

# Historical demography of bird populations from Hispaniola assessed by nuclear and mitochondrial gene sequences

Anna L. FAHEY<sup>1,4</sup>, Robert E. RICKLEFS<sup>3</sup> and J. Andrew DEWOODY<sup>1,2,5\*</sup>

<sup>1</sup> Department of Forestry and Natural Resources, Purdue University, West Lafayette, Indiana 47907, U.S.A.

<sup>2</sup> Department of Biological Sciences, Purdue University, West Lafayette, Indiana 47907, U.S.A.

<sup>3</sup> Department of Biology, University of Missouri at St. Louis, St. Louis, Missouri 63121, U.S.A.

<sup>4</sup> Missouri Southern State University, 3950 E Newman Rd, Joplin, Missouri 64801, U.S.A.

<sup>5</sup> Purdue University, David C. Pfendler Hall, 715 W. State Street, West Lafayette, Indiana 47907, U.S.A.;  
e-mail: dewoody@purdue.edu

Received 30 November 2014; Accepted 8 April 2015

**Abstract.** Molecular markers can be used to infer the demographic history of a given species, but many historic processes simultaneously impact multiple species. Thus, comparative historical demography has the potential to provide insight into drivers of evolution. In this study, we used nuclear DNA (nDNA) sequences to corroborate (or refute) demographic inferences based on earlier mitochondrial DNA (mtDNA) data from 16 species of Hispaniola birds. Our previously published analysis suggested population expansion in five of six migratory species (following glacial retreat in North America), with less evidence of expansion in non-migratory species. Additional molecular markers should reduce locus-specific bias, and so we generated sequence data for several nuclear loci. Test statistics associated with the nDNA provided only equivocal evidence for population expansion in 10 of the 16 species. Discordance between mtDNA and nDNA is not uncommon because the two genomes are exposed to different selective pressures and have different effective population sizes and modes of inheritance. The nDNA analyses reported here cast some doubt on our earlier mtDNA inferences. They also suggest that the signal to noise ratio of demographic statistics is typically low because of the inherent variability in selective regimes and coalescence across loci.

**Key words:** intron, neutrality statistics, migration, Illumina MiSeq, amplicon sequencing

## Introduction

The island of Hispaniola harbors more biodiversity than any other island in the West Indies, and the evolution of its fauna is of considerable interest because of the complex geologic and climatic history of the Antillean archipelago (Donnelly 1989, Perfit & Williams 1989, Higuera-Gundy et al. 1999, Iturralde-Vinent & MacPhee 1999). Hispaniola consists of two paleo-islands (North and South), which are still partially sundered by a former marine channel that acts as a distribution barrier for some species (Mann et al. 1991, Townsend et al. 2007). The island as a whole contains three parallel mountain ranges separated by deep valleys, and it is associated with ten nearby offshore islands. Its topology varies dramatically, resulting in an ecologically heterogeneous landscape that includes flooded grasslands, savannahs, wet forests, dry forests, and row-crop agriculture (Latta

et al. 2006). During glacial periods, the climate was cooler and drier than at present (Bonatti & Gartner 1973, Curtis et al. 2001), which presumably increased the extent of suitable habitat for montane species and restricted the extent of tropical forest species (Higuera-Gundy et al. 1999). The varied topography and associated vegetation no doubt contribute to the considerable biodiversity found on the island.

Geographic barriers to dispersal on Hispaniola are reflected in the gene pools among the most vagile vertebrates, such as the birds. Molecular studies have revealed pronounced genetic structure of some resident species across the island, and coalescent approaches indicate that divergence times between isolated populations differ dramatically among species (roughly 0.5–10 MYA, Sly et al. 2010, 2011). These and other studies suggest a complex evolutionary history of Hispaniola's avifauna.

\* Corresponding Author

One factor contributing to the complex evolutionary histories of Hispaniolan birds may be the migratory tendencies and associated population dynamics of many species. The island is close enough to continental North America to support large numbers of wintering migrants, yet isolated enough to have evolved endemic species (Ricklefs & Bermingham 2008). Some of the bird species that live year-round and breed on Hispaniola are endemic to the island, whereas other non-migratory species have ranges that extend beyond Hispaniola (we refer to these as “resident” in contrast to “endemic” species). The migratory species that overwinter on the island but breed in the United States and Canada should have experienced population expansion following the Last Glacial Maximum (LGM) due to the increased breeding area associated with the retreat of the ice sheets, unless populations of these species are limited in their wintering ranges. In contrast to migrant species, we expect non-migrants to exhibit little evidence of population expansion because the melting of ice sheets at the end of the Pleistocene caused a concomitant rise in sea level of ~120 meters that reduced somewhat the area of the island as a whole (Bonatti & Gartner 1973, Fleming et al. 1998, Higuera-Gundy et al. 1999, Curtis et al. 2001, Clark & Mix 2002).

Contemporary gene pools harbor signals of historic population demography including long-term equilibrium, growth, and population decline (Harpending et al. 1998). We recently used mitochondrial DNA (mtDNA) sequences to infer the demographic history of 16 avian species sampled on Hispaniola. The mtDNA data suggested that migratory species experienced more pronounced population growth in their recent evolutionary history than did non-migratory species, perhaps because continental breeding habitat for migratory birds increased as glaciers retreated, whereas changes to the insular breeding habitat of non-migratory birds occurred on a more modest scale, and rising sea levels reduced the total land area (Fahey et al. 2012).

These inferences relied solely on mtDNA, a common molecular marker of choice because of its inherent nucleotide variability, conserved genomic organization, uniparental mode of inheritance, and haploid condition. The latter two attributes account for the 4-fold smaller effective population size ( $N_e$ ) of mtDNA relative to nuclear genes, of which each individual has two copies, one inherited from each parent. Thus, mtDNA lineages are expected to undergo lineage sorting more rapidly than nuclear genes (Avise 2000, Zink & Barrowclough 2008).

Such lineage sorting may be stochastic (as expected for genes behaving in a neutral manner), deterministic (as expected for genes under strong selection), or some combination thereof. Multiple (independent) genetic markers are likely to produce more robust demographic inferences and help to overcome the idiosyncratic histories of individual loci (Hare 2001, Dupuis et al. 2012).

Beyond multiple loci, studies of multiple taxa often reveal more robust insights than single-taxon studies, in part because similar patterns across species imply common evolutionary processes (e.g. vicariance, Knowlton et al. 1993). For example, phylogeographic patterns of individual species can be informative, but data from multiple sympatric species often provide important context for the evolutionary process (e.g. geologic or climatic events as drivers, Avise 2000). In terms of demography, congruent histories among taxa suggest that common environmental factors may be responsible (Bos et al. 2008, Qu et al. 2010). In contrast, disparate histories among sympatric taxa could point to species-specific demographic drivers (e.g. novel predators or pathogens, Ricklefs 2010). Herein, we reevaluate Fahey et al. (2012) mtDNA findings of better support for population expansion in migratory birds than in resident birds in light of new data from six nuclear introns.

## Material and Methods

### *Field sampling and laboratory sequencing techniques*

Avian blood samples were collected via venipuncture from the Aceitillar sector of the Dominican Republic’s Sierra de Bahoruco National Park, under permits issued by the government; a description of the field work can be found in Latta & Ricklefs (2010). Ten non-migratory species of birds were used in these analyses, including six resident species – *Coereba flaveola* (CFA), *Columbina passerina* (CPA), *Elaenia fallax* (EFA), *Loxigilla violacea* (LVI), *Myiarchus stolidus* (MST), *Turdus plumbeus* (TPL) – and four endemic species – *Microligea palustris* (MPA), *Phaenicophilus palmarum* (PPA), *Spindalis dominicensis* (SDO), and *Todus subulatus* (TSU). Six migratory species, all warblers (Parulidae), were also included in the analysis: *Setophaga caeruleascens* (SCR), *Setophaga discolor* (SDI), *Setophaga palmarum* (SPA), *Setophaga tigrina* (STI), *Mniotilta varia* (MVA), and *Seiurus aurocapilla* (SAU). Species descriptions are provided in Latta et al. (2006). Note the genus *Setophaga* was still referred to as *Dendroica* in Fahey et al. (2012), and thus DCR = SCR, DDI = SDI, DPA = SPA, and DTI = STI.

Details of our DNA extraction procedures can be found in Fahey et al. (2012). We used avian PCR primers described in Kimball et al. (2009) to amplify nuclear gene fragments. Primer pairs were tested in a subset of individuals from each of the 16 species, and those that produced amplicons in all 16 species were further evaluated by dideoxy sequencing. Of the > 20 nuclear loci evaluated, six amplified well in most of our study species and harbored sequence variants; these six markers were then amplified and sequenced in a larger subset of individuals.

