ORNITOLOGÍA NEOTROPICAL
(2016) 27: 223–236
ORIGINAL ARTICLE

A MULTI-GENE ESTIMATE OF HIGHER-LEVEL PHYLOGENETIC RELATIONSHIPS AMONG NIGHTJARS (AVES: CAPRIMULGIDAE)

Noor D. White[1,2] · George F. Barrowclough[3] · Jeff G. Groth[3] · Michael J. Braun[1,2]

[1] Department of Vertebrate Zoology, National Museum of Natural History, Smithsonian Institution, Washington, DC 20560, USA.

[2] Behavior, Ecology, Evolution and Systematics Program, University of Maryland, College Park, MD 20742, USA.

[3] Department of Ornithology, American Museum of Natural History, Central Park West at 79th Street, New York, NY 10024, USA.

E-mail: Noor D. White · noordwhite@gmail.com

ABSTRACT · The higher-level phylogenetic relationships of the nightjars and nighthawks (Caprimulgidae) have been challenging for traditional systematics due to their cryptic plumage and conservative morphology. We explored these relationships by combining two previously published molecular datasets with new data to generate a complete matrix (7,104 bp) of evolutionarily disparate sequence elements from four genes for 36 taxa. We analyzed each of the genes separately for base composition heterogeneity and heterozygosity. We analyzed the concatenated matrix in a likelihood framework using seven different partitioning schemes. As the number of subsets in a given partitioning scheme increased, tree length and likelihood score also increased; however, the branching topology was little affected by increasingly complex partitioning schemes. Our best maximum likelihood tree has increased bootstrap support at 13 of 30 ingroup nodes compared with previous analyses, a result likely due to doubling the length of the sequence data. Coalescent-based species tree inference produced a tree congruent with all strongly supported nodes in the maximum likelihood tree. This topology agrees with previous molecular studies in identifying three small, early branching Old World genera (*Eurostopodus*, *Lyncornis*, and *Gactornis*) and four more speciose terminal clades, representing the New World nighthawks (genus *Chordeiles*) and three nightjar radiations centered in South America, Central America and the Old World, respectively. Increased node support across the tree reinforces a historical scenario with origins in the region surrounding the Indian Ocean, followed by diversification in the New World and subsequent recolonization and radiation in the Old World. Future work on this group should incorporate additional members of the genera *Lyncornis* and *Eurostopodus*, to determine which is the basal lineage of Caprimulgidae.

RESUMEN · Relaciones filogenéticas de más alto nivel de los atajacaminos (Aves: Caprimulgidae) en base a un análisis multigénico

Las relaciones filogenéticas de más alto nivel de los atajacaminos y añaperos (Caprimulgidae) son un reto para la sistemática tradicional, debido a que el grupo posee morfología poco variable y plumajes crípticos. Exploramos relaciones filogenéticas en el grupo combinando dos conjuntos de datos moleculares ya publicados con nuevos datos. La matriz completa (7,104 bp) se generó con cuatro genes y 36 taxones, incluyendo marcadores con distintos modelos de evolución. Se examinó cada uno de los genes por separado para determinar heterocigosidad y heterogeneidad de la composición de bases. Se analizó la matriz concatenada en un marco de máxima verosimilitud utilizando siete particiones diferentes. La longitud de los árboles filogenéticos y su verosimilitud aumentaron a la par del número de subconjuntos en una partición particular; sin embargo, la topología del árbol varió poco entre particiones. En comparación con topologías publicadas, nuestro árbol de máxima verosimilitud tuvo mejor soporte para 13 de los 30 nodos internos, resultado que podría deberse al uso del doble de los datos de secuencias. El método de árboles de especies basado en coalescencia produjo una topología congruente con la obtenida por máxima verosimilitud. Esta topología concuerda con previos estudios moleculares, identificando tres pequeños géneros del Viejo Mundo como basales en la filogenia (*Eurostopodus*, *Lync-ornis* y *Gactornis*), y cuatro clados terminales con más especies. Estos clados terminales representan los atajacaminos del Nuevo Mundo del género *Chordeiles*, y otras tres radiaciones de América del Sur, Central y del Viejo Mundo. Nuestros resultados sugieren un escenario histórico con orígenes del grupo en la región circundante al Océano Indico, seguido por la diversificación en las Américas y la posterior recolonización y radiación en el Viejo Mundo. Futuros estudios en este grupo deben incorporar miembros adicionales de los géneros *Lyncornis* y *Eurostopodus*, lo que permitirá estudiar cuál es el linaje basal de Caprimulgidae.

Key words: *Caprimulgidae* · Molecular phylogeny · Nightjars · Partitioning · *Strisores*

Receipt 11 March 2016 · First decision 8 May 2016 · Acceptance 12 October 2016 · Online publication 15 November 2016

Communicated by Kaspar Delhey

INTRODUCTION

The field of molecular phylogenetics has changed dramatically in the past 15 years. Whereas datasets of less than a kilobase from a single mitochondrial gene were frequently published in the 1990's, the size and complexity of datasets have advanced rapidly to include multiple nuclear genes, whole mitochondrial genomes and even entire nuclear genomes (e.g., Jarvis et al. 2014). Today it is relatively straightforward to amass datasets consisting of hundreds to thousands of nuclear markers for dozens of taxa, due to the advent of high-throughput sequencing platforms and the development of efficient genome reduction techniques (e.g., McCormack et al. 2012, Faircloth et al. 2012, Lemmon et al. 2012). However, drawing conclusions from analyses of these datasets requires caution, as they can produce trees with high statistical support that conflict with independent analyses bearing equally high support (e.g., note conflict between the Bayesian trees in Jarvis et al. 2014 and Prum et al. 2015, see also Hahn & Nakhleh 2015). These examples illustrate that, while analytical methods have been advancing rapidly, the growth of datasets has outpaced the development of software with which to analyze them (see discussion in Kumar et al. 2012). Thus it is important to examine the complexities of phylogenetic inference on datasets of moderate size where more comprehensive analyses can be undertaken, in order to both test emerging analytical methods and to provide topological comparisons for genome-scale work. Here we explore two important analytical issues for which relatively new software has been developed: 1) data partitioning (Bull et al. 1993, de Queiroz 1993) - whereby different models of sequence evolution are applied to distinct subsets of a data matrix evolving under different functional constraints, and 2) incongruence between gene trees and species trees (reviewed in Liu et al. 2015). We apply these methods to address the deeper relationships in the Caprimulgidae (nightjars and nighthawks), a family with a striking but understudied evolutionary history.

The Caprimulgidae were long divided into two subfamilies, nightjars (Caprimulginae) and nighthawks (Chordeilinae) based on several morphological characters including wing shape, palate structure, and rictal bristles (e.g., Oberholser 1914, Ridgway 1914, Peters 1940, Hoff 1966, Cleere 1998). However, the exact composition of the two subfamilies was never settled, with several genera (*Podager*, *Eurostopodus*, *Veles*, *Nyctiprogne*) being shifted back and forth due to presence or absence of some of these characters (e.g., Holyoak 2001, Whitney et al. 2003). Moreover, it was clear that these traits might be prone to convergence because they were associated with foraging ecology - nightjars typically sally after flying insects from an exposed perch at night, while nighthawks pursue flying insects during sustained flight at dusk and dawn. A second major issue concerned the large genus *Caprimulgus* (sensu lato), with 55–57 species and a cosmopolitan distribution, which appeared to be a grab bag of taxa with an ancestral body plan and few derived features (Cleere 1998).

Although a number of authors have commented on the morphology and anatomy of various exemplars of Caprimulgidae, most did not have sufficient taxon sampling to address relationships across the family in any detail (e.g., Oberholser 1914, Wetmore 1918, Hoff 1966, Bühler 1970, Schodde & Mason 1980, Mayr 2002, Mayr et al. 2003). An exception was Mayr (2010) who examined eight caprimulgid genera but did not find or did not analyze phylogenetically informative variation within the family. The only morphological study with truly extensive sampling of Caprimulgidae is the recent osteological analysis of Costa (2014), who examined nearly 50 species and all genera but *Veles*.

Molecular studies have begun to clarify caprimulgid phylogeny, suggesting a complex biogeographic and evolutionary history. The DNA hybridization data of Sibley & Ahlquist (1990) and mitochondrial cytochrome b (MT-CYB) sequence data of Mariaux & Braun (1996) first indicated that the two traditional subfamilies were not monophyletic. Barrowclough et al. (2006) used recombination activating gene-1 (RAG-1) to investigate the Caprimulgidae at the generic level, finding strong support for the placement of *Eurostopodus* sister to the rest of Caprimulgidae and for polyphyly of *Caprimulgus* (sensu lato). They also showed that most caprimulgid species belong to one of four major geographically-relevant clades, either restricted to the New World or the Old World. More data from MT-CYB and MYC, the cellular homolog of the myelocytomatosis viral oncogene, reinforced these conclusions (Larsen et al. 2007; Braun & Huddleston 2009).

