Extreme mitogenomic divergence between two syntopic specimens of *Arremon aurantirostris* (Aves: Emberizidae) in central Panama suggests possible cryptic species

Krisangel López1, Cody Angeli1, Celestino Aguilar2,3, Jose R. Loaiza2, Luis Fernando De León2, W. Owen McMillan3, and Matthew J. Miller1,2,3

1Department of Biology, Villanova University, Villanova, PA, USA; 2Instituto de Investigaciones Científicas y Servicios de Alta Tecnología (INDICASAT-AIP), Panamá, República de Panamá, and 3Smithsonian Tropical Research Institute, Panamá, República de Panamá

Abstract

We report the complete mitochondrial genome of two specimens of Orange-billed Sparrow (*Arremon aurantirostris*) from Colón Province, in central Panama. The two specimens were collected on the same day, and at the same locality; however, they showed substantial divergence (6.3% average pairwise divergence among coding genes). A survey of ND2 sequence variation across Panama suggests that this divergence is the result of geographic differentiation and secondary contact. This high level of mitochondrial divergence among co-occurring individuals raises the possibility of multiple biological species in Orange-billed Sparrows. Our results are yet another demonstration that much remains to be discovered regarding avian biodiversity in Panama and throughout the Neotropics.

As part of ongoing studies on speciation in Panamanian birds we generated full mitogenomes from two specimens of Orange-billed Sparrow (*Arremon aurantirostris*, Emberizidae). Genomic DNA was obtained from Smithsonian Tropical Research Institute Cryological Collection samples GMS1179 and GMS1180, which refer to University of Washington Burke Museum specimens 106 602 and 106 601. Both male birds were collected no more than 250 meters from each other on 13 March 2004 at Achiote, Colón Province in central Panama (9.22°N, 80.02°W).

Mitogenome sequences were obtained as a by-product of hybrid target capture techniques (Faircloth et al., 2012). Previously, we have reported (Aguilar et al., 2014; Loaiza et al., 2014; Miller et al., 2014) full mitogenomes obtained directly from the phyluce pipeline (https://github.com/faircloth-lab/phyluce). However, in this instance, initial assembly did not recover a complete mitochondrial genome. Instead, we used Bowtie 2 (Langmead et al., 2012) implemented in Geneious v7.0.6 (Biomatters, http://www.geneious.com) to map reads from GMS1179 to a reference mitogenome (KM078771; Paroreomyza montana, Lerner et al., 2011), which resulted in a complete mitogenome, which we then used to map reads from GMS1180 using Bowtie 2. We compared and edited the two alignments by eye. We used DOGMA (Wyman et al., 2004; http://dogma.cbb.utexas.edu/) to annotate both alignments.

Both recovered genomes were composed of 2 rRNAs, 22 tRNAs and 13 protein-coding genes in the standard avian order (Desjardins & Morais, 1990). Notably, the coding sequence for COX II differs from that of related birds (based on a BLAST search). COX II is typically 684 base pairs, ending with a TAA stop codon. However, the homologous codon for both GMS1179 and GMS1180 is CAA, confirmed by over 99% coverage in both alignments (depth of coverage: 1340 and 224 reads respectively). In the open reading frame, the first stop codon occurred nine base pairs downstream, resulting in a coding sequence of 693 base pairs that overlapped with the tRNA by eight base pairs. This finding should be investigated by targeted Sanger sequencing of related taxa. Otherwise, our results agree with other Passeriformes mitochondrial genomes. Both *Arremon* specimens had identical sizes of all protein-coding and non-coding regions with the exception of 16S rRNA which differed by one nucleotide. We found greater sequence similarity within non-coding regions than coding regions. Among coding regions, average sequence divergence was 6.3% (range: 4.5% for COX II to 8.0% for ND2). There were no differences in the start or stop codons within protein-coding genes. Annotated full mitochondrial genome sequences are deposited in Genbank: (accessions: KR780063 and KR780063). A table with mitogenome feature details is available at: http://dx.doi.org/10.6084/m9.figshare.1417972 (Supplementary Table 1).

The degree of pairwise divergence among protein-coding regions is above the level typically reported within bird species (Kerr et al., 2007; Price, 2008). High within-species mitochondrial divergence has been found to be the result of secondary contact among geographically distinctive lineages in other Panamanian birds (Miller et al., 2008, 2010, 2011, 2014). Thus we sequenced the mitochondrial ND2 gene for an additional three birds from both Bocas del Toro Province (in western...
Panama: 8.85 N, 82.17 W), Darien Province (eastern Panama: 7.76 N, 77.68 W), and three additional birds from Achiote. (Genbank accessions for these sequences: KR781504–KR781512). We generated a maximum parsimony tree in PAUP* 4.0a142 (Swofford, 2002) using a heuristic search with 100 bootstrap replicates, rooted to a sequence from Arremon schlegeli obtained from Genbank (HQ537420). The resulting tree indicated that the GMS1179 belongs to a clade containing A. aurantirostris from Bocas del Toro but not Darien, while GMS1180 belongs in a clade of birds from Darien, but not Bocas del Toro (Figure 1). Two of the three additional Achiote birds fall in the Bocas del Toro clade, while the third individual was part of the clade with the Darien birds. These results support the hypothesis that the extreme sequence divergence between the mitogenomes of A. aurantirostris reported here is the result of geographic isolation and secondary contact. Assuming the 1.8% per million year sequence divergence rate across protein-coding mitochondrial genes in Hawaiian honeycreepers (Lerner et al., 2011), our result suggests that eastern and western A. aurantirostris diverged 3.5 million years ago. Two alternatives are possible: either the two lineages of A. aurantirostris represent cryptic co-occurring lineages that do not interbreed, or they are reproductively compatible despite mitochondrial divergences that often indicate partial or complete reproductive isolation in birds (Lijtmaer et al., 2003; Price, 2008). We believe that these mitogenomes could offer an opportunity to study the role of mito-nuclear discordance in the speciation process (Hill, 2015).

Figure 1. Maximum parsimony tree of ND2 coding sequences. Bootstrap support (100 replicates) indicated at nodes.

Declaration of interest
The authors report no conflict of interest. The authors alone are responsible for the content and writing of the article. This work was funded by the Smithsonian Tropical Research Institute Office of Academic Programs and the Villanova University Biology Department. We thank Panama’s Environmental Ministry (MiAmbiente, formerly ANAM) for continued support of scientific collecting, without which, this study would not have been possible. We thank J. Klicka, G. Spellman, J. DeCosta, P. Guitton for collaboration in the field that was responsible for collecting the specimens used in this study.

References
Supplementary material available online

Supplementary Table 1