The nuclear loci used in this study included *hmgn2*, *irf2*, *mb*, *nat15*, *pcbd*, and *rho*. None of these loci are syntenic in the chicken genome and, because of the conserved chromosome structure in bird genomes, they presumably are unlinked (Derjushcheva et al. 2004, Griffin et al. 2007). PCRs were conducted in 20µl total volumes using ~40 ng template DNA, 1U NEB Taq polymerase, 0.3 µM of each primer, MgCl<sub>2</sub> (as listed in Table 1), 10 mM Tris-HCl, 50 mM KCl, 0.5 mg/ml BSA, and 0.2 mM of each dNTP. PCRs were performed in Eppendorf MasterCyclers under the following conditions: 94 °C for two minutes; 31 cycles of 94 °C for 30 seconds, primer specific annealing temperature (Table 1) for 30 seconds, and 72 °C for 30 seconds; and concluding with a five minute extension at 72 °C. Touchdown annealing protocols were used to amplify *pcbd* in PPA and MVA (30 cycles at 68°, decreasing by 0.5° each cycle, then 58° for 30 more cycles) and *mb* in MVA individuals (30 cycles at 60°, decreasing by 0.5° each cycle, then 48° for 30 more cycles). For *pcbd* amplifications in some species, two sets of primers (internal and external) were used to generate better coverage of this long sequence. A low sodium acetate precipitation was used for PCR purification prior to sequencing. Amplicons were checked for size via gel electrophoresis and subsequently quantified using a Nanodrop8000 (Thermo Scientific) spectrophotometer. Dideoxy sequencing reactions employed BigDye version 3.1 chemistry and a Prism 3730XL sequencer (Applied Biosystems, Foster City, CA, U.S.A.) using ~50-100 ng of PCR template. Sequencher 4.7 (Gene Codes Corp., Ann Arbor, MI) was used for alignment of sequences into contigs. All candidate single nucleotide polymorphisms (SNPs) were carefully inspected by eye; when necessary, poor quality sequence reads were discarded and individuals were resequenced. Gametic phase was addressed computationally with PHASE (version 2.1, Stephens et al. 2001, Stephens & Scheet 2005).

We used PHASE to assign haplotypes to individual birds because this algorithm is generally robust even

when underlying assumptions are violated (e.g. selective neutrality, Bos et al. 2008). We first used a probability threshold of 0.60 to reduce the number of unresolved haplotypes with a minimal increase in false positives (Harrigan et al. 2008, Garrick et al. 2010). For PHASE runs where site probabilities were less than 0.60 (excluding polymorphic sites unique to an individual), then cloning techniques were used to help resolve gametic phase by priming the PHASE algorithm. In those cases, a Promega T/A cloning kit (pGEM-T Easy Vector System II) was used and amplicons were sequenced directly from individual colonies; these cloned sequences were then used as “known” haplotypes in subsequent PHASE analyses. When PHASE probabilities remained less than 0.60, we then removed individual polymorphic sites for all individuals of that species. This approach is imperfect because overall haplotype diversity is reduced (i.e. some information is discarded), but culling sites is preferable to culling unresolved individuals, as the latter approach more egregiously reduces overall phylogenetic diversity and biases population genetic parameters (Garrick et al. 2010).

Because cloning and traditional dideoxy sequencing are expensive and time consuming, we also used second-generation sequencing techniques to identify haplotypes that could prime the PHASE algorithm. We used the Nextera XT DNA sample preparation kit (Illumina) to simultaneously fragment and tag the six nuclear amplicons for subsequent sequencing using the Illumina MiSeq platform. Individual samples from each species were assigned a unique index/tag and six amplicons (one for each gene) were pooled for each species. The Nextera XT kit provides 24 tags, so we used the remaining eight tags to help further resolve highly polymorphic genes/individuals by again combining amplicons, but also having only one gene per species per tag (Table 2). We were able to tag multiple individuals with one index tag because we only had one representative for each of the six genes, and the genes could easily be sorted out computationally after the Illumina run (i.e. reads from the same amplicons were assembled into the same contigs).

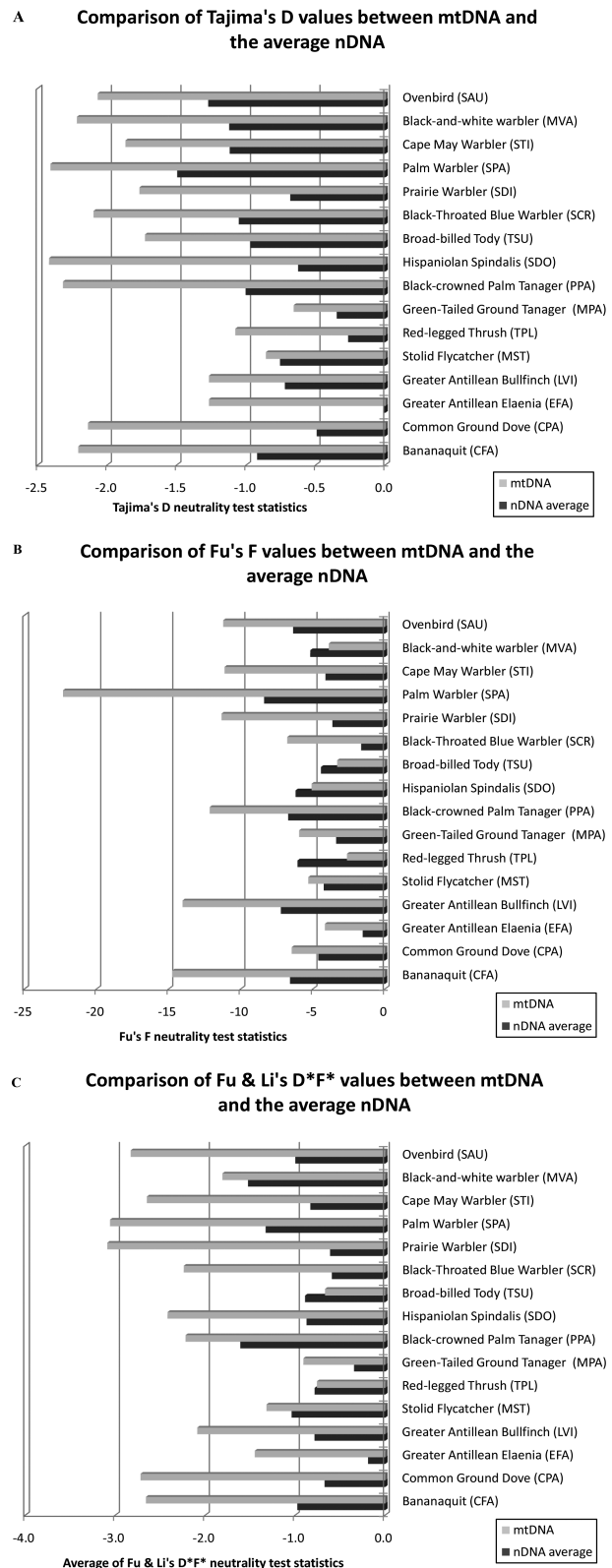
Illumina sequences were trimmed with `fastx_clipper` to remove adapters and clip poor quality bases. Filtered reads were sorted by index tags into BAM files and viewed using the program Integrative Genomic Viewer (IGV, Robinson et al. 2011). Gametic phase of polymorphic sites was determined using a sliding window of ~100 nt because this corresponds to the MiSeq read length following processing

(i.e. trimming primers and tags). Sequences from Illumina contigs were used to identify haplotypes in corroboration with the Sanger sequencing efforts and PHASE analyses. If polymorphic sites in a haplotype were separated by more than 100 nt, then only partial haplotype fragments could be determined. In such cases, we used the fragment with either the greater number of polymorphic sites or the more informative polymorphic sites for assigning “known” haplotypes in PHASE.

### Analyses of historical demography

We used Arlequin 3.1 (Excoffier et al. 2005) to compute standard population genetic statistics, including nucleotide diversity (average number of nucleotide differences between two sequences in a population,  $\pi$ ), haplotype diversity (measure of the distinctiveness of a haplotype in a population,  $h_d$ ), mean number of pairwise differences (MPD), number of haplotypes ( $h$ ), and number of polymorphic sites ( $s$ ). We also used Arlequin to quantify characteristics of haplotype mismatch distributions for each locus and species individually. This includes the raggedness index ( $r$ ), which quantifies the smoothness of the mismatch distributions (Harpending et al. 1993). Low values ( $< 0.05$ ) of  $r$  indicate the mismatch distribution is smooth and thus does not differ significantly from the null model of population expansion.

To test the hypothesis that migratory bird populations have recently expanded to a greater extent than non-migratory populations, we employed Arlequin and DnaSP (version 5.1, Rozas et al. 2003) to estimate neutrality test statistics. Tajima’s  $D$  (Tajima 1989) and Fu’s  $F$  (Fu 1997) were calculated in Arlequin. Fu’s  $F$  uses information from the distribution of haplotypes to detect population growth, whereas Tajima’s  $D$  uses the frequency of segregating sites to test for deviations from neutrality. DnaSP was used to determine Fu and Li’s  $D^*$  and  $F^*$  (Fu & Li 1993) statistics, which are similar to Tajima’s  $D$  in that they use the frequency of nucleotide substitutions to test for deviations from neutrality. Significantly negative values ( $P < 0.05$ ) of Tajima’s  $D$ , Fu’s  $F$ , and Fu and Li’s  $D^*$  and  $F^*$  indicate population expansion or positive selection. Fu’s  $F$  has been found to be the most powerful test statistic for detecting population growth and selection (Fu 1997, Ramos-Onsins & Rozas 2002, Ramírez-Soriano et al. 2008) except in the case of recombination, Fu and Li’s  $D^*$  and  $F^*$  are the most powerful at detecting background selection (Fu 1997). We used general linear models (SAS GLM procedure) to estimate the effects on  $r$ ,  $D$ , and  $F$  statistics of population status



**Fig. 1.** Comparison of nDNA (averaged across the six nuclear markers) and mtDNA test statistics for all species. Test statistics include Tajima’s  $D$  (A), Fu’s  $F$  (B), and Fu and Li’s  $D^*$  &  $F^*$  (C). As in Fahey et al. (2012), Fu and Li’s  $D^*$  and  $F^*$  were highly correlated and thus we referred to them as a unit ( $D^*F^*$ ) and took the average of the values for the figure. Note: these figures are not depicting whether the values were significant or not, just the overall values.



