Timeframes of Speciation, Reticulation, and Hybridization in the Bulldog Bat Explained Through Phylogenetic Analyses of All Genetic Transmission Elements

FARSAI ALI ANWARALIKHAN1,2,*, CALEB D. PHILLIPS1, and ROBERT J. BAKER1

1Department of Biological Sciences and the Museum, Texas Tech University, Lubbock, TX 79409, USA and 2Department of Zoology, Faculty of Resource Science and Technology, Universiti Malaysia Sarawak, 94300 Kota Samarahan, Sarawak, Malaysia

*Correspondence to be sent to: Department of Zoology, Faculty of Resource Science and Technology, Universiti Malaysia Sarawak, 94300 Kota Samarahan, Sarawak, Malaysia; E-mail: afaisalfrst.unimas.my.

Received 22 February 2013; reviews returned 19 April 2013; accepted 16 September 2013

Associate Editor: Thomas Near

Abstract.—Phylogenetic comparisons of the different mammalian genetic transmission elements (mtDNA, X-, Y-, and autosomal DNA) is a powerful approach for understanding the process of speciation in nature. Through such comparisons the unique inheritance pathways of each genetic element and gender-based processes can link genomic structure to the evolutionary process, especially among lineages which have recently diversified, in which genetic isolation may be incomplete. Bulldog bats of the genus Noctilio are an exemplar lineage, being a young clade, widely distributed, and exhibiting unique feeding ecologies. In addition, currently recognized species are paraphyletic with respect to the mtDNA gene tree and contain morphologically identifiable clades that exhibit mtDNA divergences as great as among many species. To test taxonomic hypotheses and understand the contribution of hybridization to the extant distribution of genetic diversity in Noctilio, we used phylogenetic, coalescent stochastic modeling, and divergence time estimates using sequence data from cytochrome-b, cytochrome c oxidase-I, zinc finger Y, and zinc finger X, as well as evolutionary reconstructions based on amplified fragment length polymorphisms (AFLPs) data. No evidence of ongoing hybridization between the two currently recognized species was identified. However, signatures of an ancient mtDNA capture were recovered in which an mtDNA lineage of one species was captured early in the noctilionid radiation. Among subspecific mtDNA clades, which were generally coincident with morphology and statistically definable as species, signatures of ongoing hybridization were observed in sex chromosome sequences and AFLP. Divergence dating of genetic elements corroborates the diversification of extant Noctilio beginning about 3 Ma, with ongoing hybridization between mitochondrial lineages separated by 2.5 myr. The timeframe of species’ divergence within Noctilio supports the hypothesis that shifts in the dietary strategies of gleaners insects (N. albiventris) or fish (N. loriculus) are among the most rapid instances of dietary evolution observed in mammals. This study illustrates the complex evolutionary dynamics shaping gene pools in nature, how comparisons of genetic elements can serve for understanding species boundaries, and the complex considerations for accurate taxonomic assignment. [AFLP; genetic transmission elements; hybridization; Noctilio; X-chromosome; Y-chromosome.]

The process of speciation is frequently studied through identifying and comparing patterns of variability in morphology, ecology, and genetic elements. A large portion of species-level taxonomy is based on a narrow array of data types, such as morphology and mitochondrial phylogenies, leaving much evolutionary diversity and complexity yet to be discovered. It is hypothesized that the number of valid mammalian species yet to be described remains in the thousands (Baker and Bradley 2006; Clare et al. 2007; Francis et al. 2010; Clare 2011). This estimate is based partially on the observation that detailed morphological and mitochondrial studies often identify cryptic species. However, complex evolutionary histories involving introgression and hybridization are also described, and the incorporation of genetic data from differentially inherited genetic elements is valuable for characterizing such phenomena (e.g., Good et al. 2003, 2008; Buckley et al. 2006; Linnen and Farrell 2008; Mauereir-Butler et al. 2008; Bossu and Near 2009; Larsen et al. 2010; Reid et al. 2012). Results such as these are refining and redefining our understanding of the process of speciation and the complexity of systematic biodiversity. Inference of species trees is often complicated by incongruency among gene trees (Knowles 2009). In order to understand the process of speciation it is important to identify the sources of phylogenetic incongruency (Reid et al. 2012). The lineage sorting process and the stochastic nature of genetic mutation are well-known phenomena that can result in phylogenetic incongruency (Familio and Nei 1988; Takahata 1989; Degnan and Rosenberg 2006). Incongruency based on these phenomena is not independent from effects of divergence time, drift, selection, and demography. In contrast, phylogenetic incongruency as a result of hybridization identifies on-going or past gene flow (Rusebarg et al. 2000). It is expected that incongruency arising via hybridization events can be detected in the study of recent adaptive radiations, as these complexes are more likely to exhibit incomplete reproductive isolation given recency of divergence. It is also true that lineage sorting and stochastic mutational processes can be influential in phylogenetic reconstructions of recent divergences. Owing to the different inheritance pathways of the various autosomal, sex, and organelle chromosomes, phylogenetic comparisons among these genetic elements can be used to identify and characterize hybridization. Specifically, comparative phylogenetic analysis of genetic variation distributed across genetic elements yields perspectives about the presence and nature of hybridization that is distinguishable from sorting or mutational effects on individual loci.
Studies based on ecology, morphology, and genetics indicate the bulldog bat genus Noctilio, of the monotypic family Noctilionidae, is a young clade, vagile, and widely distributed. Comparative phylogenetic analysis of the mammalian genetic transmission elements could document details of this history and timeframes of divergence and hybridization. Noctilio consists of two species, *N. albiventris* and *N. leporinus*. The two species are broadly sympatric throughout their ranges in Central and South America, and *N. leporinus* is also distributed across the Caribbean Islands. Although these species are sympatric and display minimal morphological shape differences, they can be differentiated based on size, with *N. leporinus* being considerably larger than *N. albiventris* (Hood and Pitocchelli 1983; Hood and Jones 1984). Although *N. leporinus* feeds broadly on fish, flying insects, and aquatic invertebrates, *N. albiventris* primarily is insectivorous. However, there is evidence from stomach content analyses showing that *N. albiventris* is able to consume fish and fruit (Howell and Burch 1974; Hood and Pitocchelli 1983; Hood and Jones 1984; Bordignon 2006; Gonçalves et al. 2007). As no other member of the superfamily Noctilionoidea (consisting of 7 families and an estimated 180 species) has developed a piscivorous life history strategy, and *N. leporinus* still retains a partially insectivorous dietary strategy, it is likely that piscivory has evolved recently in this genus (Lewis-Oritt et al. 2001; Pavan et al. 2013).

