LBNL-57512
Preprint

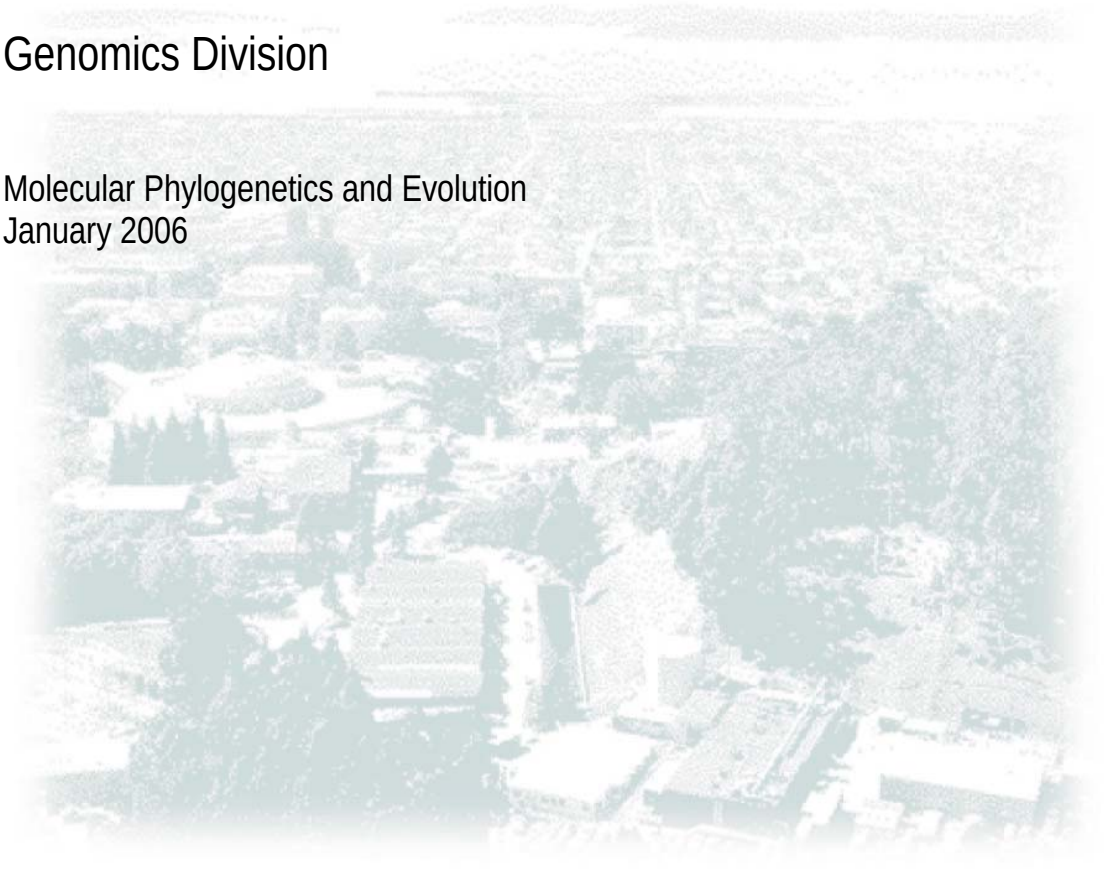
ERNEST ORLANDO LAWRENCE
BERKELEY NATIONAL LABORATORY

The phylogeny of Mediterranean tortoises and their close relatives based on complete mitochondrial genome sequences from museum specimens

James F. Parham, J. Robert Macey, Theodore J. Papenfuss, Chris R. Feldman, Oguz Turkozan, Rosa Polymeni, Jeffrey Boore

Genomics Division

Molecular Phylogenetics and Evolution
January 2006



The phylogeny of Mediterranean tortoises and their close relatives based on complete mitochondrial genome sequences from museum specimens

James F. Parham[a,b,*], J. Robert Macey[a,c], Theodore J. Papenfuss[c], Chris R. Feldman[d], Oguz Türkozan[e], Rosa Polymeni[f], Jeffrey Boore[a,g]

[a] *Department of Evolutionary Genomics, DOE Joint Genome Institute and Lawrence Berkeley National Laboratory, 2800 Mitchell Drive, Walnut Creek, CA, 94598, USA*

[b] *Museum of Paleontology, University of California, Berkeley, CA, 94720, USA*

[c] *Museum of Vertebrate Zoology, University of California, Berkeley, CA, 94720, USA*

[d] *Department of Biology, Utah State University, Logan, UT, 84322, USA*

[e] *Adnan Menderes University, Faculty of Science and Arts, Department of Biology, 09010 Aydin, Turkey*

[f] *Department of Zoology and Marine Biology, Faculty of Biology, School of Science, University of Athens, GR 157 84 Panepistimioupolis, Athens, Greece*

[g] *Department of Integrative Biology, 3060 Valley Life Science Building, University of California, Berkeley, CA, 94720, USA*

Corresponding author: James Parham, Department of Evolutionary Genomics, DOE Joint Genome Institute, 2800 Mitchell Drive, Walnut Creek, CA 94598, USA

parham@socrates.berkeley.edu; phone: 925-296-5818; Fax 510-642-1822.

Abstract

As part of an ongoing project to generate a mitochondrial database for terrestrial tortoises based on museum specimens, the complete mitochondrial genome sequences of 10 species and a ~14 kb sequence from an eleventh species are reported. The sampling of the present study emphasizes Mediterranean tortoises (genus *Testudo* and their close relatives). Our new sequences are aligned, along with those of two testudinoid turtles from GenBank, *Chrysemys picta* and *Mauremys reevesii*, yielding an alignment of 14,858 positions, of which 3,238 are parsimony informative. We develop a phylogenetic taxonomy for *Testudo* and related species based on well-supported, diagnosable clades. Several well-supported nodes are recovered, including the monophyly of a restricted *Testudo, T. kleinmanni* + *T. marginata* (the *Chersus* clade), and the placement of the enigmatic African pancake tortoise (*Malacochersus tornieri*) within the predominantly Palearctic greater *Testudo* group (*Testudona* tax. nov.). Despite the large amount of sequence reported, there is low statistical support for some nodes within *Testudona* and so we do not propose names for those groups. A preliminary and conservative estimation of divergence times implies a late Miocene diversification for the *testudonan* clade (6-12 million years ago), matching their first appearance in the fossil record. The multi-continental distribution of *testudonan* turtles can be explained by the establishment of permanent connections between Europe, Africa, and Asia at this time. The arrival of *testudonan* turtles to Africa occurred after one or more initial tortoise invasions gave rise to the diverse (>25 species) '*Geochelone* complex.' Two unusual genomic features are reported for the mtDNA of one tortoise, *M. tornieri*: (1) *nad4* has a shift of reading frame

that we suggest is resolved by translational frameshifting of the mRNA on the ribosome during protein synthesis and (2) there are two copies of the control region and *trnF*, with the latter having experienced multiple nucleotide substitutions in a pattern suggesting that each is being maintained by selection.

Keywords: Reptilia, Testudines, Testudinidae, *Testudo*, *Agrionemys*, *Indotestudo*, *Malacochersus*, mitochondrial genomes, fossils, systematics, gene duplications, taxonomy, biogeography, Africa, Asia, Europe, Mediterranean

1. Introduction

1.1 Preface

Tortoises (Testudinidae Gray, 1825) are a clade of terrestrial turtles that originated in Asia in the early Cenozoic (~60 million years ago) and then rapidly dispersed to Europe, Africa, and the New World (Holroyd and Parham, 2003) alongside early radiations of placental mammals (Beard, 1998). Despite boasting a rich fossil record (Auffenberg, 1974) and diverse living members (Ernst and Barbour, 1989), the evolutionary relationships of tortoises remain poorly known. Sadly, the uncertainties about the history of tortoises are matched by uncertainties about their future. Tortoises face serious threats to their survival throughout their range including habitat destruction and over-harvesting for food, traditional medicine, and the pet trade. Their global decline is reflected by the fact that Testudinidae is the only polytypic family of non-marine turtles afforded blanket CITES protection.

As part of an effort to generate a robust phylogeny for tortoises to aid in the reconstruction of their paleobiogeography and morphological evolution, we are assembling a large mitochondrial database generated from museum specimens representing all major tortoise lineages. Our first report of this work focuses on the turtles of the genus *Testudo* Linnaeus, 1758 and their closest relatives (Fig. 1). For this study, we report the complete mitochondrial genomes of 10 museum specimens including *Testudo* species and closely related tortoises as well as a ~14 kb sequence from an eleventh tortoise species (Appendix 1). Beyond identifying well-supported, readily-diagnosable nodes to anchor useful clade names, we also use our phylogeny to reassess

aspects of the paleobiogeography of these tortoises. Some unique genomic features of the pancake tortoise, *Malacochersus tornieri* (Siebenrock, 1903), are also discussed.

1.2. Establishing a working concept of Testudo

In the first work to use valid binomial names for animals, the 10th edition of *Systema Naturae* (Linnaeus, 1758), every turtle species was placed in the genus *Testudo*. For the rest of the 18th century, authors automatically put newly described turtle species into *Testudo* which, in the scientific literature, simply meant 'turtle.' But in 1800, Brongniart split sea turtles out of *Testudo* into their own genus: *Chelonia*. This opened the doors to over two hundred years of subjective splitting and, since that time, the content of *Testudo* has been aggressively winnowed as species groups were carved into many new genera. Given the ever-shrinking and shifting content of the genus *Testudo,* we feel the need to establish an explicit working concept of the name. Lapparent de Broin (2001) recently suggested that the genus name *Testudo* be restricted to just three terrestrial species (from the Mediterranean region, the Caucasus, and Iran) that have kinetic hinges in the ventral shell (plastron): *Testudo graeca* Linnaeus 1758 (type species), *Testudo marginata* Schoepff 1792, *Testudo kleinmanni* Lortet 1883. We provisionally refer to this group as *Testudo.* A recent phylogenetic study based on mitochondrial DNA (mtDNA) using sequences of *rrnS* (Kuyl et al., 2002) tested the relationships of these species to other tortoises and found strong support for the monophyly of *Testudo* as recognized here.