The phylogeny and classification of Caprimulgidae underwent a significant overhaul with the work of Han et al. (2010), including new generic designations for many groups. These authors used data from three genes: MT-CYB, MYC, and growth hormone (GH). Their work mostly confirmed the findings of Barrowclough et al. (2006), although the relative placement of the four major geographic clades differed. With more comprehensive taxon sampling, Han et al. (2010) detected two previously unappreciated long branches: a deep split within *Eurostopodus* (justifying resurrection of the genus *Lyncornis*) and the Malagasy endemic "*Caprimulgus*" *enarratus*, for which they erected the new genus *Gactornis*. Most recently, Sigurdsson & Cracraft (2014) studied the phylogeny of New World Caprimulgidae at the species, and in some cases subspecies, level with data from four genes (including RAG-1 and MT-CYB). Their resolution of relationships between the major geographic clades is compatible with those found by Barrowclough et al. (2006), though no study has yet found strong support for one of the alternative topologies. We will use the generic nomenclature of Han et al. (2010) and follow Sigurdsson & Cracraft (2014) in referring to the four major geographic clades as 1) the Poorwill Clade, con-

taining mostly North and Central American nightjars (genera *Antrostomus*, *Nyctiphrynus* and *Siphonorhis*); 2) the Nighthawk Clade (the New World genus *Chordeiles*); 3) the South American Clade (the Neotropical genus *Hydropsalis*); and 4) the Old World Clade, a radiation including African, Asian, Australian, and European lineages (genus *Caprimulgus* sensu stricto).

In this study, we sought to provide an improved estimate of the higher level phylogeny of Caprimulgidae by combining the datasets of Barrowclough et al. (2006) and Han et al. (2010). The expanded dataset effectively doubled the number of characters in either previous one, and incorporates heterogeneous molecular marker types, including mitochondrial and nuclear genes, introns, exons and untranslated regions (UTRs). We added new RAG-1 sequences to generate a complete data matrix for 36 taxa addressing key basal nodes in the family and the placement of all major geographic clades. We partitioned and analyzed the matrix using a variety of *a priori* and *a posteriori* partitioning methods, and attempted to identify the species tree despite gene tree-species tree incongruence with a recently described quartet assembly approach, SVDquartets (Chifman & Kubatko 2014, 2015). Lastly, we followed up on the report by Barrowclough et al. (2006) that some caprimulgids have elevated GC content and excessive heterozygosity at the RAG-1 locus by exploring these parameters in all of our nuclear loci.

METHODS

Sequencing. Barrowclough et al. (2006) previously obtained RAG-1 sequences for 24 species examined in this study. Using methods described by Groth & Barrowclough (1999), and Barrowclough et al. (2006), we sequenced 12 additional taxa for the RAG-1 exon to obtain a common set of critical species for comparison with the Han et al. (2010) study (see Table 1 for voucher and GenBank accession numbers). The placement of *Hydropsalis parvulus* differs between Han et al. (2010) and Barrowclough et al. (2006), so both of those vouchered specimens were included here to test for possible contamination or mislabeling. The new RAG-1 sequences were assembled using Sequencher software (version 5.1; Gene Codes: Ann Arbor, MI), and aligned manually. All new RAG-1 sequences were examined for 1) indels that were not a multiple of three base pairs in length, 2) unexpected stop codons in the reading frame, and 3) unexpectedly similar (chimeric or contaminated) portions of sequence between taxa before inclusion in this study.

Dataset generation. A complete matrix of four loci for 36 taxa was generated by combining the aligned dataset of RAG-1 with aligned data from Han et al. (2010) using PAUP* (version 4.0a130; Swofford 2003). The four loci include: the entire MT-CYB coding sequence; parts of exons 2 and 3, and all of intron 2 from GH; part of intron B, all of exon 3 and part of the 3' UTR of MYC; and most of the exonic region of RAG-1. Alignments of MT-CYB, GH, and MYC were initially done in Clustal X (version 1.8.3; Thompson et al. 1997), then edited manually by Han et al. (2010). The resulting 7,104 base pair (bp) aligned dataset doubles the size of either original, and combines multiple lines of genetic evidence. Representatives from every major nocturnal lineage of Strisores were included to allow outgroup rooting of all trees. The aligned data matrix is deposited in Treebase (ID # 19469).

We did not incorporate the data from Sigurdsson & Cracraft (2014) in this analysis for two reasons. First, due to their focus on New World taxa, the differences in taxon sampling - especially of outgroups - would require substantial further sequencing. Second, we do not expect the addition of mitochondrial loci to help elucidate phylogeny at this evolutionary depth (further discussed later). Therefore, including the Sigurdsson & Cracraft (2014) data would have added limited data from only one additional locus (intron 9 from the nuclear aconitase gene) to this study.

Heterozygosity and base composition. In order to determine whether the previously reported (Barrowclough et al. 2006) high GC content and heterozygosity at the RAG-1 locus extended to other loci, we estimated heterozygosity and overall base composition on our data. Barrowclough et al. (2006) noted that heterogeneity in base composition was largely driven by elevated GC content at third codon positions (GC3) in some species. However, two of the three nuclear loci used here include extensive noncoding regions. Therefore, we compared the distributions of overall GC content, rather than just GC3. We used contingency G-tests of A + T versus G + C proportions among species, separately for each of the three genes, to calculate base composition heterogeneity, and employed non-parametric Kolmogorov-Smirnov (KS) tests to calculate distributions of observed heterozygosity between genes. To visualize GC differences among the genes, we used standard Tukey (1977) box plots to summarize their distributions; medians, upper and lower quartiles (box), and ranges (whiskers) were found. Species with GC content greater than, or less than, 1.5 times the interquartile range from their respective quartile, were indicated by dots. The mitochondrial MT-CYB locus is not relevant for analyses of heterozygosity, and hence was excluded from these analyses.

Partitioning schemes and alternative models. *A priori* partitions of the sequence data were chosen based on the expectation that rates and patterns of molecular evolution will vary among loci, subcellular compartments, and distinct functional regions of genes. For example, a model of sequence evolution applied to the slowly-evolving exons may be inappropriate to apply to the more quickly-evolving introns. We tested six *a priori* partitioning schemes that subdivide the dataset in various ways, ranging from

Table 1. GenBank accession numbers and voucher information for all new sequences generated for this study. Acronyms: CONACYT = Consejo Nacional de Ciencia y Tecnología, FMNH = Field Museum of Natural History, USNM = US National Museum of Natural History, AMNH = American Museum of Natural History, KUNHM = University of Kansas Natural History Museum, ANSP = The Academy of Natural Sciences, Philadelphia.

Species	GenBank accession number	Voucher information
Antrostomus ridgwayi	KU361177	CONACYT (Mexico) 415
Caprimulgus affinis	KU361174	FMNH 358300
Caprimulgus manillensis	KU361180	USNM B6090
Chordeiles pusillus	KU361178	USNM B12993
Eurostopodus argus	KU361170	AMNH DOT2401
Gactornis enarratus	KU361171	FMNH 438654
Hydropsalis anomalus	KU361179	KUNHM 3275
Hydropsalis anthonyi	KU361173	ANSP 4580
Hydropsalis nigrescens	KU361175	USNM B4478
Hydropsalis parvulus	KU361176	KUNHM 106
Hydropsalis whitelyi	KU361181	USNM B19022
Siphonorhis brewsteri	KU361172	KUNHM 8149

unpartitioned, to very simple subdivisions, to a more complex scheme (Table 2). For comparison, we tested the *a posteriori* method of partitioning implemented in PartitionFinder (version 1.1.1; Lanfear et al. 2012). Given user specified subsets of the data, PartitionFinder uses a heuristic search to find an optimal partitioning scheme by searching for the best model for each subset, and combining subsets that conform to similar models of sequence evolution. We specified 18 possible subsets of the data, including the first, second and third positions of each exon in each gene, as well as subsets for the UTR and each intron. The best partitioning scheme was selected by PartitionFinder using the corrected Akaike information criterion (AICc; Akaike 1973, Hurvich & Tsai 1989). We employed the 'greedy' search algorithm, for the sake of computational tractability.

