1-1-2019

Genetic Analysis Reveals Strong Genetic analysis reveals strong phylogeographical divergences within the Scarlet Macaw Ara macao

Kari L. Schmidt
Columbia University

Matthew Aardema
Montclair State University, aardemam@mail.montclair.edu

George Amato
American Museum of Natural History

IBIS
BOU international journal of avian science
Ibis (2020), **162**, 735–748 doi: 10.1111/ibi.12760

Genetic analysis reveals strong phylogeographical divergences within the Scarlet Macaw *Ara macao*

KARI L. SCHMIDT,[1,2] MATTHEW L. AARDEMA[2,3]* & GEORGE AMATO[2]
[1]*Department of Ecology, Evolution, and Environmental Biology, Columbia University, New York, NY, 10027, USA*
[2]*Sackler Institute for Comparative Genomics, American Museum of Natural History, New York, NY, 10024, USA*
[3]*Department of Biology, Montclair State University, Montclair, NJ, 07043, USA*

Scarlet Macaws *Ara macao* have the largest geographical distribution of any Neotropical psittacine, occupying a variety of lowland forest habitats from Mexico to Brazil. Two subspecies, *Ara macao macao* and *Ara macao cyanoptera*, are currently recognized based on wing chord length and plumage coloration, with formal descriptions suggesting genetic introgression in southern Nicaragua and northern Costa Rica. The present study aimed to investigate the extent of genetic diversification within *A. macao* by analysing mitochondrial sequence data from contemporary and historical samples. Phylogenetic reconstruction and population aggregation analysis confirmed two distinct phylogeographical groups, with a high degree of intraspecific genetic structure and no evidence of a putative hybrid zone. Whole mitochondrial genome sequencing further confirmed substantial divergence (~ 1.8%) between the *cyanoptera* and *macao* lineages. These results demonstrate a separation of *A. macao* into two distinct evolutionary entities and highlight a non-uniform distribution of intraspecific diversity, suggesting current conservation designations may warrant re-evaluation.

Keywords: conservation, mitochondrial DNA, museum specimens, Neotropics

Scarlet Macaws *Ara macao* are among the most colourful and charismatic of all New World psittacines. With a near ubiquitous distribution from southern Mexico to Bolivia and eastern Brazil, Scarlet Macaws occur across a diverse array of habitat types. As with many large macaw species, populations undergo seasonal movements in response to staggered phenological events (Renton 2002, Karubian *et al.* 2005). Their broad habitat tolerance and high dispersal capacity highlight the potential for significant genetic connectivity across relatively large areas. However, the geographical extent of the Scarlet Macaw's preferred lowland ecosystems has undergone complex and dynamic shifts across evolutionary time, fuelled by interactions between major geological events and climatic fluctuations (Haffer 2008, Bush *et al.* 2009, Enters *et al.* 2010, Hessler *et al.* 2010). It is currently unclear how the heterogeneity and fluidity of Neotropical environments, coupled with the Scarlet Macaw's life history traits, have influenced patterns of genetic exchange and/or evolutionary isolation across their expansive range.

Described by Linnaeus in 1758, Scarlet Macaws were historically considered a monotypic species (Forshaw 2006). However, based on plumage variation and four morphometric traits among 31 museum specimens of known geographical origin, two subspecies have been described (Wiedenfeld 1994). Ranging from southern Mexico to central Nicaragua, *A. m. cyanoptera* is characterized as more robust, with a wide band of yellow and general absence of green plumage on the median and secondary wing coverts. Populations distributed from southern Nicaragua to Brazil are classified as *A. m. macao* and are identified by reduced yellow and more pronounced green coloration on wing feathers, as well as overall smaller size. Wiedenfeld (1994) also noted a steep decline in wing-chord length and intergradation of colour variants in southern Nicaragua and northern Costa Rica, prompting demarcation of this area as a natural

*Corresponding author.
Email: aardemam@montclair.edu
Twitter: @AardemaMatthew

hybrid zone between *A. m. cyanoptera* and *A. m. macao*.

The utility of subspecies in taxonomy and conservation biology remains contentious given arbitrary diagnosis criteria and ambiguous links to evolutionary history (Mayr 1963, Zink 2004, Fitzpatrick 2010). Therefore, rigorous delineation of conservation targets within Scarlet Macaws is needed to create an empirical foundation for conservation assessment, management strategies and policy decision-making (Mace 2004, Haig *et al.* 2006). Historically common throughout its global distribution, populations have experienced widespread demographic declines and local extinctions due to capture for the pet trade and habitat destruction (Wiedenfeld 1994, Forshaw 2006). Anthropogenic pressures are most severe in Central America, where high human population densities and associated resource demands put tremendous strain on the region's ecosystems. This has resulted in a 22% reduction in primary forest within a 20-year period from 1990 to 2010 (FAO 2010). Census estimates suggest fewer than 4000 Scarlet Macaws remain in Central America, scattered across the landscape in isolated forest fragments (Wiedenfeld 1994). Consequently, *A. macao* is listed under Appendix I of CITES, although it is classified as a species of 'Least Concern' by IUCN and BirdLife International.

This study examines patterns of mitochondrial DNA sequence variation among Scarlet Macaws sampled throughout their known geographical range to evaluate the hypothesis of two weakly differentiated subspecies. We also assess the congruence between mitochondrial haplotypes, subspecies designations and previously reported contact zones.

METHODS

Sample collection, DNA amplification and sequencing

We obtained genetic material from 100 Scarlet Macaw samples of birds distributed across the majority of the species' historical range (Table S1). DNA was extracted from feather, blood or tissue collected from wild individuals in the Laguna del Tigre National Park, Guatemala (n = 11; Collection Permit CONAP I-11-97), and Chiquibul National Forest Reserve, Belize (n = 2; Collection Permit CD/60/3/08(35)). A collaborator (K. Gebhardt, University of Idaho, ID, USA) provided DNA extracts from feathers acquired along clay licks in the Tambopata National Reserve, Peru (n = 8). The remaining tissues were toe-pads taken from study skins located among natural history collections throughout the USA (n = 79). We also obtained genetic material from a single Blue-and-yellow Macaw *Ara ararauna* and Green-wing Macaw *Ara chloropterus* to serve as outgroup taxa for phylogenetic reconstruction.

We extracted total genomic DNA from blood samples using DNeasy tissue extraction kits (Qiagen Inc., Valencia, CA, USA) as per the manufacturer's protocol. Modifications were made to optimize DNA yield from feather and toe-pad samples. These modifications consisted of a 48-h digestion, pre-heating the elution buffer to 70 °C, a 30-min incubation of elution buffer on spin columns prior to centrifugation, and reduced elution volumes. Whole pieces of toe-pad were subject to a bleach bath (10% bleach for 5 min) to reduce surface contaminants and three subsequent water washes prior to DNA extraction, as described above.

DNA was amplified and sequenced at four mitochondrial gene regions: 12S rDNA (12S), 16S rDNA (16S), cytochrome oxidase subunit I (COI) and cytochrome b (cytb). Primers used for PCR-amplification and sequencing are detailed in Table S2. Internal primers were designed using OLIGOANALYSER 3.1 (Integrated DNA Technologies) for each gene region to generate short overlapping amplicons from historical specimens. All polymerase chain reactions (PCRs) were performed on epGradient S Mastercycler thermocyclers (Eppendorf). We employed previously published primers to PCR fragments from modern tissue in a total reaction volume of 15 μL that included: ~ 20–50 ng DNA, 10 mM Tris-HCl (pH 9.0), 50 mM KCl, 1.5 mM $MgCl_2$, 200 μM dNTPs, 0.65 μM of each primer and 0.5 U of *Taq* DNA polymerase (Fisher Scientific, Houston, TX, USA). Thermocycler conditions were as described in Tavares *et al.* (2004), with an annealing temperature of 50 °C. A total volume of 25 μL was used for amplifications from historical tissues, and reactions included: ~ 20–50 ng DNA, 10 mM Tris-HCl (pH 8.3), 50 mM KCl, 1.5 mM $MgCl_2$, 10 μg bovine serum albumen, 200 μM dNTPs, 0.6 μM of each primer and 0.5 U of AmpliTaq Gold DNA polymerase (Applied Biosystems). PCR of short fragments was carried out under the following

conditions: 95 °C (10 min), 35 cycles of 95 °C (30 s), 50 °C (30 s), 72 °C (45 s), and a final extension of 72 °C (7 min). All gene regions were sequenced in both directions using BigDye 3.1 chemistry (PerkinElmer) on an ABI 3730xl sequencer (Applied Biosystems, Foster City, CA, USA). Sequences were visualized, edited and aligned with SEQUENCHER 4.8 (Gene Codes Corp., Ann Arbor, MI, USA).