Splitting has also occurred at the species level, with some authors (e.g., Perälä, 2001; Perälä and Bour, 2004) recognizing morphometrically diagnosable populations of

Testudo as distinct taxa. We can not evaluate the validity of these taxonomic decisions without a better understanding of the genetic variation within and among populations, especially where the ranges of proposed taxa come into contact. Ongoing studies in this area (Kuyl et al., 2002; Harris et al., 2003; Fritz et al., in press; Kuyl et al., 2005) suggest that at least some of the proposed divisions within *Testudo* are premature. Pending more data, we feel justified in employing a conservative species-level taxonomy that retains the three 'classic' species of *Testudo*, all of which are well-recognized, uncontroversial, monophyletic groups. The alternative would be to reinforce untested taxonomic conclusions that may potentially cloud the scientific and conservation literature with ephemeral taxa.

1.3. The tortoises formerly known as Testudo

The mtDNA study of Kuyl et al. (2002) found moderate support for a close relationship between *Testudo* and three Eurasian lineages: (1) *Indotestudo* Lindholm, 1929, a genus that includes three species from India and southeast Asia; (2) *Agrionemys horsfieldii* (Gray 1844) from the desert steppes of Central Asia; and (3)'*Testudo*' *hermanni* Gmelin 1789 from coastal regions of the northern Mediterranean. These three lineages represent the most recent taxonomic splits from *Testudo* (Khozatsky and Mlynasrski 1966; Bour 1980; Lapparent de Broin, 2001).

The taxonomic status of '*T.*' *hermanni* is not resolved. Some authors (Gmira, 1993, 1995; Kuyl et al., 2002) recommended placing the European tortoise, '*T.*' *hermanni*, in the genus *Agrionemys* Khozatsy and Mlynasrski, 1966 with *A. horsfieldii*, but this is based only on weak morphological or molecular support for their association to

the exclusion of other tortoises. Lapparent de Broin (2001) did not accept this scheme, but did consider '*T.*' *hermanni* as separate from *Testudo*. In this study we refrain from giving *hermanni* any official genus name, but refer to it as a '*Testudo*' (with single quotes to denote uncertainty) pending more definitive evidence of its evolutionary affinities.

1.4. The phylogenetic position of the Testudo *group among tortoises*

At a broader scale, the phylogenetic relationships of *Testudo*, and species recently split from *Testudo*, to the diverse tortoise fauna from sub-Saharan Africa are better known. It is clear that most of the sub-Saharan tortoises belong to a '*Geochelone* complex' that is separate from the greater *Testudo* group based on morphological and molecular evidence (Crumly, 1984; Meylan and Sterrer, 2000; Lapparent de Broin, 2000a, 2001; Kuyl et al., 2002; Takahashi et al., 2003; Parham et al., unpublished data). The phylogeny of this '*Geochelone* complex' will be treated elsewhere; however, we do address the phylogenetic position of one sub-Saharan species, *Malacochersus tornieri*. Crumly (1984) hypothesized that *M. tornieri* is closely related to *Testudo* and allies based on shared derived characters of head scalation and tracheal morphology. In other aspects, *M. tornieri* is highly and uniquely derived, reflecting its ecology as a crevice specialist. For example, unlike other tortoises, it has an extremely flat and flexible shell (hence the common name 'pancake tortoise'). We test Crumly's hypothesis about the possible affinity of the enigmatic *M. tornieri* to the greater *Testudo* group with molecular data.

2. Materials and Methods

2.1. Specimen information

Our sampling includes 11 museum specimens (Appendix 1). Four of these represent the three *Testudo* species, *T. graeca*, *T. marginata*, and *T. kleinmanni*. Of these samples, our *T. graeca* and *T. marginata* samples have known localities. The two specimens of *Testudo graeca* are from different parts of its range, North Africa and Asia. Kuyl et al. (2002) identified distinct genetic groups from these regions. We include two pet trade samples of *Indotestudo* that were donated to the Museum of Vertebrate Zoology without locality data. *Indotestudo* species are very similar and there was some debate as to whether there are two or three species (Iverson et al., 2001). Consequently, the identification of pet trade specimens must be done carefully. Based on their morphology, one of us (JFP) identified them as *Indotestudo elongata* (Blyth, 1853) and *I. forstenii* (Schlegel and Muller, 1844) and confirmed this by comparing their sequences to *cob* sequences of other *Indotestudo* from GenBank (Iverson et al., 2001; **AY434561**, **AY434643**). In an alignment of 1,115 positions, our *I. elongata* and *I. forstenii* were just 0.4% and 1.1% different, respectively, from those identified as the same species reported by Iverson et al. (2001). The high degree of sequence similarity between our specimens and other individuals identified by different authors gives us some confidence in the identification of our specimens as well as those used in Iverson et al. (2001). However, we recognize that this assignment is tentative because Iverson et al. (2001) lacked vouchered specimens. We also include one sample of *A. horsfieldii* from the Kopet-Dagh region of Turkmenistan and one sample of '*T.*' *hermanni* from the European part of Turkey (Thrace). Our specimen of *Malacochersus tornieri* was donated to the Museum of Vertebrate Zoology as a frozen specimen without any locality data, but we are confident about its identification because *M. tornieri* is so specialized that it can not be easily

mistaken for any other animal (Fig. 1A). For outgroups we included one sample of the '*Geochelone* complex', *Geochelone pardalis* (Bell, 1828) from Somaliland (formerly part of Somalia), and a pet trade specimen of the basal tortoise *Manouria emys* (Schlegel and Muller, 1844). As with the *M. tornieri*, we do not doubt the identification of the *Manouria emys* specimen. All specimens are preserved at either the Museum of Vertebrate Zoology, University of California, Berkeley, CA, or the California Academy of Sciences, San Francisco, CA (Appendix 1). All specimens are preserved in formalin, but only after tissues were frozen. For outgroups we used the complete mitochondrial genomes of two testudinoid turtles from GenBank (*Chrysemys picta* [Gray, 1844], **NC002073**; and *Mauremys reevesii* [Gray, 1831], **NC006082**)

2.2. Laboratory protocols

Genomic DNA was extracted from liver or muscle using the Qiagen QIAamp tissue kit. Amplification of genomic DNA was conducted using rTth long PCR enzyme (Applied Biosystems) with a denaturation at 94º C for 15 sec, annealing at 46-50º C for 20 sec, and extension at 68º C for 60 sec for a total of 38 cycles, followed by an additional extension at 72º for 12 min. Negative controls were run on all amplifications to check for contamination. Initial amplifications were conducted using primers described in Macey et al. (1997). Perfectly matching primers were then constructed for each taxon based on the DNA sequence of this fragment to complete the amplification of each mtDNA.

Amplification products were sheared randomly into fragments of approximately 1.5 kb by repeated passage through a narrow aperture using a Hydroshear device. After

end-repair, the sheared DNA was gel purified and ligated into pUC18 vector and then transformed into bacterial cells to construct a library of random fragments. Automated colony pickers introduced single clones into bacterial broth in 384-well format. These plasmid clones were processed robotically through rolling circle amplification (Dean et al., 2001; Hawkins et al., 2002), sequencing reactions, and reaction cleanup using SPRI (Elkin et al., 2002). Sequences were determined using ABI3730xl DNA sequencers and then assembled based on overlap to form deep contigs.

2.3. Phylogenetic analyses

DNA sequences were aligned manually. Protein-coding genes were constrained to align by codon and tRNA-coding genes were constrained to align by regions of potential secondary structure (Kumazawa and Nishida, 1993; Macey and Verma, 1997). We excluded highly variable regions that were ambiguously aligned that encompass all of the control region, 165 positions from other non-coding regions, 140 positions of *rrnS*, and 342 positions of *rrnL*. A total of 182 positions were excluded from the alignment of tRNA genes: the D-loop is excluded from *trnH* and *trnS*; the T-loop is excluded from *trnE*; and both the D- and T-loops are excluded from the tRNA genes for F, V, L1, I, W, K, R, T, and P. We excluded a total of 282 positions from the protein coding genes *atp8* (21), *nad5* (168), *nad6* (87), and *cob* (6). In the case of *M. tornieri* we used the *trnF* that most closely resembled that of other tortoises and was in the standard vertebrate position adjacent to *rrnS* (*trnF2*). The final alignment contains 14,858 positions and provides 3,238 parsimony informative characters.

We used maximum parsimony (MP; Farris, 1983), maximum likelihood (ML; Felsenstein, 1981), and Bayesian inference (BI; Larget and Simon, 1999) phylogenetic methods to infer phylogenetic trees. We conducted both MP and ML phylogenetic analyses in PAUP* 4.0b10 (Swofford, 2002) and BI analyses with MrBayes 3.0b4 (Huelsenbeck and Ronquist 2001). We executed MP analyses with the branch and bound search option, which guarantees an exact solution. To assess nodal support, we used the bootstrap resampling method (Felsenstein, 1985) employing 1000 pseudoreplicates of heuristic searches using TBR branch swapping and 100 random sequence additions pre replication in PAUP*. We obtained decay indices (="branch support'' of Bremer, 1994) for all nodes. As an additional test of clade significance we used the Wilcoxon rank-sum test (Templeton, 1983) as outlined by Schulte et al. (1998), Macey et al. (1999) and Lee (2000).

