676
PPM1A	protein phosphatase, Mg2+/Mn2+ dependent, 1A	TTAGACAAGCCAAAGATG	ATGTTGTCTCKRCTTCCC	52	231
PPTC7	PTC7 protein phosphatase homolog (S. cerevisiae)	ACGTTTTRGTAAAAGARGGAC	AAACTGTGCAAAARGGTGAC	53	465
PSAT1	phosphoserine aminotransferase 1	ATYAGTGTCTWGAATGAGYCA	CAGCAAAAATYACMCCAAAC	59	417
PSMA6	proteasome (prosome, macropain) subunit, alpha type, 6	AAGCACTGGTTCCCT	TCTGGAGTGATGTGAATAGT	60	295
PSMC6	proteasome (prosome, macropain) 26S subunit, ATPase, 6	TCGYCGWCAGCTKGAC	GCTTCATAATCWATTTCVCCA	60	663
PSMD14	proteasome (prosome, macropain) 26S subunit, non-ATPase, 14	GTKGTGYGGTTGGTATCACAG	ACTGRACAATRTTTGARGTCAT	60	520
RAB1A	RAB1A, member RAS oncogene family	CTDCTTAGGTTTGCMGATGA	CTTGACTGGAGTGCTCTG	60	400
RAN	RAN, member RAS oncogene family	GTAACACAAATTTGARAAGCC	CCACATTCACCTCCGAA	60	272
RPL4	ribosomal protein L4	TACTCYGARAAGGGGGA	CATDCGCTGAGASGCATA	55	413

RPL7A	ribosomal protein L7a	CTTYGTSAAATGGCCCC	CCAGTGRGRCGGATCT	60	501
RPL11	ribosomal protein L11	GARAARGAGAACCCYATGCG	AACCAGCGCATKGCYT	58	380
RPL38	ribosomal protein L38	CCTCGCAARATTGARGATC	TTCAGYTCCTTCACWGMCA	55	152
RPLP0	ribosomal protein, large, P0	ATGCCCAGGGAAGACA	GCTTCRCTGGCWCCA	60	437
RPS4X	ribosomal protein S4, X-linked	AAGAAGCAYYTGAAGCGSGT	AAAGCTGTTGCCATTKGATC	60	596
RPS6	ribosomal protein S6	TTCTATGAGAAAGCGNATGGC	TCAGAYTTRGAWGTRGARGCTC	60	589
RPS14	ribosomal protein S14	AATGGCWCHCGHAAGGG	ATRGGGGTGACATCCTCAAT	55	339
SDHB	succinate dehydrogenase complex, subunit B, iron sulfur (Ip)	GGCTCYTGYGCMATGAACAT	GTGCAGTTCATGATRGTTGT	60	387
SEC13	SEC13 homolog (S. cerevisiae)	AAYGGAGGGCAGATCC	CCCTTGTTKACATCRCTGAT	60	711
SEC62	SEC62 homolog (S. cerevisiae)	TATGTRTGGATMTATGACCC	CATTKCCACTACTRTGCTG	60	468
SGK1	serum/glucocorticoid regulated kinase 1	GGAGAGTTGTTCTACCATCT	CRCTCACATTTGGGTAAAA	60	499
SLC25A20	solute carrier family 25 (carnitine/acylcarnitine translocase), member 20	ATGACVTAYGAATGGCTGAA	AAGRAAACAKGCCGCATT	52	228
SMU1	smu-1 suppressor of mec-8 and unc-52 homolog (C. elegans)	TTTATGATGATGGATGATGC	CCRCCTTCTCTYTTACCAGA	58	493
SNX6	sorting nexin 6	GCTYCATGATTCCTTTGT	ATCCTTAGCAGCTTGWGAT	52	620
SOD2	superoxide dismutase 2, mitochondrial	GCBRCYTAYGTGAACAACCT	ATHACRTTCCARATMGCTTTYA	50	407
SRP54	signal recognition particle 54kDa	GGTATGGGBGAYATTGAAGGA	GYCARMAGYTCHTGDACATCTCT	55	316
STK39	serine threonine kinase 39	TGCCARACMAGTATGGATGA	GGAGGATAYTTGTGATAAGG	60	410
TAT	tyrosine aminotransferase	GATCTAYGGWGAYATGGTGTTT	GGRTAYTCRAARCAMGTDGCKG	55	435
UBA5	ubiquitin-like modifier activating enzyme 5	GACKGCTGAARATGYTGACAAG	ATCCTTGCATYGCRTTGT	53	502
UBE2J2	ubiquitin-conjugating enzyme E2, J2 (UBC6 homolog, yeast)	GTGGYTATTAYCATGGVAAACT	TGACYGTGTARGCRAAGGC	49	547
UBE2K	ubiquitin-conjugating enzyme E2K (UBC1 homolog, yeast)	AATMTGGCATCCBAATATTAG	TCAGTTACTCAGRAGYARTTCTGT	55	310
UBE2N	ubiquitin-conjugating enzyme E2N (UBC13 homolog, yeast)	AGGATTCCCCCTTTGAG	CTACCTTCATTCACAGTGCT	59	480
UBE2Q1	ubiquitin-conjugating enzyme E2Q family member 1	ATGAARGARCTCAGGGAYAT	ACCAGCCRTTYTTCTCAT	60	405
UBE2V2	ubiquitin-conjugating enzyme E2 variant 2	GGAAGAACTTGAAGAAGG	CATATTTTCTTTGGACATCAT	60	283
VDAC1	voltage-dependent anion channel 1	TGGAATTYACAAGCTCAGG	GGCTTWARAGTCTGAGTGATCC	60	569
VPS29	vacuolar protein sorting 29 homolog (S. cerevisiae)	TTGGTKTTRGTATTAGGAGA	CTACTTYYACATCATCTCCAAT	60	451
WTAP	Wilms tumor 1 associated protein	AAGTGTGCAATGCTTATCC	CTGTGTCATTKGANTGGTG	60	537
XPOT	exportin, tRNA (nuclear export receptor for tRNAs)	ATAAATATGCTGCTGCTAGGT	GCYTCATCTTCACCCTC	60	452

Figure 1: Flow diagram of EST/cDNA methodology used in this study for marker development.

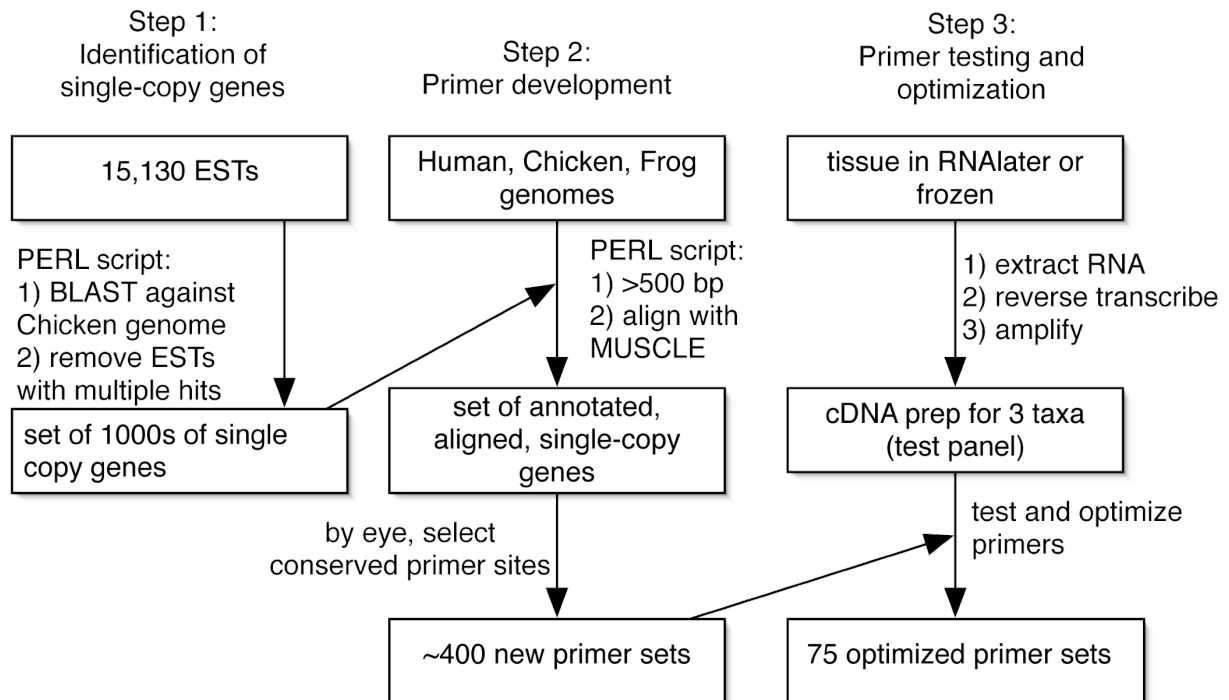


Figure 2: Comparison of the net phylogenetic informativeness (PI) of four different data types 1) NUCL (all nucleotides), 2) N12 (1st and 2nd codon positions), 3) AA (amino acids), and 4) DEGEN1 (nucleotides modified to take codon degeneracy into account). PImax denotes the point at which maximum PI is achieved. Bootstrap values are from RAxML analyses and are in the order of (NUCL/N12/AA/DEGEN1). “*” represents a bootstrap value of ≥ 95 , with a single “*” representing all four bootstrap values ≥ 95 and “-” indicates that this relationship was not recovered in the analysis.

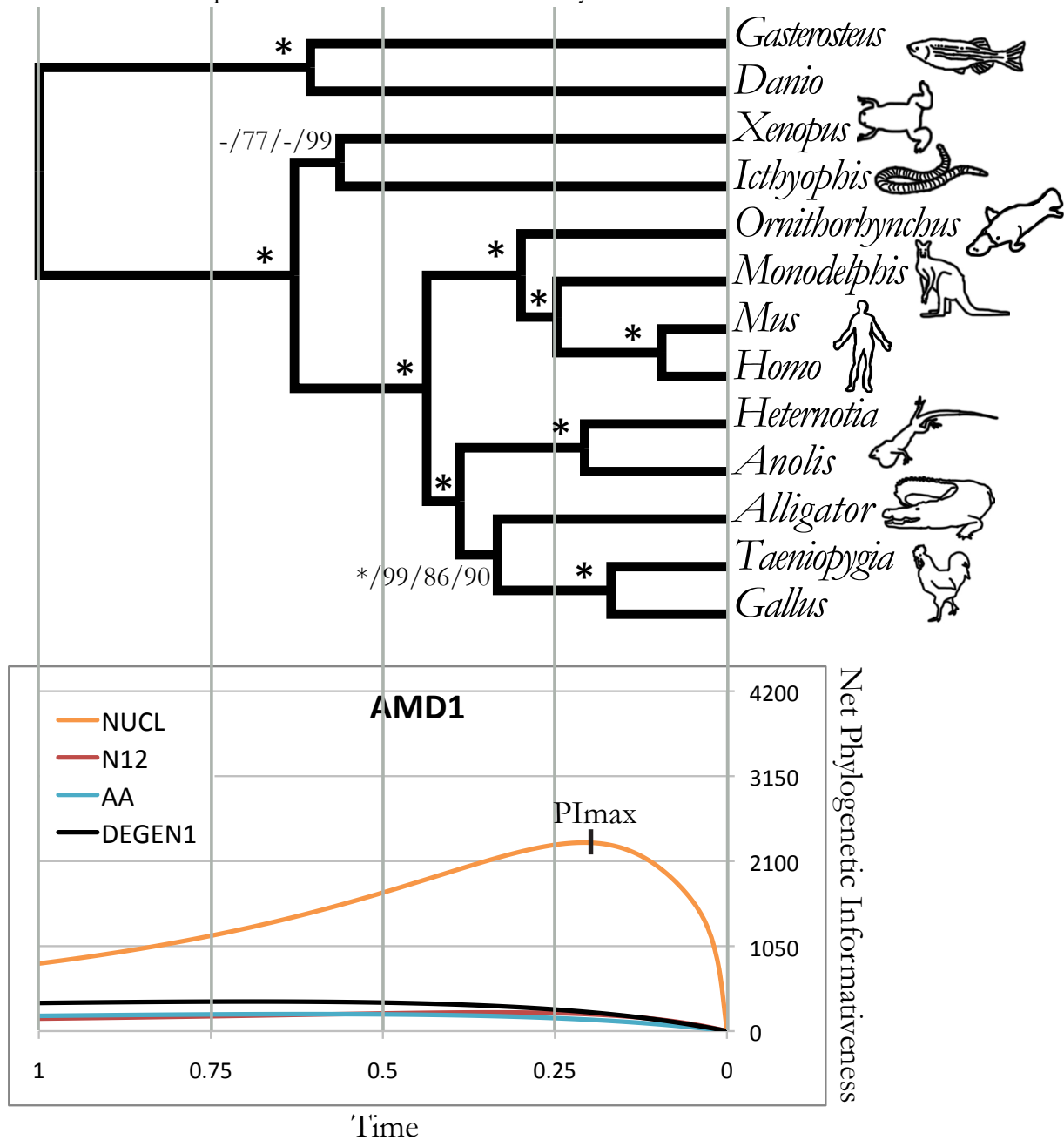


Figure 3: Scatter plot of uncorrected pair-wise distance (p-dist) between Human-Mouse against time of maximum phylogenetic informativeness (PI_{max}) ($r^2=0.02$, $\sigma=0.2$). The lack of correlation between p-dist and PI_{max} supports the conclusion that p-dist is a poor measure of phylogenetic informativeness.

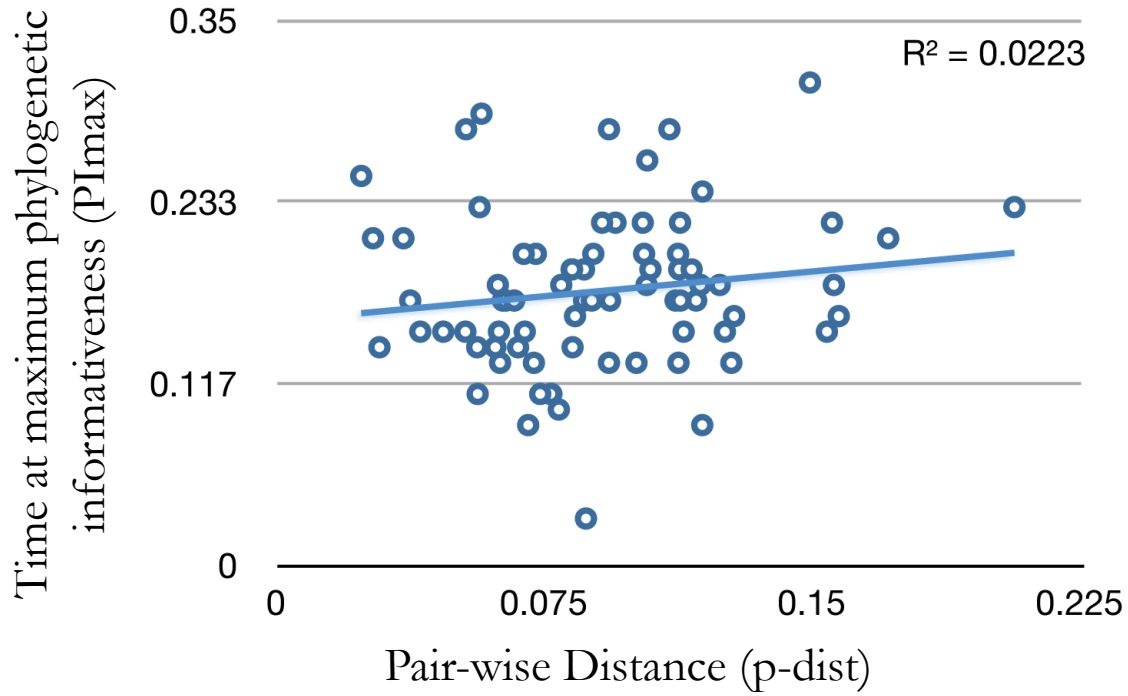


Figure 4: Net phylogenetic informativeness (PI) of the 75 markers using the NUCL (all nucleotides) dataset. Genes are sorted according to their P_{Imax}, with genes exhibiting PI at shallower times being at the top right.

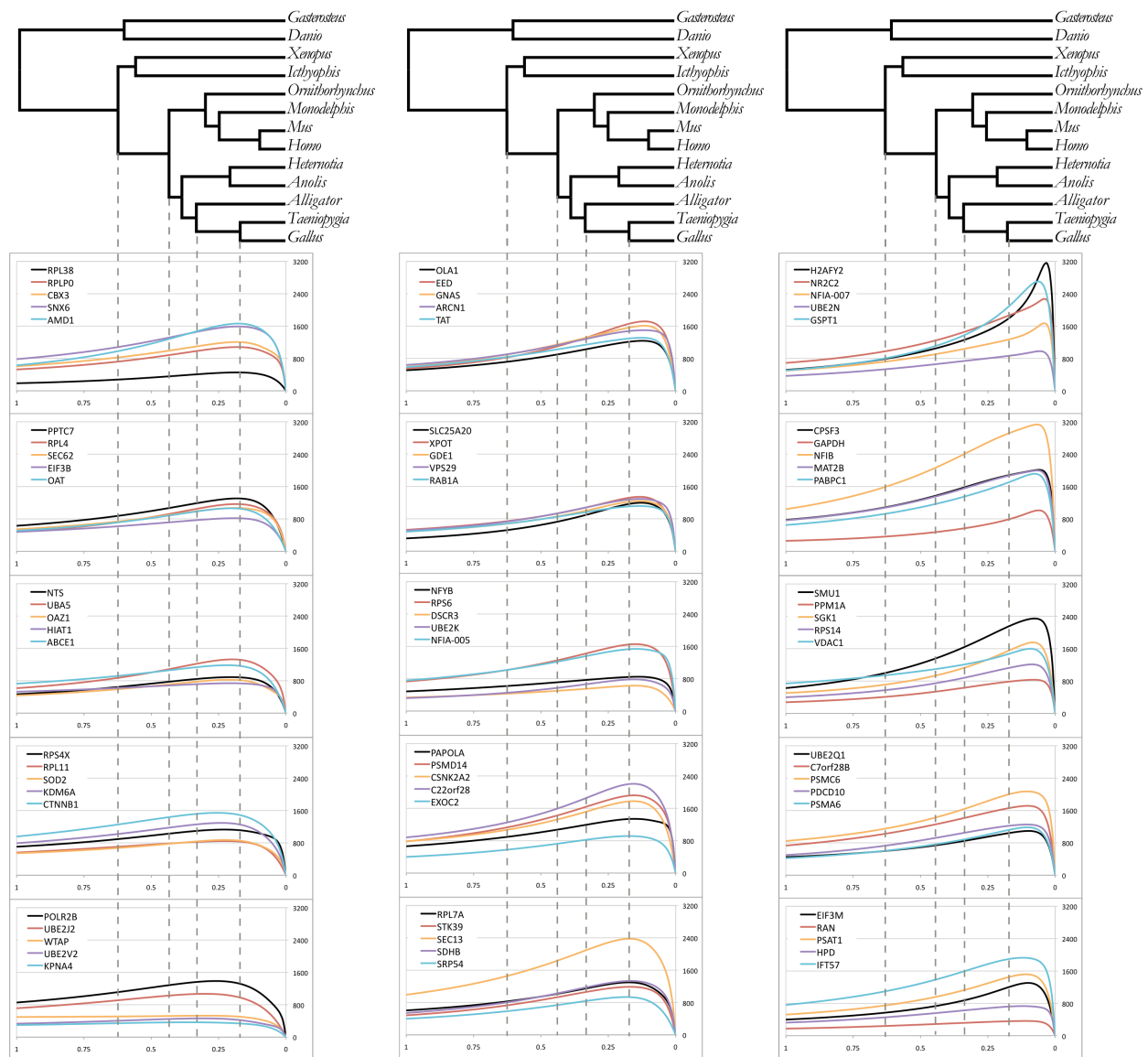


Figure 5: Net phylogenetic informativeness (PI) of the 75 markers using the N12 (1st and 2nd codon positions) dataset. Genes are sorted according to their P_{max}, with genes exhibiting PI at shallower times being at the top right.

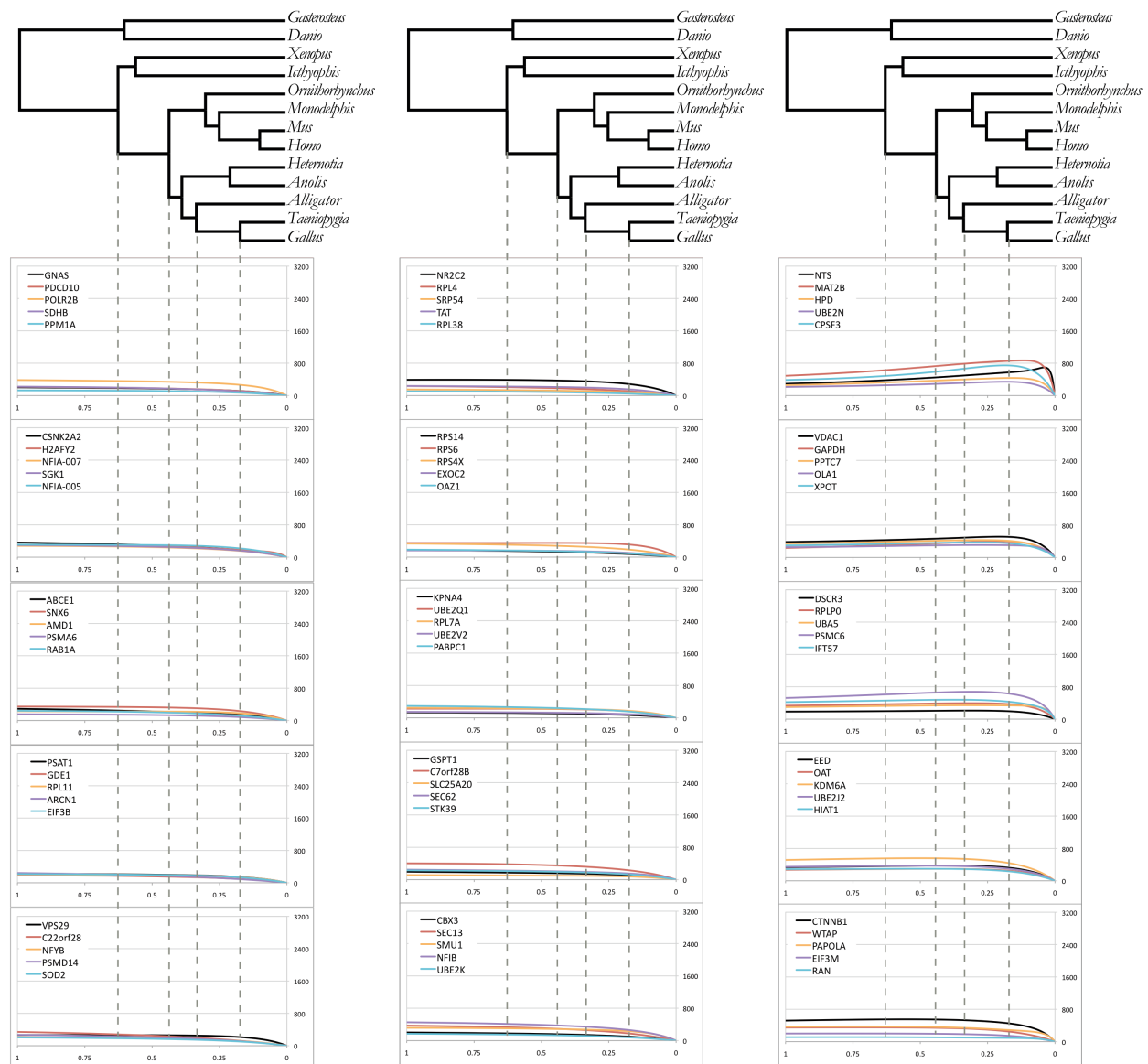


Figure 6: Net phylogenetic informativeness (PI) of 46 markers using the DEGEN1 (nucleotides taking codon degeneracy into account) dataset. Genes are sorted according to their P_{max}, with genes exhibiting PI at shallower times at the top right. PI was not profiled for all genes, as ambiguity codes prevented PI from being calculated for some markers.

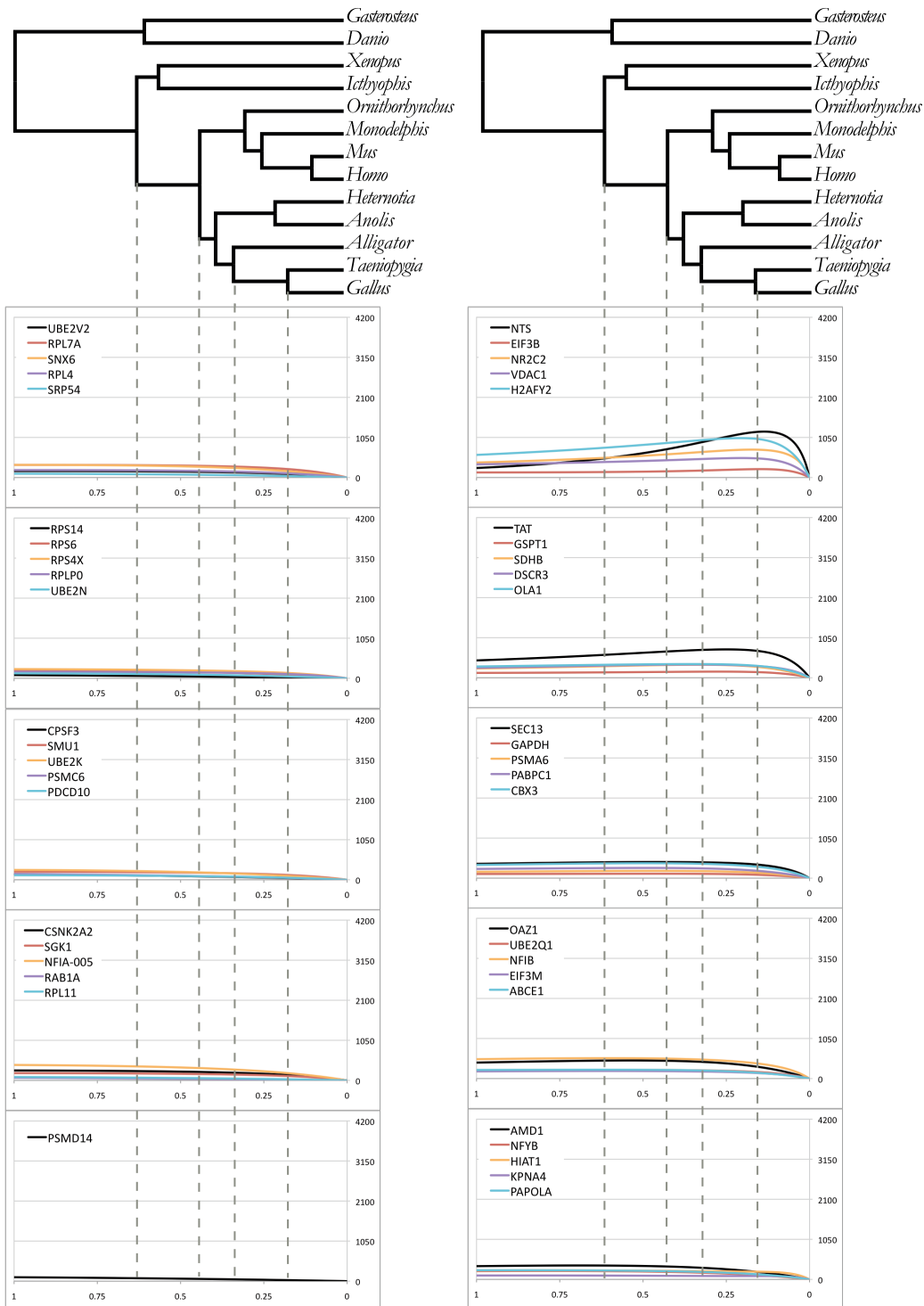


Figure 7: Net phylogenetic informativeness (PI) of the 75 markers using the AA (amino acids) dataset. Genes are sorted according to their PImax, with genes exhibiting PI at shallower times being at the top right.

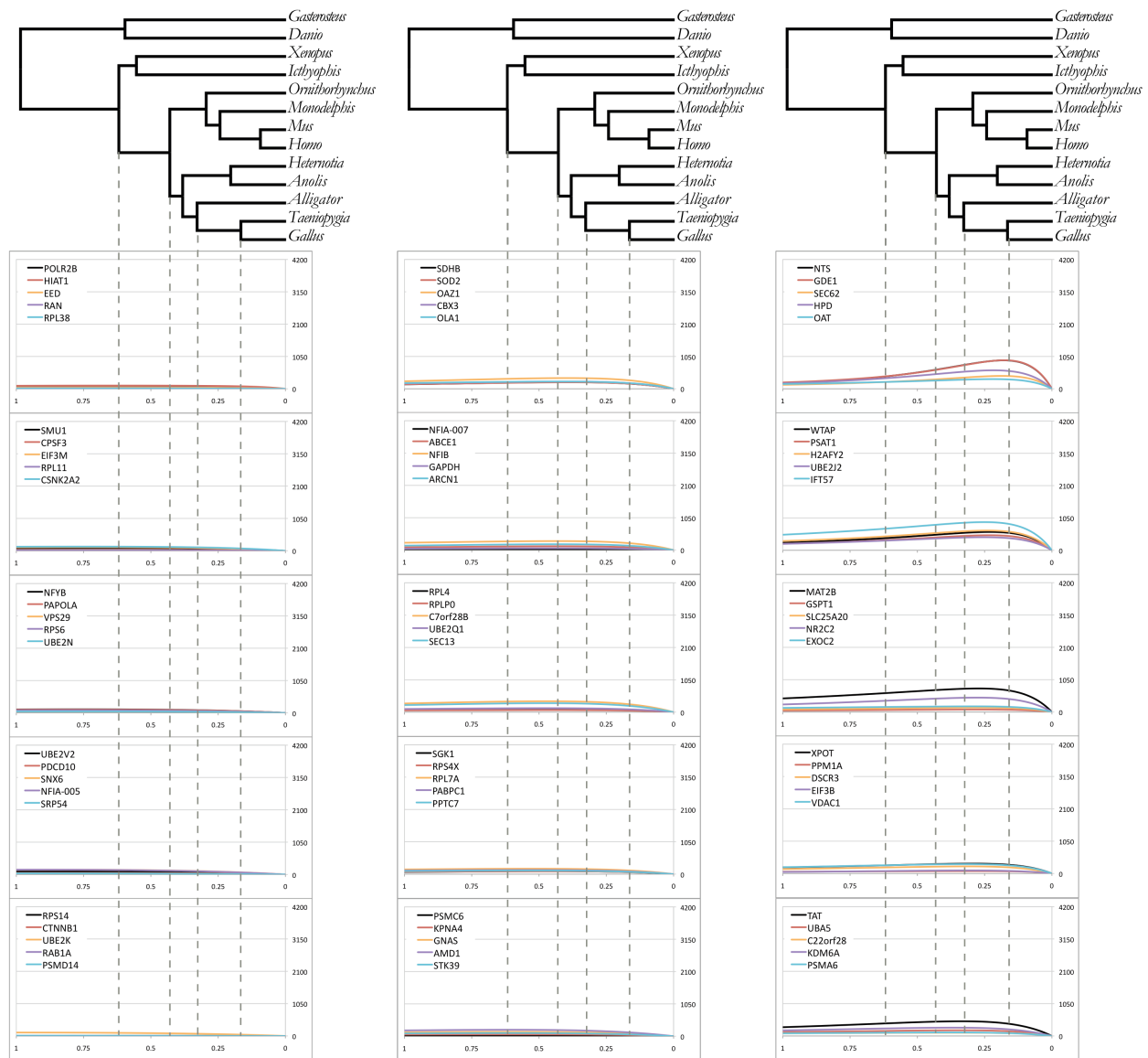
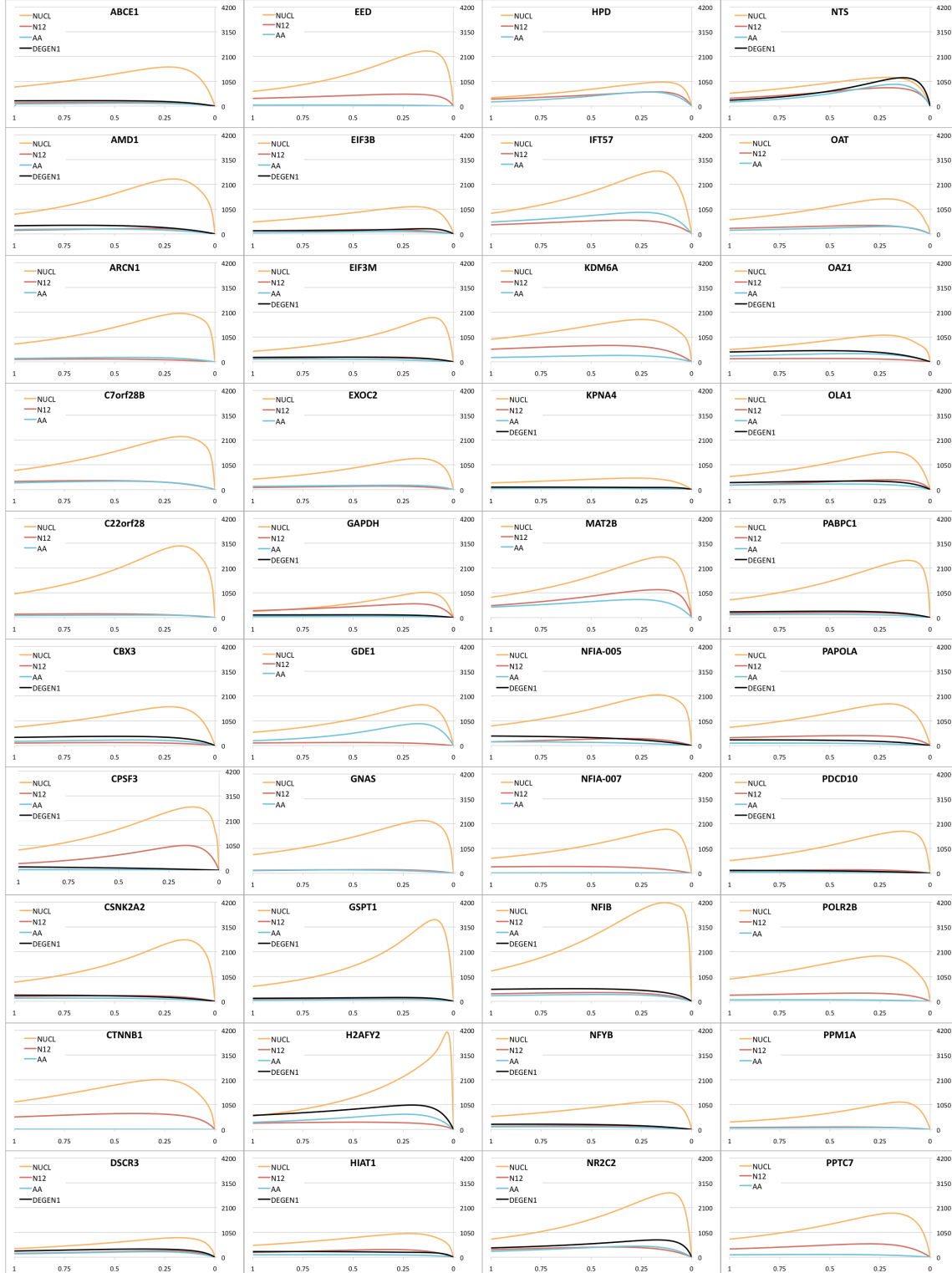
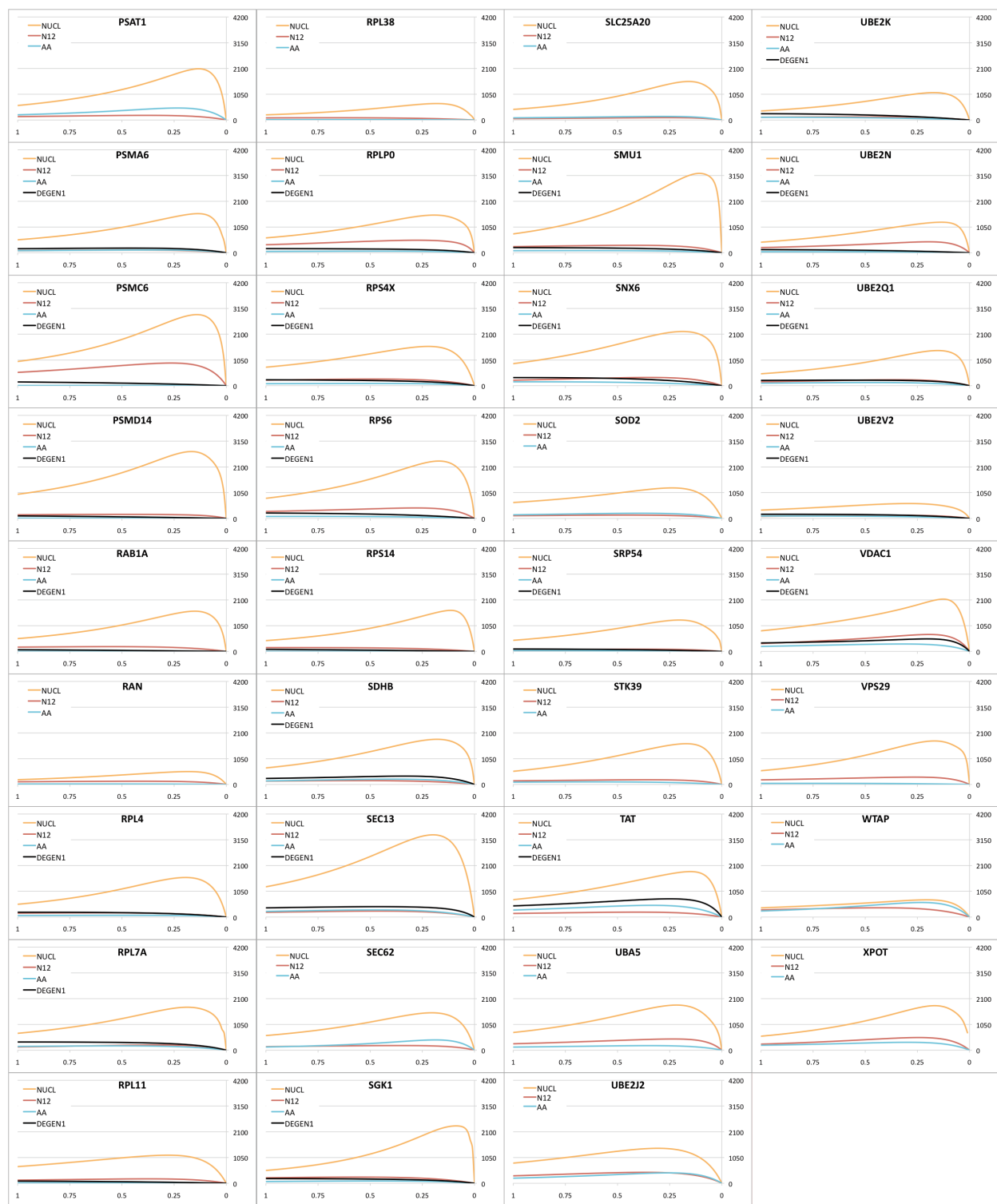


Figure 8: Individual gene comparison of the net phylogenetic informativeness (PI) of four different data types 1) NUCL (all nucleotides), 2) N12 (1st and 2nd codon positions), 3) AA (amino acids), and 4) DEGEN1 (nucleotides modified to take codon degeneracy into account). Genes are in alphabetical order. Ambiguity codes of DEGEN1 prevented PI from being calculated for some markers and are not shown.





SUPPLEMENTAL MATERIALS

Appendix A: Taxon sampling used for this study for 75 genes, including specimen numbers and identification numbers from Ensembl or Genbank (begins with JF) numbers. MVZ=Museum of Vertebrate Zoology, MKF=collection of Matthew K. Fujita.

Scientific Name	Specimen #	ABCE1	AMD1	ARCN1	C7orf28B	C22orf28	CBX3
<i>Alligator mississippiensis</i>	n/a						
<i>Alligator mississippiensis</i>	MVZ241485	JF264630					JF264634
<i>Anolis carolinensis</i>	n/a	ENSACAT0000003476	ENSACAT00000011200	ENSACAT00000006956	ENSACAT00000012252	ENSACAT00000014912	ENSACAT00000012990
<i>Danio rerio</i>	n/a	ENSDART00000015777	ENSDART00000064391	ENSDART00000047902	ENSDART00000112197	ENSDART00000111348	ENSDART00000012352
<i>Gallus gallus</i>	n/a	ENSGALT00000016194	ENSGALT00000024267	ENSGALT00000012021	ENSGALT00000005479	ENSGALT00000020590	ENSGALT00000037831
<i>Gasterosteus aculeatus</i>	n/a	ENSGACT00000024658	ENSGACT00000008748	ENSGACT00000006906	ENSGACT00000012513	ENSGACT00000000758	ENSGACT00000005157
<i>Heteronotia binoei</i>	MKF226	JF264631	JF264633				JF264635
<i>Homo sapiens</i>	n/a	ENST00000296577	ENST00000368876	ENST00000359415	ENST00000316731	ENST00000216038	ENST00000396386
<i>Icthyophis bannanicus</i>	MVZ226265	JF264632					JF264636
<i>Monodelphis domestica</i>	n/a	ENSMODT00000000569	ENSMODT00000022740	ENSMODT00000017112	ENSMODT00000010416	ENSMODT00000001751	
<i>Mus musculus</i>	n/a	ENSMUST00000080536	ENSMUST00000072796	ENSMUST00000034607	ENSMUST00000031621	ENSMUST00000001834	ENSMUST00000081455
<i>Ornithobrychus anatinus</i>	n/a	ENSOANT00000021944	ENSOANT00000022441	ENSOANT00000005959	ENSOANT00000001187	ENSOANT00000008775	ENSOANT00000014582
<i>Taeniopygia guttata</i>	n/a	ENSTGUT00000002812	ENSTGUT00000012549		ENSTGUT00000009125	ENSTGUT00000011413	ENSTGUT00000002996
<i>Xenopus tropicalis</i>	n/a	ENSXETT00000053891	ENSXETT00000046135	ENSXETT00000046606	ENSXETT00000051883	ENSXETT00000025212	ENSXETT00000026572

Scientific Name	Specimen #	CPSF3	CSNK2A2	CTNNB1	DSCR3	EED	EIF3B
<i>Alligator mississippiensis</i>	n/a						
<i>Alligator mississippiensis</i>	MVZ.241485						JF264641
<i>Anolis carolinensis</i>	n/a	ENSACAT00000001924	ENSACAT00000008468	ENSACAT00000004530	ENSACAT00000009093	ENSACAT00000002498	ENSACAT00000008991
<i>Danio rerio</i>	n/a	ENSDART00000040292	ENSDART00000045488	ENSDART00000038495	ENSDART00000036375	ENSDART00000059345	ENSDART00000083129
<i>Gallus gallus</i>	n/a	ENSGALT00000026493	ENSGALT00000001304	ENSGALT00000037585	ENSGALT00000025863	ENSGALT00000022764	ENSGALT00000006789
<i>Gasterosteus aculeatus</i>	n/a	ENSGACT00000011824	ENSGACT00000020807	ENSGACT00000008011	ENSGACT00000027638	ENSGACT00000019753	ENSGACT00000022154
<i>Heteronotia binoei</i>	MKF226	JF264637	JF264638		JF264639		
<i>Homo sapiens</i>	n/a	ENST00000460593	ENST00000262506	ENST00000447363	ENST00000288304	ENST00000327320	ENST00000314800
<i>Ichthyophis bannanicus</i>	MVZ.226265				JF264640		
<i>Monodelphis domestica</i>	n/a	ENSMODT00000018712	ENSMODT00000017930	ENSMODT00000014262	ENSMODT00000026714	ENSMODT00000005356	ENSMODT00000012445
<i>Mus musculus</i>	n/a	ENSMUST00000067284	ENSMUST00000056919	ENSMUST00000007130	ENSMUST00000023615	ENSMUST00000032850	ENSMUST00000075771
<i>Ornithorynchus anatinus</i>	n/a	ENSOANT00000013660	ENSOANT00000019189	ENSOANT00000010867	ENSOANT00000010234	ENSOANT00000014147	ENSOANT00000006843
<i>Taeniopygia guttata</i>	n/a	ENSTGUT00000013563	ENSTGUT00000004937	ENSTGUT00000004994	ENSTGUT00000005148	ENSTGUT00000013522	ENSTGUT00000008501
<i>Xenopus tropicalis</i>	n/a	ENSXETT00000018123	ENSXETT00000030055	ENSXETT00000030302	ENSXETT00000028561	ENSXETT00000022912	ENSXETT00000044504

Scientific Name	Specimen #	EIF3M	EXOC2	GAPDH	GDE1	GNAS	GSPT1
<i>Alligator mississippiensis</i>	n/a			ES3719			
<i>Alligator mississippiensis</i>	MVZ241485	JF264642					
<i>Anolis carolinensis</i>	n/a	ENSACAT00000002586	ENSACAT00000017798	ENSACAT00000006562	ENSACAT00000001563	ENSACAT00000014380	ENSACAT00000004073
<i>Danio rerio</i>	n/a	ENSDART00000010256	ENSDART00000077962	ENSDART00000063800		ENSDART00000064907	ENSDART00000016780
<i>Gallus gallus</i>	n/a	ENSGALT00000019782	ENSGALT00000020938	ENSGALT00000023323	ENSGALT00000039934	ENSGALT00000022385	ENSGALT00000041131
<i>Gasterosteus aculeatus</i>	n/a	ENSGACT00000021178	ENSGACT00000006621	ENSGACT00000013612		ENSGACT00000021968	ENSGACT00000005765
<i>Heteronotia binoei</i>	MKF226	JF264643					JF264645
<i>Homo sapiens</i>	n/a	ENST00000323213	ENST00000448181	ENST00000396858	ENST00000353258	ENST00000423027	ENST00000219627
<i>Icthyophis bannanicus</i>	MVZ226265	JF264644					JF264646
<i>Monodelphis domestica</i>	n/a	ENSMODT00000012326	ENSMODT00000010601	ENSMODT00000038970	ENSMODT00000007908	ENSMODT00000027257	ENSMODT00000005947
<i>Mus musculus</i>	n/a	ENSMUST00000111109	ENSMUST00000021785	ENSMUST00000101008	ENSMUST00000038791	ENSMUST00000076605	ENSMUST00000080030
<i>Ornithorynchus anatinus</i>	n/a	ENSOANT00000023857	ENSOANT00000010113	ENSOANT00000007125	ENSOANT00000005108	ENSOANT00000017127	ENSOANT00000014965
<i>Taeniopygia guttata</i>	n/a	ENSTGUT00000005128	ENSTGUT00000002228	ENSTGUT00000013888	ENSTGUT00000017100	ENSTGUT00000010130	ENSTGUT00000005366
<i>Xenopus tropicalis</i>	n/a			ENSXETT00000055696	ENSXETT00000047019	ENSXETT00000042648	ENSXETT00000012173

Scientific Name	Specimen #	H2AFY2	HIAT1	HPD	IFT57	KDM6A	KPNA4
<i>Alligator mississippiensis</i>	n/a			ES316653			
<i>Alligator mississippiensis</i>	MVZ241485			JF264648			
<i>Anolis carolinensis</i>	n/a	ENSACAT00000017852	ENSACAT00000003594	ENSACAT00000011308		ENSACAT00000017962	ENSACAT00000007314
<i>Danio rerio</i>	n/a	ENSDART00000058162	ENSDART00000063207	ENSDART00000020472	ENSDART00000012919		ENSDART00000034456
<i>Gallus gallus</i>	n/a	ENSGALT00000007651	ENSGALT00000008549	ENSGALT00000006930	ENSGALT00000024781	ENSGALT00000026141	ENSGALT00000015544
<i>Gasterosteus aculeatus</i>	n/a		ENSGACT00000022406	ENSGACT00000015747	ENSGACT00000020891		ENSGACT00000013375
<i>Heteronotia binoei</i>	MKF226		JF264647		JF264649	JF264650	
<i>Homo sapiens</i>	n/a	ENST00000373255	ENST00000370152	ENST00000289004	ENST00000457963	ENST00000433797	ENST00000334256
<i>Icthyophis bannanicus</i>	MVZ226265						JF264651
<i>Monodelphis domestica</i>	n/a		ENSMODT00000004616	ENSMODT00000015491	ENSMODT00000022893	ENSMODT00000039307	ENSMODT00000020758
<i>Mus musculus</i>	n/a	ENSMUST00000020283	ENSMUST00000029570	ENSMUST00000031398	ENSMUST00000046777	ENSMUST00000052368	ENSMUST00000029353
<i>Ornithorynchus anatinus</i>	n/a	ENSOANT00000019241	ENSOANT00000019102		ENSOANT00000011935	ENSOANT00000010353	ENSOANT00000011092
<i>Taeniopygia guttata</i>	n/a	ENSTGUT00000004272	ENSTGUT00000005517	ENSTGUT00000006350	ENSTGUT00000014172	ENSTGUT000000006361	ENSTGUT00000011634
<i>Xenopus tropicalis</i>	n/a	ENSXETT00000037396	ENSXETT00000028894	ENSXETT00000017017	ENSXETT00000055345	ENSXETT00000010467	ENSXETT00000019107

Scientific Name	Specimen #	MAT2B	NFIA-005	NFIA-007	NFIB	NFYB	NR2C2
<i>Alligator mississippiensis</i>	n/a						
<i>Alligator mississippiensis</i>	MVZ241485		JF264653		JF264655		
<i>Anolis carolinensis</i>	n/a	ENSACAT00000014603	ENSACAT00000007756	ENSACAT00000007756	ENSACAT00000017867	ENSACAT00000016083	ENSACAT00000012268
<i>Danio rerio</i>	n/a	ENDSART00000003978	ENDSART000000090242	ENDSART000000090237	ENDSART000000044670	ENDSART000000099770	
<i>Gallus gallus</i>	n/a	ENSGALT00000002683	ENSGALT00000017773	ENSGALT000000038014	ENSGALT00000008735	ENSGALT000000037651	ENSGALT00000013876
<i>Gasterosteus aculeatus</i>	n/a		ENSGACT000000004145	ENSGACT000000004145	ENSGACT000000026067		
<i>Heteronotia binoei</i>	MKF226		JF264654			JF264656	
<i>Homo sapiens</i>	n/a	ENST00000321757	ENST00000371185	ENST00000371187	ENST00000380934	ENST00000240055	ENST00000396874
<i>Icthyophis bannanicus</i>	MVZ226265	JF264652				JF264657	
<i>Monodelphis domestica</i>	n/a	ENSMODT00000009630	ENSMODT00000003069	ENSMODT00000003069	ENSMODT00000019093	ENSMODT00000002682	ENSMODT00000008911
<i>Mus musculus</i>	n/a	ENSMUST00000101347	ENSMUST00000107055	ENSMUST00000107056	ENSMUST00000050872	ENSMUST00000105309	ENSMUST00000006047
<i>Ornithorynchus anatinus</i>	n/a	ENSOANT00000007343	ENSOANT00000021222	ENSOANT00000019758	ENSOANT00000016052	ENSOANT00000008818	ENSOANT00000017254
<i>Taeniopygia guttata</i>	n/a	ENSTGUT00000001541	ENSTGUT00000010179	ENSTGUT00000010179	ENSTGUT00000004855	ENSTGUT00000011626	ENSTGUT00000011047
<i>Xenopus tropicalis</i>	n/a	ENSXETT00000032976	ENSXETT00000019479	ENSXETT00000019478	ENSXETT00000001859	ENSXETT000000042558	ENSXETT00000010442

Scientific Name	Specimen #	NTS	OAT	OAZ1	OLA1	PABPC1	PAPOLA
<i>Alligator mississippiensis</i>	n/a		ES316854	ES320547			
<i>Alligator mississippiensis</i>	MVZ241485			JF264658	JF264661	JF264662	
<i>Anolis carolinensis</i>	n/a		ENSACAT00000017625		ENSACAT00000017111	ENSACAT00000016610	ENSACAT00000013626
<i>Danio rerio</i>	n/a			ENS DART000000105532	ENS DART00000045299	ENS DART00000032474	ENS DART00000015154
<i>Gallus gallus</i>	n/a	ENSGALT00000018184	ENSGALT00000038658	ENSGALT00000001183	ENSGALT00000015187	ENSGALT00000023333	ENSGALT00000037556
<i>Gasterosteus aculeatus</i>	n/a		ENSGACT00000003381	ENSGACT00000016744	ENSGACT00000013368	ENSGACT00000004695	ENSGACT00000003581
<i>Heteronotia binoei</i>	MKF226			JF264659		JF264663	JF264665
<i>Homo sapiens</i>	n/a	ENST00000256010	ENST00000368845	ENST00000322297	ENST00000424954	ENST00000347137	ENST00000392990
<i>Icthyophis bannanicus</i>	MVZ226265			JF264660		JF264664	
<i>Monodelphis domestica</i>	n/a	ENSMODT00000026101	ENSMODT00000021066	ENSMODT00000005918	ENSMODT00000011836	ENSMODT00000036071	ENSMODT00000016424
<i>Mus musculus</i>	n/a	ENSMUST00000020040	ENSMUST00000084500	ENSMUST00000060987	ENSMUST00000100015	ENSMUST00000001809	ENSMUST00000101063
<i>Ornithorynchus anatinus</i>	n/a	ENSOANT00000010600	ENSOANT00000009207	ENSOANT00000001022	ENSOANT00000013706	ENSOANT00000011635	ENSOANT00000021786
<i>Taeniopygia guttata</i>	n/a	ENSTGUT00000008249	ENSTGUT00000011917	ENSTGUT00000000329	ENSTGUT00000009088	ENSTGUT00000012575	ENSTGUT00000013221
<i>Xenopus tropicalis</i>	n/a		ENSXETT00000021114	ENSXETT00000029891	ENSXETT00000018601	ENSXETT00000014537	ENSXETT00000030476

Scientific Name	Specimen #	PDCD10	POLR2B	PPM1A	PPTC7	PSAT1	PSMA6
<i>Alligator mississippiensis</i>	n/a					ES316685	ES608
<i>Alligator mississippiensis</i>	MVZ241485						JF264669
<i>Anolis carolinensis</i>	n/a	ENSACAT00000017961	ENSACAT00000002264	ENSACAT00000015231	ENSACAT00000001872	ENSACAT000000006128	ENSACAT00000001856
<i>Danio rerio</i>	n/a	ENSDART00000019063	ENSDART000000041408	ENSDART000000076527	ENSDART00000015214		ENSDART00000012859
<i>Gallus gallus</i>	n/a	ENSGALT00000015429	ENSGALT00000037965	ENSGALT00000019443	ENSGALT00000007244	ENSGALT000000024496	ENSGALT00000016353
<i>Gasterosteus aculeatus</i>	n/a	ENSGACT00000013114	ENSGACT000000021502	ENSGACT00000011157	ENSGACT000000022651	ENSGACT000000006399	ENSGACT00000016600
<i>Heteronotia binoei</i>	MKF226	JF264666				JF264668	JF264670
<i>Homo sapiens</i>	n/a	ENST00000475915	ENST00000396241	ENST00000325658	ENST00000354300	ENST00000347159	ENST00000261479
<i>Ichthyophis bannanicus</i>	MVZ226265	JF264667					JF264671
<i>Monodelphis domestica</i>	n/a	ENSMODT00000036561	ENSMODT00000026361	ENSMODT00000010930	ENSMODT00000014729	ENSMODT00000016516	ENSMODT00000016318
<i>Mus musculus</i>	n/a	ENSMUST00000029424	ENSMUST00000031167	ENSMUST00000021514	ENSMUST00000119015	ENSMUST000000025542	ENSMUST00000021412
<i>Ornithorynchus anatinus</i>	n/a	ENSOANT00000017424	ENSOANT00000017764	ENSOANT00000018794	ENSOANT00000007571	ENSOANT000000004707	ENSOANT00000022172
<i>Taeniopygia guttata</i>	n/a	ENSTGUT00000011549	ENSTGUT00000003066	ENSTGUT00000013537	ENSTGUT00000006637	ENSTGUT00000003535	ENSTGUT00000012478
<i>Xenopus tropicalis</i>	n/a	ENSXETT00000049474	ENSXETT00000011733	ENSXETT00000036927	ENSXETT00000045244	ENSXETT00000007823	ENSXETT00000054899

Scientific Name	Specimen #	PSMC6	PSMD14	RAB1A	RAN	RPL4	RPL7A
<i>Alligator mississippiensis</i>	n/a					ES316508	ES583
<i>Alligator mississippiensis</i>	MVZ241485	JF264672	JF264673		JF264675	JF264678	JF264681
<i>Anolis carolinensis</i>	n/a	ENSACAT00000006886	ENSACAT00000011931	ENSACAT00000001844	ENSACAT00000007195	ENSACAT00000009829	ENSACAT00000015040
<i>Danio rerio</i>	n/a	ENDSART00000053834	ENDSART00000091901	ENDSART00000043002		ENDSART00000060363	ENDSART00000004238
<i>Gallus gallus</i>	n/a	ENSGALT00000037197	ENSGALT00000018142	ENSGALT00000014281	ENSGALT00000004055	ENSGALT00000038315	ENSGALT00000005066
<i>Gasterosteus aculeatus</i>	n/a	ENSGACT00000007662	ENSGACT00000020049	ENSGACT00000005122		ENSGACT00000006665	ENSGACT00000021129
<i>Heteronotia binoei</i>	MKF226				JF264676	JF264679	
<i>Homo sapiens</i>	n/a	ENST00000216407	ENST00000263639	ENST00000260638	ENST00000392369	ENST00000432669	ENST00000277425
<i>Icthyophis bannanicus</i>	MVZ226265			JF264674	JF264677	JF264680	JF264682
<i>Monodelphis domestica</i>	n/a	ENSMODT00000018452	ENSMODT00000006717	ENSMODT00000004174	ENSMODT00000019816	ENSMODT00000036322	ENSMODT00000016254
<i>Mus musculus</i>	n/a	ENSMUST00000022380	ENSMUST00000028278	ENSMUST00000074741	ENSMUST00000111343	ENSMUST00000034966	ENSMUST00000102898
<i>Ornithorynchus anatinus</i>	n/a	ENSOANT00000017157	ENSOANT00000009874	ENSOANT00000030441	ENSOANT00000010295	ENSOANT00000023550	ENSOANT00000011240
<i>Taeniopygia guttata</i>	n/a	ENSTGUT00000013650	ENSTGUT00000006958	ENSTGUT00000002422	ENSTGUT00000004882	ENSTGUT00000009825	ENSTGUT00000005689
<i>Xenopus tropicalis</i>	n/a	ENSXETT00000018862	ENSXETT00000035366	ENSXETT00000030593		ENSXETT00000034688	ENSXETT00000010943

Scientific Name	Specimen #	RPL11	RPL38	RPLP0	RPS4X	RPS6	RPS14
<i>Alligator mississippiensis</i>	n/a	ES5078	ES319935	ES2340	ES321882	ES321879	ES321806
<i>Alligator mississippiensis</i>	MVZ241485	JF264683	JF264686				JF264691
<i>Anolis carolinensis</i>	n/a	ENSACAT00000005076	ENSACAT00000014866	ENSACAT00000003794	ENSACAT00000014647	ENSACAT00000017706	ENSACAT00000015018
<i>Danio rerio</i>	n/a	ENDSART000000063870	ENDSART000000021069	ENDSART000000073462	ENDSART00000007957	ENDSART000000023336	ENDSART000000053208
<i>Gallus gallus</i>	n/a	ENSGALT000000040531	ENSGALT000000002220	ENSGALT000000011731	ENSGALT00000007717	ENSGALT000000024329	ENSGALT00000007299
<i>Gasterosteus aculeatus</i>	n/a	ENSGACT000000002790	ENSGACT000000004327	ENSGACT000000011351	ENSGACT000000024824	ENSGACT000000023657	ENSGACT000000027698
<i>Heteronotia binoei</i>	MKF226	JF264684	JF264687		JF264689	JF264690	JF264692
<i>Homo sapiens</i>	n/a	ENST00000458455	ENST00000439590	ENST00000392514	ENST00000316084	ENST00000315377	ENST00000407193
<i>Ichthyophis bannanicus</i>	MVZ226265	JF264685	JF264688				JF264693
<i>Monodelphis domestica</i>	n/a	ENSMODT000000016346	ENSMODT000000008449	ENSMODT000000019339	ENSMODT000000035235	ENSMODT000000005147	ENSMODT000000012078
<i>Mus musculus</i>	n/a	ENSMUST000000078981	ENSMUST000000082092	ENSMUST000000086519	ENSMUST000000033683	ENSMUST000000102814	ENSMUST000000122279
<i>Ornithorhynchus anatinus</i>	n/a	ENSOANT00000000540		ENSOANT000000010419	ENSOANT00000004685	ENSOANT000000005142	ENSOANT00000005084
<i>Taeniopygia guttata</i>	n/a	ENSTGUT000000001261		ENSTGUT000000009872	ENSTGUT00000004981	ENSTGUT000000004518	ENSTGUT000000019180
<i>Xenopus tropicalis</i>	n/a	ENSXETT000000002557	ENSXETT000000054097	ENSXETT000000041250	ENSXETT000000056754	ENSXETT000000049958	ENSXETT000000034484

Scientific Name	Specimen #	SDHB	SEC13	SEC62	SGK1	SLC25A20	SMU1
<i>Alligator mississippiensis</i>	n/a						
<i>Alligator mississippiensis</i>	MVZ241485	JF264694	JF264697				JF264702
<i>Anolis carolinensis</i>	n/a	ENSACAT00000016271	ENSACAT00000014914	ENSACAT00000015851	ENSACAT00000010308	ENSACAT00000012163	ENSACAT00000003478
<i>Danio rerio</i>	n/a	ENDSART00000062264	ENDSART00000077063		ENDSART00000035031	ENDSART00000059140	ENDSART00000098346
<i>Gallus gallus</i>	n/a	ENSGALT00000000694	ENSGALT00000013715	ENSGALT00000015276	ENSGALT00000031766	ENSGALT00000011304	ENSGALT00000003135
<i>Gasterosteus aculeatus</i>	n/a	ENSGACT00000006777			ENSGACT00000014583	ENSGACT00000001538	ENSGACT00000022220
<i>Heteronotia binoei</i>	MKF226	JF264695	JF264698	JF264699			JF264703
<i>Homo sapiens</i>	n/a	ENST00000375499	ENST00000383801	ENST00000480708	ENST00000367857	ENST00000440964	ENST00000397149
<i>Icthyophis bannanicus</i>	MVZ226265	JF264696		JF264700	JF264701		JF264704
<i>Monodelphis domestica</i>	n/a	ENSMODT00000003442	ENSMODT00000006949	ENSMODT00000036602	ENSMODT00000022227	ENSMODT00000017045	ENSMODT00000004300
<i>Mus musculus</i>	n/a	ENSMUST00000010007	ENSMUST00000032440	ENSMUST00000029256	ENSMUST00000020145	ENSMUST00000035222	ENSMUST00000030117
<i>Ornithorynchus anatinus</i>	n/a	ENSOANT00000018599	ENSOANT00000014867	ENSOANT00000012201	ENSOANT00000020700	ENSOANT00000012061	ENSOANT00000015665
<i>Taeniopygia guttata</i>	n/a	ENSTGUT00000017297	ENSTGUT00000010897	ENSTGUT00000011446	ENSTGUT00000012019	ENSTGUT00000009491	ENSTGUT00000001519
<i>Xenopus tropicalis</i>	n/a	ENSXETT00000029442	ENSXETT0000003059	ENSXETT00000020231	ENSXETT00000000246	ENSXETT00000020560	ENSXETT00000052109

Scientific Name	Specimen #	SNX6	SOD2	SRP54	STK39	TAT	UBA5
<i>Alligator mississippiensis</i>	n/a			ES316890		ES317837	
<i>Alligator mississippiensis</i>	MVZ241485	JF264705					
<i>Anolis carolinensis</i>	n/a	ENSACAT00000002518	ENSACAT00000004516	ENSACAT00000002230	ENSACAT00000001427	ENSACAT00000007792	ENSACAT00000000616
<i>Danio rerio</i>	n/a	ENDSART000000105121	ENDSART000000062556	ENDSART000000018678		ENDSART000000101472	ENDSART000000093046
<i>Gallus gallus</i>	n/a	ENSGALT000000016282	ENSGALT000000019062	ENSGALT000000016327	ENSGALT000000017779	ENSGALT000000001314	ENSGALT000000019134
<i>Gasterosteus aculeatus</i>	n/a	ENSGACT000000005459	ENSGACT000000011956	ENSGACT000000005684	ENSGACT000000007102	ENSGACT000000000602	
<i>Heteronotia binoei</i>	MKF226	JF264706		JF264708	JF264709		
<i>Homo sapiens</i>	n/a	ENST00000362031	ENST000000337404	ENST000000216774	ENST00000355999	ENST00000355962	ENST00000494238
<i>Ichthyophis bannanicus</i>	MVZ226265	JF264707					
<i>Monodelphis domestica</i>	n/a	ENSMODT000000018472	ENSMODT000000009516	ENSMODT000000016472	ENSMODT000000009571	ENSMODT000000002821	ENSMODT000000014874
<i>Mus musculus</i>	n/a	ENSMUST000000005798	ENSMUST000000007012	ENSMUST000000021407	ENSMUST000000102715	ENSMUST000000001720	ENSMUST000000035166
<i>Ornithorynchus anatinus</i>	n/a	ENSOANT000000022156	ENSOANT000000000987	ENSOANT000000022166	ENSOANT000000013507	ENSOANT000000013005	ENSOANT000000010821
<i>Taeniopygia guttata</i>	n/a	ENSTGUT000000012425	ENSTGUT000000011861	ENSTGUT000000012459	ENSTGUT000000007884	ENSTGUT000000008728	ENSTGUT000000004628
<i>Xenopus tropicalis</i>	n/a	ENSXETT000000026895	ENSXETT000000045789	ENSXETT000000026869	ENSXETT000000035375	ENSXETT000000050583	ENSXETT000000054036

Scientific Name	Specimen #	UBE2J2	UBE2K	UBE2N	UBE2Q1	UBE2V2	VDAC1
<i>Alligator mississippiensis</i>	n/a						
<i>Alligator mississippiensis</i>	MVZ241485			JF264712			
<i>Anolis carolinensis</i>	n/a	ENSACAT00000017064	ENSACAT00000011332	ENSACAT00000012593		ENSACAT00000012314	ENSACAT00000002834
<i>Danio rerio</i>	n/a	ENDSART000000086924	ENDSART000000025770		ENDSART000000097333	ENDSART000000026578	ENDSART000000066373
<i>Gallus gallus</i>	n/a	ENSGALT00000002864	ENSGALT000000023108	ENSGALT00000018421		ENSGALT000000024586	ENSGALT00000010508
<i>Gasterosteus aculeatus</i>	n/a	ENSGACT00000010489	ENSGACT000000023759	ENSGACT000000024916	ENSGACT000000008411	ENSGACT000000023737	ENSGACT000000027664
<i>Heteronotia binoei</i>	MKF226		JF264710	JF264713		JF264715	JF264717
<i>Homo sapiens</i>	n/a	ENST00000360466	ENST00000261427	ENST00000318066		ENST00000324746	ENST00000395047
<i>Ichthyophis bannanicus</i>	MVZ226265		JF264711	JF264714		JF264716	JF264718
<i>Monodelphis domestica</i>	n/a	ENSMODT00000020911	ENSMODT00000026174	ENSMODT00000012029	ENSMODT00000015211	ENSMODT00000012991	ENSMODT00000018029
<i>Mus musculus</i>	n/a	ENSMUST00000105582	ENSMUST00000031105	ENSMUST00000099329	ENSMUST00000053212	ENSMUST00000096234	ENSMUST00000109065
<i>Ornithorynchus anatinus</i>	n/a	ENSOANT00000010778	ENSOANT00000016400	ENSOANT00000010567		ENSOANT00000011570	ENSOANT00000021286
<i>Taeniopygia guttata</i>	n/a	ENSTGUT00000004247	ENSTGUT00000009159	ENSTGUT00000008606	ENSTGUT00000003648	ENSTGUT00000011367	ENSTGUT00000001351
<i>Xenopus tropicalis</i>	n/a	ENSXETT00000002268	ENSXETT00000054669	ENSXETT00000028257	ENSXETT00000033517	ENSXETT00000010542	ENSXETT00000034818

Scientific Name	Specimen #	VPS29	WTAP	XPOT
<i>Alligator mississippiensis</i>	n/a			
<i>Alligator mississippiensis</i>	MVZ241485			
<i>Anolis carolinensis</i>	n/a	ENSACAT00000001835	ENSACAT00000004475	ENSACAT00000003191
<i>Danio rerio</i>	n/a	ENSDART000000101246		ENSDART000000082119
<i>Gallus gallus</i>	n/a	ENSGALT00000007214	ENSGALT00000019061	ENSGALT00000015988
<i>Gasterosteus aculeatus</i>	n/a	ENSGACT00000006246		
<i>Heteronotia binoei</i>	MKF226		JF264719	JF264720
<i>Homo sapiens</i>	n/a	ENST00000447578	ENST00000358372	ENST00000332707
<i>Icthyophis bannanicus</i>	MVZ226265			
<i>Monodelphis domestica</i>	n/a	ENSMODT00000014756	ENSMODT00000009313	
<i>Mus musculus</i>	n/a	ENSMUST00000031418	ENSMUST00000007007	ENSMUST00000039810
<i>Ornithorynchus anatinus</i>	n/a	ENSOANT00000007569	ENSOANT00000019058	ENSOANT00000017327
<i>Taeniopygia guttata</i>	n/a	ENSTGUT00000016168	ENSTGUT00000011878	ENSTGUT00000006723
<i>Xenopus tropicalis</i>	n/a	ENSXETT00000052892	ENSXETT00000045788	ENSXETT00000055066

CHAPTER 2
THE PHYLOGENETIC PLACEMENT OF TURTLES

ABSTRACT

With respect to higher-level vertebrate phylogeny, one of the outstanding open questions is in regard to the phylogenetic placement of turtles. Five different positions for turtles in the amniote phylogeny have been previously proposed including: turtles as basal amniotes, as basal reptiles, as the sister to lepidosaurs (tuatara, lizards, snakes, amphisbaenians), as sister to crocodilians, and as sister to archosaurs (crocodilians, birds, and other dinosaurs). To address the phylogenetic placement of turtles, I generated a vertebrate phylogeny using 128 taxa and 75 protein-coding genes. In addition, I performed analyses to identify unstable (“rogue”) taxa, and tested for the most appropriate data-type (NUCL—nucleotides, N12—first and second codon positions, DEGEN1—codons degenerated using ambiguity codes, AA—amino acids) and data partitioning strategy. Results using alternative data-types varied, with the NUCL dataset supporting a turtle-crocodilian relationship, the N12 dataset supporting a turtle-archosaur relationship, and DEGEN1 and AA datasets supporting a turtle-lepidosaur relationship. Statistical tests showed that the data are not biased by either GC content or missing data. Approximately unbiased topology tests performed on each dataset lend strongest support to the hypotheses linking turtles with archosaurs, either with turtles as the sister of extant birds plus crocodilians, or with turtles as the sister taxon of crocodilians. Further exploration of the data show that less than 200 3rd codon sites are driving the analyses towards a turtle-crocodilian relationship. A recent paleontological study provided additional morphological support for the turtle-archosaur relationship (Bhullar et al. 2009).

