Systematics of a radiation of Neotropical suboscines (Aves: Thamnophilidae: *Epinecrophylla*)

Oscar Johnson[a]*, Jeffrey T. Howard[a,b], & Robb T. Brumfield[a]

[a] *Department of Biological Sciences and Museum of Natural Science, Louisiana State University, Baton Rouge, LA 70803, USA*

[b] *Present address: Louisiana State University Health Sciences Center Shreveport, Shreveport, LA 71103, USA*

* Corresponding author

Email addresses: ojohns7@lsu.edu* (Oscar Johnson), jtylerh12@gmail.com (Jeffrey T. Howard), robb@lsu.edu (Robb T. Brumfield).

Abstract

The stipple-throated antwrens of the genus *Epinecrophylla* (Aves: Thamnophilidae) are represented by eight species primarily found in the lowlands of the Amazon Basin and the Guiana Shield. The genus has a long and convoluted taxonomic history, with many attempts made to address the taxonomy and systematics of the group. Here we employ massively parallel sequencing of thousands of ultraconserved elements (UCEs) to provide both the most comprehensive subspecies-level phylogeny of *Epinecrophylla* antwrens and the first population-level genetic analyses for most species in the genus. Most of our analyses are robust to a diversity of phylogenetic and population genetic methods, but we show that even with thousands of loci we are unable to confidently place the western Amazonian taxon *pyrrhonota*. We uncovered phylogenetic relationships between taxa and patterns of population structure that are discordant with both morphology and current taxonomy. For example, we found deep genetic breaks between taxa in the *ornata* group currently regarded as species, and in the *haematonota* and *leucophthalma* groups we found paraphyly at the species and subspecies levels, respectively. Our population genetics analyses showed extensive admixture between some taxa despite their deep genetic divergence. We present a revised taxonomy for the group, discuss the biogeographic patterns that we uncover, and suggest areas for further study.

Keywords: Systematics, Ultraconserved elements, Phylogenomics, *Epinecrophylla*, Amazonia

1. Introduction

The stipple-throated antwrens of the genus *Epinecrophylla* (Isler et al. 2006; Aves: Thamnophilidae) are represented by 21 currently recognized taxa, eight of which are considered species (*E. fulviventris*, *E. ornata*, *E. erythrura*, *E. leucophthalma*, *E. gutturalis*, *E. amazonica*, *E. spodionota*, and *E. haematonota*; Figure 1). These species are primarily found in the lowlands of the Amazon Basin and the Guiana Shield, with one (*E. fulviventris*) found west of the Andes from Ecuador to Honduras (Clements et al., 2019; Zimmer and Isler, 2003). All species are small, insectivorous dead-leaf foraging specialists, typically found in pairs or small family groups in tropical upland forest (Remsen and Parker, 1984; Wiley, 1971). The genus reaches its greatest diversity in the western Amazon Basin, with up to three species broadly co-occurring in most regions, despite similar plumage and foraging behavior between species (Remsen and Parker, 1984; Zimmer and Isler, 2003).

Multiple attempts have been made to resolve relationships in the genus with molecular data, with increasing numbers of loci and individuals used (Hackett and Rosenberg, 1990; Harvey et al., in press; Whitney et al., 2013). Long considered to be in the genus *Myrmotherula*, early molecular work using protein electrophoresis provided the first indication that the stipple-throated antwren complex was not a close relative of other *Myrmotherula* antwrens (Hackett and Rosenberg, 1990). This was further supported with Sanger sequencing of mitochondrial and nuclear loci (Bravo et al., 2014; Brumfield et al., 2007; Irestedt et al. 2004), with the studies finding that *Epinecrophylla* was most closely related to bushbirds in the genera *Neoctantes* and *Clytoctantes*. This work led to the naming of a new genus for the group, *Epinecrophylla* (Isler et al., 2006), with *E. haematonota* as the type species. Some authorities changed the common

65 names of *Epinecrophylla* antwrens to stipplethroats (Clements et al., 2019; Remsen et al., 2019)
66 to reflect this taxonomic rearrangement.
67
68 *1.1 History of taxa within Epinecrophylla*
69
70 The species-level taxonomy of the genus has undergone considerable rearrangement
71 through history (Cory and Hellmayr, 1924; Isler and Whitney, 2018; Peters, 1951; Whitney et
72 al., 2013; Zimmer, 1932a; 1932b), particularly in the *haematonota* and *leucophthalma* groups.
73 Early authors (e.g. Cory and Hellmayr, 1924) considered *E. haematonota* to include as
74 subspecies the taxa *pyrrhonota* and *amazonica* and placed both *spodionota* and *sororia* as
75 subspecies of *E. leucophthalma*, largely based on back color (rufous in the former three taxa,
76 brown in the latter three). Using this same reasoning, Todd (1927), when describing the rufous-
77 backed form *phaeonota*, treated it as a subspecies of *E. haematonota*, but considered *E.*
78 *amazonica* a species distinct from all other forms. Zimmer (1932a), however, noted that back
79 color may not be a valid species-level character and transferred *amazonica, spodionota*, and
80 *sororia* to *E. haematonota*, and *phaeonota* to *E. leucophthalma*. Zimmer (1932a) suggested the
81 possibility of species rank for the rufous-backed taxon *phaeonota*, but also noted intermediate
82 individuals between it and the adjacent brown-backed taxa *leucophthalma* and *sordida*. This
83 treatment was maintained by most authors (e.g. Meyer de Schauensee, 1970; Peters, 1951)
84 until Parker and Remsen (1987) recognized *E. spodionota* (including *sororia*) as a separate
85 species. This taxonomic treatment was augmented by the recent discovery of two range-
86 restricted taxa in the group; *E. fjeldsaai* of eastern Ecuador (Krabbe et al., 1999) and *E. dentei* of

the Aripuanã-Machado region of Brazil (Whitney et al., 2013), each described as a new species. In describing *E. dentei* Whitney et al. (2013) also estimated a mitochondrial phylogeny of the genus, including samples of most taxa, in which they found *fjeldsaai* was phylogenetically embedded within *haematonota*. Based on that work and the mitochondrial distance between taxa, Remsen et al. (2019) separated *E. haematonota* into four species: *E. fjeldsaai* (based on morphology), *E. pyrrhonota*, *E. haematonota*, and *E. amazonica* (including *dentei*), whereas other authors united *pyrrhonota*, *amazonica*, and *dentei* under *E. haematonota* while maintaining *E. fjeldsaai* as a distinct species (Dickinson and Christidis, 2014). Since then, Isler and Whitney (2018) conducted a thorough analysis of the vocalizations of *haematonota*, *fjeldsaai*, and *pyrrhonota* in which they found no vocal differences among the three taxa, leading to the recognition of the latter two taxa as subspecies of the former (Remsen et al., 2019).

Within *E. ornata*, the gray-backed Peruvian taxon *atrogularis* was long considered a separate species, leaving the rufous-backed forms *saturata* and *hoffmannsi* as subspecies of *E. ornata* (Cory and Hellmayr, 1924). This treatment was maintained until Zimmer (1932b) described the gray-backed *meridionalis* as a subspecies and united all five taxa in the group under the species *E. ornata*. This is the current treatment of most recent authors (Clements et al., 2019; Dickinson and Christidis, 2014; Remsen et al., 2019), although del Hoyo et al. (2019) consider *E. hoffmannsi* a species separate from the rest of *E. ornata* based primarily on the female plumage and minor vocal differences.

The taxonomy of the remainder of the genus has remained rather more stable through time, with the three other species *E. fulviventris*, *E. gutturalis,* and *E. erythrura*, all largely

considered independent species by most authors. *E. erythrura* and *E. leucophthalma* are currently considered polytypic, while the four taxa described in *E. fulviventris* are generally considered synonyms of the nominate subspecies (Cory and Hellmayr, 1924; Zimmer and Isler, 2003). We here follow the taxonomy of the South American Classification Committee (Remsen et al., 2019) and make taxonomic recommendations in light of the Biological Species Concept (de Queiroz, 2007; Mayr, 1942).

Much of the previous molecular work in *Epinecrophylla* has relied upon mitochondrial sequence data, although a recent phylogenomic study of all suboscine passerine birds included 1-2 samples of each species of *Epinecrophylla* using sequence capture of ultraconserved elements (UCEs) and recovered a well-resolved topology for the genus (Harvey et al., in press). Here we expand on the previous genetic work in the genus, addressing the systematics of the group with both next-generation sequencing of thousands of nuclear loci and draft mitochondrial genomes, and population-level sampling of most taxa. *Epinecrophylla* provide a unique system in which to study speciation in the Amazon Basin due to their high species diversity, documented phenotypic hybrid zones, and multiple broadly sympatric species. Our expanded sampling both of individuals and loci provides the most in-depth view of the evolutionary history, species limits, population structure, and introgression between taxa in this group.

2. Methods

131 *2.1. Sampling*

132

133 We obtained a total of 66 *Epinecrophylla* and three outgroup samples representing 18
134 of the 21 widely recognized taxa in the genus and all currently recognized species. Missing
135 ingroup taxa are *E. o. ornata*, *E. o. saturata*, and *E. leucophthalma dissita*. The outgroup species
136 we used are *Myrmorchilus strigilatus*, *Neoctantes niger*, and *Clytoctantes atrogularis*. When
137 available, we obtained samples from across the geographic range of each *Epinecrophylla* taxon,
138 with one sample chosen per geographic locality. Fifty-three tissue samples were obtained from
139 vouchered specimens housed at museums in the United States, with sequence data for the
140 remaining 16 samples obtained from Harvey et al. (in press; Table 1).
141 We extracted total DNA from the 53 tissue samples using ca. 25 mg of pectoral muscle
142 with a Qiagen DNeasy Blood and Tissue Kit (Qiagen; Hilden, Germany) and quantified DNA
143 concentration using a Qubit 2.0 fluorometer (Life Technologies Corporation; Carlsbad, CA).
144 Samples were standardized to 10 ng/uL. We then sheared samples to approximately 600 base
145 pair (bp) fragments with an Episonic 1100 bioprocessor (EpiGentek; Farmingdale, NY) and
146 assessed fragment length using a High Sensitivity DNA Assay on an Agilent 2100 Bioanalyzer
147 (Agilent Technologies; Santa Clara, CA). We generated DNA libraries using a KAPA Biosystems
148 Hyper Prep kit (Wilmington, Massachusetts, USA) and enriched UCEs using a set of 5,742
149 probes that target 5,060 loci in vertebrates ("Tetrapods-UCE-5Kv1"; uce-5k-probes.fasta)
150 following the protocol of Faircloth et al. (2012). Enriched samples were pooled at equimolar
151 ratios and paired-end sequencing was conducted on one lane of a HiSeq 3000 sequencer at
152 Oklahoma Medical Research Foundation Clinical Genomics Center (OMRF; Oklahoma City,

153 Oklahoma, USA). The sequencing lane contained DNA libraries used in other projects. The 16
154 samples obtained from Harvey et al. (in press) were enriched using a custom probe set
155 consisting of 2,500 vertebrate UCEs and 96 exons.
156
157 *2.2. Contig assembly*
158
159 OMRF demultiplexed sequence reads using custom scripts. We trimmed raw reads of
160 adapter contamination and low-quality bases using illumiprocessor (Faircloth, 2013) and
161 trimmomatic (Bolger et al., 2014) with default settings. We then subsampled all read files to 2
162 million reads per individual to decrease computation time for contig assembly and to normalize
163 assemblies across samples. Read data were assembled with Itero (https://github.com/faircloth-
164 lab/itero). Because samples were sequenced with two different probe sets, we opted to match
165 contigs to the "Tetrapods-UCE-2.5Kv1" (uce-2.5k-probes.fasta) probe set which consists of
166 2,560 baits targeting 2,386 UCE loci, and is a subset of the other probe sets used in sequencing.
167 For the samples sequenced with the "Tetrapods-UCE-5Kv1" probe set, we also separately
168 matched assembled contigs to this probe set.
169
170 *2.3. Sample identification and locus filtering*
171
172 To confirm the identifications of samples we used the Phyluce 1.6.7 (Faircloth, 2015)
173 tool *match_contigs_to_barcodes* to match contigs from each sample to a mitochondrial
174 Cytochrome c oxidase subunit I (COI) barcode sequence of *Epinecrophylla pyrrhonota* obtained

175 from GenBank (JN801852.1) and map those contigs against the Barcode of Life Database (BOLD;
176 Ratnasingham and Hebert, 2007). We then used the Phyluce 1.6.7 (Faircloth, 2015) tool
177 *get_trinity_coverage* to calculate per-locus read coverage for all contigs matching either UCE
178 and mitochondrial loci. Three samples contained mitochondrial loci with high coverage (>30x)
179 that matched the incorrect species in BOLD, suggesting either sample misidentification or high
180 levels of contamination, and were eliminated from our dataset (Table S1). Nine additional
181 samples contained high-coverage mitochondrial contigs matching the expected species in
182 BOLD, but with a small number of low-coverage mitochondrial contigs matching the incorrect
183 species. We used the maximum coverage of 8.05x of these potentially contaminated low-
184 coverage mitochondrial contigs as a filter and removed all UCE contigs across all samples that
185 had an average read depth below this threshold.
186
187 *2.4. Mitochondrial genome assembly*
188
189 We used off-target reads from the UCE sequencing to assemble draft mitochondrial
190 genomes. We assembled mitochondrial genomes in MITObim 1.9 (Hanh et al., 2013), which is a
191 Perl wrapper for MIRA 4.0.2 (Chevreux et al., 1999), using as a reference the complete
192 mitochondrial genome of *Myrmoderus loricatus* (G. Bravo, unpublished data) and the --*quick*
193 option. We annotated the assembled mitochondrial genomes using the MITOchondrial genome
194 annotation Server (MITOS) 2 (Bernt et al., 2013) and aligned the 13 mitochondrial protein
195 coding genes in MAFFT (Katoh et al., 2002) implemented in Geneious 10.2.3
196 (https://www.geneious.com) to create a final partitioned draft mitogenome alignment.