To determine the most appropriate model of DNA substitution for reconstructing tortoise relationships under ML, we evaluated the fit of various models of molecular evolution to our data via the Akaike Information Criterion (AIC; Akaike, 1974) with the program Modeltest 3.06 (Posada and Crandall, 1998). The AIC has recently been shown to be a superior method of model selection than the hierarchical likelihood ratio test (Posada and Buckley, 2004). We performed ML analyses under the optimal model (Appendix 2) with the heuristic search algorithm using TBR branch swapping with 10 random sequence additions, simultaneously estimating parameter values (with 10 Γ rate categories) and tree topology (i.e., no initial parameter estimates or starting tree). We then successively re-estimated parameter values and searched for trees until we obtained a stable topology and ML score (Wilgenbusch and de Queiroz, 2000). We assessed nodal

support with 10 bootstrap pseudoreplicates using TBR branch swapping and 10 random sequence additions.

We also performed ML-based BI analyses to search for additional tree topologies. Because MrBayes can perform singular phylogenetic analyses using different models of evolution we assessed the best fit model of evolution for each mtDNA locus via the AIC with the program MrModeltest 2.1 (Nylander 2004). However, to avoid over-parameterization, we combined mitochondrial loci into the same data partition if they belonged the same functional type (either rRNA, tRNA, or coding DNA) and conformed to the same model of evolution (Appendix 2). We then performed mixed-model BI tree searches, allowing separate parameter estimates under the chosen models of DNA substitution for each data partition. We did not specify nucleotide substitution model parameters or a topology *a priori*. We ran BI analyses for 3×10^6 generations using the default temperature (0.2) with four Markov chains per generation, sampling trees every 100 generations. To determine when Markov chains had converged on stable likelihood values, we plotted the -lnL scores against the number of generations (Huelsenbeck and Ronquist, 2001). We then computed a 50 % majority rule consensus tree after excluding those trees sampled prior to the stable equilibrium (after the first 1×10^5 generations). Nodal support is given by the frequency of the recovered clade, which corresponds to the posterior probability of that clade under the assumed models of sequence evolution (Rannala and Yang, 1996; Huelsenbeck and Ronquist, 2001).

2.4. Polytomy tests

We tested whether the low statistical support for some of our nodes was the result of a "hard polytomy" (Maddison, 1989) resulting from a simultaneous set of branching events vs. a "soft polytomy" which simply indicates an inability to resolve the true pattern of bifurcating branches. We analyzed the sequence data with the bootstrap version of the polytomy test developed by Jackman et al. (1999) and later implemented by McGuire (2001). This test is based on the bootstrap values of every possible four taxon data set comprised of one outgroup (here *Geochelone pardalis*) and three of the taxa involved in the clade that garnered low statistical support (i.e., excluding well-supported nodes such as those between *Testudo* and *Indotestudo* samples). The principle assumption of this test is that if the recovered polytomy truly represents approximately simultaneous branching events, then restricted data sets should not recover higher than random statistical support because the removal of taxa would not create significantly longer branches. To generate the critical bootstrap value for our data set we randomized the characters within taxa 100 times in MacClade (Maddison and Maddison, 2001), generated bootstrap values for each node (n = 300), and then determined which value represented the 95^{th} percentile.

3. Results

3.1. Structural features of tortoise mitochondrial genomes

The 10 complete mtDNA sequences range in size from 16,455 to 19,438 nucleotides. The bulk of this length variation occurs in the control regions which range from 948 to 3,885 nucleotides, with the larger control regions caused predominantly by repeated sequences. We note two very unusual genomic features. First, all of these

tortoises share a nucleotide insertion at an identical position in *nad3* that would be predicted to cause a shift in the reading frame, exactly as has been reported previously for the turtle *Chrysemys picta* and some birds (Mindell et al., 1998). The phenomenon of translational frameshifting, whereby the ribosome accommodates by passing over the additional out-of-frame nucleotide in the mRNA has been studied in some systems (Pande et al., 1995; Farabaugh and Vimaladithan, 1998; Hansen et al. 2003), and perhaps this is the mechanism here, although we cannot rule out the alternative of RNA editing to remove this nucleotide from the transcript. This frameshift does not appear in the mtDNAs of alligators and squamate reptiles, so one must infer that the insertion occurred in the same position twice, in the lineage leading to turtles and that leading to birds or, alternatively, that this is ancestral to the larger group and has been repeatedly lost in various lineages. In addition, one tortoise (*M. tornieri*) has an additional frameshift in *nad4*; the reading frame is interrupted by one out-of-frame nucleotide, the sequence determined is unambiguous and from several independent clones as well as two separate amplifications, and the conceptual translation in the two reading frames before and after this nucleotide insertion point are well conserved.

M. tornieri also is unusual in having a tandem duplication of the control region and *trnF* (Fig. 2A,B). This is the first gene duplication reported for any turtle mtDNA. The two control regions are divergent and only the 1683 bp one adjacent to *trnP* appears functional (correpsonds to other turtle control regions in GenBank). The other ‘control region’ (2017 bp) does not correspond well to any sequence on GenBank. However, most unusually, it appears that each of the duplicated tRNAs is being maintained by selection, since all of the 16 nucleotide differences between the two have occurred without

disrupting potential secondary structure (Fig. 2C) or altering the anticodon. There are 34 paired, three anticodon, and 32 unpaired nucleotides, so to assume the alternative, that one is a pseudogene and that the nucleotide substitutions have occurred randomly, requires accepting that there were 16 random substitutions hitting 32 particular positions without hitting any of the other 37 that are presumed to be essential. The odds of this are approximately one in 10 million. It is generally thought that one of the descendents of a gene duplication will become a pseudogene and decay away by mutational processes unless (1) additional gene dosage is provided and selected for, (2) one of the descendents adopts a new function, or (3) "subfunctionalization" occurs, whereby the two copies divide the original function in such a way that the organism depends on having both copies (Lynch and Force, 2000). In order to test the commonality of this phenomenon, we queried GenBank for all mitochondrial genomes that have the same gene annotation appearing more than once, then examined these individually in detail. Many were simple annotation errors, but some correctly show gene duplications. Six species (Campbell and Barker, 1999; Townsend and Larson, 2002; Kumazawa and Endo, 2004; Lavrov et al., 2004; Yokobori et al. 2004; see also GenBank record **AY636151**) have copies that are identical or nearly so for sequence, so are presumably of very recent origin and are of uncertain fate. Another 10 (Kumazawa et al. 1998; Beagley et al., 1999; Hrbek and Larson, 1999; Macey et al. 2000; Eberhard et al., 2001; Dowton et al, 2003; Mueller et al., 2004; see also GenBank record **AJ421396**) have only one copy that appears to be functional, with others being evident pseudogenes. Some urochordates have a second copy of *trnG* (Yokobori et al., 1999, 2003; Gissi et al., 2004), but in this case there is no evidence that they are the result of a *trnG* duplication, but rather this is to mediate a

change in the genetic code. The only cases we could identify like the one for this turtle, where both copies appear to be maintained by selection, are for *trnV* in the Manila clam, *Venerupis (Ruditapes) philippinarum* (Adams and Reeve, 1850) (**NC_003354**), although this remains undescribed by any publication, and *trnM* in several independent lineages (Le et al. 2000; Passamonti et al. 2003; Yokobori et al., 2003; Boore et al., 2004; Gissi et al., 2004; see also GenBank records **NC_005055** and **NC_003354**), for which one might presume that subfunctionalization into separate roles as initiator of protein translation (with formyl-methionine) vs. elongator for internal peptide positions (with methionine) causes both to be necessary. (The second *trnM* was not recognized by the authors for the platyhelminths, but is clearly present in the sequence. Also, the authors of submission **NC_005055** for the rice frog, *Fejervarya limnocharis* (Gravenhorst, 1829), label the second *trnM* as a pseudogene, although we judge otherwise.) It is not obvious how duplicated copies of *trnF* would subfunctionalize or adopt a novel role, or why any increased gene dosage would be beneficial, since phenylalanine is not a commonly coded amino acid in these mitochondrial genes. It is interesting that the pancake tortoise, *M. tornieri* not only has these two unusual genomic features (a duplicated gene and control region plus an additional translational frameshift), but is also the most unusual morphologically and has an unusually rapid rate of sequence evolution (see below).

3.2. Phylogenetic relationships

Parsimony analyses resulted in two most parsimonious trees of 11,297 steps (CI = 0.726, RI = 0.859). The two parsimony trees conflict regarding the relative placement of ‘*T.*’ *hermanni* and *A. horsfieldii*. In one tree these taxa are sister (Fig. 3), but in the

second tree (not shown) they are paraphyletic. The ML analysis (-lnL = 67939.28) and BI mixed-model analysis with 16 data partitions (-lnL = 67700.28) (Appendix 2) recover '*T.*' *hermanni* and *A. horsfieldii* as sister taxa. In most other respects the MP, ML, and BI trees are identical. The only exception is the placement of *M. tornieri*. In the MP and BI analyses, *M. tornieri* is placed as the sister taxon to the *Indotestudo* species with low statistical support; in the ML analysis *M. tornieri* is placed as sister taxon the rest of the ingroup (*Testudo*, *Indotestudo*, '*T.*' *hermanni*, and *Agrionemys*). In order to compare the MP and BI placement of *M. tornieri* to the ML topology we performed an SH test (Shimodaira and Hasegawa, 1999) against an ML search in which the placement of *M. tonieri* was constrained as the sister taxon to *Indotestudo* (as in the MP and BI) analysis. The constrained ML search under the GTR + Γ + I model produced a single tree (-lnL = 67940.19) with a log-likelihood value that is not significantly different than the unconstrained ML topology using a one-tailed multiple-comparisons LRT (Shimodaira and Hasegawa, 1999) with 100 RELL bootstrap pseudoreplicates ($\delta = 1.8324$, $P = 0.4$). Therefore, although *M. tornieri* is clearly a member of the clade that includes *Testudo* and its close allies, the weak support for the placement of *M. tornieri* from MP, ML, and BI analyses as well as the Wilcoxon rank-sum and SH tests shows that the placement of this taxon within the ingroup is not yet resolved (See Sections 3.3. and 3.4 below).