**Table 1.** The six nuclear loci used for this study with the PCR amplification details and references. See Kimball et al. (2009) for more details.

Locus	Primer name	[Mg <sup>++</sup> ]	Annealing temp°	Reference
<i>rho</i>	Rhod.3F Rhod.4R	1.5 mM	60°	Kimball et al. (2009)
<i>hmg2</i>	HMG17. 3F HMG17. 4R	1.5 mM	58°	Kimball et al. (2009)
<i>pcbd</i>	PCBD.2F/ PCBD.4R (external) PCBD.3F/ PCBD.3R (internal)	1.5 mM	64°	Kimball et al. (2009)
<i>mb</i>	MYO2 MYO3F	2.0 mM	52° or 50°	Slade et al. (1993) Heslewood et al. (1998)
<i>nat15</i>	NAT.4F NAT.5R	2.0 mM	62° or 58°	Kimball et al. (2009)
<i>irf2</i>	IRF2.2F IRF2.3R	2.0 mM	55.5°	Kimball et al. (2009)

**Table 2.** Illumina sequencing templates and the number of reads per tag (after trimming). Individuals in bolded italics (n = 3) had read depths < 100×, all others (n = 129) had read depths > 100×. CFA = *Coereba flaveola*, CPA = *Columbina passerina*, SCR = *Setophaga caerulea*, SDI = *Setophaga discolor*, SPA = *Setophaga palmarum*, STI = *Setophaga tigrina*, EFA = *Elaenia fallax*, LVI = *Loxia violacea*, MPA = *Microligea palustris*, MVA = *Mniotilta varia*, MST = *Myiarchus stolidus*, PPA = *Phaenicophilus palmarum*, SAU = *Seiurus aurocapilla*, SDO = *Spindalis dominicensis*, TPL = *Turdus plumbeus*, and TSU = *Todus subulatus*.

Tag	<i>rho</i>	<i>pcbd</i>	<i>nat15</i>	<i>mb</i>	<i>irf2</i>	<i>hmg2</i>	# of reads
1	CFA-DR_176	CFA-DR_130	CFA-DR_281	CFA-DR_438	CFA-DR_287	CFA-DR_750	164, 562
2	CPA-DR_484	CPA-DR_408	CPA-DR_408	<b>CPA-DR2_051</b>	CPA-DR_433	CPA-DR_413	194, 608
3	SCR-DR_978	SCR-DR_978	SCR-PUVI_15	SCR-PUVI_15	SCR-DR_032	SCR-DR_164	199, 986
4	SDI-DR_089	SDI-DR2_54	SDI-DR_064	SDI-DR_062	SDI-DR2_087	SDI-DR2_088	246, 472
5	SPA-DR_104	SPA-DR_106	SPA-DR_090	SPA-DR_116	SPA-DR_341	SPA-DR_096	202, 276
6	STI-DR_860	STI-DR_047	STI-DR_173	STI-DR_173	STI-DR2_153	STI-DR_947	302, 450
7	EFA-DR_999	EFA-DR2_042	EFA-DR_787	EFA-DR_982	EFA-DR_021	EFA-DR_982	249, 178
8	LVI-DR_238	LVI-DR_161	LVI-DR_330	LVI-DR_239	LVI-DR_425	LVI-DR_151	174, 258
9	MPA-DR2_194	MPA-DR_940	MPA-DR_813	MPA-DR2_134	MPA-PUVI_06	MPA-DR_476	377, 572
10	MST-DR_379	MST-DR2_240	MST-DR_870	MST-DR2_050	MST-DR_907	MST-DR_733	326, 256
11	MVA-DR_807	MVA-DR_277	MVA-DR2_032	MVA-DR_159	MVA-DR_781	MVA-DR_270	283, 578
12	PPA-DR_226	PPA-DR_181	PPA-DR_257	PPA-DR_045	PPA-DR_210	PPA-DR_431	612, 302
13	SAU-DR_819	SAU-DR_118	SAU-DR_313	SAU-DR2_343	SAU-DR2_305	SAU-DR2_305	208, 374
14	SDO-DR_081	SDO-DR2_309	SDO-DR_015	SDO-DR_039	SDO-DR2_319	SDO-DR2_009	221, 362
15	TPL-DR_295	TPL-DR_867	TPL-DR2_325	TPL-DR_727	TPL-DR_867	TPL-DR_763	292, 840
16	TSU-DR_480	TSU-DR_271	TSU-DR_820	TSU-DR2_103	TSU-DR_857	TSU-DR_838	234, 128
17	CFA-DR_729	SDI-DR_450	CPA-DR_421	CFA-DR_320	SPA-DR_378	CPA-DR_998	200, 220
18	<b>CPA-DR_408</b>	STI-DR_947	SPA-DR_116	<b>CPA-DR_421</b>	STI-DR_983	SDI-DR2_059	246, 342
19	LVI-DR_160	EFA-DR_123	STI-DR_896	LVI-DR_151	MPA-DR_888	SPA-DR_453	324, 062
20	MPA-DR2_145	LVI-DR_197	MPA-DR_752	MPA-DR2_119	MVA-DR_807	STI-DR_860	348, 138
21	MVA-DR_071	MST-DR_170	MST-DR_499	MST-DR2_004	SAU-DR_122	LVI-DR_288	182, 104
22	PPA-DR_249	MVA-DR_781	PPA-DR_470	SDO-DR_080	SDO-DR_968	MST-DR_496	241, 482

(endemic, resident, migrant), genome (nuclear versus mitochondrial), and individual genes.

## Results

We sequenced a mean of 18.3 (range 4-32) individuals per species (n = 16) across a mean of 5.4 nuclear genes (Table 3). We successfully cloned and sequenced 47

amplicons using conventional (dideoxy) approaches. We also sequenced 132 amplicons using the Illumina MiSeq (Table 2); the mean number of trimmed Illumina reads for each tag/sample was 265, 116 (range 164, 562-612, 302). For most of the genes, the mean coverage exceeded 1000× and 98 % of the amplicons (129 of 132) yielded > 100× coverage. This

**Table 3.** Genetic diversity and neutrality test statistics. C = category of species [R = regional residents (species that breed on the island but are found on other nearby islands), E = endemics (endemic to Hispaniola), and M = migrants (breed in continental North America)], length of the sequence (nt), sample size (n), the number of haplotypes (h), and the number of polymorphic sites (s). Genetic diversity and neutrality statistics for each species:  $\pi$  = nucleotide diversity, hd = haplotype diversity, MPD = mean number of pairwise differences between haplotypes, r = raggedness index. Italicized and bolded values are indicative of population expansion. CFA = *Coereba flaveola*, CPA = *Columbina passerina*, SCR = *Setophaga caerulea*, SDI = *Setophaga discolor*, SPA = *Setophaga palmarum*, STI = *Setophaga tigrina*, EFA = *Elaenia fallax*, LVI = *Loxigilla violacea*, MPA = *Microligea palustris*, MST = *Myiarchus stolidus*, MVA = *Mniotilta varia*, PPA = *Phaenicophilus palmarum*, SAU = *Seiurus auropigilla*, SDO = *Spindalis dominicensis*, TPL = *Turdus plumbeus*, and TSU = *Todus subulatus*. The ND2 mitochondrial data are from Fahey et al. (2012). \* =  $p < 0.05$ , \*\* =  $p < 0.01$ .