To test the effect of alternative models on phylogenetic analysis, we conducted maximum-likelihood (ML) analyses of each partitioning scheme and each individual gene under two models. First, an independent general-time reversible model with estimated proportion of invariant sites and gamma-distributed rate variation among sites (GTR + I + G; the most highly parameterized model available) was applied to each subset. Second, we set user-defined partitioning schemes, and allowed PartitionFinder to select the best model for each subset in a given scheme. All PartitionFinder runs used unlinked branch lengths, and searched all models of sequence evolution implemented in the program. Model descriptions may be found in Posada (2008). The number of parameters in each partitioning scheme was calculated automatically by PartitionFinder. To calculate that number for our schemes under GTR + I + G, we ran each scheme through PartitionFinder with both the scheme and model (GTR + I + G for each subset) user-defined.

Phylogenetic analyses. ML tree searches were conducted for each individual gene and all partitioning schemes of the combined dataset using GARLI (version 2.01) with partitions unlinked (Zwickl 2006). In order to ensure thorough searches for optimal tree topologies, we used the 'searchreps' option to vary the number of search replicates performed within a GARLI run. After 100 GARLI runs with a given number of search replicates, we compared topologies between best trees from each run using the 'treedist' function in PAUP* (version 4.0a130; Robinson & Foulds 1981). When the best topologies were identical for all 100 runs, we assumed the number of search replicates was sufficient to produce the optimal tree topology. The number of search replicates required to satisfy this criterion was two in all cases except the individual gene trees for GH and MT-CYB, which required 40 search replicates. Overall tree lengths were calculated in PAUP* (version 4.0a146) using the function 'describetrees/brlens'.

To evaluate nodal support, 100 non-parametric bootstrap datasets were generated and subjected to GARLI tree searches, with 1 search replicate for each bootstrap run. Nodal support values were tabulated as the number of bootstrap runs in which a particular node appears, and plotted on the optimal topology for each gene or partitioning scheme using the SumTrees program (version 3.3.1) in the python library DendroPy (version 3.12.0; Sukumaran & Holder 2010).

Table 2. Partitioning schemes used to study phylogenetic relationships across 36 species of Caprimulgiformes. The length of each partition is given in nucleotide base pairs (bp). Model chosen by PartitionFinder for each subset, as well as individual genes, are given. Model descriptions may be found in Posada (2008).

Scheme	Partitions	Partition length (bp)	PartitionFinder model
Unpartitioned	n / a	7.104	GTR + I + G
Coding vs. non-coding	Codon positions 1 & 2	3.137	GTR + I + G
	Codon position 3	1.565	GTR + I + G
	Introns & UTR	2.402	TVMef + G
By gene	MYC	1.318	TVM + I + G
	MT-CYB	1.143	GTR + I + G
	GH	1.765	SYM + G
	RAG-1	2.878	GTR + I + G
Coding positions vs. non-coding	Codon position 1	1.569	SYM + I + G
	Codon position 2	1.568	GTR + I + G
	Codon position 3	1.565	GTR + I + G
	Introns & UTR	2.402	TVMef + G
Nuclear vs. mito vs. non-coding	Nuclear codon positions 1 & 2	2.375	GTR + I + G
	Nuclear codon position 3	1.184	TVM + I + G
	MT-CYB positions 1 & 2	762	TVM + I + G
	MT-CYB position 3	381	TIM + I + G
	Introns & UTR	2.402	TVMef + G
Nuclear vs. mito vs. introns vs. UTR	Nuclear codon positions 1 & 2	2.375	GTR + I + G
	Nuclear codon position 3	1.184	TVM + I + G
	MT-CYB positions 1 & 2	762	TVM + I + G
	MT-CYB position 3	381	TIM + I + G
	Intron	2.012	TVMef + G
	UTR	390	TIM + G
PartitionFinder	Introns	2.012	TVMef + G
	MYC UTR, MYC exon 1st pos, MYC exon 3rd pos, RAG-1 2nd pos, RAG-1 3rd pos, GH exon1 1st pos, GH exon1 3rd pos	2.764	TVM + I + G
	MYC exon 2nd pos, MT-CYB 1st pos	573	GTR + I + G
	MT-CYB 2nd pos, GH exon1 2nd pos	391	TrN + I + G
	MT-CYB 3rd pos	380	TIM + I + G
	RAG-1 1st pos, GH exon2 2nd pos	984	SYM + G
Individual genes			
MYC	n / a	1.318	TVM + I + G
MT-CYB	n / a	1.143	GTR + I + G
GH	n / a	1.765	SYM + G
RAG-1	n / a	2.878	GTR + I + G

To address gene tree / species tree discordance, we applied a new coalescent-based species tree method which uses the full data matrix directly, without estimating individual gene trees or utilizing computationally-inefficient Bayesian statistics (SVD-quartets; Chifman & Kubatko 2014, 2015). This method computes the probability distribution of site patterns at the tips of the tree by integrating over the

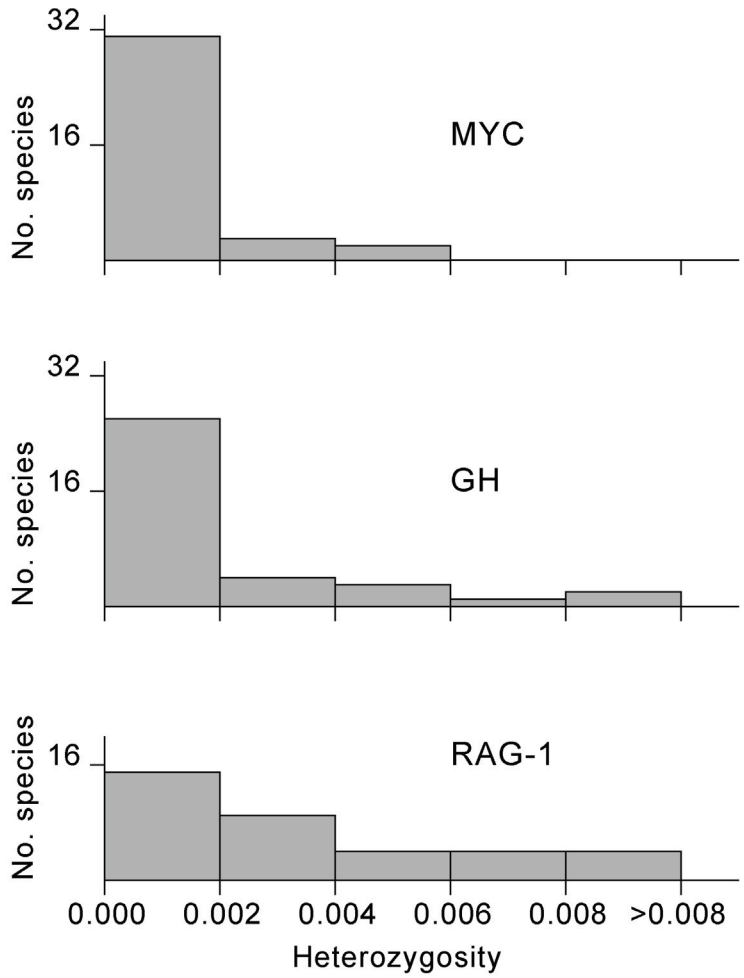

Figure 1. Distributions of observed heterozygosity for 36 species of caprimulgiform birds at three nuclear loci; distribution for RAG-1 is significantly different from that of both GH ($P < 0.05$) and MYC ($P < 0.01$); distributions for GH and MYC do not differ ($P > 0.05$): Kolmogorov-Smirnov two-sample tests.

probability distribution of gene trees under the coalescent model. It was designed for single nucleotide polymorphism data, but has been demonstrated to perform well on multi-locus datasets, such as ours. We employed the SVDquartets code implemented in PAUP* (version 4.0a147), conducting exhaustive sampling of quartets ('eval=all') and 100 non-parametric bootstrap replicates. All trees were rooted to known outgroups representing each of the other nocturnal families of Strisores, and have been deposited in Treebase (ID # 19469).

RESULTS & DISCUSSION

Heterozygosity and base composition. In an earlier report on RAG-1 variation in Caprimulgiformes, Barrowclough et al. (2006) found that GC3 composition was correlated with heterozygosity in a clade of Old World nightjars. We extended that investigation to the two additional nuclear loci examined here to see if the RAG-1 results represented a general, perhaps genome-wide, phenomenon. The number of species showing high levels of heterozygosity for RAG-1 was greater than for GH, which in turn exceeded that of MYC (Figure 1). The RAG-1 distribution of heterozygosity was significantly different from that of GH ($P < 0.05$) and of MYC ($P < 0.01$), but the distributions for GH and MYC were not statistically divergent ($P > 0.05$).