In all analyses, our samples of *Indotestudo* group together with high statistical support. A clade that includes *M. tornieri*, *Indotestudo*, '*T.*' *hermanni* and *A. horsfieldii* is weakly supported by MP and ML, but strongly supported by BI. That clade is sister to an unambiguously well-supported clade including all *Testudo* species. Within *Testudo*, two well-supported clades are recognized: 1) *T. marginata* + *T. kleinmanni*; 2) African and

Asian *T. graeca*. The monophyly of the *Testudo* + *Indotestudo* + *M. tornieri* + *A. horsfieldii* + '*T.*' *hermanni* relative to the lineages sampled here is strong with *G. pardalis* more closely related to Palearctic tortoises than to *Manouria emys* as predicted by all morphological studies (Crumly, 1984; Meylan and Sterrer, 2000; Takahashi et al., 2003). Finally, the monophyly of tortoises is well-supported by MP and BI, but weakly supported by the likelihood analysis.

3.3. Polytomy tests

The critical value for our polytomy test was 77%, significantly lower than that found by Jackman et al. (1999) and McGuire (2001). We attribute our lower critical value to the higher number of parsimony informative characters presented here or lack of a weighting scheme. Twenty-three of the 37 relevant four taxon statements yielded bootstrap values greater than the critical value, with 16 of these yielding bootstraps even higher than 95% (the 99.66th percentile). Given these results, we can confidently reject the null hypothesis of a hard polytomy for *A. horsfieldii* + *Indotestudo* + *M. tornieri* + *Testudo* + '*T.*' *hermanni* relationships. Thus, the hard polytomy test suggests that uncertainty surrounding relationships within this clade is not simply due to a rapid radiation yielding short internodes characterized by few synapomorphies, but may instead be an artifact of several taxa dividing a potentially large branch. A commonly invoked solution for such a soft polytomy is that additional data should be able to reconstruct the sequence of branching events (Maddison, 1989). Since our alignment is based on complete or nearly complete mitochondrial genomes of all the major lineages within this clade, additional data for testing the poorly-supported nodes should be derived from other

markers (e.g., nuclear DNA). However, we explore another explanation for the effective polytomy presented here: significantly different rates of mitochondrial evolution (see Section 3.4 below).

3.4. Relative rate tests

The pairwise divergences between *M. tornieri* and other taxa of the *A. horsfieldii* + *Indotestudo* + *M. tornieri* + *Testudo* + '*T.*' *hermanni* clade (10.8-12.0%) exceed that found between any other species of this lineage (maximum of 10.5%). Although statistical support for the specific arrangement of *M. tornieri* within this group is weak, the monophyly of the *A. horsfieldii* + *Indotestudo* + *M. tornieri* + *Testudo* + '*T.*' *hermanni* clade is not in question, implying a relatively higher rate of mtDNA evolution in *M. tornieri*. Because significantly disparate evolutionary rates between taxa can adversely affect phylogenetic reconstruction (Felsenstein, 1978; Swofford et al., 1996, Kolaczkowski and Thornton, 2004), we examined differences in evolutionary rates within the *A. horsfieldii* + *Indotestudo* + *M. tornieri* + *Testudo* + '*T.*' *hermanni* clade. We compared differences in substitution rates between members of this clade using relative-rate tests (Sarich and Wilson 1973; Wu and Li 1985). In RRTree 1.1.9 (Robinson-Rechavi and Huchon 2000) we compared rates in a pairwise fashion between each member of the *A. horsfieldii* + *Indotestudo* + *M. tornieri* + *Testudo* + '*T.*' *hermanni* clade using K2P distances (Kimura 1980) and treating sequences as noncoding to include tRNA data. In all comparisons we used *G. pardalis* as the sister group to make rate comparisons to a third group. The differences in substitution rates between *M. tornieri* and each main lineage of the *A. horsfieldii* + *Indotestudo* + *M. tornieri* + *Testudo* + '*T.*'

hermanni clade are significant (Table 1). Specifically, *M. tornieri* displays a more rapid rate of sequence evolution compared to other members of its clade. All other pairwise rate comparisons were not significant, except for rate variation between *A. horsfieldii* and *Testudo*, which does not scale with the rate divergence seen between *M. tornieri* and other taxa (Table 1).

The elevated rate of mitochondrial evolution of *M. tornieri* coincides with other unusual genomic features (the translational frameshift in *nad4* and tandem duplication of *trnF*), suggest either a relaxation of selection on the mt genome, a less accurate replication or repair mechanism, or an environment conducive to mutagenesis. As mentioned previously, *M. tornieri* is also special among tortoises in its bizarre morphology associated with its ecology as a rock crevice specialist. Omland (1997) postulated that molecular and morphological evolutionary rates are often correlated, a phenomenon that appears consistent with our observations of *M. tornieri*. However, the underlying mechanisms responsible for associated rates of molecular and morphological evolution remain unclear, and such correlations are still debated (e.g., Bromham et al., 2002; Marko and Moran, 2002).

When we additional analyses excluding *M. tornieri*, the resulting topology shows increased support for a clade that includes *A. horsfieldii*, '*T.*' *hermanni*, and *Indotestudo* (99 MP bootstrap from 58; 30 decay index from 4; 90 ML bootstrap from 60; BI posterior probabilities stay at 100). These results support the hypothesis that the inclusion of the more rapidly evolving *M. tornieri* sequence is partially responsible for the basal ingroup polytomy.

4. Discussion

4.1. The monophyly, diagnosis, and definition of clades

4.1.1. Phylogenetic taxonomy protocols

Some of the groups recognized by morphological workers (Crumly, 1984; Gmira, 1993, 1995) are strongly supported by our molecular data (e.g., *Testudo* species, *Indotestudo* species, and *T. kleinmanni* + *T. marginata*; Fig. 3). Because we are confident in the monophyly of these clades and feel that other biologists will want to discuss them in a phylogenetic context, we feel justified in proposing phylogenetically-defined names for them. Note that in our study, the origin and conversion of clade names should be attributed to the first author (Parham) and not all seven contributing authors. We adopt the recommendation of PhyloCode (2003) that all clade names should be distinguished from most ranked taxon names governed by the ICZN (1999) through the use of italics. All phylogenetic definitions are given in section 4.1.3.

4.1.2. The monophyly and diagnosis of some clades

Crumly (1984) was the first to hypothesize a monophyletic clade that included members of the *Testudo* group and *Malacochersus*, exclusive of members of the '*Geochelone* complex.' His hypothesis was based on the observation that the *Testudo* group and *Malacochersus* share a unique pattern of head scales and a shortened trachea (although this latter character is homoplastic within *Testudinidae*). The head scale character was first noted by Loveridge and Williams (1957), but was considered to have evolved independently in *M. tornieri*. We are confident in the exclusive monophyly of *Malacochersus* and the greater *Testudo* group within tortoises. Because the *Testudo*

group and *Malacochersus* is a well-supported, diagnosable clade, we phylogenetically define it as *Testudona* (tax. nov.).

Our data support the hypothesis that the genus *Indotestudo* (represented by species *elongata* [type species], *forstenii*, and *travancorica*) is a reciprocally monophyletic group. Although we lack any samples of *I. travancorica*, other studies (Iverson et al., 2001; Spinks et al., 2004) show that the three species of the *Indotestudo* form a closely-related, monophyletic group, so the hypothesized position for *I. elongata* and *I. forstenii* can confidently be extended to *I. travancorica*. In order to stabilize the use of these names we convert the ICZN genus *Indotestudo* into a phylogenetically-defined name.

The three species of *Testudo* form a reciprocally monophyletic clade relative to taxa considered separate genera (e.g., *Agrionemys*, *Indotestudo*). *Testudo graeca* (type species of *Testudo*), *T. marginata*, and *T. kleinmanni* are also diagnosed by a conspicuous morphological trait, a kinetic hinge in the plastron. In order to stabilize the name *Testudo*, we convert the ICZN genus name *Testudo* into a phylogenetically-defined clade name.

Within *Testudo*, our study and Kuyl et al. (2002) found strong support for a clade that includes *T. marginata* and *T. kleinmanni* exclusive of other *Testudo*. Gmira (1993, 1995) found the same result working with morphological data and proposed that both species be placed in a separate subgenus, *Chersus* Wagler, 1830 (type species *T. marginata*). The use of the name *Chersus* was supported by Kuyl et al. (2002) and is a logical choice for an unranked clade name for the node that unites *marginata* and *kleinmanni*. Living *Chersus* can be diagnosed from other living *Testudo* by the presence

of black triangles on the plastron (Fig. 1B) as well as other, more subtle, scalation characters (Gmira 1993, 1995).

4.1.3. Phylogenetic definitions of tortoise clades

We provide phylogenetic definitions for four tortoise clades that are supported by molecular and morphological data. A rank-free hierarchical representation of the taxonomy proposed here is shown in Table 2. All group names are node-based crown groups. For each of these crown names, the prefix 'pan' can be used to refer to non overlapping stem groups (e.g., '*Pantestudo*') following Gauthier and de Queiroz (2001) and Joyce et al. (2004).

Testudona is defined as the crown clade arising from *Malacochersus* [orig. *Testudo*] *tornieri* (Siebenrock, 1903), *Testudo graeca* Linnaeus, 1758, *Agrionemys* [orig. *Testudo*] *horsfieldii* (Gray, 1844), *Indotestudo* [orig. *Testudo*] *elongata* (Blyth, 1853).

