OXFORD UNIVERSITY PRESS | Systematic Biology

Speciation with gene flow in North American *Myotis* bats.

Journal:	*Systematic Biology*
Manuscript ID	USYB-2015-275.R2
Manuscript Type:	Regular Manuscript
Date Submitted by the Author:	n/a
Complete List of Authors:	Morales, Ariadna; The Ohio State University, Evolution Ecology and Organismal Biology Jackson, Nathan; University of Tennessee, EEB Dewey, Tanya; University of Michigan, Department of Ecology and Evolutionary Biology O'Meara, Brian; University of Tennessee, EEB Carstens, Bryan; The Ohio State University, Evolution Ecology and Organismal Biology
Keywords:	P2C2M, phylogeographic model selection, Myotis, PHRAPL, speciation with gene flow

Speciation with gene flow in North American *Myotis* bats

Ariadna E. Morales[1*], *Nathan D. Jackson*[2], *Tanya A. Dewey*[3], *Brian C. O'Meara*[2], *Bryan C. Carstens*[1*]

[1]Department of Evolution, Ecology and Organismal Biology, Ohio State University, 318 W. 12th Avenue, Columbus, OH, 43210

[2]Department of Ecology and Evolutionary Biology, University of Tennessee, Knoxville, 442 Hesler Biology Building, Knoxville, TN, 37996

[3]Department of Biology, Colorado State University, 1878 Campus Delivery, Fort Collins, CO, 80523

*Authors for correspondence: carstens.12@osu.edu, moralesgarcia.1@osu.edu.

[Abstract]

Growing evidence supports the idea that species can diverge in the presence of gene flow. However, most methods of phylogeny estimation do not consider this process, despite the fact that ignoring gene flow is known to bias phylogenetic inference. Furthermore, studies that do consider divergence-with-gene-flow typically do so by estimating rates of gene flow using a isolation-with-migration model (IM), rather than evaluating scenarios of gene flow (such as divergence-with-gene flow or secondary contact) that represent very different types of diversification. In this investigation, we aim to infer the recent phylogenetic history of a clade of western long-eared bats while evaluating a number of different models that parameterize gene flow in a variety of ways. We utilize PHRAPL, a new tool for phylogeographic model selection, to compare the fit of a broad set of demographic models that include divergence, migration, or both among *Myotis evotis*, *M. thysanodes* and *M. keenii*. A genomic dataset consisting of 808 loci of ultraconserved elements (UCEs) was used to explore such models in three steps using an incremental design where each successive set was informed by, and thus more focused than, the previous set of models. Specifically, the three steps were to (i) assess whether gene flow should be modeled and identify the best topologies, (ii) infer directionality of migration using the best topologies, and (iii) estimate the timing of gene flow. The best model (AIC model weight ~0.98) included two divergence events ((*Myotis evotis*, *M. thysanodes*), *M. keenii*) accompanied by gene flow at the initial stages of divergence. These results provide a striking example of speciation-with-gene-flow in an evolutionary lineage.

[phylogeographic model selection; *Myotis* bats; PHRAPL; P2C2M; speciation with gene flow]

INTRODUCTION

Species can diverge in spite of gene flow (*e.g.* Jónsson et al. 2014). Growing evidence shows that gene flow does not necessarily stop instantly or completely during speciation (Nosil 2008; Pinho and Hey 2010), but can continue to occur as lineages diverge (e.g., Papadopoulos et al. 2011; Martin et al. 2013). The growing realization that gene flow can occur among diverging lineages implies not only that phylogenetic studies of closely related clades should not ignore this process, but also that they should consider the geographic ranges occupied by the diverging lineages, since the presence of gene flow implies that individuals from these lineages have come into physical contact. Failure to account for interspecific gene exchange when estimating a species tree may result in biased estimates of topology, divergence time, and population size (Eckert and Carstens 2008; Leaché et al. 2014). For example, if some of the shared polymorphism among lineages results from gene flow, divergence times may be underestimated and effective population sizes may be overestimated when this polymorphism is attributed to incomplete lineage sorting (e.g., Leaché et al. 2014). This places empiricists in a difficult position, because the methods that co-estimate population divergence and gene flow (e.g., Hey 2010) require a known topology, but since approaches to species tree inference do not account for gene flow, the estimates from these methods may be inaccurate if alleles that result from gene flow are present in the data. Furthermore, because many evolutionary processes can generate gene tree discordance (such as gene flow, incomplete lineage sorting, hybridization and/or errors in phylogenetic reconstructions; Knowles and Kubatko 2010), it is difficult to identify whether shared polymorphism among lineages is due to either retained ancestral polymorphism or gene flow (Slatkin and Maddison 1989), even when both processes are considered.

Co-estimating gene flow and species trees is a difficult task due to the large number of possible parameters in model space. For a system of three lineages (either populations or species), possible models include the fully resolved trees and no gene flow (three possible topologies), or each of these trees containing up to 8 migration parameters each for present and ancestral gene flow (combination of these parameters results in 877 models per topology). If one is also willing to consider that no divergence has occurred (if there is no evidence of coalescence, i.e., n-island or partial island models), or that divergence occurred simultaneously (i.e. a polytomy), the total number of models that consider all possible topologies, divergence events, and migration parameters would surpass 128,600. As taxa are added, this number increases factorially. Given that no full likelihood method exists for jointly estimating all relevant parameters, here we adopt complementary approaches to, first, evaluate the fit of the species tree model to our data, and secondly, to identify the specific model(s) that offer the best fit. First, we assess the statistical fit of the multispecies coalescent model (MSCM) to our genomic data using P2C2M (Gruenstaeudl et al. 2015). Second, we utilize PHRAPL, a model selection framework developed by Jackson et al. (*in review*) to evaluate the statistical fit of models that include both divergence and gene flow to the empirical dataset. Our explicit consideration of model fit allows us to jointly estimate the phylogenetic relationships and gene flow in a taxonomically complex system, the western long-eared bats of North America.

Myotis *bats*

The genus *Myotis* is among the most speciose mammalian genera, with more than 100 living species distributed worldwide (Simmons 2005). *Myotis* species in the New World are part of a monophyletic group that diverged around 10 - 15 Ma from European and Asian lineages

(Stadelmann et al. 2007). Phylogenetic relationships and species boundaries among many New World *Myotis* species are controversial, mostly due to cryptic morphological variation and poor sampling in phylogenetic studies (Bickham et al. 2004; Larsen et al. 2012; Ruedi et al. 2013). Consequently, some species are difficult to identify in the field due to morphological similarities that are not necessarily indicative of common descent (e.g. Ruedi and Mayer 2001; Dewey 2006). Phylogenetic studies using mitochondrial DNA (mtDNA; cytochrome B) and nuclear (Rag2) loci suggest that *Myotis* bats consist of geographically delineated lineages with unclear morphological differences among some species (Ruedi and Mayer 2001; Stadelmann et al. 2007). Ruedi and Mayer (2001) suggested that *Myotis* bats are prone to evolve into sets of similar ecological and morphological niches wherever they occur.

Despite a substantial effort to infer relationships and species limits among many *Myotis* species, important questions remain regarding the patterns of phylogeographic diversity, species boundaries, as well as the evolutionary processes that promote divergence in this group. Of particular interest is the mode of speciation that has prevailed (i.e., sympatric vs. allopatric) during the evolution of *Myotis*. Even though geographic barriers such as deserts, sea channels, or mountain ranges may reduce dispersal and promote strong genetic structure between some species (e.g. Castella et al. 2000), other species are largely distributed along heterogeneous biomes, sometimes in sympatry. For example, the western long-eared *Myotis* complex is found across most of North America. *Myotis evotis* and *M. thysanodes* have extensive ranges from northwestern US to central Mexico, and may co-occur in part of their distribution (Fig. 1). *Myotis keenii* has a small distribution in coastal areas of the Pacific Northwest, but it may also co-occur with *M. evotis* and *M. thysanodes*, specially in the southern and southeastern part of its distribution (Fig. 1). Inferring species relationships among the western long-eared bats has been

a challenge for evolutionary biologists. For instance, Stadelmann et al. (2007) used Rag1 and cytochrome B and one individual per species to estimate a phylogeny of *Myotis* bats; their results showed that *M. evotis* and *M. keenii* are sister species which form a group that is sister to *M. thysanodes*. Larsen et al. (2012) later confirmed this topology by sampling cytochrome B from Neotropical *Myotis* species (1-3 individuals per species). Dewey (2006) extensively sampled cytochrome B from across the entire distribution of these three taxa in North America and estimated a topology in which *M. keenii* is sister to the group *M. evotis*-*M. thysanodes*. However, she also observed that relationships among *M. evotis*, *M. thysanodes* and *M. keenii* were paraphyletic and inferred that mitochondrial introgression had occurred among species (at least in cytochrome B), despite their being morphologically distinct. Carstens and Dewey (2010) added multilocus data from six anonymous nuclear loci to the dataset analyzed by Dewey (2006), and their results showed that although *M. evotis*, *M. thysanodes* and *M. keenii* form a single monophyletic group, they are paraphyletic with each other at multiple loci, which is likely in part due to rampant gene flow among these lineages. The sister species of this complex is believed to be *M. lucifugus* (e.g., Stadelmann et al. 2007; Carstens and Dewey 2010; Larsen et al. 2012) and it is estimated that most of the diversification within *M. lucifugus*/western long-eared *Myotis* group occurred during the Pleistocene (between 1-1.5 Ma; Carstens and Dewey 2010; Ruedi et al. 2013).

