Molecular Phylogenetics and Evolution 122 (2018) 102–109

Contents lists available at ScienceDirect

Molecular Phylogenetics and Evolution

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

Mitogenomics supports an unexpected taxonomic relationship for the extinct diving duck *Chendytes lawi* and definitively places the extinct Labrador Duck

Janet C. Buckner[a], Ryan Ellingson[a], David A. Gold[b], Terry L. Jones[c], David K. Jacobs[a,*]

[a] Department of Ecology and Evolutionary Biology, University of California, Los Angeles, CA 90095, United States
[b] Division of Biology and Biological Engineering, California Institute of Technology, Pasadena, CA 91125, United States
[c] Department of Social Sciences, California Polytechnic State University, San Luis Obispo, CA 9340, United States

ARTICLE INFO

Keywords:
Biodiversity
Anatidae
Ancient DNA
Labrador duck
Mitochondrial genome
California paleontology
Duck evolution

ABSTRACT

Chendytes lawi, an extinct flightless diving anseriform from coastal California, was traditionally classified as a sea duck, tribe Mergini, based on similarities in osteological characters. We recover and analyze mitochondrial genomes of *C. lawi* and five additional Mergini species, including the extinct Labrador Duck, *Camptorhynchus labradorius*. Despite its diving morphology, *C. lawi* is reconstructed as an ancient relictual lineage basal to the dabbling ducks (tribe Anatini), revealing an additional example of convergent evolution of characters related to feeding behavior among ducks. The Labrador Duck is sister to Steller's Eider which may provide insights into the evolution and ecology of this poorly known extinct species. Our results demonstrate that inclusion of full length mitogenomes, from taxonomically distributed ancient and modern sources can improve phylogeny reconstruction of groups previously assessed with shorter single-gene mitochondrial sequences.

1. Introduction

True ducks (subfamily Anatinae) are a distinct clade of birds whose evolutionary history is valuable for understanding past and present environments. Unfortunately, the phylogenetic relationships within this group remain problematic, making it difficult to reconstruct the life history of several interesting but extinct duck species. We reconstruct the Anatinae phylogeny to systematically place the extinct diving duck *Chendytes lawi*. Miller (1925) erected the genus *Chendytes* based on Holocene fossil material from the California coast and nearby Channel Islands (Fig. 1). Two species are known. The goose-sized *C. lawi* has more degeneration of the wing elements than the smaller *C. milleri,* which may represent an intermediate form between a flying ancestor and the flightless *C. lawi* (Howard, 1955). Known only from the Pleistocene of San Nicolás Island, *C. milleri* is more limited in abundance and geography than *C. lawi,* which has an extensive Holocene record extending from northern Baja to southern Oregon (Jones et al., 2008a; Gruhn and Bryan, 2006). Carbon dating and the frequent recovery of material from middens suggest that the latter species was eventually lost to human exploitation, but unlike many other extinct Pleistocene lineages it persisted until as recently as 2400 years ago (Jones et al., 2008a; Grayson, 2008).

Chendytes was traditionally classified as a sea duck, tribe Mergini. Miller (1925) allied it with the Surf scoter (*Melanitta perspicillata*), but an extended study by Livezey (1993) suggested placement in the eider genus *Somateria*. Despite uncertainty regarding the modern genus closest to *Chendytes,* previous authors consistently placed it amongst the Mergini based on osteological characters and proportions (Howard, 1947, 1955, 1964; Livezey, 1993; Miller, 1930; Miller et al., 1961). Nevertheless, several characters used for phylogenetic placement of *Chendytes* were found to be convergent, as they also occur in other diving Anatinae clades such as *Tachyeres* (steamer ducks) and Aythyini (scaups/pochards), as well as in more basal diving anseriforms such as *Oxyura* (e.g. Ruddy duck) (Livezey, 1993; Miller, 1930).