Species common name	C	Gene	nt	n	h	s	$\pi$	hd (SD)	MPD	r	Tajima's D	Fu's F	Fu & Li's D*	Fu & Li's F*
Banaquit (CFA)	R	HMG	373	7	13	21	0.0272	0.989 (0.03)	10.154	<b>0.024</b>	0.26	<b>-4.18**</b>	0.431	0.342
		IRF2	325	17	14	16	0.0090	0.929 (0.02)	2.930	0.069*	-0.83	<b>-4.91*</b>	-1.564	-1.723
		MB	605	26	16	16	0.0034	0.777 (0.05)	2.069	<b>0.028</b>	-1.28	<b>-7.95**</b>	-1.437	-1.458
		NAT	462	21	24	20	0.0069	0.949 (0.02)	3.201	<b>0.045</b>	-0.97	<b>-17.70**</b>	-0.506	-0.717
		PCBD	486	28	9	17	0.0074	0.849 (0.02)	3.609	0.086	-1.13	1.33	-1.019	-1.165
Common Ground Dove (CPA)	R	RHO	469	22	11	11	0.0025	0.620 (0.08)	1.180	0.069	<b>-1.59*</b>	<b>-6.12*</b>	-1.014	-1.146
		ND2	1041	58	16	17	0.0024	0.634 (0.07)	1.013	<b>0.042</b>	<b>-2.21**</b>	<b>-14.75**</b>	<b>-2.40*</b>	<b>-2.66*</b>
		HMG	348	25	20	16	0.0049	0.845 (0.04)	1.717	0.059	<b>-1.61*</b>	<b>-16.97**</b>	-1.697	-1.838
		IRF2	630	18	10	15	0.0040	0.683 (0.08)	2.533	0.069	-0.97	-1.67	-0.802	-0.882
		MB	746	5	4	4	0.0015	0.533 (0.18)	1.111	0.109	-0.822	-0.66	0.65	0.6
Greater Antillean Elaenia (EFA)	R	PCBD	572	19	14	13	0.0048	0.882 (0.03)	2.727	<b>0.049</b>	-0.37	<b>-4.86*</b>	-0.567	-0.576
		RHO	462	5	2	1	0.0012	0.533 (0.10)	0.533	0.289	1.30	1.03	N/A	N/A
		ND2	1040	35	11	15	0.0012	0.696 (0.08)	1.257	<b>0.049</b>	<b>-2.14**</b>	<b>-6.47**</b>	<b>-2.52*</b>	<b>-2.72*</b>
		HMG	679	16	11	10	0.0047	0.835 (0.05)	3.173	0.098	0.86	-1.86	0.0021	0.078
		IRF2	655	19	3	2	0.0009	0.558 (0.04)	0.615	0.204	0.56	0.65	N/A	N/A
Greater Antillean Bullfinch (LVI)	R	MB	718	12	5	4	0.0011	0.659 (0.08)	0.801	0.194	-0.69	-1.47	-1.094	-1.113
		NAT	871	13	13	11	0.0027	0.895 (0.04)	2.320	0.212**	-0.65	<b>-6.66**</b>	-0.814	-0.897
		PCBD	884	12	15	30	0.0106	0.946 (0.03)	9.351	<b>0.027</b>	1.25	-1.67	1.322	1.17
		RHO	472	4	3	8	0.0046	0.464 (0.20)	2.179	0.458	-1.42	1.76	N/A	N/A
		ND2	1041	20	8	7	0.0020	0.774 (0.08)	1.211	0.111	-1.27	<b>-4.17**</b>	-1.36	-1.45
Greater Antillean Bullfinch (LVI)	R	HMG	201	13	16	14	0.0147	0.957 (0.02)	2.948	<b>0.029</b>	-1.23	<b>-9.60**</b>	-1.189	-1.391
		IRF2	652	31	11	10	0.0025	0.797 (0.04)	1.646	0.068	-0.62	-3.28	-1.043	-1.27
		MB	685	18	8	10	0.0014	0.578 (0.09)	0.937	0.084	<b>-1.54*</b>	<b>-3.86**</b>	-1.36	-1.455
		NAT	175	19	18	14	0.0154	0.926 (0.02)	2.694	<b>0.030</b>	-0.61	<b>-10.35**</b>	-0.587	-0.736
		PCBD	259	30	20	14	0.0111	0.919 (0.02)	2.885	<b>0.036</b>	-0.12	<b>-9.09**</b>	0.778	0.529
Stolid Flycatcher (MST)	R	RHO	413	29	22	22	0.0107	0.816 (0.05)	4.417	<b>0.033</b>	-0.22	<b>-7.17*</b>	-0.428	-0.396
		ND2	1041	59	28	36	0.0053	0.958 (0.01)	4.807	<b>0.036</b>	-1.27	<b>-14.04**</b>	-1.9	-2.09
		HMG	663	22	18	22	0.0054	0.890 (0.04)	3.598	<b>0.012</b>	-0.77	<b>-6.50**</b>	-1.748	-1.754
		IRF2	565	19	7	6	0.0010	0.549 (0.09)	0.636	0.149	-1.50	<b>-4.10*</b>	-1.399	-1.49
		MB	692	9	5	5	0.0014	0.601 (0.11)	0.948	<b>0.046</b>	-1.10	-1.39	-1.124	-1.156
Stolid Flycatcher (MST)	R	NAT	665	19	18	26	0.0089	0.929 (0.02)	5.940	<b>0.010</b>	0.21	-3.65	-0.04	-0.003
		PCBD	687	17	18	17	0.0054	0.927 (0.03)	3.684	<b>0.016</b>	-0.38	<b>-8.23**</b>	-0.814	-0.801
		RHO	502	20	4	3	0.0008	0.354 (0.09)	0.379	0.193	-1.01	-1.62	N/A	N/A
		ND2	1040	31	11	9	0.0020	0.893 (0.03)	1.626	0.132	-0.86	<b>-5.32**</b>	-1.1	-1.32

	HMG	334	20	18	15	0.0088	0.935 (0.02)	2.951	<b>0.039</b>	-0.46	-9.01**	-1.282	-1.388
	IRF2	390	14	12	8	0.0044	0.868 (0.04)	1.704	0.104	-0.53	-6.88**	-0.053	-0.177
	MB	596	24	14	8	0.0035	0.879 (0.03)	2.067	0.051	0.40	-5.79**	0.022	-0.057
Red-legged	NAT	622	17	17	17	0.0055	0.900 (0.04)	3.433	<b>0.021</b>	-0.84	-7.46**	-1.053	-1.163
Thrush (TPL)	PCBD	717	23	16	19	0.0043	0.766 (0.07)	3.067	<b>0.034</b>	-0.78	-5.31*	-1.141	-1.105
	RHO	402	19	10	8	0.0057	0.850 (0.03)	2.300	<b>0.032</b>	0.60	-1.94	-0.834	-0.821
	ND2	1041	32	10	13	0.0034	0.851 (0.04)	2.159	<b>0.026</b>	-1.08	-2.64	-0.56	-0.76
	HMG	253	7	11	20	0.0218	0.956 (0.05)	5.516	0.067	-0.60	-3.46*	-0.282	-0.415
	IRF2	503	32	17	31	0.0126	0.895 (0.02)	6.356	<b>0.023</b>	-0.23	-0.43	-1.425	-1.423
Green-Tailed	MB	527	17	5	5	0.0012	0.367 (0.10)	0.656	0.240	-1.23	-1.62	0	0
Ground Tanager	NAT	187	18	17	21	0.0198	0.897 (0.03)	3.700	<b>0.017</b>	-0.83	-6.41**	-0.764	-0.894
(MPA)	PCBD												
	RHO	168	20	17	13	0.0241	0.915 (0.02)	4.051	<b>0.033</b>	1.14	-5.04*	1.073	0.999
	ND2	1040	58	12	8	0.0013	0.809 (0.04)	1.301	0.116	-0.66	-5.96**	-0.75	-0.91
	HMG												
	IRF2	659	24	12	20	0.0039	0.845 (0.04)	2.581	<b>0.031</b>	-1.22	-2.48	-1.772	-1.873
Black-crowned	MB	566	24	9	9	0.0045	0.673 (0.07)	2.559	0.159	0.74	-0.32	-0.107	-0.041
Palm Tanager	NAT	574	18	14	20	0.0047	0.852 (0.05)	2.697	<b>0.027</b>	-1.62*	-5.19*	-2.247	-2.393
(PPA)	PCBD	454	22	27	34	0.0083	0.965 (0.01)	3.757	0.056	-1.27	-20.36**	-1.554	-1.735
	RHO	486	16	10	11	0.0026	0.756 (0.06)	1.284	0.076	-1.66*	-5.22*	-1.88	-2.01
	ND2	1041	56	12	13	0.0023	0.412 (0.08)	0.602	0.155	-2.32**	-12.16**	-1.99	-2.22
	HMG	171	22	9	9	0.0100	0.738 (0.05)	1.715	0.081	-0.26	-1.99	-0.731	-0.895
	IRF2	663	19	18	13	0.0048	0.910 (0.03)	3.178	<b>0.026</b>	0.09	-8.69**	0.37	0.317
	MB	749	28	13	10	0.0020	0.768 (0.05)	1.494	<b>0.039</b>	-0.88	-6.37**	-0.617	-0.734
Hispaniolan	NAT	306	21	18	21	0.0107	0.933 (0.02)	3.268	<b>0.033</b>	-0.71	-7.67**	-1.403	-1.52
Spindalis (SDO)	PCBD	276	25	15	11	0.0092	0.909 (0.02)	2.545	<b>0.040</b>	0.11	-5.24*	-0.386	-0.507
	RHO	495	26	9	10	0.0011	0.404 (0.08)	0.560	0.139	-2.12**	-7.25**	-1.73	-1.91
	ND2	1041	32	13	15	0.0050	0.776 (0.08)	2.488	<b>0.041</b>	-2.42**	-5.09**	-2.23*	-2.42*
	HMG	284	13	13	18	0.0110	0.871 (0.06)	3.129	<b>0.034</b>	-1.40	-4.72*	-1.199	-1.431
	IRF2	418	27	16	17	0.0069	0.804 (0.05)	2.871	0.058	-0.71	-5.07*	-0.053	-0.177
	MB	735	24	6	6	0.0008	0.305 (0.09)	0.560	0.357	-1.50*	-2.92*	-0.416	-0.43
Broad-billed	NAT	424	10	11	20	0.0180	0.916 (0.04)	7.632	<b>0.046</b>	0.17	-0.25	-0.661	-0.752
Tody (TSU)	PCBD	970	27	14	15	0.0016	0.662 (0.07)	1.558	0.061	-1.58*	-7.51**	-1.342	-1.499
	RHO	355	10	10	7	0.0042	0.795 (0.09)	1.489	0.080	-0.81	-6.19**	-0.913	-1.058
	ND2	1041	29	4	3	0.0014	0.200 (0.10)	0.207	0.404	-1.73*	-3.32**	-0.75	-0.67
	HMG	449	7	10	15	0.0094	0.945 (0.05)	4.209	<b>0.040</b>	-1.01	-3.16*	-0.721	-0.874
	IRF2												
Black-Throated	MB	717	10	6	8	0.0019	0.637 (0.12)	1.384	<b>0.049</b>	-0.37	-1.30	0.0713	0.093
Blue Warbler	NAT	807	12	5	13	0.0016	0.312 (0.12)	1.312	0.334	-2.01**	-0.18	-0.289	-0.314
(SCR)	PCBD												
	RHO	467	14	8	9	0.0037	0.762 (0.07)	1.706	0.083	-0.83	-1.98	-1.136	-1.28
	ND2	1041	20	9	9	0.0019	0.705 (0.11)	0.989	0.124	-2.10**	-6.78**	-2.06*	-2.24*