Although the paraphyletic patterns observed in the western long-eared bats complex have been assumed to result from introgression, gene flow was not explicitly considered in any of the aforementioned studies, Thus, in this study we apply a model selection framework to a new genomic dataset in order to infer relationships among *M. evotis*, *M. thysanodes* and *M. keenii* while also explicitly accounting for the presence of gene flow.

METHODS

Sampling

Tissue samples were acquired from all western long-eared bat species from across the ranges of these species, including sympatric zones (i.e. individuals collected at the same locality; Fig. 1; Supplemental Table S1, http://datadryad.org/review?doi=doi:10.5061/dryad.b0q2g). Some of the samples were donated by museum collections (Supplemental Table S1). Our sampling included 80 individuals from *M. evotis* (34), *M. thysanodes* (20), *M. keenii* (12), and *M. lucifugus* (14), with the latter chosen as an outgroup based on results from Carstens and Dewey (2010), who found that *M. lucifugus* is sister to the western long-eared complex. To identify each species, the Mammal Species of the World classification was followed (Simmons 2005), which is typically based on descriptions of body measurements and fur color variation. All specimens included in this study were carefully identified in the field and confirmed by T. Dewey at the University of Michigan Museum of Zoology or by curators at museum collections (see Dewey 2006 and Suppelemntary Table S1).

Capture probe design

A set of 2560 probes for UCE was previously designed to infer phylogenetic relationships among placental mammals (McCormack *et al.* 2012). To increase the applicability of this probe set to *Myotis* species, we aligned probe sequences with the *Myotis lucifugus* genome (Ensembl release 59; Flicek *et al.* 2014) using LASTZ v.1.03.54 (Harris 2007) and retained 1608 sequences that shared at least 92.5% identity across 100 or more base pairs. Duplicate sequences were discarded such that our final UCE probe set contained 1573 sequences of 120 bp each. A

customized sequence capture probe library was purchased from MYcroarray, Inc (Ann Arbor, USA).

Library preparation and sequencing

Genomic DNA was isolated using a DNeasy blood and tissue kit (QIAGEN). To increase the amount of DNA in some of our samples, whole genome amplification was performed using REPLI-g kits (QIAGEN). DNA was fragmented using an Ultrasonic Processor (Fisher Scientific) to size distribution of 400-600 bp. TruSeq libraries were prepared using KAPA library kits (Kapa Biosystems) and custom dual indexes (*i.e.* adapters) with unique combinations were ligated to each sample. Eight samples per pre-enrichment library were pooled at equimolar concentrations. During the enrichment, the protocol provided by MYcroarray was followed, which is based on the workflow described by Gnirke *et al.* (2009). Post-enriched libraries were pooled at equimolar ratios. The final library was sequenced at the Georgia Genomics Facility using a full lane of an Illumina NextSeq 150 bp paired-end Mid Output Flow Cell run.

Contigs assembly, mapping and haplotype reconstruction

Raw reads were demultiplexed using Casava (Illumina, Inc.) and a custom Perl script (available as supplemental material). To assess the quality of the reads and trim the adapters, we followed the Illumiprocessor workflow described by Faircloth (2013). Illumiprocessor uses Trimmomatic v.0.32 (Bolger *et al.* 2014), which trims adapter contamination from reads, and generates *de novo* consensus contigs in each sample using VelvetOptimizer v.2.2.5 (http://www.vicbioinformatics.com/software.velvetoptimiser.shtml last accessed May 11, 2015), which runs as a wrapper script for the Velvet assembler v.1.2.09 (Zerbino and Birney 2008). To

increase the length of our target UCEs loci, we identified the coordinates of each UCEs probe in the *M. lucifugus* genome, and extracted sequences with 500 bp extra on each side. Consensus contigs were then aligned to UCEs target loci using LASTZ (Harris 2007). Any contigs that did not match the target UCEs locus or that matched more than one locus were removed. Retained contigs were mapped against an index of target UCEs loci, using BWA-MEM v. 0.7.8-r455 (Li and Durbin 2009). Individual consensus sequences were generated using SAMTools v.0.1.19 (Li *et al.* 2009). Files were imported from SAM to BAM format, sorted, and indexed. Files in VCF and FASTQ format (with hard-masked low quality bases <Q20) were then generated and seqtk was used to covert output-masked FASTQ files to FASTA (https://github.com/lh3/seqtk last accessed May 11, 2015).

Each locus was aligned using MAFFT v7.123b (Katoh and Standley 2013). To remove sites with more than 50% of missing data, each locus was filtered using mothur v.1.34.4 (Schloss *et al.* 2009). To resolve gametic phase in multiple heterozygous sites, PHASE v.2.1.1 was used (Stephens *et al*. 2001; Stephens and Donnelly, 2003). After phasing, we realigned each locus and removed sites in which gametic phase resolution was not possible (at an 80% accuracy). PGDSpider v.2.0.8.2 (Lischer and Excoffier 2012) was used to convert FASTA files to many different formats (*e.g.* nexus, phylip).

Summary statistics

To assess the level of polymorphism of phased and filtered UCEs, population genetic summary statistics were calculated for each locus and species using DnaSp v.5.10.1 (Librado and Rozas 2009). We analyzed only those loci that were amplified in at least 5 individuals per species across all loci. Overall, 808 UCEs were used (unless otherwise noted) to estimate the

total number of segregating sites (S), number of haplotypes (Hap), haplotype diversity (Hd), nucleotide diversity (π), Watterson's θ, and Tajima's D.

Species tree reconstruction and MSCM fit evaluation

We utilized *BEAST v.1.8.1 (Heled and Drummond 2010; Drummond *et al.* 2012) to estimate the posterior distribution of gene trees and the species tree. When the complete data set of 808 loci was analysed, we were unable to obtain results from *BEAST in a reasonable amount of time (less than 8 weeks). Therefore, we independently analyzed 10 sets of 50 loci chosen at random. This number was chosen based on exploratory analyses using our data as well as on previous studies that have suggested 50 loci as an upper threshold for *BEAST (*e.g.* O'Neill *et al.* 2012). All analyses included at least three individuals per ingroup species and at least two outgroup (*M. lucifugus*) samples. We used PhyML (Guindon & Gascuel 2003) to estimate models of sequence evolution for each locus. However, to reduce the number of parameters during tree estimation we decided to use a HKY model (transition/transversion rate ratio plus distinct frequencies for each nucleotide) for all loci as this model was found to be representative. For all loci, base frequencies were set to empirically estimated values and a strict molecular clock was assumed with a fixed rate of 1.0. The settings used for all *BEAST analyses were as follows: a Yule process on the species tree prior with a random starting tree, improper priors for species.popMean and species.yule.birthRate, 100 million generations sampled every 10000 steps and a burn-in of 10%. MCMC convergence and independent sampling of generations were assessed in Tracer v.1.6 (Rambaut et al. 2014). Maximum Clade Credibility trees were built using Tree Annotator v.1.8.1 (Heled and Drummond 2010).

To evaluate the fit of the MSCM used to estimate gene trees and the species tree in *BEAST, we used posterior predictive simulation (PPS) as implemented in P2C2M v.0.7.6 (Gruenstaeudl et al. 2015). All analyses were performed using the *BEAST input and output files, with 100 replicates. P2C2M calculates several summary statistics to compare posterior and posterior predictive distributions of each gene tree using PPS. These summary statistics include the number of deep coalescences (*ndc;* Maddison 1997), the coalescent likelihood (*coal*: Rannala and Yang 2003; Liu and Yu 2010), and the likelihood of coalescent waiting times (*lcwt*: Rannala and Yang 2003; Reid *et al*. 2014). The discrepancy between posterior and posterior predictive distributions is then measured by calculating the difference between summary statistics from each distribution (in every gene tree and the species tree). If the MSCM has a good fit to the data, the expectation is no difference between these distributions (centered on zero). A substantial deviation from this expectation demonstrates a poor fit of the MSCM to the data.