Here, we address the systematic placement of *Chendytes lawi* using molecular data. We generated mitochondrial genome sequences for *C. lawi* and five additional sea duck species, including the extinct Labrador Duck, *Camptorhynchus labradorius*. We analyzed these in combination with other anatid mitochondrial sequences. Using maximum likelihood and Bayesian inference methods, we compare phylogenetic results from three alternative data matrices: (1) maximized taxonomic sampling with missing data (2) mitogenomes only with limited taxonomic sampling and (3) a two-gene matrix with maximized taxonomic sampling and zero missing data. Our results consistently indicate that *Chendytes lawi* is not a member of any currently recognized diving duck clade but is a stem dabbling duck with convergent osteological adaptations for diving. *Camptorhynchus labradorius* is an eider that is sister to *Polysticta stelleri* (Steller's Eider) within the sea-duck tribe Mergini. The

* Corresponding author at: 2155 Terasaki Life Sciences Building, 610 Charles E. Young Drive East, Los Angeles, CA 90095-7246, United States.
E-mail address: djacobs@ucla.edu (D.K. Jacobs).

https://doi.org/10.1016/j.ympev.2017.12.008
Received 20 June 2017; Received in revised form 27 October 2017; Accepted 5 December 2017
Available online 13 December 2017
1055-7903/

combination of mitogenomic and single mitochondrial gene sequences improve estimates of phylogeny within the Anatinae.

2. Methods

2.1. Sample collection

In this study, we follow the binomial list of Clements et al. (2017) for species – eBird/Clements checklist of birds of the world – supplemented by the higher taxonomic classifications outlined by Cracraft (2013). Samples of *Chendytes lawi* bone fragments were provided by TLJ. The bones were recovered from archaeological site, CA-SLO-2, on the central California coast in San Luis Obispo County, midway between Avila Beach and Morro Bay (Jones et al. 2008b).[1] Two toe pad samples from separate individuals of *Camptorhynchus labradorius* (Labrador Duck) were lent by the American Museum of Natural History (AMNH). One toe pad sample from each of the extant species of Mergini (*Mergus serrator*, *Mergus merganser americanus*, *Melanitta nigra* and *Melanitta fusca deglandi*) were provided by the Donald R. Dickey Collection of Birds and Mammals at the University of California, Los Angeles (UCLA DC). All other mitochondrial sequences for anatid species were acquired from Genbank. See S1 for catalog numbers, accession numbers and species information.

2.2. DNA extraction, library preparation and target enrichment

DNA was extracted from a left tibiotarsus shaft of *Chendytes lawi* in a dedicated ancient DNA facility at UCLA. Prior to extraction the outer surface of the bone was removed with a sterilized dremel tool and a new, disposable rotary head to reduce exogenous contamination. The sample was subsequently ground into a coarse powder. This powder was then incubated in a solution of EDTA pH 8.0 and proteinase-K for 24 h on a rotator followed by 3 h of incubation at 56 °C. DNA was then treated to the silica-adhesion protocol described in Rohland and Hofreiter (2007). The resulting DNA extract was quantified using an Agilent 2100 Bioanalyzer (Agilent Technologies, Santa Clara, California, USA) and then stored at –20 °C prior to downstream application.

DNA extraction from all toe pads was performed in a dedicated hood for historic DNA extraction in a pre-PCR DNA free laboratory. We followed the manufacturer's protocol for the Qiagen DNEasy extraction kit (Qiagen Inc., Valencia, California, USA) with the following modifications: adding 40 uL proteinase K and 20 uL of 1 M dithiothreitol (DTT) to the extraction buffer and incubating the sample at 50 °C until completely digested, about 48 h (Fulton et al., 2012a). Qubit (ThermoFisher) fluorometric quantification of double stranded DNA followed extraction and all subsequent steps of the protocol.

After extraction, we prepared libraries from all DNA extracts with end-repair and dual-indexed adapter ligation using the Kapa Biosystems LTP Library Preparation Kit and custom indexes from BadDNA UGA oligos, respectively, following the manufacturers' protocols. Resulting DNA was size selected with Seramag Speedbeads to exclude fragments less than 150 base pairs (bp) (Rohland and Reich, 2012). Note: Though libraries were prepared in the same way for all samples, the extraction and library preparation (indexing) of the *Chendytes lawi* specimen was completed in the dedicated ancient DNA facility before any handling or processing of the historic toe pad samples.