INTRODUCTION

The ‘great Tree of Life’ was first described by Charles Darwin in his seminal work “The Origin of Species” (Darwin 1859). Essentially all fields of biology, including health care, agriculture, and conservation biology, depend on an understanding of the tree of life (Cracraft et al. 2002, Yates et al. 2004). With the advent and widespread use of molecular phylogenetics in the 1990s, progress towards understanding the tree of life has increased rapidly. Efforts have culminated in a National Science Foundation initiative called “Assembling the Tree of Life” (AToL). AToL began in 2001, and ~115 projects have been funded spanning nearly all living organisms (<http://www.phylo.org/atol>; <http://www.nsf.gov>).

Vertebrata, comprised of approximately 62,800 currently described species (IUCN 2010), is a well-studied group (Thomson & Shaffer 2009). Great advances have been made in determining the relationships within vertebrates, but many uncertainties remain (Meyer & Zardoya 2003), including 1) monophyly versus paraphyly of Agnatha (jawless fishes), 2) relationships within Actinopterygii (ray-finned fishes), 3) the relationship between Actinistia (coelacanths), Dipnoi (lungfish), and tetrapods, 4) the monophyly versus paraphyly of amphibians, as well as relationships between the three living orders: Gymnophiona (caecilians), Caudata (salamanders), Anura (frogs), 5) placement of Testudines (turtles) within Amniota, 6) basal relationships within Aves (birds), and 7) relationships between the three major mammalian clades (monotremes, marsupials, and placentals). A major goal of this study was to build a robust vertebrate phylogeny to address as many of these outstanding phylogenetic uncertainties as possible, while providing a well-supported phylogenetic tree

describing relationships both between and within the major vertebrate clades. The focus of this chapter is the phylogenetic placement of turtles within amniotes. Relationships within other vertebrate groups are discussed in Chapters 3 (turtles) and Chapter 5 (Actinopterygii, Amphibia, Squamata, Aves, and Mammalia).

Turtle Placement

The correct phylogenetic placement of turtles (Order Testudines) within amniotes has been debated for more than 150 years (Meyer & Zardoya 2003). The morphology that characterizes turtles is sufficiently distinct from that of all other vertebrates (i.e. bony shell, anapsid skull) that it has been likened to a palimpsest (a parchment from which earlier writing has been erased to clear it for new writing; Broom 1924), obscuring evidence of ancestry. A series of morphological and molecular studies (and re-analysis of past studies) have attempted to resolve this problem (Table 1), resulting in five major competing hypotheses for turtle phylogenetic placement (Figure 1):

1) Turtles as the most basal amniotes (Figure 1A, Table 1)

Historically, amniotes were classified by the number of temporal skull openings present, with no openings (anapsid) found in turtles, a single opening (synapsid) found in mammals, and two openings (diapsid) found in non-turtle reptiles and birds (Gaffney 1980). The anapsid condition is believed to be the most plesiomorphic and, as the sole extant vertebrates with anapsid skulls, turtles were traditionally placed as the most basal amniote (sister to diapsids and synapsids). No molecular studies have supported this hypothesis.

2) Turtles as the most basal reptiles (Figure 1B, Table 1)

Contrary to Gaffney (1980), much earlier studies (Goodrich 1916, Boulenger 1918) argued for the affinity of turtles with diapsids based on the fenestrated pelvis and hooked 5th metatarsal. Subsequent morphological and paleontological studies have corroborated these findings, placing mammals as the sister group to all other amniotes. Based on these arguments, anapsid turtles were placed as the sister group to Diapsida (Carroll 1988, Gauthier et al. 1988, Laurin & Reisz 1995, Rieppel & deBraga 1996, Lee 1997, Reisz 1997, Lyson et al 2010). The majority of morphological studies, along with embryological, molecular studies, and combined morphological and molecular studies, support the position of turtles as the most basal reptile (see Table 1 for references).

Within the studies supporting this hypothesis, there is debate on the turtle's closest extinct relative: pareiasaurs (Gregory 1946, Lee 1995, 1996, 1997, Rieppel & Reisz 1999), captorhniidae (Gaffney & Meeker 1983), procolophonids (Reisz & Laurin 1991, Laurin & Reisz 1995), or Sauropterygia (daBraga & Rieppel 1997, Rieppel & Reisz 1999).

3) Turtles as the sister group to Lepidosauria (Figure 1C, Table 1)

A single molecular study (Hedges 1994) and four morphological studies (Rieppel & deBraga 1996, deBraga & Rieppel 1997, Rieppel & Reisz 1999, Hill 2005) present data that support a sister relationship between turtles and Lepidosauria. An additional embryological study (Werneburg & Sanchez-Villagra, 2009) settled on the position of turtles as basal reptiles, but their data also showed some support for turtles as the sister group of

Lepidosauria. However, reanalysis of the morphological data of Rieppel & deBraga (1996) supporting the turtle-lepidosaur relationship resulted in support for turtles as basal reptiles (Lee 1997, Wilkinson et al 1997, Lyson et al. 2010).

4) Turtles as the sister group to Archosauria (Figure 1D, Table 1)

The majority of molecular studies, along with two morphological studies, support a sister relationship between turtles and Archosauria (see Table 1 for references). However, eight of the eleven molecular studies are based on varying amounts of mitochondrial DNA (mtDNA); since mtDNA evolves as a single locus, these eight studies are effectively based on the same locus. In addition, Rieppel (2000) reviewed the morphological characters that supported the turtle-archosaur relationship, and found many of these traits were incorrectly coded, and correct coding would place turtles as basal reptiles or sister to lepidosaurs.

5) Turtles as the sister group to Crocodylia (Figure 1E, Table 1)

A series of molecular studies have supported turtles as the sister group of Crocodylia (Manen et al. 1997, Kirsch & Meyer 1988, Hedges & Poling 1999, Mannen & Li 1999, Cao et al. 2000). From a morphological standpoint, this hypothesis is problematic, as paleontologists generally accept the monophyly of Archosauria (Zardoya & Meyer 2001).

In general, the majority of neontological and paleontological morphological data supported turtles as basal reptiles, while molecular data have supported turtles as the sister group to Archosauria (Lyson et al. 2010), but most hypotheses (except 1 and 5) have had both morphological and molecular data supporting them in some studies. Identical results are expected whether morphological or molecular characters are used, since organisms have a single phylogenetic history. However, differences in results between morphological and molecular studies are possible as morphological studies are able to include fossil taxa that affect polarities throughout the tree (Gauthier et al. 1988, Conrad 2008). In addition, both morphological and molecular data are sensitive to missing data (Lyson et al. 2010), and homoplasy (Eernisse & Kluge 1993, Rieppel & Reisz 1999). Morphological data are especially sensitive to the coding of characters (Lee 1997) and lack of intermediate forms (Lee 1993), while molecular data are susceptible to rate heterogeneity (Lyson et al. 2010) and substitution bias (Marshall 1992), among other things (i.e., heterotachy, misspecified models of molecular evolution). Despite advances in the field with the implementation of new analyses, collection of more molecular data, and discovery of new fossils (i.e. Li et al. 2008), turtle ancestry remains a mystery.

In this study I use newly developed markers (Fong & Fujita submitted, Chapter 1) to construct a robust vertebrate phylogeny to tackle the placement of turtles within Amniota. These markers were developed using gene transcripts (expressed sequence tags [ESTs] and cDNA), so potential priming sites are conserved across vertebrates. Included in these phylogenetic analyses is work evaluating the impact of missing data on phylogenetic inference, different data partitioning strategies, identifying unstable (“rogue”) taxa, and utilizing different data-types for phylogenetic reconstruction.

METHODS

Taxon and Marker Sampling

Taxon sampling was selected to span vertebrates (Tetrapoda + fish outgroups; Table 2). Of the 128 taxa, genetic data were available for 46 taxa for which complete genomes are available online (Genbank and Ensembl, Hubbard et al. 2007), and the remaining 82 samples were newly sequenced for the markers developed here. Available data were skewed heavily towards mammals (36), but also include fish (5), frog (1), lizard (1), bird (2), and crocodile (1). New taxon sampling expanded representation within amphibians, reptiles, and birds (Table 2), with an emphasis on turtles to address additional questions regarding turtle systematics (Chapter 3). The tuatara (*Sphenodon punctatus*) was not included in this study because of the difficulties obtaining a sample of this rare and endangered animal. However, the omission of this taxon should have little effect on the results, as the sister taxon relationship between *Sphenodon* and Squamata (lizards, snakes, amphisbaenians) forming the group Lepidosauria is well established (Hedges & Poling 1999). A detailed list of all taxa used in this study, along with sequence identification numbers (Genbank or Ensembl), can be found in Appendix A.

There has been much discussion in the literature on choosing the appropriate number of taxa and loci for phylogenetic inference (Hillis 1996, Kim 1996, Graybeal 1998, Hillis 1998, Bremer et al. 1999, Mitchell et al. 2000, Rosenberg & Kumar 2001, Zwickl & Hillis 2002, Hillis et al. 2003, Sorenson et al. 2003). Diverse taxon sampling is important for bisecting long branches, which may be problematic in phylogenetic analyses (Hendy & Penny 1989). With the methodology employed here, I simultaneously increased both the number of taxa and loci. The final dataset includes 128 taxa and 75 protein-coding, nuclear loci.

In molecular phylogenetic studies, it is important to include multiple genes to avoid stochastic errors due to different evolutionary histories of genes (Philippe & Telford 2006). Selecting the appropriate number of loci and taxa has been a longstanding debate (Graybeal 1998, Zwickl & Hillis 2002, Hillis et al. 2003, Rokas et al. 2003, Hedtke et al. 2006, Edwards et al. 2007). The appropriate number of loci seems to be dependent on whether there is congruence between gene trees and the species tree; using 2-3 loci can resolve a phylogeny when there is gene tree/species tree congruence while as many as 120 genes are needed when there is conflict between gene trees/species tree (Edwards et al. 2007). In the case of EST loci, 1000 random EST loci are suggested (Philippe & Telford 2006). Philippe et al. (2005) recommended >10,000 positions in order to minimize the effect of missing data. The dataset compiled in this study comprises an alignment of 34,110bp.

Datasets and Data-Types

The markers used in this study were specifically developed for this project. Information on marker development and information on the markers themselves can be found in Fong & Fujita (submitted) and Chapter 1. Individual datasets for each of the 75 genes consisted of orthologous sequences of single-copy genes from online genomes (46 total) and successful sequences from the 82 new samples (Fong & Fujita submitted, Chapter 1). In addition to individual gene datasets, concatenated datasets of the 75 genes were also analyzed. Concatenation of genes was performed using a perl script (available upon request).

In Fong & Fujita (submitted) and Chapter 1, I analyzed phylogenetic informativeness (PI) of four different data-types: NUCL (nucleotide), N12 (1st and 2nd codon positions), DEGEN1 (use of nucleotide ambiguity codes to account for codon degeneracy), and AA (amino acids). The use of each of these data-types has been argued for in higher-level phylogenetics: N12 datasets exclude the 3rd codon positions, which become saturated and exhibit homoplasy, especially at deeper time scales (Edwards et al. 1991), DEGEN1 datasets may be advantageous by reducing problems with base compositional heterogeneity among taxa (Regier & Schultz 2001), and AA datasets evolve more slowly and have increased character-state space, reducing convergence (Albert et al. 1994, Fröhlich & Parker 2000, Simmons et al. 2004). I found that the NUCL dataset contained the most PI and should be used the majority of the time, but is potentially hindered by homoplasy at deeper phylogenetic scales. AA, N12, and DEGEN1 data-types had very little PI, but may be more useful at deeper time scales. Here, I tested these claims on this new, larger dataset. Data were converted to the three additional datasets with the following methodology: AA was translated in Gencious, 3rd codon positions were removed for N12 using MacClade v4.08 (Maddison & Maddison 2008), and DEGEN1 was converted using a Perl script (Regier et al. 2010; available at <http://www.phylotools.com>).

Missing Data

In addition to exploring the utility of different data-types, a subset of taxa was selected to test the impact of missing data on phylogenetic inference. A 15-taxon dataset was compiled by choosing the 2-4 taxa with the most data in each major taxonomic group (Table 3). This dataset is 81.9% complete. For the vertebrate group Crocodylia, there were two individuals of the species *Alligator mississippiensis* (from the EST database and a new sample). To minimize the amount of missing data, I combined the data from both individuals. Combination of the two individual was justified, as they are the same species and are functionally a single taxon when looking at higher-level vertebrate relationships. When present for both individuals, data were identical, further justifying the combination. For each of the two concatenated datasets (all taxa, 15-taxon), the four different data-types were analyzed: NUCL, N12, DEGEN1, and AA.

Phylogenetic Analyses

Each dataset was subjected to maximum likelihood and Bayesian phylogenetic analyses using RAxML (Stamatakis 2006) and MrBayes (Ronquist & Huelsenbeck 2003), respectfully. The assumptions and parameters employed in these analyses are described below.

MAXIMUM LIKELIHOOD ANALYSES

RAxML analyses were run on individual genes and on the eight different concatenated datasets. Nucleotide analyses were run under the GTRGAMMA model of evolution for tree inference with 1000 bootstrap replicates under the GTRCAT model of evolution, while amino acid analyses were run under the Protein Gamma model of evolution using empirical frequencies also with 1000 bootstrap replicates. All concatenated datasets were partitioned according to the optimal partitioning strategy (see below). RAxML v7.2.5

and v7.2.6 (Stamatakis 2006, Stamatakis et al. 2008) analyses were run locally and on the CIPRES portal v2.2 (<http://www.phylo.org/portal2/>; Miller et al. 2009).

BAYESIAN ANALYSES

Bayesian analyses were run on a subset of the datasets, as the computational burden for some of the datasets would require extremely long analysis times to achieve stationarity. For example, some data sets would have required as many as 2000 hours (~3 months) to complete. I decided to forgo these long Bayesian analyses, as results from RAxML and MrBayes analyses were almost identical when run. MrBayes v3.2.1 and v3.2 (Ronquist & Huelsenbeck 2003) analyses were run locally and using the BioHPC@CBSU resource at Cornell University (<http://cbsuapps.tc.cornell.edu>). All analyses were run with four chains for 10 million generations. Burn-in of MCMC chains was evaluated using the online program AWTY by evaluating cumulative plots of posterior probabilities of the 20 most variable splits (Wilgenbusch et al. 2004).

DATA PARTITIONING STRATEGIES

Since this study deals with a complex, multi-gene dataset, it is important to explore heterogeneous processes of molecular evolution through partitioning the data. Recent studies have shown the importance of partitioning data (Brandley et al. 2005, Brown & Lemmon 2007, McGuire et al. 2007). Data partitioning strategies often involve partitioning by gene and/or codon. Tests of alternative partitioning strategies were performed on the NUCL dataset only, as the N12 dataset is a subset of the NUCL dataset, and the DEGEN1 and AA datasets have information on codon position subsumed into gene partitions. For RAxML analyses of the All-NUCL dataset, three different partitioning strategies were tested: by gene (75 partitions), by gene and 1st+2nd and 3rd codon position (150 partitions), and by gene and codon position (225 partitions). Results were compared using likelihood ratio tests to choose the best partitioning strategy.

ROGUE TAXA

Rogue taxa (Wilkinson 1994, 1996) are taxa whose placement in a tree is unstable (being placed in multiple positions), which when using consensus/bootstrap methods creates polytomies and reduces support values. The instability of rogue taxa is often due to missing data or heterogeneous rates of evolution between taxa (Sanderson & Shaffer 2002). The identification of rogue taxa a priori is difficult, but several a posteriori analyses have been developed. For example, a “leaf-dropping method” was developed by Pattengale et al. (2010) and is implemented in RAxML. In general terms, this methodology gives a score to taxa based on the impact of their removal. If the score of the tree is higher excluding a set of taxa, they are identified as rogue taxa and removed. A detailed description of the algorithm follows below.

I deviate from the terminology used in Pattengale et al. (2010) to use more common phylogenetics terminology; a “leaf” refers to a taxon and an “edge” refers to a branch. Bipartitions (the resulting two subsets of taxa remaining when a branch is removed) and relative information content (RIC) are in the basis of this methodology. There are two types of bipartitions, trivial and nontrivial. Trivial bipartitions are ones which separate a single

taxon from the tree (i.e. a branch leading to a tip of the tree), whereas a nontrivial bipartition refers to the removal of an internal branch to create two subsets of taxa with membership >1 . The ratio of the total number of bipartitions (trivial + nontrivial) to the number of bipartitions in a fully resolved tree (no polytomies) is called the RIC. Since the numerator of the RIC includes trivial and nontrivial bipartitions, this methodology penalizes the removal of taxa (trivial bipartitions) and rewards for greater resolution of the tree (more nontrivial bipartitions) (Pattengale et al. 2010). This algorithm can be run on majority-rule and strict consensus trees.

Pattengale et al. (2010) uses a greedy heuristic algorithm that discards a taxon only if its removal provides an improvement in resolution that outweighs the loss of removing more taxa. Various dropsets (sets of removed taxa) are given a score; this score is the difference between the number of branches created and the number of taxa removed. The dropset with the highest score is used, and the corresponding taxa are removed. When there are multiple dropsets with the same highest score, one can opt to select the first dropset (which allows for repeatability) or use a random selection. The analysis iterates until no further benefit is gained from removing taxa.

Determining which rogue taxa to remove was a multi-step process that was run separately for each dataset (representing the four data-types). First, a preliminary RAxML analysis was run with all taxa (128). Based on this phylogeny, taxa that were obviously in incorrect phylogenetic positions (i.e. turtle placed in the mammal clade) were removed. This was repeated until all remaining taxa were placed in the correct clades. Next, the bootstrap results from the RAxML analysis were used in rogue taxa analyses (Pattengale et al. 2010). The rogue taxa analysis was run 10 times using the “random” option to select a dropset for both the strict and majority rules consensus trees. To maximize the number of taxa retained and identify the most unstable taxa, taxa were considered rogue if they were identified in ≥ 5 of the 10 analyses under either strict or majority rules consensus. These taxa were removed and another RAxML analysis was run. Rounds of rogue taxon/RAxML analyses were repeated until there was no improvement in bootstrap values.

Post-hoc Analyses

To better understand and interpret the results from the phylogenetic analyses, two sets of additional tests were performed, approximately unbiased topology tests and statistical analyses of the data using R.

APPROXIMATELY UNBIASED TOPOLOGY TESTS

Approximately unbiased topology tests (AU tests; Shimodaira 2002) were used to test whether sub-optimal trees were significantly worse than the Maximum Likelihood tree. AU tests are similar to the more commonly used Shimodaira-Hasegawa (SH) tests in that they both adjust the selection bias overlooked in the standard use of the bootstrap probability and Kishino-Hasegawa tests. However, AU tests are better than SH tests, as it has been shown that SH tests are more biased (Shimodaira 2002). AU tests were performed to compare the five alternative hypotheses of turtle placement for each of the four datasets. Constrained RAxML analyses were run for each of the five different topologies using the GTRGAMMA model of sequence evolution, and per-site log likelihoods calculated. These

per-site log likelihoods were then input into the program CONSEL (Shimodaira & Hasegawa 2001).

STATISTICAL ANALYSES USING R

The R computer language and environment (<http://www.r-project.org/>) was used to further explore the NUCL dataset. Per-site log likelihoods (LLS) from constrained analyses with the turtle-crocodilian and turtle-archosaur relationships were compared using R, under the GTRGAMMA substitution model. Parameters, such as average GC content (%GC) and average missing data (%missing) have been hypothesized to bias phylogenetic inference (Eyre-Walker 1998, Wiens 2003). “%GC” and “% missing” were calculated per-site and per-gene for different subsets of taxa (crocodiles, archosaurs, turtles) to determine if there were any trends and biases in the data in relation to the LLS.

RESULTS

Datasets and Data-Types

Aligned datasets for the four data-types (NUCL, N12, DEGEN1, AA) resulted in a total of 34,110bp, 22,739bp, 29,739bp, and 11,360aa positions respectively. The entire dataset consists of 3,928 sequences and is 40.9% complete. Of these sequences, 878 are newly sequenced, while the remaining 3,050 are from Genbank and Ensembl. The 15 taxa dataset contained 921 sequences, representing an 81.9% complete dataset.

Data Partitioning

The results from the likelihood ratio tests (LRT) of the alternative partitioning strategies of the NUCL dataset are shown in Table 4. As expected, a greater number of data partitions resulted in higher likelihood scores, despite topologies being identical. However, when comparing the scores of χ^2 tests, partitioning by gene/codon (225 partitions) was not significantly different from partitioning by gene (75 partitions) or by gene and n1+n2 and n3 (n12n3; 175 partitions), but the n12n3 partitioning strategy was significantly better than partitioning by gene. Therefore, the best partitioning strategy identified by the LRT was n12n3, and this partitioning strategy was used for subsequent analyses with the NUCL datasets. For the N12 dataset, since this dataset excludes the 3rd codon position, partitioning by gene matched the optimal partitioning strategy for the NUCL dataset.

Rogue Taxa

The 128 taxa included in this study varied in their completeness of data (average = 31 genes, range = 1-75 genes). As the amount of missing data is relatively high in this dataset, and missing data have been shown to compromise phylogenetic results (Wilkinson 1994, 1996), it was important to identify and remove rogue taxa. Rogue taxa analyses were run separately for each of the four data-types and considered complete when no additional rogue taxa were identified or there were no improvements in bootstrap values. Taxa were removed from datasets either because they were clearly placed in the wrong phylogenetic position or the rogue taxa algorithm identified them. Table 5 is a list of the rogue taxa that were identified and removed from each dataset. For the NUCL dataset, there was one round of taxa removal due to wrong phylogenetic position and two rounds of the rogue taxa

algorithm. The final result was 18 taxa removed and 110 remaining taxa. For the N12 dataset, there was one round of taxa removal due to wrong phylogenetic position, and three rounds of the rogue taxa algorithm. The final result was 20 taxa removed and 108 remaining taxa. For the DEGEN1 dataset, there was one round of taxa removal due to wrong phylogenetic position, and three rounds of the rogue taxa algorithm. The final result was 30 taxa removed and 88 remaining taxa. For the AA dataset, there were two rounds of taxa removal due to wrong phylogenetic position, and three rounds of the rogue taxa algorithm. The final result was 29 taxa removed and 89 remaining taxa. None of the sets of rogue taxa for the four datasets were identical, but there is significant overlap between the datasets (Table 5).

To illustrate the improvement in bootstrap values after rogue taxa removal, support values of the reptile clade were plotted for each successive round of analysis for each dataset (Figure 2). Over the course of removing taxa, each dataset showed significant gains in bootstrap values. For each of the datasets (NUCL, N12, DEGEN1, AA), bootstrap values increased by 38, 55, 98, and 89 points, respectively, from the first to final round of analysis.

Phylogenetic Analyses

To study the phylogenetic position of turtles, Bayesian and/or RAxML analyses were performed on eight different datasets (All-NUCL, 15tax-NUCL, All-N12, 15tax-N12, All-DEGEN1, 15tax-DEGEN1, All-AA, 15tax-AA). For datasets where available, results from both RAxML and Bayesian analyses were similar, so only RAxML trees are shown (Figure 3). Phylogenies have been simplified to show only the relationships between the higher-level vertebrate groups.

When looking at the relationships between higher-level vertebrate taxa, the relationships within amphibians and placement of turtles vary between datasets (Figure 3). Amphibian relationships are discussed in Chapter 5. For the placement of turtles, three different results were recovered. The NUCL dataset inferred turtles to be the sister group of Crocodylia, the N12 dataset places turtles as the sister group to Archosauria (birds and crocodilians), and the DEGEN1 and AA datasets infer turtles to be the sister group to Squamata (lizards, snakes, amphisbaenians). In addition, individual gene analyses recover the four main hypotheses (Figure 4). When sorted by the rate of gene evolution, there does not seem to be any pattern in the relationship recovered (Figure 4C).

Approximately Unbiased Tests and Statistical Tests

To test the statistical significance of the results, I performed Approximately Unbiased (AU) tests using CONSEL. The hypothesis placing turtles as basal amniotes was excluded for all data-types, whereas turtles as basal reptiles was excluded for the NUCL and N12 dataset (Table 6). In addition, the NUCL dataset excluded the turtle-lizard relationship from statistically likely trees (Table 6). Because previous work (Fong & Fujita submitted, Chapter 1) demonstrated that NUCL datasets provide the most phylogenetic informativeness, I place more reliability of the results of the NUCL dataset; turtles are either the sister group to archosaurs or crocodilians.

Using R, I compared site-wise and gene-wise log likelihood scores (LLS) of two constrained topologies (turtle-archosaur, turtle-crocodilian) to two parameters, %GC and

%missing, to determine whether there were any biases in the data. When comparing %GC between turtles and the two potential sister groups (archosaur, crocodilian), there was no difference (Figures 5-6). For %missing, turtles have more missing data compared to archosaurs (Figure 7), but similar amounts to crocodilians (Figure 8). Because of this difference in missing data between archosaurs and crocodilians, I tested whether missing data has an effect on per-site LLS; the plot of missing data and LLS shows there is no effect (Figure 9). Lastly I explored LLS between the two alternative trees, turtle-archosaur and turtle-crocodilian (Figure 10). The Lowess (locally weighted scatterplot smoothing) trend line (not shown) is very close to a value of zero. Based on this graph, it can be seen that of the 31,440 sites, there are a few sites with high LLS that seem to be driving the results of the analyses. I chose two arbitrary cut-off points, to investigate high scoring sites (>0.1 , <-0.1 , referred to as “0.1”) and the highest scoring sites (>0.25 , <-0.25 , referred to as “0.25”) in more detail. Sites were classified by their codon position and cumulative LLS scores for each codon position were calculated (Tables 7, 8). In terms of the number of sites supporting each relationship, they were relatively similar in number; for the 0.25 sites, 160 supported the turtle-crocodilian relationship, whereas 123 supported the turtle-archosaur relationship (Table 8A). For the 0.1 sites, 443 supported the turtle-crocodilian relationship, whereas 437 supported the turtle-archosaur relationship (Table 8A). Sorted by codon position, we see that third codon positions drive the results (Tables 7B, 8B), supporting the turtle-crocodilian relationship. And as more sites are included, the support for the turtle-crocodilian relationship decreases (0.25 LLS=-40.48, 0.1 LLS=-30.43, all data=-27.23 [not in table]). When third codon positions were removed (N12 dataset), the LLS became positive, and support switched in favor of the turtle-archosaur relationship.

DISCUSSION

Previous studies of amniote phylogenetics have placed turtles in five different positions in the phylogeny (Table 1, Figure 1). Based on my results here, it is easy to see why so many different hypotheses have been previously supported, as my results differ depending on the gene and data-type analyzed. For individual gene analyses, each of the major hypotheses was recovered (Figure 4), and these results do not seem to be correlated with the rate of gene evolution (Figure 4C). For concatenated datasets, there is no support for two traditional hypotheses for the placement of turtles (as basal amniotes, or as basal reptiles) so these two possibilities can be excluded. However, the different data-types recover different placements for turtles: sister to crocodilians (NUCL), sister to archosaurs (N12), and sister to squamates (DEGEN1, AA) (Figure 3). The phylogenetic placement of turtles, therefore, becomes a question of data interpretation and analysis of which results are most reliable.

In order to interpret these results, I evaluated my data (1) for phylogenetic informativeness (PI), (2) using Approximately Unbiased (AU) topology tests, and (3) using methods implemented in the statistical program R to assess taxon-dependent biases in the data. First, in Fong & Fujita (submitted) and Chapter 1 of this dissertation, I compared the PI of the four different data-types (NUCL, N12, DEGEN1, AA), and found that NUCL datasets should be used in most cases because of the high levels of PI, whereas AA datasets were rarely useful at this phylogenetic scale. However, when dealing with deeper divergences

or difficult phylogenetic problems with short branches, NUCL datasets have a high potential for homoplasy. In these situations, N12 and DEGEN1 datasets could be more useful because the levels of homoplasy are much lower. The placement of turtles in the amniote phylogeny deals with a deep divergence, making the N12 and DEGEN1 datasets potentially more useful than the NUCL dataset. These results place more reliability on the NUCL dataset, placing turtles as the sister group to archosaurs.

Approximately Unbiased topology tests are used to obtain the confidence set of trees. For tests of all data-types, the phylogenetic hypothesis of turtles as basal amniotes was always rejected, and the NUCL and N12 datasets also excluded the hypothesis of turtles as basal reptiles (Table 6). Additionally, the NUCL dataset excluded the turtle-squamate hypothesis. Again, since the NUCL dataset has been shown to be more reliable, results show that turtles have an affinity with archosaurs, either in the form of a turtle-archosaur or turtle-crocodilian sister group relationship.

Lastly, the statistical program R was used to analyze the genetic data further to determine if there are any taxon-specific biases in the data. For example, base composition (%GC) and missing data (%missing) have been hypothesized to bias phylogenetic inference (Eyre-Walker 1998, Wiens 2003). Comparisons of %GC, %missing and log Likelihood scores (LLS) were performed for individual sites and genes. As seen from scatterplots, both %GC and %missing do not seem to bias the results of the phylogenetic analyses (Figures 5-9). When comparing the per-site LLS supporting the alternative hypotheses, I find that for the 34,110bp dataset, the results are strongly driven by a much smaller subset of the data (0.25LLS=283bp, 0.1LLS=890bp, Tables 7-8). When the data are separated according to codon position, it can be seen that 3rd codon positions strongly drive the analysis to favor the turtle-crocodilian sister group relationship. However, when third codon positions are removed, these data support the turtle-archosaur relationship. These results corroborate the results from the different data-types, with the NUCL dataset supporting the turtle-crocodilian relationship and the N12 dataset (3rd codon positions removed) supporting the turtle-archosaur relationship. In fact, as more data are included in the analysis, the support for the turtle-crocodilian relationship weakens from -40.48 (0.25 total LLS) to -30.44 (0.1 total LLS) to -27.23 (all data, not in table). In fact, the difference between the turtle-archosaur and turtle-crocodilian hypotheses is being driven by ~0.4% of the data (~139bp, Table 7A). Thus, the analyses of data biases implemented in R suggest that the placement of turtles as the sister taxon of archosaurs is more reliable than the competing hypothesis of a turtle-crocodilian sister taxon arrangement.

Taking a paleontological perspective, the question of the placement of turtles is no more resolved, with various studies supporting turtles as the sister group to reptiles, lepidosaurs, and archosaurs (Table 1). This confusion is not surprising, as turtles have a very derived morphology. Most morphological analyses support the position of turtles as the sister taxon of reptiles (i.e. Lyson et al. 2010) and some paleontologists are “skeptical that turtles are derived diapsids” (Brochu 2001). A major criticism of the turtle-archosaur and turtle-crocodile hypotheses is that there is very little morphological data supporting these phylogenies. However, turtles are very derived when they first appear in the fossil record, and the lack of transitional fossils and/or inappropriate use of ingroup versus outgroup taxa may be biasing the morphological analyses (Brochu 2001). Despite this, a recent study

discovered that the laterosphenoid ossification (also called pleurosphenoid), a character apparently misclassified in a previous study (Gaffney 1990), is a potential synapomorphy for stem turtles and Archosauriformes (Bhullar & Bever 2009). Additionally, when comparing the turtle-archosaur and turtle-crocodilian hypotheses, morphological data favor the turtle-archosaur hypothesis, as a monophyletic Archosauria (birds and crocodiles) is generally accepted (Gauthier 1986, Gauthier & Padian 1985, Brochu 2001, Zardoya & Meyer 2001).

CONCLUSION

In the vertebrate phylogeny, one of the outstanding problems is the phylogenetic placement of turtles. The phylogenetic position of turtles has been difficult to resolve, which can be seen by the short internodes within reptiles. As a result, there are published analyses of molecular and/or morphological data that support five different hypotheses of turtle phylogenetic placement (Table 1, Figure 1). Based on my molecular analyses, I recovered three of these hypotheses with a single large molecular dataset depending on how the data were scored (i.e., by data-type). The NUCL dataset supported the turtle-archosaur relationship, the N12 dataset supported the turtle-crocodilian relationship, and the DEGEN1 and AA datasets supported the turtle-lepidosauria relationship. Statistical explorations of the data show that these results were not biased by either GC content or the amount of missing data, and that a small number of sites are driving key results. In deciding between these alternative results, tests of data-type, AU topology tests, and analyses using R support the affinity of turtles with archosaurs, either as turtle-archosaur or turtle-crocodilian relationships, but slightly favoring the turtle-archosaur relationship. Although most morphological data do not favor a turtle-archosaur relationship, there are morphological characters that are consistent with this topology.

ACKNOWLEDGEMENTS

This work would not have been possible without the help of many people. First, I would like to thank the individuals and institutions that provided tissue samples for this project: Drs. Carol Spencer and Jim McGuire of the Museum of Vertebrate Zoology, UC Berkeley, Dr. Matthew K Fujita of the Museum of Comparative Zoology, Harvard University, Dr. Jerome Fuchs of the California Academy of Sciences, Dr. Juan L. Parra of State University of New York, Stony Brook, and Dr. Ricardo Pereira of the Universidade de Porto. Next I would like to thank those who helped with data analyses: Dr. James F Parham of the University of Alabama, Dr. Adam Leache of the University of Washington, Dr. Matthew C Brandley of the University of Sydney, Dr. Matthew K Fujita of the Museum of Comparative Zoology, Harvard University, Drs. Bastien Boussau and Jeremy Brown of UC Berkeley, Dr. Nicholas Pattengale of the University of New Mexico, Dr. Alexandros Stamatakis, Simon Berger, and Andre Aberer of the Heidelberg Institute for Theoretical Studies, and Dr. Jeffrey Townsend and Francesc Lopez of Yale University. In addition, the development of this project would not have been possible without endless conversations with past and present members of the MVZ and UC Berkeley, especially Jim McGuire, Jeffrey Boore, Adam Leache, Matt Fujita, Guin Wogan, Shobi Lawalata, Roberta Damasceno, Jose Wellington Alves dos Santos, Angela Ribeiro, Ana Almeida. Funding for

this project came from the National Science Foundation Doctoral Dissertation Improvement Grant (#08-564) and the Museum of Vertebrate Zoology.

CITATIONS

Albert VA, Backlund A, Bremer K, Chase MW, Manhart JR, Mishler BD, Nixon KC (1994) Functional constraints and rbc L evidence for land plant phylogeny. *Annals of the Missouri Botanical Garden* 81: 534-567.

Benton MJ (1990) Phylogeny of the major tetrapod groups: morphological data and divergence dates. *Journal of Molecular Evolution* 30: 409-424.

Bhullar B-AS, Bever GS (2009) An Archosaur-like laterosphenoid in early turtles (Reptilia: Pantestudines). *Breviora* 518: 1-11.

Boulenger GA (1918) Sur la place des cheloniens dans la classification. *Comptes Rendues A l'Academie Des Sciences, Paris* 169: 605-607.

Bremer B, Jansen RK, Oxelman B, Backlund M, Lantz H, Kim KJ (1999) More characters or more taxa for a robust phylogeny-case study from the coffee family (Rubiaceae). *Systematic Biology* 48: 413-435.

Brandley MC, Schmitz A, Reeder TW (2005) Partitioned Bayesian analyses, partition choice, and the phylogenetic relationships of scincid lizards. *Systematic Biology* 54: 373-390.

Brochu CA (2001) Progress and future directions in archosaur phylogenetics. *Journal of Paleontology* 75: 1185-1201.

Broom R (1924) On the classification of the reptiles. *Bulletin of the American Museum of Natural History* 51: 39-65.

Brown J, Lemmon, AR (2007) The importance of data partitioning and the utility of bayes factors in Bayesian phylogenetics. *Systematic Biology* 56: 643-655.

Cao Y, Adachi J, Hasegawa M (1998) Comment on the quartet puzzling method for finding maximum-likelihood tree topologies. *Molecular Biology and Evolution* 15: 87-89.

Cao Y, Sorenson MD, Kumazawa Y, Mindell DP, Hasegawa M (2000) Phylogenetic position of turtles among amniotes: evidence from mitochondrial and nuclear genes. *Gene* 259: 139-148.

Carroll RL (1988) *Vertebrate paleontology and evolution*. Freeman, New York, New York. 698p.

Caspers GJ, Reinder GJ, Leunissen JAM, Wattel J, de Jong WW (1996) Protein sequences indicate that turtles branched off from the amniote tree after mammals. *Journal of Molecular Evolution* 42: 580-586.

Cracraft J, Donoghue M, Dragoo J, Hillis D, Yates T (eds.) (2002). Assembling the tree of life. National Science Foundation Report. http://www.phylo.org/atol/pdf_docs/atol.pdf

Darwin C (1859) *On the origin of species*. John Murray, London.

deBraga M, Rieppel O (1997) Reptile phylogeny and the interrelationships of turtles. *Zoological Journal of the Linnean Society* 120: 281-354.

Edwards SV, Arctander P, Wilson AC (1991) Mitochondrial resolution of a deep branch in the genealogical tree for perching birds. *Proceedings: Biological Sciences* 243: 99-107.

Edwards SV, Liu L, Pearl DK (2007) High-resolution species trees without concatenation. *Proceedings of the National Academy of Sciences USA* 104: 5936-5941.

Eernisse DJ, Kluge AG (1993) Taxonomic congruence versus total evidence, and amniote phylogeny inferred from fossils, molecules, and morphology. *Molecular Biology and Evolution* 10: 1170-1195.

Eyre-Walker A (1998) Problems with parsimony in sequences of biased base composition. *Journal of Molecular Evolution* 47: 686-690.

Fitch WM, Margoliasch E (1967) Construction of phylogenetic trees. *Science* 155: 279-284.

Frohlich MW, Parker DS (2000) The mostly male theory of flower evolutionary origins: from genes to fossils. *Systematic Botany* 25: 155-170.

Gaffney, ES (1980) Phylogenetic relationships in the major groups of amniotes. In Panchen, AL (ed.). *The terrestrial environment and the origin of land vertebrates*. Academic Press, London, New York. Pp 593-610.

Gaffney ES, Meeker LJ (1983) Skull morphology of the oldest turtles: a preliminary description of *Proganochelys quenstedti*. *Journal of Vertebrate Paleontology* 3: 25-28.

Gauthier JA, Padian K (1985) Phylogenetic, functional, and aerodynamic analyses of the origin of birds and their flight. In Hecht, MK, Ostrom, JH, Viohl, G, Wellnhofer, P (Eds.) *The beginning of birds*. Freunde des Jura-Museum, Eichstatt. Pp 185-197.

Gauthier JA (1986) Saurischian monophyly and the origin of birds. *Memoirs of the California Academy of Sciences* 8: 1-55.

- Gauthier J, Kluge AG, Row T (1988) Amniote phylogeny and the importance of fossils. *Cladistics* 4: 105-209.
- Goodrich ES (1916) On the classification of Reptilia. *Proceedings of the Royal Society of London B* 81: 261-276.
- Graybeal A (1998) Is it better to add taxa or characters to a difficult phylogenetic problem? *Systematic Biology* 47: 9-17.
- Gregory WK (1946) Pareiasaurs versus placodonts as near ancestors to the turtles. *Bulletin of the American Museum of Natural History* 86: 275-326.
- Hedges SB (1994) Molecular evidence for the origin of birds. *Proceedings of the National Academy of Sciences USA* 91: 2621-2624.
- Hedges SB, Poling LL (1999) A molecular phylogeny of reptiles. *Science* 283: 998-1001.
- Hedtke SM, Townsend TM, Hillis DM (2006) Resolution of phylogenetic conflict in large data sets by increased taxon sampling. *Systematic Biology* 55: 522-529.
- Hendy MD, Penny D (1989) A framework for the quantitative study of evolutionary trees. *Systematic Zoology* 38: 297-309.
- Hennig W (ed.) (1983) *Stammesgeschichte der chordaten*. P Parey, Hamburg. Pp 132-139.
- Hill RV (2005) Integration of morphological data sets for phylogenetic analysis of Amniota: the importance of integumentary characters and increased taxonomic sampling. *Systematic Biology* 54: 530-547.
- Hillis DM (1996) Inferring complex phylogenies. *Nature* 383: 130-131.
- Hillis DM (1998) Taxonomic sampling, phylogenetic accuracy, and investigator bias. *Systematic Biology* 47: 3-8.
- Hillis DM, Pollock DD, McGuire JA, Zwickl DJ (2003) Is sparse taxon sampling a problem for phylogenetic inference? *Systematic Biology* 52: 124-126.
- Hugall AF, Foster R, Lee MSY (2007) Calibration choice, rate smoothing, and the pattern of tetrapod diversification according to the long nuclear gene RAG-1. *Systematic Biology* 56: 543-563.
- IUCN (2010) IUCN red list of threatened species 2010.4. [<http://www.iucnredlist.org/>].

- Iwabe N, Hara Y, Kumazawa Y, Shibamoto K, Saito Y, Miyata T, Katoh K (2005) Sister group relationship of turtles to the bird-crocodylian clade revealed by nuclear DNA-coded proteins. *Molecular Biology and Evolution* 22: 810-813.
- Janke A, Erpenbeck D, Nilsson M, Arnason U (2001) The mitochondrial genomes of the iguana (*Iguania iguania*) and the caiman (*Caiman crocodylus*): implications for amniote phylogeny. *Proceedings of the Royal Society of London B* 268: 623-631.
- Kim J (1996) General inconsistency conditions for maximum parsimony: effects of branch lengths and increasing numbers of taxa. *Systematic Biology* 45: 363-374.
- Kirsch JA, Mayer GC (1988) The platypus is not a rodent: DNA hybridization, amniote phylogeny and the palimpsest theory. *Philosophical Transactions of the Royal Society B* 353: 1221-1237.
- Kumazawa Y, Nishida M (1999) Complete mitochondrial DNA sequences of the green turtle and blue-tailed mole skink: statistical evidence for archosaurian affinity of turtles. *Molecular Biology and Evolution* 16: 784-792.
- Laurin M, Reisz R (1995) A reevaluation of early amniote phylogeny. *Zoological Journal of the Linnean Society* 113: 165-223.
- Laurin M, Reisz R (1997) A new perspective on tetrapod phylogeny. In Sumida, SS, Martin, KLM (Eds.). *Amniote Origins completing the transition to land*. Academic Press, Inc. San Diego, CA.
- Lee MSY (1993) The origin of the turtle body plan: bridging a famous morphological gap. *Science* 261: 1716-1720.
- Lee MSY (1995) Historical burden in systematics and the interrelationships of 'Parareptiles'. *Biological Reviews* 70: 459-547.
- Lee MSY (1996) Correlated progression and the origin of turtles. *Nature* 379: 812-815.
- Lee MSY (1997) Reptile relationships turn turtle... *Nature* 389: 245-246.
- Lee MSY (2001) Molecules, morphology, and monophyly of diapsid reptiles. *Contributions to Zoology* 70. <http://dpc.uba.uva.nl/ctz/vol70/nr01/art01>.
- Li C, Wu X, Rieppel O, Wang L, Zhao L (2008) An ancestral turtle from the Late Triassic of southwestern China. *Nature* 456: 497-501.
- Lyson TR, Bever GS, Bhullar B-AS, Joyce WG, Gauthier JA (2010) Transitional fossils and the origin of turtles. *Biology Letters*. doi:10.1098/rsbl.2010.0371.

- Mannen H, Tsoi SC, Krushkal JS, Li WH, Li SS (1997) The cDNA cloning and molecular evolution of reptiles and pigeon lactate dehydrogenase isozymes. *Molecular Biology and Evolution* 14: 1081-1087.
- Mannen H, Li SS (1999) Molecular evidence for a clade of turtles. *Molecular Phylogenetics and Evolution* 13: 144-148.
- Marshall CR (1992) Substitution bias, weighted parsimony, and amniote phylogeny as inferred from 18S rRNA sequences. *Molecular Biology and Evolution* 9: 370-373.
- McGuire JA, Witt CC, Altschuler DL, Remsen JV (2007) Phylogenetic systematics and biogeography of hummingbirds: Bayesian and maximum likelihood analyses of partitioned data and selection of an appropriate partitioning strategy. *Systematic Biology* 56: 837-856.
- Meyer A, Zardoya R (2003) Recent advances in the (molecular) phylogeny of vertebrates. *Annual Review of Ecology Evolution and Systematics* 34: 311-338.
- Miller MA, Holder MT, Vos R, Midford PE, Liebowitz T, Chan L, Hoover P, Warnow T (2009) The CIPRES Portals. CIPRES. 2009-08-04. URL: http://www.phylo.org/sub_sections/portal. Accessed: 2009-08-04. (Archived by WebCite(r) at <http://www.webcitation.org/5imQlJeQa>)
- Mindell DP, Sorenson MD, Dimcheff DE, Hasegawa M, Ast JC, Yuri T (1999) Interordinal relationships of birds and other reptiles based on mitochondrial genomes. *Systematic Biology* 48: 138-152.
- Mitchell A, Mitter C, Regier JC (2000) More taxa or more characters revisited: combining data from nuclear protein-encoding genes for phylogenetic analyses of Noctuoidea (Insecta: Lepidoptera). *Systematic Biology* 49: 202-224.
- Pattengale ND, Swenson KM, Moret BME (2010) Uncovering hidden phylogenetic consensus. *Bioinformatics Research and Applications* 6053: 128-139.
- Philippe H, Lartillot N, Brinkmann H (2005) Multigene analyses of bilaterian animals corroborate the monophyly of Ecdysozoa, Lophotrochozoa, and Protostomia. *Molecular Biology and Evolution* 22: 1246-1253.
- Platz JE, Conlon JM (1997) ...and turn back again. *Nature* 389: 246.
- Philippe H, Telford MJ (2006) Large-scale sequencing and the new animal phylogeny. *TRENDS in Ecology and Evolution* 21: 614-620.

Regier JC, Schultz JW (2001) Elongation factor-2: a useful gene for arthropod phylogenetics. *Molecular Phylogenetics and Evolution* 20: 136-148.

Regier JC, Shultz JW, Zwick A, Hussey A, Ball B, Wetzer R, Martin JW, Cunningham CW (2010) Arthropod relationships revealed by phylogenomic analysis of nuclear protein-coding sequences. *Nature* 463: 1079-1083.

Reisz RR, Head JJ (2008) Turtle origins out to sea. *Nature* 456: 450-451.

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

Reisz RR (1997) The origin and early evolutionary history of amniotes. *Trends in Ecology and Evolution* 12: 218-222.

Rest JS, Ast JC, Austin CC, Waddell PJ, Tibbetts EA, Hay JM, Mindell DP (2003) Molecular systematics of primary reptilian lineages and the tuatara mitochondrial genome. *Molecular Phylogenetics and Evolution* 29: 289-297.

Rieppel O, DeBraga M (1996) Turtles as diapsid reptiles. *Nature* 384: 453-455.

Rieppel O, Reisz R (1999) The origin and early evolution of turtles. *Annual Review of Ecology and Systematics* 30: 1-22.

Rieppel O (2000) Turtles as diapsid reptiles. *Zoologica Scripta* 29: 199-212.

Rokas A, Carroll SB (2005) More genes or more taxa? The relative contribution of gene number and taxon number to phylogenetic accuracy. *Molecular Biology and Evolution* 22: 1137-1344.

Rokas A, Williams BL, King N, Carroll SB (2003) Genome-scale approaches to resolving incongruence in molecular phylogenetics. *Nature* 425: 798-804.

Romer AS (1966) *Vertebrate paleontology*, 3rd ed. University of Chicago Press, Chicago.

Ronquist F, Huelsenbeck JP (2003) MrBayes 3: Bayesian phylogenetic inference under mixed models. *Bioinformatics* 19: 1572-1574.

Rosenberg MS, Kumar S (2001) Incomplete taxon sampling is not a problem for phylogenetic inference. *Proceedings of the National Academy of Sciences USA* 98: 10751-10756.

Sanderson MJ, Shaffer HB (2003) Troubleshooting molecular phylogenetic analyses. *Annual Review of Ecology and Systematics* 33: 49-72.

- Shimodaira H (2002) An approximately unbiased test of phylogenetic tree selection. *Systematic Biology* 51: 492-508.
- Shimodaira H, Hasegawa M (2001) CONSEL: for assessing the confidence of phylogenetic tree selection. *Bioinformatics* 17: 1246-1247.
- Simmons MP, Carr TG, O'Neill K (2004) Relative character-state space, amount of potential phylogenetic information, and heterogeneity of nucleotide and amino acid characters. *Molecular Phylogenetics and Evolution* 32: 913-926.
- Sorenson MD, Oneal E, Garcia-Moreno J, Mindell DP (2003) More taxa, more characters: the Hoatzin problem is still unresolved. *Molecular Biology and Evolution* 20: 1484-1498.
- Stamatakis A (2006) RAxML-VI-HPC: maximum likelihood-based phylogenetic analyses with thousands of taxa and mixed models. *Bioinformatics* 22:2688–2690
- Stamatakis A, Hoover P, Rougemont J (2008) A rapid bootstrap algorithm for the RAxML web-servers. *Systematic Biology* 57:758–771
- Strimmer K, von Haeseler A (1996) Quartet puzzling: a quartet maximum-likelihood method for reconstructing tree topologies. *Molecular Biology and Evolution* 13: 564-569.
- Thomson RC, Shaffer HB (2010) Rapid progress on the vertebrate tree of life. *BMC Biology* 8:19.
- van de Peer Y, Neefs JM, de Rijk P, De Wachter R (1993) Reconstructing evolution from eukaryotic small-ribosomal-subunit RNA sequences: calibration of the molecular clock. *Journal of Molecular Evolution* 37: 221-232.
- Werneburg I, Sanchez-Villagra R (2009) Timing of organogenesis support basal position of turtles in the amniote tree of life. *BMC Evolutionary Biology* 9: 82.
- Wiens JJ (2003) Missing data, incomplete taxa, and phylogenetic accuracy. *Systematic Biology* 52: 528-538.
- Wilkinson M, Thorley J, Benton MJ (1997) Uncertain turtle relationships. *Nature* 387: 466.
- Wilgenbusch JC, Warren DL, Swofford DL (2004) AWTY: A system for graphical exploration of MCMC convergence in Bayesian phylogenetic inference. <http://ceb.csit.fsu.edu/awty>.

Wilkinson M (1994) Common cladistic information and its consensus representation: reduced Adams and reduced cladistic consensus trees and profiles. *Systematic Biology* 43: 343-368.

Wilkinson M (1996) Majority-rule reduced consensus trees and their use in bootstrapping. *Molecular Biology and Evolution* 13: 437-444.

Yates TL, Salazar-Bravo J, Dragoo JW (2004) The importance of the tree of life to society. In: Cracraft J, Donoghue MJ. *Assembling the tree of life*. Oxford University Press, New York.

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

Zardoya R, Meyer A (2001) On the origin and phylogenetic relationships among living amphibians. *Proceedings of the National Academy of Sciences USA* 98: 7380-7383.

Zardoya R, Malaga-Trillo E, Veith M, Meyer A (2003) Complete nucleotide sequence of the mitochondrial genome of a salamander *Mertensiella luschani*. *Gene* 317: 17-27.

Zwickl DJ, Hillis DM (2002) Increased taxon sampling greatly reduces phylogenetic error. *Systematic Biology* 51: 588-598.

Table 1: Summary table of past studies on turtle placement in the amniote phylogeny.

HYPOTHESIS	CITATION	MORPH	DNA
1) Basal Amniote	Gaffney 1980	X	
2) Basal Reptile	Romer 1966	X	
	Gregory 1946	X	
	Gaffney & Meeker 1983	X	
	Gauthier et al 1988	X	
	Benton 1990	X	
	Reisz & Laurin 1991	X	
	Eernisse & Kluge 1993	X	X
	Lee 1993	X	
	van de Peer et al 1993		X
	Laurin & Reisz 1995	X	
	Lee 1995	X	
	Caspers et al 1996		X
	Lee 1996	X	
	Strimmer & von Haeseler 1996		X
	deBraga & Rieppel 1997	X	
	Laurin & Reisz 1997	X	
	Lee 1997	X	
	Wilkinson et al 1997	X	
	Cao et al 1998		X
	Rieppel & Reisz 1999	X	
	Lee 2001	X	X
	Li et al 2008	X	
	Reisz & Head 2008	X	
	Werneburg & Sanchez-Villagra 2009	X	
	Lyson et al 2010	X	
3) Sister group to Lepidosauria (FOR)	Hedges 1994		X
	Rieppel & deBraga 1996	X	
	deBraga & Rieppel 1997	X	
	Rieppel & Reisz 1999	X	
	Hill 2005	X	
(AGAINST)	Lee 1997	X	
	Wilkinson et al 1997	X	
4) Sister group to Archosauria (FOR)	Fitch & Margoliash 1967		X
	Hennig 1983	X	
	Platz & Conlon 1997		X
	Zardoya & Meyer 1998		X
	Kumazawa & Nishida 1999		X
	Mindell et al 1999		X
	Cao et al 2000		X
	Janke et al 2001		X
	Zardoya & Meyer 2001		X
	Rest et al 2003		X
	Zardoya et al 2003		X
	Iwabe et al 2005		X
	Hugall et al 2007		X
	Bhullar & Bever 2009	X	
(AGAINST)	Rieppel 2000	X	
5) Sister group to Crocodylia	Mannen et al 1997		X
	Kirsch & Mayer 1998		X
	Hedges & Poling 1999		X
	Mannen & Li 1999		X
	Cao et al 2000		X

Table 2: Summary of the taxon sampling used in this study.

	EST/Genomes	New Sampling	TOTAL
Fish	5	0	5
Amphibian	1	8	9
Mammal	36	0	36
Tuatara	0	0	0
Lizard	1	15	16
Snake	0	3	3
Crocodile	1	1	2
Bird	2	10	12
Turtle	0	45	45
TOTAL	46	82	128

Table 3: Taxa included in the reduced 15 taxa dataset.

Group	Taxon
Fish (2)	<i>Danio rerio</i> <i>Gasterosteus aculeatus</i>
Amphibian (2)	<i>Ichthyophis bannanicus</i> <i>Xenopus tropicalis</i>
Lizard (2)	<i>Anolis carolinensis</i> <i>Heteronotia binoei</i>
Crocodile (1)	<i>Alligator mississippiensis</i>
Turtle (2)	<i>Malaclemys terrapin</i> <i>Testudo hermanni</i>
Bird (2)	<i>Gallus gallus</i> <i>Taeniopygia guttata</i>
Mammal (4)	<i>Homo sapiens</i> <i>Monodelphis domestica</i> <i>Mus musculus</i> <i>Ornithorhynchus anatinus</i>

Table 4: Likelihood scores and results from likelihood ratio tests testing different data partitioning strategies. The most optimal partitioning strategy for the NUCL dataset was partitioning by gene and by 1st+2nd and 3rd codon positions (n12n3).

	GENE	n12n3
n12n3	LnL=-383863.5 $\chi^2=13699.61$ D.F.= 8700 P=0.0	
GENE/CODON	LnL=-370163.9 $\chi^2=15939.32$ D.F.= 17400 P=1.0	LnL=367924.2 $\chi^2=2239.72$ D.F.= 8700 P=1.0

Table 5: List of rogue taxa removed for each data-type. Highlighted taxa were removed from all four datasets. (NUCL=nucleotide, N12=1st and 2nd codon positions, DEGEN1=nucleotides accounting for codon degeneracy, AA=amino acids).