Indotestudo Lindholm, 1929 is defined as the crown clade arising from the last common ancestor of *Indotestudo* [orig. *Testudo*] *elongata* (Blyth, 1853), *Indotestudo* [orig. *Testudo*] *forstenii* (Schlegel and Muller, 1844), and *Indotestudo* [orig. *Testudo*] *travancorica* (Boulenger, 1907).

Testudo Linnaeus, 1758 is defined as the crown clade arising from the last common ancestor of *Testudo graeca* Linnaeus 1758, *Testudo marginata* Schoepff, 1792, and *Testudo kleinmanni* Lortet, 1883.

Chersus Wagler 1930 is defined as the crown clade arising from the last recent common ancestor of *Testudo marginata* Schoepff, 1792 and *Testudo kleinmanni* Lortet, 1883.

4.1.4. The taxonomy of ‘Testudo’ hermanni

Although ‘*Testudo*’ *hermanni* is clearly not a member of the *Testudo* clade as defined above, we do not propose a new ICZN binomial for this species. Some authors have suggested that ‘*T.*’ *hermanni* should be placed in the genus *Agrionemys*. A close relationship with the type species of *Agrionemys*, *A. horsfieldii*, is supported by a morphological analysis (Gmira, 1993, 1995). However, Lapparent de Broin (2000a, 2001) claims that this result is complicated by the addition of new fossil data. From a molecular perspective, although the ML and BI analyses recovers a *hermanni* + *horsfieldii* clade, this relationship is only represented in one of the two shortest parsimony trees. Given the lack of strong evidence for the phylogenetic position of ‘*T.*’ *hermanni*, we do not recommend its inclusion in *Agrionemys* and refrain from converting this ICZN genus into a phylogenetically-defined clade. ‘*Testudo*’ *hermanni* has never been used as a type species for an ICZN genus and so no older names are readily available. Meanwhile, Lapparent de Broin (pers. comm. to JFP) is planning to propose a new ICZN genus for ‘*T.*’ *hermanni*.

4.2. Paleobiogeography of Testudona

4.2.1. The paleobiogeography of the Chersus *clade*

The two species of the *Chersus* clade, *T. marginata* and *T. kleinmanni*, occur on either side of the Mediterranean (Fig. 1B). The natural distribution of *T. marginata* is restricted to mainland and archipelago Greece and extreme southern Albania, although it has been introduced by humans into other Mediterranean islands such as Sardinia.

Populations of *T. kleinmanni* are scattered along the coasts of northern Africa from Tripolotania to the Sinai. This allopatric distribution of extant *Chersus* lineages can be explained by several competing hypotheses including: (1) the extinction of intermediate populations; (2) vicariance by the division of an ancestral *Chersus* clade by the filling of the Mediterranean basin by seawater from the Atlantic ocean at ~5.33 million years ago (Krijgsman et al., 1999); or (3) dispersal from one side of the Mediterranean to the other.

The "extinction hypothesis" is not supported by any data as there are no known fossils of *Chersus* outside its present range. The importance of "oceanic" dispersal to explain animal distributions has been underestimated (de Queiroz, 2005). Tortoises in particular are known to have traversed greater distances over salt water than the dispersal event proposed here (e.g., the Galapagos tortoises and others; see Meylan and Sterrer, 2000 for a review). Although the European *Testudo* fossil record is ~10 million years old (Lapparent de Broin, 2001; Danilov, in press), there are no fossils that confirm that stem *Chersus* turtles occurred on southern side of the Mediterranean basin prior to 5.33 million years ago. The oldest fossil of confidently identified *testudonans* from North Africa are less than three million years old (Lapparent de Broin, 2000b: 54-55). The 'vicariance hypothesis' would also be supported by the discovery of *Chersus* fossils in North Africa that predate the formation of the modern Mediterranean Sea (e.g., >5.33 million years old).

4.2.2. The age and paleobiogeography of the testudonan *diversification*

In the absence of a *testudonan*-specific clock (see section 4.2.3 below), we used previously calibrated rates from other vertebrates (not based on fossils, see Weisrock et

al., 2001) to the divergences of the basal *testudonan* polytomy yields age estimates ranging from 6.0-9.4 million years old (myo) (pairwise distances = 8.24-10.73% for the same region calibrated by Weisrock et al. [2001], not including *M. tornieri*). These age estimates are in the late Miocene. Although fossil tortoises over 50 million years old are known from Europe and Asia, the fossil record of *testudonan*-like turtles begins ~10 myo (reviewed in Lapparent de Broin, 2000b, 2001; Danilov, in press) also in the late Miocene. Therefore, both the fossil and tentative molecular age estimates suggest that *Testudona* probably originated as recently as the late Miocene (~5-15 myo).

The Miocene origin of *Testudona* is considerably younger than the oldest tortoise from Africa, *Gigantochersina ammon* (Andrews, 1903), confidently dated at 35.4-35.6 myo (late Eocene; Holroyd and Parham, 2003). Although its phylogenetic position is uncertain, *G. ammon* might represent part of the lineage that gave rise to the "*Geochelone* complex," a radiation that was already diverse in Africa by ~20 myo (the early Miocene; Lapparent de Broin, 2000b, 2003) and is still represented by over 25 extant species. Given the antiquity of African tortoise fossils, testudonans must have invaded in Africa after one or more tortoise lineages were already established on the continent.

Today, *testudonan* tortoises have a wide geographic distribution that includes the Palearctic (*Testudo, A. horsfieldii*, '*T.' hermanni*), Oriental (*Indotestudo*), and Ethiopian (*M. tornieri*) biogeographic realms (Fig. 1A). The dispersal of the *Indotestudo* and *M. tornieri* lineages from a Palearctic center of diversification was made possible by the ongoing collision of Arabia into Anatolia and Iran that established land bridges that connected Europe, Africa, and southern Asia in the late Miocene (Rögl, 1999). Although the fossil record of Palearctic tortoises is very rich (Auffenberg, 1974; Crumly, 1983; Ye,

1994; Lapparent de Broin, 2000a, b, 2001; Danilov, in press), these specimens have never been analyzed in an explicit phylogenetic framework or in light of the major tectonic and environmental changes occurring throughout the Miocene (Agusti et al., 1999). A more detailed study of Miocene *testudonan* fossils integrated into the molecular phylogenetic context presented here will yield additional clues to the timing and patterns of intercontinental dispersal events as well as the initial division of lineages.

4.2.3. Problems with the application of fossil data to molecular clocks

We report a single molecular age estimate in section 4.2.2 (above), and only in a very conservative fashion (i.e., based on the entire range from across disparate vertebrate taxa). By doing so we only hope to achieve a rough approximation of the antiquity of the ingroup (e.g., within 5 or 10 million years), but concede that there are major obstacles preventing us from exploring a more precise *testudonan* molecular clock. Namely, there are the caveats raised by Graur and Martin (2004). First, a detailed study of *testudonan* molecular clocks would require a more detailed understanding of the dating of *testudonan*-bearing strata than is presently available. This is primarily because uncertainties associated with the dating of fossils need to be included in all calibrated age estimates. This simple point is often ignored my molecular clock studies that use fossil calibrations (e.g., Near et al., 2005). Secondly, as far as we know, all molecular clock studies that rely on fossils calibrations have not reported if their fossils specimens can be placed in a phylogenetic context with any degree of statistical significance. Consequently, the appearance of statistical rigor presumed by these studies is misleading. In our case, the *testudonan* fossils mentioned here have never been studied in a

phylogenetic context. So while we can be reasonably sure the fossils mentioned here are at least *pantestudonans*, we can't confidently attribute them to, or exclude them from, extant lineages. In any case, we don't have the multiple calibration points necessary to generate reliable estimates.

Appendix 1: Voucher and GenBank information for the sequences used in our study. Permit information (including CITEs) for imported specimens is on file at the Museum of Vertebrate Zoology and California Academy of Sciences. 1) *Chrysemys picta* **GenBank NC002073**; 2) *Mauremys reevesii* **GenBank NC006082**; 3) *Manouria emys* MVZ 238129, **DQ080040**. Pet trade, no locality data; 4) *Geochelone pardalis* MVZ 241333, **DQ080041**. Awdal Region, Somaliland; 5) '*Testudo*' *hermanni* MVZ 244866, **DQ080046**. Thrace, Turkey; 6) *Agrionemys horsfieldi* CAS 184468, **DQ080045**. Kopet-Dagh, Turkmenistan; 7) *Malacochersus tornieri* MVZ 234632, **DQ080042**. Pet trade, no locality data; 8) *Indotestudo elongata* MVZ 234627, **DQ080043**. Pet trade, no locality data; 9) *Indotestudo forstenii* MVZ 234627, **DQ080044**. Pet trade, no locality data; 10) *Testudo marginata* MVZ 247484, **DQ080047**. Athens, Greece; 11) *Testudo kleinmanni* MVZ 230361, **DQ080048**. Pet trade, no locality data; 12) *Testudo graeca* MVZ 235707, **DQ080049**. Nabul Governerate, Tunisia; 13) *Testudo graeca* CAS 218245, **DQ080050**. Gaziantep, Turkey.