We synthesized an 80mer bait set with 8× tiling at MYcroarray based on eight published anatid mitogenomes (*Dendrocygna javanica, Cygnus atratus, Mergus squamatus, Cairina moschata, Aythya americana, Anas formosa, Anas crecca, Anas platyrhynchos*) to target the full mitochondrial genome (see S1 for accession numbers). Following library preparation and quantification, we performed target capture of mitogenomes following the manufacturer's protocol except that hybridization and subsequent washes were carried out at a temperature of 55 °C. Dual-end sequencing (2 × 300 bp) was performed on the pooled, enriched libraries on a MiSeq instrument (Illumina Inc., San Diego, California, USA).

2.3. Read processing

Reads were de-multiplexed and sequence quality was evaluated using FastQC (Andrews, 2010). As the insert lengths from our ancient and historic DNA were on average several base pairs shorter than our read length, we created a first set of "QC reads" by cutting the sequences down to the first 70 bp. We also used the Trimmomatic pipeline enabled on the online platform Galaxy (version 0.32.3) on the raw reads to remove adapter contamination and sequencing artifacts to create a second set of QC reads (Bolger et al., 2014). Poor quality leading and trailing ends were removed and sequences were trimmed based on a sliding window of 5 base pairs where windows with an average quality less than 30 were removed. Reads from Trimmomatic of fewer than 30 bp were excluded. We mapped these two sets of processed reads for each of our museum samples from extant species as well as the Labrador Duck to the single published mitogenome for a Mergini species, *Mergus squamatus,* using Geneious Pro 9.0.5 (Kearse et al., 2012). Reads from *Chendytes lawi* were verified by eye and suspected contaminants were identified using NCBI's nucleotide BLAST and removed. Once all samples were processed, contigs were assembled to produce two mitogenome sequences for each species based on the 70mer and Trimmomatic processed reads. We compared the mappings for the two datasets for each species visually by aligning the assembled contigs and looking for discrepancies. When discrepancies were found, the highest base call quality was used to decide between nucleotides. When there was no difference in base call quality, the Trimmomatic mappings were chosen over the 70mer mappings.

2.4. Validation of historic DNA

We used mapDamage 2.0 (Jónsson et al., 2013) to identify nucleotide mis-incorporations in our ancient and historic DNA samples (S2). Such damage patterns are typical and expected from these historical sources of genetic data and thus also serve as a method of confirming that ancient DNA (aDNA) sequences are from the target specimen and not from modern contaminants. We used Bowtie (Langmead et al., 2009) to map the MiSeq reads back to the assembled mitogenomes of *C. lawi* and *M. f. deglandi*. The resulting alignment (BAM) files were used as input files for mapDamage.

2.5. Phylogenetic reconstruction

We aligned mitogenome data to previously published Anatid mitogenomes and mitochondrial gene sequences available from GenBank using the MUSCLE algorithm in Geneious Pro 9.0.5. Each taxon had between one and four gene fragments, or a complete mitogenome, represented in the alignment. In all, 32 taxa had complete mitogenomes, 14 had four gene fragments, 17 had three fragments, 30 had two fragments and 11 taxa had one fragment and were present in the combined matrix (matrix A) of 104 total taxa (S1). Two additional partial matrices of 32 taxa with complete genomes (matrix B), and 72 taxa with both Cyt b and COI fragments without missing data (matrix C) were generated. In matrix C, one taxon with missing data was included, *Polysticta stelleri*, in order to confirm relationships to extinct species recovered from analyses of matrix A. Note that the control region was

[1] [1]The *C.lawi* fossils were excavated in 1968 (Greenwood, 1972) from a 3.4 m deep shell midden deposit, and were identified in 2007 by Judy Porcasi who relied on comparative materials at Los Angeles County Museum of Natural History (Paleontology Department.) and the Department of Biology at California State University, Long Beach. Thirty-three radiocarbon determinations show that the archaeological site was occupied from 10,300 to 300 cal BP. The specimen that produced the sequence was recovered from unit S4/W2, from a depth of 270–280 cm (Jones et al., 2008b). Eight radiocarbon determinations show that the levels between 280 and 200 cm date between 9000 and 5000 cal BP.