Systematic studies have indicated a complex yet poorly defined evolutionary history for Noctilio. This complexity has been partially characterized by multiple subspecific phylogenetic clades within *N. albiventris*, some of which may be valid species and retain a primarily insectivorous dietary strategy (Lewis-Oritt et al. 2001).

*Noctilio leporinus* consists of two moderately diverged (≤ 3.2% cytochrome-b genetic distance) clades (Lewis-Oritt et al. 2001; Clare et al. 2007). Taxonomic revisions of both species have been based on bivariate analyses of morphological characteristics (Davis 1973, 1976; Timm and Genoways 2003). These studies divided *N. albiventris* into the four subspecies *N. a. affinis*, *N. a. albiventris*, *N. a. cabrerai*, and *N. a. major*. However, *N. a. affinis* was later subsumed and considered as the junior synonym of *N. a. albiventris* based on the finding of specimens with intermediate morphology from French Guiana (Simmons and Voss 1998; Simmons 2005; Gardner 2008).

*Noctilio leporinus* has been partitioned by morphology into the three distinct subspecies *N. l. leporinus*, *N. l. mastivus*, and *N. l. refuscens* (Davis 1973, 1976). Results from these studies of subspecific variation document the process of a dynamic and ongoing evolutionary radiation in bulldog bats, and the extent of ongoing gene flow among these evolutionary clades remains unknown.

Detailed phylogenetic characterization of this complex would not only yield an improved estimate of biodiversity for critical ecoregions, but would also improve our understanding of the process of speciation in recently radiated and highly vagile mammalian lineages. The study of all genetic transmission elements in this evolutionary background would inform timeframes of divergence across the different genetic transmission elements and would characterize hybridization among diverging lineages. To this end, patterns of genetic variation at mitochondrial, X- and autosomal chromosome loci were compared using several phylogenetic measures. These comparisons provide a method for identifying and characterizing the evolutionary legacies of stochastic sorting, introgression, and hybridization in recent radiations that avoids the pitfalls of less comprehensive genomic sampling.

**MATERIALS AND METHODS**

**Taxonomic Sampling**

Tissue samples were obtained from the following collections: Museum of Texas Tech University (TT, Lubbock), Royal Ontario Museum (ROM, Toronto), Louisiana Museum of Natural History (LSUMZ, Baton Rouge), American Museum of Natural History (AMNH, New York), National Museum of Natural History (USNM, Washington, D.C.), Museum of Vertebrate Zoology (MVZ, Berkeley), Sam Noble Oklahoma Museum of Natural History (SNOMNH, Norman), Museum of Southwestern Biology (MSB, Albuquerque) and Carnegie Museum of Natural History (CMNH, Pittsburgh). Specimens were selected based on each museum’s recorded locality information in order to sample the geographic range of the four subspecies of *N. albiventris* and three subspecies of *N. leporinus* (Davis 1973, 1976; Fig. 1). Online Appendix 1 (doi: 10.5061/dryad.77c1h) contains information about museum voucher specimen/tissue number and associated GenBank accession numbers for cytochrome-b (Cyt b), cytochrome c oxidase-I (COI), zinc finger X (ZFX), zinc finger Y (ZFY), and Amplified Fragment Length Polymorphisms (AFLPs). *Carollia perspicillata*, *Mormoops megalophylla*, *Pteronotus parnellii*, *P. personatus*, *P. quadridens* were also sequenced and genotyped, and used as outgroups (Online Appendix 1).

**Mitochondrial DNA Sequencing**

Total genomic DNA was extracted from muscle or liver tissues by standard phenol/chloroform methods (Sambrook et al. 1989). The entire Cyt b gene (1140 bp; n = 83) and a portion of the COI gene (657 bp; n = 82) were PCR amplified and sequenced to explore mitochondrial (mtDNA) variation within the sampling of *N. albiventris* and *N. leporinus*. Amplification for Cyt b was performed using primers LGL765/LGL766 (Bickham et al. 2004). Amplifications of COI included multiple primer pairs (VF1/VR1, VF1i/VR1i, and VF1d/VR1d; Ivanova et al. 2006). Twenty-five μL reaction volumes included ~200 ng DNA, 0.12 μM of each primer, 1.5 mM MgCl2, 0.012 mM deoxynucleoside triphosphates, 1X reaction buffer, 0.32 mg/ml Bovine Serum Albumin, and 0.625U
California) and final products were reconstituted in QiaQuick Gel Extraction Kits (Qiagen Inc., Chatsworth, California). ZFY amplicons were gel excised and purified using M5L, (Hoffmann and Baker, 2001); MVZ04, (Smith and Patton, 1993). COI amplicons were sequenced using ABI Big Dye chemistry version 3.1, and an ABI 3100-Avant Genetic Analyzer (PE Applied Biosystems, Foster City, California). The entire Cyt b gene was sequenced for 30 µL of buffer EB. Portions of ZFX (819–847 bp) and ZFY (733–754 bp) were sequenced using internal sequencing primers developed for Noctilio: ZFX-NocF 5′TGGAATGAAAAATCCTCAA3′ and ZFX-NocR 5′TTTCTGAGATGGAAATGGA3′ for ZFX, and ZFY-NocF: 5′GGCATATGCGGATCAT3′ and ZFY-NocR: 5′CTGAAACATTTACTGGTTCCAAGA3′ for ZFY. Allele-specific priming was not required for ZFX sequences, as only one individual was observed to be heterozygous.