Appendix 2: Best fit models of evolution for each mtDNA locus selected by AIC in MrModeltest. For the BI analysis we combined loci into the same data partition if they belonged the same functional group (rRNA, tRNA, or coding DNA) and conformed to the same model of evolution, resulting in 16 data partitions. For the ML analysis we used the model of evolution chosen by AIC in Modeltest for the entire mitochondrial genome (GTR + I + Γ). Partition 1: *rrnL*, *rrnS*= GTR + I + Γ; Partition 2: *cox1*, *cox3*, *nad1*, *nad3*, *nad4*, *nad4L*, *nad5*, *nad6*= GTR + I + Γ; Partition 3: *nad2*, *atp6*= HKY + I + Γ; Partition 4: *cob*= GTR + Γ; Partition 5: *cox2*= GTR + I; Partition 6: *cox2*= HKY + Γ; Partition 7:

trnF= GTR + I + Γ; Partition 8: *trnR*= GTR + Γ; Partition 9: *trnQ*= GTR + I; Partition 10: *trnW*= SYM + I + Γ; Partition 11: *trnK*= SYM + Γ; Partition 12: *trnN*= SYM + I; Partition 13: *trnA*, *trnC*, *trnD*, *trnG*, *trnL2*, *trnE*, *trnP*= HKY + Γ; Partition 14: *trnV*, *trnM*, *trnY*, *trnT*= HKY + I; Partition 15: *trnH*, *trnS1*= HKY; Partition 16: *trnI*, *trnL1*, *trnS2*= K80 + Γ.

Acknowledgments

Thanks to Jennifer Kuehl, Dave Engle**,** W. Brian Simison, and Jonathan Fong of the Evolutionary Genomics Department of the Joint Genome Institute for their advice and assistance to JFP in the laboratory, to Jeff Froula from the same group for help gathering data on gene rearrangements, to Meredith Mahoney for helpful advice to JFP regarding analyses, to Carla Cicero, David Wake, and Rochelle Germano of the Museum of Vertebrate Zoology and Jens Vindum of the California Academy of Sciences for their assistance with the accession of specimens into museum and the loaning of tissues, to Suleiman Ahmed Gulaid, President of Amoud University, Borama, Somaliland, for coordinating field work in Somaliland for TJP, to Sahat Shammakov of the Ministry of Nature Protection of Turkmenistan for hosting JRM and TJP, to Michael Thomas and Luobin Yang of Idaho State University's Evolutionary Genomics Group for computational assistance, to James Buskirk and U. Fritz for insightful conversations about *Testudo* and related tortoises, to the Scientific and Technical Research Council of Turkey (TUBITAK TBAG-2206 [102T104]), to Randall Irmis of the University of California Museum of Paleontology for many useful discussions about the application of the fossil record to molecular clock studies, and to Sarah Rieboldt for continuous support throughout this project. Most of the images for Figure 1 were generously donated by others; thanks to James Buskirk for helping arrange much of this. The photo credits are Paula Elis. Morris (*M. tornieri*), Remy Amman and Jerome Maran (*Indotestudo*, '*T.' hermanni*, both *T. graeca*, and *T. marginata*), JFP (*Agrionemys*), Sherif Baha El Din (*T. kleinmanni*), Kinji Hayashi (*T. marginata* plastron view), and James Harding (*T. kleinmanni* plastron view). This work is LBNL-57512 and was performed under the

auspices of the U.S. Department of Energy, Office of Biological and Environmental Research, by the University of California, Lawrence Berkeley National Laboratory, under contract No. DE-AC03-76SF00098. This is UCMP Contribution # 1879.

References

Adams, A., Reeve, L., 1850. Mollusca. In: Adams, A., The Zoology of the Voyage of H.M.S. Samarang under the command of Captain Sir Edward Belcher during the years 1843-1846. Part 3. Reeve, Benham and Reeve, London.

Agusti, J., Rook, L., Andrews, P., (Eds.), 1999. The Evolution of Neogene Terrestrial Ecosystems in Europe. Cambridge Univ. Press, Cambridge.

Akaike, H., 1974. A new look at the statistical model identification. IEEE Trans. Auto. Cont. 19, 716-723.

Andrews, C.W., 1903. *Testudo ammon* sp. nov. In: Andrews, C.W., Beadnell, H.J.L., 1903. A Preliminary Notice of a Land Tortoise from the Upper Eocene of the Fayûm, Egypt (With a Note on the Method Employed in Excavation, Strengthening, and Transport). National Printing Dept., Cairo.

Auffenberg, W., 1974. Checklist of fossil terrestrial tortoises (Testudinidae). Bull. Florida St. Mus. Biol. Sci. 18, 121–251.

Avise, J.C., Bowen, B.W., Lamb, T., Meylan, A.B., 1992. Mitochondrial DNA evolution at a turtle's pace: Evidence for low genetic variability and reduced microevolutionary rate in the Testudines. Mol. Biol. Evol. 9, 457-473.

Beagley, C.T., Okimoto, R., Wolstenholme, D.R., 1999. Mytilus mitochondrial DNA contains a functional gene for a tRNA-ser(UCN) with a dihydrouridine arm-replacement loop and a pseudo=tRNA-ser(UCN) gene. Genetics 152, 641-652.

Beard, K.C., 1998. East of Eden: Asia as an important center of taxonomic origination in mammalian evolution. Bulletin of Carnegie Museum of Natural History 34, 5-39.

Bell, T., 1828. Description of three new species of land tortoise. Zool. J. London. 3:419-421.

Blyth, E., 1853. Notices and descriptions of various reptiles, new or little known. J. Asiat. Soc. Bengal Calcutta 22, 639-655.

Boore, J.L., Medina, M., Rosenberg, L.A., 2004. Complete sequences of two highly rearranged molluscan mitochondrial genomes, those of the scaphopod *Graptacme eborea* and of the bivalve *Mytilus edulis*. Mol. Biol. Evol. 21(8), 1492–1503.

Boulenger, G.A., 1907. A new tortoise from Travancore. J. Bombay. Nat. Hist. Soc. 17, 560-564.

Bour, R., 1980. Essai sur la taxinomie des Testudindae actuels (Reptilia, Chelonii). Bull. Mus. Natl. Hist. Natur. Paris (4)A, 541-546.

Bour, R., 2004. Répartition géographique des différents taxons actuellement reconnus dans le genre *Testudo* (s.l.). Manouria 22, back cover.

Bremer, K., 1994. Branch support and tree stability. Cladistics 10, 295-304.

Bromham, L., Woolfit, M., Lee, M.S.Y., Rambaut, A., 2002. testing the relationship between morphological and molecular rates of change along phylogenies. Evol. 56, 1921-1930.

Brongniart, A., 1800. Essai d'une classification naturelle des reptiles, par le citoyen. Premiere Partie. Etablissement des ordres. Bull. des Sci. Soc. Philomat. Paris 2, 81–82.

Campbell, N.J., Barker, S.C., 1999. The novel mitochondrial gene arrangement of the cattle tick, *Boophilus microplus*: Fivefold tandem repetition of a coding region. Mol. Biol. Evol. 16(6), 732-740.

Crumly, C.R., 1983. An annotated checklist of the fossil tortoises of China and Mongolia. Proc. Biol. Soc. Wash. 96, 567-580.

Crumly, C.R., 1984. A hypothesis for the relationships of land tortoises genera (Family Testudinidae). Studia Geológica Salmanticensia vol. especial 1, Studia Palaeocheloniologica I, 115-124.

Danilov, I.G., In press. Die fossilen Schildkroten Europas. In: Fritz, U. (Ed.), Handbuch der Reptilien und Amphibien Europas. Schildkroten II (Cheloniidae, Dermochelyidae, Fossile Schildkroten). Aula-Verlag, Wiebelsheim.

de Queiroz, A., 2005. The resurrection of oceanic dispersal in historical biogeography. Trends Ecol. Evol. 20, 68-73.

Dean, F.B., Nelson, J.R., Giesler, T.L., Lasken, R.S., 2001. Rapid amplification of plasmid and phage DNA using Phi 29 DNA polymerase and multiply-primed rolling circle. Genome Res. 11, 1095–1099.

Dowton, M., Castro, L.R., Campbell, S.L., Bargon, S.D., Austin, A.D., 2003. Frequent mitochondrial gene rearrangements at the Hymenopteran nad3-nad5 junction J. Mol. Evol. 56(5), 517-526.

Eberhard, J.R., Wright, T.F., Bermingham, E., 2001. Duplication and concerted evolution of the mitochondrial control region in the parrot genus Amazona. Mol. Biol. Evol. 18(7), 1330-1342.

Elkin, C., Kapur, H., Smith, T., Humphries, D., Pollard, M., Hammon, N., Hawkins, T., 2002. Magnetic bead purification of labeled DNA fragments for high-throughput capillary electrophoresis sequencing. Biotech. 32, 1296–1302.

Ernst, C.H., Barbour, R.W., 1989. Turtles of the World. Smithsonian Institution Press,

Washington D.C.

Farabaugh, P.J., Vimaladithan, A., 1998. Effect of frameshift-inducing mutants of elongation factor 1alpha on programmed +1 frameshifting in yeast. RNA 4(1), 38-46.

Farris, J.S., 1983. The logical basis of phylogenetic analysis. In: Platnick, N., Funk, V.A. (Ed.), The logical basis of phylogenetic analysis. Columbia University Press, New York, pp. 7-36.

Felsenstein, J., 1978. Cases in which parsimony and compatibility methods will be positively misleading. Syst. Zool. 27, 401-410.

Felsenstein, J., 1981. Evolutionary trees from DNA sequences: A maximum likelihood approach. J. Mol. Evol. 17, 368–376.

Felsenstein, J., 1985. Confidence limits on phylogenies: an approach using the bootstrap. Evol. 39, 783–791.

Fritz, U., Kirok, P., Kami, H., Wink, M., In press. Environmentally caused dwarfism or a valid species—Is *Testudo weissingeri* Bour, 1996 a distinct evolutionary lineage? New evidence from mitochondrial and nuclear genomic markers. Mol. Phyl. Evol.