We followed standard historical DNA protocols when manipulating samples from museum specimens (Gilbert *et al.* 2005, Wandeler *et al.* 2007). Specifically, DNA extractions were conducted in a workspace dedicated to historical tissues, using independent sets of reagents for modern and historical samples to reduce the risk of contamination from exogenous DNA and PCR-amplified products. Multiple negative extraction and PCR controls, along with amplification of short, overlapping fragments (~ 200 bp) were employed to screen for contamination. PCR reactions were repeated to minimize the incorporation of errors due to miscoding lesions. Our sequences have been deposited in GenBank with Accession Nos. MK641864–MK641965, MK6419 66–MK642067, MK642068–MK642169 and MK 642170–MK642271.

Phylogenetic analysis

For each of our four mitochondrial gene regions, we determined the best partitioning scheme and nucleotide substitution model using PARTITIONFINDER 2.1.1 (Lanfear *et al.* 2016) with the small sample size-corrected version of the Akaike information criterion (AICc). The first, second and third codon positions for the two coding genes (COI and cytb) were examined separately. We used two methods of phylogenetic inference to evaluate hierarchical relationships among haplotypes, maximum likelihood (ML) and Bayesian inference (BI). Both analyses were conducted with the full concatenated dataset, with data partitions and mutation models implemented based on the results from our PARTITIONFINDER analysis (Table S3).

We conducted ML phylogenetic reconstruction in RAxML 8.2.12 (Stamatakis 2014). Measures of nodal support for ML analyses were generated from 1000 non-parametric bootstrap replicates and each run was initiated from a random starting tree. MRBAYES 3.2.6 (Ronquist & Huelsenbeck 2003, Altekar *et al.* 2004) was used to reconstruct phylogenetic relationships using BI. A Markov chain Monte Carlo process was set for five simultaneous chains with 10 million generations, each starting from a random tree and using the default heating scheme. Markov chains were sampled every 100 generations, with the initial 25% of trees discarded as burn-in.

Population aggregation analysis

We used the discrete character-based methodology of population aggregation analysis (PAA; Davis & Nixon 1992) to test the current hypothesis of intraspecific diversity in Scarlet Macaws. With this approach, populations sharing a suite of fixed nucleotide differences are aggregated into diagnosably distinct taxa (Cracraft 1983). Character fixation provides a contextual framework to infer monophyly by demonstrating that terminal taxa are united through hierarchical, rather than reticulate, genealogies (Goldstein *et al.* 2000). To evaluate the *a priori* hypothesis of two distinct units within Scarlet Macaws (i.e. *A. m. cyanoptera* and *A. m. macao*) proposed by Wiedenfeld (1994), we generated a character matrix using MACCLADE 4.06 (Sinauer Associates Inc., Sunderland, MA, USA). The matrix was screened for the presence or absence of fixed and alternate character differences among putative subspecific taxa.

Molecular diversity

Corrected genetic distances (K2P; Kimura 1980) were calculated among each aggregated population identified in this study using the four genetic regions described above, as executed in MEGA 7.0 (Kumar *et al.* 2008) . The number of haplotypes (h), haplotype (H_d) and nucleotide (π) diversity, number of segregating sites (S) and number of singleton mutations were quantified in DnaSP 5.10.01 (Librado & Rozas 2009) to further investigate patterns of intraspecific genetic variation. We constructed median-joining networks (Bandelt *et al.* 1999) using POPART 1.7 (Leigh & Bryant 2015) to visualize relationships between mitochondrial haplotypes. Given regions of the nucleotide matrix with gaps or missing data are excluded from inferences of molecular diversity, we created a truncated dataset including only samples with complete concatenated haplotypes to maximize the number of sites under consideration and provide a more comprehensive overview of nucleotide variation.

Whole mitochondrial genome sequencing

To assess further taxonomic distinctiveness and evaluate the extent of divergence between *A. m. cyanoptera* and *A. m. macao*, we sequenced a complete mitochondrial genome from each subspecies. Sequenced samples came from two captive males of the Ara Project (http://thearaproject.org/) in Costa Rica. The Macaws within the Ara Project consist of confiscated birds, pets that have been donated by their owners, and birds that have been raised in the breeding centre. Thus, the source locations of the individuals are generally unknown, although they are assumed to be native to Costa Rica. Prior to sequencing the whole mitochondrial genomes of these two birds, their mitochondrial haplotypes were assessed as described above to assign taxonomic status (also see Results). These two birds were not part of the previous population aggregations and broad sampling phylogenetic analyses.

We extracted whole genomic DNA from these two samples as described above for contemporary blood samples. These extractions were used to construct standard paired-end libraries (150 bp) with barcoded adapters using the Illumina TruSeq Library Preparation kit with the standard protocol (Illumina, San Diego, CA, USA) after fragmentation on a Covaris device (Woburn, MA, USA). Following library preparation, we pooled the two whole genome sequence (WGS) libraries with three other libraries (unrelated to this project) and this multiplexed library was sequenced across five lanes of an Illumina HiSeq X Ten. Library preparation and sequencing were done at the New York Genome Center (NYGC).

Reads were de-multiplexed according to each sample's unique barcoded adapter at the NYGC. We trimmed the sample-specific raw reads using TRIMGALORE 0.4.1 (https://www.bioinformatics.babraham.ac.uk/projects/trim_galore/), excluding regions with a quality score less than 30 or remnant TruSeq adapter sequence, and then discarded reads that were less than 20 base pairs in length after quality trimming.

To reduce the potential for mapping biases and/or mis-mapped reads from nuclear DNA sequences of mitochondrial origin ('numts', Hazkani-Covo *et al.* 2010), we mapped our trimmed reads to the full Scarlet Macaw reference genome (GenBank Accession Number: GCA_000400695.1, Seabury *et al.* 2013) using the BURROWS-WHEELER ALIGNER (BWA) 0.7.12 with 'MEM' option (Li & Durbin 2009). After mapping, PCR duplicates were removed with PICARD TOOLS 1.119 (https://broadinstitute.github.io/picard/). We then used the GENOME ANALYSIS TOOLKIT 3.6 ('GATK'; McKenna *et al.* 2010) to realign and mask insertions and deletions (indels) and identify divergent sites. Indel realignment was done using IndelRealigner and variants were called with HaplotypeCaller, both with GATK. The FastaAlternateReferenceMaker tool in GATK was used to make a final FASTA file for each of the two samples. From these sample-specific assembles, we obtained mitochondrial genomes for subsequent analyses. These genomes have been deposited in GenBank with Accession Nos. MK351783 and MK351784.

We compared the mitochondrial genomes of our two sequenced samples with all other *Ara* species with published mitochondrial genomes, as well as additional Psittacidae (Table S4). Alignment of the mitochondrial genomes was performed using CLUSTAL OMEGA (Sievers *et al.* 2011). We determined the best partitioning schemes and nucleotide substitution models for the mitochondrial genomes simultaneously using PARTITIONFINDER 2.1.1 (Lanfear *et al.* 2016) with the small sample sized-corrected version of the AICc. Our data blocks included the first, second and third codon positions for each of the 13 protein coding regions, two rRNAs and 22 tRNAs (Table S5). Phylogenetic analysis was carried out using this partitioning scheme using RAxML 8.2.12 (Stamatakis 2014). To determine branch support values, we performed 1000 pseudo-bootstrap replicates.

