DNA from a 100-year-old holotype confirms the validity of a potentially extinct hummingbird species

Jeremy J. Kirchman1,*, Christopher C. Witt2, Jimmy A. McGuire3 and Gary R. Graves4

1 New York State Museum, Albany, NY 12230, USA
2 Department of Biology and Museum of Southwestern Biology, University of New Mexico, Albuquerque, NM 87131, USA
3 Museum of Vertebrate Zoology, University of California, Berkeley, CA 94720, USA
4 National Museum of Natural History, Smithsonian Institution, PO Box 37012, Washington, DC 20013-7012, USA
*Author for correspondence (jkirchma@mail.nysed.gov).

We used mtDNA sequence data to confirm that the controversial 100-year-old holotype of the Bogotá sunangel (*Heliangelus zusii*) represents a valid species. We demonstrate that *H. zusii* is genetically well differentiated from taxa previously hypothesized to have given rise to the specimen via hybridization. Phylogenetic analyses place *H. zusii* well differentiated from taxa previously hypothesized to have given rise to the species via hybridization. Phylogenetic analyses place *H. zusii* as sister to a clade of mid- to high-elevation Andean species currently placed in the genera *Taphrolesbia* and *Aglaiocercus*. *Heliangelus zusii*, presumed extinct, has never been observed in nature by biologists. We infer that the species occupied a restricted distribution between the upper tropical and temperate zones of the northern Andes and that it was most probably driven to extinction by deforestation that accompanied human population growth during the nineteenth and early twentieth centuries. We demonstrate the feasibility of obtaining DNA from nearly microscopic tissue samples from old hummingbird specimens and suggest that these methods could be used to resolve the taxonomy of dozens of avian taxa known only from type specimens.

Keywords: ancient DNA; Andes Mountains; hummingbird; hybrid; extinction

1. INTRODUCTION

Cataloguing the biota of biologically unexplored regions often follows closely on the heels of human population expansion and subsequent waves of anthropogenically caused extinctions (Fuller 2001). Many species that were rare when first encountered by biologists have not been observed since the type specimens were collected. Species known only from unique specimens are rightfully viewed with caution because they may represent hybrids or aberrant phenotypes. Living populations of these biological mysteries are occasionally ‘rediscovered’ in nature many decades or even centuries after their scientific descriptions, providing hope to conservationists and new data to


2. MATERIAL AND METHODS

(a) Laboratory procedures

Skin cells were scraped from the feet of ANSP 159261 using a sterile scalpel blade at the National Museum of Natural History, Smithsonian Institution. Extraction and PCR setup (with negative controls and replications) were performed in the dedicated aDNA laboratory at the New York State Museum, where no previous work on hummingbirds had been performed. We employed a modified version of the silica-based DNA extraction method of Höss & Pålbo (1993). We designed six PCR primer pairs to target short segments of the mitochondrial ND2 and ND4 genes and their flanking tRNAs. Extraction protocol, primer sequences and PCR conditions are described in the electronic supplemental material. In a separate laboratory at University of California, Berkeley, we obtained sequences from frozen tissues of 95 hummingbird species and from five swifts (Apodidae), a tree swift (Hemiprocnidae) and an owlet-nightjar (Aegothelidae), which were used to root phylogenetic trees. Taxon sampling within the hummingbird family Trochilidae included species from all nine major trochilid clades and all 18 genera within the ‘coquette’ clade (Lophornithini) (McGuire et al. 2008), including an additional 16 species beyond those included in McGuire et al. (2007) because we expected *H. zusii* to be nested within that group. Methods and primers for amplification and sequencing from frozen samples are available in McGuire et al. (2007). Specimen details and GenBank (NCBI) accession numbers are in the electronic supplemental material, table S1.

(b) Phylogenetic analyses

Four of the six primer pairs produced PCR products from the mitochondrial ND2, ND4, tRNA-Has and tRNA-Ser genes, totalling 356 base pairs (bp) of concatenated sequence. Phylogenies were reconstructed using maximum parsimony (MP), maximum likelihood (ML) and Bayesian approaches to analyse both a taxonomically broad sample of hummingbirds for only those characters for which we had data for *H. zusii* (95 hummingbird species plus outgroups, 356 bp), and a larger alignment of 4133 bp of the above
mitochondrial genes plus the nuclear β-fibrinogen intron 7 and adenylate kinase 1 intron 5 genes for 33 species of montane coquettes, one lowland coquette, *H. zusii* and outgroups. This combination of analyses allowed us to initially place *H. zusii* within the broad phylogenetic context of all hummingbird lineages, and then to bring as much data as possible to bear on the relationship of *H. zusii* to its closest relatives. We used the program NONA (Goloboff 1995) for MP and to compute Bremer support values (Bremer 1994). ML tree searches and bootstrapping (1000 replicates) were run in PHYML (Guindon & Gascuel 2003). Bayesian analysis using MRBAYES (Huelsenbeck & Ronquist 2001) included two runs of 10 million generations, sampling every 10 000 generations and excluding the first one million generations as burn-in. Details of substitution model selection are given in the electronic supplementary material.