Y- and X-Chromosome Gene Sequencing

A total of 28 male Noctilio were sequenced for a portion of the paternally inherited ZFY gene, and 33 (8 female and 25 male) were sequenced for a portion of the bi-parentally inherited ZFX. Sampling included at least one representative from each phylogroup identified in the mtDNA data set, except for ZFX, and only fragments (50–400 bp in length) with a signal larger than 50 relative fluorescent units were retained as confident genotype calls. GeneMapper version 4.0 software (Applied Biosystems) was used to manually edit and verify all data. From the resulting binary data matrix, descriptive statistics were calculated in GeneAIEx version 6 software (Peakall and Smouse 2006). Error rates were calculated following (Bonis et al., 2004) using 50 replicated samples (~15% of the overall sample size).

Sequence Variation and Study Comparisons

Initial analysis of the data addressed various genetic summary statistics, substitution trends, and overall comparisons of the geographic distribution of variation. To this end, DnaSPv5 (Librado and Rozas 2009) was used to resolve the number of unique haplotypes and polymorphic sites for each locus. Substitution saturation was evaluated for each locus using the Xia index, calculated in DAMBE version 5.2.69 (Xia and Xie 2001; Xia et al. 2003). For the mtDNA portion of the data set (Cyt b and COI), Kimura 2-parameter (K2P) distance matrices were generated in MEGA version 5 (Tamura et al. 2011). K2P distance values for these genes were calculated to describe genetic distances separating major clades, and overall comparisons of the geographic distribution of variation. For this analysis, mtDNA clades were compared to sequences from Cyt b and COI previously deposited on GenBank (Lewis-Oritt et al. 2001; Clare et al. 2007; Cyt b: EF330794-806; COI: EF080519–522, 524, 525, 527, 528, 530–533, JF448375) were incorporated to increase sampling and improve assessment of the geographic range for each phylogroup.
identified and relation to proposed distributions of subspecies (Davis 1973, 1976).

Phylogenetic Analyses

Phylogenetic reconstructions were initially performed for each mtDNA gene separately to assess phylogenetic congruence among gene tree topologies. Mitochondrial genes were subsequently concatenated into unique composite haplotypes. Maximum-Likelihood and Bayesian phylogenetic reconstructions were generated for each gene, the concatenated mtDNA data set, and sex chromosome markers. Appropriate models of molecular evolution for each gene, concatenated mtDNA data set, and sex chromosome markers were determined using the Bayesian Information Criterion in MEGA version 5 (Tamura et al. 2011), and the best-fit model implemented the Bayesian Information Criterion in MEGA version 5 and sex chromosome markers were determined using the Bayesian Information Criterion in MEGA version 5 for each gene, concatenated mtDNA data set, and sex chromosome markers. Maximum-Likelihood and Bayesian phylogenetic analyses were conducted in PhyML version 3.0 (Guindon and Gascuel 2003) and MrBayes. Phylogenetic analyses were conducted in PhyML version 3.0 (Guindon and Gascuel 2003) and MrBayes. Maximum-Likelihood analyses were initiated using the BioNJ (Gascuel 1997) starting trees. Bootstrap support values were obtained through 1000 subsampling replicates. Bayesian phylogenetic reconstructions were performed using four independent MCMC chains consisting of two independent runs at 10 million generations each, and trees were logged every 100th iteration. As recommended by the program authors, iterations were performed until the average standard deviation of split frequencies was less than 0.01 (Ronquist et al. 2005). The stability of log likelihoods for 5000 trees was examined in TRACER version 1.5 (Rambaut and Drummond 2007) and the first 10,000 trees were discarded as burn-in.

Given the strong morphological signal separating N. albiventris from N. leporinus (Davis 1973, 1976; Timm and Genoways 2003), it was possible that previous phylogenetic analyses on this group reporting polyphyly (Lewis-Oritt et al. 2001; Pavan et al. 2013) were the result of outgroup confounding effect, given the evolutionary distance of available outgroups for Noctilio (Holland et al. 2003). To examine this condition, additional Bayesian phylogenetic analyses were conducted in MrBayes without outgroups (Ronquist et al. 2005), and implementing a strict molecular clock. Huelsenbeck et al. 2002 have previously demonstrated the ability of this method to determine proper rooting.

The AFLP binary matrix was transformed into Nei–Li genetic distances and used to generate a Neighbor-Joining tree (NJ) with 1000 bootstrap replicates in PAUP version 4.0b10 (Swofford 2002). Bayesian analysis was performed for the AFLP data set using the restriction site model in MrBayes that allowed correction for coding bias (Ronquist et al. 2005). We used the option “noabsencesites”, which most closely fits the bias in AFLP data sets (Ronquist et al. 2005). The Dirichlet prior for the state frequencies was set to (0.38, 0.62), matching the actual 0/1 frequencies in the data set. Iterations and final phylogeny estimation followed the same steps described for the DNA sequence data. Principal coordinate analysis (PCoA) was performed on the AFLP data matrix using GenAIEx to describe multidimensional relationships and major patterns of genetic variance within and among mtDNA lineages. The first three components from PCoA were used in generating a 3D scatterplot with the minimum spanning network option in MYSTAT version 12 (Systat Software, Inc., USA). Ellipses were drawn around major geographic clusters.

Because phylogenetic comparisons across loci disclosed patterns of hybridization among subspecies within each species, a formal statistical test for the presence of hybrid individuals was conducted using the AFLP data matrix. For this analysis, NewHybrids version 1.1 (Anderson and Thompson 2002) was used to determine posterior probabilities that individuals belonged to a hybrid category (F1, F2, backcrosses). Hybrid testing was conducted for N. leporinus and N. albiventris separately. For N. leporinus this analysis conformed to a two-parental-type system as only two major genetic clades were disclosed through all phylogenetic analyses. Within N. albiventris for mtDNA, clades were resolved and phylogenetic reconstructions based on AFLP indicated mixed ancestry. Therefore hybrid analysis for this species was also conducted as a two-parental-type system. Although this design was potentially too simplistic for N. albiventris, a history which is most robustly diagnosed through phylogenetic and variable reduction methods (described above), formal hybrid classification is limited to a two-parental-type system. The analysis consisted of a total of 5000 sweeps, in which the first 500 sweeps were removed as burn-in. The default recommendation of a uniform prior was imposed for priors on mixing proportions and allele frequencies, and no a priori assumptions were made for class membership of samples.