Taxon Removed	Dataset			
	NUCL	N12	DEGEN1	AA
<i>Accipiter cooperii</i>	X	X	X	X
<i>Agrionemys horsfieldii</i>		X	X	
<i>Aglaiocercus kingi</i>	X			
<i>Anas cyanoptera</i>	X	X	X	X
<i>Aptenodytes patagonicus</i>	X	X	X	X
<i>Aspidoscelis tigris</i>		X	X	X
<i>Bipes biporus</i>			X	X
<i>Casuarinus casuarinus</i>	X	X	X	X
<i>Chamaeleo senegalensis</i>	X	X	X	X
<i>Chelodina longicollis</i>				X
<i>Chelonoidis denticulata</i>			X	
<i>Cuora galbinifrons</i>			X	X
<i>Cuora pani</i>	X		X	X
<i>Dermatemys mawii</i>		X	X	
<i>Dermochelys coriacea</i>	X		X	X
<i>Emydura subglobosa</i>	X	X	X	X
<i>Emys marmorata</i>		X	X	X
<i>Ensatina eschscholtzii</i>	X	X	X	X
<i>Geotrypetes seraphini</i>	X	X	X	X
<i>Heloderma horridum</i>		X	X	X
<i>Hynobius arisanensis</i>	X	X	X	X
<i>Indotestudo elongata</i>	X		X	X
<i>Kinixys belliana</i>	X			
<i>Leucocephalon yuwonoi</i>	X		X	X
<i>Lissemys punctata</i>		X	X	X
<i>Macaca mulatta</i>		X		
<i>Malacochersus tornieri</i>	X		X	X
<i>Mauremys caspica</i>		X	X	X
<i>Mesalina brevirostris</i>			X	X
<i>Microcebus murinus</i>			X	
<i>Nectarinia olivacea</i>		X	X	X
<i>Pelodiscus sinensis</i>			X	
<i>Phrynosps sp.</i>			X	X
<i>Pipa pipa</i>			X	
<i>Platemys platycephala</i>			X	X
<i>Platysternon megacephalum</i>			X	X
<i>Rana livida</i>	X	X	X	X
<i>Rhinemra floridana</i>			X	X
<i>Sacalia quadriocellata</i>			X	X
<i>Terrapene carolina</i>			X	X
<i>Trachemys scripta</i>			X	X
<i>Tupaia belangeri</i>	X			
<i>Varanus yemenensis</i>			X	X
<i>Vicugna pacos</i>			X	
<i>Xantusia vigilis</i>		X	X	X

Table 6: Results from the Approximately Unbiased tests using CONSEL for the five alternative placements of turtles in the amniote phylogeny. Analyses were performed separately for three of the data-types. “X” denotes the best tree, and statistically identical trees to this best tree are highlighted in orange.

	NUCL	N12	DEGEN1	AA
ARCHOSAUR	p=0.155	X	p=0.648	p=0.347
CROCODILE	X	p=0.474	p=0.138	p=0.155
LIZARD	p=1e-4	p=0.113	X	X
REPTILE	p=2e-4	p=0.03	p=0.65	p=0.092
AMNIOTE	p=0.007	p=1e-8	p=0.004	p=0.001

Table 7: Sites with the highest per-site log Likelihood scores (LLS), >0.25 or <-0.25.

A) Number of total sites supporting the alternative hypotheses, sorted by codon position.

B) Calculation of cumulative LLS score per codon position. A positive value indicates support for the turtle-archosaur relationship, while a negative value indicates support for the turtle-crocodilian relationship.

A)

0.25	1st	2nd	3rd	TOTAL
CROC	17	4	139	160
ARCH	20	4	99	123

B)

0.25	1st	2nd	3rd	totalLLS	meanLLS
CROC	-18.61928	-4.33631	-102.23331	-125.1889	-0.7824306
ARCH	18.15983	5.153907	61.394413	84.70815	0.6886841
Δ	-0.45945	0.817597	-40.838897	-40.48075	

Table 8: Sites with high per-site log Likelihood scores (LLS), >0.1 or <-0.1. A) Number of total sites supporting the alternative hypotheses, sorted by codon position. B) Calculation of cumulative LLS score per codon position. A positive value indicates support for the turtle-archosaur relationship, while a negative value indicates support for the turtle-crocodilian relationship.

A)

0.1	1st	2nd	3rd	TOTAL
CROC	58	19	366	443
ARCH	55	22	360	437

B)

0.1	1st	2nd	3rd	totalLLS	meanLLS
CROC	-13.16675	-6.48385	-145.0843	-164.7349	-0.371862
ARCH	23.53337	8.030259	102.732471	134.2961	0.3073138
Δ	10.36662	1.546409	-42.351829	-30.4388	

Figure 1: Consensus backbone phylogeny of Amniota (Zardoya & Meyer 2001) with alternative placement of turtles. A) Turtles as the most basal amniote, B) Turtles as the most basal reptile, C) Turtles as sister group to Lepidosauria, D) Turtles as sister group to Archosauria, E) Turtles as sister group to Crocodylia.

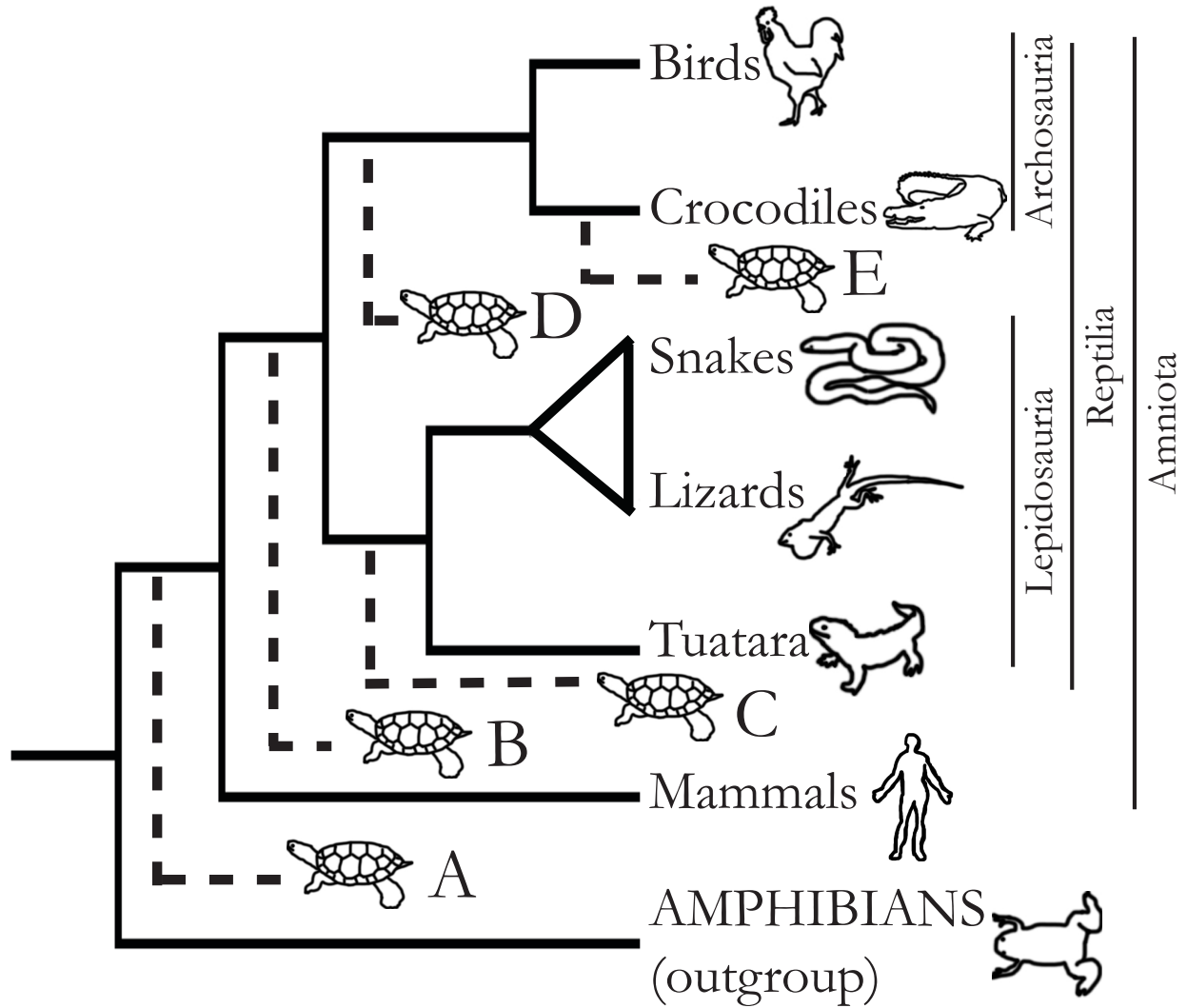


Figure 2: Changes in bootstrap values of the reptile clade after successive rounds of rogue taxa removal. The last point of each graph indicates the last round of rogue taxa removal.

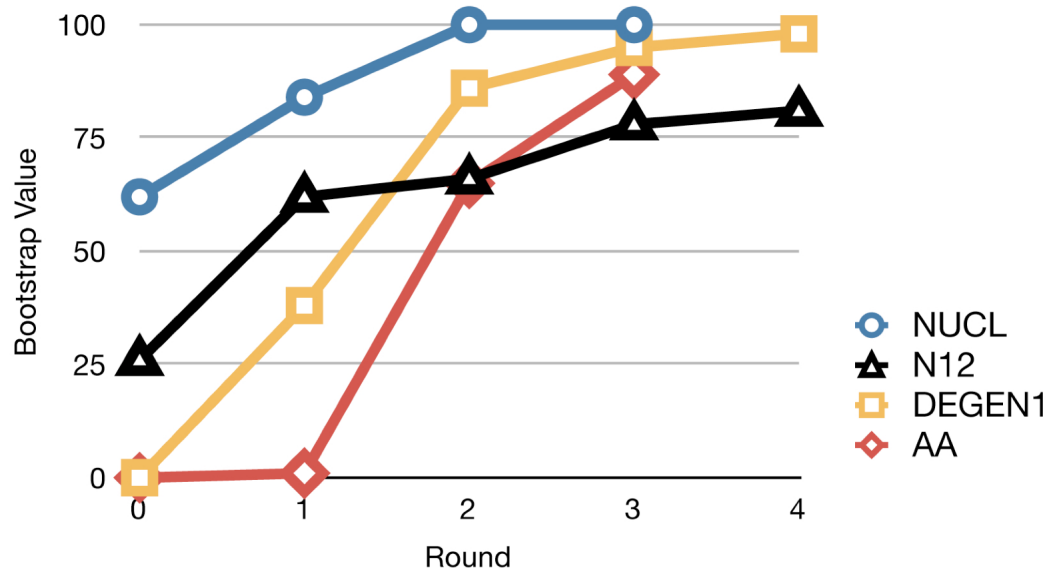


Figure 3: Maximum Likelihood phylogenies of the different datasets. Phylogenies have been simplified to only show higher-level relationships within vertebrates. A) All taxa-NUCL dataset, B) 15 taxa-NUCL dataset, C) All taxa-N12 dataset, D) 15 taxa-N12 dataset (note: the Bayesian analysis recovered a turtle-crocodile relationship), E) All taxa-DEGEN1 dataset, F) 15 taxa-DEGEN1 dataset, G) All taxa-AA dataset, H) 15 taxa-AA dataset. Support values for phylogenies with all taxa (A,C,E,G) show RAxML bootstrap values only if ≥ 50 . Support values for phylogenies with 15 taxa show support values in the form of RAxML bootstrap/Bayesian posterior probability. An * indicates full support.

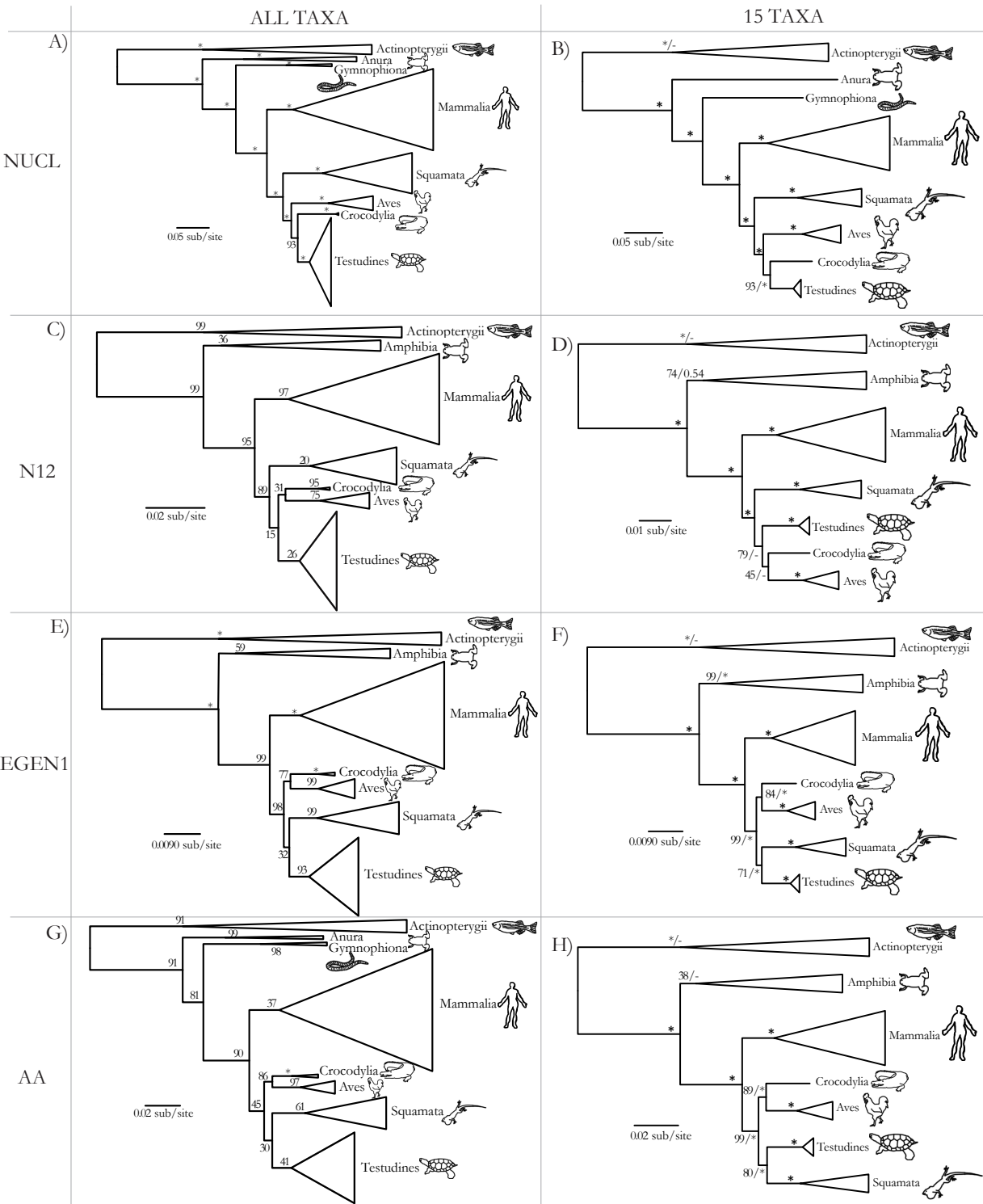


Figure 4: Phylogenetic position of turtles based on the results of individual gene analyses. Genes were divided into two groups A) all major taxonomic groups present, and B) missing a crocodilian representative. Percentages represent the proportion of genes that show the particular relationship. Relationships in the gene-by-gene analyses are often poorly supported due to the reduced amount of data (152-789bp), so topologies irrespective of bootstrap values are noted. C) Topologies organized by the rate of evolution of genes (fastest on left, slowest on right), to qualitatively evaluate whether the placement of turtles is influenced by rate of evolution. Colors correspond to the pie charts. It seems there is no obvious relationship between placement of turtles and rate of evolution.

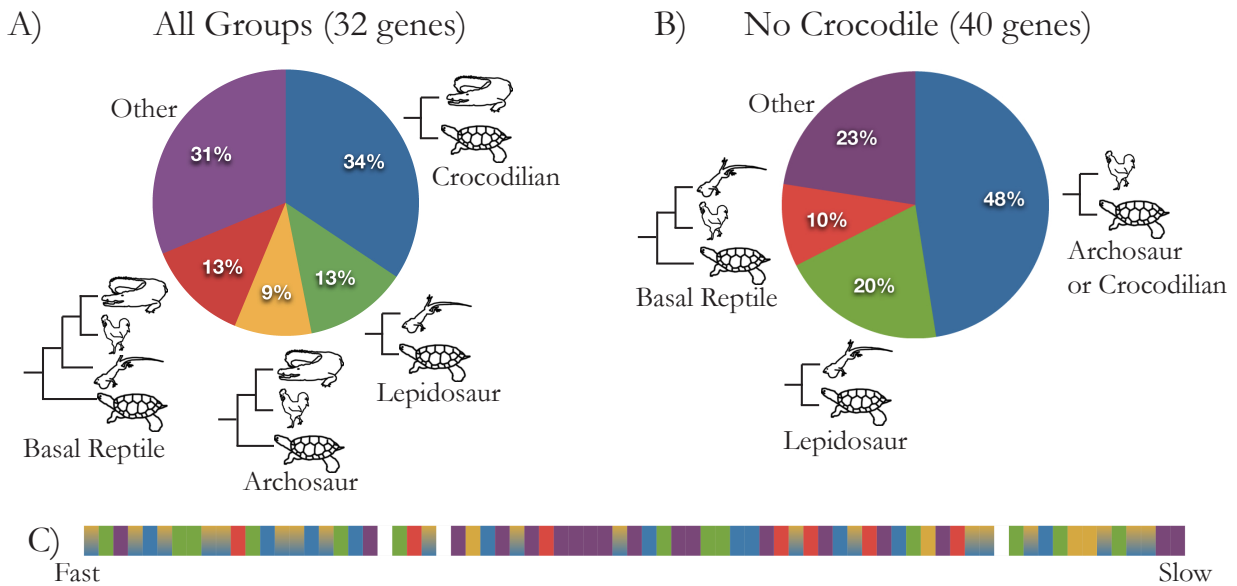


Figure 5: Plot comparing average percentage of GC (%GC) between turtles and archosaurs across all sites. %GC for each site was calculated for each group, and the difference calculated for each site. A value of “0” indicates no difference, while a positive value indicates that turtles have a higher %GC, and a negative value indicates that archosaurs have a higher %missing. The red line represents the Lowess trend line of the data.

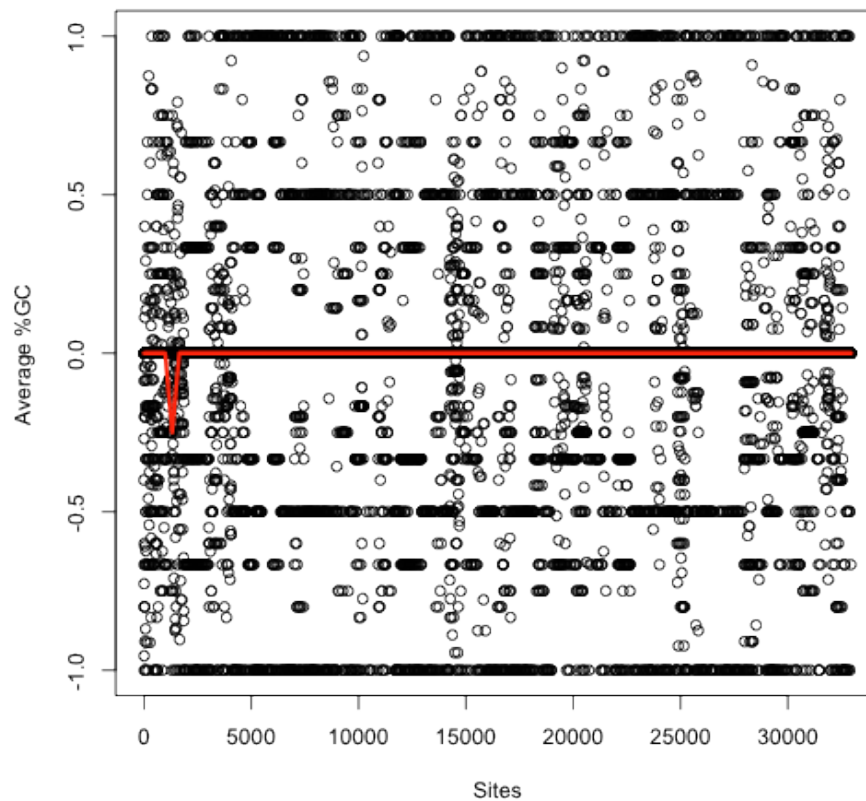


Figure 6: Plot comparing average percentage GC (%GC) between turtles and crocodiles across all sites. %GC for each site was calculated for each group, and the difference calculated for each site. A value of “0” indicates no difference, whereas a positive value indicates that turtles have a higher %GC, and a negative value indicates that crocodiles have a higher %GC. The red line represents the Lowess trend line of the data.

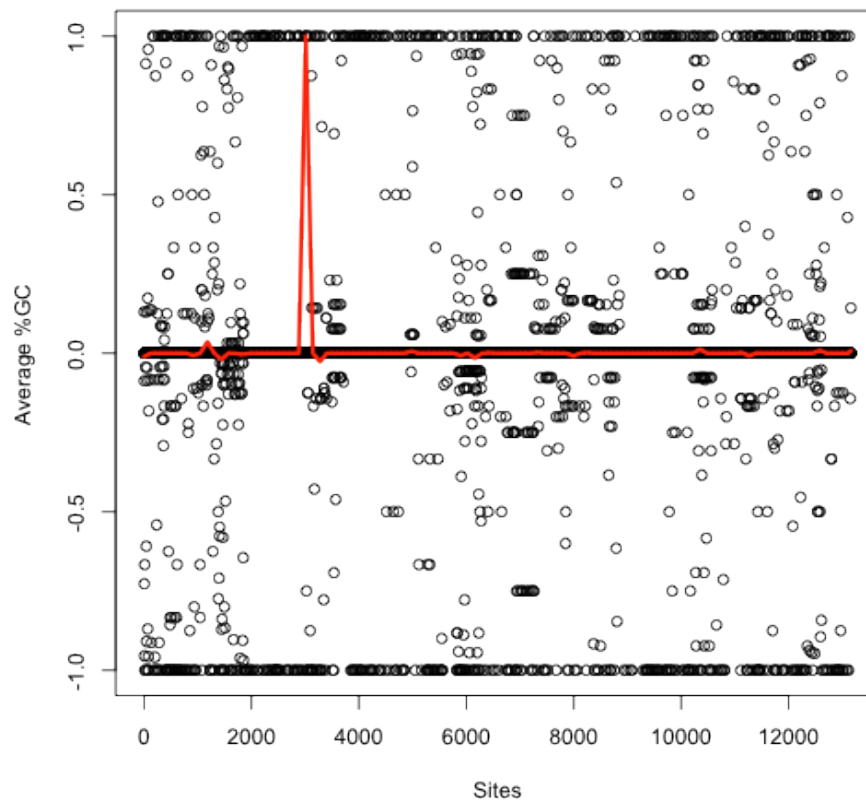


Figure 7: Plot comparing average percentage of missing data (%missing) between turtles and archosaurs across all sites. %missing for each site was calculated for each group, and the difference calculated for each site. A value of “0” indicates no difference, whereas a positive value indicates that turtles have a higher %missing, and a negative value indicates that archosaurs have a higher %missing. The red line represents the Lowess trend line of the data.

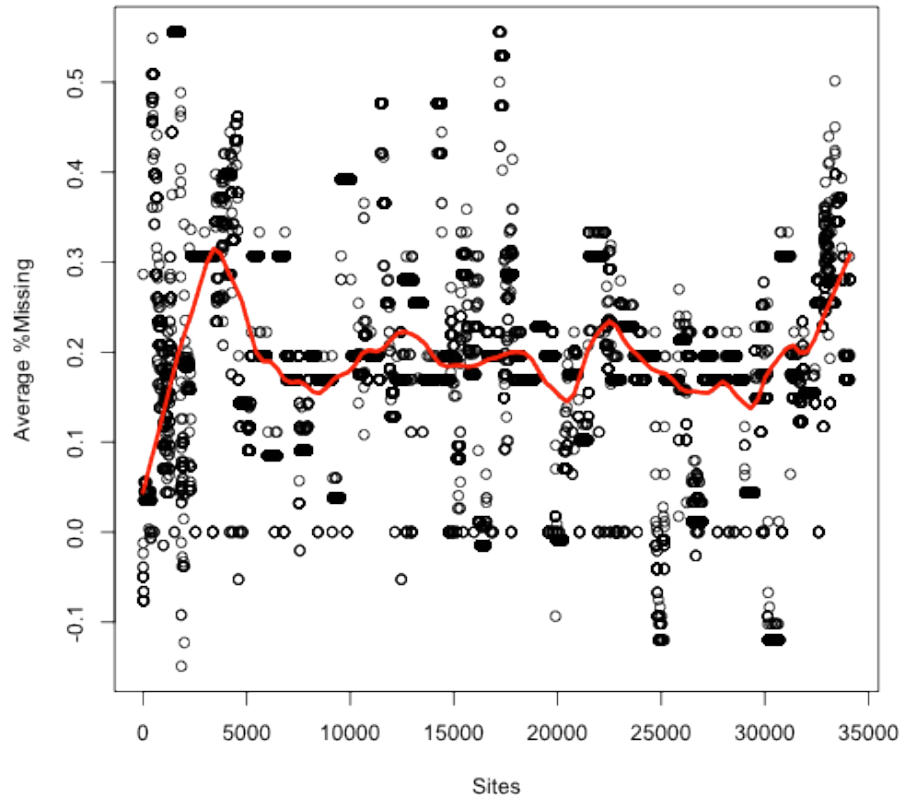


Figure 8: Plot comparing average percentage missing data (%missing) between turtles and crocodilians across all sites. %missing for each site was calculated for each group, and the difference calculated for each site. A value of “0” indicates no difference, whereas a positive value indicates that turtles have a higher %missing, and a negative value indicates that crocodiles have a higher %missing. The red line represents the Lowess trend line of the data.

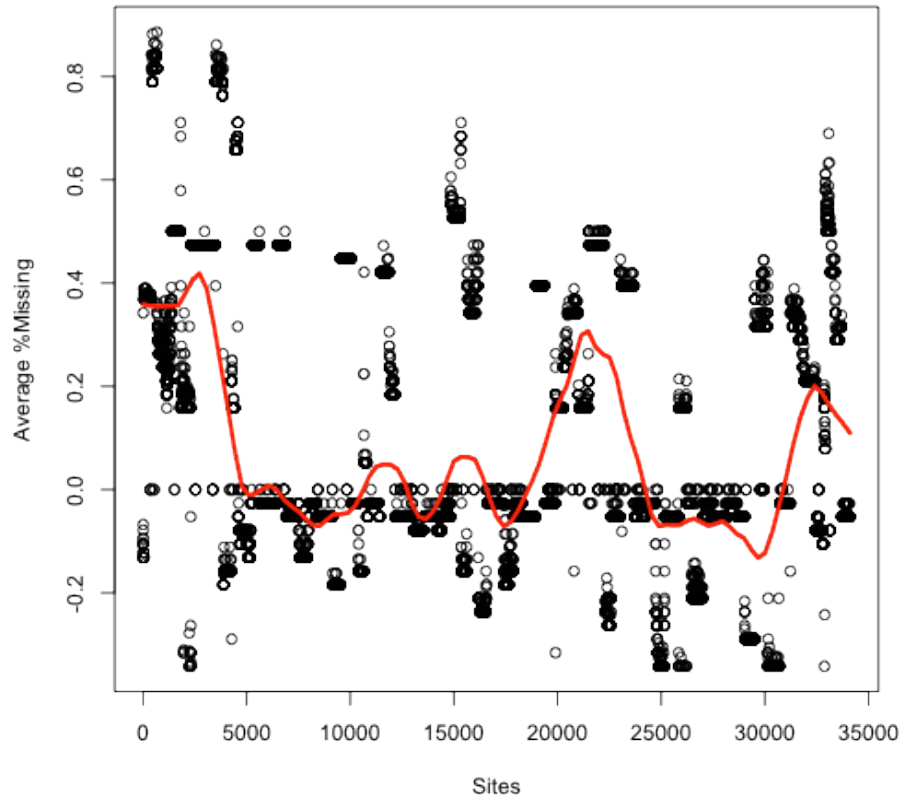


Figure 9: Comparison of per-site log Likelihood scores (LLS) to missing data (%missing). The red line represents the Lowess trend line of the data. This Lowess line shows that there is no effect of missing data on the LLS.

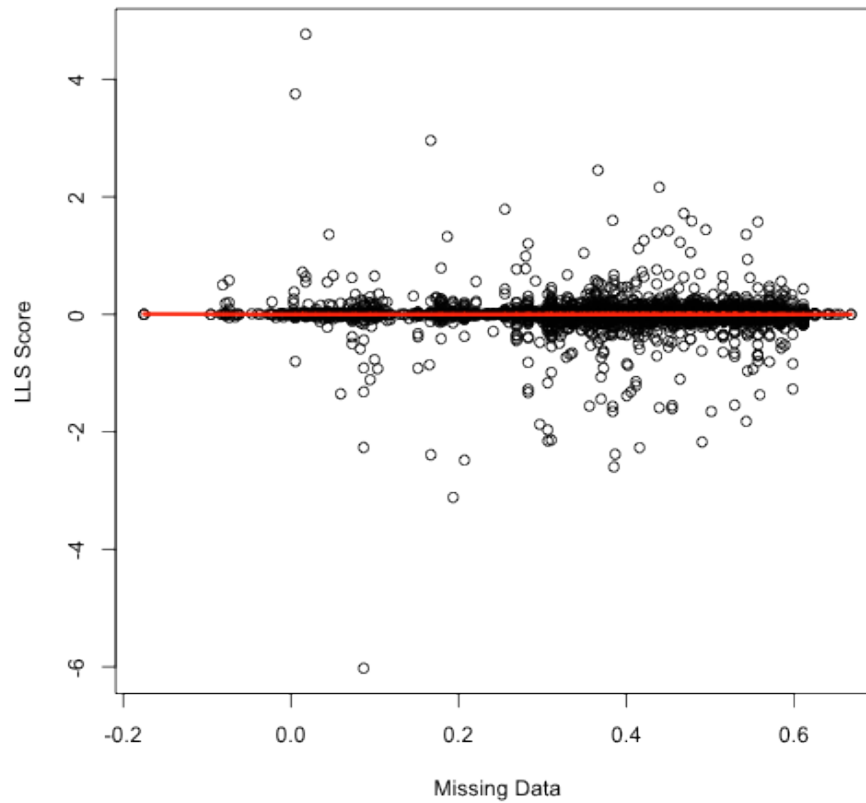
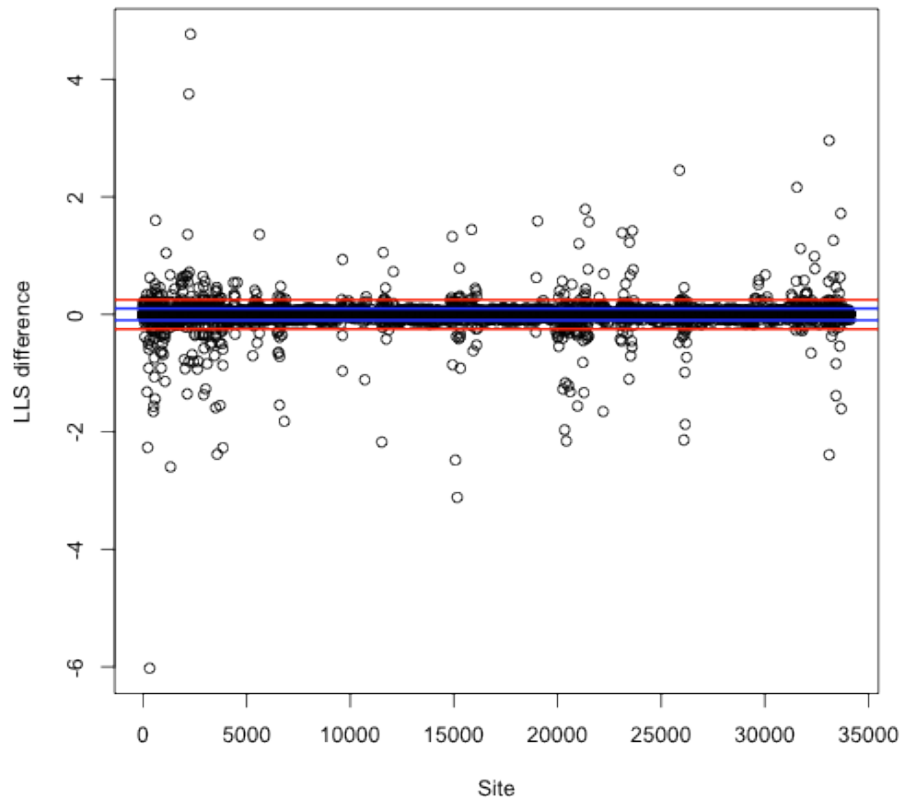


Figure 10: Comparison of the per-site log Likelihood scores (LLS) for the turtle-archosaur and turtle-crocodilian phylogenies. A value of “0” indicates no difference in LLS, whereas a positive value indicates that the turtle-archosaur tree has a higher LLS, and a negative value indicates that the turtle-crocodile has a higher LLS. Red lines delineate values ± 0.25 , and blue lines delineate values ± 0.1 .



Appendix A: Complete list of taxa included in this study, including their museum specimen numbers and identification numbers for sequences. Identification numbers for *Alligator mississippiensis* are from Genbank’s EST database, while the remaining numbers are from Ensembl. Cells with “*” are newly sequenced, but do not have Genbank numbers.

Scientific Name	Specimen #	ABCE1	AMD1	ARCN1	C7orf28B	C22orf28
<i>Accipiter cooperii</i>	MVZ181729					
<i>Agelaiocercus kingi</i>	MVZ183274					
<i>Agrionomys borsfieldii</i>	MVZ245929					
<i>Alethe castanea</i>	JF1706					
<i>Alligator mississippiensis</i>	n/a					
<i>Alligator mississippiensis</i>	MVZ241485	*				
<i>Anas cyanoptera</i>	MVZ179315					
<i>Anniella pulchra</i>	MVZ247492					
<i>Anolis carolinensis</i>	n/a	ENSACAT00000003476	ENSACAT00000011200	ENSACAT00000006956	ENSACAT00000012252	ENSACAT00000014912
<i>Apalone ferrox</i>	MVZ241535		*			
<i>Aptenodytes patagonicus</i>	MVZ178921					
<i>Aspidoscelis tigris</i>	MVZ243345					
<i>Bipes biporus</i>	MVZ236260					
<i>Bos taurus</i>	n/a	ENSBTAT00000048747	ENSBTAT00000004338	ENSBTAT00000008799	ENSBTAT00000004259	ENSBTAT00000014699
<i>Callithrix jacchus</i>	n/a	ENSCJAT00000005088	ENSCJAT00000010193	ENSCJAT00000026136	ENSCJAT00000029767	ENSCJAT00000012014
<i>Canis familiaris</i>	n/a	ENSCAFT00000012170	ENSCAFT00000006286	ENSCAFT00000020001	ENSCAFT00000024533	ENSCAFT00000002802
<i>Casuarus casuarus</i>	MVZ178927					
<i>Cavia prellus</i>	n/a	ENSCPOT00000007439	ENSCPOT00000008859	ENSCPOT00000025891	ENSCPOT00000014932	ENSCPOT00000014606
<i>Chamaeleo senegalensis</i>	MVZ249659					
<i>Chelodina longicollis</i>	MVZ230488					
<i>Chelonia mydas</i>	MVZ233240					
<i>Chelonoidis denticulata</i>	MVZ175391					
<i>Chelydra serpentina</i>	MVZ250691					
<i>Choloepus hoffmanni</i>	n/a	ENSCHOT00000004550	ENSCHOT00000009195	ENSCHOT00000006340		ENSCHOT00000009147
<i>Chrysemys picta</i>	MVZ241506	*			*	
<i>Claudius angustatus</i>	MVZ164771					
<i>Cuora galbinifrons</i>	MVZ230933					
<i>Cuora pani</i>	MVZ230513					
<i>Cyrtodactylus</i>	JAM8710	*	*			
<i>Danio rerio</i>	n/a	ENDSART00000015777	ENDSART000000064391	ENDSART00000047902	ENDSART000000112197	ENDSART000000111348
<i>Dasyatis novemcinctus</i>	n/a	ENDSDNOT00000011817	ENDSDNOT00000006865	ENDSDNOT00000010006	ENDSDNOT00000025597	ENDSDNOT00000012428
<i>Dendropicus gabonensis</i>	JF1513		*			
<i>Dermatemys mawii</i>	MVZ233618					
<i>Dermochelys coriacea</i>	MVZ149844					
<i>Dibamus sp.</i>	MVZ224113		*			
<i>Dierochelys reticularia</i>	MVZ230923	*				
<i>Dipodomys ordii</i>	n/a	ENDSDORT00000016691		ENDSDORT00000005138	ENDSDORT00000013286	ENDSDORT00000010857
<i>Draco beccarii</i>	JAM8961		*			*
<i>Draco sumatranus</i>	SZL71					
<i>Echinops telfairi</i>	n/a	ENSETET00000011259		ENSETET00000014914		ENSETET00000001987
<i>Emydura subglobosa</i>	MVZ230496					
<i>Emys marmorata</i>	MVZ164996					
<i>Ensatina eschscholtzii</i>	RP158					
<i>Equus caballus</i>	n/a	ENSECAT00000025599	ENSECAT00000011822	ENSECAT00000021276	ENSECAT00000012151	ENSECAT00000025737

Scientific Name	Specimen #	ABCE1	AMD1	ARCNI	C7orf28B	C22orf28
<i>Eretmochelys imbricata</i>	MVZ233242					
<i>Erinaceus europaeus</i>	n/a	ENSEEUT00000009701		ENSEEUT00000015606		ENSEEUT00000006261
<i>Fejervarya limncharis</i>	MVZ251769		*			
<i>Felis catus</i>	n/a	ENSFCAT00000019340		ENSFCAT00000011798		ENSFCAT00000009799
<i>Gallus gallus</i>	n/a	ENSGALT00000016194	ENSGALT00000024267	ENSGALT00000012021	ENSGALT00000005479	ENSGALT00000020590
<i>Gasterosteus aculeatus</i>	n/a	ENSGACT00000024658	ENSGACT00000008748	ENSGACT00000006906	ENSGACT00000012513	ENSGACT00000000758
<i>Geochelone sulcata</i>	MVZ238888	*			*	
<i>Geomyda spengleri</i>	MVZ257081					
<i>Geotrypetes seraphini</i>	MVZ252475					
<i>Gorilla gorilla</i>	n/a		ENSGGOT00000006493	ENSGGOT00000006194	ENSGGOT00000002478	ENSGGOT00000011447
<i>Hardella thurjii</i>	MVZ248392					
<i>Heloderma horridum</i>	MVZ238489					
<i>Heteronotia binoei</i>	MKF226	*	*			
<i>Homo sapiens</i>	n/a	ENST00000296577	ENST00000368876	ENST00000359415	ENST00000316731	ENST00000216038
<i>Hynobius arisanensis</i>	MVZ251776					
<i>Icthyophis bannanicus</i>	MVZ226265	*				
<i>Indotestudo elongata</i>	MVZ234627					
<i>Kachuga smithii</i>	MVZ248394					
<i>Kinixys belliana</i>	MVZ234631					
<i>Kinosternon scorpioides</i>	MVZ247584					
<i>Leptotyphlops dulcis</i>	MVZ236621					
<i>Leucocephalon yuwonoi</i>	MVZ241499					
<i>Lissemys punctata</i>	MVZ248398					
<i>Loxodonta africana</i>	n/a	ENSLAFT00000017972	ENSLAFT00000013205			ENSLAFT00000002082
<i>Macaca mulatta</i>	n/a	ENSM MUT00000006286	ENSM MUT00000000779	ENSM MUT00000003620	ENSM MUT00000011151	ENSM MUT00000009413
<i>Macropus eugenii</i>	n/a	ENSMEUT00000006989	ENSMEUT00000010659	ENSMEUT00000001683		ENSMEUT00000007832
<i>Malaclemys terrapin</i>	MVZ257705	*	*	*	*	*
<i>Malacocheilus tomeri</i>	MVZ234632					
<i>Malayemys subtrijuga</i>	MVZ230474					
<i>Mauremys caspica</i>	MVZ237439					
<i>Mauremys mutica</i>	MVZ230487	*				
<i>Mesalina brevirostris</i>	MVZ246016					
<i>Microcebus murinus</i>	n/a	ENSMICT00000015362	ENSMICT00000011317	ENSMICT00000011434	ENSMICT00000012753	ENSMICT00000013304
<i>Monodelphis domestica</i>	n/a	ENSMODT00000000569	ENSMODT00000022740	ENSMODT00000017112	ENSMODT00000010416	ENSMODT00000001751
<i>Mus musculus</i>	n/a	ENSMUST000000080536	ENSMUST00000072796	ENSMUST00000034607	ENSMUST00000031621	ENSMUST00000001834
<i>Myotis lucifugus</i>	n/a	ENSMLUT00000008622	ENSMLUT00000005668	ENSMLUT00000014119		ENSMLUT00000007803
<i>Nectarinia olivacea</i>	JF1603					
<i>Nerodia rhombifer</i>	MVZ246092					
<i>Ochotona princeps</i>	n/a	ENSOPRT00000011558	ENSOPRT00000012582	ENSOPRT00000016051	ENSOPRT00000001733	ENSOPRT00000003498
<i>Ornithorhynchus anatinus</i>	n/a	ENSOANT00000021944	ENSOANT00000022441	ENSOANT00000005959	ENSOANT00000001187	ENSOANT00000008775
<i>Oryzotylagus cuniculus</i>	n/a	ENSOCUT00000001822		ENSOCUT00000010163	ENSOCUT00000007106	ENSOCUT00000008271
<i>Oryzias latipes</i>	n/a	ENSORL T00000003593	ENSORL T00000019051	ENSORL T00000019192	ENSORL T00000004269	ENSORL T00000024934
<i>Osaecaia ochrocephala</i>	MVZ232864				*	
<i>Otolemur garnettii</i>	n/a	ENSOGAT00000006909	ENSOGAT00000016359	ENSOGAT00000014612		ENSOGAT00000001426
<i>Pan troglodytes</i>	n/a	ENSPTRT00000030635	ENSPTRT00000034187	ENSPTRT00000050286	ENSPTRT00000034921	ENSPTRT00000026693

Scientific Name	Specimen #	ABCE1	AMD1	ARCNI	C7orf28B	C22orf28
<i>Pelodiscus sinensis</i>	MVZ226472					
<i>Pelomedusa subrufa</i>	MVZ236628					
<i>Phrynosoma sp.</i>	MVZ247578					
<i>Pipa pipa</i>	MVZ247516					
<i>Platymys platycephala</i>	MVZ247579					
<i>Platysternon megacephalum</i>	MVZ230486					
<i>Plestiodon gilberti</i>	MVZ249915					
<i>Pongo pygmaeus</i>	n/a	ENSPPYT00000017546	ENSPPYT00000019695	ENSPPYT00000004689	ENSPPYT00000020217	ENSPPYT00000013644
<i>Procavia capensis</i>	n/a	ENSPCAT00000006206	ENSPCAT00000006808	ENSPCAT00000001138		ENSPCAT00000013646
<i>Pseudemys concinna</i>	MVZ241520					
<i>Pteropus vampyrus</i>	n/a	ENSPVAT00000013171	ENSPVAT00000003440		ENSPVAT00000010918	ENSPVAT00000002381
<i>Rana livida</i>	MVZ241436					
<i>Rattus norvegicus</i>	n/a	ENSRNOT00000024753	ENSRNOT00000067493	ENSRNOT00000018598	ENSRNOT00000001365	ENSRNOT00000006570
<i>Rhinemra floridana</i>	MVZ233342					
<i>Rhinoclemmys pulcherrima</i>	MVZ230924	*				
<i>Sacalia quadriocellata</i>	MVZ230485					
<i>Sorex araneus</i>	n/a	ENSSART00000006278	ENSSART00000014392	ENSSART00000014532	ENSSART00000011035	ENSSART00000004807
<i>Spermophilus tridecemlineatus</i>	n/a	ENSSTOT00000004630	ENSSTOT00000000205	ENSSTOT00000014448	ENSSTOT00000006082	ENSSTOT00000003885
<i>Sternotherus carinatus</i>	MVZ238123					
<i>Sternotherus minor</i>	MVZ241528	*	*		*	
<i>Stigmocelys pardalis</i>	MVZ241333					
<i>Sus scrofa</i>	n/a	ENSSSCT00000009903	ENSSSCT00000004855		ENSSSCT00000008337	ENSSSCT00000000167
<i>Sylvietta virens</i>	JF1484					
<i>Taeniopygia guttata</i>	n/a	ENSTGUT00000002812	ENSTGUT00000012549		ENSTGUT00000009125	ENSTGUT00000011413
<i>Takijfuga rubripes</i>	n/a	ENSTRUT00000045023	ENSTRUT00000046885	ENSTRUT00000026344	ENSTRUT00000043189	ENSTRUT00000024909
<i>Tarsius syrichta</i>	n/a	ENSTSYT00000011351	ENSTSYT00000009698	ENSTSYT00000003067	ENSTSYT00000009425	ENSTSYT00000012320
<i>Terpsiphone batesi</i>	JF1488					
<i>Terrapene carolina</i>	MVZ241538					
<i>Testudo graeca</i>	MVZ234282					
<i>Testudo hermanni</i>	MVZ244866	*	*		*	*
<i>Tetraodon nigroviridis</i>	n/a	ENSTNIT00000007439	ENSTNIT00000005456	ENSTNIT00000017599	ENSTNIT00000001495	ENSTNIT00000005180
<i>Trachemys scripta</i>	MVZ241524					
<i>Tapia belangeri</i>	n/a	ENSTBET00000005012	ENSTBET00000015143	ENSTBET00000002082		ENSTBET00000004412
<i>Tursiops truncatus</i>	n/a	ENSTTRT00000012189	ENSTTRT00000006171	ENSTTRT00000004308	ENSTTRT00000005278	ENSTTRT00000006417
<i>Urocyon yemenensis</i>	MVZ236610					
<i>Vicugna pacos</i>	n/a	ENSVPAT00000003337	ENSVPAT00000000959	ENSVPAT00000005600		ENSVPAT00000005375
<i>Xantusia vigilis</i>	MVZ236373					
<i>Xenopeltis unicolor</i>	MVZ239387					
<i>Xenopus tropicalis</i>	n/a	ENSXETT000000053891	ENSXETT000000046135	ENSXETT000000046606	ENSXETT000000051883	ENSXETT000000025212

Scientific Name	Specimen #	CBX3	CPSF3	CSNK2A2	CTNNB1	DSCR3
<i>Accipiter cooperii</i>	MVZ181729					
<i>Agelaiocercus kingi</i>	MVZ183274					*
<i>Agrionemys borsfieldii</i>	MVZ245929					
<i>Alethe castanea</i>	JF1706					
<i>Alligator mississippiensis</i>	n/a					
<i>Alligator mississippiensis</i>	MVZ241485	*				
<i>Anas cyanoptera</i>	MVZ179315					
<i>Anniella pulchra</i>	MVZ247492					
<i>Anolis carolinensis</i>	n/a	ENSACAT00000012990	ENSACAT00000001924	ENSACAT00000008468	ENSACAT00000004530	ENSACAT00000009093
<i>Apalone ferox</i>	MVZ241535				*	*
<i>Aptenodytes patagonicus</i>	MVZ178921					
<i>Aspidoscelis tigris</i>	MVZ243345					
<i>Bipes biporus</i>	MVZ236260					
<i>Bos taurus</i>	n/a	ENSBAT00000007283	ENSBAT00000026303	ENSBAT00000018322	ENSBAT00000021838	ENSBAT00000006292
<i>Callithrix jacchus</i>	n/a	ENSCJAT00000015514	ENSCJAT00000012298	ENSCJAT00000022702	ENSCJAT00000035036	ENSCJAT00000005032
<i>Canis familiaris</i>	n/a	ENSACFT00000004632	ENSACFT00000005417	ENSACFT00000013508	ENSACFT00000008400	ENSACFT00000015549
<i>Casuarus casuarinus</i>	MVZ178927					
<i>Cavia prellus</i>	n/a	ENSCPOT00000026791	ENSCPOT00000011405	ENSCPOT00000008513	ENSCPOT00000002957	ENSCPOT00000013271
<i>Chamaeleo senegalensis</i>	MVZ249659					
<i>Chelodina longicollis</i>	MVZ230488					
<i>Chelonia mydas</i>	MVZ233240					
<i>Chelonoidis denticulata</i>	MVZ175391					
<i>Chelydra serpentina</i>	MVZ250691	*				
<i>Choloepus hoffmanni</i>	n/a	ENSCHOT00000013883	ENSCHOT00000000182	ENSCHOT00000010511	ENSCHOT00000011331	ENSCHOT00000004991
<i>Chrysemys picta</i>	MVZ241506	*				*
<i>Claudius angustatus</i>	MVZ164771					
<i>Cuora galbinifrons</i>	MVZ230933					
<i>Cuora pani</i>	MVZ230513					
<i>Cyrtodactylus</i>	JAM8710	*				*
<i>Danio rerio</i>	n/a	ENDSART00000012352	ENDSART00000040292	ENDSART00000045488	ENDSART00000038495	ENDSART00000036375
<i>Dasyops novemcinctus</i>	n/a	ENDSDNOT00000006243	ENDSDNOT00000008942	ENDSDNOT00000024550	ENDSDNOT00000008445	ENDSDNOT00000009922
<i>Dendropicos gabonensis</i>	JF1513					*
<i>Dermatemys mawii</i>	MVZ233618					
<i>Dermochelys coriacea</i>	MVZ149844					
<i>Dibamus sp.</i>	MVZ224113		*	*		
<i>Dierochelys reticularia</i>	MVZ230923	*				
<i>Dipodomys ordii</i>	n/a	ENDSDORT00000005485	ENDSDORT00000014264	ENDSDORT00000001894		ENDSDORT00000008988
<i>Draco beccarii</i>	JAM8961	*	*	*		*
<i>Draco sumatranus</i>	SZL71					
<i>Echinops telfairi</i>	n/a	ENSETET00000018578	ENSETET00000003895	ENSETET00000008464	ENSETET00000006456	
<i>Emydura subglobosa</i>	MVZ230496					
<i>Emys marmorata</i>	MVZ164996					
<i>Ensatina eschscholtzii</i>	RP158					
<i>Equus caballus</i>	n/a	ENSECAT00000013241	ENSECAT00000021609	ENSECAT00000014950	ENSECAT00000007477	ENSECAT00000025337

Scientific Name	Specimen #	CBX3	CPSF3	CSNK2A2	CTNNB1	DSCR3
<i>Eretmochelys imbricata</i>	MVZ233242	*				
<i>Erinaceus europaeus</i>	n/a	ENSEEUT00000007233	ENSEEUT00000015425	ENSEEUT00000000928	ENSEEUT00000011915	ENSEEUT00000011349
<i>Fejervarya limncharis</i>	MVZ251769					
<i>Felis catus</i>	n/a	ENSFCAT00000015228	ENSFCAT00000007854	ENSFCAT00000002695	ENSFCAT00000003470	ENSFCAT00000012834
<i>Gallus gallus</i>	n/a	ENSGALT00000037831	ENSGALT00000026493	ENSGALT00000001304	ENSGALT00000037585	ENSGALT00000025863
<i>Gasterosteus aculeatus</i>	n/a	ENSGACT00000005157	ENSGACT00000011824	ENSGACT000000020807	ENSGACT000000008011	ENSGACT000000027638
<i>Geochelone sulcata</i>	MVZ238888					
<i>Geomyda spengleri</i>	MVZ257081					
<i>Geotrypetes seraphini</i>	MVZ252475					
<i>Gorilla gorilla</i>	n/a	ENSGGOT00000017133	ENSGGOT00000001770		ENSGGOT00000001852	ENSGGOT00000001591
<i>Hardella thurjii</i>	MVZ248392	*				
<i>Heloderma horridum</i>	MVZ238489					
<i>Heteronotia binoei</i>	MKF226	*	*	*		*
<i>Homo sapiens</i>	n/a	ENST00000396386	ENST00000460593	ENST00000262506	ENST00000447363	ENST00000288304
<i>Hynobius arisanensis</i>	MVZ251776					
<i>Icthyophis bannanicus</i>	MVZ226265	*				*
<i>Indotestudo elongata</i>	MVZ234627					
<i>Kachuga smithii</i>	MVZ248394	*				
<i>Kinixys belliana</i>	MVZ234631					
<i>Kinosternon scorpioides</i>	MVZ247584					
<i>Leptotyphlops dulcis</i>	MVZ236621					
<i>Leucocephalon yuwonoi</i>	MVZ241499					
<i>Lissemys punctata</i>	MVZ248398					
<i>Loxodonta africana</i>	n/a	ENSLAFT00000006905	ENSLAFT00000009468	ENSLAFT00000013535	ENSLAFT00000018292	ENSLAFT00000000818
<i>Macaca mulatta</i>	n/a	ENSMMUT00000009535		ENSMMUT00000022339	ENSMMUT00000020513	ENSMMUT00000009734
<i>Macropus eugenii</i>	n/a		ENSMEUT00000013169	ENSMEUT00000013459	ENSMEUT00000005909	ENSMEUT00000007940
<i>Malaclemys terrapin</i>	MVZ257705	*	*	*		*
<i>Malacochersus tornieri</i>	MVZ234632					
<i>Malayemys subtrijuga</i>	MVZ230474	*				
<i>Mauremys caspica</i>	MVZ237439					
<i>Mauremys mutica</i>	MVZ230487					
<i>Mesalina brevirostris</i>	MVZ246016					
<i>Microcebus murinus</i>	n/a	ENSMICT00000006997	ENSMICT00000002370	ENSMICT00000008215	ENSMICT00000000358	ENSMICT000000006451
<i>Monodelphis domestica</i>	n/a		ENSMODT00000018712	ENSMODT00000017930	ENSMODT00000014262	ENSMODT000000026714
<i>Mus musculus</i>	n/a	ENSMUST000000081455	ENSMUST000000067284	ENSMUST000000056919	ENSMUST000000007130	ENSMUST000000023615
<i>Myotis lucifugus</i>	n/a	ENSMLUT00000007577	ENSMLUT00000010650	ENSMLUT00000000282	ENSMLUT000000008777	ENSMLUT000000005501
<i>Nectarinia olivacea</i>	JF1603					
<i>Nerodia rhombifer</i>	MVZ246092					
<i>Ochotona princeps</i>	n/a	ENSOPRT00000011694	ENSOPRT00000008918	ENSOPRT00000011854	ENSOPRT00000013563	
<i>Ornithorhynchus anatinus</i>	n/a	ENSOANT00000014582	ENSOANT00000013660	ENSOANT00000019189	ENSOANT00000010867	ENSOANT00000010234
<i>Oryzotylagus cuniculus</i>	n/a	ENSOCUT00000002464	ENSOCUT00000003728	ENSOCUT00000005074	ENSOCUT00000001468	ENSOCUT00000014107
<i>Oryzias latipes</i>	n/a	ENSORLT00000015476	ENSORLT00000023538	ENSORLT00000015609	ENSORLT00000007358	ENSORLT00000017827
<i>Osaecaiba ochrocephala</i>	MVZ232864					
<i>Otolemur garnettii</i>	n/a	ENSOGAT00000016585	ENSOGAT00000001681	ENSOGAT00000003283	ENSOGAT00000016166	ENSOGAT00000005529
<i>Pan troglodytes</i>	n/a	ENSPTRT000000035134	ENSPTRT000000021694	ENSPTRT00000015073	ENSPTRT000000065237	ENSPTRT0000000025921

Scientific Name	Specimen #	CBX3	CPSF3	CSNK2A2	CTNNB1	DSCR3
<i>Pelodiscus sinensis</i>	MVZ226472					
<i>Pelomedusa subrufa</i>	MVZ236628					
<i>Phrynosps sp.</i>	MVZ247578					
<i>Pipa pipa</i>	MVZ247516					
<i>Platenys platycephala</i>	MVZ247579					
<i>Platysternon megacephalum</i>	MVZ230486					
<i>Plestiodon gilberti</i>	MVZ249915					
<i>Pongo pygmaeus</i>	n/a	ENSPPYT00000020641	ENSPPYT00000014714	ENSPPYT00000008708	ENSPPYT00000016265	ENSPPYT00000013244
<i>Procarvia capensis</i>	n/a	ENSPCAT00000004870	ENSPCAT00000015059	ENSPCAT00000007858	ENSPCAT00000004160	ENSPCAT00000005518
<i>Pseudemys concinna</i>	MVZ241520	*				
<i>Pteropus rampyrus</i>	n/a	ENSPVAT00000006297	ENSPVAT00000005325	ENSPVAT00000004620	ENSPVAT00000008910	ENSPVAT00000016995
<i>Rana livida</i>	MVZ241436					
<i>Rattus norvegicus</i>	n/a	ENSRNOT00000015828	ENSRNOT00000009652	ENSRNOT00000016386	ENSRNOT00000026016	ENSRNOT00000002284
<i>Rhinemra floridana</i>	MVZ233342					
<i>Rhinoclemmys pulcherrima</i>	MVZ230924	*				*
<i>Sacalia quadriocellata</i>	MVZ230485					
<i>Sorex araneus</i>	n/a	ENSSART00000007998	ENSSART00000005233	ENSSART00000002629	ENSSART00000012911	
<i>Spermophilus tridecemlineatus</i>	n/a	ENSSTOT00000006221	ENSSTOT00000004483	ENSSTOT00000003542	ENSSTOT00000016133	ENSSTOT00000004451
<i>Sternotherus carinatus</i>	MVZ238123	*				
<i>Sternotherus minor</i>	MVZ241528	*				*
<i>Stigmocelys pardalis</i>	MVZ241333					
<i>Sus scrofa</i>	n/a	ENSSSCT00000018192		ENSSSCT00000003104	ENSSSCT00000012342	
<i>Sylvietta virens</i>	JF1484					
<i>Taeniopygia guttata</i>	n/a	ENSTGUT00000002996	ENSTGUT00000013563	ENSTGUT00000004937	ENSTGUT00000004994	ENSTGUT00000005148
<i>Takijfugu rubripes</i>	n/a	ENSTRUT000000031019	ENSTRUT00000017341	ENSTRUT00000019433	ENSTRUT00000008312	ENSTRUT00000016913
<i>Tarsius syrichta</i>	n/a	ENSTSYT00000008141	ENSTSYT00000002989	ENSTSYT00000011481	ENSTSYT00000011129	ENSTSYT00000002699
<i>Terpsiphone batesi</i>	JF1488					
<i>Terrapene carolina</i>	MVZ241538					
<i>Testudo graeca</i>	MVZ234282					
<i>Testudo bermanni</i>	MVZ244866	*	*	*	*	*
<i>Tetraodon nigroviridis</i>	n/a	ENSTNIT00000013186	ENSTNIT00000005201	ENSTNIT00000006632	ENSTNIT00000022527	ENSTNIT00000010640
<i>Trachemys scripta</i>	MVZ241524					
<i>Tapia belangeri</i>	n/a		ENSTBET00000002719	ENSTBET00000004226	ENSTBET00000012285	ENSTBET00000004239
<i>Tursiops truncatus</i>	n/a	ENSTTRT00000000711	ENSTTRT00000000862	ENSTTRT00000000842	ENSTTRT000000006094	ENSTTRT000000007721
<i>Varanus yemenensis</i>	MVZ236610					
<i>Vicugna pacos</i>	n/a	ENSVPAT00000010418	ENSVPAT000000008042	ENSVPAT00000003902	ENSVPAT000000007752	ENSVPAT000000009785
<i>Xantusia vigilis</i>	MVZ236373					
<i>Xenopeltis unicolor</i>	MVZ239387	*				
<i>Xenopus tropicalis</i>	n/a	ENSXETT00000026572	ENSXETT00000018123	ENSXETT00000030055	ENSXETT00000030302	ENSXETT00000028561

Scientific Name	Specimen #	EED	EIF3B	EIF3M	EXOC2	GAPDH
<i>Accipiter cooperii</i>	MVZ181729					
<i>Agelaiocercus kingi</i>	MVZ183274					
<i>Agrionemys borsfieldii</i>	MVZ245929					
<i>Aluthe castanea</i>	JF1706			*		
<i>Alligator mississippiensis</i>	n/a					ES3719
<i>Alligator mississippiensis</i>	MVZ241485		*	*		
<i>Anas cyanoptera</i>	MVZ179315					
<i>Anniella pulchra</i>	MVZ247492	*				
<i>Anolis carolinensis</i>	n/a	ENSACAT00000002498	ENSACAT00000008991	ENSACAT00000002586	ENSACAT00000017798	ENSACAT00000006562
<i>Apalone ferox</i>	MVZ241535					
<i>Aptenodytes patagonicus</i>	MVZ178921					
<i>Aspidoscelis tigris</i>	MVZ243345					
<i>Bipes biporus</i>	MVZ236260					
<i>Bos taurus</i>	n/a	ENSBTAT00000010324	ENSBTAT00000009832	ENSBTAT00000007021	ENSBTAT00000026920	ENSBTAT000000042879
<i>Callithrix jacchus</i>	n/a	ENSCJAT000000027694	ENSCJAT00000002720	ENSCJAT00000011435	ENSCJAT000000038403	
<i>Canis familiaris</i>	n/a	ENSACAT000000007206	ENSACAT000000026037	ENSACAT00000011839	ENSACAT00000014607	ENSACAT000000028274
<i>Casuarinus casuarinus</i>	MVZ178927					
<i>Cavia prellus</i>	n/a	ENSCPOT00000001836	ENSCPOT00000012661	ENSCPOT00000003890	ENSCPOT00000003649	ENSCPOT000000021753
<i>Chamaeleo senegalensis</i>	MVZ249659					
<i>Chelodina longicollis</i>	MVZ230488					
<i>Chelonia mydas</i>	MVZ233240					
<i>Chelonoidis denticulata</i>	MVZ175391					
<i>Chelydra serpentina</i>	MVZ250691					
<i>Choloepus hoffmanni</i>	n/a	ENSCHOT00000000744	ENSCHOT00000006169	ENSCHOT00000011332	ENSCHOT00000009936	
<i>Chrysomys picta</i>	MVZ241506			*		*
<i>Claudius angustatus</i>	MVZ164771					
<i>Cuora galbinifrons</i>	MVZ230933					
<i>Cuora pani</i>	MVZ230513					
<i>Cyrtodactylus</i>	JAM8710			*		
<i>Danio rerio</i>	n/a	ENDSART000000059345	ENDSART000000083129	ENDSART00000010256	ENDSART000000077962	ENDSART000000063800
<i>Dasyatis novemcinctus</i>	n/a	ENDSNOT00000001628	ENDSNOT00000009093	ENDSNOT00000012907	ENDSNOT00000018300	
<i>Dendropicus gabonensis</i>	JF1513			ENDSDORT00000004139		
<i>Dermatemys mawii</i>	MVZ233618					
<i>Dermochelys coriacea</i>	MVZ149844					
<i>Dibamus sp.</i>	MVZ224113					
<i>Dierochelys reticularia</i>	MVZ230923			*		*
<i>Dipodomys ordii</i>	n/a	ENDSDORT000000004240	ENDSDORT00000001402	ENDSDORT00000004139	ENDSDORT000000008787	
<i>Draco beccarii</i>	JAM8961					
<i>Draco sumatranus</i>	SZ1.71		*			
<i>Echinops telfairi</i>	n/a	ENSETET00000017442	ENSETET00000001719	ENSETET00000013437	ENSETET00000018422	ENSETET00000013748
<i>Emydura subglobosa</i>	MVZ230496					
<i>Emys marmorata</i>	MVZ164996					
<i>Ensatina eschscholtzii</i>	RP158					
<i>Equus caballus</i>	n/a	ENSECAT00000017377	ENSECAT00000014767	ENSECAT00000024368	ENSECAT00000029083	ENSECAT00000003799