197

198 *2.5. Nuclear locus phasing, alignment, and SNP calling*

199

200 To phase UCE loci we selected as a reference the individual from our sampling that
201 contained the greatest number of UCE loci after filtering; *Epinecrophylla leucophthalma*
202 LSUMNS 42670. We phased UCE loci using the *seqcap_pop* pipeline
203 (https://github.com/mgharvey/seqcap_pop; Faircloth, 2015; Harvey et al., 2016) to obtain a
204 Single Nucleotide Polymorphism (SNP) dataset and followed Andermann et al. (2019) to
205 obtained phased alignments. The *seqcap_pop* pipeline utilizes tools from the Phyluce package
206 (Faircloth, 2015), SAMtools 0.1.19 (Li et al., 2009), Picard (Broad Institute, Cambridge, MA),
207 BWA 0.7.17 (Li and Durbin, 2009), and GATK 3.3.0 (McKenna et al., 2010) to process next-
208 generation sequence data for population-level genetic analyses. Briefly, *seqcap_pop* maps
209 sequencing reads to the reference individual to obtain a pileup, adds read groups and marks
210 PCR duplicate reads for each individual, merges bam files within each species, calls indels and
211 single-nucleotide polymorphisms on merged bam files, and phases high-quality indels and SNPs
212 to produce vcf files of phased SNPs. We further filtered this dataset using vcftools 0.1.16
213 (Danecek et al., 2011) to remove SNPs with quality scores less than 30 and read depth less than
214 5.5, those with greater than 75% missing data, restricted to bi-allelic loci, and removed indels.
215 We refer to this dataset as the "linked SNP dataset", as it contains multiple SNPs per locus. We
216 then sampled at random one SNP per UCE locus to obtain a final dataset of putatively unlinked
217 SNPs, which we refer to as the "unlinked SNP dataset". To obtain phased alignments we used
218 Phyluce 1.6.7 (Faircloth, 2015) to phase UCE loci following Andermann et al. (2019), phasing

data by mapping reads against the reference individual using the Phyluce tools *snp_bwa_align* and *snp_phase_uces*. This pipeline maps raw reads against contigs of a reference individual using BWA 0.7.17 (Li and Durbin, 2009), and then sorts and phases alleles in SAMtools 0.1.19 (Li and Durbin, 2009) and Picard (Broad Institute, Cambridge, MA). We used MAFFT 7.130b (Katoh and Standley, 2013) in the Phyluce tool *align_seqcap_align* to align and edge-trim the contigs output by this pipeline, treating the two alleles as separate individuals and allowing ambiguous sites in alignments. We produced a final alignment using the Phyluce 1.6.7 (Faircloth, 2015) tool *get_only_loci_with_min_taxa* to produce a 75% complete data matrix. This tool calculates the data matrix completeness as the percentage of individuals in the dataset with sequence data for each locus.

To investigate fine-scale patterns of population structure within each species we called SNPs within each species or species complex to obtain an additional six SNP datasets. We grouped samples based on the clades in the Exabayes phylogeny estimated from the 75% complete UCE data matrix (see section 2.6). Three clades corresponded to species (*ornata*, *leucophthalma*, and *gutturalis*) and a fourth to a set of closely related taxa that have undergone considerable taxonomic rearrangement through history (*dentei*, *amazonica*, *spodionota*, *sororia*, *pyrrhonota*, *haematonota*, and *fjeldsaai*). This latter clade is hereafter referred to as the "*haematonota s.l.*" clade. Within the *haematonota s.l.* clade we additionally subdivided taxa into two clades for SNP calling: one containing *dentei*, *amazonica*, *spodionota*, and *sororia* (hereafter the "*amazonica* + *spodionota*" clade), and the other containing *pyrrhonota*, *haematonota*, and *fjeldsaai* (hereafter the "*haematonota* + *pyrrhonota*" clade). For each

dataset we selected as a reference the individual with the highest number of assembled contigs after filtering (Supplemental Table 4) and repeated the *seqcap_pop* pipeline described above.

2.6. Phylogenetic estimation

From the 2,386 locus UCE dataset we estimated a phylogenetic tree with all samples using a Bayesian analysis in Exabayes 1.5 (Aberer et al., 2014) using the 75% complete concatenated phased alignment. We conducted 4 independent runs for 2 million generations each, discarding the first 25% of trees as burn-in. After checking in Tracer 1.7.1 (Rambaut et al., 2018) that samples had converged based on ESS values greater than 200, we generated an extended majority-rule consensus tree using the topologies from the four independent runs. We used this topology to estimate a time-calibrated phylogenetic tree in BEAST 2.5.2 (Bouckaert et al., 2019). For this analysis we took the tree estimated in Exabayes and trimmed it to include only the samples for which we obtained draft mitochondrial genomes (the Exabayes UCE phylogeny containing all samples is available in Figure S1), and one allele per individual. We then made this tree ultrametric using the *chronos* function in ape 5.2 (Paradis and Schliep, 2018) and used the resulting topology as a constraint tree through the entire BEAST analysis. No fossils are available for *Epinecrophylla* or its close relatives to allow for a fossil calibration of our phylogenetic tree. We therefore used three combinations of mitochondrial alignments and substitution rates to estimate the branch lengths for this topology. We first used the concatenated alignment of the 13 mitochondrial protein-coding genes and the widely used mitochondrial cytochrome B (CytB) 2.1% substitution rate obtained from Hawaiian

honeycreepers (Lerner et al., 2011; substitution rate: 0.01105 substitutions/site/Myr, 95% Confidence Interval [CI]: 0.00425–0.01785). The second method used the alignment of cytochrome B (CytB) extracted from the draft mitochondrial genomes and the same CytB substitution rate. For the final analysis we analyzed the third codon position of CytB and used a mass-calibrated substitution rate following (Nabholz et al., 2016) using the third codon position and their calibration set 2 (substitution rate: 0.0208 substitutions/site/Myr, 95% CI: 0.0170–0.0263). Body mass estimates were obtained from specimens at the Louisiana State University Museum of Natural Science (mean = 9.30 g, SD = 1.18, n = 229), representing all species in the genus. We used the date estimate from a dated phylogeny of all suboscine passerines to constrain the root of our phylogeny (Harvey et al., in press), and set *E. fulviventris* as the outgroup, with the date of 9.28 Ma (95% CI: 8.60–11.07 Ma) as the prior on the root node. For each analysis we used the GTR + γ model of rate variation and a proportion of invariant sites, a log-normal relaxed clock, a yule model on the tree, and default settings for other priors. We ran each analysis for 100 million generations, sampling every 10,000, and a burn-in of 10%. We checked in Tracer 1.7.1 (Rambaut et al., 2018) that all parameters reached convergence with ESS values over 200. Both analyses using the CytB alignment were unable to converge and had multiple priors with ESS values well below 200, indicating inappropriate priors and/or insufficient generations. The analysis using the draft mitochondrial genome alignment, however, converged with ESS values over 200 for all parameters. For the mitogenome analysis we calculated a maximum clade credibility (MCC) tree in TreeAnnotator 2.5.1 from the posterior of trees, implemented in BEAST 2.5.2 (Bouckaert et al., 2019), and visualized the resulting tree in FigTree 1.4.4 (Rambaut, 2009).

We used SNAPP 1.4.2 (Bryant et al., 2012) implemented in BEAST 2.5.2 (Bouckaert et al., 2019) to calculate a species tree directly from SNP data in a full-coalescent analysis. This site-based method has the advantage of bypassing gene tree estimation and minimizing error due to the low information content of individual UCE loci. Initial runs using all samples in our dataset were unable to converge in a reasonable amount of time, either with individuals as separate tips or with individuals pooled to tips by clade identified from the Exabayes 75% UCE phylogeny. Therefore, we called SNPs again using two sample-filtering and locus-filtering strategies, all with *Myrmorchilus strigilatus* as an outgroup; the first using up to two samples from each clade identified in the Exabayes 75% phylogeny for a total of 19 samples, and the second using one sample from each widely recognized species in the genus for a total of 11 samples. In both datasets we used the individual from each taxon that had the greatest number of loci recovered to maximize the number of SNPs recovered and allowed 5% missing data. Otherwise we followed the *seqcap_pop* and SNP filtering pipeline described in section 2.5. After SNP filtering we selected at random one SNP per locus to minimize issues with linkage within a locus. In all runs we calculated the mutation rates from the data and used default priors. We ran all analyses for 2 million generations, storing every 1,000 generations, and a burn-in of 10%, and checked that the run converged in Tracer 1.7.1 (Rambaut et al., 2018) based on ESS values over 200. From the posterior of species trees we calculated a maximum clade credibility (MCC) tree in TreeAnnotator 2.5.1 implemented in BEAST 2.5.2 (Bouckaert et al., 2019). We used Densitree 2.0.1 (Bouckaert, 2010) to visualize the posterior tree set of the SNAPP runs and FigTree 1.4.4 (Rambaut, 2009) to visualize the MCC tree.

305 In addition to the analyses outlined above, we conducted a variety of phylogenetic
306 analyses, each with its own assumptions, strengths, and weaknesses relative to treating sources
307 of phylogenetic variation. Details and results for these analyses are available in the
308 Supplementary Materials.
309
310 *2.7. Population genetics and introgression*
311
312 In addition to our phylogenetic analyses we utilized SNPs to investigate patterns of
313 population-level genetic structure and also introgression within and between clades. We used
314 STRUCTURE (Pritchard et al., 2000) and Discriminant Analysis of Principal Components (DAPC)
315 to analyze patterns of population structure within each clade, and implemented each analysis
316 on all six clade-level SNP datasets described above. For STRUCTURE we analyzed the linked SNP
317 datasets and implemented the *linked* model, providing the distance in base pairs between SNPs
318 within each locus, and ran analyses for 2 million generations, discarding the first 50,000 as
319 burn-in. We ran 10 replicate analyses for each value of K from one to 10, or until the likelihood
320 value of K decreased significantly. We selected the best K value based on the ΔK method of
321 Evanno (Evanno et al., 2005), implemented in STRUCTURE HARVESTER (Earl and vonHoldt,
322 2012).
323 DAPC uses sequential k-means clustering of principal components to infer the number
324 of genetic clusters in a dataset. We conducted a DAPC analysis in *adegenet* (Jombart and
325 Ahmed, 2011), following the recommendations of Jombart et al. (2010), and selected the best
326 number of clusters based on the lowest Bayesian Information Criterion (BIC) score. In addition,

we conducted a Principal Components Analysis (PCA) analysis, with samples coded by DAPC group assignments. Although BIC scores for the *haematonota s.l.* clade indicated that K values greater than 2 had a worse fit to the data than K=2, we calculated DAPC group assignments for K values from 3–5 to investigate finer-scale patterns of population structure, due to the greater number of described taxa in this clade.

We calculated pairwise distance estimates between all taxa in the group from both the mitochondrial and nuclear DNA data. For the mitochondrial distances we used the alignment of the 13 mitochondrial protein-coding genes. On a neighbor-joining tree reconstructed from the raw *p*-distance matrix in PAUP* 4.0 (Swofford, 1999), we estimated the proportion of invariant sites (0.590355) and the gamma shape parameter (1.82626). These values were then fixed for calculations of a distance matrix under the GTR + γ + I finite-sites substitution model. For the nuclear data we estimated the weighted fixation index (F_{st}) between each pair of taxa using the method of Weir and Cockerham (1984) implemented in vcftools 0.1.16 (Danecek et al., 2011) using the unlinked SNP dataset. For all calculations we also report the average within-taxon distance estimates as a measure of intra-specific genetic structure.

3. Results

3.1. Sequencing results and sample identification

Illumina sequencing of UCEs resulted in an average of 3.8 million reads per individual, and an average read length of 130 bp after trimming. After removing potentially contaminated

or misidentified samples our dataset contained 63 *Epinecrophylla* samples and two outgroups. Including the three potentially contaminated *Epinecrophylla* samples (based on BOLD results) in a phylogenetic tree estimated in RAxML 8.2.12 (Stamatakis, 2014), two grouped with the correct taxon but sat on abnormally long terminal branches, suggesting contamination, and a third grouped with the outgroup samples, suggesting sample misidentification (Figure S1H). After assembly and locus filtering we obtained an average of 2,195 UCE loci per sample (range 1,234—2,306 loci), with a mean locus length of 652 bp (range 234–1,283 bp) and mean read depth of UCE loci of 22.5x (SD: 43.0x). Missing data had a strong effect on the number of UCE loci retained in the alignment, and the alignment that including no missing data contained 330 UCE loci and was not analyzed further. The 95% complete phased alignment contained 1,659 UCE loci and an aligned matrix of 1,140,275 bp and the 75% complete phased alignment contained 2,149 UCE loci and an aligned matrix of 1,401,699 bp.

We obtained partial or complete mitochondrial genomes for 59 ingroup samples and both outgroups, including at least one sample per species (Table S2). Three samples, including one of the outgroups, contained greater than 40% missing data and were removed from the analysis (Table S2). The average mitochondrial genome size was 17,253 bp (range 16,017–17, 930 bp) and had a mean read depth of 304.4x (SD: 780.0x). The aligned dataset of 56 individuals using the 13 mitochondrial protein-coding genes was 11,488 bp in length (range 9,921–11,396 bp), contained a total of 635,429 bp, and 4.8% missing data.

