Molecular Phylogenetics and Evolution 65 (2012) 241–250

Contents lists available at SciVerse ScienceDirect

Molecular Phylogenetics and Evolution

journal homepage: www.elsevier.com/locate/ympev

Marine turtle mitogenome phylogenetics and evolution

Sebastián Duchene [a], Amy Frey [a], Alonzo Alfaro-Núñez [b], Peter H. Dutton [a], M. Thomas P. Gilbert [b], Phillip A. Morin [a,*]

[a] Protected Resources Division, Southwest Fisheries Science Center, National Marine Fisheries Service, NOAA, 9801 La Jolla Shores Dr., La Jolla, CA 92037, USA
[b] Centre for GeoGenetics, Natural History Museum of Denmark, University of Copenhagen, Øster Voldgade 5-7, 1350 Copenhagen, Denmark

ARTICLE INFO

Article history:
Received 7 March 2012
Revised 11 June 2012
Accepted 13 June 2012
Available online 26 June 2012

Keywords:
Sea turtle
Molecular clock
Mitogenome
Molecular adaptive evolution
Mitochondrial phylogenetics

ABSTRACT

The sea turtles are a group of cretaceous origin containing seven recognized living species: leatherback, hawksbill, Kemp's ridley, olive ridley, loggerhead, green, and flatback. The leatherback is the single member of the Dermochelidae family, whereas all other sea turtles belong in Cheloniidae. Analyses of partial mitochondrial sequences and some nuclear markers have revealed phylogenetic inconsistencies within Cheloniidae, especially regarding the placement of the flatback. Population genetic studies based on *D-Loop* sequences have shown considerable structuring in species with broad geographic distributions, shedding light on complex migration patterns and possible geographic or climatic events as driving forces of sea-turtle distribution. We have sequenced complete mitogenomes for all sea-turtle species, including samples from their geographic range extremes, and performed phylogenetic analyses to assess sea-turtle evolution with a large molecular dataset. We found variation in the length of the *ATP8* gene and a highly variable site in *ND4* near a proton translocation channel in the resulting protein. Complete mitogenomes show strong support and resolution for phylogenetic relationships among all sea turtles, and reveal phylogeographic patterns within globally-distributed species. Although there was clear concordance between phylogenies and geographic origin of samples in most taxa, we found evidence of more recent dispersal events in the loggerhead and olive ridley turtles, suggesting more recent migrations (<1 Myr) in these species. Overall, our results demonstrate the complexity of sea-turtle diversity, and indicate the need for further research in phylogeography and molecular evolution.

Published by Elsevier Inc.

1. Introduction

The sea turtles comprise seven extant species grouped into two families: Dermochelidae, with the leatherback (*Dermochelys coriacea*) as the single extant species, and Cheloniidae, with six species: hawksbill, Kemp's ridley, olive ridley, loggerhead, green, and flatback turtles (*Eretmochelys imbricata*, *Lepidochelys kempii*, *L. olivacea*, *Caretta caretta*, *Chelonia mydas*, and *Natator depressus*, respectively). Their phylogenetic placement has been somewhat debated, with different molecular data sets supporting different groupings within Cheloniidae. The placement of *N. depressus* has been particularly problematic, with different data supporting it as the sister taxon either to a clade comprising the genera *Eretmochelys*, *Caretta*, and *Lepidochelys* (Dutton et al., 1996; Iverson et al., 2007), or to *Chelonia* only (Naro-Maciel et al., 2008).

Most sea turtles (except *L. kempii* and *N. depressus*) have a pantropical distribution across a wide latitudinal range from Canada to South Africa, Southern Argentina and Chile (Hirth et al., 1997). Genetic studies based on the mitochondrial *D-Loop* of *C. mydas* (Encalada et al., 1996), *D. coriacea* (Dutton et al., 1999), and *L. olivacea* (Bowen et al., 1991; Karl and Bowen, 1999) suggest differentiation of Indo-Pacific and Atlantic groups. This implies that South and Central America and the Isthmus of Panama represents a stronger geographic barrier to gene flow than do colder waters in the southern tip of Africa (Avise et al., 1992; Dutton et al., 1999), at least in these three species.