RESULTS

Data quality

Mitochondrial haplotypes consisted of 2245 aligned nucleotide characters distributed across four gene regions: 12S (357 bp), 16S (538 bp), COI (600 bp) and cytb (750 bp). There were no indels in the protein coding genes. Seventy haplotypes were detected among concatenated mitochondrial DNA (mtDNA) sequences from the 100 Scarlet Macaws sampled across the contemporary and historical species' range. A

total of 112 variable sites (~ 5%) were observed within the nucleotide alignment, including a single indel (1 bp) within the 16S sequence. Complementary haplotypes were also successfully generated from a Blue-and-gold Macaw and Green-wing Macaw. Single-banded

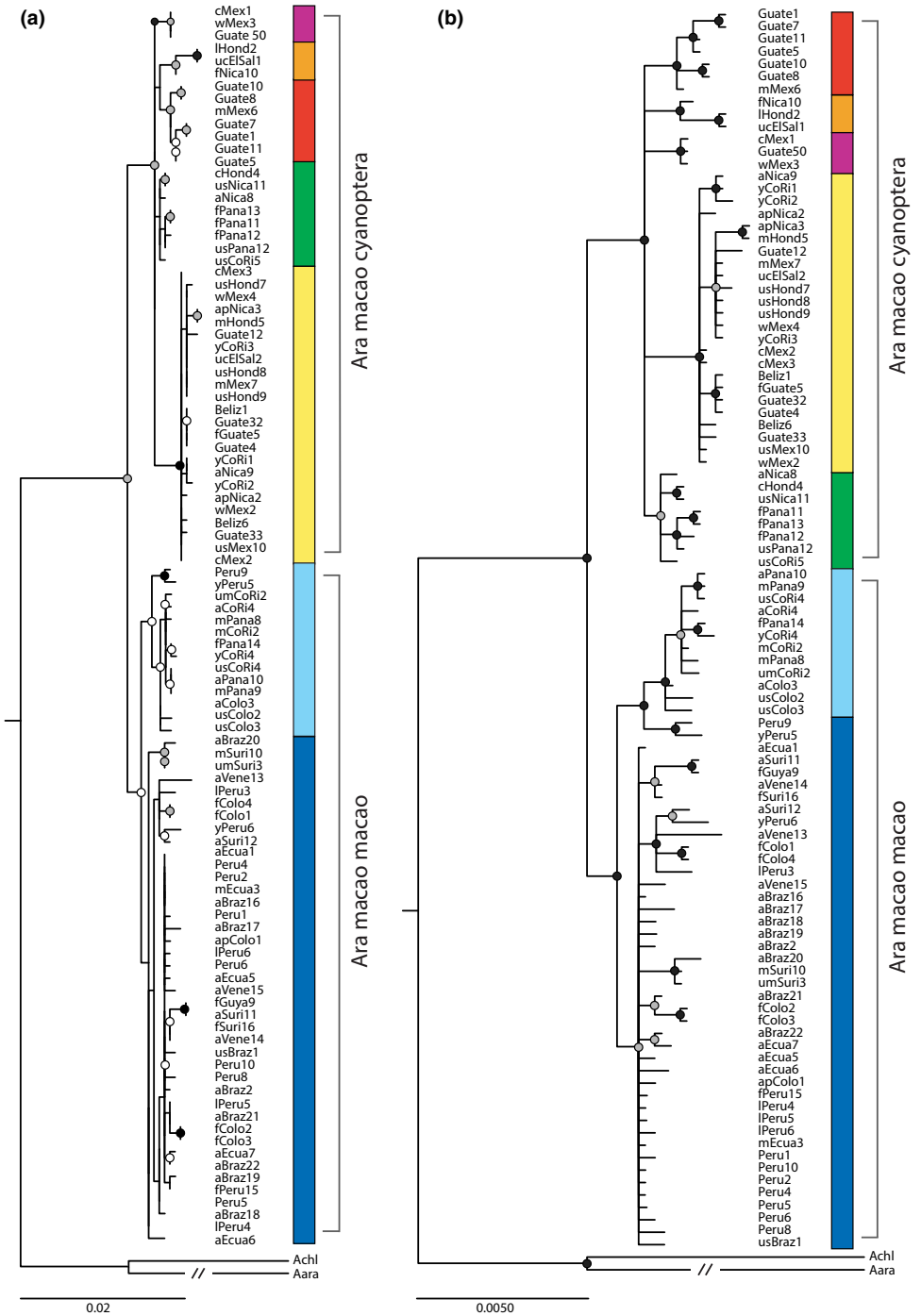

Figure 1. Maximum likelihood (a) and Bayesian (b) phylogenetic trees. Coloured circles designate bootstrap support and posterior probabilities. White: > 60/0.60; grey: > 80/0.80; black: > 95/0.95. Colour scheme is consistent with haplogroup designations given in Fig. 2.

PCR products of the expected size were obtained for all amplifications with previously published external primer pairs (modern samples) and novel internal primer pairs for small overlapping fragments (historical samples). Multiple independent DNA extractions and PCR amplifications from museum tissues resulted in 3–8× sequence coverage for amplicons exhibiting unique single base pair mutations.

Phylogenetic analysis

Phylogenetic reconstruction based on our 2245-bp concatenated sequence dataset revealed two major monophyletic groups, with near identical topologies recovered for both ML and BI (Fig. 1). Sampled haplotypes within *A. m. cyanoptera* form a well-supported clade (ML bootstrap support (BS_{ML}) = 82, and BI posterior probability (PP_{BI}) = 1.0; Fig. 1). Each *cyanoptera* haplogroup clustered into a monophyletic group, although in some cases support was low (BS_{ML} = 58–97, PP_{BI} = 0.94–1.0; Fig. 1). Evolutionary relationships among the five haplogroups were poorly resolved: ML denotes Haplo2 as sister to the other four *cyanoptera* haplogroups, whereas all five haplogroups collapse into a single polytomy in the BI analysis (Fig. 1). ML analysis revealed low nodal support for *A. m. macao* (BS_{ML} = 62; Fig. 1), yet BI yielded high posterior probabilities (PP_{BI} = 0.96; Fig. 1). Similarly, *trans*-Andean haplotypes form a monophyletic group within the otherwise paraphyletic Haplo4 clade with moderate bootstrap support (BS_{ML} = 79; Fig. 1) and high posterior probabilities (PP_{BI} = 0.98; Fig. 1). Haplo7 haplotypes of Central American origin are monophyletic with similar nodal support (BS_{ML} = 71, PP_{BI} = 0.93; Fig. 1), nested within the broader *trans*-Andean clade.

Population aggregation analysis

Examination of the character matrix yielded unambiguous support for two distinct evolutionary units, as evidenced by four diagnostic characters distributed across gene regions: 12S (one), COI (one) and cytb (two). No diagnostic characters correspond with the hypothesized subspecies boundary in central Nicaragua. Instead, the geographical limit appears to occur further south in Costa Rica (Fig. 2). Forty-six mitochondrial haplotypes were identified among the 66 individuals with complete (2245 bp) sequences. Scarlet Macaw subspecies form two distinct clusters within the median-joining network, differentiated by eight mutational steps; no haplotypes are shared between subspecies (Fig. 2). Inspection of mtDNA sequence data also revealed seven clusters of closely related haplotypes (herein referred to as haplogroups and numbered in order of their discovery).

Five clusters were identified within *A. m. cyanoptera*, consistent with the phylogenetic reconstruction (Haplo1, Haplo2, Haplo3, Haplo5 and Haplo6). Each haplogroup pair is differentiated by three to eight fixed nucleotide characters, generally consisting of one or two shared haplotypes and few singleton mutations, with considerable geographical overlap (Fig. 2). *Ara m. macao* was characterized by two haplogroups (Haplo4 and Haplo7), distinguished from each other by two fixed mutations, with discrete geographical ranges delimited by the Andean cordilleras (Fig. 2). The *trans*-Andean populations in Colombia and lower Central America belong to Haplo7, whereas Haplo4 is distributed throughout the Amazon Basin.