Divergence Time Estimates

Divergence times were estimated separately for unique haplotypes from each locus including 67 concatenated mtDNA haplotypes, 8 ZFY haplotypes, and 13 ZFX haplotypes. Molecular clock testing was performed using Likelihood Ratio Tests (Felsenstein 1988) implemented in MEGA. Following the results of these tests a relaxed lognormal molecular clock was implemented for the mtDNA data set and strict molecular clocks for ZFX and ZFY. BEAST version 1.7.1 (Drummond and Rambaut 2007) was used to estimate timescales of diversification for Noctilio using both paleontological and molecular constraints. The oldest fossil indistinguishable from modern N. albiventris
is known from the middle Miocene La Venta fauna, ∼12.5 Ma in Colombia (Czaplewski 1996; Czaplewski et al. 2003). However, this fossil was not used to define a node prior, as it would have constrained the time of diversification of Noctilio, which is uncertain due to gaps in N. leporinus fossil record (Meredith et al. 2011). Fossil priors demarcating the Mormoopidae/Phyllostomidae split recovered from the Whitneyan land deposits in Florida were used as an outgroup node prior (dated late Oligocene; Morgan 2002). This node was constrained with an exponentially distributed prior having a minimum age of 28.5 Ma with an exponential mean of 3.2, allowing the maximum age of this divergence to occur ∼40.6 Ma following (Meredith et al., 2011). The molecular date (= 43.01; SD = 1.5) for the node shared by the families Furipteridae, Mormoopidae, Noctilionidae, and Phyllostomidae in (Meredith et al., 2011) was also used in our analyses and was described by a normally distributed prior.

Phylogeny and time to most recent common ancestor (tmrca) for Noctilio nodes were estimated for the concatenated mtDNA data matrix partitioned by gene. Previously estimated models of evolution (Cyt b: HKY+I+G, COI: HKY+G) were applied with substitution rate parameters, rate heterogeneity model, and base frequencies unlinked across codon positions. Analyses for ZFX and ZFY intronic sequences implemented the HKY+I and HKY models of evolution, respectively. Cardilla, Mormoops, and Pirostomus were used as outgroups in phylogenetic reconstructions. Node dates were estimated using a birth–death process prior as proposed for reconstructing phylogenies not including fossil lineages (Gernhard 2008). Analyses consisted of two independent runs of 5 × 10^6 iterations for the mtDNA data set, and two independent runs of 10 × 10^6 iterations for the ZFX and ZFY data sets, with every 1000th iteration logged for all analyses. Independent runs were combined for each data set using LogCombiner version 1.7 (Drummond and Rambaut 2007). TRACER version 1.5 (Rambaut and Drummond 2007) was used to determine appropriate burn-in (∼10%), examine convergence, effective sample sizes (ESSs), and 95% highest probability density intervals (HPD) of constrained priors. To compare effects of implementing different speciation models, additional analyses were conducted using a Yule specie prior (Yule 1925) for the mtDNA (including only two random unique haplotypes per clade) and ZFY data sets (including all unique haplotypes) with 50 × 10^6 iterations performed for each data set, with every 1000th iteration logged. Results from both speciation models were compared for molecular dating incongruences.

Species Delimitations Using Coalescent Model
In order to identify independently coalescing lineages that may correspond to species without a priori assumptions of population membership, we used the General Mixed Yule-Coalescent Model (Pons et al. 2006). We used bGMYC, a Bayesian implementation of the model that accounts for uncertainty in the phylogeny and model parameters by sampling trees and conducting MCMC (Reid and Carstens 2012). For this analysis the ultrametric concatenated mtDNA trees for unique haplotypes from the BEAST analyses were analysed. The two original tree files with 50 million generations were combined after discarding the first 5 million as burn-in. All parameters for all replicates retained ESSs greater than 200. The combined tree file was resampled at a lower frequency resulting in 90 trees (one tree every 1,000,000 generations). Initial bGMYC analyses using only a single tree from the population of trees suggested that the MCMC reached stationarity after 10,000 steps. All 90 trees were analysed for 50,000 generations, discarding the first 40,000 as burn-in and sampling every 100th iteration. MCMC samples from each tree were pooled together to calculate the probabilities that any two leaves in the phylogeny were conspecific. We used a probability of being conspecific of 0.5 or higher to cluster individuals into putative species. The 50% threshold value was selected because the number of clades recognized as conspecifics using this threshold can be readily supported by existing morphological data, and corresponds to well-supported genetic clades. This was considered a conservative measure as compared with implementing larger threshold values which resulted in identifying considerably shallower coalescing clades as conspecific.

RESULTS
Mitochondrial Variation and Phylogenies
Cyt b and COI displayed similar levels of genetic variation, with between 28 and 37 haplotypes recovered for each gene within each species (Table 1). Tests of substitution saturation yielded non-significant results for both genes (Supplementary Table S1, doi: 10.5061/dryad.77c1h). Average K2P sequence divergence (Supplementary Table S2) within N. albiventris was 0.05–0.07. Four mtDNA clades (subsequently referred to as Albiventris 1–4; see below for phylogenetic results) within currently recognized N. albiventris were recovered with a pairwise divergence value of 0.05 substitutions/site or greater. The geographic range for these lineages generally corresponds to previously recognized subspecies (Figs. 1 and 2). Clades Albiventris 1, 2, and 4 were restricted to South America, whereas Clade 3 was restricted mainly to Central America with two samples recorded from eastern Venezuela (El Manteco, Bolivar). Two monophyletic mtDNA lineages were identified within N. leporinus (Leporinus 1–2) with an estimated sequence divergence value of 0.024 substitutions/site. Nectilio leporinus genetic lineages correspond to two of the three proposed subspecies ranges but with extensive geographic overlap (Figs. 1 and 2). Leporinus 1 was restricted to Central America, western Andes, and Jamaica (Greater Antilles). Leporinus 2 was restricted
to South America, Lesser Antilles (Dominica, Grenada, Grenadines, and Montserrat), and Panama.