Gauthier, J., de Queiroz, K., 2001. Feathered dinosaurs, flying dinosaurs, crown dinosaurs, and the name ‘‘Aves.’’ In: Gauthier, J., Gall, L.F. (Eds.), New Perspectives on the Origin and Early Evolution of Birds: Proceedings of the International Symposium in Honor of John H. Ostrom. Peabody Museum of Natural History, New Haven, pp 7-41.

Gissi, C., Iannelli, F, Pesole, G., 2004. Complete mtDNA of *Ciona intestinalis* reveals extensive gene rearrangement and the presence of an *atp8* and an extra *trnM* gene

in ascidians. J. Mol. Evol. 58, 376-389.

Gmelin, J.F., 1789. *Testudo hermanni* sp. nov. In: Linnaeus, C., Systema Naturae, Volume 1 (thirteenth edition). N. p., Lipsia, pp. 1309.

Gmira, S., 1993. Une nouvelle espéce de tortue Testudininei (*Testudo kenitrensis* n. sp.)de l'inter Amirien-Tensiftien de Kénitra (Maroc). C. R. Acad. Sci. Paris. Sér. II, 316, 701-707.

Gmira, S., 1995. Étude des Chéloniens Fossiles du Maroc. Cahier de Paléontologie, Paris.

Graur, D. Martin, W. 2004. Reading the entrails of chickens: molecular time-scales of evolution and the illusion of precision. Trends Genet. 20, 80–86.

Gravenhorst, J.L.C., 1829. Deliciae Musei Zoologici Vratislaviensis Fasciculus Primus continens Chelonios et Batrachia. Leopold Voss, Leipzig.

Gray, J.E., 1825. A synopsis of the genera of reptiles and amphibia, with a description of some new species. Ann. Philos., 10, 193-217.

Gray, J.E., 1831. Synopsis Reptilium, Pt. 1, Cataphracta. Tortoises, Crocodiles, Enaliosaurians. Treuttel, Wurtz, London

Gray, J.E., 1844. Catalogue of the Tortoises, Crocodiles, and Amphisbaenians in the Collections of the British Museum. British Museum of Natural History, London.

Gray, J.E., 1852. Description of a new genus and some new species of tortoises. Proc. Zool. London. 1852, 133-135.

Hansen, T.M., Baranov, P.V., Ivanov, I.P., Gesteland, R.F., Atkins, J.F., 2003. Maintenance of the correct open reading frame by the ribosome. EMBO Rep. 4(5), 499-504.

Harris, D.J., Znari, M., Mace, J.-C., Carretero, M.A., 2003. Genetic variation in *Testudo*

graeca from Morocco estimated using 12S rRNA sequencing. Rev. Esp. Herp. 17, 5-9.

Hawkins, T.L., Detter, J.C., Richardson, P.M., 2002. Whole genome amplification-applications and advances. Curr. Opin. Biotechnol. 13, 65–67.

Holroyd, P.A., Parham, J.F. 2003. The antiquity of African tortoises. J. Vert. Paleo. 23, 688–690.

Hrbek, T., Larson, A. 1999. The evolution of diapause in the killfish family Rivulidae (Atherinomorpha, Cyprinodontiformes): A molecular phylogenetic and biogeographic perspective. Evolution 53, 1200-1216.

Huelsenbeck, J.P., and Ronquist, F., 2001. MrBayes: Bayesian inference of phylogenetic trees. Bioinf. 17, 754-755.

International Code of Zoological Nomenclature., 1999 (fourth edition). International Trust for Zoological Nomenclature, London.

Iverson, J.B., Spinks, P.Q., Shaffer, H.B., McCord, W.P., Das, I. 2001. Phylogenetic relationships among the Asian tortoises of the genus *Indotestudo* (Reptilia: Testudines: Testudinidae). Hamadryad 26, 272-275.

Jackman, T.R., Larson, A., de Queiroz, K., Losos, J.B. 1999. Phylogenetic relationships and tempo of early diversification in *Anolis* lizards. Syst. Biol. 48, 254–285.

Janke, A., Arnason, U. 1997. The complete mitochondrial genome of *Alligator mississippiensis* and the separation between recent Archosauria (birds and crocodiles). Mol. Biol. Evol. 14(12), 1266-1272.

Joyce, W.G., Parham, J.F., Gauthier, J.A., 2004. Developing a protocol for the conversion of rank-based taxon names to phylogenetically defined clade names,

as exemplified by turtles. J. Paleontol. 78, 989-1013.

Khozatsky, L.K., and Mlynasrski, M., 1966. *Agrionemys* – nouveau genre de tortue terrestres (Testudinidae). Bull. Acad. Polonaise Sci. 14, 123-125.

Kimura, M., 1980. A simple method for estimating evolutionary rate of base pair substitution through comparative studies of nucleotide sequence. J. Mol. Evol. 16, 111-120.

Kolaczkowski, B., Thornton, J.W., 2004. Performance of maximum parsimony and likelihood phylogenetics when evolution is heterogeneous. Nature 431, 980-984.

Krijgsman, W., Hilgen, F.J., Raffi, I., Sierro, F.J., Wilson, D.S., 1999. Chronology, causes and progression of the Messinian salinity crisis. Nature 400, 652-655.

Kumazawa, Y., Endo, H., 2004. Mitochondrial genome of the Komodo Dragon: Efficient sequencing method with reptile-oriented primers and novel gene rearrangements. DNA Res. 11(2), 115-125.

Kumazawa, Y., Nishida, M., 1993. Sequence evolution of mitochondrial tRNA genes and deep-branch animal phylogenetics. J. Mol. Evol. 37, 380-398.

Kumazawa, Y., Ota, H., Nishida, M., Ozawa, T., 1998. The complete nucleotide sequence of a snake (*Dinodon semicarinatus*) mitochondrial genome with two identical control regions. Genetics 150(1), 313-329.

Kuyl, A.C. van der, Ballasina, D.L.P., Dekker. J.T., Maas, J., Willemsen, R.E., Goudsmit, J., 2002. Phylogenetic relationships among the species of the genus *Testudo* (Testudines: Testudinidae) inferred from mitochondrial 12S rRNA gene sequences. Mol. Phyl. Evol. 22, 174-183.

Kuyl, A.C. Van der, Ballasina, D.L.P., Zorgdrager, F., 2005. Mitochondrial haplotype

diversity in the tortoise species *Testudo graeca* from North Africa and the Middle East. BMC Evol. Biol. 5, 1-8.

Lapparent de Broin, F. de, 2000a. Les Chéloniens de Sansan. Mém. Mus. Nat. d'Hist. Natur. 183, 219-261.

Lapparent de Broin, F. de, 2000b. African chelonians from the Jurassic to the Present. A preliminary catalog of the African fossil chelonians. Palaeont. Afric. 36, 43-82.

Lapparent de Broin, F. de, 2001. The European turtle fauna from the Triassic to the Present. Dumerilia 4, 155-216.

Lapparent de Broin, F. de, 2003. The Miocene chelonians from the southern Namibia. Mem. Geol. Surv. Namib. 19, 67-102.

Larget, B., and Simon, D.L., 1999. Markov chain monte carlo algorithms for the bayesian analysis of phylogenetic trees. Mol. Biol. Evol. 16(6), 750–759.

Lavrov, D., Brown, W.M., Boore, J.L., 2004. Phylogenetic position of the Pentastomida and (pan)crustacean relationships. Proceedings of the Royal Society of London B 271(1538), 537-544.

Le, T.H., Blair, D., Agatsuma, T., Humair, P.F., Campbell, N.J., Iwagami, M., Littlewood, D.T., Peacock, B., Johnston, D.A., Bartley, J., Rollinson, D., Herniou, E.A., Zarlenga, D.S., McManus, D.P., 2000. Phylogenies inferred from mitochondrial gene orders-a cautionary tale from the parasitic flatworms. Mol. Biol. Evol. 17(7), 1123-1125.

Lee, M.S.Y. 2000. Tree robustness and clade significance. Syst. Biol.49, 829-836.

Lindholm, W.A., 1929. Revidiertes Verzeichnis der Gattungen der rezentenSchildkröten nebst Notizen zur Nomenklatur einiger Arten. Zoolog. Anz. 81, 275–272.

Linnaeus, C., 1758. Systema Naturae, Volume 1 (tenth edition). Laurentius Salvius, Holmia.

Lortet, L., 1883. Etudes zoologiques sur la faune de Lac de Tiberiade. Arch. Mus. Hist. Natur., Lyon 3, 99-189.

Loveridge, A., Williams, E.E., 1957. Revision of African tortoises and turtles of the suborder Cryptodira. Bull. Mus. Comp. Zool. Harvard 115, 163-557.

Lynch, M., Force, A., 2000. The probability of duplicate gene preservation by subfunctionalization. Genetics 154(1), 459-473.

Macey, J.R., Larson, A., Ananjeva, N.B., Fang, Z., Papenfuss, T.J., 1997. Two novel gene orders and the role of light-strand replication in rearrangement of the vertebrate mitochondrial genome. Mol. Biol. Evol. 14, 91–104.

Macey, J.R., Schulte, J.A. II, Larson, A., 2000. Evolution and phylogenetic information content of mitochondrial genomic structural features illustrated with acrodont lizards. Syst. Biol. 49(2), 257-277.

Macey, J.R., Schulte, J.A. II, Larson, A., Tuniyev, B. S., Orlov, N., Papenfuss, T. J., 1999. Molecular phylogenetics, tRNA evolution, and historical biogeography in anguid lizards and related taxonomic families. Mol. Phyl. Evol. 12, 250272.

Macey, J.R., Verma, A., 1997. Homology in phylogenetic analysis: alignment of transfer RNA genes and the phylogenetic position of snakes. Mol. Phyl. Evol. 7, 272–279.