Model selection

We calculated and compared approximate likelihoods across a wide range of demographic models using PHRAPL (Jackson et al. *in revi*ew; https://github.com/bomeara/phrapl last accessed September 8, 2016). PHRAPL implements a framework to examine a set of models that include divergence, migration, or both without *a priori* constraints on topology or direction of gene flow. Gene tree distributions simulated under a given model are used to approximate the log-likelihood of the data given that model. An information theoretic approach is then used to select the model(s) that best fit(s) the data (Anderson 2008). PHRAPL uses Akaike's Information Criterion (AIC), which can be interpreted as a measure of lack of model fit. To better interpret these relative values, Akaike weights

(wAIC) are used to compare models. These weights are analogous to model probabilities because the sum of all wAIC values in a given set of models is equal to 1. PHRAPL serves as an exploratory tool for phylogenetic and phylogeographic-scale questions (Jackson et al. *in review*), and can compare the fit of a range of isolation-only (IO), migration-only (MO), isolation-with-migration (IM), or mixed models (MX) to the data. Demographic models were designed as follows: IO = models with fully resolved trees and no migration between lineages, MO = n-inland models (no coalescent events) with symmetric or asymmetrical migration between two or more lineages, IM = models with fully-resolved trees and symmetric or asymmetrical migration between lineages, and MX = models which were intermediate to IM and MO models, including one coalescent event and migration between some, but not all species. After an appropriate model or set of models is selected using AIC, population sizes, divergence times, and migration rates can be estimated under this (these) model(s) either using PHRAPL (with a refined grid with more points per parameter values) or another available analytical tool.

Between any two populations, PHRAPL can examine models that 1) exclude migration (0 rates), 2) allow migration in a single direction (1 rate), 3) allow symmetrical migration in both directions (1 rate), or 4) allow asymmetrical migration (2 rates). Because the number of possible models is extremely large when three groups are included in the analyses (more than 128,600 models), we performed model selection in two separate steps. We first rendered the model space more manageable by initially only considering models in which all nonzero migration rates were set to be the same (i.e., only a single free migration parameter was allowed for symmetric migration). The resulting model set, which contained all possible combinations of possible topologies and migration rate assignments, included 81 models, with a mix of IO, IM, MO and MX models. After inferring the best models in this initial model set, we generated a second

model set that allowed for asymmetrical migration between populations (i.e., two free migration parameter were allowed), but only included those topologies observed in the best models from the initial search. This allowed us to explore subtle and complex migration histories by only focusing on the most likely coalescent histories. This second model set included 320 models and contained the topologies of the two best models (under full sampling).

For the PHRAPL analysis, we assigned samples to one of three groups following the species classification of *M. evotis*, *M. thysanodes*, and *M. keenii*. Input gene trees for all 808 loci were estimated in RAxML v.8.1.16 (Stamatakis 2014). Each locus was analyzed with two independent runs, a GTR+GAMMA model, and a rapid-hill climbing algorithm. Then, using PHRAPL we subsampled gene trees, taking four alleles per species (for 200 replicates) because simulation testing indicates that this number represents the optimal trade-off between computational effort and accuracy (Jackson et al. *in review*). The log-likelihood (*ln*L) and AIC of each model was then calculated based on the proportion of matches between simulated and empirical trees, where 100,000 trees were simulated for each model. Simulation of gene trees was conducted using a grid of parameter values for both population divergence (τ= 0.30, 0.58, 1.11, 2.12, 4.07, 7.81 and, 15.00) and migration (m= 0.10, 0.22, 0.46, 1.00, 2.15, 4.64) that was designed to encompass the full range of potential values for western long-eared bats. Models were analyzed in parallel runs and results were then combined such that model averaged parameter values could be calculated for τ and m.

Testing secondary contact vs. speciation with gene flow

To test whether gene flow between lineages (if any) has been constant over the history of divergence or, alternatively, only occurred either recently (secondary contact) or at early stages

of divergence (speciation with gene flow), we explored two additional models in PHRAPL that vary in the mode of migration. In these models, topology and direction of gene flow between species were fixed to match the best model inferred from the previous analyses. First, in a model positing secondary contact, migration was set to only occur at the tips (i.e., between 0.1 * θ generations and the present, where $\theta = 4Ne\mu$). Second, in a model of speciation with gene flow, migration was initiated only after $\tau - \tau/5$ generations had passed. The fit of these two models was then compared to the fit of the equivalent IM (assuming constant gene flow along branches) and IO models that had been analyzed previously.

Sensitivity analyses

This study is the first to analyze hundreds of loci collected from an empirical system using PHRAPL, and thus exploratory analyses were conducted to assess the information content of the data and to understand how much data are required to assess model fit. Since we were interested in exploring how the rank or wAIC for each model changed as a function of the amount of data, we conducted PHRAPL analyses for all models explored above (i.e., 81, 320 and 4 sets), starting with 8 loci and incrementally adding sets of 8 loci until 808 loci were reached (for a total of 101 subsets). We repeated this procedure starting with 101 loci and adding sets of 101 loci until 808 loci were reached (for a total of 8 subsets). Instead of running PHRAPL 8 or 101 times, the output match vector (*i.e.* number of times the observed subsample was found in the simulated distribution of trees for each simulation cycle) was extracted to re-calculate the lnL and AIC from only those loci specified in each subset. wAIC values were plotted per model per subset of loci using *g*plots (Warnes *et al.* 2014).

To assess whether the selection of the best model might be driven by the information content of each locus (i.e. nucleotide diversity, segregating sites, or percentage of missing data), we analyzed one locus at a time using the set of 81 models, and calculated the strength of evidence for the best model against the second-best model (i.e. measured as the ratio of the two models with the highest wAIC; Anderson 2008). Regression analyses were then performed between information content of each locus and the wAIC ratio of the two best models. Box-Cox approximations were used to determine whether transformations of one or two of the variables were needed (Kutner *et al*. 2004).

Parameter estimation

To estimate divergence time, migration rate and population size parameters under the optimal model (i.e., inferred with PHRAPL) a full probabilistic framework was used, as implemented in IMa2 (Hey and Nielsen 2007; Hey 2010). Due to computational constraints, we restricted our IMa2 analyses to the same sets of loci (10 sets of 50) analyzed in *BEAST. The parallel version of IMa2 was used (i.e., IMa2p; Sethuraman and Hey 2015) with the following settings: 120 chains distributed in 24 processors, a linear heating scheme (-ha 0.98), 100000 burn-in steps, and 1000 genealogies sampled per locus every 100 steps. Uniform distributions were set as prior for all parameters, with maximum parameter values according to the largest values estimated by PHRAPL. Divergence time, population size, and migration estimates were then scaled using a nucleotide mutation rate of 2.2×10^{-9} per year (genomic mutation rate calculated for mammals; Kumar and Subramanian 2002) and a generation length of 3-5 years (Carstens and Dewey 2010).

Finally, simulation studies have shown that when a model ignores interspecific gene flow, divergence times tend to be underestimated, resulting in lineages that appear to be younger than they actually are (Leaché et al. 2014). To contrast divergence time estimates between models with and without gene flow, values from the IM model were compared with those calculated from the *BEAST analyses (which ignored gene flow).

RESULTS

Capture probes and summary statistics

The Illumina NextSeq lane produced more than 240 million paired-end reads for the samples analyzed here. After processing, 52 million contigs were assembled, and contigs with coverage greater than 100x that aligned to UCEs were retained. Only those loci that were amplified in at least three individuals per species were aligned and phased. After filtering, 808 loci that contained <1% missing data were analyzed. These loci averaged 422 base pairs (range: 227–743), 39 segregating sites (range: 0–132), and 15 haplotypes (range: 1–61; Table 1; Supplemental Fig. S2). The average value of haplotype diversity was 0.4487 (range: 0.033–0.951), whereas the average value of nucleotide diversity was 0.0066 (range: 0.0001– 0.1034). *M. keenii* showed the highest haplotype and nucleotide diversity values while *M. evotis* showed the lowest values. Mean values of Tajima's D differed between species, but none of them were significantly different.

Species tree reconstruction and MSCM fit evaluation

Three species trees were estimated in the *BEAST analysis across the 10 independent datasets (Supplemental Fig. S3): ((*M. evotis, M. thysanodes), M. keenii)* was inferred in five

analyses, ((*M. keenii, M. thysanodes), M. evotis)* was inferred in two analyses, and ((*M. evotis, M. keenii), M. thysanodes)* was inferred in three analyses. Each species tree was highly supported with posterior probabilities close to 1.0 (Supplemental Table S4), and there was nothing to indicate that any of these results were questionable (e.g. observed ESS > 200 in most cases). However, results from the P2C2M demonstrated that most of these loci have a poor fit to the MSCM. In extreme cases, none of the loci showed a good fit to the MSCM, and in all cases the number of loci with a demonstrably good fit to the data did not exceed 50% (Supplemental Fig. S3).