Independent phylogenetic analyses for mtDNA genes recovered minor differences in tree topology, especially for the position of Albiventris 1 clade, and relationships among Albiventris 2, 3, and 4. Through Bayesian and Maximum-Likelihood analyses of Cyt b, Albiventris 1 was placed sister to the rest of Noctilio lineages with low statistical support, whereas Albiventris 2 is sister to a clade containing Albiventris 3 and 4, but only supported in Maximum-Likelihood analyses. Through phylogeny reconstruction using COI, Albiventris 1 also positioned sister to the rest of Noctilio with low statistical support, whereas Albiventris 4 was sister to a clade containing Albiventris 2 and 3, but only supported in Maximum-Likelihood analyses (Supplementary Fig. S1). Bayesian strict clock analyses of Cyt b and COI phylogeny positioned Albiventris 1 as sister to Leporinus 1 and 2 with high support values. Maximum-Likelihood and Bayesian analyses of the concatenated data set provided very similar results (Fig. 2), retaining relatively higher support values at terminal nodes as compared with deeper nodes. These phylogenies had $<50\%$ support for reciprocal monophyly of N. albiventris and N. leporinus as currently recognized (Fig. 2). In contrast, Bayesian strict clock analyses of the concatenated data set yielded strong support for almost all terminal and deeper nodes and provided strong statistical support for a paraphyletic relationship of currently recognized species (Fig. 3).

### X- and Y-variation and Phylogenies

For 33 individuals sequenced for ZFX, 16 variable positions resulted in seven and six haplotypes for N. albiventris and N. leporinus, respectively (Table 1). For 28 individuals sequenced for ZFY, 26 variable positions resulted in six and two haplotypes for N. albiventris and N. leporinus, respectively. The total number of haplotypes recovered for ZFX was greater than that observed for ZFY. ZFY retained higher variability (3.2% vs. 1.9%) and number of polymorphic sites (2.9% vs. 1.4%) as compared with ZFX. Statistical testing indicated neither of these markers were significantly saturated (Supplementary Table S1).

Differences in branching order between ZFX and ZFY were recovered, and both Maximum-Likelihood and Bayesian reconstructions produced the same branching order for each locus. Phylogenetic reconstructions for neither marker fully recovered mtDNA defined lineages into exclusive sex chromosome clades. Although sister relationships between species were not recovered, N. leporinus was supported through both sex-linked phylogenies. Relationships among N. albiventris clades differed between phylogenies and were poorly supported. For both species, a pattern in both phylogenies was sharing of sex loci alleles within species. Figure 2 contains information on geographic localities for samples found to contain sex loci sequences indicating a history of hybridization or incomplete lineage sorting. Similar to the rooted phylogenies, Bayesian strict clock analysis for both sex chromosome loci recovered monophyly of N. leporinus and low statistical support for monophyly of N. albiventris (Fig. 3).

### AFLP-Nuclear Marker Phylogeny and Analysis

Three hundred thirty-nine AFLP loci were scored for 56 samples. An error rate of 4.9% was estimated based on replicate amplification and genotyping, with most discrepancies originating from poor amplification of low quality DNA. Highest variability was observed for Leporinus 2 (31.3%), and the lowest was observed for Albiventris 1 (12.7%). The average number of private bands per clade across the entire data set was found to be 10 (ranging from 1 within Leporinus 1, to 20 within Albiventris 4). AFLP unbiased Nei genetic distance values ranged from 3.3% (Albiventris 2 vs. Albiventris 4) to 20.1% (Albiventris 4 vs. Leporinus 1).

The AFLP unrooted Bayesian phylogeny supported reciprocal monophyly of species, provided overall similar patterns as those recovered through the phylogenetic analyses of other loci, and were most similar at the mtDNA clade level to the sex chromosome phylogenies (Fig. 2). For example, members from
Figure 2. Phylogenetic relationships among Noctilionidae hypothesized from Maximum-Likelihood consensus trees for three different data sets: Concatenated Cytb and COI (top left), ZFY (top right), ZFX (bottom right), and Bayesian topology for unrooted AFLP nuclear loci (bottom left). The distributions of mitochondrial phylogroups across all phylogenies are identified with dotted vertical lines and corresponding symbols. Node values for clades are Maximum-Likelihood bootstrap support followed by Bayesian posterior probabilities. AFLP node support values refer to Neighbor-Joining bootstrap values followed by Bayesian posterior probabilities. Dashes (-) indicate nodes with low statistical support for particular analyses.

Separate hybrid classifications for *N. albiventris* and *N. leporinus* subspecies identified individuals belonging to parental forms or the F2 hybrid category. In all instances, individuals were assigned to these classifications with at least a 0.91 posterior probability, and in 95% of classifications assignments were made with >0.99 posterior probability. For *N. albiventris*, individuals classified as F2 hybrids were among those positioned on intermediate branches in the AFLP phylogeny among different mtDNA clade members, and were samples collected within areas of sympatry following genetic and morphological data (this study; Davis 1973, 1976). Similarly, for *N. leporinus* the F2 hybrid classification included samples indicated in phylogenetic analyses as being of hybrid ancestry, were collected within the Guiana Shield, or were collected...
within the proposed geographic distribution of the other clade/subspecies. Two distantly collected Paraguayan samples were also identified as F2 hybrids. Figure 2 indicates which samples were identified as F2 hybrids for both species.