	HMG	391	15	21	27	0.0164	0.977 (0.01)	6.405	<b>0.011</b>	-0.22	<b>-8.82**</b>	-0.305	-0.372
	IRF2	613	22	9	16	0.0062	0.600 (0.08)	3.803	0.176	0.11	1.03	-0.088	-0.235
	MB	744	32	4	3	0.0002	0.122 (0.06)	0.124	0.590	<b>-1.57*</b>	<b>-4.05*</b>	-0.754	-0.675
Prairie Warbler (SDI)	NAT	497	24	12	15	0.0044	0.864 (0.03)	2.192	<b>0.032</b>	-0.97	-3.32	-1.167	-1.287
	PCBD	404	8	12	17	0.0101	0.967 (0.03)	4.067	<b>0.030</b>	-0.72	<b>-5.23**</b>	-1.109	-1.257
	RHO	469	25	5	4	0.0013	0.534 (0.06)	0.602	0.167	-0.74	-1.49	0.168	0.15
	ND2	1041	57	18	20	0.0018	0.840 (0.04)	1.877	<b>0.036</b>	<b>-1.77**</b>	<b>-11.34**</b>	<b>-2.95*</b>	<b>-3.09*</b>
	HMG	254	18	21	17	0.0129	0.940 (0.03)	3.286	<b>0.040</b>	-0.92	<b>-13.76**</b>	-0.767	-0.972
	IRF2	379	25	15	20	0.0060	0.785 (0.05)	2.282	<b>0.023</b>	<b>-1.56*</b>	<b>-6.06*</b>	-1.043	-1.27
	MB	759	30	5	5	0.0003	0.190 (0.07)	0.230	0.447	<b>-1.84**</b>	<b>-4.18**</b>	-1.124	-1.156
Palm Warbler (SPA)	NAT	320	20	19	19	0.0070	0.919 (0.03)	2.251	<b>0.049</b>	<b>-1.50*</b>	<b>-13.62**</b>	-1.991	-2.093
	PCBD												
	RHO	469	25	5	4	0.0005	0.225 (0.08)	0.235	0.355	<b>-1.67*</b>	<b>-4.28**</b>	-1.124	-1.156
	ND2	1041	58	21	24	0.0026	0.696 (0.07)	1.257	<b>0.041</b>	<b>-2.41**</b>	<b>-22.33**</b>	<b>-2.78*</b>	<b>-3.06*</b>
	HMG	313	20	17	18	0.0105	0.951 (0.01)	3.291	<b>0.023</b>	-0.64	<b>-6.68**</b>	-1.055	-1.183
	IRF2	627	25	18	26	0.0039	0.758 (0.06)	2.454	<b>0.012</b>	<b>-1.90**</b>	<b>-9.29*</b>	-1.616	-1.823
	MB	757	29	4	3	0.0004	0.276 (0.07)	0.285	0.276	-1.12	<b>-2.04*</b>	-0.754	-0.675
Cape May Warbler (STI)	NAT	703	18	9	10	0.0033	0.773 (0.05)	2.348	0.061	-0.32	-1.17	-0.498	-0.58
	PCBD	363	10	13	23	0.0128	0.932 (0.04)	4.637	0.079	-1.08	<b>-4.16*</b>	-0.581	-0.802
	RHO	469	25	7	12	0.0025	0.540 (0.08)	1.173	0.069	<b>-1.66*</b>	-1.44	0.08	0.071
	ND2	1041	43	15	15	0.0023	0.846 (0.04)	1.402	<b>0.050</b>	<b>-1.87**</b>	<b>-11.13**</b>	<b>-2.46*</b>	<b>-2.65*</b>
	HMG												
	IRF2	557	17	16	25	0.0090	0.863 (0.05)	4.993	<b>0.019</b>	-0.55	-3.59	-1.425	-1.423
	MB	744	17	6	5	0.0010	0.613 (0.07)	0.742	0.165	-1.04	<b>-2.41*</b>	-0.791	-0.851
Black-and-white warbler (MVA)	NAT	262	8	12	22	0.0167	0.942 (0.05)	4.383	0.055	-1.63	<b>-4.84**</b>	-2.098	-2.219
	PCBD	384	10	15	24	0.0127	0.963 (0.03)	4.879	<b>0.022</b>	-1.01	<b>-6.75**</b>	-1.516	-1.576
	RHO	491	16	15	16	0.0047	0.901 (0.04)	2.329	<b>0.038</b>	-1.39	<b>-8.32**</b>	-1.362	-1.56
	ND2	1041	20	8	12	0.0031	0.590 (0.13)	1.289	0.126	<b>-2.22**</b>	<b>-3.90**</b>	-1.65	-1.81
	HMG	145	11	17	21	0.0240	0.900 (0.03)	3.483	<b>0.029</b>	-1.02	<b>-7.34**</b>	-1.176	-1.339
	IRF2	471	7	5	7	0.0029	0.506 (0.16)	1.363	0.346	-1.41	-0.86	0.083	0.086
	MB	679	22	9	8	0.0013	0.554 (0.09)	0.885	0.061	-1.45	<b>-5.04**</b>	-1.164	-1.262
Ovenbird (SAU)	NAT												
	PCBD												
	RHO	351	17	17	15	0.0060	0.900 (0.03)	2.094	0.063	-1.22	<b>-12.25**</b>	-1.417	-1.483
	ND2	1040	42	23	47	0.0043	0.899 (0.04)	4.520	0.100*	<b>-2.07**</b>	<b>-11.22**</b>	<b>-2.57*</b>	<b>-2.83*</b>



depth of coverage helped us decipher the gametic phase of complex haplotypes. The mean and median PHASE probabilities for each gene and species are listed in Table 4.

Across species, the intron *hmgn2* (mean length of 347 nt) had a mean of 15.4 haplotypes and 17.4 polymorphic sites; *irf2* (540 nt) averaged 12.2 haplotypes and 15.5 polymorphic sites; *mb* (689 nt) had the smallest number of haplotypes (7.4) and polymorphic sites (6.5); *nat15* (528 nt) had a mean of 14.1 haplotypes and 17.5 polymorphic sites; *pcbd* (538 nt) had the highest mean number of haplotypes (15.7) and polymorphic sites (19.5); and finally, *rho* (434 nt) had a mean of 9.7 haplotypes and 9.6 polymorphic sites. The mean haplotype diversity for each gene across the 16 species ranged from 0.533-0.909, with *mb* having the lowest value and *hmgn2* the highest. The average nucleotide diversity for each gene over all species ranged from 0.0016-0.013, again with *mb* having the lowest value and *hmgn2* as the highest (Table 3).

Nearly half (41 of 87) of the raggedness index (*r*) values for each species and gene were < 0.05, consistent with population expansion (Table 3). Nearly all *F* values (83 of 87) were negative, and 57 of 87 (66 %) were significantly so. Most Tajima's *D* values were also negative (78 out of 87), but only 16 of 87 values (18.3 %) were significantly negative

and thus consistent with population expansion. As in Fahey et al. (2012), Fu and Li's *D\** and *F\** were highly correlated and thus for simplicity we refer to them as a unit (*D\*F\**). We computed *D\*F\** for only 74 genes instead of 87 because at least four haplotypes are required for this assessment. No *D\*F\** values were significantly negative, and 12 out of 74 (16.2 %) had positive values (though none significantly positive). This was surprising, as Fahey et al. (2012) found that all mtDNA values of Tajima's *D* and *D\*F\** were negative and approximately half were significantly negative. When comparing nDNA values (averaged across all loci) against the mtDNA results for any given test statistic, 87.5 % of the mtDNA statistics were less than (more negative) than the nDNA averages (Fig. 1A-C). Lower figures in these test statistics are more likely to be significant, thus with the mtDNA statistics lower than the average nDNA statistics it is not surprising that we did not see as many significant values for the nDNA test statistics.

Our mtDNA dataset suggested that migratory species in our sample exhibited more pronounced population expansion than did resident or endemic species (Fahey et al. 2012). Mismatch distributions based on nuclear introns revealed some evidence of population expansion in three resident (LVI, MST, and TPL), two endemic (MPA and SDO), and two migrant (SPA, and MVA) species. Each of these

**Table 4.** The mean and median PHASE reconstruction probabilities for each gene and species. Bolded entries have four sites with probabilities < 0.60, italicized entries are those with 1-3 sites with probabilities < 0.60 and a mean or median not above 0.90.