Scientific Name	Specimen #	EED	EIF3B	EIF3M	EXOC2	GAPDH
<i>Ereimochelys imbricata</i>	MVZ233242					
<i>Erinaceus europaeus</i>	n/a	ENSEEUT00000005691	ENSEEUT00000014385	ENSEEUT00000001445	ENSEEUT00000012220	ENSEEUT00000011472
<i>Fagelaryia limnocharis</i>	MVZ251769					
<i>Felis catus</i>	n/a	ENSFCAT00000001593	ENSFCAT000000005393	ENSFCAT000000000310	ENSFCAT00000019319	ENSFCAT000000006876
<i>Gallus gallus</i>	n/a	ENSGALT000000022764	ENSGALT000000006789	ENSGALT00000019782	ENSGALT000000020938	ENSGALT000000023323
<i>Gasterosteus aculeatus</i>	n/a	ENSGACT00000019753	ENSGACT000000022154	ENSGACT000000021178	ENSGACT000000006621	ENSGACT00000013612
<i>Geochelone sulcata</i>	MVZ238888			*		
<i>Geomyda spengleri</i>	MVZ257081			*		
<i>Geotrypetes seraphini</i>	MVZ252475					
<i>Gorilla gorilla</i>	n/a	ENSGGOT00000007556	ENSGGOT00000016351	ENSGGOT00000002812	ENSGGOT00000005456	ENSGGOT00000009591
<i>Hardella thurjii</i>	MVZ248392					
<i>Heloderma horridum</i>	MVZ238489					
<i>Heteronotia binoei</i>	MKF226			*		
<i>Homo sapiens</i>	n/a	ENST00000327320	ENST00000314800	ENST00000323213	ENST00000448181	ENST00000396858
<i>Hynobius arisanensis</i>	MVZ251776					
<i>Icthyophis bannanicus</i>	MVZ226265			*		
<i>Indotestudo elongata</i>	MVZ234627					
<i>Kachuga smithii</i>	MVZ248394					
<i>Kinixys belliana</i>	MVZ234631					
<i>Kinosternon scorpioides</i>	MVZ247584					
<i>Leptotyphlops dulcis</i>	MVZ236621					
<i>Leuccephalon yuwonoi</i>	MVZ241499					
<i>Lissemys punctata</i>	MVZ248398					
<i>Loxodonta africana</i>	n/a	ENSLAFT00000002676	ENSLAFT00000001542	ENSLAFT000000022348	ENSLAFT00000012951	ENSLAFT00000002780
<i>Macaca mulatta</i>	n/a	ENSM MUT000000027559	ENSM MUT000000021777	ENSM MUT00000012612	ENSM MUT000000033140	ENSM MUT000000026261
<i>Macropus eugenii</i>	n/a		ENSM EUT00000000898	ENSM EUT00000000782	ENSM EUT00000005888	ENSM EUT00000015320
<i>Malaclemys terrapin</i>	MVZ257705	*	*	*	*	*
<i>Malacocheilus tornieri</i>	MVZ234632					
<i>Malayemys subtrijuga</i>	MVZ230474					
<i>Mauremys caspica</i>	MVZ237439					
<i>Mauremys mutica</i>	MVZ230487					
<i>Mesalina brevirostris</i>	MVZ246016					
<i>Microcebus murinus</i>	n/a	ENSM ICT00000010724	ENSM ICT00000010116	ENSM ICT00000011866	ENSM ICT00000008014	ENSM ICT00000006543
<i>Monodelphis domestica</i>	n/a	ENSM ODT00000005356	ENSM ODT00000012445	ENSM ODT00000012326	ENSM ODT00000010601	ENSM ODT00000038970
<i>Mus musculus</i>	n/a	ENSM UST00000032850	ENSM UST00000075771	ENSM UST00000111109	ENSM UST00000021785	ENSM UST00000101008
<i>Myotis lucifugus</i>	n/a	ENSM LUT00000014694	ENSM LUT00000013747	ENSM LUT00000014733	ENSM LUT00000014504	ENSM LUT00000014322
<i>Nectarinia olivacea</i>	JF1603					
<i>Nerodia rhombifer</i>	MVZ246092					
<i>Ochotona princeps</i>	n/a	ENSOPRT00000013846	ENSOPRT00000014528		ENSOPRT00000012712	ENSOPRT00000012511
<i>Ornithorhynchus anatinus</i>	n/a	ENSOANT00000014147	ENSOANT000000006843	ENSOANT00000023857	ENSOANT00000010113	ENSOANT00000007125
<i>Oryzotolagus cuniculus</i>	n/a		ENSOCUT00000010296	ENSOCUT00000004484	ENSOCUT00000016353	
<i>Oryzias latipes</i>	n/a	ENSORLT000000000076	ENSORLT00000011118	ENSORLT00000012359	ENSORLT00000015864	ENSORLT00000015301
<i>Oscacilia ochrocephala</i>	MVZ232864			*		
<i>Oryzomys latipes</i>	n/a	ENSOGAT00000013959	ENSOGAT000000006323	ENSOGAT00000015917	ENSOGAT00000014586	ENSOGAT00000010025
<i>Pan troglodytes</i>	n/a	ENSPTRT000000007719	ENSPTRT000000034838	ENSPTRT000000006554	ENSPTRT000000032644	ENSPTRT000000067000

Scientific Name	Specimen #	EED	EIF3B	EIF3M	EXOC2	GAPDH
<i>Pelodiscus sinensis</i>	MVZ226472					
<i>Pelomedusa subrufa</i>	MVZ236628					
<i>Phrynosops sp.</i>	MVZ247578					
<i>Pipa pipa</i>	MVZ247516					
<i>Platenys platycephala</i>	MVZ247579					
<i>Platysternon megacephalum</i>	MVZ230486					
<i>Plestiodon gilberti</i>	MVZ249915			*		
<i>Pongo pygmaeus</i>	n/a	ENSPPYT00000004458	ENSPPYT00000020167	ENSPPYT00000004031	ENSPPYT00000018800	ENSPPYT00000004964
<i>Procarina capensis</i>	n/a	ENSPCAT00000001374	ENSPCAT00000010543	ENSPCAT00000006971	ENSPCAT00000003759	ENSPCAT00000006908
<i>Pseudomys concinna</i>	MVZ241520			*		
<i>Pteropus vampyrus</i>	n/a	ENSPVAT00000007530	ENSPVAT00000013736	ENSPVAT00000015083	ENSPVAT00000000109	ENSPVAT00000012569
<i>Rana livida</i>	MVZ241436					
<i>Rattus norvegicus</i>	n/a	ENSRNOT000000024082	ENSRNOT000000060444	ENSRNOT000000064134	ENSRNOT000000040371	ENSRNOT000000025351
<i>Rhinerea floridana</i>	MVZ233342					
<i>Rhinoclemmys pulcherrima</i>	MVZ230924			*		
<i>Sacalia quadriocellata</i>	MVZ230485					
<i>Sorex araneus</i>	n/a	ENSSART000000002294	ENSSART000000006054	ENSSART00000010796	ENSSART000000006865	
<i>Spermophilus tridecemlineatus</i>	n/a	ENSSTOT000000004679		ENSSTOT00000010505		
<i>Sternotherus carinatus</i>	MVZ238123					
<i>Sternotherus minor</i>	MVZ241528					
<i>Stigmochelys pardalis</i>	MVZ241333					
<i>Sus scrofa</i>	n/a	ENSSSCT00000016271	ENSSSCT000000008307	ENSSSCT00000014541	ENSSSCT00000001075	
<i>Sylvietta virens</i>	JF1484					
<i>Taeniopygia guttata</i>	n/a	ENSTGUT00000013522	ENSTGUT000000008501	ENSTGUT00000005128	ENSTGUT00000002228	ENSTGUT00000013888
<i>Takifugu rubripes</i>	n/a	ENSTRUT000000034217	ENSTRUT000000028035	ENSTRUT00000015463	ENSTRUT000000004206	ENSTRUT000000008749
<i>Tarsius syrichta</i>	n/a	ENSTSYT000000005255	ENSTSYT000000008248	ENSTSYT000000000471	ENSTSYT000000002615	ENSTSYT000000002176
<i>Terpsiphone batesi</i>	JF1488					
<i>Terrapene carolina</i>	MVZ241538					
<i>Testudo graeca</i>	MVZ234282					
<i>Testudo hermanni</i>	MVZ244866			*		
<i>Tetraodon nigroviridis</i>	n/a	ENSTNIT000000002175	ENSTNIT000000007458	ENSTNIT00000017023	ENSTNIT00000011235	ENSTNIT00000011046
<i>Trachemys scripta</i>	MVZ241524					
<i>Tupaia belangeri</i>	n/a	ENSTBET000000008426	ENSTBET00000012885	ENSTBET00000003873	ENSTBET000000003680	ENSTBET00000001511
<i>Tursiops truncatus</i>	n/a	ENSTTRT00000012266	ENSTTRT00000014713	ENSTTRT00000002116	ENSTTRT00000001852	ENSTTRT00000016094
<i>Varamus yemenensis</i>	MVZ236610					
<i>Vicugna pacos</i>	n/a	ENSVPAT00000001723		ENSVPAT00000007791		
<i>Xantusia vigilis</i>	MVZ236373					
<i>Xenopeltis unicolor</i>	MVZ239387			*		
<i>Xenopus tropicalis</i>	n/a	ENSXETT000000022912	ENSXETT000000044504			ENSXETT000000055696

Scientific Name	Specimen #	GDE1	GNAS	GSPT1	H2AFY2	HIAT1
<i>Accipiter cooperii</i>	MVZ181729					
<i>Agelaiocercus kingi</i>	MVZ183274					*
<i>Agrionemys borsfieldii</i>	MVZ245929					
<i>Aluthe castanea</i>	JF1706					
<i>Alligator mississippiensis</i>	n/a					
<i>Alligator mississippiensis</i>	MVZ241485					
<i>Anas cyanoptera</i>	MVZ179315					
<i>Anniella pulchra</i>	MVZ247492			*	*	
<i>Anolis carolinensis</i>	n/a	ENSACAT00000001563	ENSACAT000000014380	ENSACAT00000004073	ENSACAT000000017852	ENSACAT00000003594
<i>Apalone ferox</i>	MVZ241535					
<i>Aptenodytes patagonicus</i>	MVZ178921					
<i>Aspidoscelis tigris</i>	MVZ243345					
<i>Bipes biporus</i>	MVZ236260					
<i>Bos taurus</i>	n/a	ENSBAT00000002710	ENSBAT00000002747	ENSBAT00000004656	ENSBAT00000001547	ENSBAT000000020947
<i>Callithrix jacchus</i>	n/a	ENSCJAT000000019019		ENSCJAT000000035572	ENSCJAT000000031934	
<i>Canis familiaris</i>	n/a	ENSACAT000000028730		ENSACAT000000029935	ENSACAT000000022230	ENSACAT000000031872
<i>Casuarinus casuarinus</i>	MVZ178927					
<i>Cavia prellus</i>	n/a	ENSCPOT000000013878	ENSCPOT00000003047	ENSCPOT000000009466	ENSCPOT000000006393	ENSCPOT000000006696
<i>Chamaeleo senegalensis</i>	MVZ249659					
<i>Chelodina longicollis</i>	MVZ230488					
<i>Chelonia mydas</i>	MVZ233240					
<i>Chelonoidis denticulata</i>	MVZ175391					
<i>Chelydra serpentina</i>	MVZ250691					
<i>Choloepus hoffmanni</i>	n/a	ENSCHOT000000011285	ENSCHOT000000008518	ENSCHOT000000003871	ENSCHOT000000002772	ENSCHOT000000010495
<i>Chrysomys picta</i>	MVZ241506		*	*		
<i>Claudius angustatus</i>	MVZ164771					
<i>Cuora galbinifrons</i>	MVZ230933					
<i>Cuora pani</i>	MVZ230513					
<i>Cyrtodactylus</i>	JAM8710			*		
<i>Danio rerio</i>	n/a		ENDSART000000064907	ENDSART000000016780	ENDSART000000058162	ENDSART000000063207
<i>Dasyatis novemcinctus</i>	n/a	ENDSDNOT000000013937	ENDSDNOT000000005586		ENDSDNOT000000011042	ENDSDNOT000000001703
<i>Dendropicus gabonensis</i>	JF1513					
<i>Dermatemys mawii</i>	MVZ233618					
<i>Dermochelys coriacea</i>	MVZ149844					
<i>Dibamus sp.</i>	MVZ224113					
<i>Dierochelys reticularia</i>	MVZ230923			*		
<i>Dipodomys ordii</i>	n/a	ENDSDORT000000000802	ENDSDORT000000000843		ENDSDORT000000011220	ENDSDORT000000007700
<i>Draco beccarii</i>	JAM8961		*	*	*	
<i>Draco sumatranus</i>	SZL71					
<i>Echinops telfairi</i>	n/a	ENSETET000000015649		ENSETET000000018250	ENSETET000000014537	
<i>Emydura subglobosa</i>	MVZ230496					
<i>Emys marmorata</i>	MVZ164996					
<i>Ensatina eschscholtzii</i>	RP158					
<i>Equus caballus</i>	n/a	ENSECAT000000010646	ENSECAT000000006514	ENSECAT000000023029	ENSECAT000000020715	ENSECAT000000006182

Scientific Name	Specimen #	GDE1	GNAS	GSPT1	H2AFY2	HIAT1
<i>Eremochelys imbricata</i>	MVZ233242					
<i>Erinaceus europaeus</i>	n/a			ENSEEUT00000008871	ENSEEUT00000003817	ENSEEUT00000013864
<i>Fagiaraya limnocharis</i>	MVZ251769					
<i>Felis catus</i>	n/a	ENSFCAT00000008371		ENSFCAT00000010118	ENSFCAT00000001225	ENSFCAT00000005792
<i>Gallus gallus</i>	n/a	ENSGALT00000039934	ENSGALT00000022385	ENSGALT00000041131	ENSGALT00000007651	ENSGALT00000008549
<i>Gasterosteus aculeatus</i>	n/a		ENSGACT00000021968	ENSGACT00000005765		ENSGACT00000022406
<i>Geochelone sulcata</i>	MVZ238888			*		
<i>Geomyda spengleri</i>	MVZ257081					
<i>Geotrypetes seraphini</i>	MVZ252475					
<i>Gorilla gorilla</i>	n/a	ENSGGOT00000002495		ENSGGOT00000004096	ENSGGOT00000005976	ENSGGOT00000000914
<i>Hardella thurjii</i>	MVZ248392					
<i>Heloderma horridum</i>	MVZ238489					
<i>Heteronotia binoei</i>	MKF226			*		*
<i>Homo sapiens</i>	n/a	ENST00000353258	ENST00000423027	ENST00000219627	ENST00000373255	ENST00000370152
<i>Hynobius arisanensis</i>	MVZ251776					
<i>Icthyophis bannanicus</i>	MVZ226265			*		
<i>Indotestudo elongata</i>	MVZ234627					
<i>Kachuga smithii</i>	MVZ248394					
<i>Kinixys belliana</i>	MVZ234631					
<i>Kinosternon scorpioides</i>	MVZ247584					
<i>Leptotyphlops dulcis</i>	MVZ236621					
<i>Leuccephalon yuwonoi</i>	MVZ241499					
<i>Lissemys punctata</i>	MVZ248398					
<i>Loxodonta africana</i>	n/a	ENSLAFT00000004616			ENSLAFT00000015624	
<i>Macaca mulatta</i>	n/a	ENSMMUT00000018712			ENSMMUT00000008973	ENSMMUT00000006013
<i>Macropus engertii</i>	n/a	ENSMEUT00000000404			ENSMEUT00000004739	ENSMEUT00000013365
<i>Malaclemys terrapin</i>	MVZ257705	*	*	*	*	*
<i>Malacocheilus tornieri</i>	MVZ234632					
<i>Malayemys subtrijuga</i>	MVZ230474			*		
<i>Mauremys caspica</i>	MVZ237439					
<i>Mauremys mutica</i>	MVZ230487					
<i>Mesalina brevirostris</i>	MVZ246016					
<i>Microcebus murinus</i>	n/a	ENSMICT00000011054			ENSMICT00000001628	ENSMICT00000005274
<i>Monodelphis domestica</i>	n/a	ENSMODT00000007908	ENSMODT00000027257	ENSMODT00000005947		ENSMODT00000004616
<i>Mus musculus</i>	n/a	ENSMUST00000038791	ENSMUST00000076605	ENSMUST00000080030	ENSMUST00000020283	ENSMUST00000029570
<i>Myotis lucifugus</i>	n/a	ENSMLUT00000002907	ENSMLUT00000011128	ENSMLUT00000000790	ENSMLUT00000006596	ENSMLUT00000002113
<i>Nectarinia olivacea</i>	JF1603					
<i>Nerodia rhombifer</i>	MVZ246092					
<i>Ochotona princeps</i>	n/a	ENSOPRT00000002156	ENSOPRT00000003954	ENSOPRT00000014247	ENSOPRT00000016479	ENSOPRT00000017102
<i>Ornithorhynchus anatinus</i>	n/a	ENSOANT00000005108	ENSOANT00000017127	ENSOANT00000014965	ENSOANT00000019241	ENSOANT00000019102
<i>Oryzotylagus cuniculus</i>	n/a	ENSOCUT00000003597	ENSOCUT00000005256	ENSOCUT00000004759	ENSOCUT00000004773	ENSOCUT00000007819
<i>Oryzias latipes</i>	n/a		ENSORLT00000024592	ENSORLT00000003583	ENSORLT00000001347	ENSORLT00000022528
<i>Oscacilia ochrocephala</i>	MVZ232864			*		
<i>Otolemur garnettii</i>	n/a	ENSOGAT00000010052			ENSOGAT00000005038	ENSOGAT00000005616
<i>Pan troglodytes</i>	n/a	ENSPTRT00000014437	ENSPTRT00000018131		ENSPTRT00000004853	ENSPTRT00000001873

Scientific Name	Specimen #	GDE1	GNAS	GSPT1	H2AFY2	HIAT1
<i>Pelodiscus sinensis</i>	MVZ226472					
<i>Pelomedusa subrufa</i>	MVZ236628					
<i>Phrynosoma sp.</i>	MVZ247578					
<i>Pipa pipa</i>	MVZ247516					
<i>Platenys platycephala</i>	MVZ247579					
<i>Platysternon megacephalum</i>	MVZ230486					
<i>Plestiodon gilberti</i>	MVZ249915					
<i>Pongo pygmaeus</i>	n/a	ENSPPYT00000008415	ENSPPYT00000010497	ENSPPYT00000008377	ENSPPYT00000002853	ENSPPYT00000001344
<i>Procarlia capensis</i>	n/a	ENSPCAT000000013921			ENSPCAT000000010544	ENSPCAT000000002219
<i>Pseudemys concinna</i>	MVZ241520					
<i>Pteropus vampyrus</i>	n/a	ENSPVAT000000002700		ENSPVAT00000005010	ENSPVAT000000000092	ENSPVAT000000017428
<i>Rana livida</i>	MVZ241436					
<i>Rattus norvegicus</i>	n/a			ENSRNOT000000003213	ENSRNOT000000033529	ENSRNOT0000000030281
<i>Rhinemys floridana</i>	MVZ233342					
<i>Rhinoclemmys pulcherrima</i>	MVZ230924			*		
<i>Sacalia quadriocellata</i>	MVZ230485					
<i>Sorex araneus</i>	n/a	ENSSART000000002560			ENSSART000000008502	ENSSART0000000002319
<i>Spermophilus tridecemlineatus</i>	n/a	ENSSTOT000000008616	ENSSTOT000000011777	ENSSTOT000000015416	ENSSTOT000000012746	ENSSTOT000000013446
<i>Sternotherus carinatus</i>	MVZ238123			*		
<i>Sternotherus minor</i>	MVZ241528			*		
<i>Stigmochelys pardalis</i>	MVZ241333					
<i>Sus scrofa</i>	n/a	ENSSSCT000000004220	ENSSSCT000000008247		ENSSSCT000000011227	ENSSSCT000000007523
<i>Sylvietta virens</i>	JF1484					
<i>Taeniopygia guttata</i>	n/a	ENSTGUT000000017100	ENSTGUT000000010130	ENSTGUT000000005366	ENSTGUT000000004272	ENSTGUT000000005517
<i>Takifugu rubripes</i>	n/a		ENSTRUT000000025276	ENSTRUT000000032126		ENSTRUT0000000020970
<i>Tarsius syrichta</i>	n/a	ENSTSYT000000001942	ENSTSYT000000009100	ENSTSYT000000013812	ENSTSYT000000013878	ENSTSYT000000010792
<i>Terpsiphone batesi</i>	JF1488					
<i>Terrapene carolina</i>	MVZ241538					
<i>Testudo graeca</i>	MVZ234282					
<i>Testudo hermanni</i>	MVZ244866			*		*
<i>Tetraodon nigroviridis</i>	n/a		ENSTNIT000000018119	ENSTNIT000000013303		ENSTNIT000000018172
<i>Trachemys scripta</i>	MVZ241524					
<i>Tupaia belangeri</i>	n/a		ENSTBET000000011317		ENSTBET000000006438	ENSTBET000000004674
<i>Tursiops truncatus</i>	n/a	ENSTTRT000000002454	ENSTTRT000000003699	ENSTTRT000000011636	ENSTTRT000000006610	ENSTTRT000000015285
<i>Varamus yemenensis</i>	MVZ236610					
<i>Vicugna pacos</i>	n/a	ENSVPAT000000003409		ENSVPAT000000006904	ENSVPAT000000002627	ENSVPAT000000001031
<i>Xantusia vigilis</i>	MVZ236373					
<i>Xenopeltis unicolor</i>	MVZ239387					
<i>Xenopus tropicalis</i>	n/a	ENSXETT000000047019	ENSXETT000000042648	ENSXETT000000012173	ENSXETT000000037396	ENSXETT000000028894

Scientific Name	Specimen #	HPD	IIFT57	KDM6A	KPNA4	MAT2B
<i>Accipiter cooperii</i>	MVZ181729					
<i>Agelaiocercus kingi</i>	MVZ183274				*	
<i>Agriomys borsfieldii</i>	MVZ245929					
<i>Aluthe castanea</i>	JF1706				*	
<i>Alligator mississippiensis</i>	n/a	ES316653				
<i>Alligator mississippiensis</i>	MVZ241485	*				
<i>Anas cyanoptera</i>	MVZ179315					
<i>Anniella pulchra</i>	MVZ247492					
<i>Anolis carolinensis</i>	n/a	ENSACAT00000011308		ENSACAT00000017962	ENSACAT00000007314	ENSACAT00000014603
<i>Apalone ferox</i>	MVZ241535					
<i>Aptenodytes patagonicus</i>	MVZ178921					
<i>Aspidoscelis tigris</i>	MVZ243345					
<i>Bipes biporus</i>	MVZ236260					
<i>Bos taurus</i>	n/a	ENSBAT00000005469		ENSBAT00000004874	ENSBAT00000023150	ENSBAT00000016399
<i>Callithrix jacchus</i>	n/a	ENSCJAT00000032883	ENSCJAT00000009621	ENSCJAT00000021421	ENSCJAT00000005654	ENSCJAT00000030042
<i>Canis familiaris</i>	n/a	ENSACAT00000012931	ENSACAT00000015740	ENSACAT00000023263	ENSACAT00000006953	ENSACAT00000027206
<i>Casuarinus casuarinus</i>	MVZ178927					
<i>Cavia prellus</i>	n/a	ENSCPOT00000000198	ENSCPOT00000008178		ENSCPOT00000010768	ENSCPOT00000002060
<i>Chamaeleo senegalensis</i>	MVZ249659					
<i>Chelodina longicollis</i>	MVZ230488					
<i>Chelonia mydas</i>	MVZ233240					
<i>Chelonoidis denticulata</i>	MVZ175391					
<i>Chelydra serpentina</i>	MVZ250691	*				
<i>Choloepus hoffmanni</i>	n/a		ENSCHOT00000003396	ENSCHOT00000006019	ENSCHOT00000008588	ENSCHOT00000013677
<i>Chrysomys picta</i>	MVZ241506	*				*
<i>Claudius angustatus</i>	MVZ164771					
<i>Cuora galbiniifrons</i>	MVZ230933					
<i>Cuora pani</i>	MVZ230513					
<i>Cyrtodactylus</i>	JAM8710		*		*	
<i>Danio rerio</i>	n/a	ENSADART000000020472	ENSADART00000012919		ENSADART00000034456	ENSADART00000003978
<i>Dasyatis novemcinctus</i>	n/a	ENSADNOT00000007264	ENSADNOT00000007812		ENSADNOT00000024704	ENSADNOT00000024947
<i>Dendropicus gabonensis</i>	JF1513	*			*	
<i>Dermatemys mawii</i>	MVZ233618					
<i>Dermochelys coriacea</i>	MVZ149844					
<i>Dibamus sp.</i>	MVZ224113		*		*	
<i>Dierochelys reticularia</i>	MVZ230923	*				
<i>Dipodomys ordii</i>	n/a	ENSADORT000000005249	ENSADORT00000006979	ENSADORT00000013577		ENSADORT00000011685
<i>Draco beccarii</i>	JAM8961		*	*	*	
<i>Draco sumatranus</i>	SZ1.71				*	
<i>Echinops telfairi</i>	n/a	ENSETET000000005015	ENSETET00000006069	ENSETET00000007805	ENSETET00000019546	
<i>Emydura subglobosa</i>	MVZ230496					
<i>Emys marmorata</i>	MVZ164996					
<i>Ensatina eschscholtzii</i>	RP158					
<i>Equus caballus</i>	n/a	ENSECAT00000012335	ENSECAT00000012110	ENSECAT00000008686	ENSECAT00000025442	ENSECAT00000022680

Scientific Name	Specimen #	HPD	IFT57	KDM6A	KPNA4	MAT2B
<i>Eremochelys imbricata</i>	MVZ233242					
<i>Erinaceus europaeus</i>	n/a	ENSEEUT00000007220	ENSEEUT00000005809	ENSEEUT00000011577	ENSEEUT00000001251	ENSEEUT000000014221
<i>Fagelaryia limnocharis</i>	MVZ251769					
<i>Felis catus</i>	n/a	ENSFCAT000000012108				ENSFCAT000000009027
<i>Gallus gallus</i>	n/a	ENSGALT000000006930	ENSGALT000000024781	ENSGALT000000026141	ENSGALT000000015544	ENSGALT000000002683
<i>Gasterosteus aculeatus</i>	n/a	ENSGACT000000015747	ENSGACT000000020891		ENSGACT000000013375	
<i>Geochelone sulcata</i>	MVZ238888					
<i>Geomyda spengleri</i>	MVZ257081					
<i>Geotrypetes seraphini</i>	MVZ252475					
<i>Gorilla gorilla</i>	n/a	ENSGGOT000000011056	ENSGGOT000000006838	ENSGGOT000000006651	ENSGGOT000000004810	
<i>Hardella thurjii</i>	MVZ248392	*			*	
<i>Heloderma horridum</i>	MVZ238489					
<i>Heteronotia binoei</i>	MKF226		*	*		
<i>Homo sapiens</i>	n/a	ENST000000289004	ENST000000457963	ENST000000433797	ENST000000334256	ENST000000321757
<i>Hynobius arisanensis</i>	MVZ251776					
<i>Icthyophis bannanicus</i>	MVZ226265				*	*
<i>Indotestudo elongata</i>	MVZ234627					
<i>Kachuga smithii</i>	MVZ248394					
<i>Kinixys belliana</i>	MVZ234631					
<i>Kinosternon scorpioides</i>	MVZ247584					
<i>Leptotyphlops dulcis</i>	MVZ236621					
<i>Leuccephalon yuwonoi</i>	MVZ241499					
<i>Lissemys punctata</i>	MVZ248398					
<i>Loxodonta africana</i>	n/a		ENSLAFT00000005208	ENSLAFT000000014817	ENSLAFT000000005976	ENSLAFT000000017578
<i>Macaca mulatta</i>	n/a	ENSMMUT000000023105	ENSMMUT000000012276	ENSMMUT000000004694	ENSMMUT000000025402	ENSMMUT000000019623
<i>Macropus eugenii</i>	n/a				ENSMEUT000000016776	ENSMEUT000000006168
<i>Malaclemys terrapin</i>	MVZ257705	*	*	*	*	*
<i>Malacocheilus tornieri</i>	MVZ234632					
<i>Malayemys subtrijuga</i>	MVZ230474					
<i>Mauremys caspica</i>	MVZ237439					
<i>Mauremys mutica</i>	MVZ230487					
<i>Mesalina brevirostris</i>	MVZ246016					
<i>Microcebus murinus</i>	n/a	ENSMICT000000013411	ENSMICT000000003764	ENSMICT000000015043	ENSMICT000000008819	ENSMICT000000005152
<i>Monodelphis domestica</i>	n/a	ENSMODT000000015491	ENSMODT000000022893	ENSMODT000000039307	ENSMODT000000020758	ENSMODT000000009630
<i>Mus musculus</i>	n/a	ENSMUST000000031398	ENSMUST000000046777	ENSMUST000000052368	ENSMUST000000029353	ENSMUST0000000101347
<i>Myotis lucifugus</i>	n/a	ENSMLUT000000000751	ENSMLUT000000010956	ENSMLUT000000001733	ENSMLUT000000013424	
<i>Nectarinia olivacea</i>	JF1603					
<i>Nerodia rhombifer</i>	MVZ246092				*	
<i>Ochotona princeps</i>	n/a	ENSOPRT000000016579	ENSOPRT000000017035	ENSOPRT000000014591	ENSOPRT000000009932	ENSOPRT000000009406
<i>Ornithorhynchus anatinus</i>	n/a		ENSOANT000000011935	ENSOANT000000010353	ENSOANT000000011092	ENSOANT000000007343
<i>Oryzotylagus cuniculus</i>	n/a	ENSOCUT000000011818	ENSOCUT000000003976	ENSOCUT000000014748	ENSOCUT000000001818	ENSOCUT000000008073
<i>Oryzias latipes</i>	n/a	ENSORLT000000013972	ENSORLT000000008271		ENSORLT000000007832	
<i>Oscacilia ochrocephala</i>	MVZ232864					
<i>Otalemur garnettii</i>	n/a	ENSOGAT000000008857	ENSOGAT000000002563	ENSOGAT000000015543	ENSOGAT000000012433	ENSOGAT000000007649
<i>Pan troglodytes</i>	n/a	ENSPTRT000000010220	ENSPTRT000000028401	ENSPTRT000000046054	ENSPTRT000000029060	ENSPTRT0000000032362

Scientific Name	Specimen #	HPD	IFT57	KDM6A	KPNA4	MAT2B
<i>Pelodiscus sinensis</i>	MVZ226472					
<i>Pelomedusa subrufa</i>	MVZ236628					
<i>Phrynosoma sp.</i>	MVZ247578					
<i>Pipa pipa</i>	MVZ247516					
<i>Platanus platycephala</i>	MVZ247579					
<i>Platysternon megacephalum</i>	MVZ230486					
<i>Plestiodon gilberti</i>	MVZ249915	*				
<i>Pongo pygmaeus</i>	n/a	ENSPPYT00000005981	ENSPPYT00000015783	ENSPPYT00000023647	ENSPPYT00000016575	ENSPPYT00000018648
<i>Procarina capensis</i>	n/a	ENSPCAT00000000285	ENSPCAT00000006641	ENSPCAT00000015177	ENSPCAT00000006665	ENSPCAT00000001601
<i>Pseudomys concinna</i>	MVZ241520					
<i>Pteropus vampyrus</i>	n/a	ENSPVAT00000001263	ENSPVAT00000011770	ENSPVAT00000004378	ENSPVAT00000004075	ENSPVAT00000017898
<i>Rana livida</i>	MVZ241436					
<i>Rattus norvegicus</i>	n/a	ENSRNOT00000001809	ENSRNOT00000002684	ENSRNOT00000005742	ENSRNOT00000014384	ENSRNOT00000004269
<i>Rhinoceros floridana</i>	MVZ233342					
<i>Rhinoceros pulcherrima</i>	MVZ230924	*			*	
<i>Sacalia quadriocellata</i>	MVZ230485					
<i>Sorex araneus</i>	n/a	ENSSART00000010050	ENSSART00000007898	ENSSART00000003566	ENSSART00000008968	ENSSART00000003274
<i>Spermophilus tridecemlineatus</i>	n/a		ENSSTOT00000013729	ENSSTOT00000001607	ENSSTOT00000004527	ENSSTOT00000009453
<i>Sternotherus carinatus</i>	MVZ238123	*				
<i>Sternotherus minor</i>	MVZ241528	*				
<i>Stigmochelys pardalis</i>	MVZ241333					
<i>Sus scrofa</i>	n/a	ENSSSCT00000010747	ENSSSCT00000013066			
<i>Sylvietta virens</i>	JF1484				*	
<i>Taeniopygia guttata</i>	n/a	ENSTGUT00000006350	ENSTGUT00000014172	ENSTGUT00000006361	ENSTGUT00000011634	ENSTGUT00000001541
<i>Takifugu rubripes</i>	n/a				ENSTRUT00000016095	ENSTRUT000000023598
<i>Tarsius syrichta</i>	n/a	ENSTSYT00000014714	ENSTSYT00000001905	ENSTSYT00000007332	ENSTSYT00000012099	ENSTSYT000000011210
<i>Terapia batesi</i>	JF1488	*				
<i>Terrapene carolina</i>	MVZ241538					
<i>Testudo graeca</i>	MVZ234282					
<i>Testudo hermanni</i>	MVZ244866			*		
<i>Tetraodon nigroviridis</i>	n/a	ENSTNIT00000011578			ENSTNIT00000010707	
<i>Trachemys scripta</i>	MVZ241524					
<i>Tupaia belangeri</i>	n/a	ENSTBET00000010098	ENSTBET00000010457	ENSTBET00000005077	ENSTBET00000004958	ENSTBET000000011513
<i>Tursiops truncatus</i>	n/a	ENSTTRT00000004197	ENSTTRT00000016533	ENSTTRT00000006761	ENSTTRT00000000079	ENSTTRT00000001154
<i>Varamus yemenensis</i>	MVZ236610					
<i>Vicugna pacos</i>	n/a	ENSVPAT00000006078	ENSVPAT00000003535	ENSVPAT00000003242	ENSVPAT00000008773	ENSVPAT00000008345
<i>Xantusia vigilis</i>	MVZ236373					
<i>Xenopeltis unicolor</i>	MVZ239387				*	
<i>Xenopus tropicalis</i>	n/a	ENSXETT00000017017	ENSXETT000000055345	ENSXETT00000010467	ENSXETT00000019107	ENSXETT000000032976

Scientific Name	Specimen #	NFIA-005	NFIA-007	NFIB	NFYB	NR2C2
<i>Accipiter cooperii</i>	MVZ181729					
<i>Agelaiocercus kingi</i>	MVZ183274					
<i>Agrionemys borsfieldii</i>	MVZ245929					
<i>Aluthe castanea</i>	JF1706					
<i>Alligator mississippiensis</i>	n/a					
<i>Alligator mississippiensis</i>	MVZ241485	*		*		
<i>Anas cyanoptera</i>	MVZ179315					
<i>Anniella pulchra</i>	MVZ247492					
<i>Anolis carolinensis</i>	n/a	ENSACAT00000007756	ENSACAT00000007756	ENSACAT00000017867	ENSACAT00000016083	ENSACAT00000012268
<i>Apalone ferox</i>	MVZ241535	*			*	
<i>Aptenodytes patagonicus</i>	MVZ178921					
<i>Aspidoscelis tigris</i>	MVZ243345					
<i>Bipes biporus</i>	MVZ236260					
<i>Bos taurus</i>	n/a	ENSBAT00000033431	ENSBAT00000033431	ENSBAT00000039408	ENSBAT00000043454	ENSBAT00000018228
<i>Callithrix jacchus</i>	n/a	ENSCJAT00000031451	ENSCJAT00000031433	ENSCJAT00000023615	ENSCJAT00000028469	ENSCJAT00000033321
<i>Canis familiaris</i>	n/a	ENSACAT00000029928	ENSACAT00000002388	ENSACAT00000002387	ENSACAT00000012000	ENSACAT00000007359
<i>Casuarinus casuarinus</i>	MVZ178927					
<i>Cavia prellus</i>	n/a	ENSCPOT00000010182	ENSCPOT00000028383	ENSCPOT00000003969	ENSCPOT00000003256	ENSCPOT00000001532
<i>Chamaeleo senegalensis</i>	MVZ249659					
<i>Chelodina longicollis</i>	MVZ230488					
<i>Chelonia mydas</i>	MVZ233240					
<i>Chelonoidis denticulata</i>	MVZ175391					
<i>Chelydra serpentina</i>	MVZ250691					
<i>Choloepus hoffmanni</i>	n/a	ENSCHOT00000007995		ENSCHOT00000002022		ENSCHOT00000002724
<i>Chrysomys picta</i>	MVZ241506	*			*	
<i>Claudius angustatus</i>	MVZ164771					
<i>Cuora galbiniifrons</i>	MVZ230933					
<i>Cuora pani</i>	MVZ230513					
<i>Cyrtodactylus</i>	JAM8710	*			*	
<i>Danio rerio</i>	n/a	ENDSART000000090242	ENDSART000000090237	ENDSART000000044670	ENDSART000000099770	
<i>Dasyatis novemcinctus</i>	n/a	ENDSDNOT00000002189	ENDSDNOT00000002189	ENDSDNOT00000019957	ENDSDNOT00000009978	ENDSDNOT00000006387
<i>Dendropicus gabonensis</i>	JF1513					
<i>Dermatemys mawii</i>	MVZ233618					
<i>Dermochelys coriacea</i>	MVZ149844					
<i>Dibamus sp.</i>	MVZ224113					
<i>Dierochelys reticularia</i>	MVZ230923	*			*	*
<i>Dipodomys ordii</i>	n/a	ENDSDORT00000005149	ENDSDORT00000005149	ENDSDORT00000004277	ENDSDORT00000001031	ENDSDORT00000007037
<i>Draco beccarii</i>	JAM8961				*	
<i>Draco sumatranus</i>	SZL71					
<i>Echinops telfairi</i>	n/a	ENSETET00000017068	ENSETET00000017068	ENSETET00000016780	ENSETET00000018215	ENSETET00000005089
<i>Emydura subglobosa</i>	MVZ230496					
<i>Emys marmorata</i>	MVZ164996					
<i>Ensatina eschscholtzii</i>	RP158					
<i>Equus caballus</i>	n/a	ENSECAT00000016287	ENSECAT00000016287	ENSECAT00000026549	ENSECAT00000009277	ENSECAT00000005581

Scientific Name	Specimen #	NFIA-005	NFIA-007	NFIB	NFYB	NR2C2
<i>Eremochelys imbricata</i>	MVZ233242					
<i>Erinaceus europaeus</i>	n/a	ENSEEUT00000015089	ENSEEUT00000015089	ENSEEUT00000012990	ENSEEUT00000008187	ENSEEUT00000001368
<i>Fagelaryia limnocharis</i>	MVZ251769					
<i>Felis catus</i>	n/a		ENSFCAT00000005998	ENSFCAT00000008475	ENSFCAT00000014087	ENSFCAT00000003538
<i>Gallus gallus</i>	n/a	ENSGALT00000017773	ENSGALT00000038014	ENSGALT00000008735	ENSGALT00000037651	ENSGALT00000013876
<i>Gasterosteus aculeatus</i>	n/a	ENSGACT00000004145	ENSGACT00000004145	ENSGACT00000026067		
<i>Geochelone sulcata</i>	MVZ238888					
<i>Geomyda spengleri</i>	MVZ257081					
<i>Geotrypetes seraphini</i>	MVZ252475					
<i>Gorilla gorilla</i>	n/a				ENSGGOT00000008904	ENSGGOT00000015534
<i>Hardella thurjii</i>	MVZ248392				*	
<i>Heloderma horridum</i>	MVZ238489					
<i>Heteronotia binoei</i>	MKF226	*			*	
<i>Homo sapiens</i>	n/a	ENST00000371185	ENST00000371187	ENST00000380934	ENST00000240055	ENST00000396874
<i>Hynobius arisanensis</i>	MVZ251776					
<i>Icthyophis bannanicus</i>	MVZ226265				*	
<i>Indotestudo elongata</i>	MVZ234627					
<i>Kachuga smithii</i>	MVZ248394					
<i>Kinixys belliana</i>	MVZ234631					
<i>Kinosternon scorpioides</i>	MVZ247584					
<i>Leptotyphlops dulcis</i>	MVZ236621					
<i>Leuccephalon yuwonoi</i>	MVZ241499					
<i>Lissemys punctata</i>	MVZ248398					
<i>Loxodonta africana</i>	n/a	ENSLAFT00000023105	ENSLAFT00000023105	ENSLAFT00000000868	ENSLAFT00000003821	ENSLAFT00000010226
<i>Macaca mulatta</i>	n/a	ENSMMUT00000019425	ENSMMUT00000047035	ENSMMUT00000020609	ENSMMUT00000031753	ENSMMUT00000045979
<i>Macropus eugenii</i>	n/a	ENSMEUT00000015107	ENSMEUT00000015107	ENSMEUT00000015979	ENSMEUT00000010247	ENSMEUT00000011849
<i>Malaclemys terrapin</i>	MVZ257705	*	*	*	*	*
<i>Malacocheilus tornieri</i>	MVZ234632					
<i>Malayemys subtrijuga</i>	MVZ230474					
<i>Mauremys caspica</i>	MVZ237439					
<i>Mauremys mutica</i>	MVZ230487					
<i>Mesalina brevirostris</i>	MVZ246016					
<i>Microcebus murinus</i>	n/a	ENSMICT00000005071	ENSMICT00000005071	ENSMICT00000005830	ENSMICT00000002044	
<i>Monodelphis domestica</i>	n/a	ENSMODT00000003069	ENSMODT00000003069	ENSMODT00000019093	ENSMODT00000002682	ENSMODT00000008911
<i>Mus musculus</i>	n/a	ENSMUST00000107055	ENSMUST00000107056	ENSMUST00000050872	ENSMUST00000105309	ENSMUST00000006047
<i>Myotis lucifugus</i>	n/a	ENSMLUT00000008386	ENSMLUT00000008386	ENSMLUT00000000243	ENSMLUT00000005468	ENSMLUT00000010069
<i>Nectarinia olivacea</i>	JF1603					
<i>Nerodia rhombifer</i>	MVZ246092					
<i>Ochotona princeps</i>	n/a	ENSOPRT00000015285	ENSOPRT00000015285	ENSOPRT00000000686	ENSOPRT00000008110	ENSOPRT00000009348
<i>Ornithorhynchus anatinus</i>	n/a	ENSOANT00000021222	ENSOANT00000019758	ENSOANT00000016052	ENSOANT00000008818	ENSOANT00000017254
<i>Oryzotylagus cuniculus</i>	n/a	ENSOCUT00000005892		ENSOCUT00000004248		ENSOCUT00000006977
<i>Oryzias latipes</i>	n/a	ENSORLT00000004401	ENSORLT00000004401	ENSORLT00000006543	ENSORLT00000014995	
<i>Oscacilia ochrocephala</i>	MVZ232864				*	
<i>Orolemmur garnettii</i>	n/a	ENSOGAT00000008608	ENSOGAT00000008608	ENSOGAT00000016285	ENSOGAT00000006281	ENSOGAT00000014040
<i>Pan troglodytes</i>	n/a	ENSPTRT000000043693	ENSPTRT00000001513	ENSPTRT00000056683	ENSPTRT00000009878	ENSPTRT000000064857

Scientific Name	Specimen #	NFIA-005	NFIA-007	NFIB	NFYB	NR2C2
<i>Pelodiscus sinensis</i>	MVZ226472					
<i>Pelomedusa subrufa</i>	MVZ236628					
<i>Phrynosoma sp.</i>	MVZ247578					
<i>Pipa pipa</i>	MVZ247516					
<i>Platemys platycephala</i>	MVZ247579					
<i>Platysternon megacephalum</i>	MVZ230486					
<i>Plestiodon gilberti</i>	MVZ249915				*	
<i>Pongo pygmaeus</i>	n/a	ENSPPYT00000001553	ENSPPYT00000001553	ENSPPYT00000022399	ENSPPYT00000005797	ENSPPYT000000015596
<i>Procarina capensis</i>	n/a	ENSPCAT000000012769	ENSPCAT000000012769	ENSPCAT00000002414	ENSPCAT000000012585	ENSPCAT000000011042
<i>Pseudemys concinna</i>	MVZ241520				*	
<i>Pteropus vampyrus</i>	n/a	ENSPVAT000000005805	ENSPVAT000000005805	ENSPVAT00000007630	ENSPVAT00000007284	ENSPVAT000000017233
<i>Rana livida</i>	MVZ241436					
<i>Rattus norvegicus</i>	n/a	ENSRNOT000000009651	ENSRNOT000000009505	ENSRNOT000000067845	ENSRNOT000000067075	ENSRNOT000000014353
<i>Rhinoceros floridana</i>	MVZ233342					
<i>Rhinoceros pulcherrima</i>	MVZ230924	*			*	
<i>Sacalia quadriocellata</i>	MVZ230485					
<i>Sorex araneus</i>	n/a	ENSSART000000004463	ENSSART000000004463	ENSSART000000003975	ENSSART000000008993	ENSSART000000004499
<i>Spermophilus tridecemlineatus</i>	n/a	ENSSTOT000000002732	ENSSTOT000000002732	ENSSTOT000000002242	ENSSTOT000000014413	ENSSTOT000000005254
<i>Sternotherus carinatus</i>	MVZ238123					
<i>Sternotherus minor</i>	MVZ241528					
<i>Stigmochelys pardalis</i>	MVZ241333					
<i>Sus scrofa</i>	n/a			ENSSSCT000000005721	ENSSSCT000000000922	ENSSSCT000000012690
<i>Sylvietta virens</i>	JF1484					
<i>Taeniopygia guttata</i>	n/a	ENSTGUT000000010179	ENSTGUT000000010179	ENSTGUT00000004855	ENSTGUT000000011626	ENSTGUT000000011047
<i>Takifugu rubripes</i>	n/a			ENSTRUT000000003505	ENSTRUT000000046107	
<i>Tarsius syrichta</i>	n/a	ENSTSYT000000014719	ENSTSYT000000014719	ENSTSYT000000008733	ENSTSYT000000014482	ENSTSYT000000014210
<i>Terpsiphone batesi</i>	JF1488					
<i>Terrapene carolina</i>	MVZ241538					
<i>Testudo graeca</i>	MVZ234282	*				
<i>Testudo hermanni</i>	MVZ244866	*	*			
<i>Tetraodon nigroviridis</i>	n/a			ENSTNIT000000001367		
<i>Trachemys scripta</i>	MVZ241524					
<i>Tupaia belangeri</i>	n/a	ENSTBET000000013446	ENSTBET000000013446	ENSTBET000000001938	ENSTBET000000005718	ENSTBET000000002428
<i>Tursiops truncatus</i>	n/a	ENSTTRT000000011221	ENSTTRT000000011221	ENSTTRT000000010572	ENSTTRT000000004628	ENSTTRT000000009878
<i>Varamus yemenensis</i>	MVZ236610					
<i>Vicugna pacos</i>	n/a	ENSVPAT000000004054	ENSVPAT000000004054	ENSVPAT000000005805	ENSVPAT000000007874	
<i>Xantusia vigilis</i>	MVZ236373					
<i>Xenopeltis unicolor</i>	MVZ239387	*				
<i>Xenopus tropicalis</i>	n/a	ENSXETT000000019479	ENSXETT000000019478	ENSXETT00000001859	ENSXETT000000042558	ENSXETT000000010442

Scientific Name	Specimen #	NTS	OAT	OAZ1	OLA1	PABPC1
<i>Accipiter cooperii</i>	MVZ181729					
<i>Agelaiocercus kingi</i>	MVZ183274					
<i>Agriomys borsfieldii</i>	MVZ245929					
<i>Aluthe castanea</i>	JF1706			*		
<i>Alligator mississippiensis</i>	n/a		ES316854	ES320547		
<i>Alligator mississippiensis</i>	MVZ241485			*	*	*
<i>Anas cyanoptera</i>	MVZ179315					
<i>Anniella pulchra</i>	MVZ247492			*	*	*
<i>Anolis carolinensis</i>	n/a		ENSACAT00000017625		ENSACAT00000017111	ENSACAT00000016610
<i>Apalone ferox</i>	MVZ241535					
<i>Aptenodytes patagonicus</i>	MVZ178921					
<i>Aspidoscelis tigris</i>	MVZ243345					
<i>Bipes biporus</i>	MVZ236260					
<i>Bos taurus</i>	n/a	ENSBTAT00000006980	ENSBTAT00000009097	ENSBTAT00000024656	ENSBTAT00000052195	
<i>Callithrix jacchus</i>	n/a	ENSCJAT00000022335	ENSCJAT00000024507	ENSCJAT00000025682	ENSCJAT00000015380	ENSCJAT00000030824
<i>Canis familiaris</i>	n/a	ENSACAT00000009648	ENSACAT00000020197	ENSACAT00000030861	ENSACAT00000020924	ENSACAT00000000885
<i>Casuarinus casuarinus</i>	MVZ178927					
<i>Cavia prellus</i>	n/a	ENSCPOT00000006642	ENSCPOT00000010434	ENSCPOT00000004747	ENSCPOT00000026056	ENSCPOT00000013042
<i>Chamaelo senegalensis</i>	MVZ249659					
<i>Chelodina longicollis</i>	MVZ230488			*		
<i>Chelonia mydas</i>	MVZ233240					
<i>Chelonoidis denticulata</i>	MVZ175391					
<i>Chelydra serpentina</i>	MVZ250691		*	*		*
<i>Choloepus hoffmanni</i>	n/a	ENSCHOT00000012681	ENSCHOT00000002345	ENSCHOT00000009350	ENSCHOT00000005061	ENSCHOT00000002088
<i>Chrysomys picta</i>	MVZ241506		*	*	*	*
<i>Claudius angustatus</i>	MVZ164771					
<i>Cnora galbiniifrons</i>	MVZ230933					
<i>Cnora pani</i>	MVZ230513					
<i>Cyrtodactylus</i>	JAM8710			*	*	*
<i>Danio rerio</i>	n/a			ENDART00000105532	ENDART00000045299	ENDART00000032474
<i>Dasylops novemcinctus</i>	n/a	ENDNOT00000005891	ENDNOT00000006180	ENDNOT00000001344	ENDNOT00000010627	ENDNOT000000008724
<i>Dendropicus gabonensis</i>	JF1513		*	*		
<i>Dermatemys mawii</i>	MVZ233618					
<i>Dermochelys coriacea</i>	MVZ149844					
<i>Dibamus sp.</i>	MVZ224113			*	*	*
<i>Dierochelys reticularia</i>	MVZ230923		*	*		*
<i>Dipodomys ordii</i>	n/a	ENDDORT00000016283	ENDDORT00000013083	ENDDORT00000016529	ENDDORT00000008680	ENDDORT00000013935
<i>Draco beccarii</i>	JAM8961			*	*	*
<i>Draco sumatranus</i>	SZL71			*		*
<i>Echinops telfairi</i>	n/a	ENSETET00000013992		ENSETET00000017945	ENSETET00000012453	ENSETET00000002795
<i>Emydura subglobosa</i>	MVZ230496					
<i>Emys marmorata</i>	MVZ164996					
<i>Ensatina eschscholtzii</i>	RP158					
<i>Equus caballus</i>	n/a	ENSECAT00000011287	ENSECAT00000015791		ENSECAT00000017411	ENSECAT00000022353

Scientific Name	Specimen #	NTS	OAT	OAZ1	OLA1	PABPC1
<i>Ereimochelys imbricata</i>	MVZ233242				*	
<i>Erinaceus europaeus</i>	n/a	ENSEEUT0000005207	ENSEEUT0000005324		ENSEEUT00000015211	ENSEEUT00000002703
<i>Fejervarya limncharis</i>	MVZ251769			*		
<i>Felis catus</i>	n/a	ENSFCAT00000011896	ENSFCAT00000013851		ENSFCAT00000007672	ENSFCAT00000001874
<i>Gallus gallus</i>	n/a	ENSGALT00000018184	ENSGALT00000038658	ENSGALT00000001183	ENSGALT00000015187	ENSGALT00000023333
<i>Gasterosteus aculeatus</i>	n/a		ENSGACT00000003381	ENSGACT00000016744	ENSGACT00000013368	ENSGACT00000004695
<i>Geochelone sulcata</i>	MVZ238888			*		
<i>Geomyda spengleri</i>	MVZ257081					
<i>Geotrypetes seraphini</i>	MVZ252475					
<i>Gorilla gorilla</i>	n/a	ENSGGOT00000010280	ENSGGOT0000000852	ENSGGOT00000007606	ENSGGOT00000005845	ENSGGOT00000006814
<i>Hardella thurjii</i>	MVZ248392	*	*			
<i>Heloderma horridum</i>	MVZ238489					
<i>Heteromolia binoei</i>	MKF226			*		*
<i>Homo sapiens</i>	n/a	ENST00000256010	ENST00000368845	ENST00000322297	ENST00000424954	ENST00000347137
<i>Hynobius arisanensis</i>	MVZ251776					
<i>Icthyophis bannanicus</i>	MVZ226265			*		*
<i>Indotestudo elongata</i>	MVZ234627					
<i>Kachuga smithii</i>	MVZ248394			*		
<i>Kinixys belliana</i>	MVZ234631					
<i>Kinosternon scorpioides</i>	MVZ247584					
<i>Leptotyphlops dulcis</i>	MVZ236621			*		
<i>Leucocephalon yunnanoi</i>	MVZ241499					
<i>Lissemys punctata</i>	MVZ248398					
<i>Loxodonta africana</i>	n/a	ENSLAFT00000005892	ENSLAFT00000006790		ENSLAFT00000002707	
<i>Macaca mulatta</i>	n/a	ENSMMUT00000032574	ENSMMUT00000014670	ENSMMUT00000025219	ENSMMUT00000009899	ENSMMUT00000024997
<i>Macropus eugenii</i>	n/a	ENSMEUT00000001840	ENSMEUT00000005670	ENSMEUT00000009144	ENSMEUT00000010004	
<i>Malaclemys terrapin</i>	MVZ257705	*	*	*	*	*
<i>Malacochersus tornieri</i>	MVZ234632					
<i>Malayemys subtrijuga</i>	MVZ230474			*		
<i>Mauremys caspica</i>	MVZ237439					
<i>Mauremys mutica</i>	MVZ230487		*			
<i>Mesalina brevirostris</i>	MVZ246016					
<i>Microcebus murinus</i>	n/a	ENSMICT00000007941	ENSMICT00000003893	ENSMICT00000015472	ENSMICT00000000353	ENSMICT000000009124
<i>Monodelphis domestica</i>	n/a	ENSMODT000000026101	ENSMODT000000021066	ENSMODT00000005918	ENSMODT00000011836	ENSMODT00000036071
<i>Mus musculus</i>	n/a	ENSMUST000000020040	ENSMUST000000084500	ENSMUST000000060987	ENSMUST00000100015	ENSMUST00000001809
<i>Myotis lucifugus</i>	n/a	ENSMLUT000000008725	ENSMLUT00000007043	ENSMLUT000000004445		ENSMLUT000000012255
<i>Nectarinia olivacea</i>	JF1603					
<i>Nerodia rhombifer</i>	MVZ246092			*		
<i>Ochotona princeps</i>	n/a	ENSOPRT00000003086	ENSOPRT00000014499		ENSOPRT000000005557	ENSOPRT000000009688
<i>Ornithorhynchus anatinus</i>	n/a	ENSOANT00000010600	ENSOANT00000009207	ENSOANT00000001022	ENSOANT00000013706	ENSOANT00000011635
<i>Oryzotylagus cuniculus</i>	n/a	ENSOCUT00000010331	ENSOCUT00000013545	ENSOCUT00000014861	ENSOCUT00000003576	ENSOCUT000000004355
<i>Oryzias latipes</i>	n/a		ENSORLT00000012067	ENSORLT00000006348	ENSORLT00000017293	ENSORLT00000003765
<i>Oscacilia ochrocephala</i>	MVZ232864			*		*
<i>Otolemur garnettii</i>	n/a	ENSOGAT00000003212		ENSOGAT00000006374	ENSOGAT00000016561	ENSOGAT00000007928
<i>Pan troglodytes</i>	n/a	ENSPTRT000000009700	ENSPTRT00000005772	ENSPTRT00000018791	ENSPTRT000000048724	ENSPTRT000000037886

Scientific Name	Specimen #	NTS	OAT	OAZ1	OLA1	PABPC1
<i>Pelodiscus sinensis</i>	MVZ226472					
<i>Pelomedusa subrufa</i>	MVZ236628			*		
<i>Phrynosops sp.</i>	MVZ247578					
<i>Pipa pipa</i>	MVZ247516			*		
<i>Platemys platycephala</i>	MVZ247579					
<i>Platysternon megacephalum</i>	MVZ230486					
<i>Plestiodon gilberti</i>	MVZ249915		*	*		
<i>Pongo pygmaeus</i>	n/a	ENSPPYT00000005698	ENSPPYT00000003305	ENSPPYT00000010900	ENSPPYT00000015039	ENSPPYT00000021919
<i>Procarina capensis</i>	n/a	ENSPCAT00000003374	ENSPCAT00000015456		ENSPCAT00000014883	ENSPCAT00000007115
<i>Pseudemys concinna</i>	MVZ241520					
<i>Pteropus vampyrus</i>	n/a	ENSPVAT00000006900	ENSPVAT00000010705	ENSPVAT00000013312	ENSPVAT00000000982	ENSPVAT00000016640
<i>Rana livida</i>	MVZ241436					
<i>Rattus norvegicus</i>	n/a	ENSRNOT00000005706	ENSRNOT00000022628	ENSRNOT00000026341	ENSRNOT00000041026	ENSRNOT00000012775
<i>Rhineura floridana</i>	MVZ233342			*		
<i>Rhinoclemmys pulcherrima</i>	MVZ230924		*	*		
<i>Sacalia quadriocellata</i>	MVZ230485					
<i>Sorex araneus</i>	n/a	ENSSART00000010512	ENSSART00000011210	ENSSART00000009634	ENSSART00000010738	ENSSART00000004437
<i>Spermophilus tridecemlineatus</i>	n/a	ENSSTOT00000010246	ENSSTOT00000004515		ENSSTOT00000015549	ENSSTOT00000016398
<i>Sternotherus carinatus</i>	MVZ238123		*	*		
<i>Sternotherus minor</i>	MVZ241528		*			*
<i>Stigmochelys pardalis</i>	MVZ241333					
<i>Sus scrofa</i>	n/a	ENSSSCT00000001018	ENSSSCT00000011746	ENSSSCT00000014704		ENSSSCT00000000652
<i>Sylvietta virens</i>	JF1484			*		
<i>Taeniopygia guttata</i>	n/a	ENSTGUT00000008249	ENSTGUT00000011917	ENSTGUT00000000329	ENSTGUT00000009088	ENSTGUT00000012575
<i>Takifugu rubripes</i>	n/a		ENSTRUT00000025192	ENSTRUT00000045925	ENSTRUT00000045305	ENSTRUT00000008899
<i>Tarsius syrichta</i>	n/a	ENSTSYT00000005374	ENSTSYT00000007487		ENSTSYT00000005532	ENSTSYT00000004231
<i>Terpsiphone batesi</i>	JF1488					
<i>Terrapene carolina</i>	MVZ241538					
<i>Testudo graeca</i>	MVZ234282			*		
<i>Testudo hermanni</i>	MVZ244866		*	*	*	*
<i>Tetraodon nigroviridis</i>	n/a		ENSTNIT00000022858	ENSTNIT00000008520	ENSTNIT00000020577	ENSTNIT00000012204
<i>Trachemys scripta</i>	MVZ241524					
<i>Tupaia belangeri</i>	n/a		ENSTBET00000013312		ENSTBET00000007570	ENSTBET00000009326
<i>Tursiops truncatus</i>	n/a	ENSTTRT00000004424	ENSTTRT00000004237	ENSTTRT00000003115	ENSTTRT00000016999	ENSTTRT00000004689
<i>Varanus yemenensis</i>	MVZ236610					
<i>Vigna paxos</i>	n/a	ENSVPAT00000008090	ENSVPAT00000011447		ENSVPAT00000009661	ENSVPAT00000003864
<i>Xantusia vigilis</i>	MVZ236373					
<i>Xenopeltis unicolor</i>	MVZ239387		*	*		
<i>Xenopus tropicalis</i>	n/a		ENSXETT00000021114	ENSXETT00000029891	ENSXETT00000018601	ENSXETT00000014537

Scientific Name	Specimen #	PAPOLA	PDCD10	POLR2B	PPM1A	PPTC7
<i>Accipiter cooperii</i>	MVZ181729					
<i>Agelaiocercus kingi</i>	MVZ183274					
<i>Agrionemys borsfieldii</i>	MVZ245929					
<i>Aluthe castanea</i>	JF1706		*			
<i>Alligator mississippiensis</i>	n/a					
<i>Alligator mississippiensis</i>	MVZ241485					
<i>Anas cyanoptera</i>	MVZ179315					
<i>Anniella pulchra</i>	MVZ247492		*			
<i>Anolis carolinensis</i>	n/a	ENSACAT00000013626	ENSACAT00000017961	ENSACAT00000002264	ENSACAT00000015231	ENSACAT00000001872
<i>Apalone ferox</i>	MVZ241535		*			
<i>Aptenodytes patagonicus</i>	MVZ178921					
<i>Aspidoscelis tigris</i>	MVZ243345					
<i>Bipes biporus</i>	MVZ236260					
<i>Bos taurus</i>	n/a	ENSBAT00000005300	ENSBAT00000044960	ENSBAT00000025800	ENSBAT00000024128	ENSBAT00000024393
<i>Callithrix jacchus</i>	n/a	ENSCJAT00000012864	ENSCJAT00000005295	ENSCJAT00000008049	ENSCJAT00000037427	ENSCJAT00000017964
<i>Canis familiaris</i>	n/a	ENSACAT00000028223	ENSACAT00000023131	ENSACAT00000003641	ENSACAT00000039670	ENSACAT00000013491
<i>Casuarinus casuarinus</i>	MVZ178927					
<i>Cavia prellus</i>	n/a	ENSCPOT00000002682	ENSCPOT00000012071	ENSCPOT00000013132	ENSCPOT00000002498	ENSCPOT00000020347
<i>Chamaeleo senegalensis</i>	MVZ249659					
<i>Chelodina longicollis</i>	MVZ230488					
<i>Chelonia mydas</i>	MVZ233240					
<i>Chelonoidis denticulata</i>	MVZ175391					
<i>Chelydra serpentina</i>	MVZ250691		*			
<i>Choloepus hoffmanni</i>	n/a		ENSCHOT00000007924	ENSCHOT00000005282		ENSCHOT00000001244
<i>Chrysemys picta</i>	MVZ241506		*			
<i>Claudius angustatus</i>	MVZ164771					
<i>Cuora galbinifrons</i>	MVZ230933					
<i>Cuora pani</i>	MVZ230513					
<i>Cyrtodactylus</i>	JAM8710	*	*			
<i>Danio rerio</i>	n/a	ENDSART00000015154	ENDSART00000019063	ENDSART00000041408	ENDSART00000076527	ENDSART00000015214
<i>Dasyatis novemcinctus</i>	n/a	ENDSNOT00000016266	ENDSNOT00000015853	ENDSNOT00000007716	ENDSNOT00000005117	ENDSNOT00000003222
<i>Dendropicus gabonensis</i>	JF1513	*				
<i>Dermatemys mawii</i>	MVZ233618					
<i>Dermochelys coriacea</i>	MVZ149844					
<i>Dibamus sp.</i>	MVZ224113		*			
<i>Dierochelys reticularia</i>	MVZ230923		*			
<i>Dipodomys ordii</i>	n/a	ENDSORT00000013416	ENDSORT00000005503	ENDSORT00000010079	ENDSORT00000007776	ENDSORT00000010344
<i>Draco beccarii</i>	JAM8961		*			
<i>Draco sumatranus</i>	SZL71		*			
<i>Echinops telfairi</i>	n/a	ENSETET00000004894	ENSETET00000006489	ENSETET00000011734	ENSETET00000002065	ENSETET00000007009
<i>Emydura subglobosa</i>	MVZ230496					
<i>Emys marmorata</i>	MVZ164996					
<i>Ensatina escholtzii</i>	RP158					
<i>Equus caballus</i>	n/a	ENSECAT00000022852	ENSECAT00000019099	ENSECAT00000014599	ENSECAT00000017811	ENSECAT00000006678