Median GC composition was highest in GH (0.49) and lowest in MYC (0.44) (Figure 2). Each of the pairwise comparisons of the overall distributions was significant at the 0.01 level (Figure 2A). There was significant heterogeneity in GC composition among taxa for RAG-1 ($\chi^2 = 67.6$, df = 35, $P < 0.01$), but not for either MYC ($\chi^2 = 7.4$, df = 35, $P > 0.5$) or GH ($\chi^2 = 38.8$, df = 35, $P > 0.1$). Although heterogeneity in base composition was not statistically significant for GH, its range (0.064) was nearly equal to that of RAG-1 (0.068), and the correlation in base composition between the two loci was significant ($P = 0.01$; Figure 2B). However, the amount of variation explained by the correlation was not very large ($R^2 = 0.18$).

As reported by Barrowclough et al. (2006), correlation analysis verified that heterozygosity and base composition were correlated across species for RAG-1

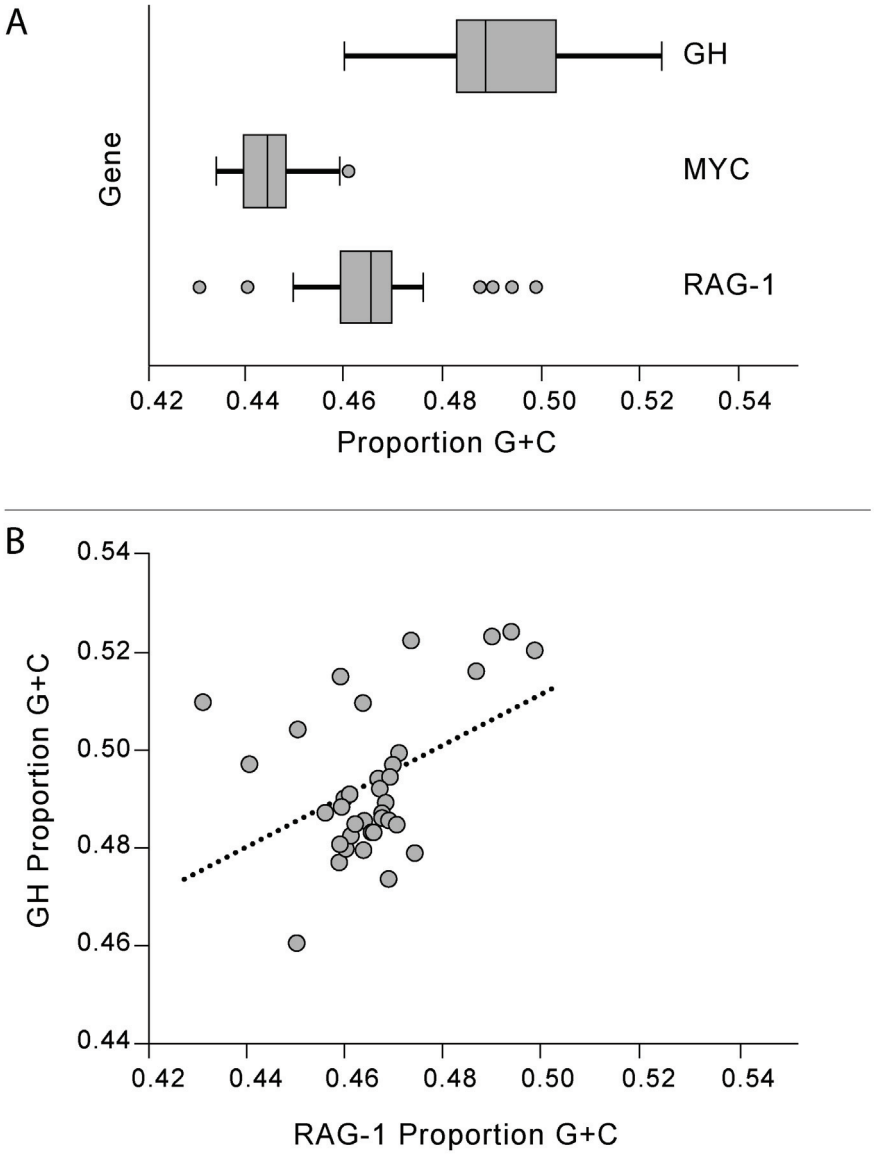

Figure 2. Base composition at three nuclear loci in 36 caprimulgiform birds. A: Box plots of distribution of GC proportions at each locus (box indicates first to third quartiles, interior line is the median; whiskers extend up to an additional 1.5 interquartile ranges; points farther from a quartile are indicated by dots); all three comparisons are statistically significant ($P < 0.01$: Kolmogorov-Smirnov two-sample tests). B: Scatter plot and linear regression ($R^2 = 0.18$) of GC proportion at GH gene versus RAG-1 gene.

($R^2 = 0.26$, $P < 0.01$), but the correlation did not extend to the other two genes (MYC: $R^2 = 0.02$, $P > 0.1$; GH: $R^2 = 0.00$, $P > 0.5$). These results strongly suggest that the condition identified in RAG-1 is not a genome-wide phenomenon attributable to such causes as larger effective population sizes, interspecific hybridization, etc. In the Domestic Fowl (*Gallus gallus*), these three loci reside on separate chromosomes and so it appears that one of the more probable causes is increased mutation associated with a GC-rich isochore encompassing the region surrounding the RAG-1 gene in some taxa. This is indicated by the comparatively long branch lengths seen in the Old World Clade of the RAG-1 gene tree relative to the other caprimulgid clades (Supplemental Material Figure 1), and perhaps to a lesser extent in the GH gene tree (Supplemental Material Figure 2). This information would potentially be of considerable importance in phylogenetic reconstruction as it informs us about the substitution process for RAG-1. Unfortunately, most current phylogenetic inference algorithms, such as those used here, utilize DNA sequence partitions taken across the entire set of taxa in the study. Partition heterogeneity across taxa (i.e., failure of base composition stationarity) is a much more difficult process to model for all but the smallest datasets (Galtier & Gouy 1998). In this case, the fact that the individual phylogenies produced by the three nuclear genes are largely concordant suggests the lack of stationarity is a small effect relative to the overall historical signal (see also Supplemental Material Figure 3 for the MYC gene tree).

Table 3. Results of alternate model analysis for each partitioning scheme, as well as individual genes used to determine phylogenetic relationships across 36 species of Caprimulgiformes. Number of parameters per scheme calculated by PartitionFinder. Maximum likelihood scores (-lnL) are those reported by GARLI. Overall tree length is presented in substitutions per site summed over the whole tree.

	PartitionFinder Model			GTR + I + G		
Scheme	**# of params**	**-lnL**	**Tree length**	**# of params**	**-lnL**	**Tree length**
Unpartitioned	79	38,507.37	1.315	79	38,507.37	1.315
Coding vs. non-coding	232	37,740.03	2.240	237	37,733.67	2.238
By gene	311	37,362.41	2.213	316	37,361.02	2.212
Coding positions vs. non-coding	308	37,631.55	2.231	316	37,623.16	2.230
Nuclear vs. mito vs. non-coding	386	36,642.52	2.374	395	36,635.04	2.391
Nuclear vs. mito vs. introns vs. UTR	462	36,520.87	2.400	474	36,515.67	2.425
PartitionFinder	459	36,250.03	2.465	474	36,240.87	2.487
Individual genes						
MYC	78	5,088.23	0.545	79	5,087.70	0.547
MT-CYB	79	12,448.95	12.588	79	12,448.95	12.588
GH	75	7,339.50	1.125	79	7,338.93	1.125
RAG-1	79	12,085.61	0.550	79	12,085.61	0.550

Data partitioning and alternative models. Despite testing seven different partitioning schemes ranging between one (unpartitioned) to six subsets of the concatenated dataset, the resulting trees were very similar in topology and bootstrap support across all treatments. The major effects of our partitioning tests were on tree length and likelihood score, which both increased with increased number of partitions (Table 3). This can be attributed to an improvement in the ability of a more parameterized model to account for sequence evolution. With our data, applying as few as three partitions (roughly tripling the number of parameters) was enough to improve estimation of the model. The greatest likelihood score was seen using the automated partitioning software on our dataset. The scheme chosen by PartitionFinder had the same number of data subsets as our most complex *a priori* scheme (6), but it fit a model that was slightly less complex (459 vs. 462 parameters), and produced a substantial improvement in likelihood score (~ 270 units). This demonstrates the value of the automated search for combination of subsets, and of sorting data by quantitative patterns that may not have been expected given preconceived notions of molecular evolution. With the exception of the 'Unpartitioned' scheme, and individual genes MT-CYB and RAG-1, the models chosen by PartitionFinder all had fewer parameters than GTR + I + G. Despite this, differences in likelihood score and tree length for a given partitioning scheme were modest (Table 3), and identical tree topologies were found under both models tested. Similarly, both the partitioning schemes and the models applied had slight but inconsistent effects on bootstrap support values (Supplemental Material Table S1, Supplemental Material Figures S1–S4).