**Divergence Dating**

Phylogeny, divergence times, and 95% HPD for the radiation of *Noctilio* estimated using BEAST are reported in Figure 4. ESSs for all statistics in both data sets were more than 1000. MtDNA, ZFX, and
FIGURE 4. Bayesian divergence time estimations of Noctilionidae for mtDNA (Cyt b and COI; top), ZFX (middle), and ZFY (bottom). MtDNA chronogram clades with high posterior probabilities are collapsed and represented as solid triangles. Node values are mean tmrca. Stars demark nodes with high posterior probabilities (>0.95). Gray bars across nodes represent 95% HPD intervals for divergence estimates. Mio = Miocene. Outgroups are not shown to conserve space.

ZFY analyses dated the basal radiation of Noctilio to 3.3 Ma (95% HPD: 2.3 and 4.3 Ma), 3.41 Ma (95% HPD: 1.8 and 5.2 Ma), 3.42 Ma (95% HPD: 2.1 and 4.9 Ma), respectively. The mean rate of mtDNA sequence evolution was estimated at 0.014 substitutions/site/my (95% HPD: 0.011–0.018). The clock rate for ZFX and ZFY was estimated at 0.0016 substitutions/site/my (95% HPD: 0.0013–0.0019), and 0.0030 substitutions/site/my (95% HPD: 0.0025–0.0034), respectively. Mitochondrial diversification among subspecies began as early as 2.8 Ma (Albiventris 2, 3, and 4) and as recent as 1.0 Ma (Leporinus 1 and 2). In comparison, ZFX diversification among subspecies started around 2.6 Ma (Leporinus 1, 2, and Albiventris 1) and as recent as 0.34 Ma (Albiventris 1, 2, and 4). However, diversification of paternal subspecific lineages was estimated to have occurred as long ago as 2.7 Ma (Albiventris 1, 2, 3, and 4) and as recent as 0.56 Ma (Albiventris 1 and 4). Independent analyses using either birth–death or Yule species priors resulted in similar tree topology and divergence times (results not shown).

Species Delimitations

The bGMYC analyses identified seven clades containing species-level probabilities >0.95 (Fig. 5). These clades corresponded to major mitochondrial clades identified within the recognized species in addition to a lineage represented by two samples (with identical haplotypes) from Venezuela, which was also identified as a presumed species. Dated to ~1 Ma, the most recent node representing a speciation event was the node separating Albiventris 3 (Central America) from the Venezuelan lineage.

DISCUSSION

Recent evolutionary radiations are difficult to resolve phylogenetically and frequently exhibit signatures interpretable as effects of the lineage sorting process. Implicating lineage sorting is often sufficient to explain patterns derived from single gene investigations (e.g., mitochondrial), which have been the foundation for much of the DNA sequence-based descriptions for parts of the tree-of-life. However, species trees can contain a history of hybridization among lineages. Characterizing these events informs temporal components of the process of speciation. Comparative phylogenetics of the genetic transmission elements indicates timeframes of divergence and genetic exchange via hybridization and can implicate gender-biased processes underlying observed patterns. By whichever combination of mechanisms genetic divergences between populations arise, phylogenetic methods can be used to identify instances of genome capture, ongoing gene flow, and gender-biased processes. Phylogenetic reconstructions of the recent radiation of Noctilio chronicle a dynamic evolutionary history. Signatures of reticulation, mitochondrial genome capture, ongoing hybridization, divergence, and gender-biased dispersal all contribute to the evolutionary
history of extant Noctilio lineages. Such occurrences were documented across species and among morphotypes that conform to monophyletic mtDNA lineages. Gene-to-gene comparisons are important for understanding the extent to which genomic components have sorted or the extent to which genetic exchange still occurs across genetic elements. These comparisons provide resolution at the level of incipient speciation and describe ongoing gene flow among lineages.

Species Divergence and Mitochondrial Capture

Recent systematic studies of the genus have been unable to recover monophyletic relationships of these ecologically and morphologically distinct species (Lewis-Oritt et al. 2001; Pavan et al. 2013). The results of the current study indicate that difficulty in understanding basal Noctilio relationships has been a consequence of lack of suitable outgroups taxa and a likely instance of ancient mtDNA capture between species. The genus Noctilio belongs to a monotypic family separated from the most recent common ancestor, Furipteridae, by ~33 myr (Teeling et al. 2005; Meredith et al. 2011). Increasing evolutionary time separating ingroup from outgroup can lead to the well-documented phenomenon of long-branch attraction in which recurrent substitution drives clade misplacement among ingroup taxa (Bergsten 2005 and references therein). Current data indicated this
phenomenon through discrepancies in mtDNA clade relationships inferred using outgroups versus Bayesian strict clock analysis. The use of outgroups in our analysis often resulted in an unresolved polytomy at the base of *Noctilio*, through uncertainty in placement of the Albiventris 1 clade. Bayesian strict clock analysis placed this clade sister to *N. leporinus* and separate from the remaining *N. albiventris* with high statistical support. Outgroup rooted reconstructions of sex chromosome data sets also yielded poor basal resolution. Bayesian strict clock rooting analyses of sex chromosomes resulted in low resolution for the monophyly of *N. albiventris*, but the placement of Albiventris 1 within *N. albiventris* was strongly supported. AFLP analysis recovered monophyly of *N. albiventris*. Thus, comparison of phylogenies based on all genetic transmission elements provides strong evidence for mitochondrial genome capture as opposed to incomplete lineage sorting. Specifically, the genome capture event as depicted in the current data involved the acquisition of an early diverging *N. leporinus* mtDNA genome lineage by Albiventris 1. Although evidence for ongoing hybridization among subspecies was also documented in this study, this instance of mitochondrial capture is unique in that it represents relatively ancient reticulation in which the mitochondrial genome lineage has been retained in only the capturing species. Given the basal relationship of the captured genome lineage relative to *N. leporinus*, and the timescale of ∼2 Ma between these clades, we hypothesize that this mitochondrial capture is a likely remnant signature of ancient hybridization during early *Noctilio* species formation.