3.2. Phylogenetic estimation

From the nuclear UCE data, we recovered a phylogeny with strong support for relationships among taxa (Figure 2, Figure 3, Figure S1). The deepest split in the tree occurred across the Andes, dividing *E. fulviventris* from the remainder of the genus. Although our sampling included just two samples of *E. fulviventris*, one of which (sample 1) is from a population occasionally separated as the subspecies *costaricensis* (del Hoyo et al., 2019; Todd, 1927), our phylogenies indicated a relatively shallow divergence between the two samples (Figure S1). The next split separated *E. ornata* from the remaining taxa (Figure 1A, Figure 2). Although we lacked samples for two taxa within this species (*E. o. ornata* and *E. o. saturata*), the two parapatric gray-backed taxa occurring in Peru (*E. o. meridionalis* and *E. o. atrogularis*) are recovered as non-sister lineages, with the southern *meridionalis* sister to the rufous-backed *hoffmannsi* of eastern Brazil, and *atrogularis* sister to those two. The next split contained the sister species *E. erythrura* and *E. leucophthalma*, which together are sister to the remaining taxa (Figure 1B, Figure 2). These two species are reciprocally monophyletic, but within *E. leucophthalma* we recovered the nominate subspecies as paraphyletic. Within this nominate subspecies of *E. leucophthalma*, the western populations (samples 13–16) showed a deep divergence from the remainder of the species. The final clade contained eight parapatric taxa (*gutturalis*, *dentei*, *amazonica*, *spodionota*, *sororia*, *pyrrhonota*, *haematonota*, and *fjeldsaai*) that together range across the majority of the Amazon Basin (Figure 1C, Figure 2). The Guiana Shield taxon *E. gutturalis* was sister to the rest of the taxa in this clade, but contains minimal genetic structure in the phylogeny (Figure 2). The remaining taxa can be divided into three groups with similar divergence times between them (Figure 2). The first group contained the southeastern Amazonian *E. amazonica* (including the subspecies *dentei*) and the Andean

393 foothill *E. spodionota* (including the subspecies *sororia*), the second is the northwestern
394 Amazonian taxon *E. haematonota pyrrhonota*, and the third is the western Amazonian *E.*
395 *haematonota haematonota* (including the subspecies *fjeldsaai*). The taxon *fjeldsaai* was
396 embedded within *haematonota* in all analyses. Our dated phylogeny placed *pyrrhonota* sister to
397 *haematonota*, with *amazonica* and *spodionota* sister to those. Analyses of nuclear data with a
398 variety of phylogenetic methods (see Supplementary Material) largely supported the above
399 results. However, some phylogenetic analyses of nuclear and mitochondrial data indicated
400 support for two alternate topologies with regard to the placement of *E. h. pyrrhonota* (Figure
401 4), with varying degrees of node support (Figure S1, S2). Branch lengths in the dated phylogeny
402 produced some results slightly discordant with those from other phylogenetic analyses,
403 suggesting that the oldest splits within *E. ornata* and *E. leucophthalma* are as old or older than
404 some of the species-level splits within the *haematonota s.l.* clade (Figure 2).
405 The site-based MCC phylogeny estimated in SNAPP using one sample per species
406 produced the same inter-specific topology as that recovered in the dated phylogeny, but with
407 low support for two nodes (Figure 3B). The Densitree representation of the posterior of species
408 trees showed that these nodes contained a primary topology the same as that recovered in the
409 other nuclear analyses, but with minor alternate topologies at nodes with lower support in the
410 MCC tree (Figure 3A). These are the same nodes that showed low support in the mitochondrial
411 phylogeny or in which the mitochondrial phylogeny differed from the nuclear phylogeny (Figure
412 S1, S2). The expanded SNAPP analysis using 19 samples recovered the same topology in the
413 MCC tree, but with much lower support for multiple nodes (Figure S4B). This was reflected in
414 considerable uncertainty in the Densitree representation of the posterior of species trees,

which showed many alternate topologies both near the root of the tree and with regards to the placement of *pyrrhonota* (Figure S4A).

3.3. Population genetics

DAPC analyses with k-means cross-validation estimated a best fit model of K=2 for each of three clades: *E. leucophthalma*, *E. ornata*, and *haematonota s.l.*, and a model of K=1 for *E. gutturalis* (Figure 5). For *E. leucophthalma* this divided the species into the "*leucophthalma* west" clade (samples 13–16) and the remainder of the eastern taxa, including the "*leucophthalma* east" clade. For *E. ornata*, DAPC separated the central Peruvian *atrogularis* from the two eastern taxa. Lastly, for the *haematonota s.l.* group, the best fit model of K=2 separated *pyrrhonota* from the remainder of the group. The worse-fit models of K>2 (based on BIC scores) for the *haematonota s.l.* clade first separated the *E. spodionota* + *E. amazonica* clade at K=3, then *E. spodionota* from *E. amazonica* at K=4, and the western-most sample (#54) of *E. h. pyrrhonota* at K=5. PCA plots with points labeled by taxon and sample number are shown in Figure S5.

STRUCTURE results (Figure 5) largely recapitulated those from DAPC but provided a more in-depth view of individuals with potential genetic backgrounds from multiple populations (i.e. potential introgression). Results from the Evanno method based on the ΔK value were unambiguous in all cases. However, in all cases the STRUCTURE plot for the "best" value of K from the Evanno method added a population that was evenly assigned across all individuals. Therefore, we consider the STRUCTURE plot for the "best" K minus 1 to be a more

437 biologically realistic representation of the data and report all STRUCTURE plots >1 that have
438 high likelihood values, following the recommendation of Meirmans (2015). Because the Evanno
439 method is unable to calculate a ΔK value at K=1 and because all individuals of *E. gutturalis* were
440 approximately equally assigned to both populations at K=2, we consider a K=1 to be the best-fit
441 model for that species. For *E. leucophthalma*, K=2 separated the "*leucophthalma* west" clade
442 from the remainder of the eastern taxa, but with all individuals containing a small percentage of
443 genetic membership from the other clade. Within *E. ornata*, results were similar to those of
444 DAPC, with *atrogularis* the most distinct at K=2, but with all individuals having a proportion of
445 their ancestry assigned to both populations. The pattern in the STRUCTURE plots for the
446 *haematonota s.l.* clade was more complex. At K=2, STRUCTURE assignments largely separated
447 *E. h. pyrrhonota* from *E. spodionota* + *E. amazonica*, with all individuals of the *E. h.*
448 *haematonota* + *E. h. fjeldsaai* group having about equal ancestry between the two groups. This
449 pattern was also reflected also in the intermediate position of the *E. h. haematonota* + *E. h.*
450 *fjeldsaai* group along the first principal component of the PCA results. At K=3 STRUCTURE
451 separated three groups that corresponded to taxonomy, with most individuals showing only a
452 small proportion of shared population assignments: 1) *E. h. pyrrhonota*, 2) *E. spodionota* + *E.*
453 *amazonica*, and 3) *E. h. haematonota* + *E. h. fjeldsaai*. The "best" value of K=4 provided only a
454 slight suggestion of differentiation between *E. spodionota* and *E. amazonica*, with *E. a. dentei*
455 genetically indistinguishable from *E. a. amazonica* and *E. s. sororia* genetically indistinguishable
456 from *E. s. spodionota*.
457
458 4. DISCUSSION

459

460 *4.1. Phylogeny and population genetics*

461 Our analyses of nuclear and mitochondrial data largely resolved the evolutionary
462 relationships among *Epinecrophylla* taxa and recovered three broadly sympatric species
463 complexes in the Amazon Basin. We consider the topology of the phylogeny illustrated in Figure
464 2 to be the best representation of relationships in the group based on the consistently high
465 support values across multiple methods that recovered this topology (Figure 3, Figure 4, Figure
466 S1A-D, Figure S1G). Although there was some disagreement among methods and data types
467 regarding the placement of the taxon *pyrrhonota* (Figure 4), most of the topologies that
468 disagreed with the sister relationship of *pyrrhonota* and *haematonota* received low support for
469 that node, often in conjunction with a very short subtending branch (Figure S1E-F). These short
470 branches suggest that the divergence between the three clades in the *haematonota s.l.* clade
471 was likely very rapid, which may explain the conflicting signal across methods and the support
472 for alternate topologies in the SNAPP posterior distributions (Figure 3A, Figure S4A). Because a
473 strictly bifurcating tree is likely not an appropriate model for intraspecific relationships in cases
474 of high levels of gene flow (Eckert and Carstens, 2008), our population genetics results may be a
475 better representation of the evolutionary relationships among individuals at this fine scale. The
476 evolutionary patterns recovered with these population genetics analyses indicate little
477 differentiation between most subspecies and even between some taxa currently recognized as
478 species (e.g. *E. spodionota* and *E. amazonica*). The three Amazonian species complexes that we
479 recovered in our phylogeny are sympatric across much of the western Amazon Basin, but are
480 represented by one species each in the eastern parts of the Basin. These species complexes are

1) *E. ornata*, 2) *E. leucophthalma* and *E. erythrura*, and 3) *E. gutturalis* and the "*haematonota s.l.*" group. Each complex contains taxa that are either allopatric or largely parapatric, with distributions typically bounded by large rivers (Figure 1). We discuss the phylogenetic results and taxonomic implications of each species separately.

4.2. Epinecrophylla fulviventris

Epinecrophylla fulviventris was recovered as sister to the remainder of the genus in all analyses and with relatively shallow divergence between our two samples in most phylogenies. We lack the geographic sampling or morphological data to make any taxonomic recommendations for this species, and thus suggest maintaining the current treatment of a monotypic *E. fulviventris* (e.g. Clements et al., 2019; Zimmer and Isler, 2003).

4.3. Epinecrophylla ornata

The only predominantly gray-bodied species in the genus, our results for this morphologically distinctive group are hampered by the lack of samples of the nominate taxon and the geographically adjacent taxon *saturata*. The species was described from a "Bogota" skin and is thus of uncertain provenance, although typically assumed to be from the lowlands of southern Amazonian Colombia (Peters, 1951). However, without samples from Ecuador or southern Colombia, we are unable to fully resolve the relationships within this species or recommend taxonomic changes. Despite the lack of these samples, we discovered deep splits

and high population structure among all three subspecies in our phylogenetic analyses, suggesting that multiple species-level taxa occur in the group. The most genetically distinct of the three taxa in phylogenetic and population genetic analyses was *E. o. atrogularis* of central Peru, which all analyses placed as sister to *E. o. meridionalis* + *E. o. hoffmannsi*, thus contradicting the opinion of some authors (e.g. del Hoyo et al., 2019) that *hoffmannsi* represents a species distinct from the other four taxa in *E. o. ornata*. This relationship is surprising given the phenotypic similarity of *atrogularis* and *meridionalis*, which both lack the rufous back of the other three taxa in the *ornata* group and differ from each other primarily in the slightly duller underparts of female *atrogularis*. *E. o. atrogularis* and *E. o. meridionalis* potentially come into contact in the Ucayali Region of southern Peru, and further research on this contact zone is of interest given the deep genetic split between the two taxa presented here. Reports of specimens of *meridionalis* with some rufous on the back from southern Peru in Cusco and Madre de Dios have been suggested to be either variation within that taxon or evidence of introgression with one of the rufous-backed forms (Ridgely and Tudor, 1994). Based on the population genetic results (Figure 5D) presented here we suspect that the latter is a more likely explanation given that our STRUCTURE results show individuals with shared population assignments between all three subspecies that we sampled, despite the deep genetic splits among them.

4.4. Epinecrophylla leucophthalma and *E. erythrura*

524 Ours is the first study to suggest a sister relationship between these two species. The split
525 between the two species is quite deep, and the two species have largely parapatric
526 distributions (Figure 1B), but are locally sympatric in Peru (Schulenberg et al., 2007; Álvarez
527 Alonso, 2002) without showing any morphological signs of introgression, thus confirming their
528 status as species. Notably, we found that the nominate subspecies of *E. leucophthalma* is
529 paraphyletic as currently recognized, with western populations of *E. leucophthalma* sister to a
530 group containing the subspecies *sordida* and *phaeonota* and the eastern populations of the
531 nominate subspecies. This deep genetic divergence within *E. leucophthalma* is comparable to
532 some divergences among taxa considered to be species within the *haematonota s.l.* clade, in
533 particular between *E. spodionota* and *E. amazonica*, but to our knowledge no morphological
534 characters have been proposed to diagnose this western population. Excluding that western
535 population, the remainder of the *E. leucophthalma* samples in our analysis showed extremely
536 low divergence among them, although most of our nuclear phylogenies grouped samples into
537 the subspecies *sordida*, *phaeonota*, and the eastern clade of *leucophthalma*. Because the
538 geographically intermediate taxon *phaeonota* is morphologically distinctive (rufous back in
539 *phaeonota* vs brown in all other *leucophthalma* taxa), we support the continued treatment of
540 the three *leucophthalma* taxa as separate subspecies despite the lack of genetic divergence
541 between them. The type locality of *E. leucophthalma* is on the right bank of the Madeira River
542 at Salto do Jirau, Rondônia, Brazil (Pelzeln and Natterer, 1871), in the same interfluve and just
543 150 km to the north of our sample #20 (Figure 1B), thus suggesting that the name
544 *leucophthalma* applies to the eastern clade of *E. l. leucophthalma* and that the Madeira River
545 may correspond to the deep genetic break within the species. The southern Andean foothill

546 taxon *dissita* comes into contact with our western clade of *E. l. leucophthalma* in southern Peru
547 (Figure 1B), so the name *dissita* could potentially be expanded to include the rest of Peru and
548 Pando, Bolivia (i.e. our "*leucophthalma* west" clade). Alternatively, a new name might be
549 necessary for the western population of *E. l. leucophthalma*. However, genetic samples of
550 *dissita* are needed to confirm which of these alternative treatments is appropriate. Estimates of
551 the finite-sites mitochondrial distance and weighted nuclear F_{st} between the western clade of *E.*
552 *l. leucophthalma* and all eastern populations are 6.3% and 0.29, respectively (Table S3, S4).
553
554 *4.5. Epinecrophylla haematonota* group
555
556 This clade contains eight taxa that have undergone many taxonomic rearrangements
557 throughout their history (Cory and Hellmayr, 1924; del Hoyo et al., 2019; Dickinson and
558 Christidis, 2014; Isler and Whitney, 2018; Peters, 1951; Remsen et al., 2019; Whitney et al.,
559 2013; Zimmer, 1932a; Zimmer and Isler, 2003). Our phylogenetic analyses indicate that *E.*
560 *gutturalis* is part of this species complex and is sister to the rest of the clade. All of our
561 phylogenetic and STRUCTURE analyses showed no population structure within *E. gutturalis*
562 across its range. DAPC results showed low levels of structure, but still indicated a K=1 based on
563 BIC scores. The five samples that showed slight divergence from the main cluster in the PCA
564 results (Figure 5C; sample numbers 32, 34, 35, 36, and 40) did not cluster based on geography.
565 While there was some disagreement among analyses on the relationships between the rest of
566 the taxa in this group, particularly with regard to the placement of *pyrrhonota*, most of our
567 analyses agree with the topology in Figure 2. Following that topology, the next split in this clade

568 presents an interesting biogeographic pattern, separating the western lowland Amazonian taxa
569 *pyrrhonota*, *haematonota*, and *fjeldsaai*, from a clade containing the Andean foothill taxa
570 *spodionota* and *sororia* and the southeastern Amazonian lowland *amazonica* and *dentei*. The
571 only known sympatry between any taxa in this group occurs on the east slope of the Andes in
572 southern Colombia where Salaman et al. (2002) reported *pyrrhonota* and *spodionota* being
573 captured in the same mist-nets, thus necessitating at a minimum the separation of *pyrrhonota*
574 and *E. amazonica* + *E. spodionota* as biological species. Given the similar divergence times
575 between *pyrrhonota*, *E. amazonica* + *E. spodionota*, and *E. haematonota* + *E. fjeldsaai* we
576 suspect that these three lineages likely represent separate biological species.
577 Isler et al. (1998) developed a yardstick-based system to evaluate species limits in
578 Thamnophilidae based on vocalizations and applied this to *haematonota*, *pyrrhonota*, and
579 *fjeldsaai*, finding that the three taxa did not differ in vocalizations and were best regarded as
580 three subspecies of *E. haematonota* (Isler and Whitney, 2018). Our results suggest that the
581 divergence between *pyrrhonota* and *E. amazonica* + *E. spodionota* (mtDNA 6.1%, weighted F_{st}
582 0.34) is comparable to that between *pyrrhonota* and *E. h. haematonota* + *E. h. fjeldsaai* (mtDNA
583 5.7%%, weighted F_{st} 0.25), and some analyses indicating a closer relationship between the
584 former groups than the latter (Figure 4), albeit with weak support. This, combined with the
585 results from our DAPC and STRUCTURE analyses which indicate that *pyrrhonota* is the most
586 divergent taxon in this group, and in particular more so than *E. amazonica* is from *E.*
587 *spodionota*, that *E. pyrrhonota* is best regarded as a distinct species. A more thorough
588 evaluation of the utility of this yardstick system between all *Epinecrophylla* taxa is desirable,

589 although the results presented here indicate that it may not be a reliable indicator of species
590 status in all cases.