Scientific Name	Specimen #	PAPOLA	PDCD10	POLR2B	PPM1A	PPTC7
<i>Ereimochelys imbricata</i>	MVZ233242					
<i>Erinaceus europaeus</i>	n/a	ENSEEUT00000013684	ENSEEUT00000015411	ENSEEUT00000009865	ENSEEUT00000004481	ENSEEUT00000015474
<i>Fagelaryia limncharis</i>	MVZ251769					
<i>Felis catus</i>	n/a	ENSFCAT00000006785	ENSFCAT00000011297	ENSFCAT00000003906		
<i>Gallus gallus</i>	n/a	ENSGALT00000037556	ENSGALT00000015429	ENSGALT00000037965	ENSGALT00000019443	ENSGALT00000007244
<i>Gasterosteus aculeatus</i>	n/a	ENSGACT00000003581	ENSGACT00000013114	ENSGACT00000021502	ENSGACT00000011157	ENSGACT00000022651
<i>Geochelone sulcata</i>	MVZ238888		*			
<i>Geomyda spengleri</i>	MVZ257081					
<i>Geotrypetes seraphini</i>	MVZ252475					
<i>Gorilla gorilla</i>	n/a	ENSGGOT00000006323	ENSGGOT00000008660	ENSGGOT00000007151	ENSGGOT00000010120	ENSGGOT00000006447
<i>Hardella thurjii</i>	MVZ248392		*			
<i>Heloderma horridum</i>	MVZ238489					
<i>Heteronotia binoei</i>	MKF226	*	*			
<i>Homo sapiens</i>	n/a	ENST00000392990	ENST00000475915	ENST00000396241	ENST00000325658	ENST00000354300
<i>Hynobius arisanensis</i>	MVZ251776					
<i>Icthyophis bannanicus</i>	MVZ226265		*			
<i>Indotestudo elongata</i>	MVZ234627					
<i>Kachuga smithii</i>	MVZ248394		*			
<i>Kinixys belliana</i>	MVZ234631					
<i>Kinosternon scorpioides</i>	MVZ247584					
<i>Leptotyphlops dulcis</i>	MVZ236621					
<i>Leuccephalon yuwonoi</i>	MVZ241499					
<i>Lissemys punctata</i>	MVZ248398					
<i>Loxodonta africana</i>	n/a	ENSLAFT00000000946	ENSLAFT00000006821	ENSLAFT00000003198		ENSLAFT00000006186
<i>Macaca mulatta</i>	n/a	ENSMUT000000026163	ENSMUT000000021997	ENSMUT00000019661	ENSMUT00000014636	ENSMUT00000018458
<i>Macropus eugenii</i>	n/a	ENSMEUT00000007865	ENSMEUT00000015795	ENSMEUT00000014237	ENSMEUT00000008920	ENSMEUT00000003116
<i>Malaclemys terrapin</i>	MVZ257705	*	*	*	*	*
<i>Malacochersus tornieri</i>	MVZ234632					
<i>Malayemys subtrijuga</i>	MVZ230474		*			
<i>Mauremys caspica</i>	MVZ237439					
<i>Mauremys mutica</i>	MVZ230487					
<i>Mesalina brevirostris</i>	MVZ246016					
<i>Microcebus murinus</i>	n/a	ENSMICT00000010353	ENSMICT00000007003	ENSMICT00000015183		ENSMICT00000004944
<i>Monodelphis domestica</i>	n/a	ENSMODT00000016424	ENSMODT00000036561	ENSMODT00000026361	ENSMODT00000010930	ENSMODT00000014729
<i>Mus musculus</i>	n/a	ENSMUST00000101063	ENSMUST00000029424	ENSMUST00000031167	ENSMUST00000021514	ENSMUST00000119015
<i>Myotis lucifugus</i>	n/a		ENSMLUT00000016981	ENSMLUT00000016759		ENSMLUT00000015566
<i>Nectarinia olivacea</i>	JF1603					
<i>Nerodia rhombifer</i>	MVZ246092					
<i>Ochotona princeps</i>	n/a	ENSOPRT00000007644	ENSOPRT00000011048	ENSOPRT00000008133	ENSOPRT00000010103	ENSOPRT0000001047
<i>Ornithorhynchus anatinus</i>	n/a	ENSOANT000000021786	ENSOANT00000017424	ENSOANT00000017764	ENSOANT00000018794	ENSOANT00000007571
<i>Oryzotylagus cuniculus</i>	n/a		ENSOCUT00000015411	ENSOCUT00000016363	ENSOCUT00000009955	ENSOCUT00000011936
<i>Oryzias latipes</i>	n/a	ENSORLT00000001962		ENSORLT00000025493	ENSORLT00000019754	ENSORLT00000011363
<i>Oscacilia ochrocephala</i>	MVZ232864		*			
<i>Otolemur garnettii</i>	n/a	ENSOGAT00000004618	ENSOGAT00000013890	ENSOGAT00000016009	ENSOGAT00000003691	ENSOGAT00000006998
<i>Pan troglodytes</i>	n/a	ENSPTRT00000012343	ENSPTRT00000029089	ENSPTRT000000064227	ENSPTRT00000011735	ENSPTRT000000049582

Scientific Name	Specimen #	PAPOLA	PDCD10	POLR2B	PPM1A	PPTC7
<i>Pelodiscus sinensis</i>	MVZ226472					
<i>Pelomedusa subrufa</i>	MVZ236628					
<i>Phrynosoma sp.</i>	MVZ247578					
<i>Pipa pipa</i>	MVZ247516					
<i>Platenys platycephala</i>	MVZ247579					
<i>Platysternon megacephalum</i>	MVZ230486					
<i>Plestiodon gilberti</i>	MVZ249915					
<i>Pongo pygmaeus</i>	n/a	ENSPPYT00000007241		ENSPPYT00000017139	ENSPPYT00000006935	ENSPPYT00000005867
<i>Procarina capensis</i>	n/a	ENSPCAT00000004242	ENSPCAT00000010346	ENSPCAT00000001623	ENSPCAT00000007683	ENSPCAT00000012130
<i>Pseudomys concinna</i>	MVZ241520					
<i>Pteropus vampyrus</i>	n/a	ENSPVAT00000013926	ENSPVAT00000014524	ENSPVAT00000010952	ENSPVAT00000002783	ENSPVAT00000003512
<i>Rana livida</i>	MVZ241436					
<i>Rattus norvegicus</i>	n/a	ENSRNOT00000006982	ENSRNOT00000013585	ENSRNOT00000039252	ENSRNOT00000008238	ENSRNOT000000036137
<i>Rhinerea floridana</i>	MVZ233342					
<i>Rhinoclemmys pulcherrima</i>	MVZ230924					
<i>Sacalia quadriocellata</i>	MVZ230485					
<i>Sorex araneus</i>	n/a	ENSSART00000013812	ENSSART00000000631	ENSSART00000001158	ENSSART00000000390	ENSSART00000000556
<i>Spermophilus tridecemlineatus</i>	n/a	ENSSTOT00000011484		ENSSTOT00000003855	ENSSTOT00000014584	ENSSTOT00000012014
<i>Sternotherus carinatus</i>	MVZ238123		*			
<i>Sternotherus minor</i>	MVZ241528		*			
<i>Stigmochelys pardalis</i>	MVZ241333		*			
<i>Sus scrofa</i>	n/a	ENSSSCT00000002783	ENSSSCT00000012850		ENSSSCT00000005604	ENSSSCT00000010768
<i>Sylvietta virens</i>	JF1484					
<i>Taeniopygia guttata</i>	n/a	ENSTGUT00000013221	ENSTGUT00000011549	ENSTGUT00000003066	ENSTGUT00000013537	ENSTGUT00000006637
<i>Takifugu rubripes</i>	n/a	ENSTRUT00000015756	ENSTRUT00000040743	ENSTRUT00000021737	ENSTRUT00000026595	ENSTRUT00000016082
<i>Tarsius syrichta</i>	n/a	ENSTSYT00000012389	ENSTSYT00000000708	ENSTSYT00000006892	ENSTSYT00000003951	ENSTSYT00000012871
<i>Terpsiphone batesi</i>	JF1488					
<i>Terrapene carolina</i>	MVZ241538					
<i>Testudo graeca</i>	MVZ234282					
<i>Testudo hermanni</i>	MVZ244866		*	*		*
<i>Tetraodon nigroviridis</i>	n/a	ENSTNIT00000020472	ENSTNIT00000020563	ENSTNIT00000007718	ENSTNIT00000020718	ENSTNIT00000012152
<i>Trachemys scripta</i>	MVZ241524					
<i>Tupaia belangeri</i>	n/a	ENSTBET00000000195	ENSTBET00000000261	ENSTBET00000014621	ENSTBET00000016824	ENSTBET00000013543
<i>Tursiops truncatus</i>	n/a		ENSTTRT00000016315	ENSTTRT00000010840	ENSTTRT00000008935	ENSTTRT00000015095
<i>Varamus yemenensis</i>	MVZ236610					
<i>Vicugna pacos</i>	n/a	ENSVPAT00000006464	ENSVPAT00000011655	ENSVPAT00000000469	ENSVPAT00000007553	ENSVPAT00000008409
<i>Xantusia vigilis</i>	MVZ236373					
<i>Xenopeltis unicolor</i>	MVZ239387		*			
<i>Xenopus tropicalis</i>	n/a	ENSXETT00000030476	ENSXETT00000049474	ENSXETT00000011733	ENSXETT00000036927	ENSXETT00000045244

Scientific Name	Specimen #	PSAT1	PSMA6	PSMC6	PSMD14	RAB1A
<i>Accipiter cooperii</i>	MVZ181729					
<i>Agelaiocercus kingi</i>	MVZ183274			*		
<i>Agriomys borsfieldii</i>	MVZ245929					
<i>Aluthe castanea</i>	JF1706		*			*
<i>Alligator mississippiensis</i>	n/a	ES316685	ES608			
<i>Alligator mississippiensis</i>	MVZ241485		*	*	*	
<i>Anas cyanoptera</i>	MVZ179315					
<i>Anniella pulchra</i>	MVZ247492		*		*	*
<i>Anolis carolinensis</i>	n/a	ENSACAT00000006128	ENSACAT00000001856	ENSACAT00000006886	ENSACAT00000011931	ENSACAT00000001844
<i>Apalone ferox</i>	MVZ241535		*			
<i>Aptenodytes patagonicus</i>	MVZ178921					
<i>Aspidoscelis tigris</i>	MVZ243345					
<i>Bipes biporus</i>	MVZ236260					
<i>Bos taurus</i>	n/a	ENSBAT00000018554	ENSBAT00000012773	ENSBAT00000020886	ENSBAT00000046166	ENSBAT00000044736
<i>Callithrix jacchus</i>	n/a	ENSCJAT00000006370	ENSCJAT00000018417	ENSCJAT00000013991	ENSCJAT00000010797	ENSCJAT00000013205
<i>Canis familiaris</i>	n/a	ENSACAT00000002468	ENSACAT000000021278	ENSACAT000000023514	ENSACAT00000015861	ENSACAT00000005128
<i>Casuarinus casuarinus</i>	MVZ178927					
<i>Cavia prellus</i>	n/a	ENSCPOT00000013888	ENSCPOT00000003150	ENSCPOT00000009563	ENSCPOT00000005942	ENSCPOT00000015588
<i>Chamaeleo senegalensis</i>	MVZ249659					
<i>Chelodina longicollis</i>	MVZ230488					
<i>Chelonia mydas</i>	MVZ233240					
<i>Chelonoidis denticulata</i>	MVZ175391		*			
<i>Chelydra serpentina</i>	MVZ250691		*			
<i>Choloepus hoffmanni</i>	n/a	ENSCHOT00000003832	ENSCHOT00000002862	ENSCHOT00000003337	ENSCHOT00000002567	ENSCHOT00000005084
<i>Chrysomys picta</i>	MVZ241506		*	*		*
<i>Claudius angustatus</i>	MVZ164771					
<i>Cuora galbinifrons</i>	MVZ230933					
<i>Cuora pani</i>	MVZ230513					
<i>Cyrtodactylus</i>	JAM8710	*	*			*
<i>Danio rerio</i>	n/a		ENDSART00000012859	ENDSART00000053834	ENDSART00000091901	ENDSART00000043002
<i>Dasyurus novemcinctus</i>	n/a		ENDSNOT00000003768	ENDSNOT00000002661	ENDSNOT00000009795	ENDSNOT00000018236
<i>Dendropicus gabonensis</i>	JF1513		*		*	
<i>Dermatemys mawii</i>	MVZ233618					
<i>Dermochelys coriacea</i>	MVZ149844					
<i>Dibamus sp.</i>	MVZ224113					
<i>Dierochelys reticularia</i>	MVZ230923		*			
<i>Dipodomys ordii</i>	n/a	ENDSORT00000016698	ENDSORT00000010060	ENDSORT00000016408	ENDSORT00000011046	ENDSORT00000002112
<i>Draco beccarii</i>	JAM8961	*				
<i>Draco sumatranus</i>	SZ1.71	*				
<i>Echinops telfairi</i>	n/a	ENSETET00000004156	ENSETET00000004891		ENSETET00000017961	ENSETET00000002202
<i>Emydura subglobosa</i>	MVZ230496					
<i>Emys marmorata</i>	MVZ164996					
<i>Ensatina escholtzii</i>	RP158					
<i>Equus caballus</i>	n/a	ENSECAT00000019292	ENSECAT00000023920	ENSECAT00000019144	ENSECAT00000021497	ENSECAT00000010149

Scientific Name	Specimen #	PSAT1	PSMA6	PSMC6	PSMD14	RAB1A
<i>Ereimochelys imbricata</i>	MVZ233242		*			
<i>Erinaceus europaeus</i>	n/a	ENSEEUT00000008468	ENSEEUT00000010610	ENSEEUT00000002738	ENSEEUT00000006460	ENSEEUT00000011705
<i>Fagelaryia limnocharis</i>	MVZ251769					
<i>Felis catus</i>	n/a	ENSFCAT00000007766		ENSFCAT00000012979		ENSFCAT00000002536
<i>Gallus gallus</i>	n/a	ENSGALT00000024496	ENSGALT00000016353	ENSGALT00000037197	ENSGALT00000018142	ENSGALT00000014281
<i>Gasterosteus aculeatus</i>	n/a	ENSGACT00000006399	ENSGACT00000016600	ENSGACT00000007662	ENSGACT00000020049	ENSGACT00000005122
<i>Geochelone sulcata</i>	MVZ238888		*			
<i>Geomyda spengleri</i>	MVZ257081		*			
<i>Geotrypetes seraphini</i>	MVZ252475					
<i>Gorilla gorilla</i>	n/a	ENSGGOT00000015535	ENSGGOT00000013251	ENSGGOT00000016401	ENSGGOT00000015800	
<i>Hardella thurjii</i>	MVZ248392		*			
<i>Heloderma horridum</i>	MVZ238489					
<i>Heteronotia binoei</i>	MKF226	*	*			
<i>Homo sapiens</i>	n/a	ENST00000347159	ENST00000261479	ENST00000216407	ENST00000263639	ENST00000260638
<i>Hynobius arisanensis</i>	MVZ251776					
<i>Icthyophis bannanicus</i>	MVZ226265		*			*
<i>Indotestudo elongata</i>	MVZ234627					
<i>Kachuga smithii</i>	MVZ248394		*			
<i>Kinixys belliana</i>	MVZ234631					
<i>Kinosternon scorpioides</i>	MVZ247584					
<i>Leptotyphlops dulcis</i>	MVZ236621					
<i>Leuccephalon yuwonoi</i>	MVZ241499		*			
<i>Lissemys punctata</i>	MVZ248398					
<i>Loxodonta africana</i>	n/a	ENSLAFT00000017916	ENSLAFT00000017359	ENSLAFT00000014825	ENSLAFT00000004620	ENSLAFT00000016664
<i>Macaca mulatta</i>	n/a	ENSM MUT00000016630	ENSM MUT00000017383	ENSM MUT00000039843	ENSM MUT00000016683	ENSM MUT00000030563
<i>Macropus eugenii</i>	n/a	ENSMEUT00000007805	ENSMEUT00000016583	ENSMEUT00000007110	ENSMEUT00000010015	ENSMEUT00000000598
<i>Malaclemys terrapin</i>	MVZ257705	*	*	*	*	*
<i>Malacochersus tornieri</i>	MVZ234632					
<i>Malayemys subtrijuga</i>	MVZ230474		*			
<i>Mauremys caspica</i>	MVZ237439					
<i>Mauremys mutica</i>	MVZ230487		*			
<i>Mesalina brevirostris</i>	MVZ246016					
<i>Microcebus murinus</i>	n/a	ENSMICT00000003009	ENSMICT00000000082			ENSMICT00000012587
<i>Monodelphis domestica</i>	n/a	ENSMODT00000016516	ENSMODT00000016318	ENSMODT00000018452	ENSMODT00000006717	ENSMODT00000004174
<i>Mus musculus</i>	n/a	ENSMUST00000025542	ENSMUST00000021412	ENSMUST00000022380	ENSMUST00000028278	ENSMUST00000007471
<i>Myotis lucifugus</i>	n/a	ENSMLUT00000010270	ENSMLUT00000010220	ENSMLUT00000009192	ENSMLUT00000006003	ENSMLUT000000009479
<i>Nectarinia olivacea</i>	JF1603					
<i>Nerodia rhombifer</i>	MVZ246092					
<i>Ochotona princeps</i>	n/a		ENSOPRT00000014572	ENSOPRT00000014549	ENSOPRT00000005119	ENSOPRT000000006593
<i>Ornithorynchus anatinus</i>	n/a	ENSOANT00000004707	ENSOANT00000022172	ENSOANT00000017157	ENSOANT000000009874	ENSOANT000000030441
<i>Oryzotylagus cuniculus</i>	n/a	ENSOCUT00000007763	ENSOCUT00000015631	ENSOCUT00000012638	ENSOCUT00000001418	ENSOCUT00000016999
<i>Oryzias latipes</i>	n/a			ENSORLT00000019327	ENSORLT00000004837	ENSORLT00000001168
<i>Oscacilia ochrocephala</i>	MVZ232864		*			
<i>Otolemur garnettii</i>	n/a	ENSOGAT00000003048	ENSOGAT00000008496	ENSOGAT00000010014	ENSOGAT00000004936	
<i>Pan troglodytes</i>	n/a	ENSPTRT000000038918	ENSPTRT00000011490	ENSPTRT00000011629	ENSPTRT00000023304	ENSPTRT000000022309

Scientific Name	Specimen #	PSAT1	PSMA6	PSMC6	PSMD14	RAB1A
<i>Pelodiscus sinensis</i>	MVZ226472		*			
<i>Pelomedusa subrufa</i>	MVZ236628		*			
<i>Phrynosoma sp.</i>	MVZ247578					
<i>Pipa pipa</i>	MVZ247516					
<i>Platemys platycephala</i>	MVZ247579					
<i>Platysternon megacephalum</i>	MVZ230486					
<i>Plestiodon gilberti</i>	MVZ249915	*	*			
<i>Pongo pygmaeus</i>	n/a	ENSPPYT00000022521		ENSPPYT00000006875	ENSPPYT00000014960	ENSPPYT00000014342
<i>Procarina capensis</i>	n/a	ENSPCAT000000013192	ENSPCAT00000006480	ENSPCAT00000000450	ENSPCAT00000001942	ENSPCAT00000002440
<i>Pseudemys concinna</i>	MVZ241520		*			
<i>Pteropus vampyrus</i>	n/a	ENSPVAT00000009719	ENSPVAT00000014168	ENSPVAT00000001320	ENSPVAT00000005414	
<i>Rana livida</i>	MVZ241436					
<i>Rattus norvegicus</i>	n/a	ENSRNOT000000065063	ENSRNOT00000009666	ENSRNOT00000009649	ENSRNOT00000006605	ENSRNOT000000020209
<i>Rhinoceros floridana</i>	MVZ233342					
<i>Rhinoclemmys pulcherrima</i>	MVZ230924		*			
<i>Sacalia quadriocellata</i>	MVZ230485					
<i>Sorex araneus</i>	n/a	ENSSART000000011156	ENSSART00000007759	ENSSART00000002577	ENSSART00000009511	ENSSART00000000376
<i>Spermophilus tridecemlineatus</i>	n/a	ENSSTOT00000013629	ENSSTOT00000011773	ENSSTOT00000014492	ENSSTOT00000014892	ENSSTOT00000011485
<i>Sternotherus carinatus</i>	MVZ238123					
<i>Sternotherus minor</i>	MVZ241528		*	*		
<i>Stigmochelys pardalis</i>	MVZ241333		*			
<i>Sus scrofa</i>	n/a	ENSSSCT00000005824	ENSSSCT00000002183	ENSSSCT00000005554	ENSSSCT00000017304	ENSSSCT00000009153
<i>Sylvietta virens</i>	JF1484		*			
<i>Taeniopygia guttata</i>	n/a	ENSTGUT00000003535	ENSTGUT00000012478	ENSTGUT00000013650	ENSTGUT00000006958	ENSTGUT00000002422
<i>Takifugu rubripes</i>	n/a		ENSTRUT00000012572	ENSTRUT00000023378	ENSTRUT000000044090	ENSTRUT00000000963
<i>Tarsius syrichta</i>	n/a	ENSTSYT000000000092	ENSTSYT000000000215	ENSTSYT00000010135	ENSTSYT000000004734	ENSTSYT000000006012
<i>Terpsiphone batesi</i>	JF1488					
<i>Terrapene carolina</i>	MVZ241538					
<i>Testudo graeca</i>	MVZ234282		*			
<i>Testudo hermanni</i>	MVZ244866		*	*	*	
<i>Tetraodon nigroviridis</i>	n/a	ENSTNIT00000019270	ENSTNIT00000012781	ENSTNIT00000008088	ENSTNIT00000012541	
<i>Trachemys scripta</i>	MVZ241524					
<i>Tupaia belangeri</i>	n/a	ENSTBET00000012945	ENSTBET00000011248		ENSTBET00000014121	ENSTBET000000003316
<i>Tursiops truncatus</i>	n/a	ENSTTRT00000013911	ENSTTRT00000007968	ENSTTRT00000003882	ENSTTRT00000006632	ENSTTRT000000008280
<i>Varamus yemenensis</i>	MVZ236610					
<i>Vicugna pacos</i>	n/a	ENSVPAT00000004044	ENSVPAT00000005690	ENSVPAT00000012227	ENSVPAT00000002718	ENSVPAT00000005639
<i>Xantusia vigilis</i>	MVZ236373					
<i>Xenopeltis unicolor</i>	MVZ239387			*		
<i>Xenopus tropicalis</i>	n/a	ENSXETT00000007823	ENSXETT000000054899	ENSXETT00000018862	ENSXETT000000035366	ENSXETT000000030593

Scientific Name	Specimen #	RAN	RPL4	RPL7A	RPL11	RPL38
<i>Accipiter cooperii</i>	MVZ181729					*
<i>Agelaiocercus kingi</i>	MVZ183274	*				*
<i>Agriomys borsfieldii</i>	MVZ245929		*	*		*
<i>Aluthe castanea</i>	JF1706	*		*		*
<i>Alligator mississippiensis</i>	n/a		ES316508	ES583	ES5078	ES319935
<i>Alligator mississippiensis</i>	MVZ241485	*	*	*	*	*
<i>Anas cyanoptera</i>	MVZ179315					*
<i>Anniella pulchra</i>	MVZ247492	*	*		*	*
<i>Anolis carolinensis</i>	n/a	ENSACAT00000007195	ENSACAT00000009829	ENSACAT00000015040	ENSACAT00000005076	ENSACAT00000014866
<i>Apalone ferox</i>	MVZ241535		*	*		*
<i>Aptenodytes patagonicus</i>	MVZ178921					*
<i>Aspidoscelis tigris</i>	MVZ243345				*	*
<i>Bipes biporus</i>	MVZ236260					*
<i>Bos taurus</i>	n/a	ENSBTAT00000017722	ENSBTAT00000021206	ENSBTAT00000015358	ENSBTAT00000027850	ENSBTAT00000042547
<i>Callithrix jacchus</i>	n/a	ENSCJAT00000007631	ENSCJAT00000014706	ENSCJAT00000028865	ENSCJAT00000015155	ENSCJAT00000029222
<i>Canis familiaris</i>	n/a	ENSACAT000000010749	ENSACAT00000008039	ENSACAT00000028043	ENSACAT00000035556	ENSACAT00000038274
<i>Casuarinus casuarinus</i>	MVZ178927					*
<i>Cavia prellus</i>	n/a		ENSCPOT00000003096	ENSCPOT0000001343	ENSCPOT00000008509	ENSCPOT00000009070
<i>Chamaelo senegalensis</i>	MVZ249659				*	
<i>Chelodina longicollis</i>	MVZ230488			*	*	
<i>Chelonia mydas</i>	MVZ233240			*		*
<i>Chelonoidis denticulata</i>	MVZ175391					*
<i>Chelydra serpentina</i>	MVZ250691		*	*	*	*
<i>Choloepus hoffmanni</i>	n/a	ENSCHOT00000011705	*			
<i>Chrysomys picta</i>	MVZ241506	*	*	*	*	*
<i>Claudius angustatus</i>	MVZ164771	*	*			*
<i>Cnora galbiniifrons</i>	MVZ230933					*
<i>Cuora pani</i>	MVZ230513					
<i>Cyrtodactylus</i>	JAM8710	*	*	*	*	*
<i>Danio rerio</i>	n/a		ENDSART000000060363	ENDSART00000004238	ENDSART000000063870	ENDSART00000021069
<i>Dasyatis novemcinctus</i>	n/a	ENDSNOT00000009227	ENDSNOT00000016287	ENDSNOT00000024638	ENDSNOT00000011222	ENDSNOT00000004818
<i>Dendropicus gabonensis</i>	JF1513	*	*	*	*	*
<i>Dermatemys mawii</i>	MVZ233618	*				*
<i>Dermochelys coriacea</i>	MVZ149844					*
<i>Dibamus sp.</i>	MVZ224113	*	*		*	*
<i>Dierochelys reticularia</i>	MVZ230923	*	*	*		*
<i>Dipodomys ordii</i>	n/a		*	ENDSNOT00000006470	ENDSNOT00000012495	ENDSNOT00000013145
<i>Draco beccarii</i>	JAM8961		*	*		*
<i>Draco sumatranus</i>	SZL71			*	*	*
<i>Echinops telfairi</i>	n/a	ENSETET00000011873	*	ENSETET00000019555	ENSETET00000014824	ENSETET00000001872
<i>Emydura subglobosa</i>	MVZ230496					
<i>Emys marmorata</i>	MVZ164996		*			
<i>Ensatina eschscholtzii</i>	RP158		*			
<i>Equus caballus</i>	n/a	ENSECAT00000021906	ENSECAT00000025071	ENSECAT00000008856	ENSECAT00000022577	ENSECAT00000019668

Scientific Name	Specimen #	RAN	RPL4	RPL7A	RPL11	RPL38
<i>Ereimochelys imbricata</i>	MVZ233242	*			*	*
<i>Erinaceus europaeus</i>	n/a			ENSEEUT0000000203	ENSEEUT00000011663	
<i>Fejervarya limncharis</i>	MVZ251769		*		*	*
<i>Felis catus</i>	n/a		*	ENSFCAT00000005827		ENSFCAT00000010853
<i>Gallus gallus</i>	n/a	ENSGALT00000004055	ENSGALT00000038315	ENSGALT00000005066	ENSGALT00000040531	ENSGALT00000002220
<i>Gasterosteus aculeatus</i>	n/a		ENSGACT00000006665	ENSGACT00000021129	ENSGACT00000002790	ENSGACT00000004327
<i>Geochelone sulcata</i>	MVZ238888	*	*	*	*	*
<i>Geomyda spengleri</i>	MVZ257081	*	*		*	*
<i>Geotrypetes seraphini</i>	MVZ252475					
<i>Gorilla gorilla</i>	n/a	ENSGGOT00000015394	ENSGGOT00000003719	ENSGGOT00000014494	ENSGGOT00000014678	ENSGGOT00000000799
<i>Hardella thurjii</i>	MVZ248392	*	*	*		*
<i>Heloderma horridum</i>	MVZ238489					
<i>Heteronotia binoei</i>	MKF226	*	*		*	*
<i>Homo sapiens</i>	n/a	ENST00000392369	ENST00000432669	ENST00000277425	ENST00000458455	ENST00000439590
<i>Hynobius arisanensis</i>	MVZ251776				*	*
<i>Icthyophis bannanicus</i>	MVZ226265	*	*	*	*	*
<i>Indotestudo elongata</i>	MVZ234627					
<i>Kachuga smithii</i>	MVZ248394		*	*		*
<i>Kinixys belliana</i>	MVZ234631			*		*
<i>Kinosternon scorpioides</i>	MVZ247584	*	*			
<i>Leptotyphlops dulcis</i>	MVZ236621		*			*
<i>Leucocephalon yuwonoi</i>	MVZ241499				*	
<i>Lissemys punctata</i>	MVZ248398					*
<i>Loxodonta africana</i>	n/a	ENSLAFT00000017669	ENSLAFT00000006497	ENSLAFT00000006436		ENSLAFT00000000352
<i>Macaca mulatta</i>	n/a	ENSMMUT00000029433	ENSMMUT00000023666	ENSMMUT00000043520	ENSMMUT00000003486	ENSMMUT00000028954
<i>Macropus eugenii</i>	n/a	ENSMEUT00000004588	ENSMEUT00000004005	ENSMEUT00000014590	ENSMEUT00000013776	ENSMEUT00000015957
<i>Malaclemys terrapin</i>	MVZ257705	*	*	*		*
<i>Malacochersus tornieri</i>	MVZ234632					
<i>Malayemys subtrijuga</i>	MVZ230474	*	*	*		*
<i>Mauremys caspica</i>	MVZ237439		*			
<i>Mauremys mutica</i>	MVZ230487	*	*			*
<i>Mesalina brevirostris</i>	MVZ246016					*
<i>Microcebus murinus</i>	n/a	ENSMICT00000000652	ENSMICT000000004420	ENSMICT000000009944	ENSMICT000000009572	ENSMICT000000008484
<i>Monodelphis domestica</i>	n/a	ENSMODT00000019816	ENSMODT00000036322	ENSMODT00000016254	ENSMODT00000016346	ENSMODT000000008449
<i>Mus musculus</i>	n/a	ENSMUST00000111343	ENSMUST00000034966	ENSMUST00000102898	ENSMUST00000078981	ENSMUST00000082092
<i>Myotis lucifugus</i>	n/a		ENSMLUT00000011807	ENSMLUT00000004668	ENSMLUT00000008868	ENSMLUT00000014453
<i>Nectarinia olivacea</i>	JF1603					*
<i>Nerodia rhombifer</i>	MVZ246092		*			
<i>Ochotona princeps</i>	n/a	ENSOPRT00000002617	ENSOPRT00000007791	ENSOPRT00000005714	ENSOPRT00000014975	ENSOPRT00000011884
<i>Ornithorhynchus anatinus</i>	n/a	ENSOANT00000010295	ENSOANT00000023550	ENSOANT00000011240	ENSOANT00000000540	
<i>Oryzotylagus cuniculus</i>	n/a		ENSOCUT00000008791	ENSOCUT00000007822		
<i>Oryzias latipes</i>	n/a		ENSORLRT000000008646	ENSORLRT00000004955	ENSORLRT000000009500	ENSORLRT00000013827
<i>Oscacilia ochrocephala</i>	MVZ232864	*	*		*	
<i>Otolemur garnettii</i>	n/a		ENSOGAT00000015885	ENSOGAT00000014031		ENSOGAT00000000002
<i>Pan troglodytes</i>	n/a	ENSPTRT000000067505	ENSPTRT00000013280	ENSPTRT00000039851	ENSPTRT00000000659	ENSPTRT00000049947

Scientific Name	Specimen #	RAN	RPL4	RPL7A	RPL11	RPL38
<i>Pelodiscus sinensis</i>	MVZ226472	*	*			*
<i>Pelomedusa subrufa</i>	MVZ236628			*	*	*
<i>Phrynos sp.</i>	MVZ247578					
<i>Pipa pipa</i>	MVZ247516				*	*
<i>Platemys platycephala</i>	MVZ247579					
<i>Platysternon megacephalum</i>	MVZ230486					*
<i>Plestiodon gilberti</i>	MVZ249915	*				*
<i>Pongo pygmaeus</i>	n/a	ENSPPYT00000006065	ENSPPYT00000016414	ENSPPYT00000023036	ENSPPYT00000002072	ENSPPYT00000010034
<i>Procarina capensis</i>	n/a		ENSPCAT00000000622	ENSPCAT00000002830	ENSPCAT00000004457	ENSPCAT00000009744
<i>Pseudemys concinna</i>	MVZ241520		*	*	*	*
<i>Pteropus vampyrus</i>	n/a	ENSPVAT00000014742	ENSPVAT00000000067	ENSPVAT00000017044	ENSPVAT00000004282	ENSPVAT00000017521
<i>Rana livida</i>	MVZ241436	*				
<i>Rattus norvegicus</i>	n/a	ENSRNOT00000001247	ENSRNOT00000013461	ENSRNOT000000042119	ENSRNOT00000039585	ENSRNOT00000048963
<i>Rhineura floridana</i>	MVZ233342					*
<i>Rhinoclemmys pulcherrima</i>	MVZ230924	*	*			*
<i>Sacalia quadriocellata</i>	MVZ230485					*
<i>Sorex araneus</i>	n/a			ENSSART00000010506	ENSSART00000010893	
<i>Spermophilus tridecemlineatus</i>	n/a	ENSSTOT00000003979	ENSSTOT000000009218	ENSSTOT00000003674	ENSSTOT00000011900	
<i>Sternotherus carinatus</i>	MVZ238123		*	*		*
<i>Sternotherus minor</i>	MVZ241528	*	*	*		*
<i>Stigmochelys pardalis</i>	MVZ241333		*	*		*
<i>Sus scrofa</i>	n/a	ENSSSCT00000010685	ENSSSCT00000005457	ENSSSCT00000005122		ENSSSCT00000018768
<i>Sylvietta virens</i>	JF1484	*				*
<i>Taeniopygia guttata</i>	n/a	ENSTGUT00000004882	ENSTGUT00000009825	ENSTGUT00000005689	ENSTGUT00000001261	
<i>Takifugu rubripes</i>	n/a		ENSTRUT000000037609	ENSTRUT000000047235	ENSTRUT00000011593	ENSTRUT000000046202
<i>Tarsius syrichta</i>	n/a		ENSTSYT00000005581	ENSTSYT00000005458	ENSTSYT00000001349	ENSTSYT000000004591
<i>Terpsiphone batesi</i>	JF1488		*			*
<i>Terrapene carolina</i>	MVZ241538					*
<i>Testudo graeca</i>	MVZ234282	*	*		*	*
<i>Testudo hermanni</i>	MVZ244866	*	*	*	*	*
<i>Tetraodon nigroriviridis</i>	n/a		ENSTNIT00000011445	ENSTNIT00000010491	ENSTNIT00000005315	ENSTNIT00000017691
<i>Trachemys scripta</i>	MVZ241524					*
<i>Tupaia belangeri</i>	n/a		ENSTBET00000014390	ENSTBET00000008599	ENSTBET00000007830	ENSTBET00000007391
<i>Tursiops truncatus</i>	n/a	ENSTTRT00000012732	ENSTTRT00000010203	ENSTTRT00000007636	ENSTTRT00000016113	ENSTTRT00000007499
<i>Varanus yemenensis</i>	MVZ236610					*
<i>Vicugna pacos</i>	n/a		ENSVPAT00000006746			ENSVPAT00000001958
<i>Xantusia vigilis</i>	MVZ236373					*
<i>Xenopeltis unicolor</i>	MVZ239387	*	*			
<i>Xenopus tropicalis</i>	n/a		ENSXETT00000034688	ENSXETT00000010943	ENSXETT00000002557	ENSXETT00000054097

Scientific Name	Specimen #	RPLP0	RPS4X	RPS6	RPS14	SDHB
<i>Accipiter cooperii</i>	MVZ181729					
<i>Agelaius kingi</i>	MVZ183274				*	*
<i>Agrionemys borsfieldii</i>	MVZ245929					
<i>Aluthe castanea</i>	JF1706				*	*
<i>Alligator mississippiensis</i>	n/a	ES2340	ES321882	ES321879	ES321806	ES321806
<i>Alligator mississippiensis</i>	MVZ241485				*	*
<i>Anas cyanoptera</i>	MVZ179315					
<i>Anniella pulchra</i>	MVZ247492				*	*
<i>Anolis carolinensis</i>	n/a	ENSACAT00000003794	ENSACAT00000014647	ENSACAT00000017706	ENSACAT00000015018	ENSACAT00000015018
<i>Apalone ferox</i>	MVZ241535				*	*
<i>Aptenodytes patagonicus</i>	MVZ178921					
<i>Aspidoscelis tigris</i>	MVZ243345				*	*
<i>Bipes biporus</i>	MVZ236260				*	*
<i>Bos taurus</i>	n/a	ENSBTAT00000055086	ENSBTAT00000025024	ENSBTAT00000019214	ENSBTAT00000011305	ENSBTAT00000011305
<i>Callithrix jacchus</i>	n/a	ENSCJAT00000017139	ENSCJAT00000043827	ENSCJAT00000040479	ENSCJAT00000041231	ENSCJAT00000041231
<i>Canis familiaris</i>	n/a	ENSACAT00000016229	ENSACAT00000027164	ENSACAT00000002536	ENSACAT00000028750	ENSACAT00000028750
<i>Casuarinus casuarinus</i>	MVZ178927					
<i>Cavia prellus</i>	n/a	ENSCPOT00000019555	ENSCPOT00000024331	ENSCPOT00000003794	ENSCPOT00000020708	ENSCPOT00000020708
<i>Chamaeleo senegalensis</i>	MVZ249659				*	*
<i>Chelodina longicollis</i>	MVZ230488					
<i>Chelonia mydas</i>	MVZ233240				*	*
<i>Chelonoidis denticulata</i>	MVZ175391				*	*
<i>Chelydra serpentina</i>	MVZ250691				*	*
<i>Choloepus hoffmanni</i>	n/a	ENSCHOT00000002758	ENSCHOT00000002501	ENSCHOT00000004638	ENSCHOT00000002359	ENSCHOT00000002359
<i>Chrysomys picta</i>	MVZ241506				*	*
<i>Claudius angustatus</i>	MVZ164771				*	*
<i>Cuora galbiniifrons</i>	MVZ230933					
<i>Cuora pani</i>	MVZ230513				*	*
<i>Cyrtodactylus</i>	JAM8710		*	*	*	*
<i>Danio rerio</i>	n/a	ENDART00000073462	ENDART00000007957	ENDART00000023336	ENDART00000053208	ENDART00000053208
<i>Dasyatis novemcinctus</i>	n/a			ENDNOT00000007237	ENDNOT00000009930	ENDNOT00000009930
<i>Dendropicos gabonensis</i>	JF1513	*			*	*
<i>Dermatemys mawii</i>	MVZ233618				*	*
<i>Dermochelys coriacea</i>	MVZ149844				*	*
<i>Dibamus sp.</i>	MVZ224113				*	*
<i>Dierochelys reticulata</i>	MVZ230923				*	*
<i>Dipodomys ordii</i>	n/a	ENDDORT00000008588	ENDDORT00000008569	ENDDORT00000012057	ENDDORT00000007145	ENDDORT00000007145
<i>Draco beccarii</i>	JAM8961		*		*	*
<i>Draco sumatranus</i>	SZL71		*		*	*
<i>Echinops telfairi</i>	n/a			ENSETET00000001583	ENSETET00000002899	ENSETET00000002899
<i>Emydura subglobosa</i>	MVZ230496				*	*
<i>Emys marmorata</i>	MVZ164996				*	*
<i>Ensatina eschscholtzii</i>	RP158					
<i>Equus caballus</i>	n/a	ENSECAT00000004097	ENSECAT00000011336	ENSECAT00000016957	ENSECAT00000026184	ENSECAT00000026184

Scientific Name	Specimen #	RPLP0	RPS4X	RPS6	RPS14	SDHB
<i>Ereimochelys imbricata</i>	MVZ233242				*	*
<i>Erinaceus europaeus</i>	n/a	ENSEEUT00000007540			ENSEEUT00000010453	ENSEEUT00000010453
<i>Fejervarya limncharis</i>	MVZ251769		*		*	*
<i>Felis catus</i>	n/a	ENSFCAT00000013358	ENSFCAT00000012043	ENSFCAT00000002541	ENSFCAT00000011087	ENSFCAT00000011087
<i>Gallus gallus</i>	n/a	ENSGALT00000011731	ENSGALT00000007717	ENSGALT00000024329	ENSGALT00000007299	ENSGALT00000007299
<i>Gasterosteus aculeatus</i>	n/a	ENSGACT00000011351	ENSGACT00000024824	ENSGACT00000023657	ENSGACT00000027698	ENSGACT00000027698
<i>Geochelone sulcata</i>	MVZ238888	*			*	*
<i>Geomyda spengleri</i>	MVZ257081				*	*
<i>Geotrypetes seraphini</i>	MVZ252475				*	*
<i>Gorilla gorilla</i>	n/a		ENSGGOT000000006139	ENSGGOT00000002668		
<i>Hardella thurjii</i>	MVZ248392				*	*
<i>Heloderma horridum</i>	MVZ238489				*	*
<i>Heteronotia binoci</i>	MKF226		*	*	*	*
<i>Homo sapiens</i>	n/a	ENST00000392514	ENST00000316084	ENST00000315377	ENST00000407193	ENST00000407193
<i>Hynobius arisanensis</i>	MVZ251776				*	*
<i>Icthyophis bannanicus</i>	MVZ226265				*	*
<i>Indotestudo elongata</i>	MVZ234627				*	*
<i>Kachuga smithii</i>	MVZ248394				*	*
<i>Kinixys belliana</i>	MVZ234631					
<i>Kinosternon scorpioides</i>	MVZ247584				*	*
<i>Leptotyphlops dulcis</i>	MVZ236621				*	*
<i>Leucocephalon yuwonoi</i>	MVZ241499					
<i>Lissemys punctata</i>	MVZ248398					
<i>Loxodonta africana</i>	n/a	ENSLAFT00000012291	ENSLAFT00000002654	ENSLAFT00000010764		
<i>Macaca mulatta</i>	n/a	ENSMMUT00000045510	ENSMMUT00000004652	ENSMMUT00000004827	ENSMMUT00000040826	ENSMMUT00000040826
<i>Macropus eugenii</i>	n/a	ENSMEUT00000016452	ENSMEUT00000007988		ENSMEUT00000014948	ENSMEUT00000014948
<i>Malaclemys terrapin</i>	MVZ257705	*	*	*	*	*
<i>Malacochersus tornieri</i>	MVZ234632				*	*
<i>Malayemys subtrijuga</i>	MVZ230474				*	*
<i>Mauremys caspica</i>	MVZ237439				*	*
<i>Mauremys mutica</i>	MVZ230487				*	*
<i>Mesalina brevirostris</i>	MVZ246016				*	*
<i>Microcebus murinus</i>	n/a	ENSMICT00000012550	ENSMICT00000016603	ENSMICT00000007255	ENSMICT00000002900	ENSMICT00000002900
<i>Monodelphis domestica</i>	n/a	ENSMODT00000019339	ENSMODT00000035235	ENSMODT00000005147	ENSMODT00000012078	ENSMODT00000012078
<i>Mus musculus</i>	n/a	ENSMUST000000086519	ENSMUST00000033683	ENSMUST00000102814	ENSMUST00000122279	ENSMUST00000122279
<i>Myotis lucifugus</i>	n/a	ENSMLUT00000001969	ENSMLUT00000011687	ENSMLUT00000007575	ENSMLUT00000013279	ENSMLUT00000013279
<i>Nectarinia olivacea</i>	JF1603				*	*
<i>Nerodia rhombifer</i>	MVZ246092				*	*
<i>Ochotona princeps</i>	n/a	ENSOPRT00000018469	ENSOPRT00000009586	ENSOPRT00000000746	ENSOPRT00000008420	ENSOPRT00000008420
<i>Ornithorhynchus anatinus</i>	n/a	ENSOANT00000010419	ENSOANT00000004685	ENSOANT00000005142	ENSOANT00000005084	ENSOANT00000005084
<i>Oryzotolagus cuniculus</i>	n/a	ENSOCUT00000005748	ENSOCUT00000015749	ENSOCUT00000003335	ENSOCUT00000011505	ENSOCUT00000011505
<i>Oryzias latipes</i>	n/a		ENSORLT00000012016	ENSORLT00000007921	ENSORLT00000016163	ENSORLT00000016163
<i>Oscacilia ochrocephala</i>	MVZ232864				*	*
<i>Otolemur garnettii</i>	n/a	ENSOGAT00000003846	ENSOGAT00000011810	ENSOGAT00000016257	ENSOGAT00000000722	ENSOGAT00000000722
<i>Pan troglodytes</i>	n/a	ENSPTRT00000010142	ENSPTRT00000040988		ENSPTRT00000032220	ENSPTRT00000032220

Scientific Name	Specimen #	RPLP0	RPS4X	RPS6	RPS14	SDHB
<i>Pelodiscus sinensis</i>	MVZ226472				*	*
<i>Pelomedusa subrufa</i>	MVZ236628				*	*
<i>Phrynos sp.</i>	MVZ247578				*	*
<i>Pipa pipa</i>	MVZ247516				*	*
<i>Platemys platycephala</i>	MVZ247579				*	*
<i>Platysternon megacephalum</i>	MVZ230486				*	*
<i>Plestiodon gilberti</i>	MVZ249915				*	*
<i>Pongo pygmaeus</i>	n/a	ENSPPYT0000005944	ENSPPYT00000023872	ENSPPYT00000022374	ENSPPYT00000018565	ENSPPYT00000018565
<i>Procarvia capensis</i>	n/a	ENSPCAT00000016408	ENSPCAT00000000728	ENSPCAT00000011794	ENSPCAT00000015915	ENSPCAT00000015915
<i>Pseudemys concinna</i>	MVZ241520				*	*
<i>Pteropus ramponius</i>	n/a	ENSPVAT00000003110	ENSPVAT00000012882	ENSPVAT00000014198	ENSPVAT00000005419	ENSPVAT00000005419
<i>Rana livida</i>	MVZ241436					
<i>Rattus norvegicus</i>	n/a	ENSRNOT00000001518	ENSRNOT000000051149	ENSRNOT00000010383	ENSRNOT00000059501	ENSRNOT00000059501
<i>Rhineura floridana</i>	MVZ233342				*	*
<i>Rhinoclemmys pulcherrima</i>	MVZ230924				*	*
<i>Sacalia quadriocellata</i>	MVZ230485				*	*
<i>Sorex araneus</i>	n/a	ENSSART00000006407	ENSSART00000010812	ENSSART00000006537	ENSSART00000014254	ENSSART00000014254
<i>Spermophilus tridecemlineatus</i>	n/a	ENSSTOT00000011324		ENSSTOT00000002220		
<i>Sternotherus carinatus</i>	MVZ238123				*	*
<i>Sternotherus minor</i>	MVZ241528				*	*
<i>Stigmochelys pardalis</i>	MVZ241333				*	*
<i>Sus scrofa</i>	n/a	ENSSSCT00000010850	ENSSSCT00000013564	ENSSSCT00000005698	ENSSSCT00000010068	ENSSSCT00000010068
<i>Sylvietta virens</i>	JF1484				*	*
<i>Taeniopygia guttata</i>	n/a	ENSTGUT00000009872	ENSTGUT00000004981	ENSTGUT00000004518	ENSTGUT00000019180	ENSTGUT00000019180
<i>Takifugu rubripes</i>	n/a	ENSTRUT000000040687	ENSTRUT000000020343	ENSTRUT000000023949	ENSTRUT00000018484	ENSTRUT00000018484
<i>Tarsius syrichta</i>	n/a	ENSTSYT00000009733	ENSTSYT00000003142	ENSTSYT00000008566		
<i>Terpsiphone batesi</i>	JF1488	*			*	*
<i>Terrapene carolina</i>	MVZ241538				*	*
<i>Testudo graeca</i>	MVZ234282					
<i>Testudo hermanni</i>	MVZ244866	*			*	*
<i>Tetraodon nigroviridis</i>	n/a	ENSTNIT00000014131	ENSTNIT00000005585	ENSTNIT00000013679	ENSTNIT00000010602	ENSTNIT00000010602
<i>Trachemys scripta</i>	MVZ241524				*	*
<i>Tupaia belangeri</i>	n/a	ENSTBET00000005503	ENSTBET00000002906	ENSTBET00000003456	ENSTBET00000006351	ENSTBET00000006351
<i>Tursiops truncatus</i>	n/a	ENSTTRT00000005658	ENSTTRT00000006164	ENSTTRT00000001786	ENSTTRT00000013032	ENSTTRT00000013032
<i>Varanus yemenensis</i>	MVZ236610				*	*
<i>Vigna paxos</i>	n/a	ENSVPAT00000012155	ENSVPAT00000010952	ENSVPAT00000006030	ENSVPAT00000003288	ENSVPAT00000003288
<i>Xantusia vigilis</i>	MVZ236373					
<i>Xenopeltis unicolor</i>	MVZ239387					
<i>Xenopus tropicalis</i>	n/a	ENSXETT000000041250	ENSXETT000000056754	ENSXETT000000049958	ENSXETT000000034484	ENSXETT000000034484

Scientific Name	Specimen #	SEC13	SEC62	SGK1	SLC25A20	SMU1
<i>Accipiter cooperii</i>	MVZ181729					
<i>Agelaiocercus kingi</i>	MVZ183274				*	
<i>Agrionemys borsfieldii</i>	MVZ245929					
<i>Aluthe castanea</i>	JF1706		*		*	
<i>Alligator mississippiensis</i>	n/a					
<i>Alligator mississippiensis</i>	MVZ241485	*				*
<i>Anas cyanoptera</i>	MVZ179315					
<i>Anniella pulchra</i>	MVZ247492		*			
<i>Anolis carolinensis</i>	n/a	ENSACAT00000014914	ENSACAT00000015851	ENSACAT00000010308	ENSACAT00000012163	ENSACAT00000003478
<i>Apalone ferox</i>	MVZ241535			*		*
<i>Aptenodytes patagonicus</i>	MVZ178921					
<i>Aspidoscelis tigris</i>	MVZ243345					
<i>Bipes biporus</i>	MVZ236260					
<i>Bos taurus</i>	n/a	ENSBAT00000023702		ENSBAT00000005592	ENSBAT00000000217	ENSBAT00000020865
<i>Callithrix jacchus</i>	n/a	ENSCJAT00000031986	ENSCJAT00000004955	ENSCJAT00000005308	ENSCJAT00000006321	ENSCJAT00000015463
<i>Canis familiaris</i>	n/a	ENSACFT000000008242	ENSACFT000000023437	ENSACFT00000000342	ENSACFT000000019184	ENSACFT000000002895
<i>Casuarinus casuarinus</i>	MVZ178927					
<i>Cavia prellus</i>	n/a	ENSCPOT00000015862	ENSCPOT00000007463	ENSCPOT00000001514	ENSCPOT000000009394	ENSCPOT000000006277
<i>Chamaeleo senegalensis</i>	MVZ249659					
<i>Chelodina longicollis</i>	MVZ230488					
<i>Chelonia mydas</i>	MVZ233240					
<i>Chelonoidis denticulata</i>	MVZ175391					
<i>Chelydra serpentina</i>	MVZ250691	*	*			
<i>Choloepus hoffmanni</i>	n/a	ENSCHOT00000010222	*	ENSCHOT00000006654	ENSCHOT00000004642	ENSCHOT000000008199
<i>Chrysomys picta</i>	MVZ241506	*		*		*
<i>Claudius angustatus</i>	MVZ164771			*		
<i>Cuora galbinifrons</i>	MVZ230933					
<i>Cuora pani</i>	MVZ230513					
<i>Cyrtodactylus</i>	JAM8710	*	*		*	
<i>Danio rerio</i>	n/a	ENDSART000000077063		ENDSART000000035031	ENDSART000000059140	ENDSART000000098346
<i>Dasylops novemcinctus</i>	n/a	ENDSDNOT00000000127	ENDSDNOT000000014077	ENDSDNOT000000017414	ENDSDNOT000000000976	ENDSDNOT000000000933
<i>Dendropicus gabonensis</i>	JF1513		*	*	*	*
<i>Dermatemys mawii</i>	MVZ233618					
<i>Dermochelys coriacea</i>	MVZ149844					
<i>Dibamus sp.</i>	MVZ224113					
<i>Dierochelys reticularia</i>	MVZ230923	*	*			*
<i>Dipodomys ordii</i>	n/a	ENDSDORT000000015325		ENDSDORT000000006575	ENDSDORT000000010459	ENDSDORT000000006012
<i>Draco beccarii</i>	JAM8961	*	*	*		*
<i>Draco sumatranus</i>	SZL71			*		
<i>Echinops telfairi</i>	n/a	ENSETET00000001957	ENSETET000000016696	ENSETET000000010202	ENSETET000000012599	ENSETET000000007954
<i>Emydura subglobosa</i>	MVZ230496					
<i>Emys marmorata</i>	MVZ164996					
<i>Ensatina eschscholtzii</i>	RP158					
<i>Equus caballus</i>	n/a	ENSECAT000000021918	ENSECAT000000018140	ENSECAT000000021724	ENSECAT000000021215	ENSECAT000000022765

Scientific Name	Specimen #	SEC13	SEC62	SGK1	SLC25A20	SMU1
<i>Eremochelys imbricata</i>	MVZ233242					
<i>Erinaceus europaeus</i>	n/a		ENSEEUT00000000927	ENSEEUT00000008771		ENSEEUT00000012107
<i>Fijieryza limnocharis</i>	MVZ251769					
<i>Felis catus</i>	n/a	ENSFCAT00000011248		ENSFCAT00000013876	ENSFCAT00000009840	ENSFCAT00000008731
<i>Gallus gallus</i>	n/a	ENSGALT00000013715	ENSGALT00000015276	ENSGALT00000031766	ENSGALT00000011304	ENSGALT00000003135
<i>Gasterosteus aculeatus</i>	n/a			ENSGACT00000014583	ENSGACT00000001538	ENSGACT00000022220
<i>Geochelone sulcata</i>	MVZ238888	*				*
<i>Geomyda spengleri</i>	MVZ257081		*			
<i>Geotrypetes seraphini</i>	MVZ252475					
<i>Gorilla gorilla</i>	n/a	ENSGGOT00000010201	*	ENSGGOT00000001505		ENSGGOT00000014859
<i>Hardella thurjii</i>	MVZ248392	*		*		*
<i>Heloderma horridum</i>	MVZ238489					
<i>Heteronotia binoei</i>	MKF226	*	*			*
<i>Homo sapiens</i>	n/a	ENST00000383801	ENST00000480708	ENST00000367857	ENST00000440964	ENST00000397149
<i>Hynobius arisanensis</i>	MVZ251776					
<i>Icthyophis bannanicus</i>	MVZ226265		*	*		*
<i>Indotestudo elongata</i>	MVZ234627					
<i>Kachuga smithii</i>	MVZ248394		*	*		*
<i>Kinixys belliana</i>	MVZ234631					
<i>Kinosternon scorpioides</i>	MVZ247584					
<i>Leptotyphlops dulcis</i>	MVZ236621					
<i>Leuccephalon yuwonoi</i>	MVZ241499					
<i>Lissemys punctata</i>	MVZ248398					
<i>Loxodonta africana</i>	n/a	ENSLAFT00000014599	ENSLAFT00000006506	ENSLAFT00000012699	ENSLAFT00000000862	ENSLAFT00000001530
<i>Macaca mulatta</i>	n/a	ENSM MUT00000028004	ENSM MUT00000030814	ENSM MUT00000011374	ENSM MUT00000021620	ENSM MUT00000002952
<i>Macropus eugenii</i>	n/a	ENSMEUT00000004845	ENSMEUT00000000554	ENSMEUT00000012358	ENSMEUT00000013233	
<i>Malaclemys terrapin</i>	MVZ257705	*	*	*		*
<i>Malacochersus tornieri</i>	MVZ234632					
<i>Malayemys subtrijuga</i>	MVZ230474			*		*
<i>Mauremys caspica</i>	MVZ237439					
<i>Mauremys mutica</i>	MVZ230487					
<i>Mesalina brevirostris</i>	MVZ246016					
<i>Microcebus murinus</i>	n/a	ENSMICT00000002399		ENSMICT00000007885	ENSMICT00000005025	ENSMICT00000012077
<i>Monodelphis domestica</i>	n/a	ENSMODT00000006949	ENSMODT00000036602	ENSMODT00000022227	ENSMODT00000017045	ENSMODT00000004300
<i>Mus musculus</i>	n/a	ENSMUST00000032440	ENSMUST00000029256	ENSMUST00000020145	ENSMUST00000035222	ENSMUST00000030117
<i>Myotis lucifugus</i>	n/a	ENSMLUT00000004900	ENSMLUT00000002553	ENSMLUT00000016793	ENSMLUT00000013935	ENSMLUT00000016098
<i>Nectarinia olivacea</i>	JF1603					
<i>Nerodia rhombifer</i>	MVZ246092					
<i>Ochotona princeps</i>	n/a	ENSOPRT00000007059	ENSOPRT00000013539	ENSOPRT00000012172	ENSOPRT00000011620	ENSOPRT00000011383
<i>Ornithorhynchus anatinus</i>	n/a	ENSOANT00000014867	ENSOANT00000012201	ENSOANT00000020700	ENSOANT00000012061	ENSOANT00000015665
<i>Oryzotylagus cuniculus</i>	n/a	ENSOCUT00000010699	ENSOCUT00000008717	ENSOCUT00000014369	ENSOCUT00000016421	ENSOCUT00000008928
<i>Oryzias latipes</i>	n/a	ENSORLT000000024784		ENSORLT00000020686		ENSORLT00000013891
<i>Oscacilia ochrocephala</i>	MVZ232864		*			
<i>Otolemur garnettii</i>	n/a	ENSOGAT00000017054	ENSOGAT00000008951	ENSOGAT00000008736	ENSOGAT00000001847	ENSOGAT00000004369
<i>Pan troglodytes</i>	n/a	ENSPTRT000000027349	ENSPTRT000000029111	ENSPTRT00000034392	ENSPTRT000000027891	ENSPTRT000000038576

Scientific Name	Specimen #	SEC13	SEC62	SGK1	SLC25A20	SMU1
<i>Pelodiscus sinensis</i>	MVZ226472					
<i>Pelomedusa subrufa</i>	MVZ236628					
<i>Phrynosoma sp.</i>	MVZ247578					
<i>Pipa pipa</i>	MVZ247516					
<i>Platenys platycephala</i>	MVZ247579					
<i>Platysternon megacephalum</i>	MVZ230486					
<i>Plestiodon gilberti</i>	MVZ249915					
<i>Pongo pygmaeus</i>	n/a	ENSPPYT00000015878	ENSPPYT00000016606	ENSPPYT00000019815	ENSPPYT00000016158	ENSPPYT00000022313
<i>Procarina capensis</i>	n/a	ENSPCAT00000009806	ENSPCAT00000011052	ENSPCAT00000010163	ENSPCAT00000000727	ENSPCAT00000012328
<i>Pseudemys concinna</i>	MVZ241520					
<i>Pteropus vampyrus</i>	n/a	ENSPVAT00000012626	ENSPVAT00000001748	ENSPVAT00000015517	ENSPVAT00000008703	ENSPVAT00000011471
<i>Rana livida</i>	MVZ241436					
<i>Rattus norvegicus</i>	n/a	ENSRNOT00000014377	ENSRNOT00000012240	ENSRNOT00000016121	ENSRNOT00000027520	ENSRNOT00000010464
<i>Rhinerea floridana</i>	MVZ233342					
<i>Rhinoclemmys pulcherrima</i>	MVZ230924	*	*	*		*
<i>Sacalia quadriocellata</i>	MVZ230485					
<i>Sorex araneus</i>	n/a		ENSSART00000010945	ENSSART00000007818	ENSSART00000008064	ENSSART00000001384
<i>Spermophilus tridecemlineatus</i>	n/a	ENSSTOT00000014021	ENSSTOT00000009370		ENSSTOT00000003325	ENSSTOT00000005219
<i>Sternotherus carinatus</i>	MVZ238123	*		*		*
<i>Sternotherus minor</i>	MVZ241528	*		*	*	*
<i>Stigmochelys pardalis</i>	MVZ241333					*
<i>Sus scrofa</i>	n/a				ENSSSCT00000012444	ENSSSCT00000012041
<i>Sylvietta virens</i>	JF1484					
<i>Taeniopygia guttata</i>	n/a	ENSTGUT00000010897	ENSTGUT00000011446	ENSTGUT00000012019	ENSTGUT00000009491	ENSTGUT00000001519
<i>Takifugu rubripes</i>	n/a	ENSTRUT00000003472		ENSTRUT00000006508	ENSTRUT000000028599	ENSTRUT00000010683
<i>Tarsius syrichta</i>	n/a	ENSTSYT00000009980	ENSTSYT0000001033	ENSTSYT00000002891	ENSTSYT00000014696	ENSTSYT00000008071
<i>Terpsiphone batesi</i>	JF1488				*	
<i>Terrapene carolina</i>	MVZ241538					
<i>Testudo graeca</i>	MVZ234282					
<i>Testudo hermanni</i>	MVZ244866	*	*	*		*
<i>Tetraodon nigroviridis</i>	n/a	ENSTNIT00000005471		ENSTNIT00000019646		ENSTNIT00000016270
<i>Trachemys scripta</i>	MVZ241524					
<i>Tupaia belangeri</i>	n/a		ENSTBET00000009656	ENSTBET00000002496	ENSTBET00000017557	ENSTBET00000000943
<i>Tursiops truncatus</i>	n/a	ENSTTRT00000004393	ENSTTRT0000001890	ENSTTRT00000009709	ENSTTRT00000017343	ENSTTRT00000007388
<i>Varamus yemenensis</i>	MVZ236610					
<i>Vicugna pacos</i>	n/a	ENSVPAT00000005489	ENSVPAT00000007986	ENSVPAT00000006275	ENSVPAT00000009655	ENSVPAT00000001300
<i>Xantusia vigilis</i>	MVZ236373					
<i>Xenopeltis unicolor</i>	MVZ239387		*			
<i>Xenopus tropicalis</i>	n/a	ENSXETT00000003059	ENSXETT00000020231	ENSXETT00000000246	ENSXETT00000020560	ENSXETT000000052109