The seven partitioning schemes tested resulted in only two topological changes (Figure 3). The two simplest schemes ('Unpartitioned' and partitioned 'By Gene') produced trees that differed by both changes, while all other schemes produced a single topology agreeing with the 'Unpartitioned' scheme in one area and the 'By Gene' partitioning scheme in the second area. All schemes with more than one partition agreed on the resolution of the three *Hydropsalis* species, but the 'By Gene' partitioning scheme differs from all others in outgroup topology, placing the Oilbird (*Steatornis caripensis*) sister to owlet-nightjars (*Aegotheles insignis*) instead of potoos (*Nyctibius grandis;* Figure 3). The 'By Gene' scheme is perhaps the least sophisticated partitioning of our dataset, and may overemphasize the signal of the mitochondrial locus relative to the other partitions. Mitochondrial loci evolve quickly, and may be too saturated to resolve phylogeny at this evolutionary depth. The other partitioning schemes tested here, as well as the unpartitioned analysis, may average out the signal of the mitochondrial locus with the nuclear loci and provide a better estimate of evolutionary history. In addition, a recent analysis that incorporates other relevant taxa from the Strisores also finds oilbird and potoos to be sister taxa (Prum et al. 2015).

The maximum likelihood topology. Our best estimate tree from the concatenated dataset is based on the 'PartitionFinder' partitioning scheme run under the model selected by that program (Figure 4). It

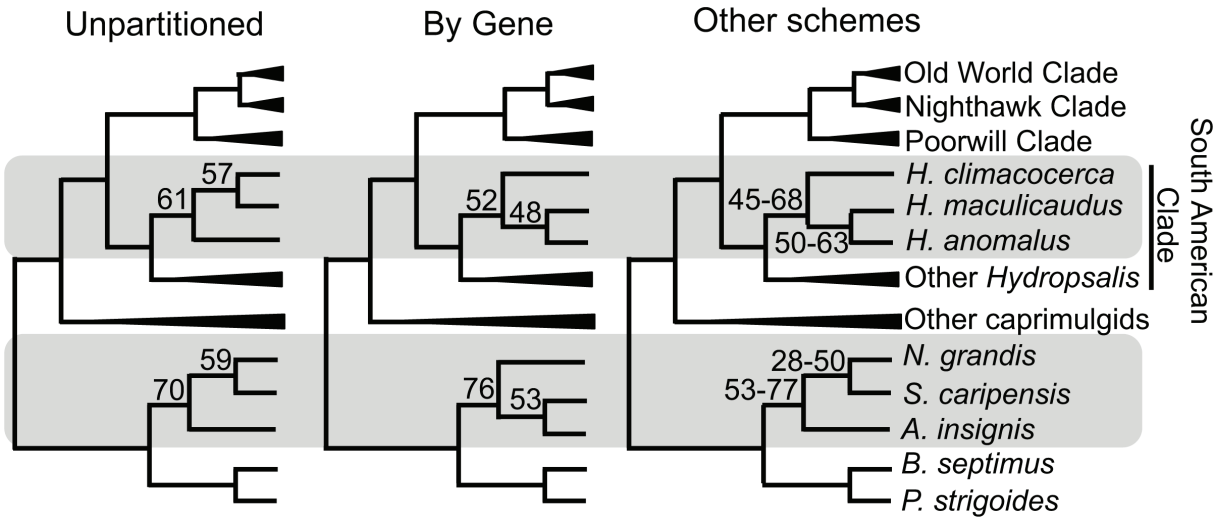

Figure 3. Cladogram representation of the three topologies found with the full dataset under alternative partitioning schemes for 36 caprimulgiform species. Branches are collapsed where topologies are identical, broadened tips represent multiple taxa. Differences between the three topologies are highlighted in grey. Bootstrap support values present on relevant nodes for 'Unpartitioned' and 'By Gene' analyses. For bootstrap support values of all partitioning schemes, see Supplemental Material Table S1.

resolves the relationships among the four major clades of caprimulgids, placing the South American Clade basal with 87% bootstrap support, and the Old World and Nighthawk Clades sister with 96% bootstrap support. *Gactornis* is firmly placed sister to the four major clades with 98% support, with *Lyncornis* and *Eurostopodus* successively more basal. *Hydropsalis rufiventris* and *H. leucopyga,* two taxa formerly considered to belong to the ''nighthawk" subfamily Chordeilinae, are firmly placed within the South American Clade as successive basal branches.

This tree shows increased resolution and bootstrap support relative to comparable prior analyses (summarized in Figure 5, tabulated in Supplementary Material Table S2). Of 30 ingroup nodes in the current tree, 12 have increased bootstrap support over the study of Han et al. (2010), while only four have decreased support. Similar increases in nodal support were seen over Barrowclough et al. (2006), and, in comparison to Sigurdsson & Cracraft (2014), support increased for nine nodes and decreased for five. Decreased support in the latter case was concentrated on shallow nodes and can be attributed to the smaller amount of mtDNA sequence included in our dataset.

The SVDquartets topology. Our SVDquartets tree (Supplementary Material Figure S5) has the same topology for the major clades of Caprimulgidae as the best ML tree, though with lower bootstrap support overall. The positions of *Eurostopodus* and *Lyncornis* are switched in this tree and the outgroup topology changes, again with lower bootstrap support. There are substantial differences within the South American Clade. *Hydropsalis maculicaudus* and *H. climacocerca* are sister as found previously in the 'Unpartitioned' analysis (Figure 3), but now with bootstrap support of 74%, the highest seen for this node in any of our analyses. *Hydropsalis rufiventris* and *H. leucopyga* are nested well within the South American Clade in the SVDquartets tree, as opposed to being successive basal branches as in the best ML tree from the concatenated dataset.

This analysis provides an interesting perspective on phylogeny in this group – it confirms some key features of the tree, but it also differs in important ways. However, we consider any gene tree / species tree-type analyses based on these data preliminary for two reasons. First, four genes is a limited sampling of loci, subject to potential sampling error. Second, one of our genes is a rapidly evolving mitochondrial locus that can be expected to have weak phylogenetic signal deep in the caprimulgid tree due to mutational saturation effects. This was documented by Larsen et al. (2007) for their caprimulgid MT-CYB data. The relative strengths and weaknesses of gene tree / species tree vs. concatenated analyses of molecular sequence data are topics of much debate (e.g., Gatesy & Springer 2014, Hahn & Nakhleh 2015, Liu et al. 2015, Simmons & Gatesy 2015, Tonini et al. 2015, Edwards et al. 2016), and we view our results here with caution.

Individual gene trees. All individual gene tree analyses identified the four major geographic clades, with the exception of the MT-CYB tree, which places *Siphonorhis brewsteri* within the South American Clade, instead of the Poorwill Clade (Supplemental Material Figure S4). The RAG-1 and GH gene trees are largely congruent, with the exception that GH places *Lyncornis* and *Eurostopodus* as sister taxa, rather