	Means/Median					
	<i>nat15</i>	<i>mb</i>	<i>hmgn2</i>	<i>rho</i>	<i>pcbd</i>	<i>irf2</i>
CFA	<i>0.73/0.70</i>	0.87/1.0	<i>0.84/0.91</i>	<i>0.89/1.0</i>	0.96/1.0	0.95/0.99
CPA	n/a	0.87/0.90	0.80/1.0	1.0/1.0	0.92/0.99	0.97/1.0
SCR	0.83/0.87	1.0/1.0	<i>0.88/0.98</i>	0.99/1.0	n/a	n/a
SDI	<b>0.91/0.98</b>	1.0/1.0	<i>0.87/0.93</i>	1.0/1.0	<i>0.78/0.78</i>	0.99/1.0
SPA	<i>0.87/1.0</i>	1.0/1.0	<i>0.78/0.76</i>	1.0/1.0	n/a	0.93/1.0
STI	0.94/1.0	1.0/1.0	0.88/0.96	0.97/1.0	<i>0.87/0.89</i>	0.89/0.99
EFA	<i>0.79/0.82</i>	1.0/1.0	0.93/0.98	0.83/0.83	0.92/0.99	0.97/1.0
LVI	0.92/1.0	0.93/1.0	0.75/0.75	0.91/0.98	<i>0.86/0.89</i>	0.97/1.0
MPA	<i>0.85/0.91</i>	0.92/1.0	<i>0.80/0.78</i>	0.92/0.99	n/a	0.98/1.0
MST	0.93/0.99	0.85/1.0	0.94/0.99	1.0/1.0	<i>0.86/0.95</i>	1.0/1.0
MVA	<i>0.78/0.75</i>	0.96/1.0	n/a	0.94/1.0	<i>0.80/0.77</i>	<i>0.89/1.0</i>
PPA	0.90/1.0	1.0/1.0	n/a	0.89/1.0	0.71/0.69	0.95/1.0
SAU	n/a	0.92/1.0	<i>0.76/0.76</i>	<i>0.75/0.75</i>	n/a	0.83/0.86
SDO	<i>0.88/0.92</i>	0.93/1.0	0.95/1.0	<i>0.89/1.0</i>	0.93/0.96	0.92/0.97
TPL	<i>0.88/0.95</i>	0.90/0.94	0.92/0.97	0.96/0.99	0.92/0.99	<i>0.84/1.0</i>
TSU	0.96/1.0	0.91/0.97	0.91/1.0	0.92/1.0	<b>0.92/1.0</b>	0.94/1.0

**Table 5.** General linear model results for differences between genomes, genes, and population status with respect to several indicators of historic demographic change. Significant F statistics and Least Squares Means (LSMeans) are indicated in bold type. Superscript letters indicate significant differences between mean values according to Duncan's multiple range test.

	r	D	Fu's F	Fu and Li's D	Fu and Li's F
Effect (F)					
Genome (1 df)	0.02	<b>26.19***</b>	<b>9.68**</b>	<b>28.4***</b>	<b>31.3***</b>
Gene (5 df)	<b>4.06*</b>	0.40	2.10	0.40	0.53
Status (2 df)	0.69	<b>6.17*</b>	0.16	1.41	1.41
LSMeans					
Endemic	0.09 <sup>a</sup>	<b>-1.23<sup>a</sup></b>	-6.80 <sup>a</sup>	-1.31 <sup>a</sup>	-1.45 <sup>a</sup>
Migrant	0.11 <sup>a</sup>	<b>-1.58<sup>b</sup></b>	-7.29 <sup>a</sup>	-1.50 <sup>a</sup>	-1.64 <sup>a</sup>
Resident	0.09 <sup>a</sup>	<b>-1.02<sup>a</sup></b>	-6.79 <sup>a</sup>	-1.21 <sup>a</sup>	-1.34 <sup>a</sup>
Mitochondrial	0.10 <sup>a</sup>	<b>-1.77<sup>a</sup></b>	<b>-8.77<sup>a</sup></b>	<b>-1.87<sup>a</sup></b>	<b>-2.05<sup>a</sup></b>
Nuclear	0.09 <sup>a</sup>	<b>-0.79<sup>b</sup></b>	<b>-5.14<sup>b</sup></b>	<b>-0.80<sup>b</sup></b>	<b>-0.90<sup>b</sup></b>
*** < 0.0001	** < 0.001	* < 0.01			

species had 4-6 independent *r* values < 0.05, but the differences between species in each of the population status categories were not significant. That said, mismatch distribution statistics are not especially powerful (Ramos-Onsins & Rozas 2002) and thus we concentrate our discussion on the neutrality statistics. For any given nDNA locus, no species had significantly negative values of all neutrality test statistics (i.e. Tajima's D, Fu's F, Fu and Li's D\*, and Fu and Li's F\*). As previously indicated, none of the nuclear genes had significantly negative values of D\*F\*. Furthermore, no species had more than three significantly negative Tajima's D values, and only four species (SPA, STI, PPA, and TSU) had two or three significantly negative Tajima's D values. Significantly negative Fu's F values at four or more genes were observed in three resident (CFA, LVI, and TPL), three endemic (PPA, SDO, and TSU), and four migrant (SPA, STI, MVA, and SAU) species. Thus 10 out of 16 species (with equal representation in migrant and non-migrants) showed evidence of population expansion in most of their Fu's F statistics, one of the most powerful tests for detecting population expansion (Fu 1997, Ramos-Onsins & Rozas 2002, Ramírez-Soriano et al. 2008). If we compare these nuclear data to our published mtDNA data, we see concordant mtDNA and nDNA evidence for historical population expansion in one resident (CFA), three endemics (PPA, SDO, and TSU), and four migrants (SPA, STI, MVA and SAU). In five cases, however, there were discordances between the nDNA and mtDNA results. Nuclear genes of CPA, SCR, and SDI showed little sign of population growth and thus do not support the mtDNA data. Similarly,

LVI and TPL also had discordances among nuclear and mitochondrial genomes: the nuclear intron sequences generally supported population expansion, yet there was no such evidence in the mtDNA sequences. Using mtDNA and nDNA, a greater proportion of migratory species (4/6) exhibited evidence of population expansion than the non-migratory species (4/10), but the difference was not significant (G-test: *P* = 0.3). The general linear model revealed different patterns of variation for each of the indices of demographic change (Table 5). The raggedness index (*r*) differed significantly among genes, but did not differ between genomes or with respect to population status (resident, endemic, migratory). The remaining statistics (D, Fu's F, and Fu and Li's D\* and F\*) differed approximately two-fold between the nuclear and mitochondrial genomes, but they did not differ with respect to population status or among nuclear genes within genome. Thus, although some species show genetic evidence of recent population expansion, this cannot be tied to commonalities in distribution or migratory status.

## Discussion

The distribution of pairwise nucleotide differences between gene sequences can highlight the evolutionary demography of a species, revealing population growth trends (Rogers & Harpending 1992). Fluctuations in population size reflect the idiosyncratic life-history characteristics of a given population, but concordant demographic fluctuations among species could also reflect long-term trends in environmental conditions. Many studies have considered the historical

demography of a single species with a single marker (usually mtDNA, Merilä et al. 1997, Zink et al. 2006, Hawley et al. 2008, Lerner et al. 2009, Norgate et al. 2009, Brace et al. 2012). More recently, authors have used multiple markers to study individual species (Firestone et al. 1999, Tchaicka et al. 2007, Bos et al. 2008, Reding et al. 2010, Nietlisbach et al. 2012). With technological advances in sequencing, more studies in the future will be able to analyze more than one marker to study multiple species (Qu et al. 2010, Sonsthagen et al. 2012), providing the most comprehensive reconstruction of evolutionary demography.

We investigated the comparative historical demography of Hispaniolian birds ( $n = 16$  species). We surveyed molecular variation at six nuclear genes to determine whether nDNA would corroborate our mtDNA results (Fahey et al. 2012). Overall, the average haplotype diversities of mtDNA and nDNA were similar (mtDNA  $h_d = 0.724$ , nDNA  $h_d = 0.755$ ) but average nucleotide diversity was lower for mtDNA (mtDNA  $\pi = 0.0026$ , nDNA  $\pi = 0.0067$ ). Although haplotype diversity for nDNA and mtDNA might be similarly impacted by population bottlenecks, nucleotide diversity might be lower for mtDNA because contemporary nDNA haplotypes represent older allelic lineages with deeper coalescences that have had more time to accumulate nucleotide substitutions. We note that the three-fold difference in nucleotide diversity between mtDNA and nDNA observed in our empirical data is reasonably consistent with theoretical expectations which predict a 4-fold difference in  $N_e$  between mtDNA and nDNA (Hare 2001).

Haplotype and nucleotide diversities did not differ between migrants, regional residents, and endemics (though endemics did have the lowest values in the mtDNA data). We did not find concordance between nDNA and mtDNA for each species, and we found relatively little nDNA support for recent population expansion in migratory species. Instead, (non-significant) positive values of Tajima's  $D$  and Fu and Li's  $D^*$  and  $F^*$  were scattered among 14 of the 16 species in our matrix of species/loci/test statistics, which could point towards stable population sizes. All other neutrality test statistics were negative, but few (~20 %) nDNA values were statistically significant. In contrast, all of the mtDNA neutrality test statistics were negative and > 50 % were significantly negative. Fu's  $F$  is generally considered the most powerful such statistic, so we identified species where the majority of the Fu's  $F$  statistics were significantly negative across nDNA loci. Using this approach, values for

three residents (CFA, LVI, and TPL), three endemics (PPA, SDO, and TSU), and four migrants (SPA, STI, MVA, and SAU) were consistent with population expansion. Eight of these ten species (CFA, SPA, STI, MVA, PPA, SAU, SDO, and TSU) exhibited significantly negative neutrality test statistics in the mtDNA, but these include both migratory and non-migratory species in similar proportions.

Disparities among mtDNA and nDNA within a species are not surprising as nuclear markers are often idiosyncratic due to selection or drift (Toews & Brelsford 2012). For example, selective sweeps of beneficial mutations will carry nearby linked loci to fixation (Galtier et al. 2009, Karl et al. 2012). This reduces genetic variation and causes demographic inferences to be based upon the last selective sweep rather than the most common recent ancestor. Fu and Li's  $D^*F^*$  statistics are more powerful at detecting background selection than Tajima's  $D$  (Fu 1997), and thus significant values of  $D^*F^*$  in conjunction with non-significant values of  $D$  suggests selection at a given locus. None of the genes in our datasets exhibited this pattern, and thus we assume that background selection has had relatively little impact on the standing levels of genetic variation at these loci.