591 Our results support the continued treatment of *fjeldsaai* as a subspecies of *E.*
592 *haematonota* due its morphological distinctiveness, but were more ambiguous with regard to
593 the taxonomic status of *dentei*. None of our analyses were able to differentiate *fjeldsaai* from
594 *haematonota*, and all phylogenetic analyses indicated that *fjeldsaai* was embedded within
595 *haematonota*. This treatment is further supported by evidence of hybridization between the
596 two taxa in northwestern Peru (LSUMNS specimens). In fact, one of our samples of *fjeldsaai*
597 (sample 62) has some rufous coloration on the lower back suggestive of hybridization. *E.* a.
598 *dentei* was placed as sister to *E. a. amazonica* in all phylogenetic analyses, but with relatively
599 shallow divergence (mtDNA 2.8%, weighted F_{st} 0.19). Our population genetic analyses including
600 all samples in the *haematonota s.l.* clade were unable to distinguish *amazonica* and *dentei*
601 (Figure 5), and in fact were largely unable to differentiate *E. amazonica* and *E. spodionota*. A
602 DAPC analysis using just the three samples of *amazonica* and the one of *dentei* did suggest a
603 K=2 was the best model based on BIC scores and separated the *dentei* sample from the rest of
604 *amazonica* (Figure S6).

605 In summary, we recommend the following 4-species treatment for the taxa in the
606 *haematonota* group: *E. haematonota* (with *fjeldsaai* as a subspecies), *E. pyrrhonota*, *E.*
607 *amazonica* (with *dentei* as a subspecies), and *E. spodionota* (with *sororia* as a subspecies). Until
608 genetic samples of *E. leucophthalma dissita*, *E. o. ornata*, and *E. o. saturata* are available for
609 study, we refrain from making taxonomic recommendations in those groups, but we suspect
610 that both *E. leucophthalma* and *E. ornata* contain multiple species-level taxa. Therefore, we

recommend the following species-level linear taxonomy for *Epinecrophylla*: *fulviventris*, *ornata*, *erythrura*, *leucophthalma*, *gutturalis*, *haematonota*, *pyrrhonota*, *amazonica*, *spodionota*.

4.6. Biogeographic patterns

Having three broadly sympatric species or species complexes distributed across the Amazon Basin provides replicated evolutionary histories across a shared landscape. Of interest is the response of each of these species or species complexes to well-known biogeographic barriers in the Amazon Basin, such as large rivers (Capparella, 1991; Wallace, 1854). The major river systems of the Amazon Basin, such as the Solimões, Negro, and Madeira all appear to have an effect on the genetic structure and range limits of *Epinecrophylla* taxa, delimiting many species and subspecies that show significant genetic breaks at those locations in our analyses. Smaller river systems, however, appear to have idiosyncratic effects on genetic structure, with some delimiting genetic groups in one species, but having little to no effect in others. For example, the Purús River is a major barrier for *E. ornata*, but has little effect on the genetic structure of other groups, while the Tapajos River is a barrier for *E. leucophthalma* but not *E. ornata*. Additionally, the distribution and genetic boundaries of phenotypically distinctive taxa such as *fjeldsaai* and *phaeonota* do not appear to always follow biogeographic barriers that affect other bird species. The brown-backed *fjeldsaai*, which we find to be phylogenetically embedded within the rufous-backed *haematonota*, hybridizes with *haematonota* within the Napo interfluve without any clear biogeographic barrier separating the two taxa. Likewise, the rufous-backed *phaeonota* is part of the otherwise brown-backed *leucophthalma* group, and

633 appears to replace the eastern populations of nominate *leucophthalma* somewhere between
634 the Juruena and Roosevelt Rivers.

635 Although all three *Epinecrophylla* species complexes overlap in much of the western
636 Amazon Basin, there is evidence that competition may play a role in their ability to coexist over
637 broad spatial scales. For example, J. Tobias (personal communication) found *E. leucophthalma*
638 and *E. amazonica* regularly at the same site in Madre de Dios, Peru, but the two species rarely
639 occurred in the same mixed species flock. Likewise, in Napo, Ecuador, Whitney (1994) found *E.*
640 *ornata* and *E. erythrura* in the same mixed species flock on just one occasion, although three
641 species of *Epinecrophylla* occurred at the site.

642 That two species complexes – *E. ornata* and *E. leucophthalma* + *E. erythrura* – are absent
643 from the Guiana Shield and the northern half of the Inambari interfluve (Figure 1A, 1B) is
644 perplexing. This pattern may be due to the vagaries of extinction, interspecific competition, or
645 habitat suitability, or some combination of those factors, all of which require further study. For
646 example, suboptimal habitat may increase competition between such closely related and
647 ecologically similar species, leading to local extinctions. Alternatively, more drastic climate
648 fluctuations in the eastern half of the basin may be driving this pattern. There is evidence that
649 the eastern Amazon Basin became much drier during the glacial periods of the Pleistocene,
650 whereas climate of the western Amazon Basin has maintained more stable over the same time
651 period, which is thought to have resulted in relatively higher losses of biodiversity in the
652 eastern half of the basin (Cheng et al., 2013). A north-south gradient in vegetation composition
653 exists in the Inambari interfluve (Tuomisto et al., 2019) and the Guiana Shield is relatively drier
654 than much of the western Amazon (Fick and Hijmans, 2017), so these species may be unable to

655 persist in these areas. Likewise, the *haematonota* group is absent from the Brazilian Shield,
656 while the other two species complexes are present there.
657
658 *4.7. Areas of potential future research*
659
660 The results of our phylogenetic and population genetic analyses suggest that multiple
661 geographic regions could produce valuable insights with both greater sampling effort and
662 natural history observations (e.g. playback experiments, surveys of contact zones, analysis of
663 vocal and morphological traits). The first is in southern Peru in the foothills of southeastern
664 Madre de Dios region, where three taxa potentially come into close geographic proximity,
665 namely *spodionota* and *amazonica*, which we recover as sister taxa in our phylogenies, and
666 *haematonota*. A second region of interest is slightly to the north in southern Ucayali region,
667 where two subspecies of *ornata*; *atrogularis* and *meridionalis*, replace each other, perhaps
668 across the Purús River. These two taxa are recovered as non-sister taxa in our phylogenies and
669 could perhaps come into contact across the headwaters of that river. Genetic samples of the
670 two northern taxa in the *ornata* group, including the type taxon, are critical to resolving
671 relationships within that clade. A third region is the headwaters of the Rio Napo in northern
672 Ecuador, where two taxa currently regarded as subspecies of *haematonota*; *pyrrhonota* and
673 *fjeldsaai*, could potentially come into contact. Analysis of a contact zone in this region would be
674 critical to resolving species limits in the *haematonota* group.
675 Despite being the most well-sampled phylogenetic study of *Epinecrophylla* to date, our
676 study lacked genetic samples from some key areas that could affect the results presented here.

677 The lack of samples for two subspecies of *E. ornata*, including the nominate, hinders our ability
678 to make any taxonomic recommendations for that species. We also lack samples of *E.*
679 *leucophthalma dissita* of the Yungas. This taxon comes into contact with the western clade of
680 nominate *leucophthalma*, and it is possible that the name *dissita* could apply to the entirety of
681 this clade of western *leucophthalma*, a population that based on our results may represent a
682 species distinct from the eastern three taxa in the *leucophthalma* group. It is worth noting that
683 additional taxa have been named within *E. leucophthalma*, but the validity of those taxa has
684 been questioned as they are generally considered not morphologically diagnosable (del Hoyo et
685 al., 2019). Two other sampling gaps bear mention; the first is the population of *E. h.*
686 *haematonota* from the north bank of the Amazon west of the Napo, which is the population
687 that presumably intergrades with *E. h. fjeldsaai*, and the second is a lack of samples for any
688 taxon from the vast region of the Brazilian Amazon west of the Madeira River and south of the
689 Amazon River, which could contain genetically distinct populations and contains the type
690 locality of *E. amazonica* (Peters, 1951).
691
692 **5. Conclusions**
693
694 As has been shown in other Neotropical avian systems (Brumfield, 2005; Musher and
695 Cracraft, 2018), our study highlights the importance of sampling populations below the species
696 level, especially in tropical regions, where the taxonomy of many groups is unresolved and
697 there may be considerable undiscovered morphological and genetic diversity. Our
698 understanding of phylogenetic relationships has grown dramatically in recent decades as

699 technological advances have allowed us to obtain and analyze sequence data for ever more
700 genetic markers and individuals, including at the population level in non-model organisms
701 (Harris et al., 2018; Tan et al., 2019; Zarza et al., 2016; Zucker et al., 2016).
702
703 **Acknowledgements**
704
705 OJ was supported by the National Science Foundation Graduate Research Fellowship under
706 grant no. DGE-1247192. JTH was supported by the National Science Foundation Research
707 Experiences for Undergraduates under grant no. DEB-1146265. We thank the curators and staff
708 at the following institutions for providing tissue samples: Gary Graves and Christopher Milensky
709 at the Smithsonian National Museum of Natural History, Richard Prum and Kristof Zyskowski at
710 the Yale Peabody Museum, Christopher Witt at the Museum of Southwestern Biology, Joel
711 Cracraft and Paul Sweet at the American Museum of Natural History, John Bates and Ben Marks
712 at the Field Museum of Natural History, and Rob Moyle and Mark Robbins at the University of
713 Kansas Biodiversity Institute & Natural History Museum. Donna Dittmann assisted with tissue
714 sample processing at the Louisiana State University Museum of Natural Science. Van Remsen,
715 Daniel F. Lane, Nelson Buainain Neto, and members of the Brumfield Lab provided invaluable
716 feedback on versions of this manuscript. Marco A. Rego provided assistance with the creation
717 of Figure 1.
718
719 **References**
720

Aberer, A.J., Kobert, K., Stamatakis, A., 2014. ExaBayes: massively parallel Bayesian tree inference for the whole-genome era. Mol Biol Evol 31, 2553-2556. 10.1093/molbev/msu236.

Álvarez Alonso, J., 2002. Characteristic avifauna of white-sand forests in northern Peruvian Amazonia. Baton Rouge, Louisiana, Louisiana State University.

Andermann, T., Fernandes, A.M., Olsson, U., Topel, M., Pfeil, B., Oxelman, B., Aleixo, A., Faircloth, B.C., Antonelli, A., 2019. Allele phasing greatly improves the phylogenetic utility of ultraconserved elements. Syst Biol 68, 32-46. 10.1093/sysbio/syy039.

Bernt, M., Donath, A., Jühling, F., Externbrink, F., Florentz, C., Fritzsch, G., Pütz, J., Middendorf, M., Stadler, P.F., 2013. MITOS: Improved de novo metazoan mitochondrial genome annotation. Mol Phylogenet Evol 69, 313-319. 10.1016/j.ympev.2012.08.023.

Bolger, A.M., Lohse, M., Usadel, B., 2014. Trimmomatic: a flexible trimmer for Illumina sequence data. Bioinformatics 30, 2114-2120. 10.1093/bioinformatics/btu170.

Bouckaert, R., 2010. DensiTree: making sense of sets of phylogenetic trees. Bioinformatics 26, 1372-1373. 10.1093/bioinformatics/btq110.

Bouckaert, R., Vaughan, T.G., Barido-Sottani, J., Duchene, S., Fourment, M., Gavryushkina, A., Heled, J., Jones, G., Kuhnert, D., De Maio, N., Matschiner, M., Mendes, F.K., Muller, N.F., Ogilvie, H.A., du Plessis, L., Popinga, A., Rambaut, A., Rasmussen, D., Siveroni, I., Suchard, M.A., Wu, C.H., Xie, D., Zhang, C., Stadler, T., Drummond, A.J., 2019. BEAST 2.5: An advanced software platform for Bayesian evolutionary analysis. PLoS Comput Biol 15, e1006650. 10.1371/journal.pcbi.1006650.

Bravo, G.A., Remsen, J.V., Jr., Brumfield, R.T., 2014. Adaptive processes drive ecomorphological convergent evolution in antwrens (Thamnophilidae). Evolution 68, 2757-2774. 10.1111/evo.12506.

Brumfield, R.T., 2005. Mitochondrial variation in Bolivian populations of the Variable Antshrike (*Thamnophilus caerulescens*). The Auk 122, 414-432.

Brumfield, R.T., Tello, J.G., Cheviron, Z.A., Carling, M.D., Crochet, N., Rosenberg, K.V., 2007. Phylogenetic conservatism and antiquity of a tropical specialization: army-ant-following in the typical antbirds (Thamnophilidae). Mol Phylogenet Evol 45, 1-13. 10.1016/j.ympev.2007.07.019.

Bryant, D., Bouckaert, R., Felsenstein, J., Rosenberg, N.A., RoyChoudhury, A., 2012. Inferring species trees directly from biallelic genetic markers: bypassing gene trees in a full coalescent analysis. Mol Biol Evol 29, 1917-1932. 10.1093/molbev/mss086.

Capparella, A.P., 1991. Neotropical avian diversity and riverine barriers. Acta Congressus Internationalis Ornithologici 20, 307-316.

Cheng, H., Sinha, A., Cruz, F.W., Wang, X., Edwards, R.L., d'Horta, F.M., Ribas, C.C., Vuille, M., Stott, L.D., Auler, A.S., 2013. Climate change patterns in Amazonia and biodiversity. Nat Commun 4, 1411. 10.1038/ncomms2415.