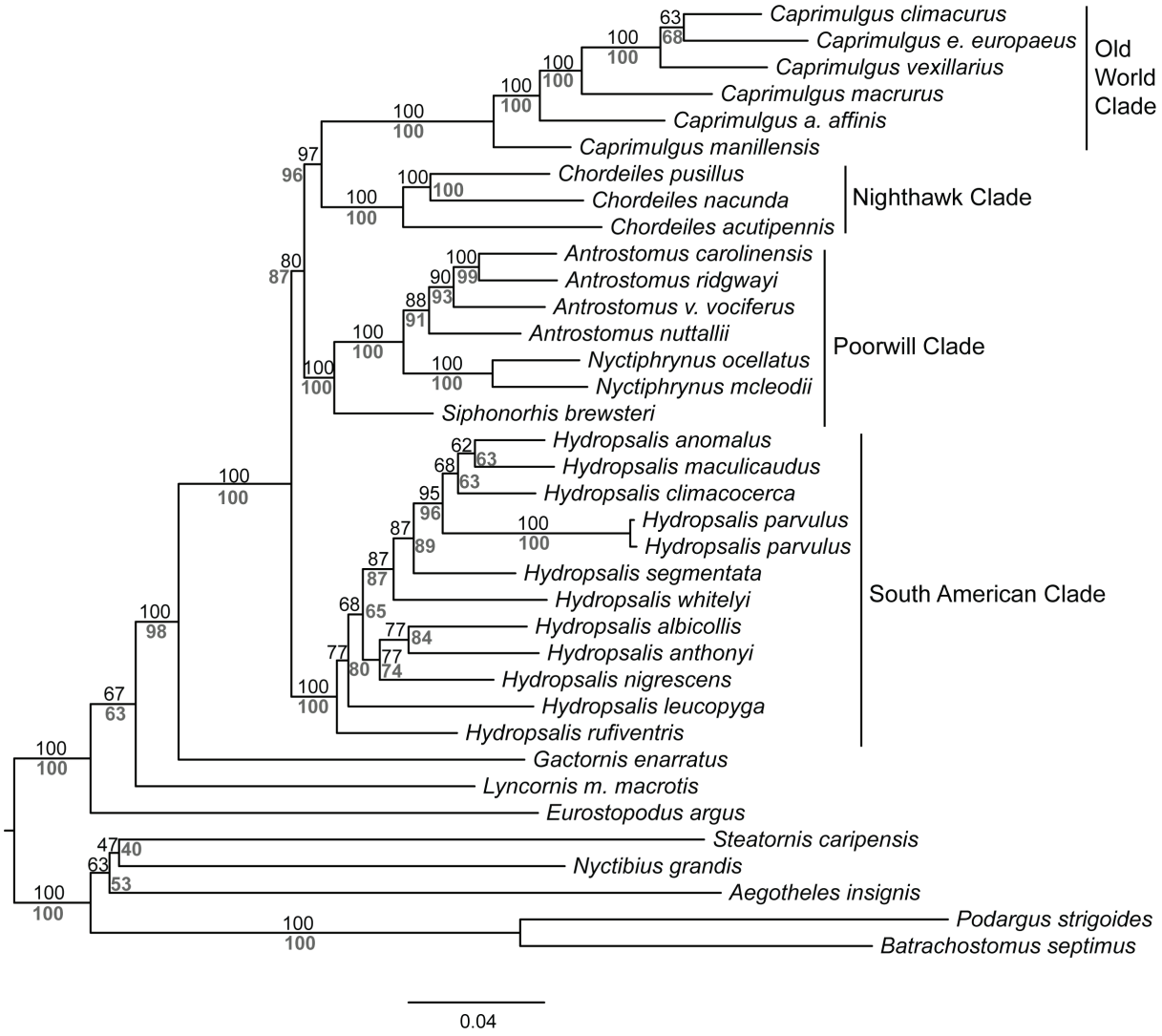

Figure 4. Best estimate of the phylogenetic tree for 36 caprimulgiform species from the data matrix partitioned by PartitionFinder as chosen by maximum likelihood score. Bootstrap support values shown for analysis under GTR + I + G above branches (black), below for analysis under model selected by PartitionFinder (grey) by AICc. Topologies were identical. Scale units are substitutions per site.

than successive basal branches (Supplemental Material Figures S1, S2). The MYC gene tree has the branching order of *Lyncornis* and *Eurostopodus* reversed from that seen in the RAG-1 tree, and fails to resolve the branching order of the Old World, Nighthawk and Poorwill Clades (Supplemental Material Figure S3). Overall, analyses of MT-CYB yield a very different topology than the nuclear genes, with much lower support (several nodes < 50% bootstrap). The lack of phylogenetic signal in this locus (at this depth) can be attributed to substitutional saturation. Also of note is the extremely long branch leading to the frogmouths *Batrachostomus septimus* and *Podargus strigoides* in the MT-CYB tree, suggesting rapid evolution of mtDNA in Podargidae. Earlier authors have also noted very high levels of mtDNA divergence in Podargidae (Cleere et al. 2007). The same phenomenon exists to a lesser extent in the MYC tree, but not in the RAG-1 or GH trees.

Basal taxa. We confirm that the genera *Lyncornis* and *Eurostopodus* are the earliest branching taxa in the Caprimulgidae. These genera were formerly lumped together in the genus *Eurostopodus,* and Sibley & Ahlquist (1990) suggested that this group be treated as a separate family. However, with increased taxon sampling, Han et al. (2010) uncovered additional complexity, detecting another early branching lineage, *Gactornis*, and the deep split in *Eurostopodus sensu lato* that justified resurrection of *Lyncornis.* While the 63% bootstrap support we found for *Eurostopodus sensu stricto* as the earliest branch is the highest support seen in any study so far, it still does not confidently resolve the trichotomy with *Lyncornis,* and our SVDquartets tree has the branching order reversed (Figure 5). The morphological evidence presented in Costa (2014; Figures 4 and 5 of that study) places *Lyncornis* as the most basal lineage with 31 % bootstrap support, with a paraphyletic

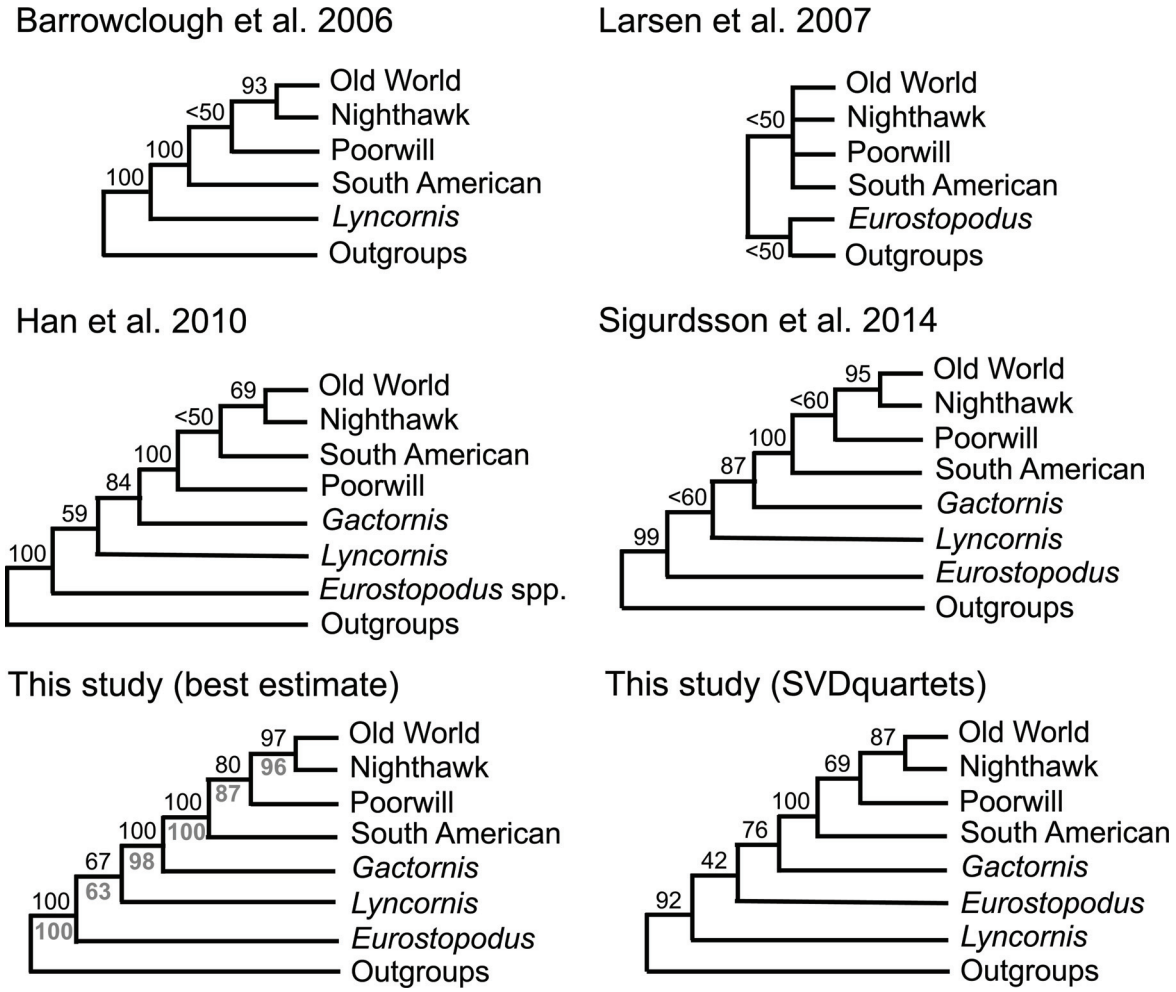

Figure 5. Caprimulgiform backbone trees with bootstrap support values from prior publications and this study obtained using maximum likelihood and coalescent (SVDquartets) analyses. Best estimate from this study depicted with bootstrap support values from analysis under GTR + I + G shown above branches (black), and with models selected by PartitionFinder shown below branches (grey). The word 'Clade' is omitted from major clade names for simplicity.

group of three *Eurostopodus* species branching next, and *Gactornis enarratus* sister to the Old World Clade. Thus, separation of any of these genera as distinct families or subfamilies cannot yet be justified. Both *Lyncornis* and *Eurostopodus* are on relatively long branches in all molecular trees, so sampling additional species from both genera may facilitate resolution.