Selection aside, the mtDNA data generally supported expanding migratory populations whereas the nDNA data did not. We see three possible explanations for this incongruity: technical artifacts, incomplete lineage sorting, and statistical issues. One possible source of technical noise is haplotype inference. Our nuclear sequences were highly polymorphic, but such variability is a double-edged sword in that it provides increased power for inferring past selection and population change, but makes haplotype phase inference more difficult. We inferred phase computationally from direct sequencing of PCR products, from TA cloning, and from Illumina sequencing. We used PHASE probability thresholds to trim ambiguous regions; this eliminates some polymorphic sites, but it is preferable to removing the unresolved haplotype because removal can bias population genetic statistics. For example, removing unresolved haplotypes can result in overestimation of neutrality test statistics (because this removes rare alleles and the excess of rare alleles leads to more negative values, which is indicative of population expansion) or underestimation of  $\Theta_w$  and  $\pi$  (Garrick et al. 2010). Trimming can also lead to overestimation of neutrality test statistics if the deleted polymorphism(s) is from a novel haplotype and incorrect haplotype

reconstructions could cause overestimation of neutrality test statistics. Thus it is possible that the relatively few nDNA neutrality test statistics with significantly negative values could be an artifact of genotype reconstruction. Haplotype reconstruction was more difficult for some genes (*pcbd*, *nat15*, and *hmg2*), but the number of significantly negative Fu's *F* values did not differ in these genes compared to the other three genes. Therefore, there was no obvious difference between genes that we might expect to have been biased due to reconstruction errors and genes that were more confidently deduced.

The lack of a clear signal in our data might also reflect biological factors such as incomplete lineage sorting and the longer coalescence times associated with diploid nDNA when compared to mtDNA genes (Hare & Avise 1996). The maternal inheritance and haploid nature of mtDNA means that its effective population size is  $\frac{1}{4}$  that of nDNA, which means that coalescence of mitochondrial markers requires  $\frac{1}{4}$  of the time for nDNA; thus, mtDNA is more informative of the recent past (Moore 1995, Zink & Barrowclough 2008). The  $4\times$  larger effective population size of the nDNA makes it more informative in the distant past, and nDNA gene genealogies are expected to be unresolved over intermediate and very recent timescales (Zink & Barrowclough 2008). Thus, the presumptive population expansions of migratory species that have occurred since the last glacial maximum may be so recent that insufficient time has elapsed for coalescence of nDNA genes. The average haplotype diversities for the mitochondrial and nuclear data were similar and moderately high, and therefore did not indicate recent bottlenecks (average overall  $hd$ : mtDNA = 0.72, nDNA = 0.75).

Finally, our results could also reflect the inherent noise associated with neutrality test statistics. This may be a plausible explanation, as other studies have also reported inconsistencies between nDNA and mtDNA data in terms of neutrality test statistics (Bensch et al. 2006, Eytan & Hellberg 2010, D'Horta et al. 2011, Lim & Sheldon 2011). If commonly employed demographic metrics are sound in the face of complex evolutionary realities, we would expect

that a large enough suite of markers should overcome the inherent statistical noise and eventually converge. The six nuclear markers we evaluated no doubt paint an incomplete picture of the whole genome, and underscore the idea that the most complete studies of demographic history will evaluate many more markers. Fortunately, advances in genome sequencing and associated technologies suggest the number of markers will soon cease to be a significant limitation.

## Conclusion

The aim of this study was to corroborate inferences from our mtDNA analysis with data from independent nuclear genes, in part to distinguish genome-wide effects (e.g. genetic drift) from selection on individual genes/genomes. Our data reveal very limited support for our hypothesis that migratory species would exhibit more pronounced indicators of population growth, and the nDNA data were more equivocal than the mtDNA data. The disparities among genomes/loci may reflect systematic (technical), stochastic (statistical), or biological (incomplete lineage sorting) noise. Ultimately, the disparities between mtDNA and nDNA probably reflect a) the deeper coalescence of nDNA and its inability to accurately capture recent demographic history and b) the noise inherent to neutrality test statistics. More empirical data and/or more extensive computer simulations are required to distinguish among these scenarios.

## Acknowledgements

Samples were collected in the field by Steve Latta at the National Aviary, funded by NSF Grants DEB-0089226 and DEB-0542390 to RER. Permission to work in the Sierra de Bahoruco was provided by the Dirección Nacional de Parques and the Departamento de Vida Silvestre, República Dominicana. Field assistance in the Dominican Republic was provided by Danilo Mejía, Vinicio Mejía, Mamonides Heredia, Miguel Angel Landestoy, Marie Abbott, and Elvis Cuevas Mendoza. Chris Rimmer and Kent McFarland of the Vermont Center for Ecostudies provided blood samples from montane broadleaf forest sites in Hispaniola. We also thank B. Dunning, K. Nichols, and members of the DeWoody lab group for their insights and comments on the paper. Finally, we thank the Provost's Office at Purdue University for funding through the University Faculty Scholar program (JAD) and the Curators of the University of Missouri and the Alexander von Humboldt Foundation for financial support (RER).

## Literature

- Avise J.C. 2000: *Phylogeography*. Harvard University Press, Cambridge, MA.
- Bensch S., Irwin D.E., Irwin J.H., Kvist L. & Åkesson S. 2006: Conflicting patterns of mitochondrial and nuclear DNA diversity in *Phylloscopus* warblers. *Mol. Ecol.* 15: 161–171.
- Bonatti E. & Gartner S., Jr. 1973: Caribbean climate during Pleistocene ice ages. *Nature* 244: 563–565.
- Bos D.H., Gopurenko D., Williams R.N. & DeWoody J.A. 2008: Inferring population history and demography using microsatellites, mitochondrial DNA, and major histocompatibility complex (MHC) genes. *Evolution* 62: 1458–1468.



- Brace S., Barnes I., Powell A., Pearson R., Woolaver L.G., Thomas M.G. & Turvey S.T. 2012: Population history of the Hispaniolan hutia *Plagiodontia aedium* (Rodentia: Capromyidae): testing the model of ancient differentiation on a geographically complex Caribbean island. *Mol. Ecol.* 21: 2239–2253.
- Clark P.U. & Mix A.C. 2002: Ice sheets and sea level of the Last Glacial Maximum. *Quat. Sci. Rev.* 21: 1–7.
- Curtis J.H., Brenner M. & Hodell D.A. 2001: Climate change in the circum-Caribbean (Late Pleistocene to Present) and implications for regional biogeography. In: Woods C.A. & Sergile F.E. (eds.), *Biogeography of the West Indies. Patterns and perspectives*. CRC Press, Boca Raton, FL: 35–54.
- D’Horta F.M., Cabanne G.S., Meyer D. & Miyaki C.Y. 2011: The genetic effects of Late Quaternary climatic changes over a tropical latitudinal gradient: diversification of an Atlantic forest passerine. *Mol. Ecol.* 20: 1923–1935.
- Derjushva S., Kurganova A., Habermann F. & Gaginskaya E. 2004: High chromosome conservation detected by comparative chromosome painting in chicken, pigeon and passerine birds. *Chromosome Res.* 12: 715–723.
- Donnelly T.W. 1989: Geologic history of the Caribbean and Central America. In: Bally A.W. & Palmer A.R. (eds.), *The geology of North America: an overview*. Geological Society of America, Boulder, CO: 299–321.
- Dupuis J.R., Roe A.D. & Sperling F.A.H. 2012: Multi-locus species delimitation in closely related animals and fungi: one marker is not enough. *Mol. Ecol.* 21: 4422–4436.
- Excoffier L., Laval L.G. & Schneider S. 2005: An integrated software package for population genetics data analysis, Arlequin ver. 3.0. *Evol. Bioinform.* 1: 47–50.
- Eytan R.I. & Hellberg M.E. 2010: Nuclear and mitochondrial sequence data reveal and conceal different demographic histories and population genetic processes in Caribbean reef fishes. *Evolution* 64: 3380–3397.
- Fahey A., Latta S.C., Ricklefs R. & DeWoody J.A. 2012: Comparative historical demography of migratory and nonmigratory birds from the Caribbean island of Hispaniola. *Evol. Biol.* 39: 400–414.
- Firestone K.B., Elphinstone M.S., Sherwin W.B. & Houlden B.A. 1999: Phylogeographical population structure of the tiger quolls *Dasyurus maculatus* (Dasyuridae: Marsupialia), an endangered carnivorous marsupial. *Mol. Ecol.* 8: 1613–1625.
- Fleming K., Johnston P., Zwart D., Yokoyama Y., Lambeck K. & Chappell J. 1998: Refining the eustatic sea-level curve since the Last Glacial Maximum using far- and intermediate-field sites. *Earth Planet. Sci. Lett.* 163: 327–342.
- Fu Y. 1997: Statistical tests of neutrality of mutations against population growth, hitchhiking and background selection. *Genetics* 147: 915–925.
- Fu Y. & Li W.H. 1993: Statistical tests of neutrality of mutations. *Genetics* 133: 693–709.
- Galtier N., Nabholz B., Glémin S. & Hurst G.D.D. 2009: Mitochondrial DNA as a marker of molecular diversity: a reappraisal. *Mol. Ecol.* 18: 4541–4550.
- Garrick R.C., Sunnucks P. & Dyer R.J. 2010: Nuclear gene phylogeography using PHASE: dealing with unresolved genotypes, lost alleles, and systematic bias in parameter estimation. *BMC Evol. Biol.* 10: 118.
- Griffin D.K., Robertson L.B.W., Tempest H.G. & Skinner B.M. 2007: The evolution of the avian genome as revealed by comparative molecular cytogenetics. *Cytogenet. Genome Res.* 117: 64–77.
- Hare M.P. 2001: Prospects for nuclear gene phylogeography. *Trends Ecol. Evol.* 16: 700–706.
- Hare M.P. & Avise J.C. 1996: Molecular genetic analysis of a stepped multilocus cline in the American oyster (*Crassostrea virginica*). *Evolution* 50: 2305–2315.
- Harpending H.C., Batzer M.A., Gurven M., Jorde L.B., Rogers A.R. & Sherry S.T. 1998: Genetic traces of ancient demography. *Proc. Natl. Acad. Sci. U. S. A.* 95: 1961–1967.
- Harpending H.C., Sherry A.R., Rogers A.R. & Stoneking M. 1993: Genetic structure of ancient human populations. *Curr. Anthropol.* 34: 483–496.
- Harrigan R.J., Mazza M.E. & Sorenson M.D. 2008: Computation vs. cloning: evaluation of two methods for haplotype determination. *Mol. Ecol. Resour.* 8: 1239–1248.
- Hawley D.M., Briggs J., Dhondt A.A. & Lovette I.J. 2008: Reconciling molecular signatures across markers: mitochondrial DNA confirms founder effect in invasive North American house finches (*Carpodacus mexicanus*). *Conserv. Genet.* 9: 637–643.
- Heslewood M.M., Elphinstone M.S., Tidemann S.C. & Baverstock P.R. 1998: Myoglobin intron variation in the Gouldian finch *Erythrura gouldiae* assessed by temperature gradient gel electrophoresis. *Electrophoresis* 19: 142–151.
- Higuera-Gundy A., Brenner M., Hodell D.A., Curtis J.H. & Leyden B.W. 1999: A 10,300 <sup>14</sup>C yr record of climate and vegetation change from Haiti. *Quat. Res.* 52: 159–170.
- Iturralde-Vinent M.A. & MacPhee R.D.E. 1999: Paleogeography of the Caribbean region: implications for Cenozoic biogeography. *Bull. Am. Mus. Nat. Hist.* 281: 1–95.
- Karl S.A., Toonen R.J., Grant W.S. & Bowen B.W. 2012: Common misconceptions in molecular ecology: echoes of the modern synthesis. *Mol. Ecol.* 21: 4171–4189.
- Kimball R.T., Braun E.L., Barker F.K. et al. 2009: A well-tested set of primers to amplify regions spread across the avian genome. *Mol. Phylogenet. Evol.* 50: 654–660.
- Knowlton N., Weigt L.A., Solorzano L.A., Mills D.K. & Bermingham E. 1993: Divergence in proteins, mitochondrial-DNA, and reproductive compatibility across the Isthmus of Panama. *Science* 260: 1629–1632.
- Latta S.C. & Ricklefs R.E. 2010: Prevalence patterns of avian Haemosporida on Hispaniola. *J. Avian Biol.* 41: 25–33.
- Latta S., Rimmer C., Keith A., Wiley J., Raffaele H., McFarland K. & Fernandez E. 2006: *Birds of the Dominican Republic and Haiti*. Princeton University Press, Princeton, NJ.
- Lerner H.R.L., Johnson J.A., Lindsay A.R., Kiff L.F. & Mindell D.P. 2009: It’s not too late for the harpy eagle (*Harpia harpyja*): high levels of genetic diversity and differentiation can fuel conservation programs. *PLoS ONE* 4: 10.