Model selection

Results from the model selection analyses indicate that gene flow is an important process to include in demographic analyses of this clade. For the smaller model set (81 models), all models with the highest wAIC include gene flow (Figs. 2 and 3; Supplemental Fig. S5). While the wAIC of any single model does not exceed 0.25, the sum of the wAIC of the top 8 models is ~0.9, while the rest of the models have very little support (Fig. 2a). The top two best models served as the basis for a second set of models (320 models total) that incorporate increased complexity of migration history. In contrast to results from the initial model set, analysis of the second model set resulted in a single model that contained the vast majority of the total model probability given all the data. This model (labeled as model 96b; Fig. 4), has a wAIC score of nearly 1.0. Additionally, the second-ranked model has the same topology as the best model and the same parameterization of gene flow between contemporaneous populations, but also includes ancestral migration (model 160b; Fig. 4). Thus, results from PHRAPL indicate that divergence with gene flow best describes the demographic history of western long-eared *Myotis* bats

Testing secondary contact vs. speciation with gene flow

The speciation-with-gene flow model was the best supported by our analyses (Fig. 5). With 50 loci or fewer, it was difficult to distinguish between constant-gene-flow and speciation-with-gene-flow models. However, once more than 100 loci were analyzed, the weight of the speciation-with-gene flow model was high. Neither isolation-only nor secondary-contact models achieved high support.

Sensitivity analyses

Regardless of the analysis (i.e., the 81 or 320 model, or timing of gene flow), our sensitivity analysis demonstrates the importance of sampling multiple loci. In the 81-model analysis, the relative model weights of the models change as a function of the number of analyzed loci (Fig. 3). When all loci were analyzed, the highest wAIC (0.25; Fig. 3a) was attained by an IM model in which all three species both coalesce simultaneously and share migrants (model 22). Model 22 was ranked the highest when subsets of 184 to 808 loci were analyzed (wAIC ranging from 0.158 to 0.246). However, when fewer loci were analyzed (between 8 and 176), a MO model with migration between all species (model 7) achieved the highest wAIC value (wAIC ranging from 0.117 to 0.164). The second-best model varied depending on the number of loci analyzed, but always included migration (Fig. 3b). With the most comprehensive sampling—when anywhere from 496 to all 808 loci were analyzed—an IM model (model 34: (*M. evotis, M. thysanodes*), *M. keenii*), with migration among all species) always garnered second-best support (wAIC ranging from 0.144 to 0.209).

The results are markedly different in the larger (320 model) analysis. An IM model (model 96b: (*M. evotis, M. thysanodes*), *M. keenii*), with a mix of symmetric and asymmetric

migration among species) was always ranked as having the highest probability when at least 40 loci were analyzed (wAIC was ~1.0 once more than 200 loci were analyzed; Fig. 4a).

Results from regression analyses indicated that the wAIC ratio of the two best models and the number of segregating sites are not correlated (F= 1.798, df= 1 and 806, p= 0.1803, r^2= 0.0022; Supplemental Fig. S6), and that the log-wAIC ratio between the two best models and nucleotide diversity per locus are not correlated (F=0.2718, df= 1 and 742, p=0.6023, r^2=0.0003; Supplemental Fig. S6). The log-likelihood of the best model and the squared relative number of missing data are slightly correlated (F=8.018, df= 1 and 806, p=0.004748, r^2=0.0098; Supplemental Fig. S6).

Parameter estimation

Timing of divergence among species was estimated using both IMa2 (gene flow included) and *BEAST (gene flow excluded). Estimated divergence times were smaller when gene flow was excluded from the model (Fig. 6). Assuming a nucleotide mutation rate of 2.2 x 10^{-9} per year per site, the divergence time between *M. evotis* and *M. thysanodes* in the presence of gene flow can be placed about 66,800 generations before present, with the divergence between the common ancestor of *M. evotis-thysanodes* and *M. keenii* at some 418,800 generations before present. In contrast, divergence time estimates made under the species tree model in *BEAST were very recent (1,723 and 3,662 generations ago, respectively).

DISCUSSION

Results from numerous empirical studies suggest that speciation can occur with gene flow (*e.g.* Martin et al. 2013; Papadopoulos et al. 2014; Jónsson et al. 2014). However, in most

cases, these studies base their inferences on migration rates that are estimated under a full IM model (i.e., a resolved tree with a distinct migration rate parameter for each species pair and direction), without assessing (i) whether an IO (i.e., species tree) model is a good fit to their data or (ii) which of several possible IM models is the most appropriate model for the data. In these cases, even when sister species display ongoing migration rates greater than 0, it is unclear whether such results are biologically meaningful or an artifact of poor model fit. An objective assessment of model fit should occur concomitantly with parameter estimation (Koopman and Carstens 2010). Here we have adopted two approaches to test whether a model of divergence with migration best represents the evolutionary history of western long-eared *Myotis* bats. The first approach was to assess the fit of an IO (i.e., species tree) model to the data, using posterior predictive simulation (Reid et al. 2014), as simulation testing suggests that poor fit to the MSCM can be caused by gene flow (Gruenstaeudl et al. 2015). Our second approach was to conduct phylogeographic model selection using PHRAPL to directly compare a broad range of models that include various IO and IM models, as well as MO and MX models. This enabled us to estimate divergence times, population sizes, and migration rates using a model that best fits the data.

Our results based on 808 UCEs clearly show that an IO model is not a good fit to our dataset for western long-eared *Myotis* bats. The MSCM, which assumes that gene tree discordance results only from incomplete lineage sorting, has a demonstrably poor fit to each of 10 replicate data sets consisting of 50 loci each, regardless of the summary statistic used. PPS results thus suggest that genomic data collected from these bat species should not be analysed using a species tree model that ignores gene flow. When we ignored this finding, and estimated species trees in *BEAST using the replicate datasets, all three possible ingroup tree topologies

were inferred in separate analyses. This result is curious given the results of simulation testing (Heled and Drummond 2012), which suggest that 50 loci should be adequate to accurately estimate the species tree. We suspect that our results reflect the differences in complexity between simulated and empirical data. However, with the same amount of data, although PHRAPL was unable to identify a single model with high support, it did rule out IO models since no model without migration achieved even appreciable support. Notably, none of the *BEAST analyses produced any indication that the resulting phylogeny was anything but a robust estimate of the species tree. For example nodal support was high and convergence metrics suggested that the MCMC had effectively searched parameter space. However, despite the topological differences in the species tree, the inferred timing of divergence was shallow in all cases.

In addition to our assessment of model fit of the MSCM, we employed PHRAPL to objectively choose among a large number of possible phylogeographic models. In total, we explored 405 models in three stages to determine what type of model (i.e., IO, MO, IM, and MX) is the most supported given our data. Unlike *BEAST, PHRAPL can incorporate all 808 loci into a single analysis, although this is made possible by the fact that PHRAPL does not conduct a full evaluation of parameter space (to our knowledge, no existing software does). We thus considered possible models in three stages because analyzing all possible models is logistically formidable, as there are >128,600 possible models for three taxa, a number that would increase if we allowed models to have migration rates that varied through time. Rather than devoting several years to an exhaustive search of model space, we simplified this search using an incremental design where each successive set was informed by, and thus more focused than, the previous set of models. Specifically, the three steps were to (i) assess whether gene flow should be modeled and identify the best topologies, (ii) infer directionality of migration using the best topologies, and (iii)

estimate the timing of gene flow. With this approach, as with all model selection studies, there is always a chance that the very best model will be excluded from the set that is being tested. However such a systematic approach ensures that the inferred model, if not the theoretical best, will likely be similar to this optimal model. When the first set of models was explored, all scenarios that achieved the highest probability included migration. When the second set of models was analyzed, a fully resolved tree with symmetric and asymmetric gene flow offered the best fit to the *Myotis* data. Counterintuitively, our sensitivity analysis suggested that fewer loci were needed to identify the optimal model in this larger, but more focused model set. Finally, when the optimal IM model was analyzed under different modes of migration (constant, secondary contact, or speciation-with-gene flow), the speciation-with-gene flow model achieved strong support once more than 200 loci were included in the analysis. These results demonstrate the application of PHRAPL in model selection when an appropriate set of models is designed after exploring and discarding models with poor fit.

Our estimates of the species tree under the MSCM illustrate the influence of inference error that can result when analyses are conducted using a model that is not appropriate for the data. The results from *BEAST indicate that speciation in western long-eared *Myotis* is very recent (between 1723 and 3662 generations ago), and the pattern of diversification is not consistent across analyses. Both of these errors are disturbing, but for different reasons. Gene flow has been demonstrated to cause errors in species tree estimation (Leaché et al. 2014), and the distortion of divergence time estimates towards the tips reported by these authors is similar to what we observe here. Highly supported discordance among species trees inferred using large iterative sets of loci is more troubling. While the average locus length (422 bp) and number of variable sites (39) of our dataset may be modest compared to other phylogenetic studies, each of

our analyzed sets of data is larger (in terms of number of loci) than those that have been suggested by Heled and Drummond (2010) to be adequate to produce a good estimate using *BEAST. Our results thus suggest that gene flow, even when it occurs only among sister taxa, can lead to errors in the inferred topology of the species tree and divergence time estimates.