Scientific Name	Specimen #	SNX6	SOD2	SRP54	STK39	TAT
<i>Accipiter cooperii</i>	MVZ181729					
<i>Agelaiocercus kingi</i>	MVZ183274		*			
<i>Agriomys borsfieldii</i>	MVZ245929					
<i>Aluthe castanea</i>	JF1706		*			
<i>Alligator mississippiensis</i>	n/a			ES316890		ES317837
<i>Alligator mississippiensis</i>	MVZ241485	*				
<i>Anas cyanoptera</i>	MVZ179315					
<i>Anniella pulchra</i>	MVZ247492	*		*	*	
<i>Anolis carolinensis</i>	n/a	ENSACAT00000002518	ENSACAT00000004516	ENSACAT00000002230	ENSACAT00000001427	ENSACAT00000007792
<i>Apalone ferox</i>	MVZ241535			*		
<i>Aptenodytes patagonicus</i>	MVZ178921					
<i>Aspidoscelis tigris</i>	MVZ243345					
<i>Bipes biporus</i>	MVZ236260					
<i>Bos taurus</i>	n/a	ENSBAT00000017106	ENSBAT00000008569	ENSBAT00000025405	ENSBAT00000061632	ENSBAT00000002866
<i>Callithrix jacchus</i>	n/a	ENSCJAT00000012439	ENSCJAT00000022748	ENSCJAT00000012915	ENSCJAT00000012805	ENSCJAT00000029184
<i>Canis familiaris</i>	n/a	ENSACAT00000021062	ENSACAT00000001037	ENSACAT00000021158	ENSACAT00000018748	ENSACAT00000032147
<i>Casuarinus casuarinus</i>	MVZ178927					
<i>Cavia prellus</i>	n/a	ENSCPOT00000006931	ENSCPOT00000000381	ENSCPOT00000003735	ENSCPOT00000013709	ENSCPOT00000008085
<i>Chamaeleo senegalensis</i>	MVZ249659					
<i>Chelodina longicollis</i>	MVZ230488					
<i>Chelonia mydas</i>	MVZ233240					
<i>Chelonoidis denticulata</i>	MVZ175391					
<i>Chelydra serpentina</i>	MVZ250691					
<i>Choloepus hoffmanni</i>	n/a	ENSCHOT00000012073	ENSCHOT00000003227	ENSCHOT00000009009	ENSCHOT00000001127	ENSCHOT00000001944
<i>Chrysomys picta</i>	MVZ241506	*	*	*		
<i>Claudius angustatus</i>	MVZ164771					
<i>Cuora galbinifrons</i>	MVZ230933					
<i>Cuora pani</i>	MVZ230513					
<i>Cyrtodactylus</i>	JAM8710	*				
<i>Danio rerio</i>	n/a	ENDSART00000105121	ENDSART00000062556	ENDSART00000018678		ENDSART00000101472
<i>Dasyurus novemcinctus</i>	n/a	ENDSDNOT00000002436	ENDSDNOT00000007903	ENDSDNOT00000017538	ENDSDNOT00000011604	ENDSDNOT00000015734
<i>Dendropicus gabonensis</i>	JF1513	*	*	*		*
<i>Dermatemys mawii</i>	MVZ233618					
<i>Dermochelys coriacea</i>	MVZ149844					
<i>Dibamus sp.</i>	MVZ224113	*				
<i>Dierochelys reticularia</i>	MVZ230923	*		*		
<i>Dipodomys ordii</i>	n/a	ENDSDORT00000007001	ENDSDORT00000001353	ENDSDORT00000005157	ENDSDORT00000015054	ENDSDORT00000002832
<i>Draco beccarii</i>	JAM8961	*				
<i>Draco sumatranus</i>	SZL71	*				
<i>Echinops telfairi</i>	n/a	ENSETET00000015870	ENSETET00000016400	ENSETET00000016309	ENSETET00000011386	ENSETET00000018731
<i>Emydura subglobosa</i>	MVZ230496					
<i>Emys marmorata</i>	MVZ164996					
<i>Ensatina eschscholtzii</i>	RP158					
<i>Equus caballus</i>	n/a	ENSECAT00000013453	ENSECAT00000012435	ENSECAT00000019578	ENSECAT00000015189	ENSECAT00000023194

Scientific Name	Specimen #	SNX6	SOD2	SRP54	STK39	TAT
<i>Eremochelys imbricata</i>	MVZ233242		*	*		
<i>Erinaceus europaeus</i>	n/a	ENSEEUT00000013932		ENSEEUT00000001514	ENSEEUT000000009014	ENSEEUT000000006287
<i>Fagelaryia limnocharis</i>	MVZ251769					
<i>Felis catus</i>	n/a	ENSFCAT00000018857	ENSFCAT00000002843	ENSFCAT00000015441	ENSFCAT00000011276	ENSFCAT000000004380
<i>Gallus gallus</i>	n/a	ENSGALT00000016282	ENSGALT00000019062	ENSGALT00000016327	ENSGALT00000017779	ENSGALT00000001314
<i>Gasterosteus aculeatus</i>	n/a	ENSGACT00000005459	ENSGACT00000011956	ENSGACT00000005684	ENSGACT00000007102	ENSGACT00000000602
<i>Geochelone sulcata</i>	MVZ238888		*			
<i>Geomyda spengleri</i>	MVZ257081			*		
<i>Geotrypetes seraphini</i>	MVZ252475					
<i>Gorilla gorilla</i>	n/a	ENSGGOT00000007692	ENSGGOT00000008248	ENSGGOT00000002943	ENSGGOT00000004381	ENSGGOT000000004720
<i>Hardella thurjii</i>	MVZ248392	*		*		
<i>Heloderma horridum</i>	MVZ238489					
<i>Heteronotia binoei</i>	MKF226	*		*	*	
<i>Homo sapiens</i>	n/a	ENST00000362031	ENST00000337404	ENST00000216774	ENST00000355999	ENST00000355962
<i>Hynobius arisanensis</i>	MVZ251776					
<i>Icthyophis bannanicus</i>	MVZ226265	*				
<i>Indotestudo elongata</i>	MVZ234627					
<i>Kachuga smithii</i>	MVZ248394	*				
<i>Kinixys belliana</i>	MVZ234631					
<i>Kinosternon scorpioides</i>	MVZ247584					
<i>Leptotyphlops dulcis</i>	MVZ236621					
<i>Leuccephalon yuwonoi</i>	MVZ241499					
<i>Lissemys punctata</i>	MVZ248398					
<i>Loxodonta africana</i>	n/a	ENSLAFT00000010573	ENSLAFT00000007331	ENSLAFT00000012008	ENSLAFT00000000792	ENSLAFT00000007149
<i>Macaca mulatta</i>	n/a	ENSMMUT00000017624	ENSMMUT00000008233	ENSMMUT000000021128	ENSMMUT00000008694	ENSMMUT00000002859
<i>Macropus eugenii</i>	n/a	ENSMEUT00000010615		ENSMEUT00000000141		ENSMEUT000000004596
<i>Malaclemys terrapin</i>	MVZ257705	*	*	*	*	
<i>Malacocheilus tornieri</i>	MVZ234632					
<i>Malayemys subtrijuga</i>	MVZ230474			*		
<i>Mauremys caspica</i>	MVZ237439					
<i>Mauremys mutica</i>	MVZ230487					
<i>Mesalina brevirostris</i>	MVZ246016					
<i>Microcebus murinus</i>	n/a	ENSMICT00000009335		ENSMICT00000010271	ENSMICT00000012420	ENSMICT00000001889
<i>Monodelphis domestica</i>	n/a	ENSMODT00000018472	ENSMODT00000009516	ENSMODT00000016472	ENSMODT00000009571	ENSMODT00000002821
<i>Mus musculus</i>	n/a	ENSMUST00000005798	ENSMUST00000007012	ENSMUST00000021407	ENSMUST00000102715	ENSMUST00000001720
<i>Myotis lucifugus</i>	n/a	ENSMLUT00000007045	ENSMLUT00000012243	ENSMLUT00000008994	ENSMLUT00000013109	ENSMLUT00000008660
<i>Nectarinia olivacea</i>	JF1603					
<i>Nerodia rhombifer</i>	MVZ246092					
<i>Ochotona princeps</i>	n/a	ENSOPRT00000015344	ENSOPRT00000004794	ENSOPRT00000014508	ENSOPRT00000013201	ENSOPRT00000001231
<i>Ornithorhynchus anatinus</i>	n/a	ENSOANT00000022156	ENSOANT00000000987	ENSOANT00000022166	ENSOANT00000013507	ENSOANT00000013005
<i>Oryzotylagus cuniculus</i>	n/a			ENSOCUT00000007839		ENSOCUT00000009958
<i>Oryzias latipes</i>	n/a		ENSORLT00000016614	ENSORLT00000012579		ENSORLT00000019234
<i>Oscacilia ochrocephala</i>	MVZ232864					
<i>Oryzomys latipes</i>	n/a	ENSOGAT00000008024	ENSOGAT00000017105	ENSOGAT00000012194	ENSOGAT00000007893	ENSOGAT00000002847
<i>Pan troglodytes</i>	n/a	ENSPTRT00000011469	ENSPTRT000000061574	ENSPTRT00000011483	ENSPTRT00000023347	ENSPTRT00000015355

Scientific Name	Specimen #	SNX6	SOD2	SRP54	STK39	TAT
<i>Pelodiscus sinensis</i>	MVZ226472					
<i>Pelomedusa subrufa</i>	MVZ236628					
<i>Phrynosoma sp.</i>	MVZ247578					
<i>Pipa pipa</i>	MVZ247516					
<i>Platemys platycephala</i>	MVZ247579					
<i>Platysternon megacephalum</i>	MVZ230486					
<i>Plestiodon gilberti</i>	MVZ249915					
<i>Pongo pygmaeus</i>	n/a	ENSPPYT00000006775	ENSPPYT00000019953	ENSPPYT00000006778	ENSPPYT00000014988	ENSPPYT00000008848
<i>Procarion capensis</i>	n/a	ENSPCAT00000003793	ENSPCAT00000008133	ENSPCAT00000006287	ENSPCAT00000013763	ENSPCAT00000003314
<i>Pseudemys concinna</i>	MVZ241520	*		*		
<i>Pteropus vampyrus</i>	n/a	ENSPVAT00000002793	ENSPVAT00000005300	ENSPVAT00000010056	ENSPVAT00000017091	ENSPVAT00000003644
<i>Rana livida</i>	MVZ241436					
<i>Rattus norvegicus</i>	n/a	ENSRNOT00000007300	ENSRNOT00000025794	ENSRNOT00000049937	ENSRNOT00000037133	ENSRNOT00000022721
<i>Rhinoceros floridana</i>	MVZ233342					
<i>Rhinoceros pulcherrima</i>	MVZ230924	*		*	*	
<i>Sacalia quadriocellata</i>	MVZ230485					
<i>Sorex araneus</i>	n/a			ENSSART00000002750	ENSSART00000013862	
<i>Spermophilus tridecemlineatus</i>	n/a	ENSSTOT00000011100	ENSSTOT00000008722	ENSSTOT00000016026	ENSSTOT00000013685	ENSSTOT00000003062
<i>Sternotherus carinatus</i>	MVZ238123	*		*		
<i>Sternotherus minor</i>	MVZ241528	*	*	*		
<i>Stigmochelys pardalis</i>	MVZ241333			*		
<i>Sus scrofa</i>	n/a	ENSSSCT00000002194	ENSSSCT00000004479	ENSSSCT00000002187		ENSSSCT00000003032
<i>Sylvietta virens</i>	JF1484					
<i>Taeniopygia guttata</i>	n/a	ENSTGUT00000012425	ENSTGUT00000011861	ENSTGUT00000012459	ENSTGUT00000007884	ENSTGUT00000008728
<i>Takifugu rubripes</i>	n/a	ENSTRUT00000019623	ENSTRUT00000012643	ENSTRUT00000017411	ENSTRUT00000001319	ENSTRUT00000013683
<i>Tarsius syrichta</i>	n/a	ENSTSYT00000002928	ENSTSYT00000000457	ENSTSYT00000002681	ENSTSYT00000000034	ENSTSYT000000006038
<i>Terapia batesi</i>	JF1488		*			*
<i>Terrapene carolina</i>	MVZ241538					
<i>Testudo graeca</i>	MVZ234282	*	*			
<i>Testudo hermanni</i>	MVZ244866	*	*	*		
<i>Tetraodon nigroviridis</i>	n/a	ENSTNIT00000005290	ENSTNIT00000015459	ENSTNIT00000005287	ENSTNIT00000022118	ENSTNIT00000011873
<i>Trachemys scripta</i>	MVZ241524					
<i>Tupaia belangeri</i>	n/a	ENSTBET00000001433	ENSTBET00000002087	ENSTBET00000011420	ENSTBET00000004925	ENSTBET00000007778
<i>Tursiops truncatus</i>	n/a	ENSTTRT00000002158	ENSTTRT00000014354	ENSTTRT00000000928	ENSTTRT00000000057	ENSTTRT00000000149
<i>Varamus yemenensis</i>	MVZ236610					
<i>Vicugna pacos</i>	n/a	ENSVPAT00000003272	ENSVPAT00000008200	ENSVPAT00000005671	ENSVPAT00000007148	ENSVPAT00000002907
<i>Xantusia vigilis</i>	MVZ236373					
<i>Xenopeltis unicolor</i>	MVZ239387	*		*		
<i>Xenopus tropicalis</i>	n/a	ENSXETT00000026895	ENSXETT00000004579	ENSXETT00000026869	ENSXETT00000035375	ENSXETT00000050583

Scientific Name	Specimen #	UBA5	UBE2J2	UBE2K	UBE2N	UBE2Q1
<i>Accipiter cooperii</i>	MVZ181729					
<i>Agelaiocercus kingi</i>	MVZ183274					
<i>Agrionemys borsfieldii</i>	MVZ245929					
<i>Aluthe castanea</i>	JF1706			*	*	*
<i>Alligator mississippiensis</i>	n/a					
<i>Alligator mississippiensis</i>	MVZ241485				*	
<i>Anas cyanoptera</i>	MVZ179315					
<i>Anniella pulchra</i>	MVZ247492			*		
<i>Anolis carolinensis</i>	n/a	ENSACAT00000000616	ENSACAT00000017064	ENSACAT00000011332	ENSACAT00000012593	
<i>Apalone ferox</i>	MVZ241535			*		
<i>Aptenodytes patagonicus</i>	MVZ178921					
<i>Aspidoscelis tigris</i>	MVZ243345					
<i>Bipes biporus</i>	MVZ236260					
<i>Bos taurus</i>	n/a	ENSBAT00000005894	ENSBAT00000021109	ENSBAT00000026871	ENSBAT00000057014	ENSBAT00000005253
<i>Callithrix jacchus</i>	n/a	ENSCJAT00000027241	ENSCJAT00000022014	ENSCJAT00000038164	ENSCJAT00000005692	*
<i>Canis familiaris</i>	n/a	ENSACAT00000010349	ENSACAT00000030687	ENSACAT00000025303	ENSACAT00000009993	ENSACAT00000028836
<i>Casuarinus casuarinus</i>	MVZ178927					
<i>Cavia prellus</i>	n/a	ENSCPOT00000010020	ENSCPOT00000008100	ENSCPOT00000014943	ENSCPOT00000008658	ENSCPOT00000004374
<i>Chamaeleo senegalensis</i>	MVZ249659					
<i>Chelodina longicollis</i>	MVZ230488					
<i>Chelonia mydas</i>	MVZ233240					
<i>Chelonoidis denticulata</i>	MVZ175391					
<i>Chelydra serpentina</i>	MVZ250691			*	*	
<i>Cholepus hoffmanni</i>	n/a	ENSCHOT00000003908	ENSCHOT00000001197	ENSCHOT00000001255	ENSCHOT000000009765	
<i>Chrysomys picta</i>	MVZ241506			*	*	
<i>Claudius angustatus</i>	MVZ164771				*	
<i>Cuora galbiniifrons</i>	MVZ230933					
<i>Cuora pani</i>	MVZ230513					
<i>Cyrtodactylus</i>	JAM8710			*	*	
<i>Danio rerio</i>	n/a	ENDSART000000093046	ENDSART000000086924	ENDSART000000025770		ENDSART000000097333
<i>Dasyatis novemcinctus</i>	n/a	ENDSDOT000000006657	ENDSDOT000000006193	ENDSDOT00000019831	ENDSDOT00000018951	ENDSDOT000000008889
<i>Dendropicus gabonensis</i>	JF1513	*	*	*		
<i>Dermatemys mawii</i>	MVZ233618					
<i>Dermochelys coriacea</i>	MVZ149844					
<i>Dibamus sp.</i>	MVZ224113			*	*	
<i>Dierochelys reticularia</i>	MVZ230923	*		*	*	
<i>Dipodomys ordii</i>	n/a	ENDSDOT00000014304	ENDSDOT00000014331	ENDSDOT00000001443	ENDSDOT000000006648	ENDSDOT000000007807
<i>Draco beccarii</i>	JAM8961			*	*	
<i>Draco sumatranus</i>	SZL71				*	
<i>Echinops telfairi</i>	n/a	ENSETET000000009823	ENSETET00000012964	ENSETET00000003749	ENSETET00000001880	
<i>Emydura subglobosa</i>	MVZ230496					
<i>Emys marmorata</i>	MVZ164996					
<i>Ensatina eschscholtzii</i>	RP158					
<i>Equus caballus</i>	n/a	ENSECAT000000022799		ENSECAT00000015213	ENSECAT00000020822	ENSECAT00000010147

Scientific Name	Specimen #	UBA5	UBE2J2	UBE2K	UBE2N	UBE2Q1
<i>Eremochelys imbricata</i>	MVZ233242					
<i>Erinaceus europaeus</i>	n/a	ENSEEUT00000002612	ENSEEUT00000010768	ENSEEUT00000011848		ENSEEUT00000001588
<i>Fagelaryia limnocharis</i>	MVZ251769					
<i>Felis catus</i>	n/a	ENSFCAT00000005618	ENSFCAT00000003368	ENSFCAT00000006039	ENSFCAT00000004261	
<i>Gallus gallus</i>	n/a	ENSGALT00000019134	ENSGALT00000002864	ENSGALT00000023108	ENSGALT00000018421	
<i>Gasterosteus aculeatus</i>	n/a		ENSGACT00000010489	ENSGACT00000023759	ENSGACT00000024916	ENSGACT00000008411
<i>Geochelone sulcata</i>	MVZ238888					
<i>Geomyda spengleri</i>	MVZ257081				*	
<i>Geotrypetes seraphini</i>	MVZ252475					
<i>Gorilla gorilla</i>	n/a	ENSGGOT00000006125	ENSGGOT00000005363	ENSGGOT00000003280	ENSGGOT00000009313	ENSGGOT00000001088
<i>Hardella thurjii</i>	MVZ248392	*		*		
<i>Heloderma horridum</i>	MVZ238489					
<i>Heteronotia binoei</i>	MKF226			*	*	
<i>Homo sapiens</i>	n/a	ENST00000494238	ENST00000360466	ENST00000261427	ENST00000318066	
<i>Hynobius arisanensis</i>	MVZ251776					
<i>Icthyophis bannanicus</i>	MVZ226265			*	*	
<i>Indotestudo elongata</i>	MVZ234627					
<i>Kachuga smithii</i>	MVZ248394					
<i>Kinixys belliana</i>	MVZ234631					
<i>Kinosternon scorpioides</i>	MVZ247584					
<i>Leptotyphlops dulcis</i>	MVZ236621					
<i>Leuccephalon yuwonoi</i>	MVZ241499					
<i>Lissemys punctata</i>	MVZ248398					
<i>Loxodonta africana</i>	n/a	ENSLAFT00000014713	ENSLAFT00000002223	ENSLAFT00000007161		
<i>Macaca mulatta</i>	n/a	ENSMMUT00000030632		ENSMMUT00000015859	ENSMMUT00000029030	
<i>Macropus eugenii</i>	n/a	ENSMEUT00000013522	ENSMEUT00000006421		ENSMEUT00000002947	ENSMEUT00000004789
<i>Malaclemys terrapin</i>	MVZ257705	*	*	*	*	*
<i>Malacocheilus tornieri</i>	MVZ234632					
<i>Malayemys subtrijuga</i>	MVZ230474				*	
<i>Mauremys caspica</i>	MVZ237439					
<i>Mauremys mutica</i>	MVZ230487					
<i>Mesalina brevirostris</i>	MVZ246016					
<i>Microcebus murinus</i>	n/a	ENSMICT00000004448	ENSMICT00000014160		ENSMICT00000014070	ENSMICT00000014861
<i>Monodelphis domestica</i>	n/a	ENSMODT00000014874	ENSMODT00000020911	ENSMODT00000026174	ENSMODT00000012029	ENSMODT00000015211
<i>Mus musculus</i>	n/a	ENSMUST00000035166	ENSMUST00000105582	ENSMUST00000031105	ENSMUST00000009329	ENSMUST000000053212
<i>Myotis lucifugus</i>	n/a	ENSMLUT00000000824	ENSMLUT00000010689	ENSMLUT00000011206	ENSMLUT00000009969	ENSMLUT000000009421
<i>Nectarinia olivacea</i>	JF1603					
<i>Nerodia rhombifer</i>	MVZ246092					
<i>Ochotona princeps</i>	n/a	ENSOPRT00000002350	ENSOPRT00000002908	ENSOPRT00000010075	ENSOPRT00000015597	
<i>Ornithorhynchus anatinus</i>	n/a	ENSOANT00000010821	ENSOANT00000010778	ENSOANT00000016400	ENSOANT00000010567	
<i>Oryzotylagus cuniculus</i>	n/a	ENSOCUT00000015754	ENSOCUT00000000966	ENSOCUT00000005874	ENSOCUT00000011287	ENSOCUT000000003712
<i>Oryzias latipes</i>	n/a	ENSORLT00000012900	ENSORLT00000020408	ENSORLT00000005524	ENSORLT00000016945	ENSORLT00000016420
<i>Oscacilia ochrocephala</i>	MVZ232864				*	
<i>Otolemur garnettii</i>	n/a	ENSOGAT00000004866	ENSOGAT00000000462	ENSOGAT00000008268	ENSOGAT00000000102	
<i>Pan troglodytes</i>	n/a	ENSPTRT00000049800	ENSPTRT00000000026	ENSPTRT00000029840	ENSPTRT000000065776	

Scientific Name	Specimen #	UBA5	UBE2J2	UBE2K	UBE2N	UBE2Q1
<i>Pelodiscus sinensis</i>	MVZ226472					
<i>Pelomedusa subrufa</i>	MVZ236628					
<i>Phrynosoma sp.</i>	MVZ247578					
<i>Pipa pipa</i>	MVZ247516					
<i>Platemys platycephala</i>	MVZ247579					
<i>Platysternon megacephalum</i>	MVZ230486					
<i>Plestiodon gilberti</i>	MVZ249915			*	*	
<i>Pongo pygmaeus</i>	n/a	ENSPPYT00000016407	ENSPPYT00000002404	ENSPPYT00000017066	ENSPPYT00000005723	ENSPPYT00000007874
<i>Procarina capensis</i>	n/a	ENSPCAT00000015191	ENSPCAT00000011172	ENSPCAT00000011556	ENSPCAT00000000086	ENSPCAT000000000982
<i>Pseudemys concinna</i>	MVZ241520			*		
<i>Pteropus vampyrus</i>	n/a	ENSPVAT00000012599	ENSPVAT00000001013	ENSPVAT00000013727	ENSPVAT00000013732	ENSPVAT000000008159
<i>Rana livida</i>	MVZ241436					
<i>Rattus norvegicus</i>	n/a	ENSRNOT00000015240	ENSRNOT000000026752	ENSRNOT000000065709	ENSRNOT000000059025	ENSRNOT000000019539
<i>Rhinerea floridana</i>	MVZ233342					
<i>Rhinoclemmys pulcherrima</i>	MVZ230924			*	*	
<i>Sacalia quadriocellata</i>	MVZ230485					
<i>Sorex araneus</i>	n/a	ENSSART000000005751	ENSSART000000002502	ENSSART00000013725	ENSSART000000005925	ENSSART000000004758
<i>Spermophilus tridecemlineatus</i>	n/a	ENSSTOT00000012165	ENSSTOT00000001776	ENSSTOT00000012445	ENSSTOT000000002162	ENSSTOT000000003220
<i>Sternotherus carinatus</i>	MVZ238123				*	
<i>Sternotherus minor</i>	MVZ241528			*	*	
<i>Stigmochelys pardalis</i>	MVZ241333					
<i>Sus scrofa</i>	n/a	ENSSSCT00000012728	ENSSSCT00000003701	ENSSSCT000000009618	ENSSSCT000000000992	ENSSSCT000000007169
<i>Sylvietta virens</i>	JF1484					
<i>Taeniopygia guttata</i>	n/a	ENSTGUT000000004628	ENSTGUT000000004247	ENSTGUT000000009159	ENSTGUT000000008606	ENSTGUT000000003648
<i>Takifugu rubripes</i>	n/a	ENSTRUT000000045858	ENSTRUT000000027055	ENSTRUT000000031126	ENSTRUT000000042535	ENSTRUT000000013831
<i>Tarsius syrichta</i>	n/a	ENSTSYT000000003101	ENSTSYT000000002061	ENSTSYT000000008053	ENSTSYT000000008770	
<i>Terpsiphone batesi</i>	JF1488			*		
<i>Terrapene carolina</i>	MVZ241538					
<i>Testudo graeca</i>	MVZ234282				*	
<i>Testudo hermanni</i>	MVZ244866		*	*	*	*
<i>Tetraodon nigroviridis</i>	n/a	ENSTNIT00000015950	ENSTNIT000000005278	ENSTNIT00000017439	ENSTNIT000000021886	ENSTNIT00000016925
<i>Trachemys scripta</i>	MVZ241524					
<i>Tupaia belangeri</i>	n/a	ENSTBET00000016104	ENSTBET000000004038	ENSTBET000000002631	ENSTBET000000012479	ENSTBET000000013183
<i>Tursiops truncatus</i>	n/a	ENSTTRT00000015522	ENSTTRT00000017021	ENSTTRT000000003175	ENSTTRT000000003311	ENSTTRT000000003199
<i>Varamus yemenensis</i>	MVZ236610					
<i>Vicugna pacos</i>	n/a	ENSVPAT000000003663		ENSVPAT00000012346	ENSVPAT000000002958	ENSVPAT000000009669
<i>Xantusia vigilis</i>	MVZ236373					
<i>Xenopeltis unicolor</i>	MVZ239387			*	*	
<i>Xenopus tropicalis</i>	n/a	ENSXETT000000054036	ENSXETT000000002268	ENSXETT000000054669	ENSXETT0000000028257	ENSXETT0000000033517

Scientific Name	Specimen #	UBE2V2	VDAC1	VPS29	WTAP	XPOT
<i>Accipiter cooperii</i>	MVZ181729					
<i>Agelaiocercus kingi</i>	MVZ183274					
<i>Agrionemys borsfieldii</i>	MVZ245929					
<i>Aluthe castanea</i>	JF1706	*				
<i>Alligator mississippiensis</i>	n/a					
<i>Alligator mississippiensis</i>	MVZ241485					
<i>Anas cyanoptera</i>	MVZ179315					
<i>Anniella pulchra</i>	MVZ247492	*	*			
<i>Anolis carolinensis</i>	n/a	ENSACAT00000012314	ENSACAT00000002834	ENSACAT00000001835	ENSACAT00000004475	ENSACAT00000003191
<i>Apalone ferox</i>	MVZ241535				*	
<i>Aptenodytes patagonicus</i>	MVZ178921					
<i>Aspidoscelis tigris</i>	MVZ243345					
<i>Bipes biporus</i>	MVZ236260					
<i>Bos taurus</i>	n/a	ENSBAT00000054236	ENSBAT00000053816	ENSBAT00000011076	ENSBAT00000010484	ENSBAT00000038585
<i>Callithrix jacchus</i>	n/a	ENSCJAT00000026284	ENSCJAT00000026431	ENSCJAT00000018052	ENSCJAT00000022843	ENSCJAT00000028819
<i>Canis familiaris</i>	n/a		ENSACFT00000037828	ENSACFT00000013463	ENSACFT00000001077	ENSACFT00000000554
<i>Casuarinus casuarinus</i>	MVZ178927					
<i>Cavia prellus</i>	n/a	ENSCPOT00000014558	ENSCPOT00000025252	ENSCPOT00000000537	ENSCPOT00000001690	ENSCPOT00000021044
<i>Chamaeleo senegalensis</i>	MVZ249659					
<i>Chelodina longicollis</i>	MVZ230488					
<i>Chelonia mydas</i>	MVZ233240					
<i>Chelonoidis denticulata</i>	MVZ175391					
<i>Chelydra serpentina</i>	MVZ250691	*		*		
<i>Choloepus hoffmanni</i>	n/a	ENSCHOT00000008699	ENSCHOT00000005647	ENSCHOT00000001234	ENSCHOT00000008285	ENSCHOT00000000290
<i>Chrysomys picta</i>	MVZ241506	*	*	*	*	
<i>Claudius angustatus</i>	MVZ164771	*	*			
<i>Cuora galbinifrons</i>	MVZ230933					
<i>Cuora pani</i>	MVZ230513					
<i>Cyrtodactylus</i>	JAM8710	*	*		*	
<i>Danio rerio</i>	n/a	ENDSART00000026578	ENDSART00000066373	ENDSART00000101246		ENDSART00000082119
<i>Dasylops novemcinctus</i>	n/a		ENDSDNOT00000002681	ENDSDNOT00000024756	ENDSDNOT00000018405	
<i>Dendropicus gabonensis</i>	JF1513					
<i>Dermatemys mawii</i>	MVZ233618					
<i>Dermochelys coriacea</i>	MVZ149844					
<i>Dibamus sp.</i>	MVZ224113	*				
<i>Dierochelys reticularia</i>	MVZ230923	*	*	*	*	
<i>Dipodomys ordii</i>	n/a		ENDSDORT00000000714	ENDSDORT00000002754	ENDSDORT00000011502	ENDSDORT00000000258
<i>Draco beccarii</i>	JAM8961	*	*		*	
<i>Draco sumatranus</i>	SZL71	*	*			
<i>Echinops telfairi</i>	n/a	ENSETET00000015454	ENSETET00000016336	ENSETET00000006361	ENSETET00000015347	ENSETET00000002983
<i>Emydura subglobosa</i>	MVZ230496					
<i>Emys marmorata</i>	MVZ164996					
<i>Ensatina eschscholtzii</i>	RP158					
<i>Equus caballus</i>	n/a		ENSECAT00000013217	ENSECAT00000009127	ENSECAT00000003936	ENSECAT00000019772

Scientific Name	Specimen #	UBE2V2	VDAC1	VPS29	WTAP	XPOT
<i>Ereimochelys imbricata</i>	MVZ233242					
<i>Erinaceus europaeus</i>	n/a	ENSEEUT00000014690	ENSEEUT00000005431	ENSEEUT00000004534	ENSEEUT00000006774	ENSEEUT00000009763
<i>Fagelaryia limnocharis</i>	MVZ251769					
<i>Felis catus</i>	n/a	ENSFCAT00000007602	ENSFCAT00000008423	ENSFCAT00000004136	ENSFCAT00000002844	ENSFCAT000000012364
<i>Gallus gallus</i>	n/a	ENSGALT000000024586	ENSGALT000000010508	ENSGALT00000007214	ENSGALT000000019061	ENSGALT000000015988
<i>Gasterosteus aculeatus</i>	n/a	ENSGACT000000023737	ENSGACT000000027664	ENSGACT00000006246		
<i>Geochelone sulcata</i>	MVZ238888	*		*		
<i>Geomyda spengleri</i>	MVZ257081					
<i>Geotrypetes seraphini</i>	MVZ252475					
<i>Gorilla gorilla</i>	n/a	ENSGGOT00000009486	ENSGGOT000000010920	ENSGGOT00000006432	ENSGGOT00000006069	
<i>Hardella thurjii</i>	MVZ248392		*			
<i>Heloderma horridum</i>	MVZ238489					
<i>Heteronotia binoei</i>	MKF226	*	*		*	*
<i>Homo sapiens</i>	n/a	ENST00000324746	ENST00000395047	ENST00000447578	ENST00000358372	ENST00000332707
<i>Hynobius arisanensis</i>	MVZ251776					
<i>Icthyophis bannanicus</i>	MVZ226265	*	*			
<i>Indotestudo elongata</i>	MVZ234627					
<i>Kachuga smithii</i>	MVZ248394		*			
<i>Kinixys belliana</i>	MVZ234631					
<i>Kinosternon scorpioides</i>	MVZ247584		*			
<i>Leptotyphlops dulcis</i>	MVZ236621					
<i>Leuccephalon yuwonoi</i>	MVZ241499					
<i>Lissemys punctata</i>	MVZ248398					
<i>Loxodonta africana</i>	n/a	ENSLAFT000000017748		ENSLAFT00000006175	ENSLAFT00000007334	ENSLAFT00000009829
<i>Macaca mulatta</i>	n/a	ENSMMUT000000017133	ENSMMUT000000022810	ENSMMUT000000045605	ENSMMUT000000018013	ENSMMUT000000010730
<i>Macropus eugenii</i>	n/a	ENSMEUT000000016633		ENSMEUT000000011581	ENSMEUT000000015459	ENSMEUT000000016086
<i>Malaclemys terrapin</i>	MVZ257705	*	*	*	*	*
<i>Malacocheirus tornieri</i>	MVZ234632					
<i>Malayemys subtrijuga</i>	MVZ230474		*		*	
<i>Mauremys caspica</i>	MVZ237439					
<i>Mauremys mutica</i>	MVZ230487					
<i>Mesalina brevirostris</i>	MVZ246016	*				
<i>Microcebus murinus</i>	n/a	ENSMICT000000009812	ENSMICT000000014716	ENSMICT00000005498	ENSMICT00000007974	ENSMICT000000012009
<i>Monodelphis domestica</i>	n/a	ENSMODT000000012991	ENSMODT000000018029	ENSMODT000000014756	ENSMODT000000009313	
<i>Mus musculus</i>	n/a	ENSMUST000000096234	ENSMUST0000000109065	ENSMUST000000031418	ENSMUST00000007007	ENSMUST000000039810
<i>Myotis lucifugus</i>	n/a	ENSMLUT000000003231	ENSMLUT0000000012469	ENSMLUT000000015531	ENSMLUT000000000722	ENSMLUT000000009699
<i>Nectarinia olivacea</i>	JF1603					
<i>Nerodia rhombifer</i>	MVZ246092					
<i>Ochotona princeps</i>	n/a	ENSOPRT000000010960	ENSOPRT000000006607	ENSOPRT000000001003	ENSOPRT000000003726	ENSOPRT000000013092
<i>Ornithorhynchus anatinus</i>	n/a	ENSOANT000000011570	ENSOANT0000000021286	ENSOANT000000007569	ENSOANT000000019058	ENSOANT000000017327
<i>Oryzotylagus cuniculus</i>	n/a	ENSOCUT000000006931	ENSOCUT000000009371	ENSOCUT000000011920	ENSOCUT000000013825	ENSOCUT000000012131
<i>Oryzias latipes</i>	n/a	ENSORLT0000000021848	ENSORLT000000001217	ENSORLT0000000022448		
<i>Oscacilia ochrocephala</i>	MVZ232864	*	*			
<i>Otolemur garnettii</i>	n/a			ENSOGAT000000006986	ENSOGAT000000007314	ENSOGAT000000004957
<i>Pan troglodytes</i>	n/a	ENSPTRT000000037492	ENSPTRT0000000031908	ENSPTRT000000009992	ENSPTRT0000000034633	ENSPTRT0000000057507

Scientific Name	Specimen #	UBE2V2	VDAC1	VPS29	WTAP	XPOT
<i>Pelodiscus sinensis</i>	MVZ226472					
<i>Pelomedusa subrufa</i>	MVZ236628					
<i>Phrynosoma sp.</i>	MVZ247578					
<i>Pipa pipa</i>	MVZ247516					
<i>Platenys platycephala</i>	MVZ247579					
<i>Platysternon megacephalum</i>	MVZ230486					
<i>Plestiodon gilberti</i>	MVZ249915		*			
<i>Pongo pygmaeus</i>	n/a	ENSPPYT00000021664		ENSPPYT00000005864	ENSPPYT00000019954	ENSPPYT00000005595
<i>Procarina capensis</i>	n/a	ENSPCAT00000000286	ENSPCAT00000010045	ENSPCAT00000011995	ENSPCAT00000004229	ENSPCAT00000001631
<i>Pseudomys concinna</i>	MVZ241520					
<i>Pteropus vampyrus</i>	n/a	ENSPVAT00000013812	ENSPVAT00000012423	ENSPVAT00000003506	ENSPVAT00000005304	ENSPVAT00000008205
<i>Rana livida</i>	MVZ241436					
<i>Rattus norvegicus</i>	n/a	ENSRNOT000000040575	ENSRNOT000000008477	ENSRNOT000000066290	ENSRNOT000000025936	ENSRNOT000000009645
<i>Rhinemys floridana</i>	MVZ233342					
<i>Rhinoclemmys pulcherrima</i>	MVZ230924		*		*	
<i>Sacalia quadriocellata</i>	MVZ230485					
<i>Sorex araneus</i>	n/a	ENSSART000000006973	ENSSART000000008215	ENSSART00000000549	ENSSART000000000919	ENSSART000000005826
<i>Spermophilus tridecemlineatus</i>	n/a	ENSSTOT00000014317	ENSSTOT00000007758	ENSSTOT00000011996		ENSSTOT000000008276
<i>Sternotherus carinatus</i>	MVZ238123		*			
<i>Sternotherus minor</i>	MVZ241528		*		*	
<i>Stigmochelys pardalis</i>	MVZ241333					
<i>Sus scrofa</i>	n/a		ENSSSCT00000015619	ENSSSCT00000010766	ENSSSCT000000004478	ENSSSCT00000000504
<i>Sylvietta virens</i>	JF1484	*				
<i>Taeniopygia guttata</i>	n/a	ENSTGUT00000011367	ENSTGUT0000001351	ENSTGUT00000016168	ENSTGUT00000011878	ENSTGUT000000006723
<i>Takifugu rubripes</i>	n/a	ENSTRUT00000017183	ENSTRUT00000018986	ENSTRUT00000016000		ENSTRUT000000025471
<i>Tarsius syrichta</i>	n/a		ENSTSYT00000007347	ENSTSYT00000012862	ENSTSYT00000001937	ENSTSYT000000008848
<i>Terpsiphone batesi</i>	JF1488					
<i>Terrapene carolina</i>	MVZ241538					
<i>Testudo graeca</i>	MVZ234282		*			
<i>Testudo hermanni</i>	MVZ244866		*		*	
<i>Tetraodon nigroviridis</i>	n/a	ENSTNIT00000011527	ENSTNIT00000019409	ENSTNIT00000015568		
<i>Trachemys scripta</i>	MVZ241524					
<i>Tupaia belangeri</i>	n/a	ENSTBET00000015711	ENSTBET00000003846	ENSTBET00000012747	ENSTBET00000003033	ENSTBET000000009774
<i>Tursiops truncatus</i>	n/a	ENSTTRT000000008322	ENSTTRT00000016856	ENSTTRT00000015088	ENSTTRT00000010370	ENSTTRT00000016807
<i>Varamus yemenensis</i>	MVZ236610					
<i>Vicugna pacos</i>	n/a		ENSVPAT00000010450	ENSVPAT00000008407	ENSVPAT000000008205	ENSVPAT000000004243
<i>Xantusia vigilis</i>	MVZ236373					
<i>Xenopeltis unicolor</i>	MVZ239387				*	
<i>Xenopus tropicalis</i>	n/a	ENSXETT00000010542	ENSXETT00000034818	ENSXETT00000052892	ENSXETT000000045788	ENSXETT00000055066

CHAPTER 3:
PHYLOGENETIC RELATIONSHIPS WITHIN TESTUDINES

ABSTRACT

The order Testudines (turtles) consists of 328 species in 14 families. Although there have been many recent studies focusing on resolving phylogenetic relationships within the group, several relationships remain ambiguous, such as the placement of Trionychia (softshell turtles), Chelonioidea (sea turtles), Chelydridae (North American snapping turtles), and *Platysternon* (Asian big-headed turtle). I address these outstanding relationships, as well as calculate divergence dates, by including 38 species (11/14 families) for 75 independent, protein-coding markers. Some of my phylogenetic results support previous hypotheses, such as reciprocally monophyletic Pleurodira (side-neck turtle) and Cryptodira (hidden-neck turtle), Trionychia (softshell turtles) as the basal cryptodire, Chelydridae and *Platysternon* having independently evolved similar morphologies, and Chelydridae as the sister taxon of Kinosternoidea. However, some phylogenetic relationships recovered in my analyses are novel, such as Cheloniidae (sea turtles) being the second-most basal cryptodire group, and *Platysternon* as the sister taxon of Testuguria. For divergence dating analyses, I find the origin of turtles to be in the Late Jurassic/Early Cretaceous. This is largely a result of younger fossil calibrations, which re-identify stem cryptodires to be stem turtles. Additional taxon sampling using these markers may help in further resolving relationships within turtles.

INTRODUCTION

Turtles are believed to have evolved >200 Ma in the Late Triassic (Gaffney 1990, Li et al. 2008), with the basal-most divergence in crown turtles being in the Late Jurassic (Joyce 2007). Carl Linnaeus provided the first systematic list of turtles in the first edition of “Systema Naturae”, naming four species (Linnaeus 1735), while the first four genera were later named (Brongniart 1800) and classified (Dumeril 1806). More than 200 years of research later, the number of recognized turtle species has risen to 328 placed in 14 families (Turtle Taxonomy Working Group 2010).

The order Testudines is divided into two major clades based on the mechanics of head retraction: Cryptodira (hidden-neck turtles) and Pleurodira (side-neck turtles). The reciprocal monophyly of these two clades is generally well-supported by both morphological (Gaffney & Meylan 1988, Shaffer et al. 1997) and molecular data (Shaffer et al. 1997, Cervelli et al. 2003, Fujita et al. 2004, Krenz et al. 2005, Near et al. 2005, Thomson & Shaffer 2009, Barley et al. 2010), although there are exceptions (Krenz et al. 2005, Sterli 2010).

Despite the fact that many molecular phylogenetic studies focus on turtle relationships, there is still uncertainty in the placement of a few turtle groups, including 1) Trionychia (softshell turtles), 2) Chelonioidea (sea turtles), 3) Chelydridae (North American snapping turtles), and 4) *Platysternon* (Asian big-headed turtle) (Figure 1). Poor taxon sampling and/or a modest number of molecular markers have limited these previous studies. Most studies have used a single molecular marker, but this has improved with the most recent publication (Barley et al. 2010) that included 19 taxa for 14 nuclear genes. My study more than doubles each of these figures, with 38 taxa and 75 new nuclear genes. The goals of this project are to resolve ambiguous relationships in the turtle phylogeny and determine the timing of divergence events between major turtle groups.

METHODS

Taxon and Marker Sampling

Initial taxon sampling for this project included representatives of all 14 turtle families. However, I was unable to include representatives for Podocnemidae and Carrotochylidae in analyses because of amplification and sequencing issues, probably due to RNA degradation of available samples. Additional taxa were removed from final analyses due to their instability in the phylogeny (i.e. rogue taxa, see below). Because most of the outstanding phylogenetic questions regarding within-turtle relationships are in regard to phylogenetic uncertainty in Cryptodira, taxon sampling was skewed heavily towards this group. Final taxon sampling included 38 taxa representing 11 of 14 families (Table 1), with chicken (*Gallus gallus*) as the outgroup, collected from the cDNA database on Ensembl (<http://www.ensembl.org>). All RAxML analyses were run on this dataset, while Bayesian-based analyses (MrBayes, BEAST) were run on a dataset with one fewer taxon (*Phrynopis* removed). Despite being placed in the correct clade (Pleurodira) in RAxML analyses and its membership in Pleurodira never being in question, *Phrynopis* was placed in Cryptodira in Bayesian analyses. This is most likely the result of *Phrynopis*'s instability in the phylogeny due to large amounts of missing data (74/75 genes missing), so we remove it from Bayesian analyses.

The 75 markers used in this study are from (Fong & Fujita submitted, Chapter 1). These markers are protein-coding (i.e. no introns) and were developed using EST databases, so PCR was performed on cDNA samples (instead of the standard genomic DNA). Samples were prepared as cDNA from tissues that had been appropriately preserved for RNA work, either stored in RNAlater (Ambion) or flash-frozen in liquid nitrogen. These preservation methods protect RNA by minimizing RNase activity and exposure (Sambrook & Russell 2001). Total RNA was extracted using a guanidinium thiocyanate-phenol-chloroform protocol (Chomczynski & Sacchi 1987; Sambrook & Russell 2001). Next, RNA extracts were reverse-transcribed to double-stranded cDNA using a SMART™ technology protocol (Clontech; Zhu et al. 2001). Resulting cDNA preps were used in place of genomic DNA extracts for standard PCR amplification. PCR products were purified using ExoSAP-IT (USB Corp.), cycle-sequenced using Big Dye Terminator v.3.1, and sequenced using an ABI 3730 automated sequencer. PCR conditions follow Fong & Fujita (submitted) and Chapter 1 of this dissertation.

Rogue Taxa Analyses

Rogue taxa (Wilkinson 1994, 1996) are taxa whose placement in a tree is unstable (being placed in multiple positions). In consensus/bootstrap analyses, rogue taxa create polytomies and reduce support values. The topological instability of rogue taxa is often because of missing data or heterogeneous rates of evolution between taxa (Sanderson & Shaffer 2002). The *a priori* identification of rogue taxa is difficult, but several *a posteriori* analyses have been developed. For example, a “leaf-dropping method” was developed by Pattengale et al. (2010) and is implemented in RAxML. In general terms, this methodology gives a score to taxa based on the impact of their removal. If the score of the tree is higher excluding a set of taxa, they are identified as rogue taxa and removed. A detailed description of the algorithm follows below.

I deviate from the terminology used in Pattengale et al. (2010) to use more common phylogenetics terminology; a “leaf” refers to a taxon and an “edge” refers to a branch. Bipartitions (the resulting two subsets of taxa remaining when a branch is removed) and relative information content (RIC) are in the basis of this methodology. There are two types of bipartitions, trivial and nontrivial. Trivial bipartitions are ones that separate a single taxon from the tree (i.e. a branch leading to a tip of the tree), whereas a nontrivial bipartition refers to the removal of an internal branch to create two subsets of taxa with membership >1 . The ratio of the total number of bipartitions (trivial + nontrivial) to the number of bipartitions in a fully resolved tree (no polytomies) is called the RIC. Since the numerator of the RIC includes trivial and nontrivial bipartitions, this methodology penalizes the removal of taxa (trivial bipartitions) and rewards for greater resolution of the tree (more nontrivial bipartitions) (Pattengale et al. 2010). This algorithm can be run on majority-rule and strict consensus trees.

Pattengale et al. (2010) uses a greedy heuristic algorithm that discards a taxon only if its removal provides an improvement in resolution that outweighs the loss of removing more taxa. Various dropsets (sets of removed taxa), are given a score; this score is the difference between the number of branches created and the number of taxa removed. The dropset with the highest score is used, and the corresponding taxa are removed. When there are multiple dropsets with the same highest score, one can opt to select the first dropset (which allows for repeatability) or use a random selection. The analysis iterates until no further benefit is gained from removing taxa.

Determining which rogue taxa to remove was a multi-step process that was run separately for each data set (representing the four data-types). First, a preliminary RAxML analysis was run with all taxa (45). Based on this phylogeny, taxa that were obviously in incorrect phylogenetic positions (i.e. tortoise placed in sea turtle clade) were removed. This was repeated until all remaining taxa were placed in the correct clades. Next, the bootstrap results from the RAxML analysis were used in rogue taxa analyses (Pattengale et al. 2010). The rogue taxa analysis was run 10 times using the “random” option to select a dropset for both the strict and majority rules consensus trees. To maximize the number of taxa retained and identify the most unstable taxa, taxa were considered rogue if they were identified in ≥ 5 of the 10 analyses under either strict or majority rules consensus. These taxa were removed and another RAxML analysis was run. Rounds of rogue taxon/RAxML analyses were repeated until there was no improvement in bootstrap values.

Phylogenetic Analyses

Based on analyses comparing the phylogenetic informativeness of different data-types (Fong & Fujita submitted), it was shown that the nucleotide (NUCL) data-type would be most appropriate for resolving relationships within turtles. Two different phylogenetic analyses were performed on the datasets, RAxML (Stamatakis 2006) and MrBayes (Ronquist & Huelsenbeck 2003). All analyses were run on the concatenated nucleotide dataset of the 75 genes and partitioned by gene and by codon (1st+2nd position, 3rd position), as this was found to be the most optimal partitioning strategy (Chapter 2).

RAxML analyses were run under the GTRGAMMA model of evolution for tree inference and for 1000 bootstrap replicates. RAxML v7.2.5 and v7.2.6 (Stamatakis 2006,

Stamatakis et al. 2008) analyses were run locally and on the CIPRES portal v2.2 (<http://www.phylo.org/portal2/>; Miller et al. 2009).

Models of data partitions for Bayesian analyses were determined using MrModeltest v2.3 (Nylander 2004). MrBayes v3.2.1 and v3.2 (Ronquist & Huelsenbeck 2003) analyses were run locally and using the BioHPC@CBSU resource at Cornell University (<http://cbsuapps.tc.cornell.edu>). All analyses were run with four chains for 10 million generations. Burn-in of MCMC chains were evaluated using the online program AWTY by evaluating cumulative plots of posterior probabilities of the 20 most variable splits (Wilgenbusch et al. 2004).

Divergence Dating

The program BEAST v1.6.1 was used for divergence dating analyses. Alignments were partitioned by gene, with empirical base frequencies, and models of evolution unlinked and set according to the models selected by MrModeltest v2 (Nylander 2004). Clock models for each gene partition were unlinked and set to a relaxed, uncorrelated lognormal clock. Tree priors were set to a speciation, Yule process, and the final tree from the Bayesian analysis was used as a starting tree. MCMC analyses were run locally for 300 million generations, and convergence was evaluated in Tracer v1.5 via plots of $-\ln L$ and effective sample size (ESS) scores >200. Plots of $-\ln L$ were used to determine the number of generations to remove as “burn-in”.

A survey of the scientific literature uncovered a list of reliable fossil calibrations for turtles (Table 2). Although studies such as Near et al. (2005) used as many as 21 calibrations, it has since been shown that many of these fossils are unreliable (Danilov & Parham 2006, 2008, Parham & Irmis 2008). In addition, new paleontological research have suggested a much younger origin of extant turtles (145.5-175.6 Ma; Danilov & Parham 2006, 2008, Joyce & Gauthier 2004, Joyce 2007) compared to the traditional 210 Ma (Gaffney 1990). Therefore, my set of calibrations is reduced to three reliable fossils including a younger calibration for the crown.

A prior on the date of the fossil calibration were modeled in BEAST using a truncated normal distribution, with 95% of the prior distribution being contained within the hard minimum and soft maximum (Benton & Donoghue 2007, Ho 2007).

RESULTS AND DISCUSSION

Taxon and Marker Sampling

A total number of 567 sequences were included in the dataset. Of these, 493 were of newly sequenced for the turtle ingroup (74 for the chicken outgroup). Individual taxa range in the amount of data collected, ranging between 1-71 genes (average 11 genes), and the dataset is 14.6% complete. The aligned dataset of the concatenated nucleotides was 34,110bp in length.

Rogue Taxa Analyses

The initial taxon sampling for the ingroup included 45 taxa, with representatives from 12 of 14 families. Seven taxa were identified as rogue taxa (*Cuora pani*, *Dermochelys coriacea*, *Emydura subglobosa*, *Indotestudo elongata*, *Kinixys belliana*, *Leucocephalon yuwonoi*, *Malacochersus*

tornieri), which resulted in the removal of another family (Dermochelyidae; Table 1). An additional taxon (*Phrynops*, Pleurodira) was removed from Bayesian analyses due to its placement into Cryptodira. We determine this placement to be erroneous because its membership in Pleurodira has never been questioned and RAxML analyses recover this relationship. This placement of *Phrynops* is probably due to large amounts of missing data (74/75 genes missing). The final datasets for RAxML and Bayesian analyses include 38 or 37 taxa, respectively, representing 11 of 14 turtle families.

Phylogenetic Analyses

Phylogenetic relationships recovered from RAxML and Bayesian analyses were almost identical, so only results from the RAxML analysis are shown (Figure 2). Additional phylogenetic analyses performed in BEAST recovered some contradictory relationships from those found in the RAxML/Bayes analyses, and are noted below (Figure 3).

A monophyletic Testudines (extant turtles) and a basal split between Pleurodira and Cryptodira were recovered with full support (100 bootstrap/1.0 posterior probability). Within Pleurodires, only representatives from two of the three recognized families were included. From the sequences available for Pleurodira, a paraphyletic Chelidae was recovered, but this result is questionable, as it is poorly supported (Figure 2), and results from BEAST analyses recovered a monophyletic Chelidae (Figure 3).

Because most of the outstanding phylogenetic questions regarding within-turtle relationships are in regard to phylogenetic uncertainty in Cryptodira, taxon sampling was skewed heavily towards this group. Below I address these issues in detail.

PHYLOGENETICS OF TRIONYCHIA

Trionychia (Joyce et al. 2004) comprises Trionychidae and Carettochelyidae (Meylan 1987, Engstrom et al. 2002, 2004). Trionychians were historically considered derived cryptodires (Gaffney & Meylan 1988, Shaffer et al. 1997), but later molecular studies suggested that they are basal within this assemblage (Chen et al. 1980, Shaffer et al. 1997, Wu et al. 1999, Fujita et al. 2004, Krenz et al. 2005, Near et al. 2005, Parham et al. 2006, Thomson et al. 2009, Barley et al. 2010). In addition, a study based on developmental timing found Trionychia to be sister to Emydidae+Testudinidae (Werneburg & Sanchez-Villagra 2009).

The lone representative of Carettochelyidae, *Carettochelys insculpta*, was not successfully amplified/sequenced for any of the new markers. However, the sister relationship between Trionychidae and Carettochelyidae is well supported (Iverson et al. 2007) so this does not prevent addressing the placement of Trionychia. Mirroring the majority of recent molecular studies (Chen et al. 1980, Shaffer et al. 1997, Wu et al. 1999, Fujita et al. 2004, Krenz et al. 2005, Near et al. 2005, Parham et al. 2006, Thomson et al. 2009, Barley et al. 2010), Trionychidae was inferred to be the most basal cryptodire group (Figure 2). With several independent studies recovering this same relationship of a basal Trionychia, this appears to be a reliable result. The placement of Trionychidae (or, more likely Trionychia) as the sister taxon of all other cryptodires then renders the traditional clade Trionychoidea (Trinoychia+Kinosternoidea) invalid.

Trionychidae is further subdivided into two subclades, Trionychinae and Cycloorbinae. These two clades have been shown to be monophyletic (Engstrom et al. 2004). Although taxon sampling was limited, representatives of both trionychid subclades were included in this study, and a sister relationship between the two was recovered (Figure 2).

PHYLOGENETIC PLACEMENT OF CHELONIOIDEA

Extant Chelonioidea comprises two lineages of sea turtles, Cheloniidae and *Dermochelys* (Gaffney & Meylan 1988). Relationships within sea turtles are strongly supported (Dutton et al. 1996, Bowen & Karl 1997, Parham & Fastovsky 1997), but placement of Chelonioidea is uncertain. Chelonioidea has been placed as the sister group to Kinosternoidea (Shaffer et al. 1997; Cervelli et al. 2003, Fujita et al. 2004), Chelydridae (Shaffer et al. 1997; Wu et al. 1999, Krenz et al. 2005), Kinosternoidea+Chelydridae (Cervelli et al. 2003, Near et al. 2005, Parham et al. 2006 [in part], Barley et al. 2010), Testudinoidea (Parham et al. 2006 [in part]), or as the most basal cryptodire (Werneburg & Sanchez-Villagra 2009).

Dermochelys coriacea was successfully sequenced, but was identified as a rogue taxon and removed from the final phylogenetic analyses. However, there is strong support for a monophyletic Chelonioidea that includes *Dermochelys* (Dutton et al. 1996, Bowen & Karl 1997, Parham & Fastovsky 1997, Iverson et al. 2007). Therefore, the cheloniid species in my study still allow me to address the placement of Chelonioidea. The placement of Cheloniidae was strongly supported as the second-most basal cryptodiran group, with Trionychia being the most basal (Figure 2). Previously, five different positions were inferred for Chelonioidea (Shaffer et al. 1997, Wu et al. 1999, Cervelli et al. 2003, Fujita et al. 2004, Krenz et al. 2005, Near et al. 2005, Parham et al. 2006, Werneburg & Sanchez-Villagra 2009, Barley et al. 2010), and this study proposes a novel, sixth position of sea turtles as the sister group of all cryptodires excluding softshell turtles. BEAST analyses, however, do not recover this position of sea turtles, but rather sister to Testudinoidea. Although BEAST analyses recover a previously proposed position of Chelonioidea, RAxML and Bayes analyses both strongly support this new relationship. So, the placement of Chelonioidea in the turtle phylogeny remains ambiguous, and further analysis with more complete taxon sampling may help in resolving this relationship.

PHYLOGENETIC PLACEMENT OF CHELYDRIDAE

Chelydridae are represented by two species, *Chelydra serpentina* and *Macrochelys temminckii*. Characteristic morphological features of chelydrids are a large head that cannot be retracted into the shell, strong jaws, and a reduced plastron (Ernst & Barbour 1992). Based on molecular data, Chelydridae has been placed in five different positions within Cryptodira in various studies, including sister to Chelonioidea (Krenz et al. 2005, Chandler & Janzen 2009), sister to Kinosternoidea (Cervelli et al. 2003, Near et al. 2005, Parham et al. 2006, Chandler & Janzen 2009, Barley et al. 2010), sister to Chelonioidea+Kinosternoidea (Shaffer et al. 1997, Cervelli et al. 2003), sister to Chelonioidea+Testudinoidea (Parham et al. 2006), and sister to Trionychia+Testudinoidea (Werneburg & Sanchez-Villagra 2009).

In my analyses, a single representative of the Chelydridae, *Chelydra serpentina*, was included. Based on the results from all analytical methods, Chelydridae was placed as the sister group to Kinosternoidea with moderate to strong support (75 bootstrap/0.97

posterior probability; Figure 2). This result is the same as that found in several independent studies (Cervelli et al. 2003, Near et al. 2005, Parham et al. 2006, Chandler et al. 2009, Barley et al. 2010), and should be the prevailing hypothesis of Chelydridae placement. The proposed name Chelydroidea (originally coined by Baur 1893) is recommended to refer to this chelydrid+kinosternid group (Knauss et al. 2010).

PHYLOGENETIC PLACEMENT OF *PLATYSTERNON*

Platysternon megacephalum, has been one of the most difficult turtle species to place in the turtle phylogeny, first being placed as the sister group to Chelydridae (and, indeed, treated as a member of this clade) based on morphology (Gaffney 1975, Shaffer et al. 1997, Brinkman & Wu 1999) and DNA (Shaffer et al. 1997). However, affinities of *Platysternon* to North American snapping turtles has been refuted by studies showing *Platysternon* to be closely related to some combination of Emydidae, Geoemydidae, and Testudinidae (morphology: Valliant 1894, Williams 1950, Danilov 1998; chromosomes: Bickham & Carr 1983; molecules: Cervelli et al. 2003, Krenz et al. 2005, Near et al. 2005, Parham et al. 2006, Thomson et al. 2009, Barley et al. 2010). Of these molecular studies, some placed *Platysternon* sister to Testudinoidea (Krenz et al. 2005, Near et al. 2005) or within Testudinoidea as sister to Emydidae (Parham et al. 2006, Thomson et al. 2009, Barley et al. 2010).

My results suggest a sister relationship between *Platysternon* and Testuguria (Geoemydidae+Testudinidae) with good support (85 bootstrap/0.99 posterior probability; Figure 2). This is a novel position of *Platysternon* from molecular data, although it has been previously suggested by morphologists (Valliant 1894). With several molecular studies presenting strong evidence for different placements of *Platysternon*, it is unclear which is the true evolutionary history. However, we can be relatively certain that *Platysternon* is not closely related to Chelydridae, and the “snapping-turtle morphology” (big head and long tail) evolved independently in two turtle lineages. Also, although the placement of *Platysternon* is perhaps not definitively resolved, it seems safe to say that *Platysternon* is a testudinoid.

OTHER RELATIONSHIPS

There are additional findings in this study that should be noted. First, within Emydidae, the monophyly of the two subclades, Dierochelyinae and Emydinae are uncertain, as RAxML/Bayesian analyses did not recover reciprocally monophyletic groups (Fig. 2), whereas BEAST analyses did (Fig. 3).

Next, the group Kinosternoidea (Dermatemydidae+Kinosternidae) was highly supported (bootstrap=97) as the sister group to Testudinoidea (Emydidae+Geoemydidae+Testudinidae) and *Platysternon* (Figure 2), although the placement was different in BEAST analyses (Figure 3). Within this group, Kinosternidae was paraphyletic in RAxML and BEAST analyses, with Staurotipinae being more closely related to Dermatemydidae, but with low support (bootstrap=60). Bayesian analyses recovered a polytomy with the three constituents being Kinosternidae, *Dermatemys*, and *Claudius* (posterior probability=1).

Also, the clade Testuguria (Testudinidae+Geoemydidae) was recovered in this analysis, but with low support (Fig. 2). Within Testuguria, however, a monophyletic Testudinidae was nested within Geoemydidae. Previous studies raised the question whether Testudinidae was nested within or the sister group of Geoemydidae (Shaffer 2009). My results are inconclusive

regarding this issue, as Geoemydidae was recovered as paraphyletic, but with low support. It should be noted that BEAST analyses also recovered a paraphyletic Geoemydidae in relation to Testudinidae due to the position of a single taxon (*Sacalia*) (Figure 3). However, in this phylogeny, the position of *Sacalia* seems erroneous, since a previous study with increased taxon sampling and more complete dataset firmly placed *Sacalia* within a monophyletic Geoemydidae (Spinks et al. 2004). Therefore, BEAST analyses are recovering reciprocally monophyletic Testudinidae and Geoemydidae.

Lastly, Geoemydidae has traditionally contained two subclades, Batagurinae and Geoemydinae, although the monophyly of these traditional groups as discounted by Spinks et al. (2004). A recent study has proposed a third subclade, Rhinoclemminae for the sole new world genus (Le and McCord 2008). My study did not recover the monophyly of Batagurinae and Geoemydinae, however support values within Geoemydidae were all low. The monophyly of Rhinoclemminae could not be evaluated, as only a single representative was included in this study.

Divergence Dating

Plots of $-lnL$ and ESS scores indicated that divergence dating analyses reached stationarity (results not shown). All ESS scores except one (Chelydridae-Kinosternoid clade, ESS=144) were much greater than 200. Results can be found in Table 3 and Figure 3. Based on these results, it seems that turtles evolved in the Middle to Late Jurassic (166.2-145.5 Ma), and the two major groups (Pleurodira, Cryptodira) arose by the Cretaceous (Figure 3). By the Paleogene, all the major turtle families were already present. These results were compared to those from the most recent divergence dating study on turtles (Near et al. 2005; Table 3). All inferred dates from Near et al. (2005) did not fall with the 95% HPD intervals from this study (Table 3). In addition, except for the clade for which the same calibration point was used (Testuguria), all inferred dates from this study are younger than dates recovered in Near et al. (2005), with some dates being >50 my younger (Table 3).

There is still much controversy regarding identification and use of turtle fossils in molecular divergence dating analyses (see Parham & Irmis 2008, Near et al. 2008). In particular, there has been recent controversy in the classification of basal turtles. Traditional turtle classification has dated the origin of crown group turtles to the Late Triassic (~210 Ma; Gaffney 1990, 1990, Gaffney & Jenkins 2010), but other studies find a much younger origin of the crown (~150 Ma; Joyce & Gauthier 2004, Danilov & Parham 2006, 2008, Joyce 2007, Sterli & Joyce 2007, Anquetin et al. 2008, Sterli 2008, Joyce & Sterli 2010); the former researchers classify a set of fossils to be stem cryptodires, while the latter researchers believe these same fossils are stem turtles. This study is the first to utilize these younger calibration dates. In comparing my results to those from Near et al. (2005), it can be seen how important correct paleontological information is, as classification of these fossils as either stem cryptodires or stem turtles greatly influences molecular divergence dating outcomes, changing divergence dates by more than 50 my and suggesting substantially more recent ages of origins for major turtle clades.