Chevreux, B., Wetter, T., Suhai, S., 1999. Genome sequence assembly using trace signals and additional sequence information computer science and biology. Proceedings of the German Conference on Bioinformatics 99, 45-56.

764 Clements, J.F., Schulenberg, T.S., Iliff, M.J., Billerman, S.M., Fredericks, T.A., Sullivan, B.L.,
765 Wood, C.L., 2019. The eBird/Clements Checklist of Birds of the World: v2019.
766 Downloaded from https://www.birds.cornell.edu/clementschecklist/download/
767 Cory, C.B., Hellmayr, C.E., 1924. Catalogue of birds of the Americas and the adjacent islands.
768 Part III. Pteroptochidae, Conopophagidae, Formicariidae. Field Museum of
769 Natural History Zoological Series Vol. XIII.
770 Danecek, P., Auton, A., Abecasis, G., Albers, C.A., Banks, E., DePristo, M.A., Handsaker, R.E.,
771 Lunter, G., Marth, G.T., Sherry, S.T., McVean, G., Durbin, R., Group, G.P.A., 2011.
772 The variant call format and VCFtools. Bioinformatics 27, 2156-2158.
773 10.1093/bioinformatics/btr330.
774 de Queiroz, K., 2007. Species concepts and species delimitation. Syst Biol 56, 879-886.
775 10.1080/10635150701701083.
776 del Hoyo, J., Elliott, A., Sargatal, J., Christie, D.A., Kirwan, G.e., 2019. Handbook of the Birds of
777 the World Alive. Lynx Edicions, Barcelona. (retrieved from http://www.hbw.com/
778 on 13/11/2019).
779 Dickinson, E.C., Christidis, L.E., 2014. The Howard & Moore complete checklist of the birds of
780 the world. 4th Edition. Aves Press, Eastbourne, U.K.
781 Earl, D.A., vonHoldt, B.M., 2012. STRUCTURE HARVESTER: a website and program for visualizing
782 STRUCTURE output and implementing the Evanno method. Conservation
783 Genetics Resources 4, 359-361.
784 Eckert, A.J., Carstens, B.C., 2008. Does gene flow destroy phylogenetic signal? The performance
785 of three methods for estimating species phylogenies in the presence of gene
786 flow. Mol Phylogenet Evol 49, 832-842. 10.1016/j.ympev.2008.09.008.
787 Evanno, G., Regnaut, S., Goudet, J., 2005. Detecting the number of clusters of individuals using
788 the software STRUCTURE: a simulation study. Mol Ecol 14, 2611-2620.
789 10.1111/j.1365-294X.2005.02553.x.
790 Faircloth, B.C., 2013. illumiprocessor: a trimmomatic wrapper for parallel adapter and quality
791 trimming. 10.6079/J9ILL.
792 Faircloth, B.C., 2015. PHYLUCE is a software package for the analysis of conserved genomic loci.
793 Bioinformatics 32, 786-788. 10.1093/bioinformatics/btv646.
794 Faircloth, B.C., McCormack, J.E., Crawford, N.G., Harvey, M.G., Brumfield, R.T., Glenn, T.C.,
795 2012. Ultraconserved elements anchor thousands of genetic markers spanning
796 multiple evolutionary timescales. Syst Biol 61, 717-726. 10.1093/sysbio/sys004.
797 Fick, S.E., Hijmans, R.J., 2017. Worldclim 2: New 1-km spatial resolution climate surfaces for
798 global land areas. International Journal of Climatology.
799 Hackett, S.J., Rosenberg, K.V., 1990. Comparison of phenotypic and genetic differentiation in
800 South American Antwrens (Formicariidae). The Auk 107, 473-489.
801 Hahn, C., Bachmann, L., Chevreux, B., 2013. Reconstructing mitochondrial genomes directly
802 from genomic next-generation sequencing reads—a baiting and iterative
803 mapping approach. Nucleic Acids Research 41, e129-e129. 10.1093/nar/gkt371.
804 Harris, R.B., Alström, P., Ödeen, A., Leache, A.D., 2018. Discordance between genomic
805 divergence and phenotypic variation in a rapidly evolving avian genus
806 (*Motacilla*). Mol Phylogenet Evol 120, 183–195.

807 Harvey, M.G., Bravo, G.A., Claramunt, S., Cuervo, A.M., Derryberry, G.E., Battilana, J., Seeholzer,
808 G.F., Shearer, M.J., Faircloth, B.C., Edwards, S.V., Pérez-Emán, J.L., Moyle, R.G.,
809 Sheldon, F.H., Aleixo, A., Smith, B.T., Chesser, R.T., Silveira, L.F., Cracraft, J.,
810 Brumfield, R.T., Derryberry, E.P., in press. The evolution of a tropical biodiversity
811 hotspot.
812 Harvey, M.G., Smith, B.T., Glenn, T.C., Faircloth, B.C., Brumfield, R.T., 2016. Sequence capture
813 versus restriction site associated DNA sequencing for shallow systematics. Syst
814 Biol 65, 910-924. 10.1093/sysbio/syw036.
815 Irestedt, M., Fjeldsa, J., Nylander, J.A., Ericson, P.G., 2004. Phylogenetic relationships of typical
816 antbirds (Thamnophilidae) and test of incongruence based on Bayes factors.
817 BMC Evol Biol 4, 23. 10.1186/1471-2148-4-23.
818 Isler, M.L., Isler, P.R., Whitney, B.M., 1998. Use of vocalizations to establish species limits in
819 antbirds (Passeriformes: Thamnophilidae). The Auk 115, 577-590.
820 10.2307/4089407.
821 Isler, M.L., Rodrigues Lacerda, D., Isler, P.R., Hackett, S.J., Rosenberg, K.V., Brumfield, R.T., 2006.
822 *Epinecrophylla*, a new genus of antwrens (Aves: Passeriformes: Thamnophilidae).
823 Proc Biol Soc Wash 119, 522-527.
824 Isler, M.L., Whitney, B.M., 2018. Reevaluation of the taxonomic positions of members of the
825 *Epinecrophylla fjeldsaai* (Aves: Passeriformes: Thamnophilidae) antwren complex
826 including *E. fjeldsaai* based on vocalizations. The Wilson Journal of Ornithology
827 130, 908-914.
828 Jombart, T., Ahmed, I., 2011. adegenet 1.3-1: new tools for the analysis of genome-wide SNP
829 data. Bioinformatics 27, 3070-3071. 10.1093/bioinformatics/btr521.
830 Jombart, T., Devillard, S., Balloux, F., 2010. Discriminant analysis of principal components: a
831 new method for the analysis of genetically structured populations. BMC Genetics
832 11, 94. 10.1186/1471-2156-11-94.
833 Katoh, K., Misawa, K., Kuma, K., Miyata, T., 2002. MAFFT: a novel method for rapid multiple
834 sequence alignment based on fast Fourier transform. Nucleic Acids Research 30,
835 3059-3066. 10.1093/nar/gkf436.
836 Katoh, K., Standley, D.M., 2013. MAFFT multiple sequence alignment software version 7:
837 improvements in performance and usability. Mol Biol Evol 30, 772-780.
838 10.1093/molbev/mst010.
839 Krabbe, N., Isler, M.L., Isler, P.R., Whitney, B.M., Alvarez A., J., Greenfield, P.J., 1999. A new
840 species in the *Myrmotherula haematonota* superspecies (Aves; Thamnophilidae)
841 from the western Amazonian lowlands of Ecuador and Peru. The Wilson Bulletin 111, 157-165.
842 Lerner, H.R., Meyer, M., James, H.F., Hofreiter, M., Fleischer, R.C., 2011. Multilocus resolution
843 of phylogeny and timescale in the extant adaptive radiation of Hawaiian
844 honeycreepers. Curr Biol 21, 1838-1844. 10.1016/j.cub.2011.09.039.
845 Li, H., Durbin, R., 2009. Fast and accurate short read alignment with Burrows–Wheeler
846 transform. Bioinformatics 25, 1754-1760. 10.1093/bioinformatics/btp324.
847 Li, H., Handsaker, B., Wysoker, A., Fennell, T., Ruan, J., Homer, N., Marth, G., Abecasis, G.,
848 Durbin, R., Subgroup, G.P.D.P., 2009. The Sequence Alignment/Map format and
849 SAMtools. Bioinformatics 25, 2078-2079. 10.1093/bioinformatics/btp352.
850 Mayr, E., 1942. Systematics and the origin of species. Columbia University Press, New York.

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 research 20, 1297-1303. 10.1101/gr.107524.110.

Meirmans, P.G., 2015. Seven common mistakes in population genetics and how to avoid them. Mol Ecol 24, 3223–3231.

Meyer de Schauensee, R., 1970. A guide to the birds of South America. Livingston Publishing Company, Wynnewood, Pennsylvania.

Musher, L.J., Cracraft, J., 2018. Phylogenomics and species delimitation of a complex radiation of Neotropical suboscine birds (*Pachyramphus*). Mol Phylogenet Evol 118, 204-221. 10.1016/j.ympev.2017.09.013.

Nabholz, B., Lanfear, R., Fuchs, J., 2016. Body mass-corrected molecular rate for bird mitochondrial DNA. Mol Ecol 25, 4438-4449. 10.1111/mec.13780.

Paradis, E., Schliep, K., 2018. ape 5.0: an environment for modern phylogenetics and evolutionary analyses in R. Bioinformatics.

Parker, T.A., III, Remsen, J.V., Jr., 1987. Fifty-two Amazonian bird species new to Bolivia. Bull. B.O.C. 107, 94-107.

Pelzeln, A.v., Natterer, J., 1871. Zur ornithologie Brasiliens: resultate von Johann Natterers reisen in den jahren 1817 bis 1835. Wien: A. Pichler's Witwe & Sohn.

Peters, J.L., 1951. Check-list of birds of the world, vol. 7. . Museum of Comparative Zoölogy, Cambridge, Massachusetts.

Pritchard, J.K., Stephens, M., Donnelly, P., 2000. Inference of population structure using multilocus genotype data. Genetics 155, 945–959.

Rambaut, A., 2009. FigTree v1.4.4. Available via http://tree.bio.ed.ac.uk/software/figtree/. Accessed 25/11/2018.

Rambaut, A., Drummond, A.J., Xie, D., Baele, G., Suchard, M.A., 2018. Posterior summarization in Bayesian phylogenetics using Tracer 1.7. Syst Biol 67, 901-904. 10.1093/sysbio/syy032.

Ratnasingham, S., Hebert, P.D.N., 2007. BOLD: The Barcode of Life Data System (http://www.barcodinglife.org). Molecular Ecology Notes 7, 355-364. 10.1111/j.1471-8286.2007.01678.x.

Remsen, J.V., Jr., Areta, J.I., Bonaccorso, E., Claramunt, S., Jaramillo, A., Pacheco, J.F., Ribas, C., Robbins, M.B., Stiles, F.G., Stotz, D.F., Zimmer, K.J., 2019. A classification of the bird species of South America. American Ornithological Society. http://www.museum.lsu.edu/~Remsen/SACCBaseline.htm.

Remsen, J.V., Jr., Parker, T.A., III, 1984. Arboreal dead-leaf searching birds of the Neotropics. The Condor 86, 36-41.

Ridgely, R.S., Tudor, G., 1994. The birds of South America. Vol II. The suboscine passerines. University of Texas Press, Austin, Texas.

Salaman, P.G.W., Stiles, F.G., Isabel Bohórquez, C., Álvarez-R., M., María Umaña, A., Donegan, T.M., Cuervo, A.M., 2002. New and noteworthy bird records from the east slope of the Andes of Colombia. Caldasia 24, 157-189.

Schulenberg, T.S., Stotz, D.F., Lane, D.F., O'Neill, J.P., Parker, T.A.I., 2007. Birds of Peru. Princeton University Press, Princeton, New Jersey.

Stamatakis, A., 2014. RAxML version 8: a tool for phylogenetic analysis and post-analysis of large phylogenies. Bioinformatics 30, 1312-1313. 10.1093/bioinformatics/btu033.

Swofford, D.L., 1999. Phylogenetic Analysis Using Parsimony, PAUP* 4.0, b166. Sinauer Associates, Boston, Massachusetts.

Tan, H.Z., Ng, E.Y.X., Tang, Q., Allport, G.A., Jansen, J., Tomkovich, P.S., Rheindt, F.E., 2019. Population genomics of two congeneric Palaearctic shorebirds reveals differential impacts of Quaternary climate oscillations across habitats types. Sci Rep 9, 18172. 10.1038/s41598-019-54715-9.

Todd, W.E.C., 1927. New gnateaters and antbirds from tropical America, with a revision of the genus Myrmeciza and its allies. Proc. Biol. Soc. Washington 40, 149-177.

Tuomisto, H., Van doninck, J., Ruokolainen, K., Moulatlet, G.M., Figueiredo, F.O.G., Sirén, A., Cárdenas, G., Lehtonen, S., Zuquim, G., 2019. Discovering floristic and geoecological gradients across Amazonia. Journal of Biogeography 46, 1734-1748. 10.1111/jbi.13627.

Wallace, A.R., 1854. On the Monkeys of the Amazon. Journal of Natural History Series 2 14, 451-454. 10.1080/037454809494374.

Weir, B.S., Cockerham, C.C., 1984. Estimating F-statistics for the analysis of population structure. Evolution 38, 1358-1370.

Whitney, B.M., 1994. Behavior, vocalizations, and possible relationships of four *Myrmotherula* antwrens (Formicariidae) from eastern Ecuador. The Auk 111, 469-475.

Whitney, B.M., Isler, M.L., Bravo, G.A., Aristizábal, N., Schunck, F., Fábio Silveira, L., de Q. Piacentini, V., 2013. A new species of Epinecrophylla antwren from the Aripuanã-Machado interfluvium in central Amazonian Brazil with revision of the “stipple-throated antwren” complex. Handbook of the birds of the World, original scientific descriptions, 263-267.

Wiley, R.H., 1971. Cooperative roles in mixed flocks of antwrens (Formicariidae). The Auk 88, 881-892.

Zarza, E., Faircloth, B.C., Tsai, W.L., Bryson, R.W., Jr., Klicka, J., McCormack, J.E., 2016. Hidden histories of gene flow in highland birds revealed with genomic markers. Mol Ecol 25, 5144-5157. 10.1111/mec.13813.

Zimmer, J.T., 1932a. Studies of Peruvian Birds. III. The genus *Myrmotherula* in Peru, with notes on extralimital forms. Part 1. American Museum Novitates 523, 1-19.