Core caprimulgids. Barrowclough et al. (2006) and Sigurdsson & Cracraft (2014) had the South American Clade sister to the other three core caprimulgid clades, but with < 50% and < 60% bootstrap support, respectively. In contrast, Han et al. (2010) had the Poorwill Clade sister to the other three core clades, with equally low support (Figure 5). By increasing the size of the dataset, we have achieved stronger resolution of the major caprimulgid clades, and find the South American Clade to be sister to the other three. This is likely due to the congruent signal provided by RAG-1 and MYC, which are both slowly-evolving nuclear genes. Another nuclear locus, intron 9 from the aconitase gene, also yields this topology (Sigurdsson & Cracraft 2014), as does our SVDquartets analysis, leaving the GH gene tree as the only nuclear locus that does not support it. While our MT-CYB gene tree does not have this topology, it also does not agree with the GH gene tree nor with that of another mitochondrial locus, NADH dehydrogenase subunit 2 (see Sigurdsson & Cracraft 2014, Figure 1). As in previous studies, we find the Nighthawk Clade to be sister to the Old World Clade, with increased support (Figure 5).

Comparing the core caprimulgid lineages between our best tree and the Sigurdsson & Cracraft (2014) tree, we see a few slight differences in topology that are likely due to taxon sampling and / or genes used. Rearrangements of three terminal taxa (*Caprimulgus climacurus, europaeus and vexillarius*) are apparent within the Old World Clade and three others in the Poorwill Clade (*Antrostomus vociferus, carolinensis and ridgwayi*), though the topology of the early-branching members of each respective clade are the same between the trees. Identical topology and boot-

strap support values can be seen in both studies within the Nighthawk Clade. Basal within the South American Clade, our tree has *Hydropsalis leucopyga* and *H. rufiventris* branching in succession. This topology is present in Barrowclough et al. (2006), and Han et al. (2010), but the order is reversed in Sigurdsson & Cracraft (2014). *H. maculicaudus* is sister to *H. anomalus* in our best estimate tree, but sister to *H. climacocerca* in Sigurdsson & Cracraft (2014) and in our SVDquartets tree. Most of these differences are not strong conflicts in terms of nodal support, and may be addressed by increasing the size of the data matrix.

Nighthawks vs. nightjars*. Hydropsalis leucopyga* and *H. rufiventris* were formerly placed with the nighthawks in the subfamily Chordeilinae. All of these taxa have pointed wings and reduced rictal bristles, and forage on the wing at dusk and dawn by coursing rapidly over open spaces (*H. leucopyga* over water, *H. rufiventris* over forest, and *Chordeiles* over open country). *Eurostopodus* and *Lyncornis* share some of these traits (Cleere 1998) and were also placed in Chordeilinae by some authors (e.g., Holyoak 2001). However, all molecular evidence agrees in firmly placing *Eurostopodus* and *Lyncornis* as the earliest branches in the family, and *H. leucopyga* and *H. rufiventris* within the South American Clade. Thus, the morphological similarities among these taxa are homoplasious and likely represent independently derived adaptations to aerial foraging. The osteological study of Costa (2014) also recovered *Eurostopodus* and *Lyncornis* as early branches and *H. leucopyga* and *H. rufiventris* (represented by *H. semitorquatus*) as lineages distinct from *Chordeiles* (but not monophyletic with the South American Clade).

Biogeography. The three earliest branches of the caprimulgid phylogeny have current distributions around the Indian Ocean; *Eurostopodus* in Australo-Papua, *Lyncornis* in South Asia and *Gactornis* in Madagascar (Cleere 1998). We can thus infer that the family may have originated in this general region. On the other hand, the two earliest branching of the four major caprimulgid clades are restricted to the New World (South American and Poorwill), while the Old World Clade is nested in the phylogeny sister to the Nighthawk Clade. Thus, diversification of the core caprimulgids in the New World appears likely, followed by a re-colonization and secondary radiation in the Old World. These scenarios were previously envisioned by Barrowclough et al. (2006), Han et al. (2010), and Sigurdsson & Cracraft (2014), and are reinforced here with increased resolution and support at key nodes in the phylogeny.

ACKNOWLEDGMENTS

We thank Matthew Kweskin for technical assistance on computational analyses, Laura Kubatko for discussion of SVDquartets, and David L. Swofford and Robert Lanfear for discussion of issues related to partitioning. We thank Charles Mitter for constructive comments on an early draft of this manuscript, Minh Le for assistance with some of the RAG-1 sequencing for this project, Thiago V. V. da Costa for sharing his unpublished dissertation, and Natalia Agudelo and Gustavo S. Cabanne for translating the abstract. NDW received Smithsonian Institution support from the Scholarly Studies Program, from the Consortium for Understanding and Sustaining a Biodiverse Planet, and as a Predoctoral Fellow. Computations in this paper were run on the Smithsonian Institution's High Performance Cluster.

REFERENCES

Akaike, H (1973) Information theory and extension of the maximum likelihood principle. *Proceedings of the Second International Symposium on Information Theory*: 267–281.

Barrowclough, GF, JG Groth & LA Mertz (2006) The RAG-1 exon in the avian order Caprimulgiformes: phylogeny, heterozygosity, and base composition. *Molecular Phylogenetics & Evolution* 41: 238–248.

Braun, MJ & CJ Huddleston (2009) A molecular phylogenetic survey of caprimulgiform nightbirds illustrates the utility of non-coding sequences. *Molecular Phylogenetics & Evolution* 53: 948–960.

Bühler, P (1970) Schädelmorphologie und Kiefermechanik der Caprimulgidae (Aves). *Zeitschrift für Morphologie der Tiere* 66: 337–399.

Bull, JJ, JP Huelsenbeck, CW Cunningham, DL Swofford & PJ Waddell (1993) Partitioning and combining data in phylogenetic analysis. *Systematic Biology* 42: 384–397.

Chifman, J & L Kubatko (2014) Quartet inference from SNP data under the coalescent model. *Bioinformatics* 30: 3317–3324.

Chifman, J & L Kubatko (2015) Identifiability of the unrooted species tree topology under the coalescent model with time-reversible substitution processes, site-specific rate variation, and invariable sites. *Journal of Theoretical Biology* 374: 35–47.

Cleere, N (1998) *Nightjars: a guide to the nightjars, nighthawks, and their relatives*. Yale Univ. Press, New Haven, Connecticut, USA.

Cleere, N, AW Kratter, DW Steadman, MJ Braun, CJ Huddleston, CE Filardi & G Dutson (2007) A new genus of frogmouth (Podargidae) from the Solomon Islands – results from a taxonomic review of *Podargus ocellatus inexpectatus* Hartert 1901. *Ibis* 149: 271–286.

Costa, TVV (2014) Osteology and phylogeny of the avian order Caprimulgiformes, with special emphasis on Nyctibiidae and Caprimulgidae. Ph.D. diss., Univ. de São Paulo, São Paulo, Brazil.

de Queiroz, A (1993) For consensus (sometimes). *Systematic Biology* 42: 368–372.

Edwards, SV, Z Xi, A Janke, BC Faircloth, JE McCormack, TC Glenn, B Zhong, S Wu, EM Lemmon, AR Lemmon, AD Leache, L Liu & CC Davis (2016) Implementing and testing the multispecies coalescent model: a valuable paradigm for phylogenomics. *Molecular Phylogenetics and Evolution* 94: 447–462.

Faircloth, BC, JE McCormack, NG Crawford, MG Harvey, RT Brumfield & TC Glenn (2012) Ultraconserved elements anchor thousands of genetic markers spanning multiple evolutionary timescales. *Systematic Biology* 61: 717–726.

Galtier, N & M Gouy (1998) Inferring pattern and process: maximum-likelihood implementation of a nonhomogeneous model of DNA sequence evolution for phylogenetic analysis. *Molecular Biology and Evolution* 15: 871–879.

Gatesy, J & MS Springer (2014) Phylogenetic analysis at deep timescales: unreliable gene trees, bypassed hidden support, and the coalescence / concatalescence conundrum. *Molecular Phylogenetics and Evolution* 80: 231–266.

Groth, JG & GF Barrowclough (1999) Basal divergences in birds and the phylogenetic utility of the nuclear RAG-1 gene. *Molecular Phylogenetics and Evolution* 12: 115–123.