3. RESULTS AND DISCUSSION

Initial analysis of the complete taxon sample (356 bp) placed *H. zusii* (ANSP 159261) unambiguously within the high-elevation group of coquettes (McGuire et al. 2008) in a clade consisting of species placed in *Taphrolesbia* and *Aglaiocercus*, but did a poor job recovering well-established, higher-level hummingbird clades. Our MP, ML and Bayesian analyses of the more character-rich (4133 bp) alignment of 35 coquette species plus outgroups provided strong support for the hypothesis that *H. zusii* represents a distinct lineage allied with *Taphrolesbia* and *Aglaiocercus* (figure 1), with high Bremer support (6), ML bootstrap (95%) and Bayesian posterior probabilities (1.00) for the critical node linking *H. zusii* to its sister taxa. ML branch lengths (not shown) as well as pairwise genetic distances indicate that *H. zusii* is not a hybrid, but is a distinct taxon exhibiting

Figure 1. Consensus Bayesian phylogeny of select hummingbird species showing the relationship of *H. zusii* to other high-elevation coquettes. Numbers at nodes are Bayesian posterior probabilities/ML bootstrap support (1000 replicates)/MP Bremer support. Trees were rooted with an outgroup comprising swifts, a treeswift and an owlet-nightjar.

Ancient DNA from extinct hummingbird  J. J. Kirchman et al. 113

species-level divergence from all sampled species (at least 3.3% at ND2 + ND4, 245 bp) and is 12.3–17.4% divergent from six species of *Heliangelus* (table 1).

In rejecting the hybrid hypothesis for *H. zusii*, we feel certain that all potential parental species were sampled. Although we did not sample *Aglaiocercus berlepschi* (Venezuelan sylph), this taxon is morphologically similar to *Aglaiocercus kingi*, and the two are considered conspecific by many authorities. The three species of *Aglaiocercus* (sylphs) undoubtedly form a monophyletic group based on morphology. The 19 species of coquettes not sampled are in the genera *Lophornis*, *Discosura*, *Setophoxoides*, *Oreotrichus*, *Heliangelus*, *Chalcostigma*, *Metallicula* and *Phlogophilus*; all these genera form well-supported clades that are genetically divergent from the clade containing *H. zusii*, and only *Chalcostigma* is not monophyletic (figure 1; McGuire et al. 2007). Robust support for the affinity of *H. zusii* with the *Taphrolesbia–Aglaiocercus* clade necessitates generic reassignment of *zusii*. We defer this action pending a general review of generic limits within the Trochilidae.

Our findings permit us to more accurately circumscribe the probable geographical range and habitat of *zusii*. Tens of thousands of hummingbird trade skins were exported from Bogotá for the millinery trade in the nineteenth century (Doughty 1975), and nearly all refer to species restricted to the northern Andean region of Colombia and adjacent Ecuador (Berlioz & Jouanin 1944). Graves (1993) speculated that *zusii* originated from the Eastern Cordillera of the Colombian Andes within a few hundred kilometres of Bogotá, or possibly in the Central Cordillera, in cloud forest between 1400 and 2200 m (altitude above sea level). The fact that only a single specimen is known despite the extensive collection of showy hummingbirds suggests that it had a relictual or restricted geographical distribution when collected. The three species of *Aglaiocercus* inhabit humid Andean forest (900–3000 m) from the Coastal Range of Venezuela south to northern Bolivia, whereas *Taphrolesbia griseventris* occurs in a semi-arid region (2750–3170 m) of northern Peru, characterized by cacti and other xerophytic plants (Collar et al. 1992). Assuming that the ecology of *zusii* falls between these extremes, the search for *zusii* should be expanded to include semi-arid habitat in the Andes as high as 3200 m from northwestern Venezuela south of northern Peru. We presume that *zusii* is extinct, but hummingbird species continue to be discovered (Fitzpatrick et al. 1979; Graves 1980; Cortes-Diago et al. 2007), suggesting that *zusii* may persist in an unexplored region of the Andes, such as an outlying cordillera (e.g. *Heliangelus regalis*) or an isolated peak (e.g. *Ramphomicron dorsale*).

aDNA from unique types may help resolve longstanding taxonomic problems in ornithology and other disciplines. Our finding that mtDNA from ANSP 159261 is highly divergent from potential parental species rules out a hybrid origin and precludes sequencing nuclear DNA, which is much more difficult to obtain from century-old archival material (Kirchman 2009). The use of silica-based techniques, originally developed for extractions from organic remains of Pleistocene origin, should enable genetic characterization of other extremely small samples of archival material such as those we have obtained from the tiny feet of a hummingbird museum specimen.