Maddison, W.P., 1989. Reconstructing character evolution on polytomous cladograms. Cladistics 5, 365-377.

Maddison, D.R., Maddison, W.P. 2001. MacClade 4.03. Sinauer Associates, Inc., Sunderland, MA.

Marko, P.B., Moran, A.L., 2002. Correlated evolutionary divergence of egg size and mitochondrial protein across the Isthmus of Panama. Evol. 56, 1303-1309.

McGuire, J.A., Heang, K.B., 2001. Phylogenetic systematics of Southeast Asian flying lizards (Iguania: Agamidae: *Draco*) as inferred from mitochondrial DNA sequence data. Biol. J. Linn. Soc. 72, 203-229.

Meylan, P.A., Sterrer, W., 2000. *Hesperotestudo* (Testudines: Testudinidae) from the Pleistocene of Bermuda, with comments on the phylogenetic position of the genus. Zool. J. Linn. Soc. 128, 51-76.

Mindell, D.P., Sorensen, M.D., Dimcheff, D.E. 1998. An extra nucleotide is not translated in mitochondrial ND3 of some birds and turtles. Mol. Biol. Evol. 15, 1568-1571.

Mueller, R.L., Macey, J.R., Jaekel, M., Wake, D.B., Boore, J.L., 2004. Morphological homoplasy, life history evolution, and historical biogeography of plethodontid salamanders: Novel insights from complete mitochondrial genome sequences. Proc. Natl. Acad. Sci. USA 101(38), 13820-13825.

Near, T.J., Meylan, P.A., Shaffer, H.B. 2005. Assessing concordance of fossil calibration points in molecular clock studies: an example using turtles. Am. Nat. 165, 137-146.

Nylander, J.A.A, 2004. MrModeltest 2.1. Program distributed by the author. Evolutionary Biology Centre, Uppsala.

Omland, K.E., 1997. Correlated rates of molecular and morphological evolution. Evol. 51, 1381-1393.

Pande, S., Vimaladithan, A., Zhao, H., Farabaugh, P.J., 1995. Pulling the ribosome out of frame by +1 at a programmed frameshift site by cognate binding of aminoacyl-tRNA. Mol. Cell Biol. 15(1), 298-304.

Passamonti, M., Boore, J.L., Scali,V. 2003. Molecular evolution and recombination in gender-associated mitochondrial DNAs of the Manila Clam *Tapes philippinarum*. Genetics 164(2), 603-611.

Perälä, J., 2001. A new species of *Testudo* (*Testudines*: Testudinidae) from the Middle East, with implications for conservation. J. Herp. 35, 567–582.

Perälä, J., Bour, R., 2004. Neotype of *Testudo terrestris* Forsskäl, 1775 (Testudines, Testudindae). Asiat. Herp. Res. 10, 114-119.

PhyloCode, 2003. A Phylogenetic Code of Biological Nomenclature, Version 2a. http://www.ohio.edu/phylocode/preface.html.

Posada, D., and Buckley, T.R., 2004. Model selection and model averaging in phylogenetics: advantages of Akaike Information Criterion and Bayesian approaches over likelihood ratio tests. Syst. Biol. 53, 793-808.

Posada, D., Crandall, K.A., 1998. Modeltest: testing the model of DNA substitution. Bioinf. 14, 817-818.

Rafinesque, C.S., 1832. Description of two new genera of soft shell turtles of North America. Atl. J. Friend Knowl. Phil. 1, 64-65.

Rannala, B., and Yang, Z.H., 1996. Probability distribution of molecular evolutionary trees: a new method of phylogenetic inference. J. Mol. Evol. 43, 304-311.

Robinson-Rechavi, M. Huchon, D., 2000. RRTree: Relative rates tests between groups of sequences on a phylogenetic tree. Bioinf. 16, 296-297.

Rögl, F. ,1999. Mediterranean and Paratethys Palaeogeography during the Oligocene and Miocene. In: Agusti, J., Rook, L., Andrews, P. (Eds.), The Evolution of Neogene Terrestrial Ecosystems in Europe. Cambridge Univ. Press, Cambridge, pp. 8-22.

Sarich, V. M., Wilson, A.C., 1973. Generation time and genomic evolution in primates. Science 179, 1144-1147.

Schlegel, H., Müller, S., 1844. Over de Schildpadden van den Indischen Archipel. Beschrijving einer nieuwe soort van Sumatra. In: Temminck, C.J. (Ed.), Verhandelingen over de natuurlijke geschiedenis der Nederlandsche overzeesche bezittingen, Pt. 3. S. and J. Luchtmans and van der Hoek, Leiden., pp. 29–36.

Schulte, J.A. II, Macey, J.R., Larson, A., Papenfuss, T.J., 1998. Molecular test of phylogenetic taxonomies: a general procedure and example using four subfamilies of the lizard family Iguanidae. Mol. Phyl. Evol. 10, 367-376.

Schoepff, J.D., 1792. Historia Testudinum iconibus Illustrata. Palmii, Erlangen.

Shimodaira, H., Hasegawa, M., 1999. Multiple comparisons of log-likelihoods with applications to phylogenetic inference. Mol. Biol. Evol. 16, 1114-1116.

Siebenrock, F., 1903. Über zwei seltene und eine neue Schildkröte des Berliner Museums. Sitz. Acad. Wiss. Wien Mathematur. Kl. 112, 439-445.

Spinks, P.Q., Shaffer, H.B., McCord, W.P., Iverson, J.B., 2004. Phylogenetic hypotheses for the turtle family Geoemydidae. Mol. Phyl. Evol. 32, 164–182.

Swofford, D.L., 2002. PAUP*. Phylogenetic Analysis Using Parsimony (*and Other Methods). Version 4. Sinauer Associates, Sunderland.

Swofford, D. L., Olson, G.J., Waddell, P.J., Hillis, D.M., 1996. Phylogenetic inference. In: Hillis, D.M., Moriyz, C., Mable, B.K. (Eds.), Molecular systematics. Sinauer

Associates, Sunderland, pp. 407-514.

Takahashi, A., Otsuka, H., Hirayama, R., 2003. A new species of *Manouria* (Testudines: Testudinidae) from the Upper Pleistocene of the Ryukyu Islands, Japan. Pal. Res. 7, 195-217.

Templeton, A.R., 1983. Phylogenetic inference from restriction site endonuclease cleavage site maps with particular reference to the humans and apes. Evol. 37, 221-244.

Townsend, T., Larson, A., 2002. Molecular phylogenetics and mitochondrial genomic evolution in the Chamaeleonidae (Reptilia, Squamata). Mol. Phylogenet. Evol. 23(1), 22-36.

Wagler, J., 1830. Natürliches System der Amphibien mit vorangehender Classification der Säugethiere und Vögel. J.G. Cotta'sche Buchhandlung, München.

Weisrock, D.W., Macey, J.R., Ugurtas, I.H., Larson, A., Papenfuss, T.J., 2001. Molecular phylogenetics and historical biogeography among salamandrids of the "true" salamander clade: Rapid branching of numerous highly divergent lineages in *Mertensiella luschani* associated with the rise of Anatolia. Mol. Phyl. Evol. 18, 434-448.

Wilgenbusch, J., de Queiroz, K., 2000. Phylogenetic relationships among the phrynosomatid sand lizards inferred from mitochondrial DNA sequences. Syst. Biol. 49, 592–612.

Wu, C. I., Li, W.-H., 1985. Evidence for higher rates of nucleotide substitution in rodents than in man. Proc. Natl. Acad. Sci. 82, 1741-1745.

Ye X., 1994. Fossil and Recent Turtles of China. Science Press, Beijing.

Yokobori, S., Fukuda, N, Nakamura, M. Aoyama, R., Oshima, T., 2004. Long-term conservation of six duplicated structural genes in cephalopod mitochondrial genomes. Mol. Biol. Evol. 21(11), 2034-2046.

Yokobori, S., Ueda, T., Feldmaier-Fuchs, G., Paabo, S., Ueshima, R., Kondow, A., Nishikawa, K., Watanabe,K. 1999. Complete DNA sequence of the mitochondrial genome of the ascidian *Halocynthia roretzi* (Chordata, Urochordata). Genetics 153(4), 1851-1862.

Yokobori, S., Watanabe, Y., Oshima, T., 2003. Mitochondrial genome of *Ciona savignyi* (Urochordata, Ascidiacea, Enterogona): Comparison of gene arrangement and tRNA genes with *Halocynthia roretzi* mitochondrial genome. J. Mol. Evol. 57(5), 574-587.

Figure Legends

Figure 1: Map showing the approximate distribution of the ingroup (*Testudo* and related species) modified from Bour (2004). (A) Non-*Testudo* members of the ingroup. (B) Species of *Testudo*; Arrows indicate disjunct populations of *T. graeca* (black), *T. marginata* (white, Sardinia), and *T. kleinammni* (white, N. Africa); *T. marginata* and *T. kleinmanni* include ventral views to show the characteristic black triangles mentioned in the text. See acknowledgments for image credits.

Figure 2: (A) Typical vertebrate gene arrangement from *cob* to *trnV*. (B) Same region for shown for *Malacochersus tornieri* showing the duplicated *trnF* and control region. Only the first control region appears functional. (C) Duplicated *trnF* sequences for *Malacochersus tornieri* folded into typical cloverleaf secondary structures. All sites that differ between the two are underlined and in boldface. The naming convention for five portions of the secondary structure are indicated.

Figure 3: Phylogram of one of the two most parsimonious trees recovered by our parsimony analysis of 14,858 nucleotide positions. Parsimony bootstrap/ML bootstrap/BI posterior probabilities are indicated above the stems and decay indices/p values from the Wilcoxon rank-sum test are indicated below the stems. Branch lengths are based on a delayed transformation character-state optimization.