Myotis *Systematics*

Phylogenetic inference involving groups of closely related species is particularly challenging. In many cases, clades with complex taxonomy are referred to as *complexes* of species, implying that the species boundaries are inadequately known. The species status and phylogenetic relationships among *M. evotis*, *M. thysanodes*, and *M. keenii* have been questioned due to paraphyletic patterns at some loci (Dewey 2006; Carstens and Dewey 2010). Carstens & Dewey (2010) suggest that these are independent lineages (i.e. species) experiencing rampant gene flow, but the species relationships were not resolved in their study nor did the authors explicitly test for migration. Here, we demonstrate not only that gene flow among these species has occurred, but that it was likely most important during the early stages of divergence, and may not occur in the present. After accounting for gene flow, the topology inferred was ((*Myotis evotis*, *M. thysanodes*), *M. keenii*), which matched the one inferred previously by Dewey (2006). As previously mentioned, ignoring interspecific gene flow tends to underestimate divergence times. For example, previous estimates suggested *M. lucifugus* diverged from the western long-eared species during the Pleistocene, implying that divergence among the western long-eared bat species occurred later (Carstens and Dewey 2010). Our divergence time estimates that explicitly consider gene flow are deeper than expected (between 66,782–418,785 generations ago). These

estimates indicate that divergence in the presence of gene flow among *M. evotis*, *M. thysanodes*, and *M. keenii* took place during the mid to late Pliocene (before the Pleistocene).

We are aware that some of these species are experiencing range expansions. For example, *M. keenii* was sampled outside of its described geographic range and recent museum records suggest that this species is expanding its range. Such recent expansion may facilitate genetic exchange with other species (e.g. *M. thysanodes*), but to date we have not detected signals of recent genetic exchange using neutral loci. Furthermore, secondary contact models did not achieve high probabilities, suggesting that any current genetic exchange is slight. This lack of support for secondary contact models also suggests that our species identification was largely accurate given that species misidentification would likely mimic the effects of gene flow, erroneously leading to support for secondary contact models.

In conclusion, our results suggest that both incomplete lineage sorting and gene flow should be modeled in phylogenetic inferences of closely related lineages. We have shown empirically that neglecting gene flow on phylogenetic and phylogeographic inferences leads to wrong conclusions on the mode and tempo of divergence, obscuring important processes such as speciation in the presence of genetic exchange.

SUPPLEMENTAL MATERIAL

Supplementary material, including data files and/or online-only appendices, can be found in the Dryad data repository at http://datadryad.org/review?doi=doi:10.5061/dryad.b0q2g

FUNDING

The National Science Foundation funded this research (DEB-1257784 / DEB 1257669).

The Ohio Supercomputer Center allocated resources to support part of this study (PAS1184). Support for A.M. was provided in part by a graduate fellowship at OSU funded by CONACyT (Reg. 217900 CVU 324588).

ACKNOWLEDGEMENTS

We thank museum curators who lent us tissue samples of specimens under their care: AMNH, American Museum of Natural History, New York, USA; ASNHC, Angelo State Natural History Collection, Texas, USA; CIIDIR-Durango, Colección de Mamíferos, Centro Interdisciplinario de Investigación para el Desarrollo Integral Regional, Durango, México; UABC, Colección de Mamíferos, Universidad Autónoma de Baja California, México; UMMZ, University of Michigan Museum Zoology. We thank Edwin Rice for assistance on DNA extractions and Troy Kieran for advice on library preparations. We thank members of the Carstens and O'Meara labs, and students in the first PHRAPL workshop for conversations related to this work.

REFERENCES

Anderson D.R. 2008. Model based inference in the life sciences: a primer on evidence. Springer, New York.

Barve N., Barve V., Jiménez-Valverde A., Lira-Noriega A., Maher S.P., Peterson A.T., Soberón J., Villalobos F. 2011. The crucial role of the accessible area in ecological niche modeling and species distribution modeling. Ecol. Model. 222:1810-1819.

Berthier P., Excoffier L., Ruedi M. 2006. Recurrent replacement of mtDNA and cryptic hybridization between two sibling bat species *Myotis myotis* and *Myotis blythii*. Proc. R. Soc. B. 273:101–109.

Bickham J.W., Patton J.C., Schlitter D.A., Rautenbach I.L., Honeycutt R.L. 2004. Molecular phylogenetics, karyotypic diversity, and partition of the genus *Myotis* (Chiroptera: Vespertilionidae). Molecular Phylogenetics and Evolution, 33:333–338.

Bolger A. M., Lohse M., Usadel B. 2014. Trimmomatic: A flexible trimmer for Illumina Sequence Data. Bioinformatics. http://dx.doi.org/10.1093/bioinformatics/btu170.

Broennimann O., Fitzpatrick M.C., Pearman P.B., Petitpierre B., Pellissier L., Yoccoz N.G., Thuiller W., Fortin M.-J., Randin C., Zimmermann N.E., Graham C.H., Guisan A. 2012. Measuring ecological niche overlap from occurrence and spatial environmental data. Global Ecology and Biogeography. 21:481-497.

Carstens B.C., Dewey T.A. 2010. Species delimitation using a combined coalescent and information theoretic approach: An example from North American *Myotis* bats. Syst. Biol. 59: 400-414.

Castella V., Ruedi M., Excoffier L., Ibañez C., Arlettaz R., Hausser J. 2000. Is the Gibraltar Strait a barrier to gene flow for the bat *Myotis myotis* (Chiroptera: Vespertilionidae)? Mol. Ecol. 9:1761–1772.

Dewey T.A. 2006 Systematics and phylogeography of North American *Myotis*. PhD dissertation. University of Michigan.

Drummond A.J., Suchard M.A., Xie D., Rambaut A. 2012. Bayesian phylogenetics with BEAUti and the BEAST 1.7. Mol. Biol. Evol. 29:1969-1973.

Eckert A.J., Carstens B.C. 2008. Does gene flow destroy phylogenetic signal? The performance of three methods for estimating species phylogenies in the presence of gene flow. Mol. Phylogenet. Evol. 49:832–842.

Faircloth B.C. 2013. Illumiprocessor: a trimmomatic wrapper for parallel adapter and quality trimming. http://dx.doi.org/10.6079/J9ILL.

Flicek, P. et al. 2014. Ensembl 2014. Nucleic. Acids. Res. 42:D749–D755.

Gnirke A., Melnikov A., Maguire J., Rogov P., LeProust E.M., Brockman W., Fennell T., Giannoukos G., Fisher S., Russ C., et al. 2009. Solution hybrid selection with ultra-long oligonucleotides for massively parallel targeted sequencing. Nat Biotechnol. 27:182–189.

Gruenstaeudl M., Reid N.H. Wheeler G.L., Carstens B.C. 2015. Posterior Predictive Checks of Coalescent Models: P2C2M, an R package. Mol. Ecol. Resour. DOI: 10.1111/1755-0998.12435.

Guindon S., Gascuel O. 2003. A simple, fast, and accurate algorithm to estimate large phylogenies by Maximum Likelihood. Syst. Biol. 52:696-704.

Guisan A., Petitpierre B., Broennimann O., Daehler C., Kueffer C. 2014. Unifying niche shift studies: insights from biological invasions. Trends Ecol. Evol. 29:260-269.

Harris R.S. 2007. Improved pairwise alignment of genomic DNA. Ph.D. Thesis, The Pennsylvania State University.

Heled J., Drummond A.J. 2010. Bayesian inference of species trees from multilocus data. Mol. Biol. Evol. 27:570–80.

Hey J. 2010. Isolation with Migration Models for More Than Two Populations. Mol. Biol. Evol. 27:905-920

Hey J., Nielsen R. 2007. Integration within the Felsenstein equation for improved Markov chain Monte Carlo methods in population genetics. Proc. Natl. Aca. Sci. 104:2785–2790.

Hijmans R.J., Cameron S.E., Parra J.L., Jones P.G., Jarvis A. 2005. Very high resolution interpolated climate surfaces for global land areas. Int. J. Climatol. 25:1965–1978.

Jackson N., Morales A., Carstens B.C., O'Meara B.C. (*in review*) PHRAPL: Phylogeographic Inference Using Approximate Likelihoods. Syst. Biol.

Jónsson H., Schubert M., Seguin-Orlando A., Ginolhac A., Petersen L., Fumagalli M., et al. 2014. Speciation with gene flow in equids despite extensive chromosomal plasticity. Proc. Natl. Acad. Sci. U.S.A. 111:18655–18660.

Katoh K., Standley D.M. 2013. MAFFT multiple sequence alignment software version 7: improvements in performance and usability. Mol. Biol. Evol. 30:772–780.