Zimmer, J.T., 1932b. Studies of Peruvian Birds. IV. The genus *Myrmotherula* in Peru, with notes on extralimital forms. Part 2. American Museum Novitates 524, 1-16.

Zimmer, K.J., Isler, M.L., 2003. Family Thamnophilidae (typical antbirds). In: del Hoyo, J., Elliott, A., Christie, D.A. (Eds.), Handbook of the Birds of the World. Lynx Edicions, Barcelona, pp. 448-681.

Zucker, M.R., Harvey, M.G., Oswald, J.A., Cuervo, A., Derryberry, E., Brumfield, R.T., 2016. The Mouse-colored Tyrannulet (*Phaeomyias murina*) is a species complex that includes the Cocos Flycatcher (*Nesotriccus ridgwayi*), an island form that underwent a population bottleneck. Mol Phylogenet Evol 101, 294-302. 10.1016/j.ympev.2016.04.031.

939

940 Table 1. Localities for samples used in this project. Abbreviations are as follows: LSUMNS =
941 Louisiana State University Museum of Natural Science, KU = University of Kansas Biodiversity
942 Institute & Natural History Museum, AMNH = American Museum of Natural History, MZUSP =
943 Museum of Zoology of the University of São Paulo, FMNH = Field Museum of Natural History,
944 MSB = Museum of Southwestern Biology, USNM = Smithsonian National Museum of Natural
945 History, YPM = Yale Peabody Museum. Probe set refers to the probe set used in sequencing. 5k
946 = Tetrapods-UCE-5Kv1 probe set targeting 5,060 loci, and sequenced for this study. 2.5k =
947 Tetrapods-UCE-2.5Kv1 probe set targeting 2,386 loci, sequences obtained from Harvey et al. (in
948 press).

949

950

951

952

953

954

955

956

957

958

Sample #	Taxon	Tissue #	Probe set	Locality	Latitude	Longitude
1	*Epinecrophylla fulviventris*	LSUMNS 82086	5k	Costa Rica: Limón; Reserva Biológica Hitoy-Cerere	9.65	-83.01
2	*E. fulviventris*	LSUMNS 2299	2.5k	Panamá: Darién; Cana	7.92	-77.70
3	*E. ornata atrogularis*	MSB 36505	5k	Perú: San Martín; ca 2.7 km S of Plataforma	-7.41	-76.27
4	*E. ornata atrogularis*	LSUMNS 74213	2.5k	Perú: Pasco; Provincia Oxapampa, Distrito Puerto Bermúdez, Comunidad San Juan	-10.50	-74.81
5	*E. ornata hoffmannsi*	LSUMNS 78113	2.5k	Brazil: Amazonas; Barra de São Manuel	-7.50	-58.26
6	*E. ornata hoffmannsi*	FMNH 391379	5k	Brazil: Pará; Serra dos Carajás	-6.28	-50.58
7	*E. ornata hoffmannsi*	FMNH 457051	5k	Brazil: Pará; Portel, FLONA do Caxiuanã, Plot PPBIO	-2.53	-50.85
8	*E. ornata meridionalis*	LSUMNS 9502	5k	Bolivia: Pando; Nicolás Suarez, 12 km by road S of Cobija, 8 km W on road to Mucden	-11.16	-68.78
9	*E. ornata meridionalis*	LSUMNS 1082	5k	Bolivia: La Paz; Río Beni, ca 20 km by river N Puerto Linares	-15.28	-67.50
10	*E. ornata meridionalis*	LSUMNS 78808	5k	Perú: Madre de Dios; Portillo, ca 7 km S Iberia	-11.45	-69.52
11	*E. erythrura erythrura*	ANSP 16560	2.5k	Ecuador: Morona-Santiago; Santiago	-2.72	-78.32
12	*E. erythrura septentrionalis*	LSUMNS 27716	5k	Perú: Loreto; 79 km WNW Contamana	-7.15	-75.69
13	*E. leucophthalma leucophthalma* west	LSUMNS 42670	5k	Perú: Loreto; ca 54 km NNW mouth of Río Morona on W bank	-4.29	-77.24

14	*E. leucophthalma leucophthalma* west	LSUMNS 10538	5k	Perú: Ucayali; W bank Río Shesha, 65 km ENE Pucallpa	-7.95	-74.25
15	*E. leucophthalma leucophthalma* west	LSUMNS 9173	5k	Bolivia: Pando; Nicolás Suarez, 12 km by road S of Cobija, 8 km W on road to Mucden	-11.16	-68.78
16	*E. leucophthalma leucophthalma* west	LSUMNS 75006	5k	Perú: Ucayali; Otorongo, 31.9 km ESE mouth of Río Cohengua	-10.38	-73.72
17	*E. leucophthalma sordida*	FMNH 392048	5k	Brazil: Mato Grosso do Norte; Municipio Alta Floresta, upper Rio Teles Pires-Rio Cristalino	-9.63	-55.93
18	*E. leucophthalma sordida*	FMNH 457026	5k	Brazil: Pará; Portel, FLONA do Caxiuanã, Plot PPBIO	-2.53	-50.85
19	*E. leucophthalma leucophthalma* east	FMNH 389907	5k	Brazil: Rondônia; Cachoeira Nazaré, W bank Río Jiparaná	-10.36	-61.82
20	*E. leucophthalma leucophthalma* east	LSUMNS 36628	5k	Brazil: Rondônia; Reserva Biológica Rebid Duro Preto, ca 70 km E Guajará-Mirim	-10.80	-64.69
21	*E. leucophthalma leucophthalma* east	LSUMNS 18242	2.5k	Bolivia: Santa Cruz; Provincia Velasco; PN Noel Kempff Mercado, 86 km ESE Florida	-14.85	-60.46
22	*E. leucophthalma leucophthalma* east	LSUMNS 14575	5k	Bolivia: Santa Cruz; Serranía de Huanchaca, 21 km SE Catarata Arco Iris	-13.92	-60.82
23	*E. leucophthalma leucophthalma* east	LSUMNS 12394	5k	Bolivia: Santa Cruz; Provincia Velasco, 32 km E Aserradero Moira, pre PN Noel Kempff Mercado	-14.60	-60.92
24	*E. leucophthalma phaeonota*	LSUMNS 85998	5k	Brazil: Amazonas; Río Sucunduri	-6.89	-59.07
25	*E. leucophthalma phaeonota*	LSUMNS 78380	5k	Brazil: Amazonas; Río Juruena	-11.05	-58.65
26	*E. leucophthalma phaeonota*	LSUMNS 77807	5k	Brazil: Amazonas; Barra de São Manuel, W bank Río Tapajós	-7.34	-58.09
27	*E. leucophthalma phaeonota*	LSUMNS 80818	5k	Brazil: Amazonas; right bank of Río Sucunduri, Igarapé da Cabaça	-5.70	-59.16

28	*E. leucophthalma phaeonota*	LSUMNS 35603	5k	Brazil: Pará; ca 139 km WSW Santarém, W of Río Tapajós, Alto Río Arapiuns	-3.60	-55.52
29	*E. leucophthalma phaeonota*	LSUMNS 80774	5k	Brazil: Amazonas; left bank of Río Sucunduri, left bank lower Río Acari (7 km from its mouth)	-7.15	-59.91
30	*E. gutturalis*	AMNH 12689	5k	Venezuela: Amazonas; Cerro de La Neblina base camp	1.3	-66.5
31	*E. gutturalis*	YPM 139781	5k	Suriname: upper West River Valley, Wilhelmina Mountains	3.75	-56.50
32	*E. gutturalis*	LSUMNS 71576	5k	Suriname: Sipaliwini; Nassau Bebergte	4.78	-54.60
33	*E. gutturalis*	KU 88804	5k	Guyana: Iwokrama Reserve, W bank Essequibo River, ca 72 river km SW Kurupukari	4.22	-59.17
34	*E. gutturalis*	KU 88801	5k	Guyana: Iwokrama Reserve, ca 41 road km SW Kurupukari	4.34	-58.85
35	*E. gutturalis*	USNM 609157	5k	Guyana: Essequibo Islands; West Demerara, Waruma River, E bank, ca 15 river km S Kako River	5.5	-60.6
36	*E. gutturalis*	USNM 586379	5k	Guyana: Northwest District; Baramita	7.35	-60.35
37	*E. gutturalis*	AMNH 11921	5k	Venezuela: Bolívar	6.2	-63.6
38	*E. gutturalis*	YPM 137211	5k	Suriname: Sipaliwini; Werehpai	3.65	-56.20
39	*E. gutturalis*	LSUMNS 20398	2.5k	Brazil: Amazonas; Manaus, km 34 ZF-3 Faz Esteio ca 80 km N Manaus	-2.44	-59.89
40	*E. gutturalis*	LSUMNS 55218	2.5k	Suriname: Sipaliwini; Balchuis Gebergte, ca 70 km SE Apura	4.73	-56.75

41	*E. gutturalis*	USNM 587338	5k	Guyana: Acari Mountains, N side	2.05	-57.55
42	*E. gutturalis*	YPM 139633	5k	Suriname: Para District	5.4	-55.2
43	*E. gutturalis*	YPM 101670	5k	Suriname: Tafelberg	3.78	-56.15
44	*E. amazonica dentei*	MZUSP 80591	2.5k	Brazil: Amazonas; Río Roosevelt, Esperança	-8.33	-60.99
45	*E. amazonica amazonica*	LSUMNS 9217	2.5k	Bolivia: Pando; Nicolás Suarez, 12 km by road S of Cobija, 8 km W on road to Mucden	-11.20	-68.78
46	*E. amazonica amazonica*	MZUSP J164	2.5k	Brazil: Rondônia; left bank Río Madeira, near mouth of Río Abunã	-9.63	-65.45
47	*E. amazonica amazonica*	LSUMNS 31342	5k	Brazil: Rondônia; ca 50 km NW Jaci Paraná, W bank of Río Madeira	-8.93	-64.10
48	*E. spodionota sororia*	KU 113634	5k	Perú: Cusco; ca Alto Manguriari	-13.53	-71.97
49	*E. spodionota sororia*	LSUMNS 2058	5k	Perú: Pasco; Puellas, km 41 on Villa Rica - Puerto Bermúdez highway	-10.29	-74.94
50	*E. spodionota sororia*	LSUMZ 76377	5k	Perú: Ucayali; north ridge of Quebrada Quirapokiari watershed	-10.45	-74.12
51	*E. spodionota spodionota*	IAvH-BT 234	2.5k	Colombia: Cauca; Santa Rosa, Serranía de Los Churumbelos, Río Alto Hornoyaco	1.70	-76.57
52	*E. spodionota sororia*	FMNH 474124	5k	Perú: Amazonas; Río Verde	-6.72	-77.43
53	*E. spodionota sororia*	LSUMNS 5392	5k	Perú: San Martín; 20 km by road NE Tarapoto on road to Yurimaguas	-6.36	-76.24
54	*E. haematonota pyrrhonota*	LSUMNS 4202	2.5k	Perú: Loreto; Lower Río Napo region, E bank Río Yanayacu, ca 90 km N Iquitos	-2.96	-73.24

55	*E. haematonota pyrrhonota*	FMNH 457014	5k	Brazil: Amazonas; Maraã, Lago Cumapi	-1.68	-65.83
56	*E. haematonota pyrrhonota*	AMNH 14224	5k	Brazil: Amazonas; Estrada Manacapuru-Novo Airão km 75	-3.29	-60.63
57	*E. haematonota pyrrhonota*	MZUSP 79027	2.5k	Brazil: Roraima; Paracaima, Comunidade Nova Esperança	4.43	-61.13
58	*E. haematonota pyrrhonota*	LSUMNS 7505	5k	Venezuela: Amazonas; Cerro De La Neblina, Camp VII	0.88	-65.99
59	*E. haematonota haematonota*	LSUMNS 75291	5k	Perú: Ucayali; Otorongo, 31.9 km ESE mouth of Río Cohengua	-10.38	-73.72
60	*E. haematonota haematonota*	LSUMNS 4579	2.5k	Perú: Loreto; S Río Amazonas, 10 km SSW mouth Río Napo on E bank Quebrada Vainilla	-3.52	-72.81
61	*E. haematonota haematonota*	LSUMNS 10790	5k	Perú: Ucayali; W bank Río Shesha, 65 km ENE Pucallpa	-7.94	-74.23
62	*E. haematonota fjeldsaai hybrid?*	LSUMNS 42704	5k	Perú: Loreto; ca 54 km NNW mouth of Morona on W bank	-4.29	-77.24
63	*E. haematonota fjeldsaai*	KU 873	2.5k	Perú: Loreto; San Jacinto, 1.5 km E Río Tigre	-2.32	-75.86
64	*E. haematonota haematonota*	LSUMNS 93087	5k	Perú: Loreto; Esperanza, E of Río Huallaga on Río Yuracyacu, 14.2 km E Santa Cruz	-5.54	-75.73
65	*E. haematonota haematonota*	LSUMNS 27427	5k	Perú: Loreto; NE bank upper Río Cushabatay, 84 km WNW Contamana	-7.07	-75.70
66	*E. haematonota haematonota*	LSUMNS 42282	5k	Perú: Loreto; 7km SW Jeberos	-5.34	-76.32
outgroup	*Clytoctantes atrogularis*	MZUSP 96888	2.5k	Brazil: Amazonas; Río Sucunduri (right bank) 60 km below BR-230 (point 9)	-6.25	-59.07
outgroup	*Myrmorchilus strigilatus*	LSUMNS 18722	5k	Bolivia: Santa Cruz; Provincia Cordillera, Estancia Perforación, ca 130 km E Charagua	-19.78	-61.97

outgroup	*Neoctantes niger*	LSUMNS 2749	2.5k	Perú: Loreto; 1 km N Río Napo, 157 km by river NNE Iquitos	-3.39	-73.18