**Subspecific Distributions, Hybridization, and Geographic Zones**

The geographic distributions of morphologically recognized subspecies and those of mtDNA clades were found to be generally geographically coincident. Congruence of independent marker types, such as genetics and morphology, has been proposed as support for subspecies status (Avise and Ball 1990). Congruence of marker types is also an indicator of species status, and distinguishing between subspecific and specific taxonomic rank often relies on prescribed thresholds. In this study stochastic coalescent modeling of the mtDNA data set with a 50% confidence threshold, defined seven presumed conspecific lineages. Perspectives from the other genetic transmission elements provided a powerful assessment of the genetic complexity of these lineages as incipient species (Figs. 2 and 3).

Within currently recognized *N. albiventris*, the recovered phylogenetic patterns are generally congruent with mtDNA divergences, yet also document putative hybridization. Specifically, Albiventris 1 corresponds to the geographic range of morphologically recognized *N. a. affinis* (Amazonia, northern Venezuela, and the coastal Guianas), Albiventris 2 to *N. a. albiventris* (southeastern Venezuela and southern Guyana to eastern Brazil), Albiventris 3 to *N. a. minor* (southern Mexico to northwestern Venezuela: west of the Cordillera Oriental), and Albiventris 4 to *N. a. cabralii* (from southwestern Brazil, Paraguay, and northern Argentina; Fig. 1, Davis 1976; Gardner 2008). Comparing the geographic distributions indicated by morphology and mtDNA clades identified putative areas of sympathy. Specifically, Albiventris 4 samples were obtained within the southern distribution (Bolivia) of Albiventris 1. Although lack of extensive genetic sampling in this area precludes a detailed description of gene flow in this contact zone, one of these samples was identified through AFLP analysis as a hybrid individual. In comparison, eastern Venezuela and Guyana were found to be a contact zone for Albiventris 1, 2, and 3. The occurrence of the single specimen of Albiventris 3 in eastern Venezuela indicated a range extension for this subspecies (this individual was placed within Albiventris 3 clades for all marker types indicating lineage sorting or hybridization did not produce this pattern). Comparing phylogenies for ZFX, ZFY, mtDNA, and AFLP collectively support a complex evolutionary history for *N. albiventris*. Of note is the observation that supported clade relationships varied among sex chromosome and mtDNA phylogenies. The AFLP phylogeny revealed Albiventris 1, 2, and 4 distributed and co-associated among various short branches of the phylogeny. Individuals from these clades were also frequently classified as hybrid individuals. All data combined provided good support for ongoing gene flow between *N. albiventris* subspecies. The pattern of incongruence between the mtDNA and sex chromosome phylogenies indicated a male-biased component to this hybridization. Geographic areas of hybridization were locations in which sympatry among hybridizing lineages is supported (Bolivia and Guiana Shield). In addition, the Guiana Shield is a known South American dispersal corridor, represents a broad environmental ecotone from neighboring regions (e.g., Fouquet et al. 2012), and has been previously proposed as a contact zone for *N. albiventris* subspecies (Davis 1976).

Within *N. leporinus* the clade designated Leporinus 1 corresponded to the proposed geographic range of *N. l. mastius* (Mexico and the southernmost Bahamas south to western Ecuador and Venezuela). However, none of our samples from eastern Venezuela and Lesser Antilles were recovered within the Leporinus 1 clade as proposed for *N. l. mastius*. Rather, 1 sample from Central America (Canal Zone, Panama), and 12 samples from Lesser Antilles (Dominica, Grenada, Grenadines, Montserrat) grouped inside Leporinus 2 mtDNA clade, conflicting with proposed subspecies distributions based on morphology. In addition, the geographic distribution of Leporinus 2 was inclusive of the combined geographic ranges of *N. l. leporinus* and *N. l. rufescens*, and no genetic data type presented in this study was able to discriminate between these proposed subspecies. The geographic area of hybridization occurs within the Guiana Shield, again highlighting the biological importance of this area of South America.
Similar to *N. albiventris*, patterns of incongruency between the mtDNA and sex chromosome phylogenies indicated a male-biased component to hybridization. *X*, *Y*, and AFLP data all place two Leporinus 2 individuals collected in eastern Venezuela within the Leporinus 1 clade. These findings might indicate the presence of Leporinus 1 in Guiana Shield, although not sampled in this study. Evidence of hybridization in Guiana Shield is more robust in the AFLP data, in which the above mentioned individuals occupy an intermediate position in the phylogeny. Another explanation, not mutually exclusive with the first, is that the distribution of Leporinus 2 is progressing northward, hybridizing with and displacing the Leporinus 1 lineage. The observed direction of hybridization and the occurrence of Leporinus 2 in the Lesser Antilles and in Panama are compatible with this hypothesis. Samples in addition to these three individuals also collected in the Guiana Shield were frequently identified as hybrid individuals through AFLP analysis, further suggesting extensive hybridization within this region. The observation that a few samples from distant locations (i.e., Honduras, Paraguay), or those with relatively divergent genotype frequencies, were identified as potential hybrids does not preclude the identification of hybridization in Guiana Shield, but rather presents the possibility of incorrect assignment of individuals with relatively unique genotypes.

**Diversification Timescale and Implications**

Divergence estimates following the fossil and molecular priors imposed in this study place the *tmrca* of extant *Noctilio* to ~3 Ma. This estimate sheds light on several aspects of the noctilionid diversification. The oldest reported fossil for *Noctilio* was recovered from La Venta, Colombia, dated to 12.5 Ma, which was morphologically similar to extant *N. albiventris* (Czaplewski 1996; Czaplewski et al. 2003). Another fossil of a hypothesized extinct species, *N. lacrimaelunaris*, was recovered from the Villavieja Formation and was dated to 5–9 Ma (Czaplewski 1997). The molecular divergence time estimated in this study supports the age of extant *Noctilio* is considerably younger than lineages recovered in the fossil record. Furthermore, the estimated divergence of extant *Noctilio* is placed during and/or after the uplift of the Isthmus of Panama. As the fossils were recovered in the northern Amazonian Basin, these data support a South American origin of *Noctilio* and a subsequent colonization of Central America, if not through colonization of the Caribbean Islands (see below).