Knowles L.L., Kubatko L.S. 2010. Estimating species trees: An introduction to concepts and models. In: Knowles L.L., Kubatko L.S., editors. Estimating Species Trees: Practical and Theoretical Aspects. Wiley-Blackwell. p. 1-14.

Koopman M.M., Carstens B.C. 2011. Plant genetic divergence predicts microbial community structure: Molecular phyllogeography of carnivorous pitcher plants. Microb. Ecol. 61:750-758

Kumar S., Subramanian S. 2002. Mutation rates in mammalian genomes. Proc. Natl. Acad. Sci. 99:803–808.

Kutner M.H., Nachtsheim C.J., Neter J. 2004. Applied Linear Regression Models. New York: McGraw-Hill Irwin.

Lack J.B., Roehrs Z.P., Stanley Jr C.E., Ruedi M., Van Den Bussche R.A. 2010. Molecular phylogenetics of *Myotis* indicate familial-level divergence for the genus Cistugo. J. Mammal. 91:976-992.

Larsen R.J., Knapp M.C., Genoways H.H., Khan F.A.A., Larsen P.A., Wilson D.E., Baker R.J. 2012. Genetic Diversity of Neotropical *Myotis* (Chiroptera: Vespertilionidae) with an Emphasis on South American Species. PLoS ONE, e46578.

Leaché A.D., Harris R.B., Rannala B., Yang Z. 2014. The influence of gene flow on species tree estimation: A simulation study. Syst. Biol. 63:17-30.

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

Li H., Handsaker B., Wysoker A., Fennell T., Ruan J., Homer N., Marth G., Abecasis G., Durbin R. and 1000 Genome Project Data Processing Subgroup. 2009. The Sequence alignment/map (SAM) format and SAMtools. Bioinformatics. 25:2078-2079.

Librado P., Rozas J. 2009. DnaSP v5: A software for comprehensive analysis of DNA polymorphism data. Bioinformatics. 25:1451-1452.

Lischer H.E.L., Excoffier L. 2012. PGDSpider: An automated data conversion tool for connecting population genetics and genomics programs. Bioinformatics. 28:298-299.

Liu L., Yu L. 2010. Phybase: an R package for species tree analysis. Bioinformatics. 26:962–963.

Maddison W.P. 1997. Gene trees in species trees. Syst. Biol. 46:523–536.

Martin S. H., Dasmahapatra K. K., Nadeau N. J., Salazar C., Walters J. R., Simpson F., … Jiggins, C. D. 2013. Genome-wide evidence for speciation with gene flow in Heliconius butterflies. Genome Res. 23:1817–1828.

McCormack J.E., Faircloth B.C., Crawford N.G., Gowaty P.A., Brumfield R.T., Glenn T.C. 2012. Ultraconserved Elements Are Novel Phylogenomic Markers that Resolve Placental Mammal Phylogeny when Combined with Species Tree Analysis. Genome Res. 22:746–754.

Nosil P. 2008. Speciation with gene flow could be common. Mol. Ecol. 17: 2103–2106.

O'Neil E. M., Schwartz R., Bullock C.T., Williams J.S., Shaffer H.B., Aguilar-Miguel X., Parra-Olea G., Weisrock D.W. 2013. Parallel tagged amplicon sequencing reveals major lineages and phylogenetic structure in the North American tiger salamander (*Ambystoma tigrinum*) species complex. Mol. Ecol. 22:111-129.

Olson D.M., Dinerstein, E. Wikramanayake, E.D., Burgess N.D., Powell G.V.N., Underwood E.C., D'Amico J.A., Itoua I., Strand H.E., Morrison J.C., Loucks C.J., Allnutt T.F., Ricketts T.H., Kura Y., Lamoreux J.F., Wettengel W.W., Hedao P., Kassem K.R. 2001. Terrestrial Ecoregions of the World: A New Map of Life on Earth: A new global map of terrestrial ecoregions provides an innovative tool for conserving biodiversity. BioScience. 51:933-938.

Papadopulos A.S.T., Baker W.J., Crayn D., Butlin R.K., Kynast R.G., Hutton I., Savolainen V. 2011. Speciation with gene flow on Lord Howe Island. Proc. Natl. Acad. Sci. 108:13188–13193.

Pinho C., Hey J. 2010. Divergence with Gene Flow: Models and Data. Annu. Rev. Ecol. Evol. Syst. 41:215-230.

Rambaut A., Suchard M.A., Xie D., Drummond A.J. 2014. Tracer v1.6. http://beast.bio.ed.ac.uk/Tracer

Rannala B., Yang Z. 2003 Bayes estimation of species divergence times and ancestral population sizes using DNA sequences from multiple loci. Genetics. 164: 1645–1656.

Reid N.M., Brown J.M., Satler J.D., Pelletier T.A., McVay J.D., Hird S.M., Carstens B.C. 2014. Poor fit to the multi-species coalescent model is widely detectable in empirical data. Syst. Biol., 63:322–333.

Ruedi M., Mayer F. 2001. Molecular systematics of bats of the genus *Myotis* (Vespertilionidae) suggests deterministic ecomorphological convergences. Molecular Phylogenetics and Evolution 21:436-448.

Ruedi M., Stadelmann B., Gager Y., Douzery E.J., Francis C.M., Lin L.-K., Guillén-Servent A., Cibois A. 2013. Molecular phylogenetic reconstructions identify East Asia as the cradle for the evolution of the cosmopolitan genus *Myotis* (Mammalia, Chiroptera). Mol. Phylogenet. Evol. 69:437-449.

Schloss P.D., Westcott, S.L. Ryabin T., Hall J.R., Hartmann M., Hollister E.B., Lesniewski R.A., Oakley B.B., Parks D.H., Robinson C.J., Sahl J.W., Stres B., Thallinger G.G., Van Horn D.J., Weber C.F. 2009. Introducing mothur: open-source, platform-independent, community-supported software for describing and comparing microbial communities. Appl. Environ. Microbiol. 75: 7537–7541.

Sethuraman A., Hey J. 2015. IMa2p – parallel MCMC and inference of ancient demography under the Isolation with migration (IM) model. Mol. Ecol. Resour. doi:10.1111/1755-0998.12437

Simmons N.B. 2005. Order Chiroptera: Mammal species of the world: a taxonomic and geographic reference. In: Wilson D.E, Reeder D.M., editors. Mammal species of the world. Johns Hopkins University Press, Baltimore. p 312-529.

Slatkin M., Maddison W.P. 1989. A cladistic measure of gene flow inferred from the phylogenies of alleles. Genetics 123:603–613.

Soberón J. M. 2010. Niche and area of distribution modeling: a population ecology perspective. Ecography, 33:159–167.

Soberón J., Nakamura M. 2009. Niches and distributional areas: concepts, methods, and assumptions. Proc. Natl. Aca. Sci. 106:19644-19650.

Stadelmann B., Lin L.-K., Kunz T.H., Ruedi M. 2007. Molecular phylogeny of New World *Myotis* (Chiroptera, Vespertilionidae) inferred from mitochondrial and nuclear DNA genes. Mol. Phylogenet. Evol. 43:32–48.

Stamatakis A. 2014. RAxML Version 8: A tool for Phylogenetic Analysis and Post-Analysis of Large Phylogenies. Bioinformatics. 30:1312-1313.

Stephens M., Donnelly P. 2003. A comparison of Bayesian methods for haplotype reconstruction from population genotype data. Am. J. Hum. Genet. 73:1162-1169.

Stephens M., Smith N., Donnelly P. 2001. A new statistical method for haplotype reconstruction from population data. Am. J. Hum. Genet. 68:978-989.

Warnes G.R., Ben Bolker, Lodewijk Bonebakker, Robert Gentleman, Wolfgang Huber Andy Liaw, Thomas Lumley, Martin Maechler, Arni Magnusson, Steffen Moeller, Marc

Schwartz, Bill Venables (2014) gplots: Various R programming tools for plotting data, R package version 2.6.0.

695 Zerbino, D.R., Birney E. 2008. Velvet: algorithms éfor de novo short read assembly using de Bruijn graphs. Genome Res. 18:821–829.

FIGURE LEGENDS

Figure 1. Map showing the distribution of the western long-eared bats –*Myotis evotis*, *M. thysanodes*, and *M. keenii*. Circles, squares and triangles represent sampling localities per species. Detailed information for all localities and samples can be found in Supplemental Table S1).

Figure 2. Comparison of wAIC values for the first set of 81 models (a) and the second set of 320 models (b). wAIC values for each model were calculated in subsets of loci gradually increasing. Each column represents a model (listed at the bottom) and each row represents a subset of loci (listed at the right). Values of each subset (row) sum to 1 and color increase from the lowest (light yellow) to highest (red) wAIC value. The plot at the top (a) shows 81 models sorted by degree of migration or isolation: isolation-only (IO) models are on the far left, migration-only (MO) models at the far right and isolation-with-migration models (IM) are in the middle. The plot at the bottom (b) shows 320 models with two fixed topologies and migration scenarios that may be symmetric or asymmetric. Detailed information for subsets of loci can be found in Supplemental Fig. S5).