Haplo2 was found throughout upper Central America, from Mexico to northeastern Costa Rica, whereas the remaining four haplogroups for *cyanoptera* exhibited more restricted geographical distributions. Haplo1 and Haplo6 were both detected in Mexico and Guatemala. Haplo3 and Haplo5 demonstrated similar ranges along the Pacific slope of El Salvador, Honduras and Nicaragua; however, Haplo3 was also recovered on Isla Coiba, Panama (Fig. 2). There is little resolution regarding relationships between *cyanoptera* haplotypes, with the exception of Haplo2 appearing more diverged from the other four *cyanoptera* haplogroups. Additional median-joining networks generated for each individual gene region, using both truncated (n = 66) and complete datasets (12S (n = 98), 16S (n = 90), COI (n = 90), and cytb (n = 94)), varied only in the number of peripheral nodes, suggesting the reduced sample size did not alter the underlying network structure (data not shown).

The majority of *macao* individuals demonstrate a star-like topology; although more divergent subgroups are present within Haplo4, they exhibit little geographical structure and remain closely associated with the haplogroup's broader reticulate network (Fig. 2). One aberrant Haplo4 mtDNA sequence was recovered closer to Haplo7 than to members of its own haplogroup, even though it shares the fixed nucleotide mutations characteristic of Haplo4.

Molecular diversity

Average genetic distances between the two putative subspecific taxa equalled 0.8%. Inter-haplogroup distances ranged from 0.2% (Haplo3/Haplo6) to 0.9% (Haplo2/Haplo4). Comparable levels of haplotype and nucleotide diversity were observed for *A. m. cyanoptera* (H_d = 0.97; π = 0.00312) and *A. m. macao* (H_d = 0.97; π = 0.00302). A reduction in haplotype diversity was noted for five haplogroups upon independent analysis, ranging from 0.00 to 0.87, and nucleotide diversity declined in concert (Table 1). Two haplogroups maintained high haplotype diversity when considered separately, H_d = 0.93 and 0.97 for Haplo3 and Haplo4, respectively, yet nucleotide diversity decreased for both (Table 1). A total of 55 polymorphisms (56% singleton mutations) were detected among *macao* haplotypes, in contrast to 35 polymorphisms (26% singleton mutations) within *cyanoptera*.

Whole mitochondrial genome sequencing

Our two mitochondrial genome assemblies were 16 970 nucleotides (*A. m. macao*; Haplo7) and

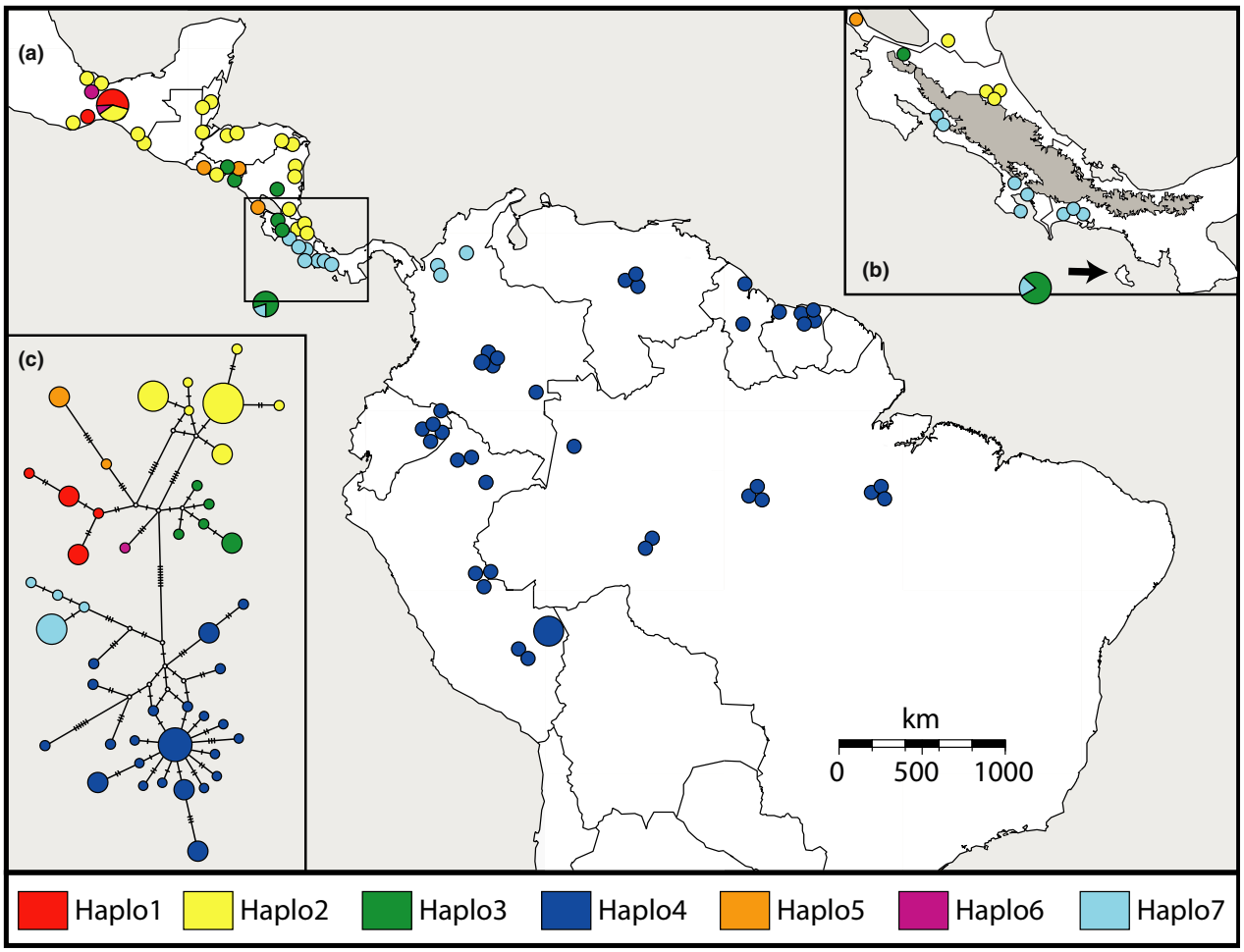

Figure 2. (a) Sampling effort and geographical distribution of haplogroups detected across Central and South America. Each dot represents a single specimen, with the exception of Laguna del Tigre National Park, Guatemala (*n* = 11); Isla Coiba, Panama (*n* = 5); and Tambopata National Reserve, Peru (*n* = 8). (Inset b) Close-up of subspecific boundary. Grey: major areas ≥ 500 m in elevation. Arrow indicates the location of Isla Coiba, Panama. (Inset c) Haplotype network showing relationships and genetic distances between samples and haplotypes. The size of the circle indicates the number of samples represented. The hashes on each branch represent the number of genetic differences between haplotypes.

Table 1. Genetic variation within and among Scarlet Macaw subspecies and haplogroups, generated from the truncated haplotype dataset.

Subspecies	Length (bp)	*n*	*h*	H_d	$\pi \times 10^{-2}$	*S*	Singletons
A. m. cyanoptera	2227	29	19	0.966 (0.018)	0.312 (0.022)	35	13
Haplo1	2231	6	4	0.867 (0.129)	0.105 (0.024)	5	2
Haplo2	2239	13	7	0.872 (0.067)	0.088 (0.015)	8	5
Haplo3	2245	6	5	0.933 (0.122)	0.095 (0.015)	5	3
Haplo5	2245	3	2	0.667 (0.314)	0.119 (0.056)	4	4
Haplo6	2245	2	1	[a]	[a]	[a]	[a]
A. m. macao	2245	37	27	0.974 (0.015)	0.302 (0.031)	55	33
Haplo4	2245	31	23	0.97 (0.020)	0.231 (0.032)	49	34
Haplo7	2245	6	4	0.800 (0.172)	0.065 (0.017)	3	1

h, number of haplotypes. H_d, haplotype diversity (SD). π, nucleotide diversity (SD). *S*, number of segregating sites. [a]Values could not be calculated.