(Iturralde-Vinent and MacPhee, 1999) hypothesized that the Caribbean colonization occurred during the existence of an ancient land bridge during the Oligocene. Our data do not support this hypothesis; but rather that of a relatively recent over-water colonization, which has been hypothesized through multiple studies (Baker and Genoways 1978; Koopman 1989; Morgan 2001). Considering the fact that closely related Leporinus 1 haplotypes were observed in Jamaica and Central America, invasion of the Greater Antilles likely occurred through Central America. Similarly for Leporinus 2, minimally diverged haplotypes were observed between South America and the Lesser Antilles, suggesting multiple and recent colonization routes to the Caribbean Islands. Additionally, the observation that haplotype diversity was lower in the Caribbean Islands as compared with mainland populations suggests the mainland as the source of the insular invasion (Larsen et al. 2007), rather than Caribbean populations as the source of Central and South American populations. Interpretation from the current data, which includes Central American sampling, is likely more robust than that of (Pavan et al., 2013), who postulated the Caribbean as the location of evolution of piscivory (and *N. leporinus*). This interpretation was based on the occurrence of haplotypes in the Caribbean which are also found on the mainland, presence of *N. leporinus* fossils in Cuba and Puerto Rico, and the absence of *N. albiventris* across the Antilles.

The estimated divergence time of between 1 and 3 Ma for *N. leporinus* supports the findings of (Lewis-Oritt et al., 2001) who postulated that piscivory in *N. leporinus* represents a recently derived morphological, ecological, and behavioral state for the genus. Based on the size of *Noctilio* fossils, which corresponds to that of *N. albiventris*, the ancestral condition was likely insectivorous (Czaplewski 1996; Czaplewski 1997; Czaplewski et al. 2003). The most parsimonious explanation is that the origin of piscivory arose subsequent to the radiation of extant *Noctilio*. Divergence estimates for the different genetic elements place the *tmrca* of *N. leporinus* to ~1 Ma. Uncertainty about the dietary condition existing on the branch leading to extant *N. leporinus* does not permit an exact time-point for the origin of piscivory. Regardless of the exact timing of this dietary evolution, such a rapid rate of change has not been documented in even the most ecologically diversified family of bats, the Phyllostomidae, which have rapidly radiated and exhibit a variety of dietary strategies (e.g., Darztmann et al. 2010; Rojas et al. 2011; Baker et al. 2012). This evolutionary achievement may have been possible because the life history characteristic of gleaners was already in place, and extensive modifications to morphology (shape) or physiological features (echolocation) were not required to transition into piscivory (Liem 1973).

**Taxonomy**

Because species are interbreeding lineages that should be evolutionarily isolated from all other such lineages, the rank of species often is recognized as the most biologically meaningful designation. Species can be morphologically, ecologically, or genetically unique, or any combination of these biological characteristics. Phylogenetic patterns from the different genetic elements are capable of
clearly identifying divergent lineages that continue to exchange genetic material. Should clades for which mtDNA and morphology are generally congruent be recognized as species although there is clear evidence for hybridization between their nuclear genomes? The observation that well over 20 species concepts exist (Mayden 1997), and that these concepts are themselves classified into conceptual categories (Wilkins 2011), illustrates the difficulty of defining species. Supplementary Table S3 provides a summary of the various characters (morphology, cytotgenetics, mitochondrial, Y-chromosome, X-chromosome, and genotypes), in addition to ecological differences that could be used to make decisions about species identification following many of the proposed species concepts. For currently recognized species of Noctilio, the striking observation is the similarity among subspecies for characteristics most focused on in many species concepts. Subspecies are shown to be genetically and reproductively compatible, exhibit little-to-no genetic divergence at autosomal and sex chromosome loci and no ecological differences have been identified. Table S3 also supports that the subsumed taxon N. a. affinis (Simmons and Voss 1998; Simmons 2005; Gardner 2008) is at least as distinct as other recognized subspecies. The intermediate morphology observed for specimens from the Guiana Shield (Simmons and Voss 1998), an established geographic region of hybridization between morphotypes, is in fact an expected characteristic of subspecies. For these reasons using all available data, although polyphyletic, the species-level taxonomy for Noctilio is at least as distinct as other recognized subspecies. For these reasons using all available data, although polyphyletic, the species-level taxonomy for Noctilio is at least as distinct as other recognized subspecies.

SUPPLEMENTARY MATERIAL
Supplementary material, including data files and/or other supplementary information related to this paper have been deposited at Dryad (http://datadryad.org/) under doi:10.5061/dryad.77c1h. Matrices and trees generated/inferred during phylogenetic analyses can be retrieved at TreeBASE: http://purl.org/phylo/treebase/phylows/study/TB2:S13946.

FUNDING
This work was funded by the Texas Biological Database Program. FAAK was supported through the length of this study by the Universiti Malaysia Sarawak, Higher Education Ministry of Malaysia and Department of Biology, Texas Tech University.

ACKNOWLEDGMENTS
We thank the Natural Science Research Laboratory (NSRL, Texas Tech University), Carnegie Museum of Natural History, Royal Ontario Museum, Museum of Vertebrate Zoology (University of California, Berkeley), National Museum of Natural History (Smithsonian), Louisiana Museum of Natural History (Louisiana State University), American Museum of Natural History, Museum of Southwestern Biology (University of New Mexico), Oklahoma Collection of Genomic Resources (The University of Oklahoma), and Universidad Nacional Mayor de San Marcos. We thank all personnel that were involved in the collection of materials deposited at these museums. We thank Miguel Pinto for requesting tissues and laboratory assistance. Tissue samples and voucher specimen loans were prepared by H. J. Garner and K. MacDonald at the NSRL of the Museum of Texas Tech University. We are grateful to L. Bradley, R. D. Bradley, J. Sullivan, N. Reid, and an anonymous reviewer for valuable comments which greatly improved this manuscript.

REFERENCES
Czaplewski N.J. 1996. Opossums (Didelphidae) and bats (Noctilionidae and Molossidae) from the late Miocene of the Amazon Basin. J. Mammal. 77:84–94.