- Lim H.C. & Sheldon F.H. 2011: Multilocus analysis of the evolutionary dynamics of rainforest bird populations in Southeast Asia. *Mol. Ecol.* 20: 3414–3438.
- Mann P., Draper G. & Lewis J.F. 1991: An overview of the geologic and tectonic development of Hispaniola. In: Mann P., Draper G. & Lewis G.F. (eds.), *Geologic and tectonic development of the North American-Caribbean Plate Boundary in Hispaniola. Geological Society of America Special Papers*, 262, Boulder, CO: 1–28.
- Merilä J., Björklund M. & Baker A.J. 1997: Historical demography and present day population structure of the greenfinch, *Carduelis chloris*- an analysis of mtDNA control-region sequences. *Evolution* 51: 946–956.
- Moore W.S. 1995: Inferring phylogenies from mtDNA variation: mitochondrial-gene trees versus nuclear-gene trees. *Evolution* 49: 718–726.
- Nietlisbach P., Arora N., Nater A., Goossens B., van Schaik C.P. & Krützen M. 2012: Heavily male-biased long-distance dispersal of orangutans (genus: *Pongo*), as revealed by Y-chromosomal and mitochondrial genetic markers. *Mol. Ecol.* 21: 3173–3186.
- Norgate M., Chamings J., Pavlova A., Bull J.K., Murray N.D. & Sunnucks P. 2009: Mitochondrial DNA indicates late Pleistocene divergence of populations of *Heteronympha merope*, an emerging model in environmental change biology. *PLoS ONE* 4: 11.
- Perfit M.R. & Williams E.E. 1989: Geological constraints and biological retrodictions in the evolution of the Caribbean Sea and its islands. In: Woods C.A. (ed.), *Biogeography of the West Indies: past, present, and future. Sandhill Crane, Gainesville, FL*: 47–102.
- Qu Y., Lei F., Zhang R. & Lu X. 2010: Comparative phylogeography of five avian species: implications for Pleistocene evolutionary history in the Qinghai-Tibetan plateau. *Mol. Ecol.* 19: 338–351.
- Ramírez-Soriano A., Ramos-Onsins S.E., Rozas J., Calafell F. & Navarro A. 2008: Statistical analysis of neutrality tests under demographic expansions, contractions, and bottlenecks with recombination. *Genetics* 179: 555–567.
- Ramos-Onsins S.E. & Rozas J. 2002: Statistical properties of new neutrality tests against population growth. *Mol. Biol. Evol.* 19: 2092–2100.
- Reding D.M., Freed L.A., Cann R.L. & Fleischer R.C. 2010: Spatial and temporal patterns of genetic diversity in an endangered Hawaiian honeycreeper, the Hawaii akepa (*Loxops coccineus coccineus*). *Conserv. Genet.* 11: 225–240.
- Ricklefs R.E. 2010: Host-pathogen coevolution, secondary sympatry, and species diversification. *Phil. Trans. R. Soc. Lond. B* 365: 1139–1147.
- Ricklefs R. & Bermingham E. 2008: The West Indies as a laboratory of biogeography and evolution. *Phil. Trans. R. Soc. Lond. B* 363: 2393–2413.
- Robinson J.T., Thorvaldsdóttir H., Winckler W., Guttman M., Lander E.S., Getz G. & Mesirov J.P. 2011: Integrative Genomics Viewer. *Nat. Biotechnol.* 29: 24–26.
- Rogers A.R. & Harpending H.C. 1992: Population growth makes waves in the distribution of pairwise genetic differences. *Mol. Biol. Evol.* 9: 552–569.
- Rozas J., Sánchez-DelBarrio J.C., Messeguer X. & Rozas R. 2003: DnaSP, DNA polymorphism analyses by the coalescent and other methods. *Bioinformatics* 19: 2496–2497.
- Slade R.W., Moritz C., Heideman A. & Hale P.T. 1993: Rapid assessment of single copy nuclear DNA variation in diverse species. *Mol. Ecol.* 2: 359–373.
- Sly N.D., Townsend A.K., Rimmer C.C., Townsend J.M., Latta S. & Lovette I.J. 2010: Phylogeography and conservation genetics of the Hispaniolan endemic palm-tanagers (Aves: *Phaenicophilus*). *Conserv. Genet.* 11: 2121–2129.
- Sly N.D., Townsend A.K., Rimmer C.C., Townsend J.M., Latta S.C. & Lovette I.J. 2011: Ancient islands and modern invasions: disparate phylogeographic histories among Hispaniola's endemic birds. *Mol. Ecol.* 20: 5012–5024.
- Sonsthagen S.A., Chesser R.T., Bell D.A. & Dove C.J. 2012: Hybridization among Arctic white-headed gulls (*Larus* spp.) obscures the genetic legacy of the Pleistocene. *Ecol. Evol.* 2: 1278–1295.
- Stephens M. & Scheet P. 2005: Accounting for decay of linkage disequilibrium in haplotype inference and missing data imputation. *Am. J. Hum. Genet.* 76: 449–462.
- Stephens M., Smith N.J. & Donnelly P. 2001: A new statistical method for haplotype reconstruction from population data. *Am. J. Hum. Genet.* 68: 978–989.
- Tajima F. 1989: Statistical method for testing the neutral mutation hypothesis by DNA polymorphism. *Genetics* 123: 585–595.
- Tchaicka L., Eizirik E., de Oliveira T.G., Cândido J.F., Jr. & Freitas T.R.O. 2007: Phylogeography and population history of the crab-eating fox (*Cerdocyon thous*). *Mol. Ecol.* 16: 819–838.
- Toews D.P.L. & Brelsford A. 2012: The biogeography of mitochondrial and nuclear discordance in animals. *Mol. Ecol.* 21: 3907–3930.
- Townsend A.K., Rimmer C.C., Latta S.C. & Lovette I.J. 2007: Ancient differentiation in the single-island avian radiation of endemic Hispaniolan chat-tanagers (Aves: *Calyptophilus*). *Mol. Ecol.* 16: 3634–3642.
- Zink R.M. & Barrowclough G.F. 2008: Mitochondrial DNA under siege in avian phylogeography. *Mol. Ecol.* 17: 2107–2121.
- Zink R.M., Drovetski S.V. & Rohwer S. 2006: Selective neutrality of mitochondrial ND2 sequences, phylogeography and species limits in *Sitta europaea*. *Mol. Phylogenet. Evol.* 40: 679–686.