16 969 nucleotides (*A. m. cyanoptera*; Haplo3) long, respectively. The *A. m. macao* assembly had one single-nucleotide insertion, one single-nucleotide deletion, 75 transitions and one transversion relative to the published *A. macao* mitochondrial genome. Our *A. m. cyanoptera* assembly had two single-nucleotide insertions, three single-nucleotide deletions, 156 transitions and 11 transversions relative to the published *A. macao* mitochondrial genome. None of the indels occurred within coding regions. In concordance with our broadly sampled phylogenetic reconstructions, our sequencing and analysis of the full mitochondrial genomes of *A. m. cyanoptera* and *A. m. macao* showed clear separation of the two taxa (Fig. 3), with a combined branch length of 0.0178 (1.8% mitochondria-wide divergence). All nodes had 100% bootstrap support except the node that joined *A. severus* with *A. tricolor*, *A. militaris* and *A. macao* (bootstrap support: 59%). The complete tree topology matches those previously published based on whole-mitochondrial genomes within the Psittacidae (e.g. Urantówka *et al.* 2017, Johansson *et al.* 2018).

DISCUSSION

Evaluation of current taxonomy

Within current taxonomic descriptions, the primary identifying characters for Scarlet Macaw subspecies are listed as a generally larger size for *A. m. cyanoptera* and the presence of a green band on the secondary wing coverts for *A. m. macao*. However, previous examinations of morphometric data revealed weak differentiation of geographical variants (Wiedenfeld 1994). Only wing chord length demonstrated statistical significance when corrected for sex (♂: $P = 0.045$; ♀: $P = 0.017$). Likewise, the extent of green tips on secondary wing coverts, an inherently ambiguous and subjective trait, exhibits a pattern of clinal variation; *A. m. cyanoptera* is described as having little to no green coloration, whereas *A. m. macao* exhibits small to substantial amounts of green plumage, with no empirical means to quantify 'little' vs. 'small'.

Relative to morphological data, a clearer picture emerges when we consider molecular genetic data. There is strong divergence between clades and sequence variation clearly differentiates Scarlet Macaw subspecies. Similarly, tree topologies representing both our broad sampling efforts and mitogenome analyses recover *A. m. cyanoptera* and *A. m. macao* as reciprocally monophyletic (Figs 1 and 3).

Ranging from Mexico to northern Costa Rica, *A. m. cyanoptera* forms a well-supported monophyletic group consisting of five reciprocally monophyletic haplogroups, with each pair differentiated by three to eight fixed nucleotide mutations. Although there is insufficient resolution within the reconstructed phylogenies or haplotype network to infer evolutionary relationships among haplogroups, extensive geographical overlap observed throughout northern Central America suggest these distinct haplogroups represent a single evolutionary entity.

Further south, in the humid lowland forests of lower Central and South America, populations of *A. m. macao* also form a monophyletic group with low nodal support. They exhibit less pronounced sequence divergence but greater geographical structure relative to *A. m. cyanoptera*. Two fixed

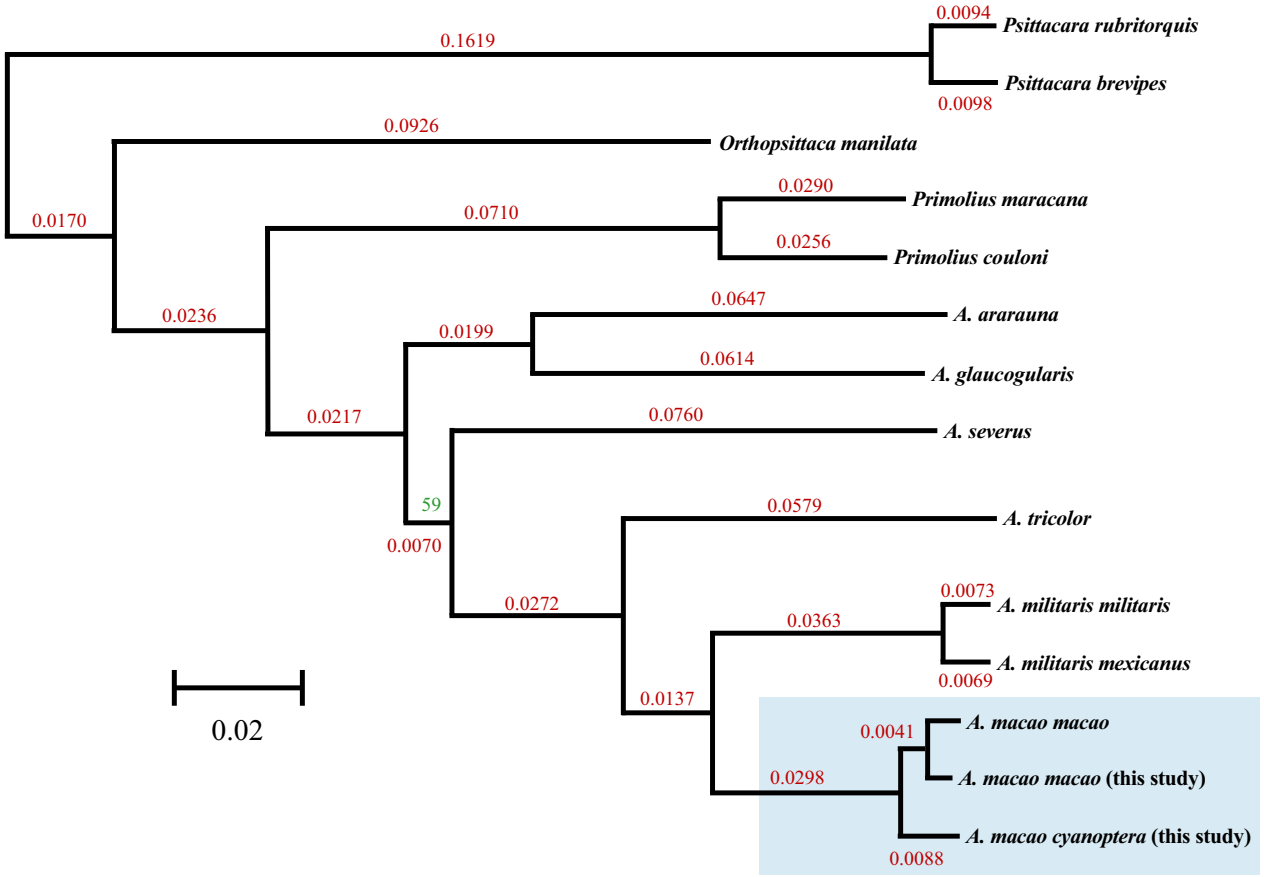

Figure 3. Maximum likelihood phylogenetic tree of full mitochondrial genomes. All nodes had 100% bootstrap support except that joining *A. severus* with *A. tricolor*, *A. militaris* and *A. macao*. This node had 56% bootstrap support (indicated in green). Red numbers indicate branch lengths under the partitioning scheme determined with PARTITIONFINDER (Table S5). The blue box highlights the *A. m. macao* and *A. m. cyanoptera* branches.

nucleotide mutations differentiate *cis*-Andean from *trans*-Andean populations, with an additional fixed nucleotide delineating *macao* populations in southern Costa Rica and Panama (Fig. 2). Interestingly, the *cis*-Andean clade is paraphyletic, probably driven by an outlying member of Haplo4 (Fig. 2), with Central American populations nested within the monophyletic *trans*-Andean clade (Fig. 1).

Short internode and terminal branches recovered from our broadly sampled phylogenetic reconstructions (Fig. 1), coupled with small genetic distances (average K2P distances: 0.8% between putative subspecies), suggest *cyanoptera* and *macao* are young lineages. When we compare the complete mitochondrial genomes of these two subspecies, we see a greater extent of divergence (1.8%), although still consistent with a recent origin. Using a conventional rate of 2.1% mitochondrial sequence divergence per million years (Weir & Schluter 2008), our data indicate these subspecific lineages diverged around 0.32 and 0.85 mya during the mid- to late Pleistocene.

Formal taxonomic descriptions include discussion of possible introgression between *A. m. cyanoptera* and *A. m. macao*, with the intergradation of colour variants coinciding with a steep decline in wing chord length across southern Nicaragua and northern Costa Rica viewed as evidence of a hybrid zone (Wiedenfeld 1994). However, the distribution of mtDNA haplotypes observed in this study demonstrates a general pattern of geographical isolation; no *cyanoptera* haplotypes were detected on mainland Central America south of the Costa Rican cordilleras ($n = 8$). The only recovered instance of *cyanoptera* haplotypes within the *macao* distribution occurs on Isla Coiba, a large island located off the Pacific slope of Panama (Fig. 2). The presence of Haplo3 over

300 km from the nearest empirically confirmed locality for *A. m. cyanoptera* is atypical for a stable hybrid zone, especially given intervening habitats are flanked by two prominent geographical barriers (i.e. central cordilleras of Costa Rica, a 22.5-km strait in the Gulf of Chiriquí).