959

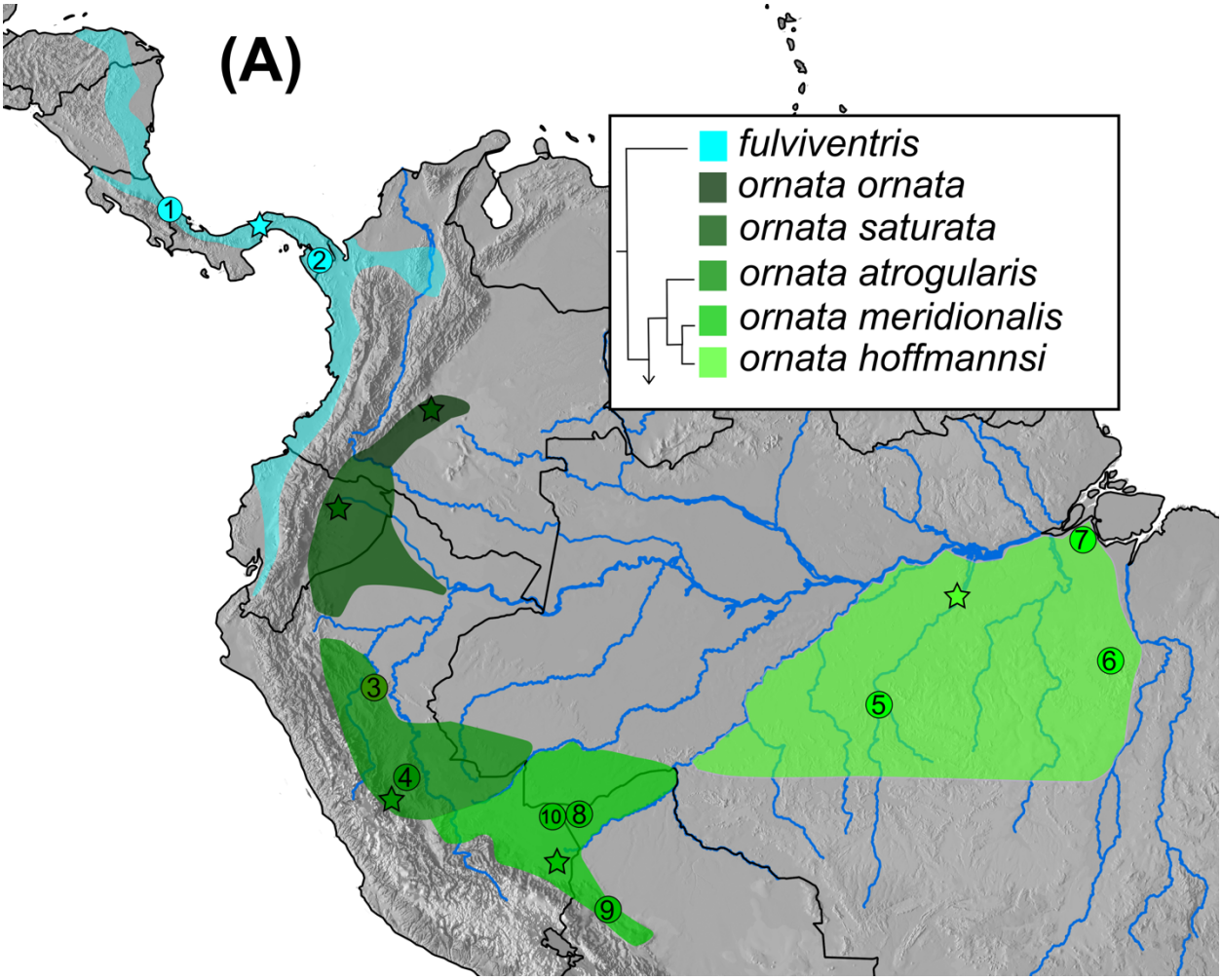
960

961

962

963 FIGURES

964



965

966

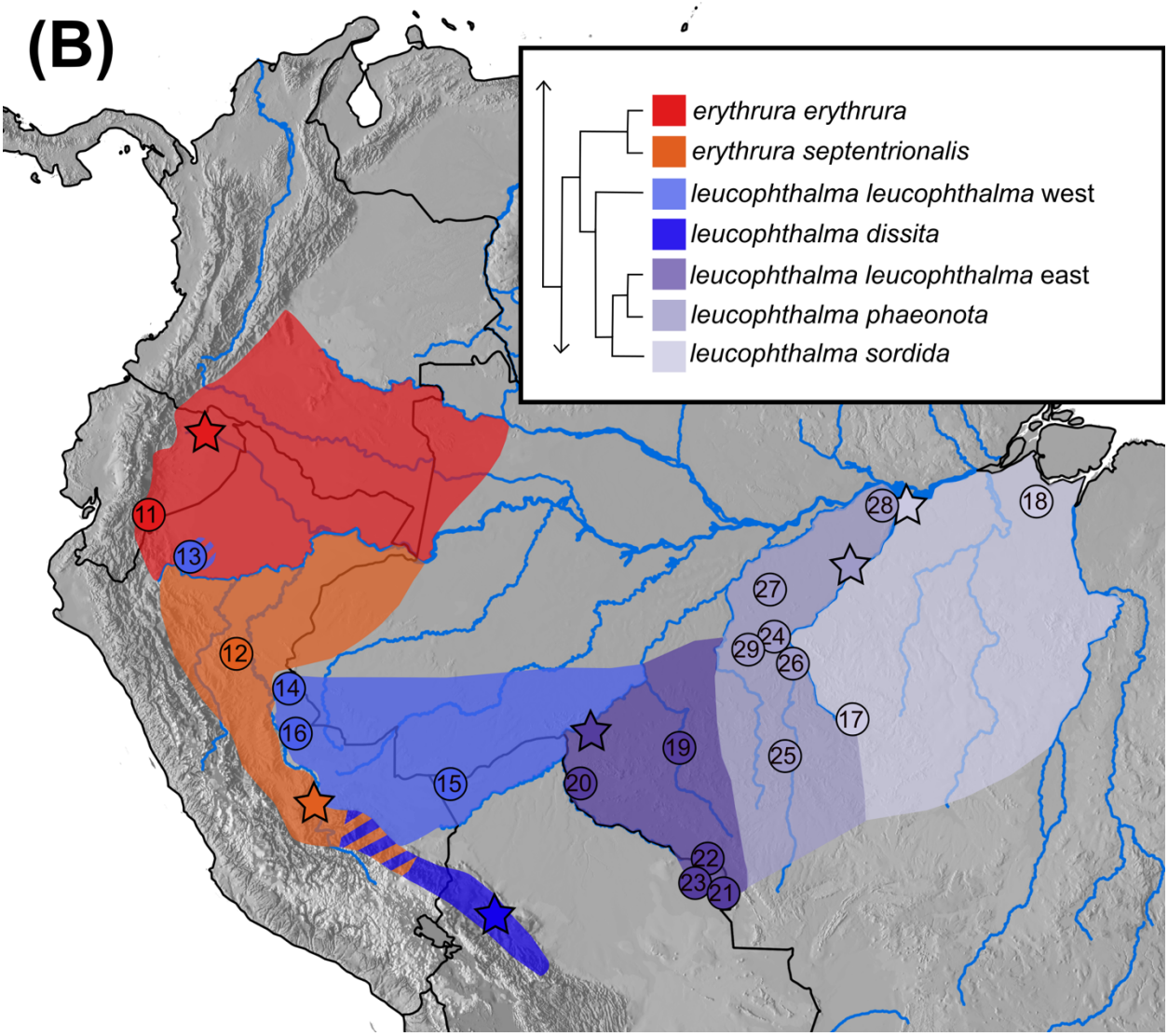
967

968

969

970

971



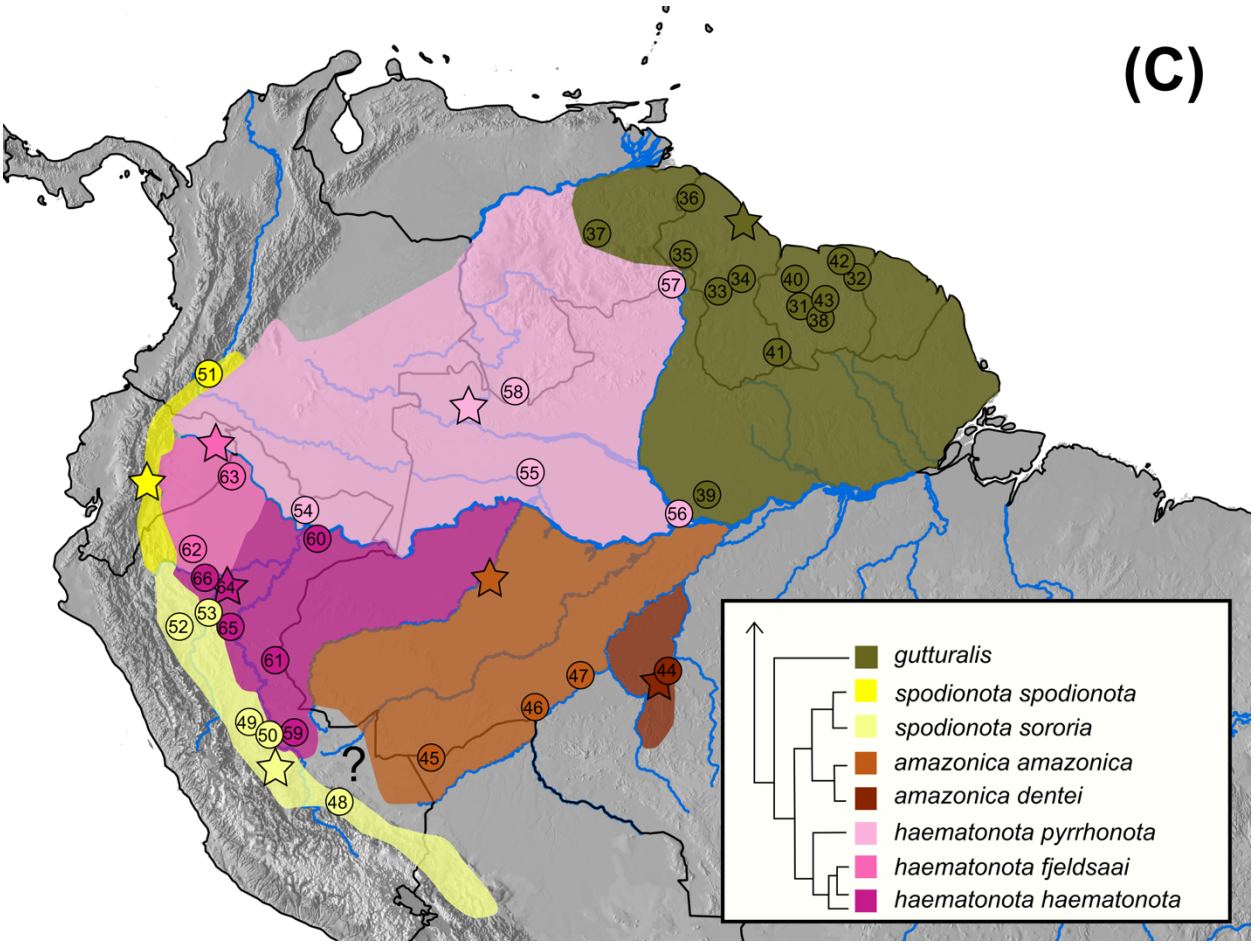
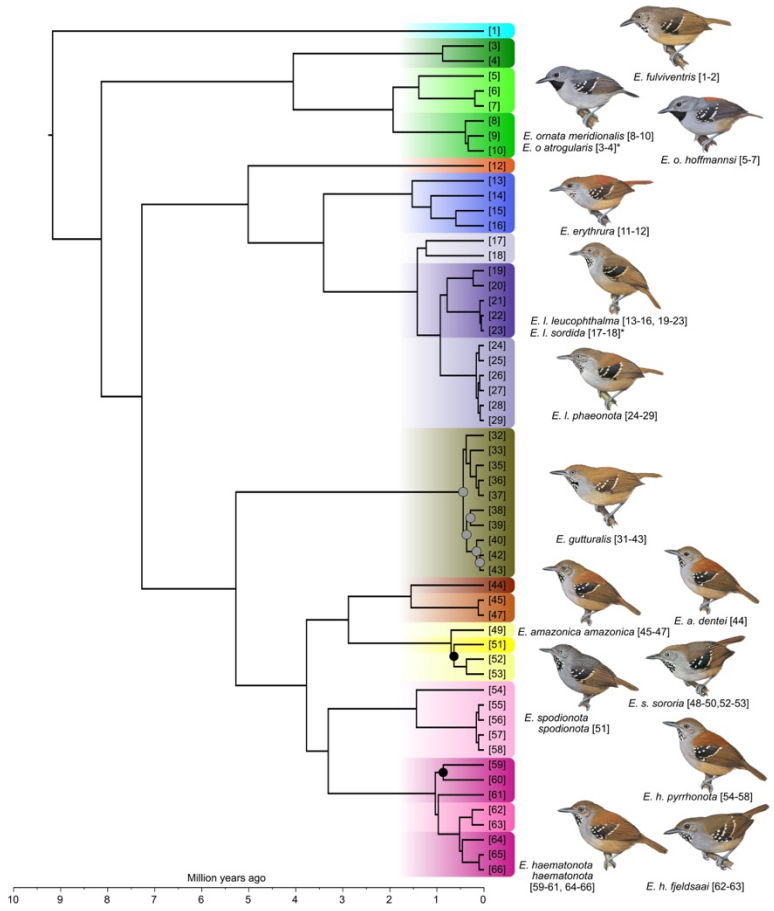


Figure 1. Maps showing taxon distributions, type localities, and sample localities used in this study. A) *Epinecrophylla fulviventris* and *ornata*. B) *E. erythrura* and *leucophthalma*. C) *E. gutturalis, pyrrhonota, dentei, amazonica, spodionota, haematonota,* and *fjeldsaai*. Country boundaries are shown in black. Major rivers are shown in blue. Locations sampled for this study are indicated with a number, corresponding to sample information in Table 1. Type localities, shown with a star, are based on Peters (1951) or type descriptions. Hashed regions indicate range overlap. Inset for each map shows a cladogram of relationships between each taxon based on the trees in Figure 2 and Figure S1H.