Figure 3: Models that achieved the highest (a) and the second highest (b) wAIC values when 81 scenarios were explored in PHRAPL. Subsets of 8 loci were gradually added until 808 were reached. At the top of each plot, models are represented. Species names are abbreviated as Mevo (*Myotis evotis*), Mthy (*M. thysanodes*), and Mkee (*M. keenii)*.

Figure 4: Models that achieved the highest (a) and the second highest (b) wAIC values when 320 scenarios were explored in PHRAPL. Subsets of 8 loci were gradually added until 808 were reached. At the top of each plot, models are represented, where symmetric gene flow (double arrow) is marked in red and asymmetric gene flow is marked in green (single arrow). Species names are abbreviated as in Fig. 3.

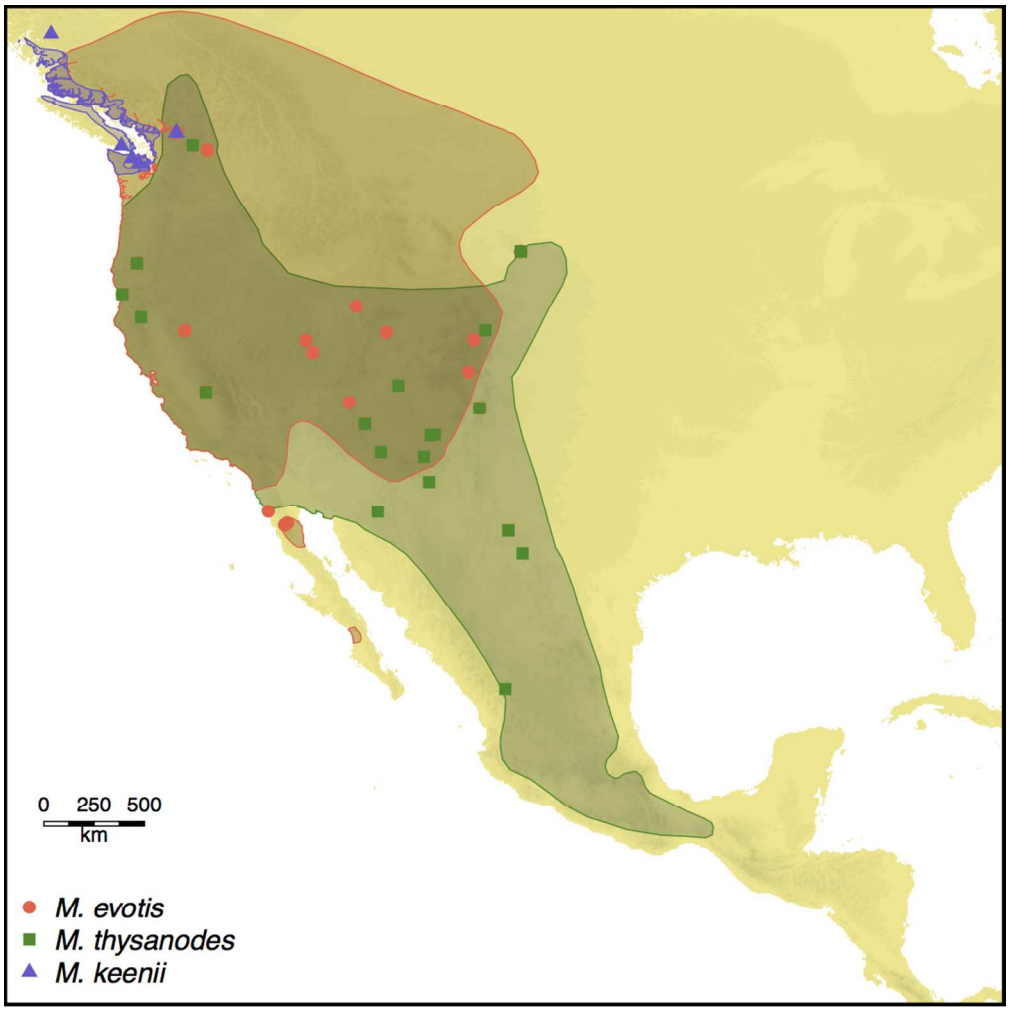
Figure 5: Comparison of wIAC values between isolation-only (IO), isolation-with-constant-migration (IM_con), isolation-with-ancestral-migration (IM_anc or speciation-with-gene-flow) and, isolation-with-recent-migration (IM_rec or secondary contact) models. As in Fig. 2, wAIC values per model were calculated with different number of loci gradually increasing. Each column represents a model (listed at the bottom) and each row represents a subset of loci (listed at the right). Values of each subset (row) sum to 1 and color increases from the lowest (light yellow) to highest (red) wAIC value. Species names are abbreviated as in Fig. 3.

Figure 6: Comparison between IMa2 (left) and *BEAST (right) divergence time estimates using 50 loci chosen at random (dataset 1). Divergence times between nodes in each figure are proportional to values calculated in IMa2 [$\tau1$=0.062 (0.049 - 0.084), $\tau2$=0.853 (0.676 - 1.044)] and *BEAST [$\tau1$=0.0016 (0.0007 – 0.0026), $\tau2$=0.0034 (0.0031 – 0.0044)]. Species names are abbreviated as in Fig. 3.

TABLES

Table 1. Mean values and standard deviation (±SD) of summary statistics for 808 UCEs from *Myotis evotis*, *M. thysanodes* and *M. keenii.*

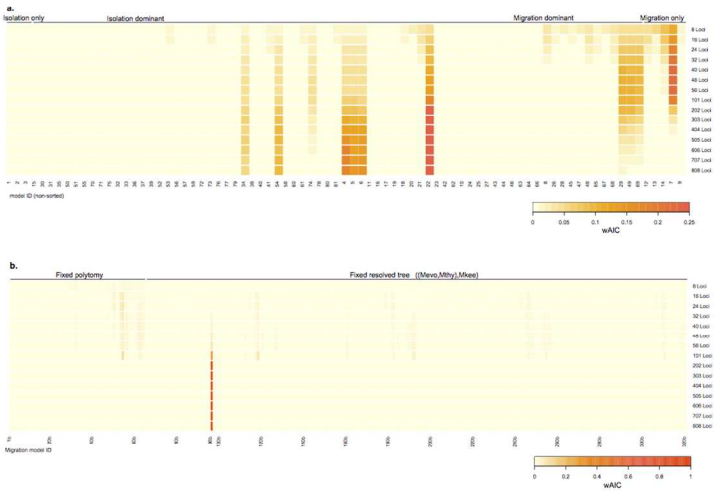
	All		Mevo		Mthy		Mkee	
N	63		34		20		9	
Sequence length [bp]	422	(71)	422	(71)	422	(71)	422	(71)
Segregating sites	39	(24)	25	(18)	16	(13)	15	(14)
Haplotypes	15	(9)	9	(5)	6	(3)	4	(2)
Haplotype diversity	0.4487	(0.2119)	0.4840	(0.2303)	0.5238	(0.2585)	0.6517	(0.3162)
Nucleotide diversity (π)	0.0066	(0.0062)	0.0068	(0.0077)	0.0065	(0.0082)	0.0114	(0.0105)
Theta per site	0.0270	(0.0159)	0.0184	(0.0132)	0.0124	(0.0114)	0.0150	(0.013)
Tajima's D	-2.4430	(0.4104)	-2.1411	(0.4751)	-1.6404	(0.06827)	-0.9513	(0.8999)



Map showing the distribution of the western long-eared bats –Myotis evotis, M. thysanodes, and M. keenii. Circles, squares and triangles represent sampling localities per species. Detailed information for all localities and samples can be found in Supplemental Table S1).