Broader examination of the Scarlet Macaw's range shows mitochondrial haplogroups are widely distributed across large geographical areas (Fig. 2), consistent with recent studies suggesting female philopatry plays a limited role in population differentiation among Neotropical psittacines (Wright *et al.* 2005, Gebhardt 2007, Faria *et al.* 2008, but see also Caparroz *et al.* 2009). Instead, this study found *cyanoptera* and *macao* haplotypes separated by the central cordilleras of Costa Rica, a volcanic mountain range varying in elevation from 500 to over 3800 m and transecting the country, with near complete isolation of the Scarlet Macaw's preferred lowland habitats along the Pacific and Caribbean slopes (Fig. 2). Furthermore, a recent study of multilocus microsatellite genotypes by Monge *et al.* (2016) found significant population genetic structure for Scarlet Macaws along the Pacific versant of Costa Rica, suggesting that, if present, a contemporary hybrid zone would probably be confined to a narrow geographical area.

While showing the potential for regional differentiation, Wiedenfeld's morphological analyses were unable to capture the strength of divergence and complexity of the Scarlet Macaw's recent evolutionary history. Our observed disparate patterns of molecular diversity and spatial distribution of mitochondrial haplotypes support *A. m. cyanoptera* and *A. m. macao* as diagnosably distinct, reciprocally monophyletic and geographically isolated evolutionary units. We recommend recognizing *cyanoptera* and *macao* as distinct subspecies, representing independent conservation units. It is also important to note that although *macao* Haplo7 is a unique and spatially discrete haplogroup, the lack of reciprocal monophyly is indicative of incomplete lineage sorting and ongoing divergence between *macao* haplogroups, advocating the characterization of *trans*-Andean populations as distinct population segments (DPS) within the broader *A. m. macao* lineage.

Patterns of diversification

Oscillations in the distribution of habitats in response to changes in temperature and precipitation are widely considered to be key factors governing the historical biogeography of lowland ecosystems across the Neotropics (Haffer 1969). Regional differences in the intensity and duration of population fragmentation and associated demographic changes provide a logical theoretical framework to explain patterns of evolutionary divergence within Scarlet Macaws.

In the case of *A. m. cyanoptera*, the subspecies' range encircles the Central American landmass, geographically constrained by an extensive system of central highlands and the coastal waters of the Atlantic and Pacific Oceans. The Scarlet Macaws' preferred humid lowland habitats are heavily dependent on the availability of fresh water, and thus shifts in palaeoecological distributions would closely track this critical resource. The genetic signatures recovered among mtDNA haplotypes may reflect the fragmentation of ancestral *A. m. cyanoptera* into allopatric refugia. The presence of several distinct *cyanoptera* haplogroups appears to inflate levels of overall sequence variation, as diversity indices decline for the majority of haplogroups when treated as separate units (Table 1). The *cyanoptera* nucleotide dataset includes relatively few segregating sites and singleton mutations, with most polymorphisms representing fixed characters differentiating haplogroups (Table 1, Fig. 2). Patterns of high inter- and low intra-haplogroup diversity are consistent with recurrent vicariance events.

A markedly different pattern of genetic variation emerges upon examination of Scarlet Macaw populations across lower Central and South America. *A. m. macao* is represented by only two haplogroups, differentiated by two fixed nucleotide bases and geographically separated by the Andes Mountains (Figs 1 and 2). Values of haplotype and nucleotide diversity recovered within Haplo4 are comparable to overall estimates of molecular variation, whereas Haplo7 exhibits a decline in diversity, the number of segregating sites and singleton mutations found among *macao* haplotypes far exceeding those of *A. m. cyanoptera* (Table 1). Low inter- and high intra-haplogroup diversity is an indication of the historical effective population size of ancestral *A. m. macao* remained relatively stable despite intense palaeoclimatic oscillations. As discussed above, the physical attributes of *A. m. macao*'s range were likely to be instrumental in driving demographic trends. Amazonia represents the nucleus of Haplo4's *cis*-Andean range, covering a geographical area an order of magnitude

greater than the entirety of Central America with relatively little surface relief. Accumulation of considerable sequence variation within mitochondrial haplotypes implies robust populations of Scarlet Macaws persisted within the complex mosaic of tropical wet forest assemblages found within Amazonia, even during the most extreme glacial cycles. The broad geographical distribution of closely related haplotypes, lack of hierarchical structure and reticulate relationships within the Haplo4 lineage (Figs 1 and 2) further support the hypothesis that geological and climatic perturbations were insufficient to disrupt genetic connectivity throughout the vast Amazonian lowlands.

Conversely, indices of intra-haplogroup diversity recovered for *macao* Haplo7 bear the hallmarks of demographic instability, similar to those seen within *A. m. cyanoptera*. Again, these may correlate with the physical characteristics of the haplogroup's *trans*-Andean distribution. In northwestern Colombia, Scarlet Macaws inhabit moist tropical ecosystems along the mid- to lower Magdalena River Valley, bounded by the Central and Oriental Cordilleras of the Northern Andes (Hilty & Brown 1986). Similarly, coastal plains in lower Central America are confined to a narrow band flanked by the central highlands and Pacific Ocean (Olson *et al.* 2001). The total geographical extent of these lowland habitats covers an area markedly smaller than either upper Central America or the Amazon Basin, with fewer major sources of fresh water. Two diagnostic nucleotide substitutions differentiate *trans*- and *cis*-Andean lineages of *A. m. macao*, with a single fixed mutation distinguishing Central from South American populations. Although the weakly divergent haplotypes and paraphyly observed within *A. m. macao* are indicative of incomplete lineage sorting, the potential for recurrent demographic bottlenecks across *trans*-Andean populations raises the possibility of introgression and/or mitochondrial sweeps producing deceptively shallow divergence estimates (McKay & Zink 2010). Expanding the sampling effort to include nuclear loci would further refine our inferences into the timing and evolutionary forces involved in the diversification of Scarlet Macaws.

Conservation implications

The confirmation of cryptic genetic diversity exposes a critical gap between conservation needs and conservation status for these charismatic parrots. At present, Scarlet Macaws are considered a species of 'Least Concern', given global demographic trends, population size and extent of available habitat are above designated thresholds for threatened status (IUCN 2015). Splitting Scarlet Macaws into two distinct conservation units, however, immediately transforms the conservation status of this group. Approximately 83% of preferred lowland habitats are located within the Amazon Basin, along with the majority of the estimated census population of 20 000–50 000 individuals for *A. macao* (BirdLife International 2011). Deforestation rates are declining throughout South America, dropping from 0.49 to 0.41% in the past decade (FAO 2010), with human population growth rates falling in concert from 1.17% to 1.07% (CEPAL 2013). Therefore, robust populations of *A. m. macao* ranging across the Amazonian lowlands may indeed qualify for 'Least Concern' status. Conversely, the situation faced by Scarlet Macaws in Central America is far more precarious. Annual human population growth rates are currently 1.59% and human densities are 4.2-fold greater relative to South America (CEPAL 2013), putting tremendous strain on the region's natural resources. Forest area in Central America declined by an average of 1.19% annually between 2000 and 2010, the highest rate reported by the United Nations (FAO 2010). The loss of important foraging and nesting habitats, coupled with intense nest poaching for the pet trade, have decimated *A. m. cyanoptera*, with fewer than 4000 wild *cyanoptera* remaining in isolated forest fragments throughout upper Central America (Juniper & Parr 1998). Recovery of five unique mitochondrial haplogroups within the *cyanoptera* lineage highlights the evolutionary significance of these populations, further advocating that *A. m. cyanoptera* be uplisted to 'Vulnerable' status. Future studies should give special consideration to Mesoamerican locales to further elucidate comparative biogeographical patterns, providing an important contextual framework for the development of conservation initiatives throughout the region.