Hahn, MW & L Nakhleh (2015) Irrational exuberance for resolved species trees. *Evolution* 70: 7–17.

Han, K-L, MB Robbins & MJ Braun (2010) A multi-gene estimate of phylogeny in the nightjars and nighthawks (Caprimulgidae). *Molecular Phylogenetics and Evolution* 55: 443–453.

Hoff KM (1966) A comparative study of the appendicular muscles of Strigiformes and Caprimulgiformes. Ph.D. diss., Washington State Univ., Pullman, Washington, USA.

Holyoak, DT (2001) Nightjars and their allies: the Caprimulgiformes. *In* Perrins, CM, J Kikkawa & WJ Bock (eds). *Bird families of the world*. Oxford University Press, New York, New York, USA.

Hurvich, CM & C-L Tsai (1989) Regression and time series model selection in small samples. *Biometrika* 76: 297–307.

Jarvis, ED, S Mirarab, AJ Aberer, B Li, P Houde, C Li, SYW Ho, BC Faircloth, B Nabholz, JT Howard, A Suh, CC Weber, RR da Fonseca, J Li, F Zhang, H Li, L Zhou, N Narula, L Liu, G Ganapathy, B Boussau, MS Bayzid, V Zavidovych, S Subramanian, T Gabaldón, S Capella-Gutiérrez, J Huerta-Cepas, B Rekepalli, K Munch, M Schierup, B Lindow, WC Warren, D Ray, RE Green, MW Bruford, X Zhan, A Dixon, S Li, N Li, Y Huang, EP Derryberry, MF Bertelsen, FH Sheldon, RT Brumfield, CV Mello, PV Lovell, M Wirthlin, MPC Schneider, F Prosdocimi, JA Samaniego, AMV Velazquez, A Alfaro-Núñez, PF Campos, B Petersen, T Sicheritz-Ponten, A Pas, T Bailey, P Scofield, M Bunce, DM Lambert, Q Zhou, P Perelman, AC Driskell, B Shapiro, Z Xiong, Y Zeng, S Liu, Z Li, B Liu, K Wu, J Xiao, X Yinqi, Q Zheng, Y Zhang, H Yang, J Wang, L Smeds, FE Rheindt, M Braun, J Fjeldsa, L Orland o, FK Barker, KA Jønsson, W Johnson, K-P Koepfli, S O'Brien, D Haussler, OA Ryder, C Rahbek, E Willerslev, GR Graves, TC Glenn, J McCormack, D Burt, H Ellegren, P Alström, SV Edwards, A Stamatakis, DP Mindell, J Cracraft, EL Braun, T Warnow, W Jun, MTP Gilbert & G Zhang (2014) Whole-genome analyses resolve early branches in the tree of life of modern birds. *Science* 346: 1320–1331.

Kumar, S, AJ Filipski, FU Battistuzzi, SL Kosakovsky Pond & K Tamura (2012) Statistics and truth in phylogenomics. *Molecular Biology and Evolution* 29: 457–472.

Lanfear, R, B Calcott, SYW Ho & S Guindon (2012) PartitionFinder: combined selection of partitioning schemes and substitution models for phylogenetic analyses. *Molecular Biology and Evolution* 29: 1695–1701.

Larsen, C, M Speed, N Harvey & HA Noyes (2007) A molecular phylogeny of the nightjars (Aves: Caprimulgidae) suggests extensive conservation of primitive morphological traits across multiple lineages. *Molecular Phylogenetics and Evolution* 42: 789–796.

Lemmon, AR, SA Emme & EM Lemmon (2012) Anchored hybrid enrichment for massively high-throughput phylogenomics. *Systematic Biology* 61: 727–744.

Liu L, Z Xi, S Wu, CC Davis & SV Edwards (2015) Estimating phylogenetic trees from genome-scale data. *Annals of the New York Academy of Sciences* 1360: 36–53.

Mariaux, J & MJ Braun (1996) A molecular phylogenetic survey of the nightjars and allies (Caprimulgiformes) with special emphasis on the potoos (Nyctibiidae). *Molecular Phylogenetics and Evolution* 6: 228–244.

Mayr, G (2002) Osteological evidence for paraphyly of the avian order Caprimulgiformes (nightjars and allies). *Journal of Ornithology* 143: 82–97.

Mayr, G, A Manegold & US Johansson (2003) Monophyletic groups within 'higher land birds' - comparison of morphological and molecular data. *Journal of Zoological Systematics and Evolutionary Research* 41: 233–248.

Mayr, G (2010) Phylogenetic relationships of the paraphyletic 'caprimulgiform' birds (nightjars and allies). *Journal of Zoological Systematics and Evolutionary Research* 48: 126–137.

McCormack, JE, BC Faircloth, NG Crawford, PA Gowaty, RT Brumfield & TC Glenn (2012) Ultraconserved elements are novel phylogenomic markers that resolve placental mammal phylogeny when combined with species-tree analysis. *Genome Research* 22: 746–754.

Oberholser, HC (1914) A monograph of the genus *Chordeiles* Swainson, type of a new family of goatsuckers. *U.S. National Museum Bulletin* 86: 1–123.

Peters, JL (1940) *Checklist of birds of the world. volume 4: Strigiformes.* Harvard Univ. Press, Cambridge, Massachusetts, USA.

Posada, D (2008) jModelTest: phylogenetic model averaging. *Molecular Biology and Evolution* 25: 1253–1256.

Prum, RO, JS Berv, A Dornburg, DJ Field, JP Townsend, EM Lemmon & AR Lemmon (2015) A comprehensive phylogeny of birds (Aves) using targeted next-generation DNA sequencing. *Nature* 526: 569–573.

Ridgway, R (1914) *The birds of North and Middle America. part VI.* Bulletin of the United States National Museum, No. 50.

Robinson, DF & LR Foulds (1981) Comparison of phylogenetic trees. *Mathematical Biosciences* 53: 131–147.

Schodde, R & IJ Mason (1980) *Nocturnal birds of Australia.* Lansdowne, Melbourne, Australia.

Sequencher® version 5.4.1 sequence analysis software, Gene Codes Corporation, Ann Arbor, Michigan, USA.

Sibley, CG & JE Ahlquist (1990) *Phylogeny and classification of birds: a study in molecular evolution*. Yale Univ. Press, New Haven, Connecticut, USA.

Sigurdsson, S & J Cracraft (2014) Deciphering the diversity and history of New World nightjars (Aves: Caprimulgidae) using molecular phylogenetics. *Zoological Journal of the Linnean Society* 170: 506–545.

Simmons, MP & J Gatesy (2015) Coalescence vs. concatenation: sophisticated analyses vs. first principles applied to rooting the angiosperms. *Molecular Phylogenetics and Evolution* 91: 98–122.

Sukumaran, J & MT Holder (2010) DendroPy: a python library for phylogenetic computing. *Bioinformatics* 26: 1569–1571.

Swofford, DL (2003) PAUP*. Phylogenetic Analysis Using Parsimony (*and other methods). Version 4. Sinauer Associates, Sunderland, Massachusetts, USA.

Thompson, JD, TJ Gibson, F Plewniak, F Jeanmougin & DG Higgins (1997) The CLUSTAL_X windows interface: flexible strategies for multiple sequence alignment aided by quality analysis tools. *Nucleic Acids Research* 25: 4876–4882.

Tonini, J, A Moore, D Stern, M Shcheglovitova & G Orti (2015) Concatenation and species tree methods exhibit statistically indistinguishable accuracy under a range of simulated conditions. *PLoS Currents* 2015 March 9; 7: ecurrents.tol.34260cc27551a527b124ec5f6334b6be; doi: 10.1371/currents.tol.34260cc27551a527b124ec5f6334b6be.

Tukey, JW (1977) *Exploratory data analysis*. Addison-Wesley, Reading, Massachusetts, USA.

Wetmore, A (1918) On the anatomy of *Nyctibius* with notes on allied birds. *Proceedings of the U.S. National Museum* 54:577–586.

Whitney, WM, JF Pacheco, PSM da Fonseca, RE Webster, GM Kirwan & JM Barnett (2003) Reassignment of *Chordeiles vielliardi* Lencioni-Neto, 1994, to *Nyctiprogne* Bonaparte, 1857, with comments on the latter genus and some presumably related chordeilines (Caprimulgidae). *Bulletin of the British Ornithology Club* 123: 103–112.

Zwickl, DJ (2006) Genetic algorithm approaches for the phylogenetic analysis of large biological sequence datasets under the maximum likelihood criterion. Ph.D. diss., Univ. of Texas, Austin, Texas, USA.