Recent advances in DNA sequencing technologies have made more molecular markers available for turtle phylogenetics. Previous studies have used as many as 14 nuclear markers across selected turtle lineages (including freshwater and terrestrial turtles; Barley et al., 2010), and five nuclear and two mitochondrial markers in marine turtles (Naro-Maciel et al., 2008). However, in terms of mitochondrial phylogenetics, only *cytochrome b* (*Cytb*) (Bowen et al., 1993), *D-Loop*, *ND4* (Dutton et al., 1996) and *12S* and *16S* (Naro-Maciel et al., 2008) regions have been used, producing highly supported trees for contrasting topologies (see Naro-Maciel et al., 2008).

In other vertebrate groups, complete mitogenomes have demonstrated an increase in phylogenetic performance in terms of branch support and divergence-time estimation relative to

* Corresponding author.
E-mail addresses: garzonsebastian@hotmail.com (S. Duchene), amy.frey@noaa.gov (A. Frey), alonzoalfaro@gmail.com (A. Alfaro-Núñez), peter.dutton@noaa.gov (P.H. Dutton), mtpgilbert@gmail.com (M. Thomas P. Gilbert), phillip.morin@noaa.gov (P.A. Morin).

1055-7903/$ - see front matter Published by Elsevier Inc.
http://dx.doi.org/10.1016/j.ympev.2012.06.010

individual mitochondrial regions, and even nuclear markers (for examples in other vertebrates see Duchene et al., 2011; Okajima and Kumazawa, 2010; Wang and Yang, 2011).

Although phylogenetic analyses using nuclear markers have made important contributions in uncovering evolutionary relationships in many taxa, branch support is sometimes low (Duchene et al., 2011). In contrast, complete mitogenomes often provide highly supported trees and precise date estimates, often congruent with nuclear data, allowing for well-supported hypotheses for the true evolutionary histories of species. However, incomplete lineage sorting, hybridization, and past gene flow can obscure evolutionary relationships, and in some taxa the addition of independent lines of evidence, such as nuclear markers or morphology is crucial to obtain reliable phylogenetic resolution. In some cases, mitochondrial trees can be poorly supported even when using complete mitogenomes (Talavera and Vila, 2011); in sea turtles, however, this has not been the case (Drosopoulou et al., 2012; Shamblin et al., 2012).

In addition to phylogenetic relationships and divergence-time estimation, evolutionary reconstructions based on the entire mitogenome can benefit from genome characterization, identification of rates of evolution, and characterization of how these rates vary along particular genomic regions. Although some non-coding regions of the mitogenome are often assumed to evolve neutrally, it is important to highlight the molecule's crucial role in cellular respiration. Therefore, the finding that some sites may be under positive selection and play an important role in environmental adaptation in other animals (Foote et al., 2011; Garvin et al., 2011) must be taken into account in phylogenetic reconstructions and inferences of evolutionary processes.

We address several important topics concerning sea turtle evolution with large amounts of new data. The first is the phylogenetic relationships among species and the distinction of several groups, including the placement of *N. depressus* in relation to *Chelonia* and the concordant phylogeographic patterns of some globally distributed species. Secondly, the timing of sea turtle speciation events is key in understanding the timescale of turtle evolution and its relation to origins of geographic barriers such as the establishment of the Isthmus of Panama and changes in temperature of southern ocean currents around Africa, as has been previously suggested for turtles (Dutton et al., 1999, 1996; Encalada et al., 1996; Naro-Maciel et al., 2008) and other marine organisms (Rosen, 1988). Lastly, particular genomic features have been found in a wide variety of taxa, and have not been thoroughly investigated in sea turtles, such as an extra base pair not translated in *ND3* in birds and terrestrial turtles (Mindell et al., 1998), as well as variation in selective constraints across the mitogenome.

We have sequenced complete mitogenomes for a set of samples of all extant sea-turtle species, and collected sequences available from GenBank to produce a large mitogenome phylogeny of these taxa. Samples from across geographic ranges have been included for several species to compile the genetic diversity and elaborate on intra-specific phylogeographic patterns and diversification events.