Role of mitochondrial genomes in phylogeograhical studies

While providing a straightforward approach to molecular analyses, it is imperative to note the inherent limitations of mtDNA and the need for

cautious interpretation of evolutionary inferences based solely on mtDNA (Moritz 1994). For example, there is a growing literature citing discrepancies when comparing mitochondrial and nuclear datasets (Toews & Brelsford 2012). One common source of discord is the uniparental mode of inheritance of mitochondria, specifically regarding susceptibility to sex-biased dispersal and introgression. The latter is particularly relevant to evolutionary studies when coupled with population bottlenecks or selective pressures, leading to possible mitochondrial sweeps replacing historical variants and decreasing divergence estimates (Kodandaramaiah *et al.* 2013, Seixas *et al.* 2018). In addition, locus to locus stochasticity in gene histories means mitochondrial genealogies reflect a single scenario of population demography and differentiation (Garrick *et al.* 2015). Despite these limitations, mtDNA remains an important resource for evolutionary biology by serving as a first line of inquiry for broad-scale phylogeographical studies, establishing an empirical framework to guide future sampling efforts, while also deepening our understanding of divergence events when combined with nuclear loci (DeSalle *et al.* 2017). Mitochondrial haplotypes have an especially critical role in historical DNA studies, even in the era of next-generation sequencing, as the molecule's high copy number allows recovery of whole mitogenomes for samples where the nuclear genome is too degraded for reliable calling of single nucleotide polymorphisms (Burrell *et al.* 2015).

We extend our gratitude to the directors and field assistants of the Wildlife Conservation Society – Guatemala Program and Friends for Conservation and Development for their assistance and logistical support during sample collection in Guatemala and Belize, respectively. We thank Donald Brightsmith and Kara Gebhardt for sharing DNA extracts from contemporary sampling localities in Tambopata, Peru. We are grateful to the curators and collections managers from the following institutions for granting access to their collections and providing historical tissue samples: Academy of Natural Sciences, American Museum of Natural History, Field Museum of Natural History, California Academy of Sciences, Carnegie Museum of Natural History, Field Museum of Natural History, Harvard Museum of Comparative Zoology, Louisiana State University Museum of Natural Sciences, National Museum of Natural History, UCLA-Dickey Collection, University of Michigan Museum of Zoology, and Yale Peabody Museum. Lastly, we thank Taylor Hains for assistance with determining the haplogroups of the two samples used in our whole mitochondrial genome sequencing. This work was made possible through graduate research funds given by the Ecology, Evolution and Environmental Biology Department (Columbia University) and the Sackler Institute for Comparative Genomics (American Museum of Natural History) to K.L.S. Additional financial support was provided by the Frank M. Chapman Memorial Fund (American Museum of Natural History) for K.L.S. M.L.A. was supported in part by a Gerstner Fellowship in Bioinformatics and Computational Biology to the American Museum of Natural History.

DATA AVAILABILITY

All data described in this paper may be accessed via GenBank. Please see the text or Supporting Information Tables S1 and S4 for corresponding accession numbers.

REFERENCES

Altekar, G., Dwarkadas, S., Huelsenbeck, J.P. & Ronquist, F. 2004. Parallel metropolis coupled Markov chain Monte Carlo for Bayesian phylogenetic inference. *Bioinformatics* **20**: 407–415.

Bandelt, H.J., Forster, P. & Röhl, A. 1999. Median-joining networks for inferring intraspecific phylogenies. *Mol. Biol. Evol.* **16**: 37–48.

BirdLife International. 2011. *Species factsheet: Scarlet Macaw Ara macao*. Cambridge: BirdLife International.

Burrell, A.S., Disotell, T.R. & Bergey, C.M. 2015. The use of museum specimens with high-throughput DNA sequencers. *J. Hum. Evol.*, **79**: 35–44.

Bush, M.B., Correa-Metrio, A.Y., Hodell, D.A., Brenner, M., Anselmetti, F.S., Ariztegui, D., Mueller, A.D., Curtis, J.H., Grzesik, D.A., Burton, C. & Gilli, A. 2009. Re-evaluation of climate change in lowland Central America during the Last Glacial Maximum using new sediment cores from Lake Petén Itzá, Guatemala. In Vimeux,F., Sylvestre,F. & Khodri,M. (eds) *Past Climate Variability in South America and Surrounding Regions*, no. 14: 113–128. Dordrecht, the Netherlands: Springer.

Caparroz, R., Miyaki, C.Y. & Baker, A.J. 2009. Contrasting phylogeographical patterns in mitochondrial DNA and microsatellites: evidence of female philopatry and male-biased gene flow among regional populations of the blue-and-yellow macaw (Psittaciformes: *Ara ararauna*) in Brazil. *Auk* **126**: 359–370.

CEPAL. 2013. Anuario estadística de América Latina y el Caribe 2013. Santiago: United Nations.

Cracraft, J. 1983. Species concepts and speciation analysis. *Curr. Ornithol.* **1**: 159–187.

Davis, J.I. & Nixon, K.C. 1992. Populations, genetic variation, and the delimitation of phylogenetic species. *Syst. Biol.* **41**: 421–435.

DeSalle, R., Schierwater, B. & Hadrys, H. 2017. MtDNA: the small workhorse of evolutionary studies. *Front. Biosci. (Landmark Ed)* **22**: 873–887.

Enters, D., Behling, H., Mayr, C., Dupont, L. & Zolitschka, B. 2010. Holocene environmental dynamics of south-eastern Brazil recorded in laminated sediments of Lago Aleixo. *J. Paleolimnol.* **44**: 265–277.

FAO 2010. *Global Forest Resources Assessment 2010*. Rome: Food and Agricultural Organization of the United Nations.

Faria, P.J., Guedes, N.M.R., Yamashita, C., Martuscelli, P. & Miyaki, C.Y. 2008. Genetic variation and population structure of the endangered Hyacinth Macaw (*Anodorhynchus hyacinthinus*): implications for conservation. *Biodivers. Conserv.* **17**: 765–779.

Fitzpatrick, J.W. 2010. Supspecies are for convenience. *Ornithol. Monogr.* **67**: 54–61.

Forshaw, J.M. 2006. *Parrots of the World: An Identification Guide*. Princeton: Princeton University Press.

Garrick, R.C., Bonatelli, I.A.S., Hyseni, C., Morales, A., Pelletier, T.A., Perez, M.F., Rice, E., Satler, J.D., Symula, R.E., Thomé, M.T.C. & Carstens, B.C. 2015. The evolution of phylogeographical data sets. *Mol. Ecol.* **24**: 1164–1171.

Gebhardt, K.J. 2007. Using molted feathers as a source of DNA to study genetic diversity and population structure of macaws in the Amazon Rainforest of Perú. In *Wildlife Resources*. Master of Science. University of Idaho.

Gilbert, M.T.P., Bandelt, H.J., Hofreiter, M. & Barnes, I. 2005. Assessing ancient DNA studies. *Trends Ecol. Evol.* **20**: 541–544.

Goldstein, P.Z., DeSalle, R., Amato, G. & Vogler, A.P. 2000. Conservation genetics at the species boundary. *Conserv. Biol.* **14**: 120–131.

Haffer, J. 1969. Speciation in Amazonian forest birds. *Science* **165**: 131–137.

Haffer, J. 2008. Hypotheses to explain the origin of species in Amazonia. *Braz. J. Biol.* **68**: 917–947.

Haig, S.M., Beever, E.A., Chambers, S.M., Draheim, H.M., Dugger, B.D., Dunham, S., Elliott-Smith, E., Fontaine, J.B., Kesler, D.C., Knaus, B.J., Lopes, I.F., Loschl, P., Mullins, T.D. & Sheffield, L.M. 2006. Taxonomic considerations in listing subspecies under the U.S. Endangered Species Act. *Conserv. Biol.* **20**: 1584–1594.

Hazkani-Covo, E., Zeller, R.M. & Martin, W. 2010. Molecular poltergeists: mitochondrial DNA copies (numts) in sequenced nuclear genomes. *PLoS Genet.* **6**: e1000834.