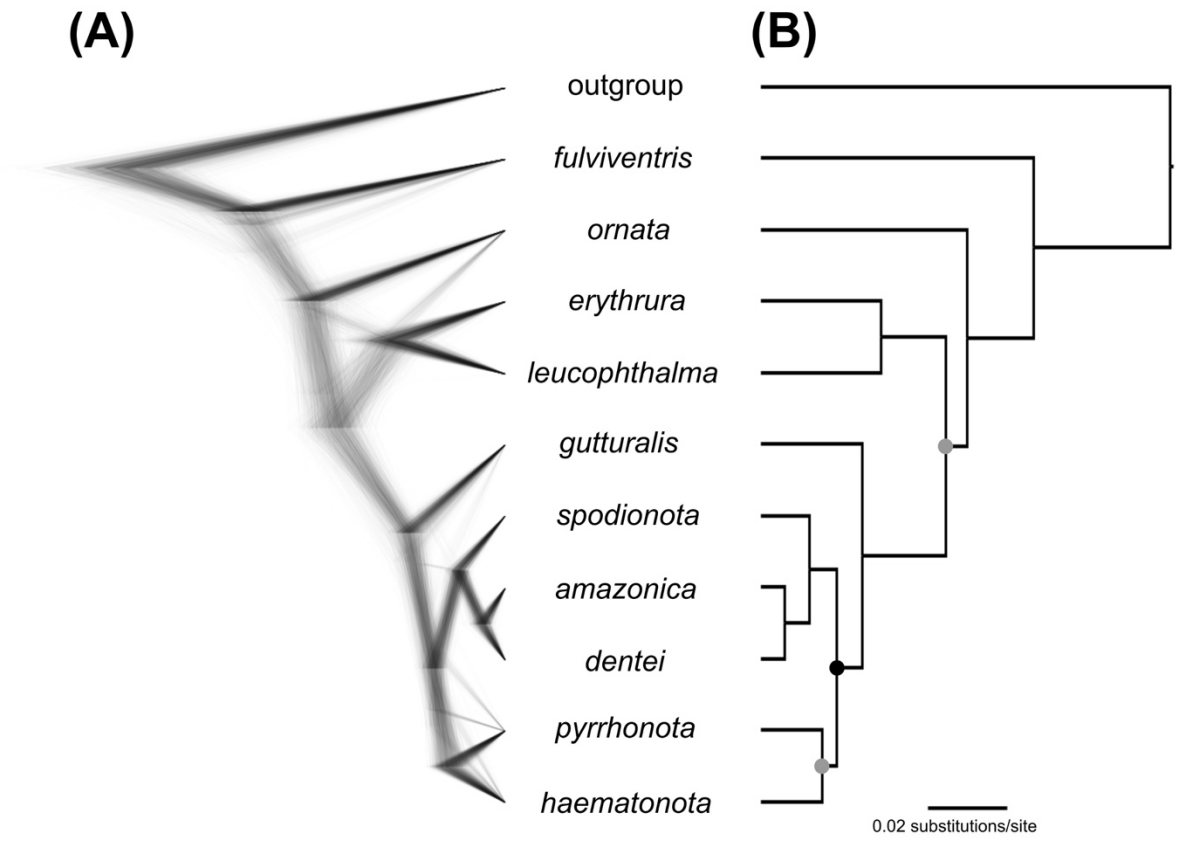


991

992 Figure 2. A dated phylogeny combining UCE and mitogenome sequence data. Topology
993 estimated in Exabayes from the 75% complete phased concatenated UCE data matrix and
994 branch lengths estimated in BEAST using the mitogenome alignments and a fixed substitution
995 rate (see section 2.6 for details). All nodes received full support unless marked with a circle.
996 Nodes with a gray circle with >0.75 posterior probability and nodes with a black circle with
997 >0.90 posterior probability. A version of this tree with 95% confidence intervals on node ages is
998 in Figure S3, and the phylogenetic tree estimated in Exabayes that contains all samples is in
999 Figure S1A. Outgroup samples have been removed for clarity. Colors and sample numbers
1000 correspond to those in Figure 1. Illustrations (all of males) reproduced by permission of Lynx

1001 Edicions. Taxa marked with an asterisk are not illustrated and are placed below the taxon they

1002 most closely resemble in plumage.



1003

1004 Figure 3. Species tree estimated in SNAPP from SNP data, using one sample per species. A) The

1005 Densitree representation of the posterior distribution of species trees and B) the Maximum

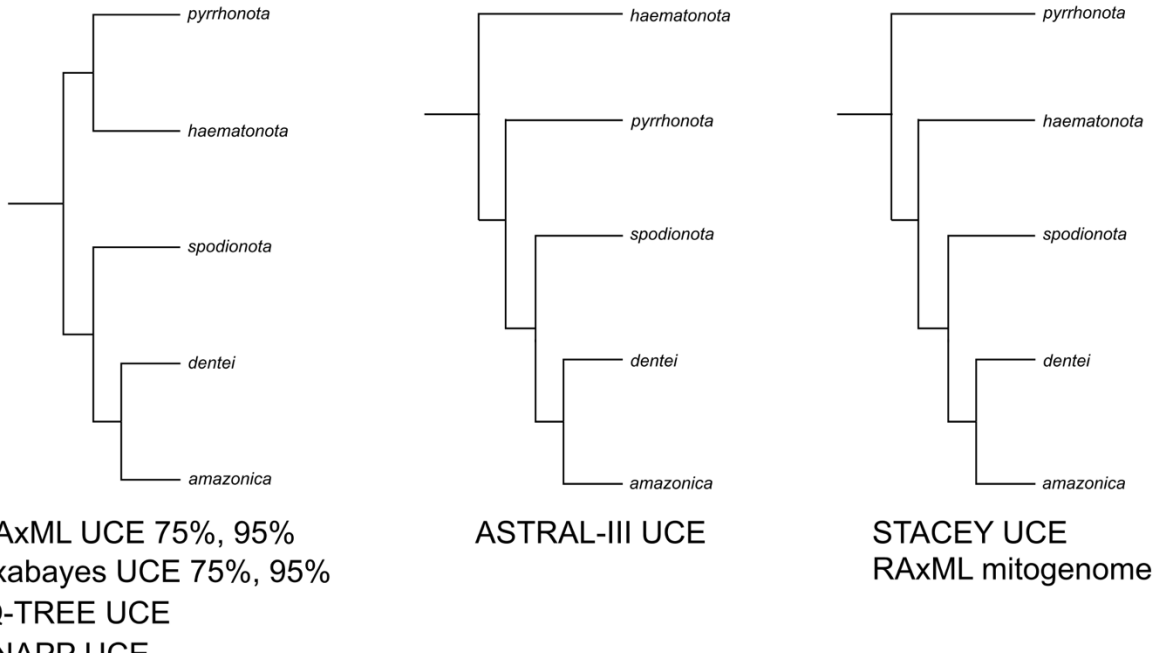
1006 Clade Credibility species tree. All nodes received full support unless marked with a circle. Nodes

1007 with a posterior probability >0.90 are marked with a black circle, and those with a posterior

1008 probability >0.75 are marked with a gray circle.

1009

1010



1011

1012

1013 Figure 4. Alternate topologies recovered across phylogenetic methods for relationships within

1014 the *Epinecrophylla haematonota s.l.* clade. Results are shown using a single individual per taxon

1015 for visualization purposes. In all cases *E. haematonota fjeldsaai* was recovered as embedded

1016 within *E. h. haematonota* and is not shown. The methods recovering each topology are shown

1017 below the topology, and the full phylogenies using each method are in Figure S1.

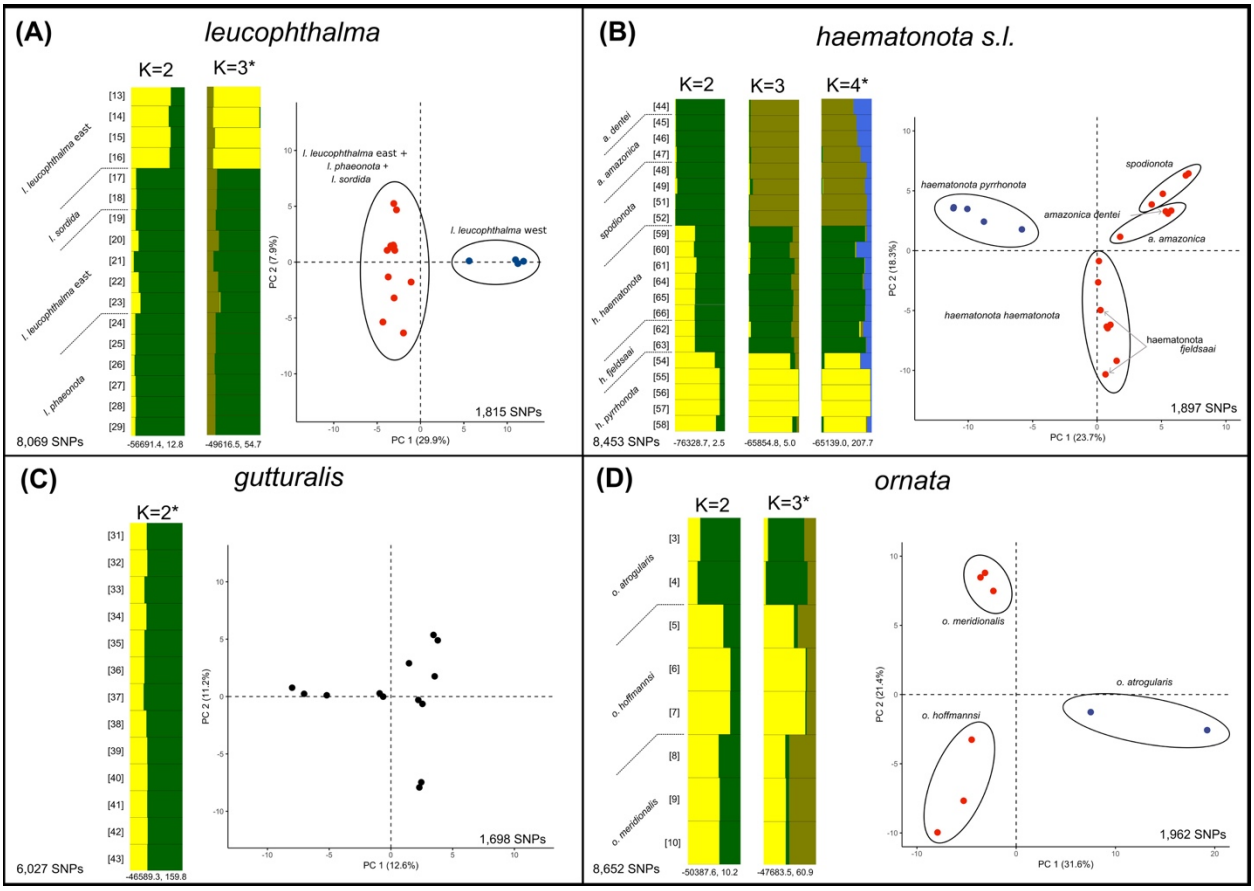


Figure 5. Intra-specific population genetic analyses. A) *Epinecrophylla leucophthalma,* B) the *E. haematonota s.l.* clade, containing *dentei*, *amazonica*, *spodionota*, *sororia*, *pyrrhonota*, *haematonota*, and *fjeldsaai*, C) *E. gutturalis*, and D) *E. ornata*. For each section is shown a Principal Components Analysis (PCA) with samples colored by the Discriminant Analysis of Principal Components Analysis (DAPC) group assignments on the right, and STRUCTURE plots for all likely values of K (i.e. those with low standard deviation across replicate runs) on the left. Not shown are results for K=1. The plot for the "best" value of K for each clade using the Evanno method is marked with an asterisk. Mean log likelihood and delta K values are shown below each STRUCTURE plot. Sample size in PCA plots refer to the number of unlinked SNPs

recovered in that clade and used in the PCA analysis. Blue and red circles denote group assignments from DAPC while black circles and text denote taxa. Sample numbers correspond to those in Figure 1 and Table 1. PCAs with sample numbers included are shown in Figure S5.

SUPPLEMENTAL CAPTIONS

Supplemental Table 1. Samples removed from analyses due to misidentification or potential contamination.

Supplemental Table 2. Samples for which we were unable to recover mitochondrial genomes either due to a failure with MITOBIM or very high amounts of missing data in the recovered mitochondrial genome.

Supplemental Table 3. Matrix of mitochondrial genetic distances between taxa in this study. Values above the diagonal represent the uncorrected *p*-distance and those below the diagonal represent the genetic distance accounting for multiple substitutions under the GTR + γ + I finite-sites substitution model. Values in the diagonal represent the average intra-taxon distance for the finite-sites distance (left) and the uncorrected *p*-distance (right). Both distance methods are based on the concatenated alignment of the 13 mitochondrial protein coding genes. All values are shown as a percentage.

Supplemental Table 4. Matrix of nuclear genetic distances between taxa in this study. Values below the diagonal are the weighted value of F_{st} between taxa, averaged across UCE loci. Values in the diagonal represent the average intra-taxon distance.

Supplemental Table 5. Individuals used as references for SNP calling.

Supplemental Figure 1. Seven estimates of the phylogeny of *Epinecrophylla*, based on UCE alignments, showing congruence of internal topologies across methods. A) Exabayes tree estimated from the 75% complete concatenated data matrix (note that this tree is also shown in Figure 2), with node support values of posterior probability. B) Exabayes tree estimated from the 95% complete concatenated data matrix, with node support values of posterior probability. C) RAxML tree estimated from the 75% complete concatenated data matrix, with node support values of bootstrap likelihood. D) RAxML tree estimated from the 95% complete concatenated data matrix, with node support values of bootstrap likelihood. E) STACEY tree estimated from the fully-partitioned alignment using the 150 loci containing the greatest number of parsimony-informative sites, with node support values of posterior probability. F) ASTRAL-III tree estimated from RAxML gene trees using the 75% complete data matrix, with node quartet support values of local posterior probability. G) IQ-TREE estimated from the 75% complete data matrix. Node support values shown in ultrafast bootstrap likelihood, gene tree concordance factors, and site concordance factors, respectively.

Supplemental Figure 2. An estimate of the phylogeny of *Epinecrophylla*, based on an alignment of draft mitochondrial genomes using a Maximum Likelihood methodology implemented in RAxML, with node support values of bootstrap likelihood.

Supplemental Figure 3. The dated phylogeny shown in Figure 1, with nodes showing the 95% highest posterior density of the divergence estimates based on 11,488 base pairs of the draft mitochondrial genomes and a fixed mitochondrial substitution rate.

Supplemental Figure 4. Species tree estimated in SNAPP from SNP data, using 1-2 samples per clade. A) The Densitree representation of the posterior distribution of species trees and B) the Maximum Clade Credibility species tree. All nodes received full support unless marked with a circle. Nodes with a posterior probability between 0.90 and 0.75 are marked with a gray circle and those <0.75 are marked with a white circle. No nodes received a posterior probability between 1 and 0.90. See Figure 3 for the SNAPP tree using one sample per species.

Supplemental Figure 5. PCA plots shown in Figure 5 with samples labeled by taxon and sample number.

Supplemental Figure 6. PCA plot using four samples of *Epinecrophylla amazonica*. DAPC results indicated a best fit model of K=2, separating the one sample of *dentei* from the three of *amazonica*.