Different molecular clock and phylogenetic frameworks were tested, and provide a basis for further mitogenomic studies in these taxa in the form of secondary calibrations (for a discussion on secondary calibrations see: Ho and Phillips, 2009; Ho et al., 2008). Furthermore, we explore particular characteristics of the mitogenome and scan for codon sites under different selective constraints from a structural and phylogenetic perspective.

2. Methods

2.1. Sampling and geographic coverage

A total of 24 sea turtle samples from known localities were sequenced and combined with additional GenBank sequences for *L. olivacea* (Tandon et al., 2006), *E. imbricata* (Tandon et al., 2006), and *C. mydas* (Okajima and Kumazawa, 2010). The geographic provenance of GenBank sequences were not publically available, so the *D-Loop* was compared to a stock assessment database to assign the most likely geographic region for these data. Table 1 lists all samples including outgroups, geographic origin, GenBank accession numbers, and bibliographic reference.

2.2. Sequencing

The complete mtDNA genomes of a green turtle from Tortugero, Costa Rica (haplotype *Cmydas T CR*); and a leatherback (haplotype *D coriacea O CR*) and Olive ridley (haplotype *L olivacea O CR*) from Ostional, Costa Rica were generated through Roche (454) FLX sequencing of PCR amplicons. The mtDNA genome was first PCR-amplified in two long overlapping 2 kb and 15 kb fragments. Subsequently the PCR products were purified, fragmented through nebulization, converted into MID-tagged sequencing libraries and sequenced as a partial fraction of an LR70 GS-FLX (Roche) run. The generated sequences were assembled into the complete mitogenome using the previous green (*Chelonia mydas*; Genbank ID AB012104), hawksbill (*E. imbricata*; Genbank ID DQ533485) and Olive ridley (*L. olivacea*; Genbank ID DQ486893) mitogenomes as reference sequences.

Genomic DNA libraries for the rest of the samples were prepared and given individual indexing sequences for multiplexing prior to pooling, library enrichment and sequencing as described in Hancock-Hanser et al. (submitted for publication). Sample libraries were pooled prior to capture array enrichment, and sample libraries for all species were enriched using sequence baits from the published mitochondrial genome of *Chelonia mydas* (Okajima and Kumazawa, 2010). The pooled, enriched library was sequenced on the Illumina Genome Analyzer II (Illumina Inc., La Jolla, CA).

2.3. Sequence assembly and mitogenome annotation

Contigs for the 24 mitogenomes were assembled using reference sea turtle mitogenomes for *L. olivacea* (Tandon et al., 2006), *C. mydas* (Okajima and Kumazawa, 2010) and *E. imbricata* (Tandon et al., 2006) found in GenBank (see Table 1) using Geneious v 4.7 (Drummond et al., 2009). All mitogenomes including GenBank references were then aligned using ClustalW v 2 (Larkin et al., 2007).

Gene identification and annotation were performed by importing GenBank sequence annotations into the newly assembled mitogenomes, followed by a complete inspection of individual gene coverage and reading-frame matching in each of the new mitogenomes.

2.4. Phylogenetic models and mitogenome characterization

Complete mitogenomes for terrestrial turtles to be used as outgroups, *Chelydra serpentina* and *Macrochelys temminckii* (Nie and Yan, 2006), were downloaded from GenBank and aligned with all sea turtles (including GenBank sequences) using Clustal W v2, producing a 17056 bp alignment.

Individual genes and non-coding regions were extracted from the alignment (producing 39 partitions) according to the imported GenBank annotations. Reading frames were visually inspected and base frequencies and proportions of variable sites were estimated using the APE package v2.8 (Paradis et al., 2004). In order to avoid possible frameshift due to gene overlap (between 3 and 10 bp) such as in *ATP8* and *ATP6*, extracted regions were concatenated after verifying their reading frames, producing a final alignment of 17094 bp. Although this procedure artificially increased the alignment length by 38 bp, due to overlapping sites, it is effective