Hessler, I., Dupont, L., Bonnefille, R., Behling, H., Gonzalez, C., Helmens, K.F., Hooghiemstra, H., Lebamba, J., Ledru, M.P., Lezine, A.M., Maley, J., Marret, F. & Vincens, A. 2010. Millennial-scale changes in vegetation records from tropical Africa and South America during the last glacial. *Quat. Sci. Rev.* **29**: 2882–2899.

Hilty, S.L. & Brown, W.L. 1986. *A Guide to the Birds of Colombia*. Princeton: Princeton University Press.

International Union for Conservation of Nature. 2015. IUCN Red List of Threatened Species. Gland: IUCN.

Johansson, U.S., Ericson, G.P., Blom, M.P.K. & Irestedt, M. 2018. The phylogenetic position of the extinct Cuban macaw *Ara tricolor* based on complete mitochondrial genome sequences. *Ibis* **160**: 666–672.

Juniper, T. & Parr, M. 1998. *Parrots: A Guide to Parrots of the World*. New Haven: Yale University Press.

Karubian, J., Fabara, J., Yunes, D., Jorgenson, J.P., Romo, D. & Smith, T.B. 2005. Temporal and spatial patterns of macaw abundance in the Ecuadorian Amazon. *Condor* **107**: 617–626.

Kimura, M. 1980. A simple method for estimating evolutionary rates of base substitutions through comparative studies of nucleotide sequences. *J. Mol. Evol.* **16**: 111–120.

Kodandaramaiah, U., Simonsen, T.J., Bromilow, S., Wahlberg, N. & Sperling, F. 2013. Deceptive single-locus taxonomy and phylogeography: *Wolbachia*-associated divergence in mitochondrial DNA is not reflected in morphology and nuclear markers in a butterfly species. *Ecol. Evol.* **3**: 5167–5176.

Kumar, S., Nei, M., Dudley, J. & Tamura, K. 2008. MEGA: A biologist-centric software for evolutionary analysis of DNA and protein sequences. *Brief Bioinform.* **9**: 299–306.

Lanfear, R., Frandsen, P.B., Wright, A.M., Senfeld, T. & Calcott, B. 2016. PartitionFinder 2: new methods for selecting partitioned models of evolution for molecular and morphological phylogenetic analyses. *Mol. Biol. Evol.* **34**: 772–773.

Leigh, J.W. & Bryant, D. 2015. PopART: full-feature software for haplotype network construction. *Methods Ecol. Evol.* **6**: 1110–1116.

Li, H. & Durbin, R. 2009. Fast and accurate short read alignment with Burrows-Wheeler transform. *Bioinformatics* **25**: 1754–1760.

Librado, P. & Rozas, J. 2009. DnaSP v5: a software for comprehensive analysis of DNA polymorphism data. *Bioinformatics* **25**: 1451–1452.

Mace, G.M. 2004. The role of taxonomy in species conservation. *Philos. Trans. R. Soc. Lond. B: Biol. Sci.* **359**: 711–719.

Mayr, E. 1963. *Animal Species and Evolution*. Cambridge: Harvard University Press.

McKay, B.D. & Zink, R.M. 2010. The causes of mitochondrial DNA gene tree paraphyly in birds. *Mol. Phylogenet. Evol.* **54**: 647–650.

McKenna, A., Hanna, M., Banks, E., Sivachenko, A., Cibulskis, K., Kernytsky, A., Garimella, K., Altshuler, D., Gabriel, S., Daly, M. & DePristo, M.A. 2010. The genome analysis toolkit: a MapReduce framework for analyzing next-generation DNA sequencing data. *Genome Res.* **20**: 1297–1303.

Monge, O., Schmidt, K., Vaughan, C. & Gutiérrez-Espeleta, G. 2016. Genetic patterns and conservation of the Scarlet Macaw (*Ara macao*) in Costa Rica. *Conserv. Genet.* **17**: 745–750.

Moritz, C. 1994. Defining 'evolutionary significant units' for conservation. *Trends Ecol. Evol.* **9**: 373–375.

Olson, D.M., Dinerstein, E., Wikramanayake, E.D., Burgess, N.D., Powell, G.V.N., Underwood, E.C., D'Amico, J.A., Itoua, I., Strand, H.E., Morrison, J.C., Loucks, C.J., Allnutt, T.F., Ricketts, T.H., Kura, Y., Lamoreux, J.F., Wettengel, W.W., Hedao, P. & Kassem, K.R. 2001. Terrestrial ecoregions of the world: a new map of life on Earth. *Bioscience* **51**: 933–938.

Renton, K. 2002. Seasonal variation in occurrence of macaws along a rainforest river. *J. Field Ornithol.* **73**: 15–19.

Ronquist, F. & Huelsenbeck, J.P. 2003. MrBayes 3: Bayesian phylogenetic inference under mixed models. *Bioinformatics* **19**: 1572–1574.

Seabury, C.M., Dowd, S.E., Seabury, P.M., Raudsepp, T., Brightsmith, D.J., Liboriussen, P., Halley, Y., Fisher,

C.A., Owens, E., Viswanathan, G. & Tizard, I.R. 2013. A multi-platform draft *de novo* genome assembly and comparative analysis for the Scarlet Macaw (*Ara macao*). *PLoS ONE* **8**: e62415.

Seixas, F.A., Boursot, P. & Melo-Ferreira, J. 2018. The genomic impact of historical hybridization with massive mitochondrial DNA introgression. *Genome Biol.* **19**: 91.

Sievers, F., Wilm, A., Dineen, D., Gibson, T.J., Karplus, K., Li, W., Lopez, R., McWilliam, H., Remmert, M., Söding, J., Thompson, J.D. & Higgins, D.G. 2011. Fast, scalable generation of high-quality protein multiple sequence alignments using Clustal Omega. *Mol. Syst. Biol.* **7**: 539.

Stamatakis, A. 2014. RAxML Version 8: a tool for phylogenetic analysis and post-analysis of large phylogenies. *Bioinformatics* **30**: 1312–1313.

Tavares, E.S., Yamashita, C. & Miyaki, C.Y. 2004. Phylogenetic relationships among some neotropical parrot genera (Psittacidae) based on mitochondrial sequences. *Auk* **121**: 230–242.

Toews, D.P. & Brelsford, A. 2012. The biogeography of mitochondrial and nuclear discordance in animals. *Mol. Ecol.* **21**: 3907–3930.

Urantówka, A.D., Kroczak, A. & Mackiewicz, P. 2017. The influence of molecular markers and methods on inferring the phylogenetic relationships between the representatives of the Arini (parrots, Psittaciformes), determined on the basis of their complete mitochondrial genomes. *BMC Evol. Biol.* **17**: 166.

Wandeler, P., Hoeck, P.E.A. & Keller, L.F. 2007. Back to the future: Museum specimens in population genetics. *Trends Ecol. Evol.* **22**: 634–642.

Weir, J.T. & Schluter, D. 2008. Calibrating the avian molecular clock. *Mol. Ecol.* **17**: 2321–2328.

Wiedenfeld, D.A. 1994. A new subspecies of Scarlet Macaw and its status and conservation. *Ornitol. Neotrop.* **5**: 99–104.

Wright, T.F., Rodriguez, A.M. & Fleischer, R.C. 2005. Vocal dialects, sex-biased dispersal, and microsatellite population structure in the parrot *Amazona auropalliata*. *Mol. Ecol.* **14**: 1197–1205.

Zink, R.M. 2004. The role of subspecies in obscuring avian biological diversity and misleading conservation policy. *Proc. R. Soc. B Biol. Sci.* **271**: 561–564.

Received 12 July 2018;
revision accepted 9 June 2019.
Associate Editor: Rheindt Frank.

SUPPORTING INFORMATION

Additional supporting information may be found online in the Supporting Information section at the end of the article.

Table S1. List of specimens included in the present study.

Table S2. Primers for PCR amplification and sequencing.

Table S3. Results from PARTITIONFINDER for the four mitochondrial regions (12S, 16S, COI and CYTB) for Bayesian inference analysis and maximum likelihood analysis.

Table S4. Whole mitochondrial genomes used in analysis (see Fig. 3).

Table S5. Results from PARTITIONFINDER for mitochondrial genomes for maximum likelihood analysis.