We are grateful to Nate Rice (ANSP) for permission to examine and sample the specimen and to Kelly B. Miller for help with parsimony analyses. We thank the following museums for providing vouchered samples that provided the basis for the original DNA sequences reported in this paper: Academy of Natural Sciences of Philadelphia, Zoological Museum of Copenhagen, Field Museum of Natural History, Louisiana State University Museum of Natural Science and University of New Mexico Museum of Southwestern Biology. Funding for aDNA laboratory work was provided by the New York State Museum. Funding for sequencing of non-ancient hummingbird samples was provided by NSF (DEB-0543556 and DEB-0330750).


**Supplementary material**

Fig. S1. Holotype of “Heliangelus” zusii (ANSP 159261)

Fig. S2. “Heliangelus” zusii and its close relatives (upper left, Aglaiocercus kingi; upper middle, Aglaiocercus coelestis; upper right, Heliangelus zusii; and lower right, Taphrolesbia griseiventris). Painting by Jon Fjeldså ©, University of Copenhagen (see next page).
Ancient DNA Extraction, Amplification, and Sequencing: Contamination controls carried out in the ancient DNA lab include negative extraction controls (containing no tissue), glove changes between handling each sample, ultraviolet irradiation of all plastics and buffers, exclusive use of aerosol-barrier pipette tips, and daily sanitation of all equipment and surfaces with 10% bleach solution. The skin sample from ANSP 159261 was incubated for 48 h with agitation at 55°C in 600 uL of an extraction buffer consisting of 7.5 M guanidinium thiocyanate, 0.1 M Tris-HCl (pH 6.4), 0.02 M EDTA, and 1.3% Triton X-100. After centrifugation, 500uL of supernatant was removed to a second tube containing 500uL of binding buffer (7.5 mol/L guanidinium thiocyanate, and 0.3M sodium acetate [pH 5.2]) and 40uL of saturated silica suspension (SiO2 in water). DNA was bound to the silica for 3 hr at 27°C, pelleted by centrifugation, washed once with binding buffer and twice with 52% ethanol supplemented with 0.125M sodium chloride, 0.01M Tris [pH 8.0] and 0.5M EDTA. Bound DNA was eluted by incubating the pellet twice in 75uL volumes of TE buffer for 30 minutes at 55°C.

PCR conditions and primer sequences: Amplifications were performed in 50uL reactions containing 5.0 uL of DNA extract, 0.5uM of each primer, 0.4mM dNTPs, 2.5mM MgCl2, 0.5mM bovine serum albumen, and 1 unit of AmpliTaq Gold DNA polymerase. Primer pairs that successfully amplified DNA from ANSP159261 were F1 (GCAACATTCCTAACAGCCTCAT) R1 (GATGTTCTCGTGAGGGTGGGTTT), F2 (AAACATTAGCCTGTGATCCTAAAAA) R2 (TCCTAAGACCAGTGATTTAACTG), F4 (TCCAAGACTCTCCAAACAAGA) R4
(GGGAGG GTGATTGTTGAGTAGT), and F5 (TCCY TRCTGCTAAACCRCAGCAA) R5
(ATGTGAGGAAGAGAA TAGTGAT). PCR products were cut from agarose gels, purified
using a Qiagen MinElute gel extraction kit, and were sequenced on an Applied Biosystems
3730XL DNA Sequencer.

doi:10.1093/bioinformatics/14.9.817) to select appropriate models for seven data partitions used
in subsequent ML and Bayesian analyses. The partitions and models are as follows:

<table>
<thead>
<tr>
<th>Data Partition</th>
<th>Number of substitution types</th>
<th>Among-site rate variation</th>
</tr>
</thead>
<tbody>
<tr>
<td>tRNAs</td>
<td>2</td>
<td>Proportion of invariant sites</td>
</tr>
<tr>
<td>ND2 1st codon position</td>
<td>6</td>
<td>Proportion of invariant sites + Gamma</td>
</tr>
<tr>
<td>ND2 2nd codon position</td>
<td>6</td>
<td>Proportion of invariant sites + Gamma</td>
</tr>
<tr>
<td>ND2 3rd codon position</td>
<td>6</td>
<td>Proportion of invariant sites + Gamma</td>
</tr>
<tr>
<td>AK1</td>
<td>6</td>
<td>Gamma distribution</td>
</tr>
<tr>
<td>bfib</td>
<td>6</td>
<td>Gamma distribution</td>
</tr>
<tr>
<td>ND4 1st codon position</td>
<td>6</td>
<td>Proportion of invariant sites + Gamma</td>
</tr>
<tr>
<td>ND4 2nd codon position</td>
<td>6</td>
<td>Proportion of invariant sites + Gamma</td>
</tr>
<tr>
<td>ND4 3rd codon position</td>
<td>6</td>
<td>Proportion of invariant sites + Gamma</td>
</tr>
</tbody>
</table>