CONCLUSION

Although relatively species poor, the relationships within turtles have been difficult to resolve. This study advances turtle systematics by combining a broad sampling of taxa covering all but three of the fourteen families of turtles with the largest set of independent, molecular markers to date (75 genes, over 34,000bp). Of the unresolved relationships between the turtle clades, these results support some previous hypotheses, such as Trionychia (softshell turtles) as the basal cryptodiran lineage, Chelydridae and *Platysternon* (snapping turtles) independently evolving similar morphologies, and Chelydridae as sister to Kinosternoidea. However, some phylogenetic relationships recovered in my analyses are novel, such as Cheloniidae (sea turtles) being the second-most basal cryptodire group, and *Platysternon* sister to Testuguria.

For accurate estimates of divergence dates, both accurate molecular phylogenies and fossil classifications, in addition to analytical methods to account for rate heterogeneity and uncertainties introduced by the fossil record, are necessary. Although tremendous strides have been made in analytical methods of divergence dating from the days of strict molecular clocks, our estimates of divergences dates are only as accurate as the fossil calibrations we use. This is the first study to use data re-identifying stem cryptodires to be stem turtles, resulting in shifting the origin of turtles 50 my younger to the late Jurassic, instead of the traditional Triassic origin hypothesis (~210 my).

This study provides an excellent database of independent, molecular markers to study turtle relationships. Future work should focus on reducing the amount of missing data in the existing data matrix, expanding taxon sampling, and including more fossil calibration points for divergence dating analyses.

ACKNOWLEDGEMENTS

I would like to thank Drs. Carol Spencer and Jim McGuire (Museum of Vertebrate Zoology, UC Berkeley) for tissue loans. Jeffrey Boore, Richard Harland, Jim McGuire, and Jim Parham helped in reviewing early drafts of this chapter. I would also like to thank Adam Leache, Matt Fujita, Matt Brandley, and Jim Parham for help with analyses. Funding for this project came from the National Science Foundation Doctoral Dissertation Improvement Grant (#08-564) and the Museum of Vertebrate Zoology.

REFERENCES

- Anquetin J, Barrett PM, Jones MEH, Moore-Fay S, Evans SE (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, Spinks PQ, Thomson RC, Shaffer, HB (2010) Fourteen nuclear genes provide phylogenetic resolution for difficult nodes in the turtle tree of life. *Molecular Phylogenetics and Evolution* 55: 1189-1194.
- Baur G (1893) Notes on the classification of Cryptodira. *American Naturalist* 27: 672-675.

- Benton MJ, Donoghue PCJ (2007) Paleontological evidence to date the tree of life. *Molecular Biology and Evolution* 24: 26-53.
- Bickham JW, Carr JL (1983) Taxonomy and phylogeny of the higher categories of cryptodiran turtles based on a cladistic analysis of chromosomal data. *Copeia* 1983: 918-932.
- Bowen BB, Karl SA (1997) Population genetics, phylogeography, and molecular evolution. In: Lutz PL, and Music JA (Eds.). *The Biology of Sea Turtles*. Boca Raton, Florida: CRC Press, pp. 29-50.
- Brinkman DB, Wu XC (1999) The skull of *Ordosemys*, an Early Cretaceous turtle from Inner Mongolia, People's Republic of China, and the interrelationships of Eucryptodira (Chelonia, Cryptodira). *Paludicola* 2: 134-147.
- Brongniart A (1800) Essai d'une classification naturelle des reptiles, par le citoyen. Seconde Partie. Formation et disposition des genres. *Bulletin des Sciences, par la Societe Philomatique de Paris* 2: 89-91.
- Cervelli M, Olivero 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.
- Chander CH, Janzen FJ (2009) The phylogenetic position of snapping turtles (Chelydridae) based on nucleotide sequence data. *Copeia* 2009: 209-213.
- Chen BY, Mao SH, Ling YH (1980) Evolutionary relationships of turtles suggested by immunological cross-reactivity of albumins. *Comparative Biochemistry and Physiology* 66B: 421-425.
- Chomczynski P, Sacchi N (1987) Single-step method of RNA isolation by acid guanadinium thiocyanate-phenol-chloroform extraction. *Analytical Biochemistry* 162: 156-159.
- Danilov IG (1998) Phylogenetic relationships of platysternid turtles [abstract]. Third Asian Herpetological Meetings.
- Danilov IG, Parham JF (2006) A redescription of 'Plestiochelys' 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 IG, Parham JF (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.

- Dumeril AMC (1806) Zoologie analytique ou Méthode naturelle de Classification des Animaux, rendue plus facile à l'aide de Tableaux synoptiques: I-XXXIII. Paris: Allais Libraire, 344p.
- Dutton PH, Davis SK, Guerra T, Owen D (1996) Molecular phylogeny for marine turtles based on sequences of the ND4-Leucine tRNA and control regions of mitochondrial DNA. *Molecular Phylogenetics and Evolution* 5: 511-521.
- Ernst CH, Barbour RW (1992) *Turtles of the World*. Smithsonian Institution Scholarly Press. 328p.
- Engstrom TN, Shaffer HB, McCord WP (2002) Phylogenetic diversity of endangered and critically endangered southeast Asian softshell turtles (Trionychidae: *Chitra*). *Biological Conservation* 104: 173-179.
- 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.
- Fong JJ & Fujita MK (submitted) Evaluating phylogenetic informativeness and data-type usage for new protein-coding genes across vertebrata. Submitted to *Molecular Phylogenetics and Evolution*.
- 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 (1975) A phylogeny and classification of the higher categories of turtles. *Bulletin of the American Museum of Natural History* 155: 391-436.
- Gaffney ES (1990) The comparative osteology of the Triassic turtle *Proganochelys*. *Bulletin of the American Museum of Natural History* 194: 1-263.
- Gaffney ES, Meylan PA (1988) A phylogeny of turtles. In: Benton, MJ (Ed.), *The phylogeny and classification of tetrapods*. Clarendon Press, Oxford, pp. 157-219.
- Gaffney ES, Jenkins FA (2010) The cranial morphology of *Kayentachelys*, an Early Jurassic cryptodire, and the early history of turtles. *Acta Zoologica* 91: 335-368.
- Ho SYW (2007) Calibrating molecular estimates of substitution rates and divergence times in birds. *Journal of Avian Biology* 38: 409-414.
- Iverson JB, Brown RM, Akre TS, Near TJ, Le, M, Thomson RC, Starkey DE (2007) In search of the tree of life for turtles. *Chelonian Research Monographs* 4: 85-106.

- Joyce WG, Gauthier JA (2004) Palaeoecology of Triassic stem turtles sheds new light on turtle origins. *Proceedings of the Royal Society of London B* 271: 1-5.
- 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: 989-10113.
- Joyce W (2007) Phylogenetic relationship of Mesozoic turtles. *Bulletin of the Peabody Museum of Natural History* 48: 3-102.
- Joyce W, Sterli J (in press) Congruence, non-homology, and the phylogeny of basal turtles. *Acta Zoologica*. doi: 10.1111/j.1463-6395.2010.00491.x
- Knauss GE, Joyce WG, Lyson TR, Pearson D (2010) A new kinosternoid from the Late Cretaceous Hell Creek Formation of North Dakota and Montana and the origin of the *Dermatemys mawii* lineage. *Palaontologische Zeitschrift* DOI: 10.1007/s12542-010-0081-x
- Krenz JG, Naylor GJP, Shaffer HB, Janzen FJ (2005) Molecular phylogenetics and evolution of turtles. *Molecular Phylogenetics and Evolution* 37: 178-191.
- Le M, McCord WP (2008) Phylogenetic relationships and biogeographical history of the genus *Rhinoclemmys* Fitzinger, 1835 and the monophyly of the turtle family Geoemydidae (Testudines: Testudinoidea). *Zoological Journal of the Linnean Society* 153: 751-767.
- Li C, Wu X, Rieppel O, Wang L, Zhao L (2008) An ancestral turtle from the Late Triassic of southwestern China. *Nature* 456: 497-501.
- Linnaeus C (1735) *Systema Naturae* 1st edition, Leiden, T. Haack, 13 unnumbered pp.
- Meylan PA (1987) The phylogenetic relationships of soft-shelled turtles (family Trionychidae). *Bulletin of the American Museum of Natural History* 186: 1-101.
- Miller MA, Holder MT, Vos R, Midford PE, Liebowitz T, Chan L, Hoover P, Warnow T (2009) The CIPRES Portals. CIPRES. 2009-08-04. URL: http://www.phylo.org/sub_sections/portal. Accessed: 2009-08-04. (Archived by WebCite(r) at <http://www.webcitation.org/5imQlJeQa>)
- Near TJ, Meylan PA, Shaffer HB (2005) Assessing concordance of fossil calibration points in molecular clock studies: an example using turtles. *The American Naturalist* 165: 137-146.
- Near TJ, Meylan PA, Shaffer HB (2008) Caveats on the use of fossil calibrations for molecular dating: a reply to Parham & Irmis. *The American Naturalist* 171: 137-140.

- Nylander JAA (2004) MrModeltest v2. Program distributed by the author. Evolutionary Biology Centre, Uppsala University, Uppsala.
- Parham JF, Fastovsky DE (1997) The phylogeny of cheloniid sea turtles revisited. *Chelonian Conservation and Biology* 2: 548-554.
- 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.
- Parham JF, Irmis RB (2008) Caveats on the use of fossil calibrations for molecular dating: a comment on Near et al.. *The American Naturalist* 171: 132-136.
- Pattengale ND, Swenson KM, Moret BME (2010) Uncovering hidden phylogenetic consensus. *Bioinformatics Research and Applications* 6053: 128-139.
- Ronquist F, Huelsenbeck JP (2003) MrBayes 3: Bayesian phylogenetic inference under mixed models. *Bioinformatics* 19: 1572-1574.
- Sambrook J, Russell DW (2001) *Molecular cloning: a laboratory manual*. Cold Springs Harbor Laboratory Press. Cold Springs Harbor, New York.
- Sanderson MJ, Shaffer HB (2003) Troubleshooting molecular phylogenetic analyses. *Annual Review of Ecology and Systematics* 33: 49-72.
- Shaffer HB, Meylan P, McKnight ML (1997) Tests of turtle phylogeny: molecular, morphological, and paleontological approaches. *Systematic Biology* 46: 235-268.
- Shaffer HB (2009) Turtles (Testudines) in SB Hedges S Kumar (eds.) *The Timetree of Life*. Oxford University Press, Pp 398-401.
- Spinks PQ, Shaffer HB, Iverson JB, McCord WP (2004) Phylogenetic hypotheses for the turtle family Geoemydidae. *Molecular Phylogenetics and Evolution* 32: 164-182.
- 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.
- Stamatakis A (2006) RAxML-VI-HPC: maximum likelihood-based phylogenetic analyses with thousands of taxa and mixed models. *Bioinformatics* 22:2688-2690
- Stamatakis A, Hoover P, Rougemont J (2008) A rapid bootstrap algorithm for the RAxML web-servers. *Systematic Biology* 57:758-771

- 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.
- Thomson RC, Shaffer HB (2009) Sparse Supermatrices for Phylogenetic Inference: Taxonomy, Alignment, Rogue Taxa, and the Phylogeny of Living Turtles *Systematic Biology* 59: 1-17.
- Turtle Taxonomy Working Group (2009) Turtles of the World, 2010 update; List of taxonomy, synonymy, distribution, and conservation. *Chelonian Research Monographs* No.5 *doi:10.3854/crm.5.000.checklist.v2.2010*
- Vaillant L (1894) Essai sur la classification generale des cheloniens. *Annales des Sciences Naturelles* 16: 331-345.
- Werneburg I, Sanchez-Villagra MR (2009) Timing of organogenesis support basal position of turtles in the amniote tree of life. *BMC Evolutionary Biology* 9: 82.
- Wilkinson M (1994) Common cladistic information and its consensus representation: reduced Adams and reduced cladistic consensus trees and profiles. *Systematic Biology* 43: 343-368.
- Wilkinson M (1996) Majority-rule reduced consensus trees and their use in bootstrapping. *Molecular Biology and Evolution* 13: 437-444.
- Williams EE (1950) Variation and selection of the cervical central articulations of living turtles. *Bulletin of the American Museum of Natural History* 94: 505-562.
- Wu P, Zhou K-Y, Yang Q (1999) Phylogeny of Asian freshwater and terrestrial turtles based on sequence analysis of 12s rRNA gene fragment. *Acta Zoologica Sinica* 45: 260-267.
- Zhu YY, Machleder EM, Chenchik A, Li R, Siebert PD (2001) Reverse transcriptase template switching: a SMART approach to full-length cDNA library construction. *BioTechniques* 30: 892-897.

Table 1: Taxon sampling for this study. Numbers in parentheses represent (# taxa included in final analyses/#taxa removed as rogue taxa). Taxa in grey font are ones that were identified as rogue taxa and removed from final phylogenetic analyses (7 taxa total).

Testudines (38/7)	Pleurodira (4/1)			Chelidae (3/1)		<i>Chelodina longicollis</i>
						<i>Emydura subglobosa</i>
						<i>Phrynops</i> sp.
						<i>Platemys platycephala</i>
						<i>Pelomedusa subrufa</i>
				Pelomedusidae (1)		
				Podocnemidae (0)		
	Cryptodira (34/6)			Chelydridae (1)		<i>Chelydra serpentina</i>
				<i>Platysternon</i> (1)		<i>Platysternon megacephalum</i>
		Chelonoidea (2/1)	Cheloniidae (2)		<i>Chelonia mydas</i>	
					<i>Eretmochelys imbricata</i>	
			Dermochlyidae (0/1)		<i>Dermochelys coriacea</i>	
		Kinosternoidea (5)	Dermatemydidae (1)		<i>Dermatemys mawii</i>	
			Kinosternidae (4)	Kinosterninae (3)	<i>Kinosternon scorpioides</i>	
					<i>Sternotherus carinatus</i>	
					<i>Sternotherus minor</i>	
				Staurotypinae (1)	<i>Claudius angustatus</i>	
		Testudinoidea (22/5)		Emydidae (7)	Dierochelyinae (5)	<i>Chrysemys picta</i>
						<i>Dierochelys reticularia</i>
						<i>Malaclemys terrapin</i>
						<i>Pseudemys concinna</i>
						<i>Trachemys scripta</i>
					Emydinae (2)	<i>Emys marmorata</i>
						<i>Terrapene carolina</i>
						<i>Cuora galbinifrons</i>
						<i>Cuora pani</i>
						<i>Geoemyda spengleri</i>
					<i>Hardella thurii</i>	
					<i>Kachuga smithii</i>	
					<i>Leucocephalon yuwonoi</i>	
				<i>Malayemys subtrijuga</i>		
				<i>Mauremys caspica</i>		
			<i>Mauremys mutica</i>			
			<i>Rhinoclemys pulcherrima</i>			
			<i>Sacalia quadriocellata</i>			
		Testudinidae (6/3)		<i>Agrionemys horsfieldii</i>		
				<i>Chelonoidis denticulata</i>		
				<i>Geochelone sulcata</i>		
				<i>Indotestudo elongata</i>		
				<i>Kinixys belliana</i>		
				<i>Malacochersus tornieri</i>		
				<i>Stigmochelys pardalis</i>		
				<i>Testudo graeca</i>		
				<i>Testudo bermanni</i>		
	Trionychia (3)	Carettochelyidae (0)				
Trionychidae (3)		Cyclanorbinae (1)	<i>Lissemys punctata</i>			
		Trionychinae (2)	<i>Apalone ferox</i>			
				<i>Pelodiscus sinensis</i>		

Table 2: Table of the fossil calibrations used in divergence dating analyses. Minimum (min) and maximum (max) dates are in million years ago (Ma).

Node	Members	min	max	Taxon	References	Notes
Testudines	Cryptodira- Pleurodira	145.5	175.6	<i>Cbengynchelys baenoides</i> , <i>Xinjiangchelys</i>	Danilov & Parham 2006, 2008; Joyce 2004, 2007	Middle to late Jurassic
Cryptodira	Trionychia- Durocryptodira	99.6	175.6	<i>Yehguia tatsuensis</i>	Danilov & Parham 2006, 2008; Joyce 2004, 2007	Middle Jurassic to early Cretaceous
Testuguria	Testudinidae- Geoemydidae	53.1	65.5	<i>Hadrianus majusculus</i>	Hutchinson 1980	

Table 3: Results of divergence dating analyses from this study (95% highest posterior density [HPD], Mean). These results are compared to divergence dates recovered in the study by Near et al. (2005). In every case all of the dates from Near et al. (2005) were not contained within the 95% HPD interval, and in all but one case, the dates recovered in this study were much younger than those found in Near et al. (2005). Node numbers correspond to phylogeny in Figure 3. All dates shown are million years ago (Ma).

	Node	95% HPD	Mean	Near et al.	Δ
Testudines	1	145.5-166.2	153.51	210 (fixed)	-56.49
Cryptodira	2	114.1-149.4	132.08	174.87	-42.80
Pleurodira	3	82.8-152.3	117.06	176.62	-59.56
<i>Chelydra</i> - Kinosternoidea	4	55.2-73.5	64.20	n/a	n/a
<i>Platysternon</i> - Testuguria	5	53.2-59.1	55.65	n/a	n/a
Testuguria	6	53.1-55.7	53.96	52 (fixed)	1.96
Trionychidae	7	18.2-64.1	39.04	100 (fixed)	-60.96
Kinosternoidea	8	25.9-47.1	36.15	65 (fixed)	-28.85
Testudinidae	9	22.3-35.3	28.85	n/a	n/a
Cheloniidae	10	8.6-43.0	24.40	50.24	-25.84
Emydidae	11	11.2-26.4	18.28	34 (fixed)	-15.72
Kinosternidae	12	10.2-27.1	17.99	50 (fixed)	-32.01

Figure 1: Diagram of unresolved relationships in the order Testudines. Questions remain on the phylogenetic placement of A) *Platysternon*, B) Trionychia, C) Chelydridae, and D) Chelonoidea.

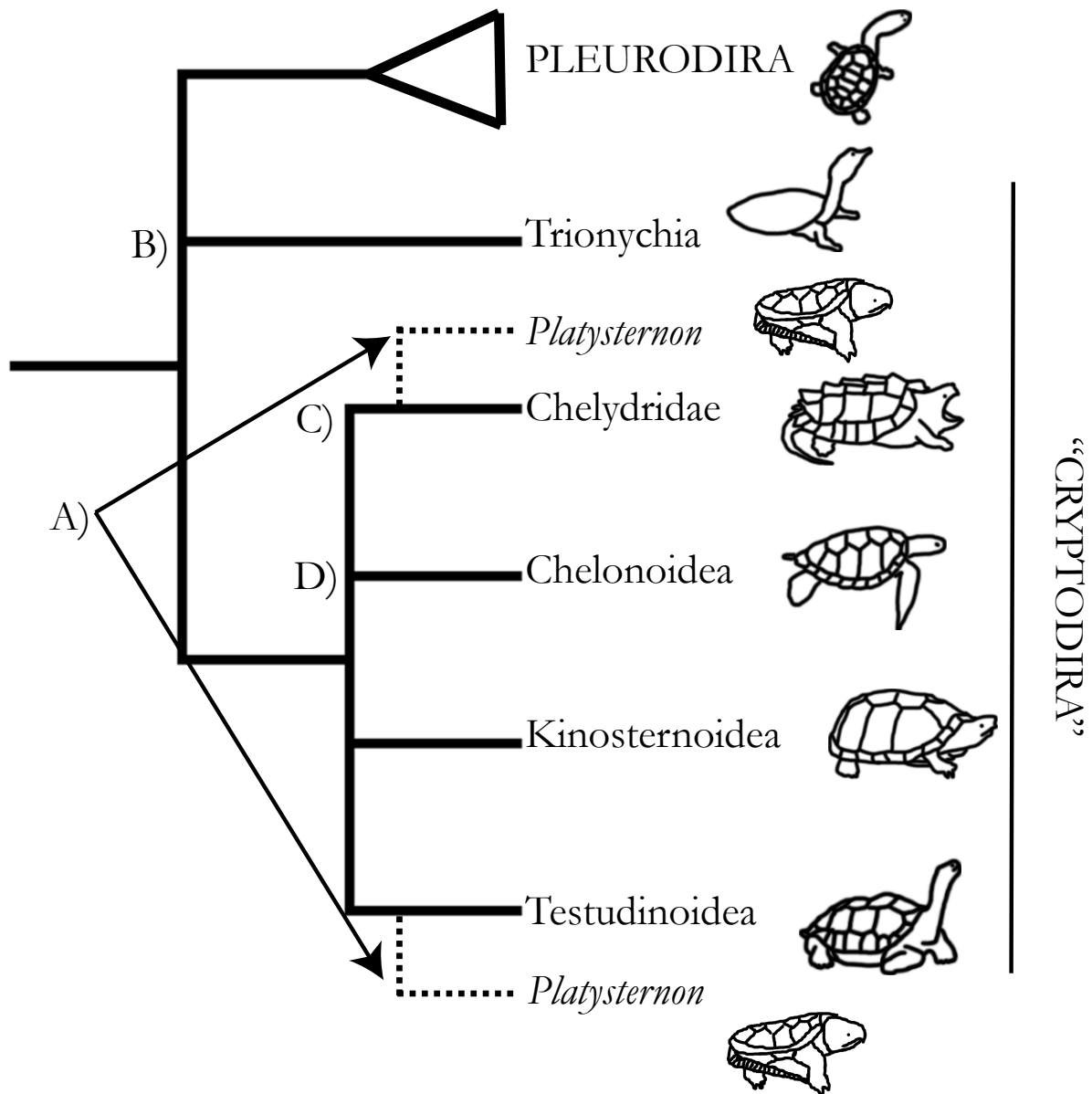


Figure 2: Phylogenetic inference of the order Testudines. Results from all analyses were similar, and the phylogeny shown is from the RAxML analysis of the concatenated nucleotide dataset. Bootstrap support values are shown next to each node as (RAxML bootstrap/Bayesian posterior probability), and are only indicated if the RAxML bootstrap is ≥ 50 . Dashed lines delineating groups (Chelidae, Testudinoidea) indicate non-monophyletic groupings. Note: *Phrynops* was removed from Bayesian analyses, as it was recovered nested within Cryptodira (see text for explanation).

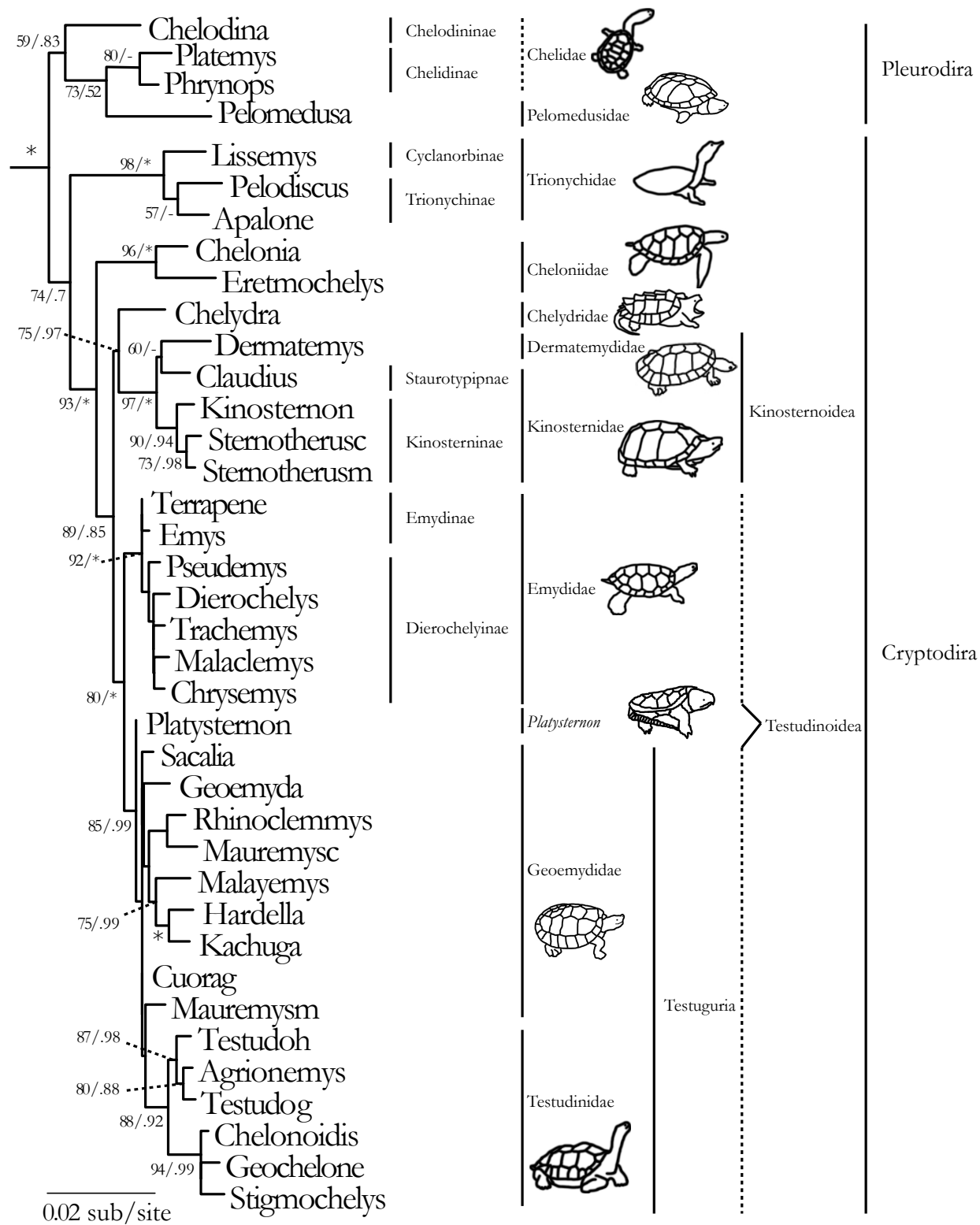
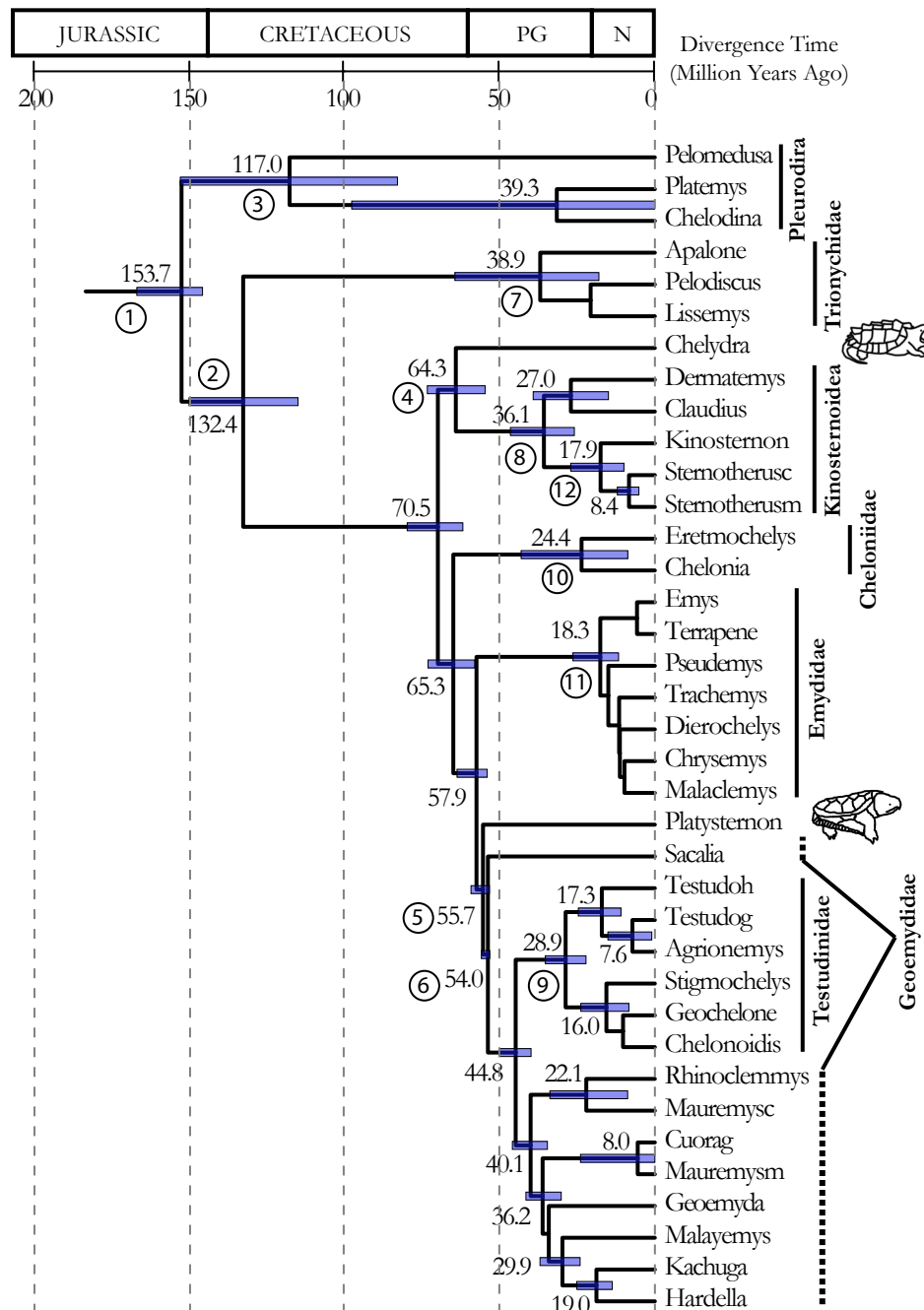


Figure 3: Phylogeny and divergence dating results from BEAST analyses. Bars at nodes represent the 95% confidence intervals, and numbers represent the mean estimation of the divergence date. Snapping turtles, Chelydridae and *Platysternon*, are represented by drawings of the turtles, and the dotted line delineating Geoemydidae indicates the group's polyphyly. Three difference are present between this phylogeny and that recovered in RAxML/Bayes analyses in Figure 2. 1) A monophyletic Chelidae, 2) Kinosternoidea the second-most basal turtle group (instead of Cheloniidae), 3) A paraphyletic Geoemydidae. PG=Paleogene, N=Neogene.



CHAPTER 4

DNA EVIDENCE FOR THE HYBRIDIZATION OF WILD TURTLES IN TAIWAN: POSSIBLE GENETIC POLLUTION FROM TRADE ANIMALS

ABSTRACT

Field surveys in Taiwan have uncovered turtles presumed to be hybrids based on their intermediate morphology. We sequenced a mitochondrial (ND4) and nuclear (R35) gene of two putative hybrid individuals, along with representatives of the potential parental species (*Mauremys mutica*, *M. reevesii*, *M. sinensis*), to determine their genetic identity. Based on our data, both individuals are hybrids, with independent, recent origins resulting from the mating of a female *M. reevesii* and a male *M. sinensis*. Since we question whether the highly traded *M. reevesii* is endemic to Taiwan, this hybridization could represent human-mediated genetic pollution. We also discuss the implications of our findings on turtle conservation in Taiwan.

INTRODUCTION

Recent Asian turtle research has uncovered the ability of distantly related species in the family Geoemydidae to hybridize (see Buskirk et al. 2005 for a review). This propensity for reticulation combined with a lack of detailed study of Asian geoemydids has sometimes confused herpetologists; at least five newly described species have turned out to be the result of recent, human-mediated hybridizations (Parham et al. 2001; Spinks et al. 2004; Stuart & Parham 2007). Asian turtles face extinction due to the turtle trade (van Dijk et al. 2000), and anthropogenic hybridization complicates the situation by polluting gene pools and exploiting precious research and conservation resources (Allendorf et al. 2001, Fong et al. 2007). Whereas most hybrids are thought to be formed in Chinese turtle farms (Parham & Shi 2001, Parham et al. 2001, Shi et al. 2007, 2008), there have been reports of turtle hybrids appearing in the wild (Otani 1995, Shi et al. 2005, Haramura et al. 2008). These studies reemphasize the point that understanding the origin of Chinese turtle hybrids may require multiple explanations (Parham et al. 2001).

Detailed surveys of turtle populations throughout Taiwan (Chen & Lue 2010) have identified two likely hybrids (Figure 1). Of the five freshwater turtle species found in Taiwan, the habitat and morphology of the probable hybrids narrowed the potential parental species to three: *Mauremys mutica*, *M. sinensis*, and *M. reevesii*. *Mauremys sinensis* is widely distributed in the low-elevation regions throughout Taiwan, while *M. mutica* mainly inhabits hilly areas in northern Taiwan. *Mauremys reevesii* is currently restricted to Kinmen Island (Figure 2; ~2km east of Fujian Province and ~270km west of Taiwan), but was historically recorded from the Taipei Basin (Horikawa 1934, Mao 1971). The goals of this study are to genetically determine whether morphologically unidentifiable individuals are hybrids, and if so, determine the parental species. We also discuss whether *M. reevesii* is endemic to Taiwan, and the implications of our findings on Taiwanese turtle conservation.

MATERIALS AND METHODS

Twenty-one specimens were collected throughout Taiwan (Figure 2, Appendix 1) and identified based on diagnostic morphological characters. Included were two suspected hybrids identified by their unusual/intermediate morphology. All individuals were genotyped for a combination of mitochondrial (mtDNA) and/or nuclear DNA (nuDNA) sequences, and combined with nine Genbank sequences for phylogenetic analyses (Appendix 1).

Genbank sequences were used as outgroups and to expand the geographical sampling of the three species. Individuals from Taiwan were not euthanized to take blood samples, since they are involved in long-term ecological surveys. DNA was extracted using a standard salt extraction protocol (Sambrook & Russell 2001). Amplification followed a standard PCR protocol to amplify an 892 bp mtDNA fragment containing the NADH dehydrogenase subunit 4 (ND4) gene and flanking tRNAs (Stuart & Parham 2004) and ~1133 bp of the RNA fingerprint protein gene intron 1 (R35) (Fujita et al. 2004). A TOPO TA cloning kit (Invitrogen) was used to recover both alleles of the R35 intron for the two putative hybrids, while the software program Phase (Stephens et al. 2001) was used to recover alleles for the parental species.

Phylogenetic analyses were run under maximum likelihood (ML) and Bayesian inference (BI). ML analyses were run using RAxML v7.0.4 (Stamatakis 2006) on the CIPRES Portal v1.15 (<http://www.phylo.org/portal>), with 1000 bootstrap replicates (Stamatakis et al. 2008). BI analyses were run in MrBayes v3.1.2. nuDNA data were run under a single model, while mtDNA data were partitioned into 1st, 2nd, 3rd codon position and tRNA, with models chosen using MrModelTest v2 (Nylander 2004) under the hLRT criterion. Two BI searches with random starting trees were run and compared using four chains, four million generations and sampling every 1000th generation. Burn-in was estimated using the online program AWTY (Wilgenbusch et al. 2004).

RESULTS

A total 29 individuals (21 new, 8 Genbank) were included for ND4 analyses and 21 individuals (15 new, 6 Genbank) for R35 analyses. Both data matrices were submitted to Treebase (#S2651; M5093-4). For ND4, uncorrected pair-wise differences within the ingroup ranged from 0-8.2%. ML analyses resulted in a single tree ($-\ln L = -2897.15$, $\alpha = 0.172$). For the BI analyses, the following models were selected for the partitions: 1st pos: HKY+ Γ ; 2nd pos: F81; 3rd pos: GTR+ Γ ; tRNA: HKY+ Γ . Both independent BI runs were almost identical, so after discarding a burn-in of 500 generations, data were combined. Both ML and BI analyses for ND4 gave similar results, so only the ML phylogenetic inference is shown (Figure 3).

For R35 data, uncorrected pair-wise differences of the ingroup ranged from 0-1.3%. ML analyses resulted in a single tree ($-\ln L = -2216.593$, $\alpha = 0.020$). For BI analyses, the HKY+I+ Γ model was selected for the entire dataset, and both independent runs were combined after removing the 500-generation burn-in due to their similar results. The ML analyses for R35 are shown in Figure 4.

We placed the two hybrid individuals into the mitochondrial and nuclear phylogenetic frameworks. Both Hybrid 1 and 2 contained *M. reevesii* mitochondrial haplotypes (Figure 3). In the nuclear analyses, Hybrid 1 and 2 possessed both *M. reevesii* and *M. sinensis* alleles (Figure 4). However, the hybrids exhibited different nuclear genotypes suggesting that they were formed through independent hybridization events.

DISCUSSION

The hybrid origin of two wild-caught individuals from Taiwan was confirmed with genetic data. Both individuals had a *M. reevesii* mother (Figure 3), and a *M. sinensis* father (Figure 4).

The nuclear data demonstrate that both hybrid individuals are not each other's closest relatives, indicating multiple hybridization events.

It is not surprising that *M. sinensis* was one of the parental species, since it was found living syntopically with Hybrid 2. Although *M. reevesii* has been documented to readily and frequently hybridize (Buskirk et al. 2005), it was unexpected that this species was the second parental species due to its rarity on Taiwan Island. *Mauremys reevesii* is currently restricted to Kinmen Island, but historical records reported *M. reevesii* from the main island of Taiwan as long ago as 1931 (Horikawa 1934) and as recently as 1971 (Mao 1971). It is possible that *M. reevesii* is currently present on the main island of Taiwan and has gone undetected. Extensive surveys of Taiwan (Chen & Lue 2010) found none, so if *M. reevesii* is present, it persists at an extremely low population density.

We raise the possibility that *M. reevesii* is not endemic to Taiwan Island, but rather the result of human introduction. *Mauremys reevesii* is only known from a few historical localities on Taiwan Island, all near developed, urban areas in the Taipei Basin. Pope (1935) noted that humans commonly moved *M. reevesii* around China because it is an important ingredient of traditional Chinese medicine. Currently, *M. reevesii* is the most common hard-shell, aquatic turtle found in commercial turtle farms (Shi & Parham 2001, Shi et al. 2008) and most widely kept as a pet (Buskirk et al. 2005), providing more opportunity for escape and release into non-native habitats. Such is the case in the Ryukyu Islands of Japan, where several feral populations have been established (Masuno et al. 1998, Ota et al. 2004). Therefore, the possibility that *M. reevesii* is not endemic to Taiwan Island is high. One way to test this hypothesis is to determine the relationship of *M. reevesii* from Taiwan to other parts of its range. This is complicated by the fact that we have not found any pure specimens on Taiwan Island, and data on *M. reevesii* are limited because genetic samples from the wild are known from just Sichuan Province and Kinmen Island (Figure 2). Despite the fact that these two localities are separated by ~1500 km, they show very limited genetic variation (0.3-1.9%, mtDNA).

The surge in occurrences of hybrid turtles appearing from the wild (Otani 1995, Shi et al. 2005, Haramura et al. 2008) should concern turtle conservation efforts. The most likely contributing factors are habitat destruction and the turtle trade. Habitat destruction is reducing the amount of suitable area for turtles, potentially forcing species that normally did not come into contact to do so, while the turtle trade moves from their native range and farmed at industrial scales in China (Shi et al. 2007, 2008). Genetically compromised captive populations from turtle farms may escape into non-native habitats. Besides, accidental escapes (Shi et al. 2004), Buddhists, pet owners, and even customs officials are releasing turtles into the wild without considering the native ranges and genetics of these species, leading towards non-natural hybridization. Stricter regulations and law enforcement need to be undertaken to protect the remaining natural habitat, as well as the farming, trading, and transport of turtles.

Mauremys reevesii is currently listed as endangered on the IUCN Redlist (IUCN 2007) and protected under the Wildlife Conservation Law as a rare and valuable species in Taiwan (Wildlife Conservation Law 1989). Based on the recent turtle surveys (Chen & Lue 2010), *M. reevesii* populations are either at extremely low population densities or extirpated. This makes *M. reevesii* a prime candidate to maintain their listing as endangered. However, further

research should be done to determine whether *M. reevesii* populations are natural or introduced. If introduced, their conservation value would be negligible. Furthermore, they should be actively removed from the wild since they are genetically polluting truly endemic species such as *M. sinensis* and *M. mutica* through hybridization (Wink et al. 2001, Buskirk et al. 2005, this study). As always, multiple explanations may be necessary to explain the origin of separate *M. reevesii* populations, so the disjunct Taiwan and Kinmen Island populations should be tested independently. If future research demonstrates that *M. reevesii* in Taiwan are introduced, this species should be removed from the Wildlife Conservation Law list to free up precious conservation resources for native, endangered species in Taiwan. However, until definitive data establish its non-native status in Taiwan, *M. reevesii* should remain on the protected list. Our work highlights the possibility that the genetic integrity of turtles endemic to Taiwan has been compromised by an alien species. There is no doubt that the food and pet trade is increasing the frequency of genetic pollution of native turtle species in Asia, and provides further reason to more carefully control and monitor this trade.

ACKNOWLEDGEMENTS

This work was made possible through funding from the National Science Foundation, EAPSI summer fellowship (JJF), the Museum of Vertebrate Zoology, UC Berkeley (JJF), and the Chelonian Research Foundation's Linnaeus Fund (JJF). We would also like to thank many people for their assistance. Drs. Kuang-yang Lue, Shou-hsien Li and their respective lab groups must be thanked for offering their laboratory facilities and assistance in labwork. James Parham helped tremendously with discussions and revisions of this manuscript. Rebecca Chong was integral with her assistance in cloning of hybrid samples and Matt Fujita assisted in Phase analyses. The samples from Taiwan were collected under the permits from Council of Agriculture, Republic of China (permit to T-HC: 0920127668, 0931614045 and 0941611580).

REFERENCES

- Allendorf FW, Leary RF, Spruell P, Wenburg JK (2001) The problems with hybrids: setting conservation guidelines. *Trends in Ecology and Evolution* 16: 613-622.
- Buskirk JR, Parham JF, Feldman CR (2005) On the hybridisation between two distantly related Asian turtles (Testudines: *Sacalia* x *Mauremys*). *Salamandra* 41: 21-26.
- Chen T-H, Lue K-Y (2010) Population status and distribution of the aquatic freshwater turtles in Taiwan. *Oryx* 44: 261-266.
- Fong JJ, Parham JF, Shi H, Stuart BL, Carter RL (2007) A genetic survey of heavily exploited, endangered turtles (*Mauremys mutica* complex): Caveats on the conservation value of trade animals. *Animal Conservation* 10: 452-460.
- Fujita MK, Engstrom TG, Starkey DE, Shaffer HB (2004) Turtle phylogeny: insights from a novel nuclear intron. *Molecular Phylogenetics and Evolution* 31: 1031-1040.

- Haramura T, Machiko Y, Akira M (2008) Preliminary survey on the turtle community in a lotic environment of the Kizu River. *Current Herpetology* 27: 101-108.
- Horikawa Y (1934) Turtles of Taiwan. *The Taiwan Jiho* 181: 7-16. (in Japanese)
- IUCN (2007) *2007 IUCN Red List of Threatened Species*. IUCN, Gland, Switzerland. <http://www.iucnredlist.org> [accessed 18 March 2008]
- Mao SH (1971) *Turtles of Taiwan*. Commercial Press, Taipei, Taiwan.
- Masuno T, Sasaki T, Yasukawa Y (1998) A record of the Reeves' pond turtle, *Chinemys reevesii* (Gray, 1831) (Testudines: Bataguridae), from Okinawajima Island, Ryukyu Archipelago. *Biological Magazine Okinawa* 36: 33-36. (in Japanese with English abstract).
- Nylander JAA (2004) MrModeltest v2. Program distributed by the author. Evolutionary Biology Centre, Uppsala University.
- Ota H, Toda M, Masunaga, G, Kikukawa A (2004). Feral populations of amphibians and reptiles in the Ryukyu Archipelago, Japan. *Global Environmental Research* 8: 133-143.
- Otani T (1995) Possible hybrids between *Geoemyda japonica* and *Cuora flavomarginata* found on Okinawajima Island, Ryukyu Archipelago. *Akamata* 11: 25-26.
- Parham JF, Shi H (2001) The discovery of *Mauremys iversoni*-like turtles at a turtle farm in Hainan Province, China: the counterfeit golden coin. *Asiatic Herpetological Research* 9: 71-76.
- Parham JF, Simison WB, 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.
- Sambrook J, Russell DW (2001) *Molecular cloning: a laboratory manual*. Cold Springs Harbor Laboratory Press. Cold Springs Harbor, New York.
- Pope CH (1935) The reptiles of China. In: *Natural history of central Asia*, Vol. X (C.A. Reeds, ed.). American Museum of Natural History, New York. 604 p.
- Shi H, Parham JF (2001) Preliminary observations of a large turtle farm in Hainan Province, People's Republic of China. *Turtle and Tortoise Newsletter* 3: 4-6.
- Shi H, Fan Z, Yin F, Yuan Z (2004) New data on the trade and captive breeding of turtles in Guangxi Province, south China. *Asiatic Herpetological Research* 10:126-128.

Shi H, Parham JF, Simison WB, Wang J, Gong S, Fu B (2005). A report on the hybridization between two species of threatened Asian box turtles (Testudines: Cuora) in the wild on Hainan Island (China) with comments on the origin of 'serrata'-like turtles. *Amphibia-Reptilia* 26: 377-381.

Shi H, Parham JF, Lau M, Chen T-H (2007) Farming endangered turtles to extinction in China. *Conservation Biology* 21: 5-6.

Shi H, Parham JF, Fan Z, Hong M, Yin F (2008) Evidence for the massive scale of turtle farming in China. *Oryx* 42: 147-150.

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

Stamatakis A (2006) RAxML-VI-HPC: maximum likelihood-based phylogenetic analyses with thousands of taxa and mixed models. *Bioinformatics* 22: 2688-2690.

Stamatakis A, Hoover P, Rougemont J (2008) A rapid bootstrap algorithm for the RAxML web-servers. *Systematic Biology* 75: 758-771.

Stephens M, Smith NJ, Donnelly, P (2001) A new statistical method for haplotype reconstruction from population data. *American Journal of Human Genetics* 68: 978-989.

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.

van Dijk, PP, Stuart BL, Rhodin AGJ (Eds.) (2000). Asian turtle trade: proceedings of a workshop on conservation and trade of freshwater turtles and tortoises in Asia. *Chelonian Research Monographs* 2.

Wildlife Conservation Law (1989) Republic of China's Wildlife Conservation Law 1-3266. Government of the Republic of China on Taiwan.

Wilgenbusch JC, Warren DL, Swofford DL (2004) AWTY: A system for graphical exploration of MCMC convergence in Bayesian phylogenetic inference. <<http://ceb.csit.fsu.edu/awty>>.

Wink M, Guicking D, Fritz U (2001) Molecular evidence for hybrid origin of *Mauremys iversoni* Pritchard & McCord, 1991, and *Mauremys pritchardi* McCord, 1997 (Reptilia: Testudines: Bataguridae). *Zoologische Abhandlungen Staatliches Museum für Tierkunde Dresden*. 51: 41-49.

Figure 1: Photographs of the two probable hybrid individuals found in the wild of Taiwan.

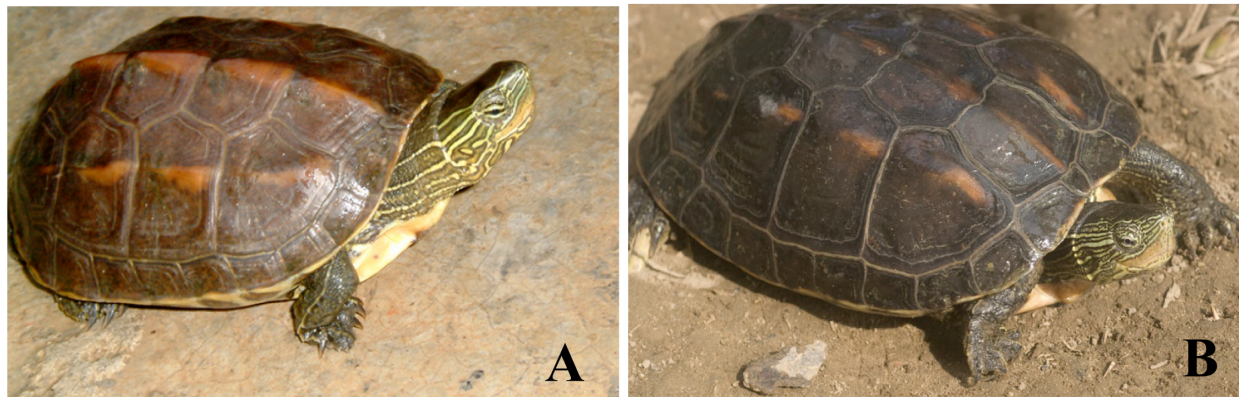


Figure 2: Map showing locality of wild specimens used in this study. M= *Mauremys mutica*, S=*Mauremys sinensis*, R=*Mauremys reevesii*, ★: hybrid. Numbers of individuals correspond to specimen numbers found in Appendix 1.

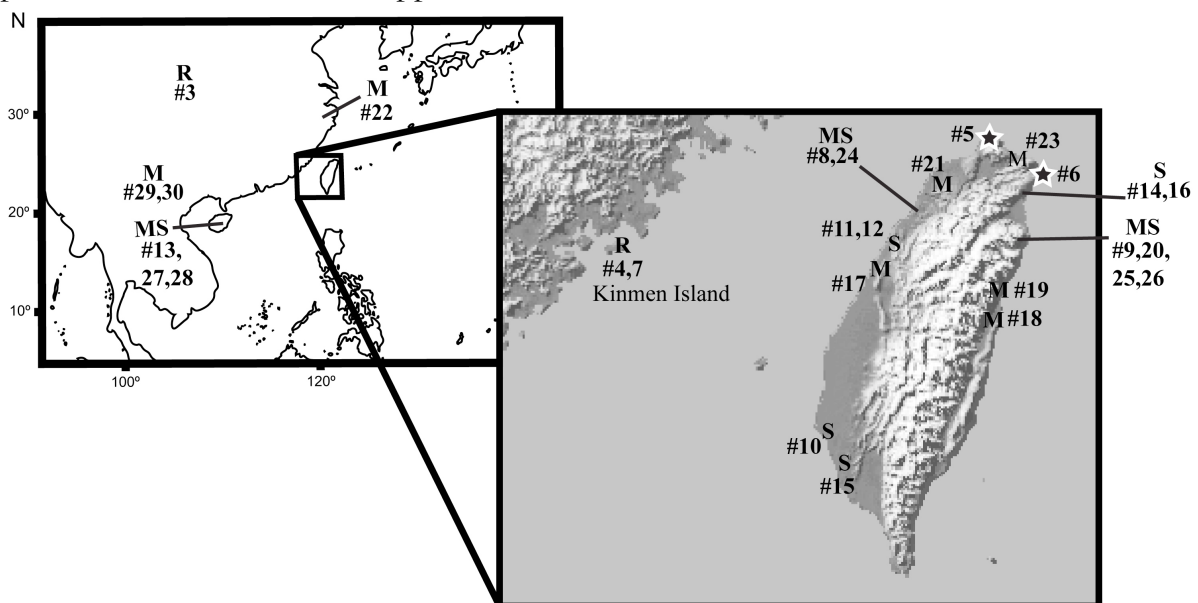


Figure 3: Maximum likelihood (ML) phylogram for mitochondrial DNA (ND4). Trees from ML and Bayesian inference resulted in almost identical trees. Nodal support is ML bootstrap/Bayesian posterior probability. Numbers of individuals correspond to specimen numbers found in Appendix 1. $\leftarrow$: Hybrid 1 haplotype, $\leftarrow$: Hybrid 2 haplotype.

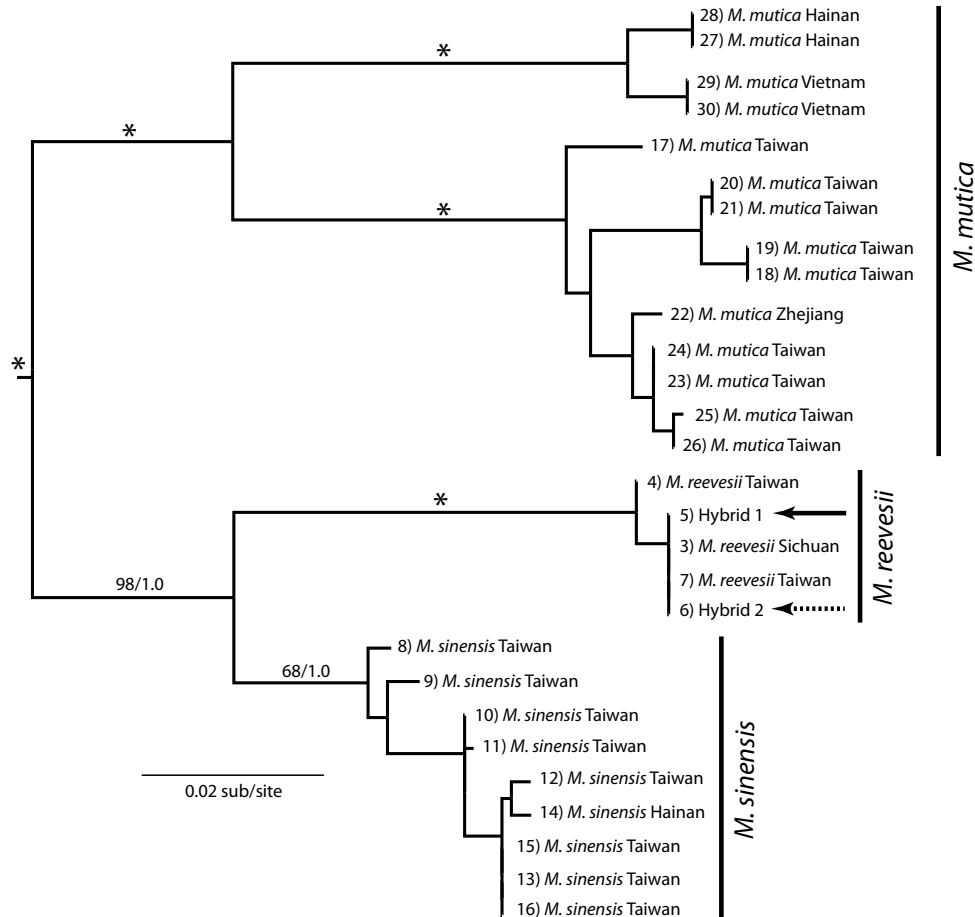
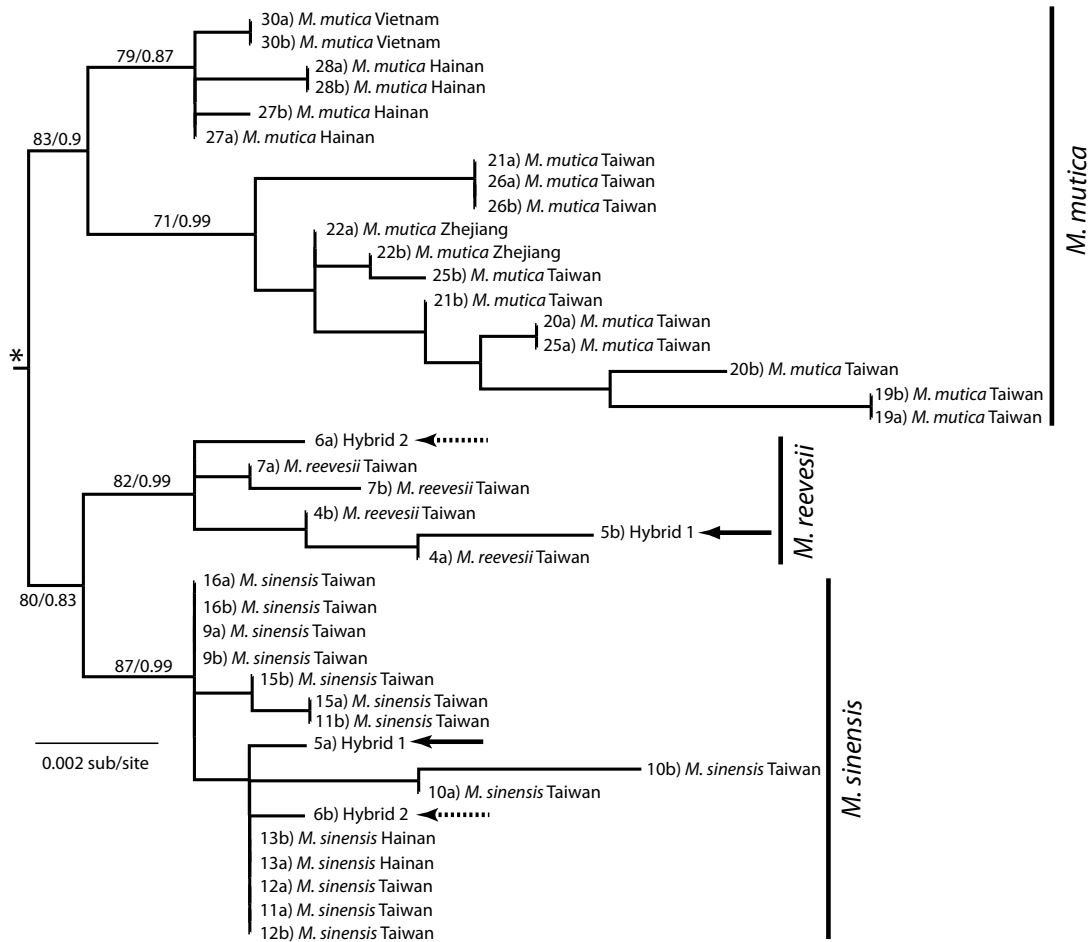


Figure 4: Maximum likelihood (ML) phylogram for nuclear DNA (R35). Trees from ML and Bayesian inference resulted in almost identical trees. Nodal support is ML bootstrap/Bayesian posterior probability. Numbers of individuals correspond to specimen numbers found in Appendix 1. ◀: Hybrid 1 alleles, ◀.....: Hybrid 2 alleles.



Appendix 1: List of individuals used in this study, along with the location of the specimen voucher and Genbank numbers. FMNH: Field Museum of Natural History, Chicago, IL, USA; MVZ: Museum of Vertebrate Zoology, Berkeley, California, USA; ROM: Royal Ontario Museum, Toronto, Ontario, Canada; All other specimens are blood samples of marked individuals (part of ecological studies) that can be found at the National Taiwan Normal University, Taipei, Taiwan.

Number	Name	Specimen #	ND4	R35
1	n/a	FMNH255695	AY364617	n/a
2	n/a	FMNH255262	n/a	DQ386654
3	<i>M. reevesii</i> Sichuan	MVZ236729	EF034110	n/a
4	<i>M. reevesii</i> Taiwan	CR	GQ259438	GQ259459
5	Hybrid 1	H1=12	GQ259439	GQ259460-1
6	Hybrid 2	H3=9	GQ259440	GQ259462-3
7	<i>M. reevesii</i> Taiwan	CR9=8	GQ259441	GQ259464
8	<i>M. sinensis</i> Taiwan	OS8=8	GQ259442	n/a
9	<i>M. sinensis</i> Taiwan	OS3=10	GQ259443	GQ259465
10	<i>M. sinensis</i> Taiwan	OS1=9	GQ259444	GQ259466
11	<i>M. sinensis</i> Taiwan	OS2=11	GQ259445	GQ259467
12	<i>M. sinensis</i> Taiwan	OS2=9	GQ259446	GQ259468
13	<i>M. sinensis</i> Hainan	MVZ230479	AY337345	DQ386678
14	<i>M. sinensis</i> Taiwan	OS8=2	GQ259447	n/a
15	<i>M. sinensis</i> Taiwan	OS1=3	GQ259448	GQ259469
16	<i>M. sinensis</i> Taiwan	OS1=11	GQ259449	GQ259470
17	<i>M. mutica</i> Taiwan	MM1=2	GQ259450	n/a
18	<i>M. mutica</i> Taiwan	MM1=15	GQ259451	n/a
19	<i>M. mutica</i> Taiwan	MM260	GQ259452	GQ259471
20	<i>M. mutica</i> Taiwan	MM3=11	GQ259453	GQ259472
21	<i>M. mutica</i> Taiwan	MM2=11	GQ259454	GQ259473
22	<i>M. mutica</i> Zhejiang	MVZ230487	AF034089	DQ386666
23	<i>M. mutica</i> Taiwan	MM9=8	GQ259455	n/a
24	<i>M. mutica</i> Taiwan	MM1=8	GQ259456	n/a
25	<i>M. mutica</i> Taiwan	MM1=11	GQ259457	GQ259474
26	<i>M. mutica</i> Taiwan	MM9=1	GQ259458	GQ259475
27	<i>M. mutica</i> Hainan	MVZ237108	EF034104	EF587915
28	<i>M. mutica</i> Hainan	MVZ251452	EF034101	EF587917
29	<i>M. mutica</i> Vietnam	ROM25613	AF348279	n/a
30	<i>M. mutica</i> Vietnam	ROM25614	AF348280	DQ386668

CHAPTER 5:

PHYLOGENETIC RELATIONSHIPS WITHIN OTHER VERTEBRATE GROUPS (ACTINOPTERYGII, AMPHIBIA, SQUAMATA, AVES, MAMMALIA)

ABSTRACT

The goals of this dissertation were to address unresolved relationships in the vertebrate phylogeny. Chapters 2 and 3 of this dissertation focused on turtle phylogenetics, namely the phylogenetic placement of turtles in the vertebrate phylogeny and the relationships within turtles, respectively. This chapter addresses relationships within the other vertebrate groups: Actinopterygii, Amphibia, Squamata, Aves, and Mammalia. Due to minimal taxon sampling, no new phylogenetic discoveries were made for the groups Actinopterygii and Aves, but these markers, with increased taxon sampling, show promise to help resolve these phylogenies. For Amphibians, questions remain regarding their origin, and whether extant amphibians are monophyletic or polyphyletic; my results point toward a diphyletic origin for the group. For squamates, there is conflict in the relationships of the major clades when using morphological or molecular data. With my sampling, I recover a phylogeny similar to previous molecular studies. In this phylogeny, the group Dibamidae+Gekkota is at the base of the phylogeny instead of an Iguania-Scleroglossa basal split (morphology), and Laterata and Lacertidae as the sister group of snakes and amphisbaenians, respectively. Lastly, for mammals, the inferred phylogeny matches the currently accepted phylogeny, with Afrotheria as the sister taxon of all remaining placental mammals. With further analyses, the enigmatic order Scandentia (*Tupaia*) is inferred to be closely related to primates, although missing data may be affecting the results.

INTRODUCTION

Resolving the Tree of Life is crucial in understanding evolution and adaptation, but also has far-reaching benefits to diverse fields such as medicine, conservation, and economics (Cracraft et al. 2002, Yates et al. 2004). Vertebrates, although only a small branch of the Tree of Life, have been the focus of intense phylogenetic research, yet many relationships remain ambiguous (Meyer & Zardoya 2003). Progress in the past decade has been rapid, with at least 50% bootstrap support for 1/4 of the nodes in the vertebrate tree (Thomson & Shaffer 2010). An ideal phylogeny would combine broad taxonomic sampling and a large set of independent molecular markers. Broad taxonomic sampling would bisect long branches, which can oftentimes be problematic in phylogenetic analyses (Hendy & Penny 1989), while a large set of independent molecular markers are needed to avoid the problem of gene trees and species tree incongruence resulting from gene duplication, loss, and lineage sorting (Page & Charleston 1997). The available data in Genbank are heterogenous, as data come from studies with different research goals and therefore have different taxon and marker sampling. There is overlap in the data, but it is primarily from a few markers: mitochondrial DNA (cytochrome b, 12s rRNA, 16s rRNA) and nuclear DNA (recombination activating gene 1 [RAG-1], β -fibrinogen, myoglobin) (Thomson & Shaffer 2010).

In this study, I utilize 75 newly developed nuclear-coding genes (Fong & Fujita submitted, Chapter 1) along with broad taxonomic sampling to help resolve relationships in the vertebrate phylogeny. Meyer & Zardoya (2003) outlined the major issues in higher-level vertebrate phylogenetics. Of these, I already addressed the phylogenetic placement of turtles in Chapter 2 of this dissertation. The goal of this chapter is to investigate the phylogenetic relationships within the major vertebrate groups except turtles, as they were the focus of

Chapter 3 of this dissertation. My taxon sampling allows me to address specific questions within certain groups. For amphibians, I address higher-level relationships in amphibians: the origin of Lissamphibia (extant amphibians) and the relationship of the three amphibian orders (Cannatella et al. 2009). For squamates, there is sharp conflict between phylogenies built using molecular and morphological data. I address the controversy by attempting to resolve interrelationships of the major clades (Iguania, Gekkota, Scincomorpha, Anguinomorpha), especially the phylogenetic placement of snakes and amphisbaenians (Conrad 2008). Lastly, for mammals, I address questions surrounding the relationship between the three major mammal clades (monotremes, marsupials, placentals; Meyer & Zardoya 2003), relationships within placentals (Murphy et al. 2001), and the phylogenetic placement of treeshrews (Family Scandentia, represented by *Tupaia*).

METHODS

Taxon and Marker Sampling

Taxon sampling was selected to span vertebrates (Tetrapoda + fish outgroups; Table 2). Of the 128 taxa, genetic data were available for 46 taxa for which complete genomes are available online (Genbank and Ensembl, Hubbard et al. 2007), and the remaining 82 samples were newly sequenced for the markers developed here. Available data were skewed heavily towards mammals (36), but also include fish (5), a frog (1), a lizard (1), birds (2), and a crocodilian (1). New taxon sampling expanded representation within amphibians, reptiles, and birds (Table 2), with an emphasis on turtles to address additional questions regarding turtle systematics (Chapter 3). Tuataras (*Spehenodon punctatus* and *S. guntheri*) were not included in this study because of the difficulties obtaining a sample of this rare and endangered animal. However, the omission of this taxon should have little effect on the results, as the sister taxon relationship between *Sphenodon* and Squamata (lizards, snakes, and amphisbaenians) is well established, forming the group Lepidosauria (Hedges & Poling 1999).

The 75 markers used in this study are from Fong & Fujita (submitted; Chapter 1). These markers are protein-coding (i.e. no introns) and were developed using EST databases, so PCR was performed on cDNA samples (instead of the standard genomic DNA). Samples were prepared as cDNA from tissues that had been appropriately preserved for RNA work, either stored in RNAlater (Ambion) or flash-frozen in liquid nitrogen. These preservation methods protect RNA by minimizing RNase activity and exposure (Sambrook & Russell 2001). Total RNA was extracted using a guanidinium thiocyanate-phenol-chloroform protocol (Chomczynski & Sacchi 1987; Sambrook & Russell 2001). Next, RNA extracts were reverse-transcribed to double-stranded cDNA using a SMART™ technology protocol (Clontech; Zhu et al. 2001). Resulting cDNA preps were used in place of genomic DNA extracts for standard PCR amplification. PCR products were purified using ExoSAP-IT (USB Corp.), cycle-sequenced using Big Dye Terminator v.3.1, and sequenced using an ABI 3730 automated sequencer. PCR conditions follow Fong & Fujita (submitted) and Chapter 1 of this dissertation.

Rogue Taxa Analyses

Rogue taxa (Wilkinson 1994, 1996) are taxa whose placement in a tree is unstable (being placed in multiple positions), which when using consensus/bootstrap methods creates polytomies and reduces support values. The instability of rogue taxa is often due to missing data or heterogeneous rates of evolution between taxa (Sanderson & Shaffer 2002). The identification of rogue taxa *a priori* is difficult, but several *a posteriori* analyses have been developed. For example, a “leaf-dropping method” was developed by Pattengale et al. (2010) and is implemented in RAxML. In general terms, this methodology gives a score to taxa based on the impact of their removal. If the score of the tree is higher excluding a set of taxa, they are identified as rogue taxa and removed. A detailed description of the algorithm follows below.

I deviate from the terminology used in Pattengale et al. (2010) to use more common phylogenetics terminology; a “leaf” refers to a taxon and an “edge” refers to a branch. Bipartitions (the resulting two subsets of taxa remaining when a branch is removed) and relative information content (RIC) are in the basis of this methodology. There are two types of bipartitions, trivial and nontrivial. Trivial bipartitions are ones which separate a single taxon from the tree (i.e. a branch leading to a tip of the tree), whereas a nontrivial bipartition refers to the removal of an internal branch to create two subsets of taxa with membership >1 . The ratio of the total number of bipartitions (trivial + nontrivial) to the number of bipartitions in a fully resolved tree (no polytomies) is called the RIC. Since the numerator of the RIC includes trivial and nontrivial bipartitions, this methodology penalizes the removal of taxa (trivial bipartitions) and rewards for greater resolution of the tree (more nontrivial bipartitions) (Pattengale et al. 2010). This algorithm can be run on majority-rule and strict consensus trees.

Pattengale et al. (2010) uses a greedy heuristic algorithm that discards a taxon only if its removal provides an improvement in resolution that outweighs the loss of removing more taxa. Various dropsets (sets of removed taxa) are given a score; this score is the difference between the number of branches created and the number of taxa removed. The dropset with the highest score is used, and the corresponding taxa are removed. When there are multiple dropsets with the same highest score, one can opt to select the first dropset (which allows for repeatability) or use a random selection. The analysis iterates until no further benefit is gained from removing taxa.

Determining which rogue taxa to remove was a multi-step process that was run separately for each data set (representing the four data-types). First, a preliminary RAxML analysis was run with all taxa (128). Based on this phylogeny, taxa that were obviously in incorrect phylogenetic positions (i.e. turtle placed in the mammal clade) were removed. This was repeated until all remaining taxa were placed in the correct clades. Next, the bootstrap results from the RAxML analysis were used in rogue taxa analyses (Pattengale et al. 2010). The rogue taxa analysis was run 10 times using the “random” option to select a dropset for both the strict and majority rules consensus trees. To maximize the number of taxa retained and identify those that were most unstable, taxa were considered rogue if they were identified in ≥ 5 of the 10 analyses under either strict or majority rules consensus. These taxa were removed and another RAxML analysis was run. If there was an improvement in

bootstrap values, a round of rogue taxon/RAxML analyses was repeated until there was no improvement in bootstrap values.

For rogue taxa analyses, identification of taxa is only the first step; it is equally important to understand why a taxon's position in a phylogeny is unstable. Research on rogue taxa is relatively new (Cranston & Rannala, 2007, Maddison & Maddison 2007, Pattengale 2010), and has focused mainly on identification. Most instability in a phylogeny results from missing data, but other biological processes may also be involved (Sanderson & Shaffer 2002). Here I attempt to gain insight into the processes responsible for rogue taxa by investigating the placement of a controversial mammal group, Scandentia (treeshrews, represented by *Tupaia*). Prior phylogenetic studies have resulted in *Tupaia* being placed as the sister-group to Glires (Schmitz et al. 2000) or Primates (Murphy et al. 2001). Because the genome for *Tupaia* has been fully sequenced (albeit with low coverage), it appears that missing data is an inadequate explanation for its unstable position. To address the phylogenetic placement of *Tupaia*, I use the program RAxML-EPA (Evolutionary Placement Algorithm; Berger & Stamatakis 2009). For the backbone phylogeny, I employed the All-NUCL (128 taxa), and 110taxa-NUCL datasets (rogue taxa removed). I partitioned the data by gene and performed 100 bootstrap replicates, with analyses run on the authors' website (<http://i12k-exelixis3.informatik.tu-muenchen.de/raxml>).

Some potential variables affecting the placement of *Tupaia* are GC content (%GC) and proportion of missing data (%missing). *Tupaia* was placed in two different positions in the phylogeny: sister to Glires and sister to Primates. I compared the %GC and %missing by using applications available in the “R” computer language and environment (<http://www.r-project.org/>). %GC and %missing were calculated for each taxon per site, and averaged for the major groups (Glires, Primates). The differences between each of these averages were calculated (Primate-Glires) and plotted, with a Lowess (locally weighted scatterplot smoothing) curve fitted to the data. These results are discussed below in the “Mammal Phylogenetics” section.

Phylogenetic Analyses

Based on analyses comparing the phylogenetic informativeness of different data-types (Fong & Fujita submitted), it was shown that the nucleotide data-type would be most appropriate for resolving relationships within the major vertebrate groups, so this is the data-type we use for all analyses. RAxML (Stamatakis 2006) was performed on the concatenated nucleotide dataset of the 75 genes and partitioned by gene and by codon (1st+2nd position, 3rd position), as this was found to be the most optimal partitioning strategy (Chapter 2).

Additionally, for relationships within amphibians, analyses of different datasets were explored. These included using different data-types (NUCL—nucleotides, N12—first and second codon positions, DEGEN1—codons degenerated using ambiguity codes, AA—amino acids) and a subset of 15 taxa. In all, eight different datasets were analyzed: All taxa-NUCL, All taxa-N12, All taxa-DEGEN1, All taxa-AA, 15 taxa-NUCL, 15 taxa-N12, 15 taxa-DEGEN1, 15 taxa-AA. RAxML was also run on these datasets.

RAxML analyses were run under the GTRGAMMA model of evolution for tree inference and for 1000 bootstrap replicates. RAxML v7.2.5 and v7.2.6 (Stamatakis 2006,

Stamatakis et al. 2008) analyses were run locally and on the CIPRES portal v2.2 (<http://www.phylo.org/portal2/>; Miller et al. 2009).

RESULTS AND DISCUSSION

Taxon and Marker Sampling

The aligned dataset of this study was 34,110bp in length, consisted of 3,928 sequences, and is 40.9% complete. Of these sequences, 878 are newly sequenced, while the rest are from Genbank and Ensembl.

Rogue Taxa Analyses

The 128 taxa included in this study varied in their completeness of data (average = 31 genes, range = 1-75 genes). As the amount of missing data is relatively high in this dataset, and missing data have been shown to affect phylogenetic analyses (Wilkinson 1994, 1996), it was important to identify and remove rogue taxa. Rogue taxa analyses were considered complete when no additional rogue taxa were identified or there were no improvements in bootstrap values. Taxa were removed from datasets either because they were clearly placed in the wrong phylogenetic position or the rogue taxa algorithm identified them. For this dataset, there was one round of taxa removal due to incorrect phylogenetic position and two rounds of the rogue taxa algorithm. The final result was 18 taxa removed and 110 remaining. The phylogenetic placement of *Tupaia* is discussed in detail below.

Actinopterygii Phylogenetics

In this study, Actinopterygii (Ray-finned fishes) was represented by five sequenced genomes: *Danio rerio*, *Gasterosteus aculeatus*, *Oryzias latipes*, *Takifugu rubripes*, and *Tetraodon nigroviridis*. All analyses recover a monophyletic Actinopterygii. Within Actinopterygii, the inferred relationships (Figure 1) are concordant with recent phylogenetic inferences of the group (Li et al. 2007, Li et al. 2008). Although no new information on phylogenetic relationships within Actinopterygii was recovered, this study provides a good test case for efficacy of these 75 new markers. Fish, which represent more than half of vertebrate diversity (IUCN 2010), have been notoriously undersampled and phylogenies poorly resolved (Thomson & Shaffer 2010). These markers, along with greater taxon sampling, hold great promise in helping to resolve relationships within fish.

Amphibia Phylogenetics

There are two outstanding questions in higher-level amphibian relationships: the origin of Lissamphibia (extant amphibians) and the relationships of the three amphibian orders. There is uncertainty in the origin of Lissamphibia, whether Lissamphibia is monophyletic or polyphyletic and the extinct ancestor of living groups. There are three hypotheses of amphibian origin: Temnospondyl hypothesis, Lepospondyl hypothesis, and polyphyletic hypothesis (reviewed in Cannatella et al. 2009). The Temnospondyl and Lepospondyl hypotheses assume a monophyletic Lissamphibia, but differ in the inferred ancestor (denoted by the name of the hypothesis). A major difference between the two hypotheses is that the Lepospondyl hypothesis requires that Lissamphibia is more closely related to amniotes than to temnospondyls. The third hypothesis, the polyphyletic

hypothesis, proposes a diphyletic origin of Lissamphibia, with caecilians being related to lepospondyls, whereas frogs and salamanders are related to temnospondyls (Cannatella et al. 2009). Many paleontological studies support a diphyletic origin of Lissamphibia (see Carroll 2007), while molecular studies support a monophyletic Lissamphibia (Meyer & Zardoya 2003).

Of these studies recovering a monophyletic Lissamphibia, there is disagreement in the relationships of the three amphibian orders (Meyer & Zardoya 2003). The Procera hypothesis places caecilians and salamanders as sister taxa, and is supported by several studies (Hedges et al. 1990, Hedges & Maxon 1993, Feller & Hedges 1998, Zhang et al. 2003). The Batrachia hypothesis, on the other hand, suggests a sister relationship between frogs and salamanders. More recent analyses support the Batrachia hypothesis, and argue that previous studies consistent with a monophyletic Procera suffered from inadequate phylogenetic signal or insufficient taxon sampling (Zardoya & Meyer 2001, Zhang et al. 2005, Hugall et al. 2007, Roleants et al. 2007).

Initially, taxon sampling included representatives from all three amphibian orders. However, rogue taxa analyses removed all salamander taxa, making a complete investigation of the three amphibian orders impossible. However, with frog and caecilian samples, I am still able to address questions regarding amphibian origins, and implicitly the relationship between the three orders. A monophyletic Amphibia was recovered with five of the datasets (All-N12, 15tax-N12, All-DEGEN1, 15tax-DEGEN1, 15tax-AA), but many with low support values, while the remaining analyses (All-NUCL, 15tax-NUCL, All-AA) recovered a polyphyletic Lissamphibia, with caecilians being more closely related to amniotes (Figure 2). Approximately Unbiased topology tests (AU tests; Shimodaira 2002) using the program CONSEL (Shimodaira & Hasegawa 2001) were used to statistically compare the two alternative hypotheses under each data-type. The N12, DEGEN1, and AA datasets were unable to statistically differentiate between a monophyletic and diphyletic Lissamphibia (Table 1). However, analysis of the NUCL dataset results show that a phylogeny with a diphyletic origin of Lissamphibia was statistically better than one with a monophyletic Lissamphibia (Table 1). Although they do not definitively solve the problem of higher-level relationships, these results seem favor a diphyletic Lissamphibian origin. However, I believe that a convincing resolution of this question will require denser taxon sampling, especially the inclusion of salamander specimens.

Squamate (lizards, snakes, amphisbaenians) Phylogenetics

There is still much debate in squamate phylogenetics, namely between morphological and molecular analyses. Traditional relationships within Squamata using morphology classified two major groups—Iguania (Iguanidae, Agamidae, Chamaeleonidae) and Scleroglossa (remaining families) (Figure 3A: Camp 1923, Figure 3B: Estes et al. 1988). Within Scleroglossa, six monophyletic groups were recovered: Gekkota (Gekkonidae, Pygopodidae), Lacertoidea (Teiidae, Gymnophthalmidae, Lacertidae, Xantusiidae), Scincoidea (Scincidae, Cordylidae), Scincomorpha (Lacertoidea, Scincoidea), Varanoidea (Helodermatidae, Varanidae), and Anguimorpha (Varanoidea, Anguinae, Xenosauridae) (Estes et al. 1988). These results were largely recovered in a recent morphological study with greater taxon and character sampling, with the addition of a novel Scincoidea clade, re-

defining the membership of the group (containing Scincidae, Dibamidae, Amphisbaenia, and Serpentes; Figure 3C: Conrad 2008). In general, morphological studies agree on the monophyly of Gekkota, Varanoidea, and Scleroglossa, but with no consensus on the relationships among these groups, and with particular problems in placing Xantusiidae, Cordyliformes, Scincidae, Dibamidae, Amphisbaenia, and Serpentes (Conrad 2008).

In sharp contrast, results from molecular studies largely conflict with results from molecular studies (Figure 3). Molecular studies drastically modify the traditional view of squamate relationships by proposing that 1) Iguania and Scleroglossa are not reciprocally monophyletic, with Iguania nested within Scleroglossa, 2) Iguania is not basal, but rather related to anguimorphans and snakes, 3) Dibamidae or Gekkota is the sister taxon of all other squamates, 4) Amphisbaenia is the sister taxon of Lacertidae, 5) Xantusiidae is related to Scincidae and Cordylidae, 5) snakes are not closely related to Varanidae, and 6) recognition of the groups Toxicofera (Anguimorpha, Iguania, Snakes) and Laterata (Teiioidea, Lacertidae, Amphisbaenia) (Vidal & Hedges 2004, Townsend et al 2004, Vidal & Hedges 2005, Fry et al. 2006). It has been pointed out that differences in results between morphological and molecular studies are expected, as morphological studies are able to include fossil taxa that affect polarities throughout the tree (Gauthier et al. 1988, Conrad 2008). I use my expanded molecular dataset to compare to past studies.

Squamate relationships from this study are shown in Figure 4. Taxon sampling for squamates did not include all families, but all the major superfamilial groups were included. Support values for the tree are generally low, probably as a result of the incompleteness of the dataset, but the majority of the relationships are concordant with the current molecular squamate phylogeny. Major groups such as Scinciformata (Xantusiidae+Scincidae) and Laterata (Teiidae, Lacertidae, Amphisbaenia) were recovered in my analyses. Also, as previously inferred, the evolution of extreme limblessness evolved independently several times across the squamate tree, as seen by distant relationships between Dibamidae, Anniellidae, Serpentes, and Amphisbaenia (Figure 4). Conrad (2008), using morphological data, also infers multiple origins of extreme limblessness, but in contrast recovers a single clade for Dibamidae, Serpentes, and Amphisbaenia (Figure 3C).

There are, however, two main points where my phylogenetic inference deviates from the recent molecular trees (Townsend et al. 2004, Hedges & Vidal 2009). First, my study recovers a Dibamidae+Gekkonidae group as the sister group of all other squamates, a hybrid result of the two recent molecular analyses; Townsend et al. (2004) recovered a basal Gekkota (Figure 3D) while Vidal & Hedges (2005) recovered a basal Dibamidae (Figure 3E). Second, although support values for the branches are low, the Toxicofera clade from Vidal & Hedges (2005) was not recovered in my analyses. Instead, snakes are supported as the sister group of Laterata. Therefore, the question of snake origins remains in question.

Recent work has advanced our understanding of squamate phylogeny, but the conflict between molecular and morphological phylogenies remains. Results from both data types are in strong opposition to one another. This is probably due to the difference in taxon sampling between molecular and morphological studies. Future work to resolve this issue will need to combine both morphological and molecular data. The 75 markers used in this study seem useful and would help bolster the molecular data in these combined analyses.

Aves (birds) Phylogenetics

Living birds in the group Aves are placed in the sub-class Neornithes. This group is further split into Paleognathae (rattites and tinamous) and Neognathae (all remaining birds; van Tuinen 2009). Neognathae is further divided into Galloanserae (land and waterfowl) and Neoaves (all remaining Neognathan orders). These higher splits within Aves are generally accepted (van Tuinen 2009). The relationships within Neoaves, however, are poorly resolved, often represented as the “neoavian comb” (Cracraft et al. 2004). Some believe that the consistent recovery of this comb reflects a rapid radiation, and therefore a hard polytomy (Poe & Chubb 2004).

Taxon sampling was selected to span the group and focus on basal relationships within Neoaves. However, after rogue taxa analyses, several taxa were removed, including *Accipiter* (Falconiformes), *Aptenodytes* (Sphenisciformes), *Casuarinus* (Paleognathae), *Anas* (Anseriformes), and *Agelaiocercus* (Apodiformes). What remained was limited sampling in Neognathae that did not allow the evaluation of basal relationships in birds. The inferred phylogenetic relationships of the remaining taxa are presented in Figure 5. These data recover the basal position of Galliformes in Neognathae. A well-supported clade of Picidae (Piciformes) and Monarchidae (Passeriformes) was found to be the basal group in Neoaves, rendering Passeriformes paraphyletic. Although taxon sampling was limited for birds in this study, these markers show promise in teasing out relationships among birds. Expanding taxon sampling to include the basal bird groups may help determine whether birds truly experienced a rapid radiation within Neoaves.

Mammal Phylogenetics

RELATIONSHIPS WITHIN MAMMALS

Mammals are a monophyletic group, characterized by traits such as the presence of mammary glands, hair, and three ossicles in the middle ear (Simpson 1959). The three major groups of mammals (Monotremata, Marsupialia, and Placentalia) have also been shown to be incontrovertibly monophyletic. Controversies remain regarding the relationships of these three orders (Meyer & Zardoya 2003), as well as for basal relationships within Placentalia (Murphy et al. 2001).

When dealing with the relationships of the three major mammal groups, two hypotheses have been proposed—the Theria and Marsupionta hypotheses. The Theria hypothesis is the traditional view, and proposes a sister relationship between marsupial and placental mammals. Support for this hypothesis comes from both morphological and molecular data (mitochondrial and nuclear) (Table 2). The Marsupionta hypothesis proposes that monotremes and marsupials form a clade to the exclusion of placental mammals. Early morphological studies proposed this relationship, and subsequent molecular studies based primarily on mitochondrial DNA supported this hypothesis as well (i.e., Zardoya & Meyer 2003). However, Phillips & Penny (2003) argued that high GC content in marsupial and monotreme genomes was responsible for their erroneous grouping as a monophyletic assemblage in the Zardoya and Meyer study. Reanalysis of these data, recoding nucleotides to RY characters, resulted in support for the Theria hypothesis (Table 2). Preliminary analysis of higher-level mammal relationships based on 13 taxa (4 mammals) further

supported the Theria hypothesis (Chapter 1). Expanding the taxon sampling to 35 mammal taxa again supported the Theria hypothesis (Figure 6).

Traditional phylogenetic relationships within placental mammals were proposed by McKenna & Bell (1997) (Table 3). However, the application of molecular phylogenetic methods have resulted in a substantially modified mammalian phylogenetic tree and associated classification. For example, the traditional Insectivora was found to be polyphyletic, and strong support has been obtained for a previously unrecognized African radiation—Afrotheria (Springer et al. 1997, Stanhope et al. 1998). Focus then turned to the identification of the basal placental clade, with alternative studies finding support for a basal Afrotheria (Murphy et al. 2001, Waddell et al. 2001), Xenarthra (Shoshani & McKenna 1998), or Atlantogenata (Afrotheria+Xenarthra) (Murphy et al. 2007, Prasad & Allard 2008).

Results from this study (Figure 6) do not support the traditional classification (McKenna & Bell, 1997), but are largely concordant with the newer mammalian phylogenetic estimates (Murphy et al. 2001, Murphy & Eizirik, 2009). In regards to basal relationships of placental mammals, my results mirror Murphy et al. (2001) and Waddell et al. (2001), and support a basal Afrotheria clade, followed by Xenarthra and Boreoeutheria.

PHYLOGENETIC PLACEMENT OF *TUPAIA*

The order Scandentia (tree shrews, represented by the genus *Tupaia*) has been a relatively problematic group with respect to its phylogenetic placement (Nishihara et al. 2002). Initially placed in the now defunct Insectivora, Scandentia has been placed as sister to Glires using data from whole mitochondrial genomes (Schmitz et al. 2000), and as sister to Primates using nuclear markers (Murphy et al. 2001). For my data, this taxon was removed from the final dataset after rogue taxa analyses. To shed insight into this problematic group, further analyses using the program RAxML-EPA were performed. RAxML-EPA was initially developed to identify and place organisms from microbial communities into a phylogenetic context (Berger & Stamatakis 2009). This algorithm utilizes a backbone phylogeny of known relationships and places these anonymous taxa within this phylogeny, calculating the optimal insertion branch under Maximum Likelihood, and the confidence of this placement is evaluated via phylogenetic bootstrapping.

The unstable placement of *Tupaia* is unlikely to have resulted from missing data, as data are from a sequenced whole nuclear genome, are present for 66 of 75 genes. Results from the RAxML-EPA analyses are depicted in Figure 7. From these analyses, we recover the same two positions of *Tupaia* previously suggested, sister to Glires and sister to Primates. Although these data more strongly support the placement near Primates (79 bootstrap), it is interesting that the alternative placement recovered was sister to Glires. It is clear from these analyses that there is conflicting phylogenetic signal in the genome grouping *Tupaia* with both Glires and Primates.

To determine if there is any systematic bias in the data, plots were used to compare the %GC and %missing between Primates and Glires (Figs. 8-9). A value of “0” indicates no difference between the groups, while a positive value indicates that Primates have a higher value, and a negative value indicates that Glires have a higher value. Based on the relatively flat Lowess curve along a value of zero, we see that there is very little difference in %GC between Primates and Glires. However, the Lowess curve for %missing is mostly negative,

indicating that Glires, on average, have more missing data than Primates. Therefore, it seems that placement of *Tupaia* as the sister group to Primates is not affected by %GC, but potentially by %missing, as Glires, on average, have more missing data than Primates. Although my data recover a Scandentia-Primate relationship, differential amounts of missing data in Glires and Primates may be biasing these results. It should be noted that the *Tupaia* nuclear genome, although fully sequenced, is only sequenced at 2x coverage. The low-quality of the genome sequencing may be affecting these results. Additional work on the molecular evolution of Scandentia, by exploring the influence of additional data parameters, increasing coverage quality of the genome, and the inclusion of more taxa, may assist in resolving this issue.

CONCLUSION

Although the main focus of my vertebrate phylogeny research was on the phylogenetic placement of and relationships within turtles, strategic taxonomic sampling allowed me to address relationships for other major vertebrate groups. For Actinopterygii, recovered relationships match the currently accepted phylogeny. Fish have been poorly sampled in tree of life projects, and additional sampling using these markers seems promising in helping to resolve the fish phylogeny. For Amphibia, my results point towards a diphyletic origin of Amphibia, but additional sampling of salamanders will be necessary to definitively answer this question. The relationships within Squamata, although poorly supported, seem to match the general molecular consensus to date. The conflict between molecular and morphological phylogenies remains. Work combining morphological and molecular data will be necessary in resolving the squamate tree. The group Neoaves within Aves remains problematic, although these markers do show promise in resolving the “avian comb”. Lastly, mammal relationships have been relatively well resolved, and my data agree with previous studies. However, the order Scandentia, although possibly closely related to Primates, remains a curious lineage as there is also signal placing it as the sister group to Glires. Although we present stronger support for Scandentia as the sister group to Primates, these results may be biased by asymmetrical amounts of missing data in Glires.

Although many questions remain, these data show that the molecular markers developed and employed here may be useful for resolving the phylogenetic positions of taxa that have proven particularly resistant to the efforts of students of vertebrate phylogeny. Denser taxon sampling and focusing on collecting a more complete dataset should help in the resolution of the vertebrate phylogeny.

ACKNOWLEDGEMENTS

I would like to thank the individuals and institutions that provided tissue samples for this project: Drs. Carol Spencer and Jim McGuire of the Museum of Vertebrate Zoology, UC Berkeley, Dr. Matthew K Fujita of the Museum of Comparative Zoology, Harvard University, Dr. Jerome Fuchs of the California Academy of Sciences, Dr. Juan L. Parra of State University of New York, Stony Brook, and Dr. Ricardo Pereira of the Universidad de Porto. I would also like to thank Jim McGuire, Jeffrey Boore, and Richard Harland for reading over drafts of this chapter. In addition, I would like to thank Adam Leaché, Matt Fujita, Matt Brandley, Alexendros Stamatakis, Bastien Boussau, Jeremy Brown, Nicholas

Pattengale, James F Parham, and Jeffrey Townsend for help with analyses. Funding for this project came from the National Science Foundation Doctoral Dissertation Improvement Grant (#08-564) and the Museum of Vertebrate Zoology.

CITATIONS

Benton MJ (1990) Phylogeny of the major tetrapod groups: morphological data and divergence dates. *Journal of Molecular Evolution* 30: 409-424.

Berger SA, Stamatakis A (2009) Evolutionary placement of short sequence reads. Exelixis-RRDR-2009-3, TU Munich.

Camp CL (1923) Classification of lizards. *Bulletin of the American Museum of Natural History* 48: 289-481.

Cannatella DC, Vieites DR, Zhang P, Wake MH, Wake DB (2009) Amphibians (Lissamphibia). In: *The timetree of Life*, Hedges, SB, Kumar, S (eds.). Oxford University Press. Pp. 353-356.

Carroll RL (2007) The Palaeozoic ancestry of salamanders, frogs, and caecilians. *Zoological Journal of the Linnean Society* 150 (suppl. 1) 1-140.

Chomczynski P, Sacchi N (1987) Single-step method of RNA isolation by acid guanadinium thiocyanate-phenol-chloroform extraction. *Analytical Biochemistry* 162: 156-159.

Conrad JL (2008) Phylogeny and systematics of Squamata (Reptilia) based on morphology. *Bulletin of the American Museum of Natural History* 310: 182pp.

Cracraft J, Donoghue M, Dragoo J, Hillis D, Yates T (eds.) (2002). *Assembling the tree of life*. National Science Foundation Report. http://www.phylo.org/atol/pdf_docs/atol.pdf

Cracraft J, Feinstein J, Garcia-Moreno J, Barker FK, Stanley S, Sorenson MD, Braun M, Cibois A, Yuri T, Harshman J, Schikler P, Mindell DP, Dyke GJ, Beresford P (2004) Phylogenetic relationships among modern birds (Neornithes): toward an avian tree of life. In: Cracraft J, Donoghue MK (eds.). *Assembling the tree of life*. New York: Oxford University Press. Pp 469-489.

Cranston KA, Rannala B (2007) Summarizing a posterior distribution of trees using agreement subtrees. *Systematic Biology* 56: 578-590.

Estes R, de Queiroz K, Gauthier J (1988) Phylogenetic relationships within Squamata. In: Estes R, Pregill G. (eds.). *Phylogenetic relationships of the Lizard families*. Stanford, California, USA: Stanford University Press. Pp 119-281.

Feller AE, Hedges SB (1998) Molecular evidence for the early history of living amphibians. *Molecular Phylogenetics and Evolution* 9: 509-516.

Fong JJ & Fujita MK (submitted) Evaluating phylogenetic informativeness and data-type usage for new protein-coding genes across vertebrata. Submitted to *Molecular Phylogenetics and Evolution*.

Fry BG, Vidal N, Norman JA, Vonk FJ, Scheib H, Ramjan SFR, Kuruppu S, Fung K, Hedges SB, Richardson MK, Hodgson WC, Ignjatovic V, Summerhayes R, Kochva E (2006) Early evolution of the venom system in lizards and snakes. *Nature* 439: 584-588.

Gauthier JA, Kluge AG, Rowe T (1988) Amniote phylogeny and the importance of fossils. *Cladistics* 4: 105-209.

Hedges SB, Moberg KD, Maxson LR (1990) Tetrapod phylogeny inferred from 18S and 28S ribosomal RNA sequences and a review of the evidence for amniote relationships. *Molecular Biology and Evolution* 7: 607-633.

Hedges SB, Maxson LR (1993) A molecular perspective on lissamphibian phylogeny. *Herpetological Monographs* 7: 27-42.

Hedges SB, Poling LL (1999) A molecular phylogeny of reptiles. *Science* 283: 998-1001.

Hedges SB, Vidal N (2009) Lizards, snakes and amphisbaenians (Squamata). In: *The timetree of Life*, Hedges, SB, Kumar, S (eds.). Oxford University Press. Pp 383-389

Hendy MD, Penny D (1989) A framework for the quantitative study of evolutionary trees. *Systematic Zoology* 38: 297-309.

Hubbard TJP, Aken BL, Ayling S, Ballester B, Beal K, Bragin E, Brent S, Chen Y, Clapham P, Clarke L, Coates G, Fairley S, Fitzgerald S, Fernandez-Banet J, Gordon L, Graf S, Haider S, Hammond M, Holland R, Howe K, Jenkinson A, Johnson N, Kahari A, Keefe D, Keenan S, Kinsella R, Kokocinski F, Kulesha E, Lawson D, Longden I, Megy K, Meidl P, Overduin B, Parker A, Pritchard B, Rios D, Schuster M, Slater G, Smedley D, Spooner W, Spudich G, Trevanion S, Vilella A, Vogl J, White S, Wilder S, Zadissa A, Birney E, Cunningham F, Curwen V, Durbin R, Fernandez-Suarez XM, Herrero J, Kasprzyk A, Proctor G, Smith J, Searle S, Flicek P (2009) Ensembl 2009. *Nucleic Acids Research* 37: D690-D697.

Hugall AF, Lee MSY (2007) The likelihood node density effect and consequences for evolutionary studies of molecular rates. *Evolution* 61: 2293-2307.

IUCN (2010) IUCN red list of threatened species 2010.4. [<http://www.iucnredlist.org/>].

Li C, Orti G, Zhang G, Lu G (2007) A practical approach to phylogenomics: the phylogeny of ray-finned fish (Actinopterygii) as a case study *BMC Evolutionary Biology* 7: 44.

Li C, Orti G (2008) Optimal data partitioning and a test case for ray-finned fishes (Actinopterygii) based on ten nuclear loci. *Systematic Biology* 57: 519-539.

Maddison WP, Maddison DR (2007) Mesquite: a modular system for evolutionary analysis [Internet]. Version 2.01. Available from <http://mesquiteproject.org>.

Meyer A, Zardoya R (2003) Recent advances in the (molecular) phylogeny of vertebrates. *Annual Review of Ecology Evolution and Systematics* 34: 311-338.

Miller MA, Holder MT, Vos R, Midford PE, Liebowitz T, Chan L, Hoover P, Warnow T (2009) The CIPRES Portals. CIPRES. 2009-08-04. URL: http://www.phylo.org/sub_sections/portal. Accessed: 2009-08-04. (Archived by WebCite(r) at <http://www.webcitation.org/5imQlJeQa>)

McKenna MC, Bell SK (1997) Classification of mammals above the species level. New York: Columbia University Press. Pp. 613.

Murphy WJ, Eizirik E, O'Brien SJ, Madsen O, Scally M, Douady CJ, Teeling E, Ryder OA, Stanhope MJ, de Jong WW, Springer MS (2001) Resolution of the early placental mammal radiation using Bayesian phylogenetics. *Science* 294: 2348-2351.

Murphy WJ, Pringle TH, Crider TA, Springer MS, Miller W (2007) Using genomic data to unravel the root of the placental mammal phylogeny. *Genome Research* 17: 413-421.

Murphy WJ, Eizirik E (2009) Placental mammals (Eutheria). In: *The timetree of Life*, Hedges, SB, Kumar, S (eds.). Oxford University Press. Pp 471-474.

Nishihara H, Terai Y, Okada N (2002) Characterization of novel Alu- and tRNA-related SINEs from the tree shrew and evolutionary implications of their origins. *Molecular Biology and Evolution*. 19: 1964-1972.

Page RDM & Charleston MA (1997) From gene to organismal phylogeny: reconciled trees and the gene tree/species tree problem. *Molecular Phylogenetics and Evolution* 7: 231-240.

Pattengale ND, Swenson KM, Moret BME (2010) Uncovering hidden phylogenetic consensus. *Bioinformatics Research and Applications* 6053: 128-139.

Phillips MJ, Penny D (2003) The root of the mammalian tree inferred from whole mitochondrial genomes. *Molecular Phylogenetics and Evolution* 28: 171-185.

- Poe S, Chubb AL (2004) Birds in a bush: five genes indicate explosive evolution of avian orders. *Evolution* 58: 404-415.
- Prasad AB, Allard MW (2008) Confirming the phylogeny of mammals by using large comparative sequence data sets. *Molecular Biology and Evolution* 25: 1795-1808.
- Roleants K, Bossuyt F (2005) Aracaeobatrachian paraphyly and Pangaeon diversification of crown-group frogs. *Systematic Biology* 54:111-126.
- Sambrook J, Russell DW (2001) *Molecular cloning: a laboratory manual*. Cold Springs Harbor Laboratory Press. Cold Springs Harbor, New York.
- Sanderson MJ, Shaffer HB (2003) Troubleshooting molecular phylogenetic analyses. *Annual Review of Ecology and Systematics* 33: 49-72.
- Schmitz J, Ohme M, Zischler H (2000) The complete mitochondrial genome of *Tupaia belangeri* and the phylogenetic affiliation of Scandentia to other eutherian orders. *Molecular Biology and Evolution* 17: 1334-1343.
- Shimodaira H (2002) An approximately unbiased test of phylogenetic tree selection. *Systematic Biology* 51: 492-508.
- Shimodaira H, Hasegawa M (2001) CONSEL: for assessing the confidence of phylogenetic tree selection. *Bioinformatics* 17: 1246-1247.
- Shoshani J, McKenna MC (1988) Higher taxonomic relationships among extant mammals based on morphology, with selected comparisons of results from molecular data. *Molecular Phylogenetics and Evolution*. 9: 572-584.
- Simpson GG (1959) Mesozoic mammals and the polyphyletic origin of mammals. *Evolution* 13: 405-414.
- Springer MS, Cleven GC, Madsen O, de Jong WW, Waddell VG, Amrine HM, Stanhope MJ (1997) Endemic African mammals shake the phylogenetic tree. *Nature* 388: 61-63.
- Stamatakis A (2006) RAxML-VI-HPC: maximum likelihood-based phylogenetic analyses with thousands of taxa and mixed models. *Bioinformatics* 22:2688–2690
- Stamatakis A, Hoover P, Rougemont J (2008) A rapid bootstrap algorithm for the RAxML web-servers. *Systematic Biology* 57:758–771
- Stanhope MJ, Waddell VG, Madsen O, de Jong W, Hedges SB, Cleven SB, Kao D, Springer MS (1998) Molecular evidence for multiple origins of Insectivora and for a new order of

endemic African insectivore mammals. *Proceedings of the National Academy of Sciences, USA* 95: 9967-9972.

Thomson RC, Shaffer HB (2010) Rapid progress on the vertebrate tree of life. *BMC Biology* 8:19.

Townsend TM, Larson A, Louis E, Macey JR (2004) Molecular phylogenetics of Squamata: the position of snakes, amphisbaenians, and dibamids, and the root of the squamate tree. *Systematic Biology* 53: 735-757.

van Tuinen M (2009) Advanced birds (Neoaves). In: *The timetree of Life*, Hedges, SB, Kumar, S (eds.). Oxford University Press. Pp 419-422.

Vidal N, Hedges SB (2004) Molecular evidence for a terrestrial origin of snakes. *Proceedings of the Royal Society of London B (suppl.)* 271: S226-S229.

Vidal N, Hedges SB (2005) The phylogeny of squamate reptiles (lizards, snakes and amphisbaenians) inferred from nine nuclear protein-coding genes. *Comptes Rendus Biologies* 328: 1000-1008.

Waddell PJ, Kishino H, Ota R (2001) A Phylogenetic foundation for comparative mammalian genomics. *Genome Informatics* 12: 141-154.

Wilkinson M (1994) Common cladistic information and its consensus representation: reduced Adams and reduced cladistic consensus trees and profiles. *Systematic Biology* 43: 343-368.

Wilkinson M (1996) Majority-rule reduced consensus trees and their use in bootstrapping. *Molecular Biology and Evolution* 13: 437-444.

Yates TL, Salazar-Bravo J, Dragoo JW (2004) The importance of the tree of life to society. In: Cracraft J, Donoghue MJ. *Assembling the tree of life*. Oxford University Press, New York.

Zardoya R, Meyer A (2001) On the origin and phylogenetic relationships among living amphibians. *Proceedings of the National Academy of Sciences USA* 98: 7380-7383.

Zhang P, Chen Y-Q, Zhou H, Wang X-L, Qu L-H (2003) The complete mitochondrial genome of a relic salamander, *Ranodon sibiricus* (Amphibia: Caudata) and implications for amphibian phylogeny. *Molecular Phylogenetics and Evolution* 28: 620-626.

Zhang P, Zhou H, Chen Y-Q, Liu Y-F, Qu L-H (2005) Mitogenomic perspectives on the origin and phylogeny of living amphibians. *Systematic Biology* 54: 391-400.

Zhu YY, Machleder EM, Chenchik A, Li R, Siebert PD (2001) Reverse transcriptase template switching: a SMART approach to full-length cDNA library construction. *BioTechniques* 30: 892-897.

Table 1: Results from the Approximately Unbiased topology tests comparing the two alternative amphibian hypotheses for the four data-types. X” denotes the best tree, and statistically identical trees to this best tree are highlighted in orange.

-lnL	NUCL	N12	DEGEN1	AA
MONOPHYLETIC	p=0.001	X	X	p=0.274
POLYPHYLETIC	X	p=0.389	p=0.347	X

Table 2: Summary table of past studies on higher mammal relationships.

	CITATION	MORPH?	DNA?
Theria (palcental-marsuapial)	Novacek 1992a	X	
	Novacek 1992b	X	
	Shoshani & McKenna 1998	X	
	Killian et al 2001		X
	Phillips & Penny 2003		X
	Hugall et al 2007		X
	Luo 2007	X	
Marsupionta (marsupial-monotreme)	Kuhne 1973	X	
	Kirsch & Mayer 1988		X
	Janke et al 1996		X
	Strimmer and von Haeseler 1996		X
	Janke et al 1997		X
	de Jong 1998		X
	Janke 2001		X
	Janke 2002		X
	Zardoya et al 2003		X

Table 3: Traditional classification of placental mammals, as proposed by McKenna & Bell (1997)

Cohort	Magnorder	Grandorder	Order
Placentalia	Xenarthra		anteaters, sloths, armadillos
	Epitheria	Anagalida	lagomorphs, rodents, elephant shrews
		Ferae	carnivorans, pangolins
		Lipotyphla	insectivorans
		Archonta	bats, primates, colugos, treeshrews
		Ungulata	aardvark
		Tublidentata	
		Eparrectocyona	cetaceans, artiodactyls
		Altungulata	perrisodactyls, elephants, manatees, hyraxes

Figure 1: Phylogenetic relationships within Actinopterygii (ray-fin fishes). Phylogeny shown is from the RAxML analysis, and bootstrap support values are shown next to each node, and are only indicated if they are ≥ 50 . An * indicates bootstrap=100. Labels at nodes indicate divergence between major groups as described from Donoghue & Benton (2007).

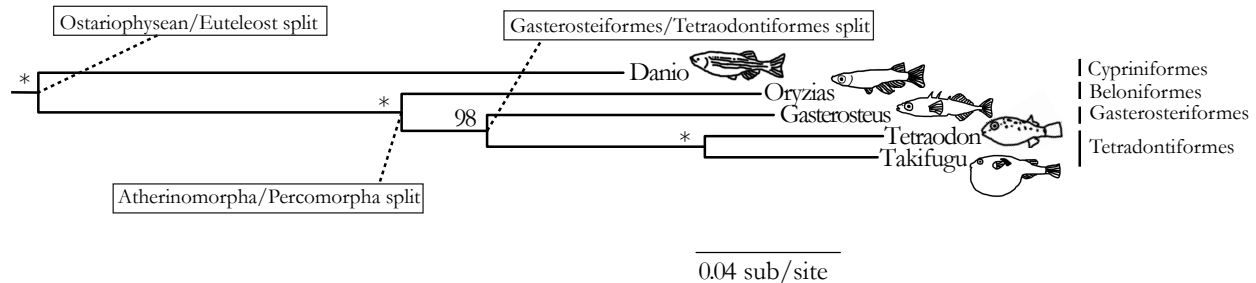


Figure 2: Phylogenetic relationships within Amphibia. RAxML analyses are shown of the A) NUCL dataset, and B) N12 dataset. Bootstrap support values are shown next to each node, and are only indicated if they are ≥ 50 . An * indicates bootstrap=100.

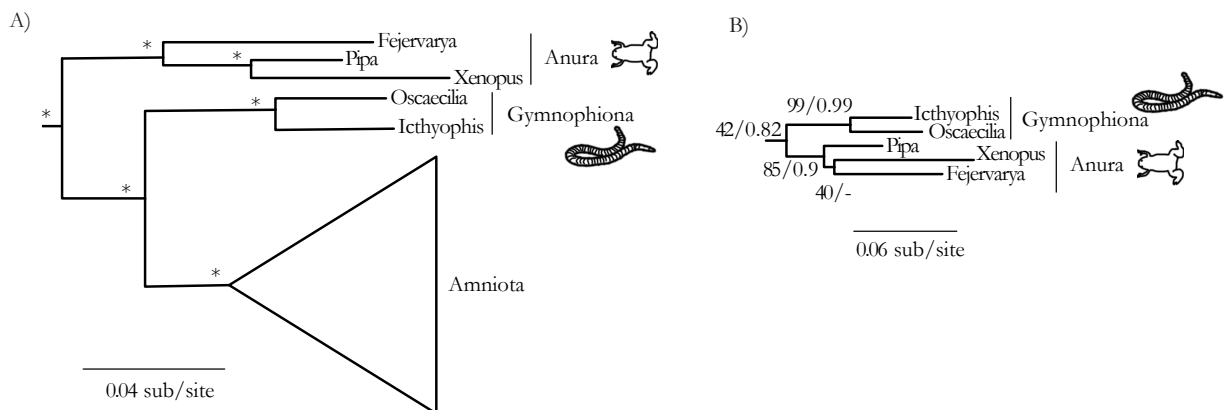


Figure 3: Phylogenetic inferences of Squamata from past studies. A) Camp (1923) based on morphology. Fossil taxa denoted by daggers (†) B) Estes et al. (1988) based on morphology. C) Conrad (2008) based on morphology. D) Townsend et al. (2004) based on molecular data. E) Vidal & Hedges (2005) based on molecular data.

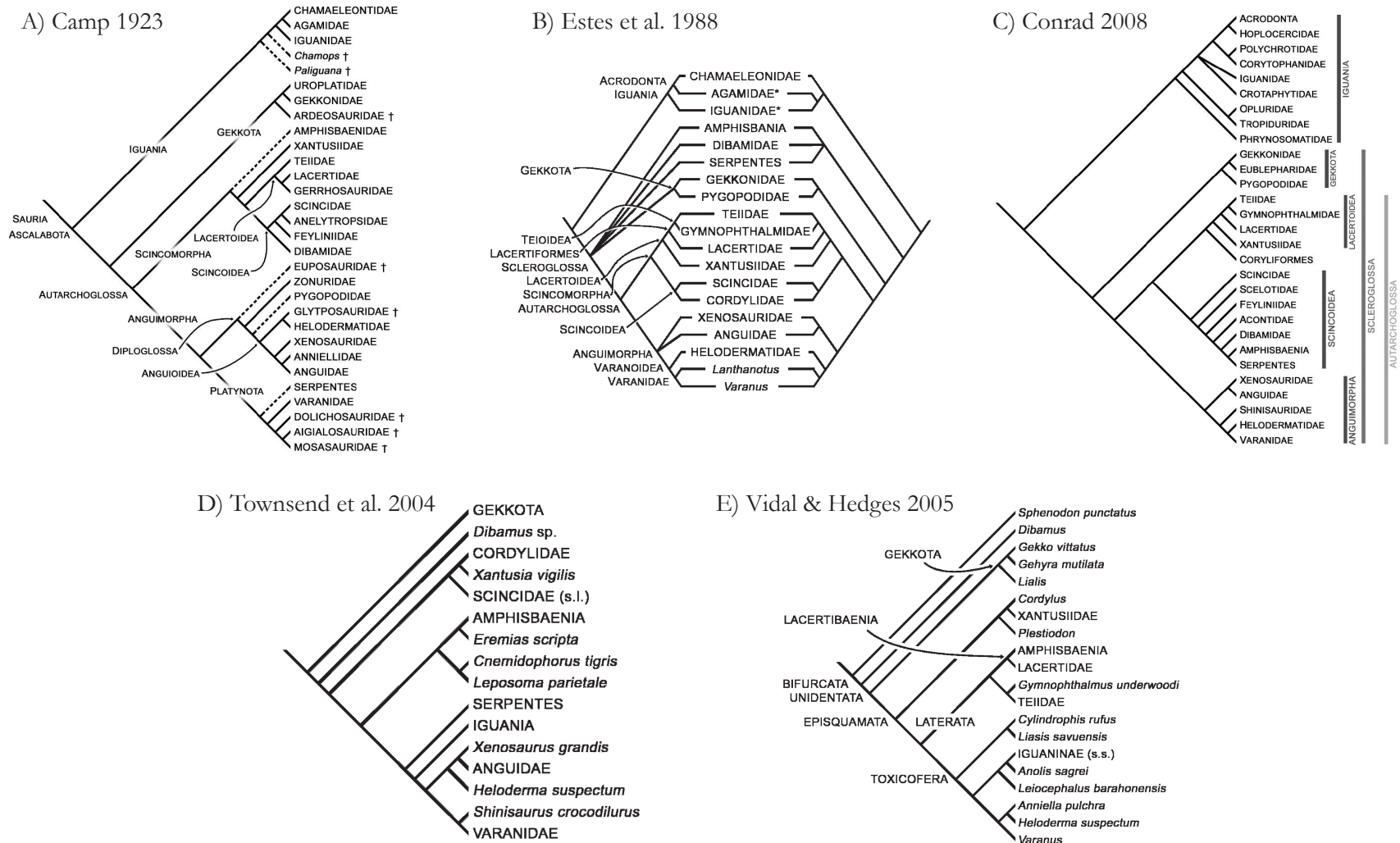


Figure 4: Phylogenetic relationships within Squamata. The tree shown is from the RAxML analysis of the NUCL dataset. Bootstrap support values are shown next to each node, and are only indicated if they are ≥ 50 . An * indicates bootstrap=100.

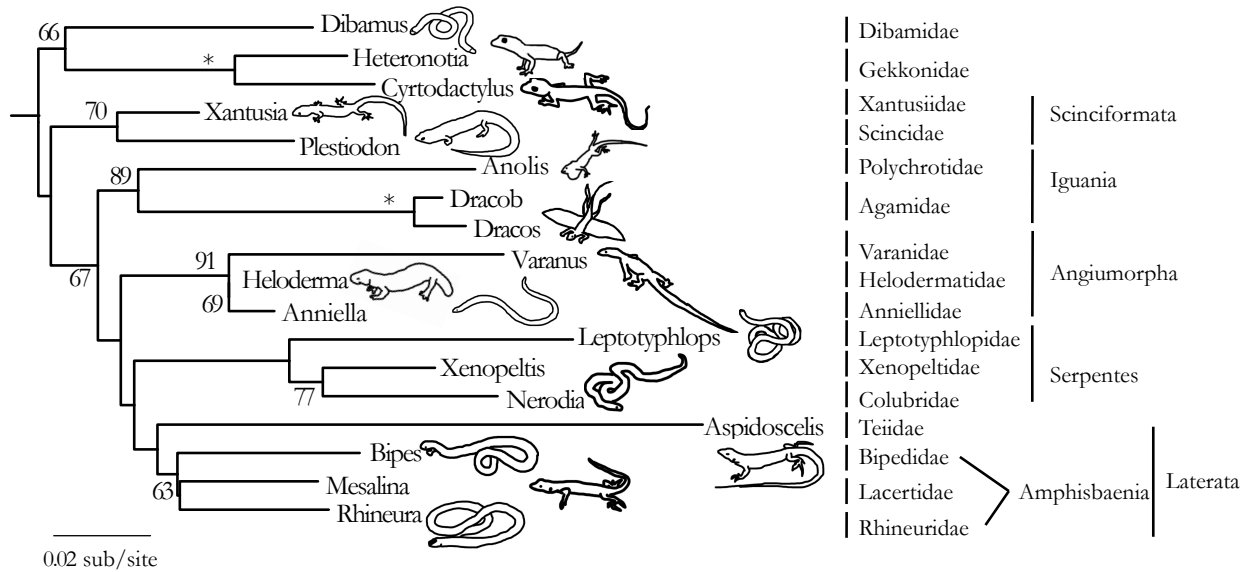


Figure 5: Phylogenetic relationships within Aves. The tree shown is from the RAxML analysis of the NUCL dataset. Bootstrap support values are shown next to each node, and are only indicated if they are ≥ 50 . An * indicates bootstrap=100.

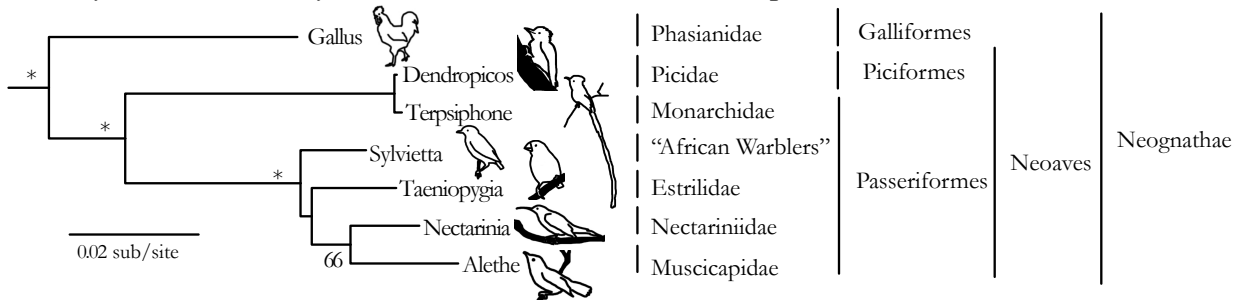


Figure 6: Phylogenetic relationships within Mammalia. The tree shown is from the RAxML analysis of the NUCL dataset. Bootstrap support values are shown next to each node, and are only indicated if they are ≥ 50 . An * indicates bootstrap=100.

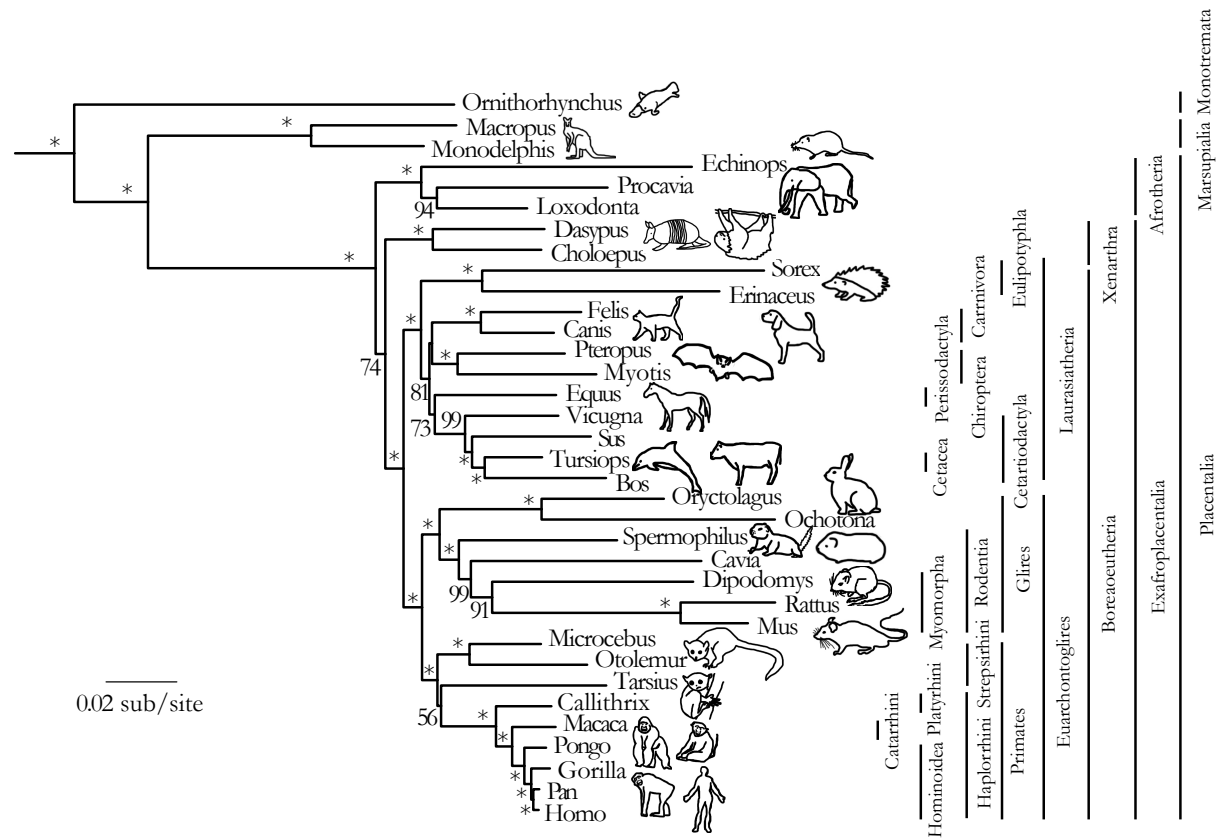


Figure 7: Phylogenetic placement of Scandentia (tree shrews, *Tupaia*) using RAxML-EPA. Two phylogenetic positions are recovered A) sister to Glires, B) sister to Primates. Numbers along branches correspond to RAxML weights (bootstrap values).

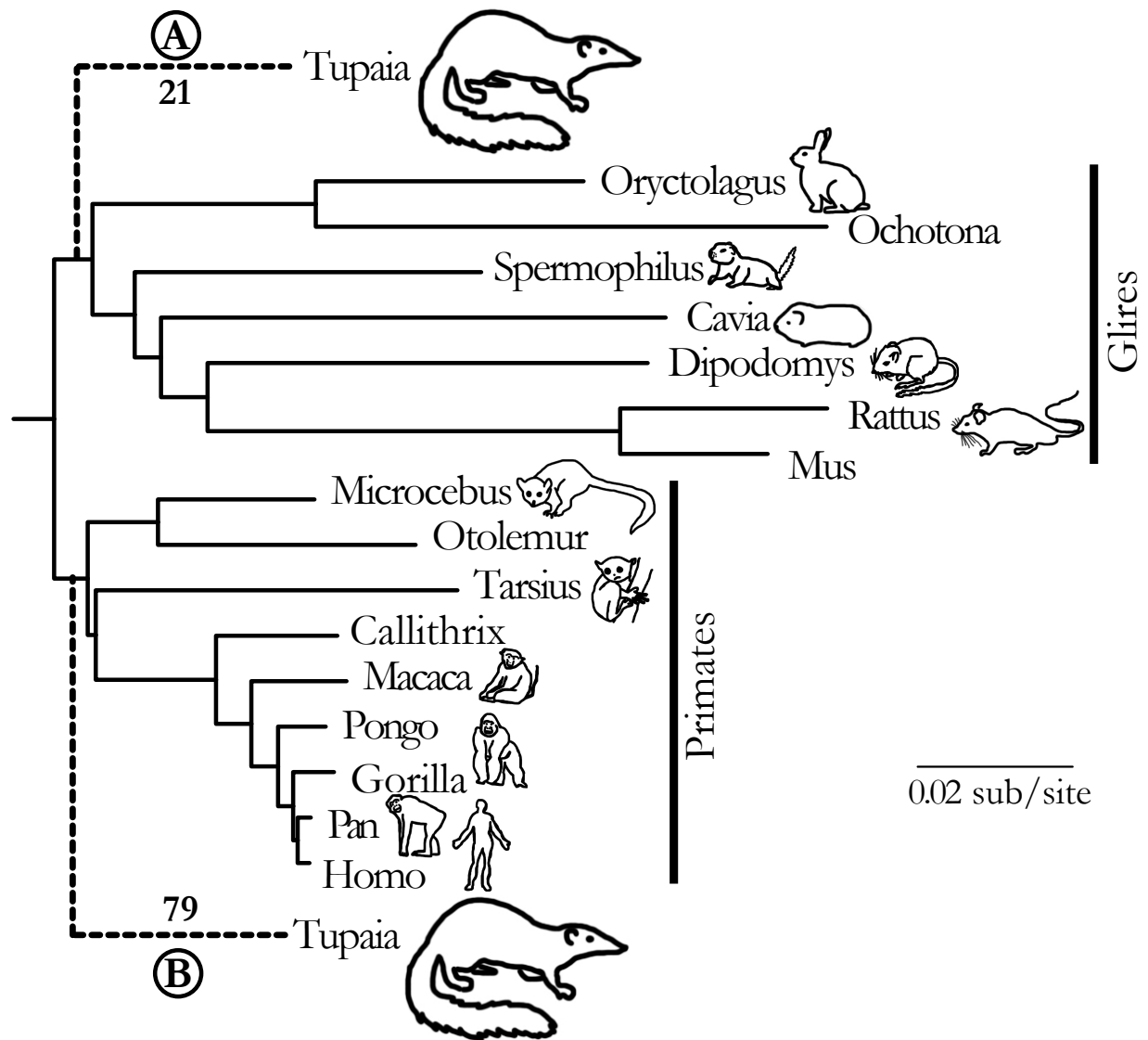


Figure 8: Plot comparing average GC content (%GC) between Primates and Glires across all sites. %GC for each site was calculated for each group, and the difference calculated for each site. A value of “0” indicates no difference, while a positive value indicates that Primates have a higher %GC, and a negative value indicates that Glires have a higher %GC. The red line represents the Lowess trend line of the data.

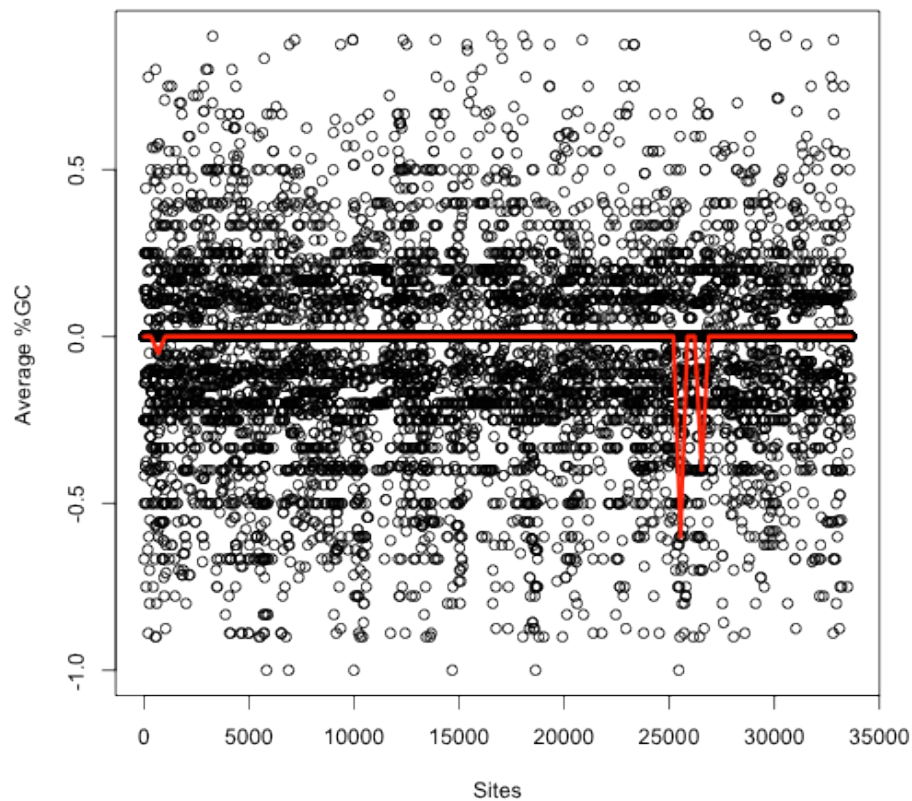


Figure 9: Plot comparing average % missing data (%missing) between Primates and Glires across all sites. %missing for each site was calculated for each group, and the difference calculated for each site. A value of “0” indicates no difference, while a positive value indicates that Primates have a higher %missing, and a negative value indicates that Glires have a higher %missing. The red line represents the Lowess trend line of the data.