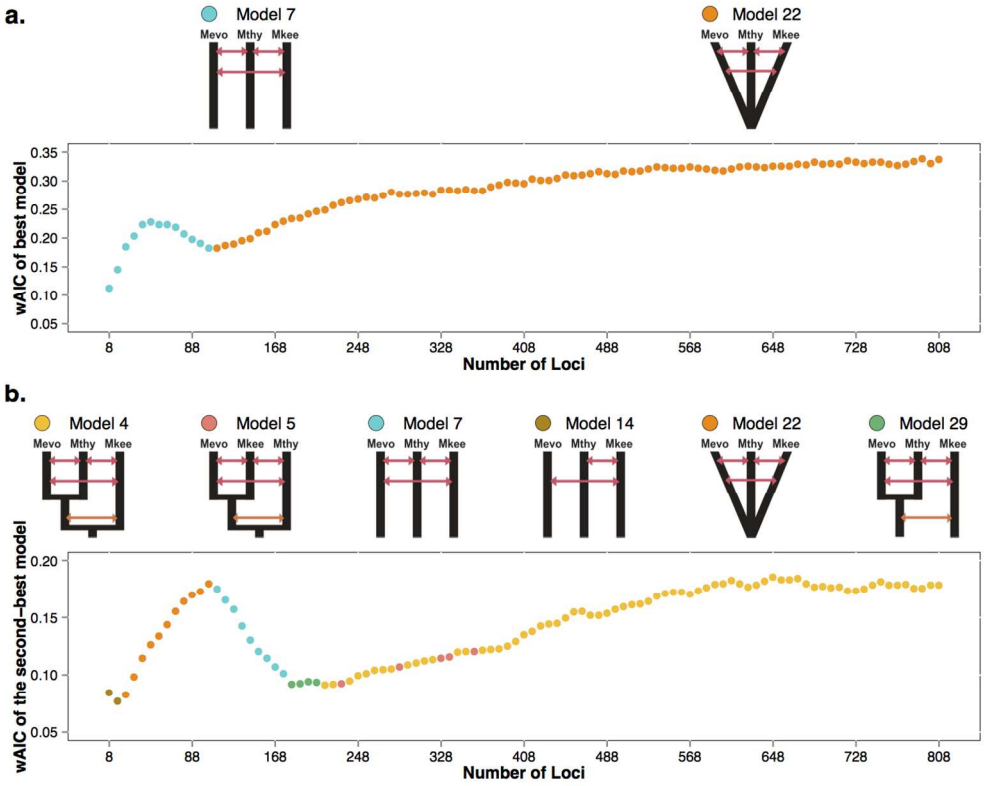
Fig. 1

203x203mm (150 x 150 DPI)



Comparison of wAIC values for the first set of 81 models (a) and the second set of 320 models (b). wAIC values for each model were calculated in subsets of loci gradually increasing. Each column represents a model (listed at the bottom) and each row represents a subset of loci (listed at the right). Values of each subset (row) sum to 1 and color increase from the lowest (light yellow) to highest (red) wAIC value. The plot at the top (a) shows 81 models sorted by degree of migration or isolation: isolation-only (IO) models are on the far left, migration-only (MO) models at the far right and isolation-with-migration models (IM) are in the middle. The plot at the bottom (b) shows 320 models with two fixed topologies and migration scenarios that may be symmetric or asymmetric. Detailed information for subsets of loci can be found in Supplemental Fig. S5).

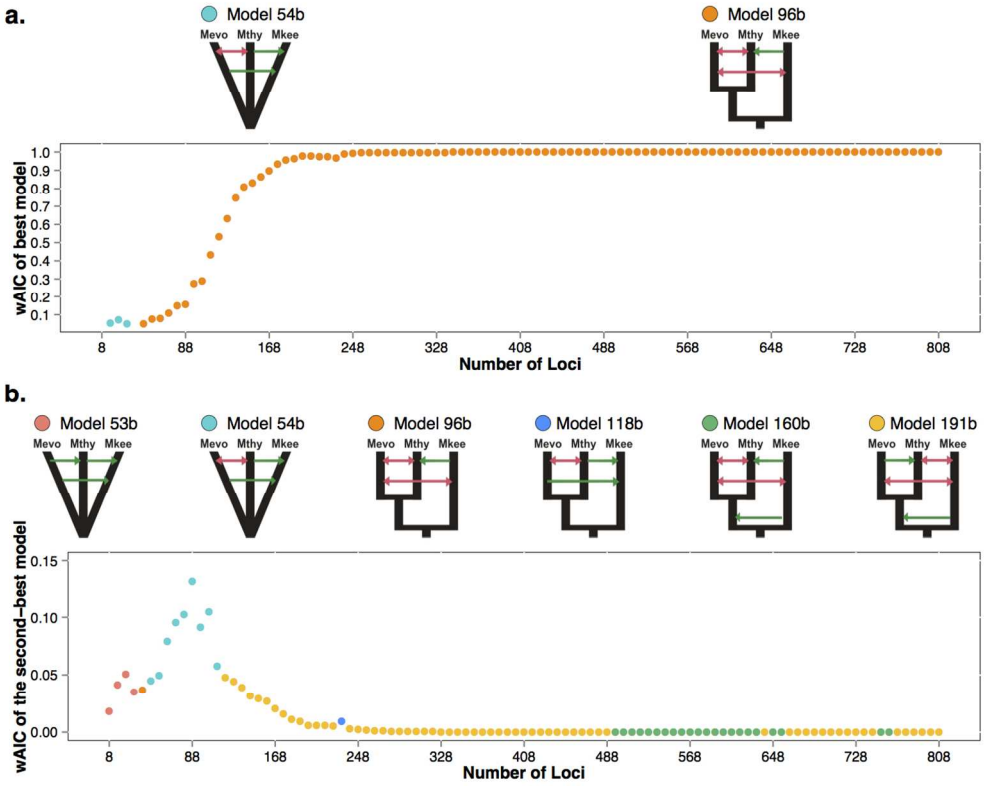
Fig. 2
215x279mm (150 x 150 DPI)



Models that achieved the highest (a) and the second highest (b) wAIC values when 81 scenarios were explored in PHRAPL. Subsets of 8 loci were gradually added until 808 were reached. At the top of each plot, models are represented. Species names are abbreviated as Mevo (Myotis evotis), Mthy (M. thysanodes), and Mkee (M. keenii).

Fig. 3

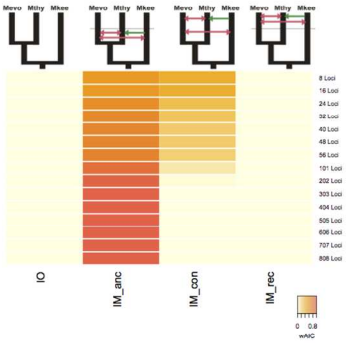
254x203mm (150 x 150 DPI)



Models that achieved the highest (a) and the second highest (b) wAIC values when 320 scenarios were explored in PHRAPL. Subsets of 8 loci were gradually added until 808 were reached. At the top of each plot, models are represented, where symmetric gene flow (double arrow) is marked in red and asymmetric gene flow is marked in green (single arrow). Species names are abbreviated as in Fig. 3.

Fig. 4

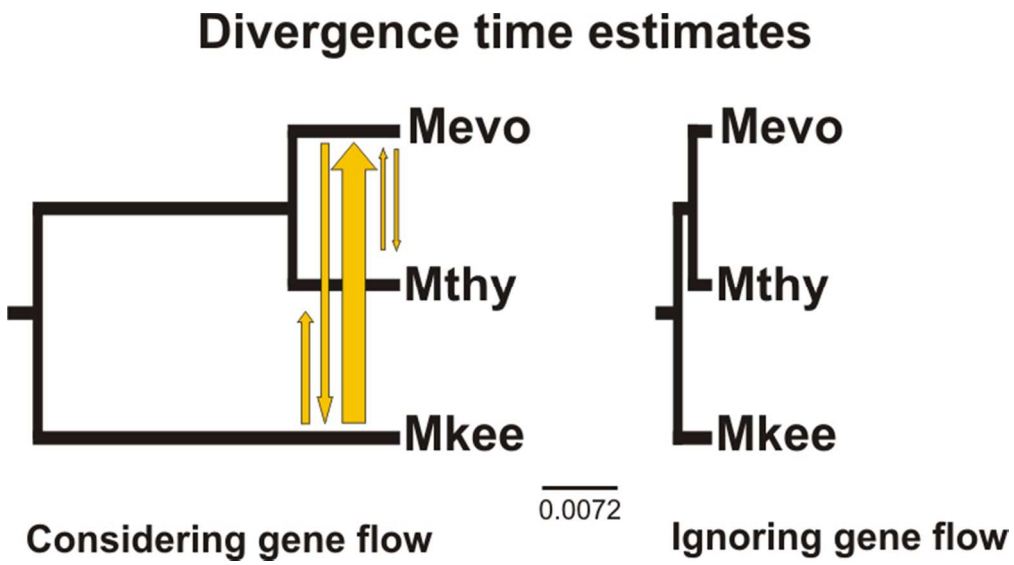
254x203mm (150 x 150 DPI)



Comparison of wIAC values between isolation-only (IO), isolation-with-constant-migration (IM_con), isolation-with-ancestral-migration (IM_anc or speciation-with-gene-flow) and, isolation-with-recent-migration (IM_rec or secondary contact) models. As in Fig. 2, wAIC values per model were calculated with different number of loci gradually increasing. Each column represents a model (listed at the bottom) and each row represents a subset of loci (listed at the right). Values of each subset (row) sum to 1 and color increases from the lowest (light yellow) to highest (red) wAIC value. Species names are abbreviated as in Fig. 3.

Fig. 5

215x279mm (150 x 150 DPI)



Comparison between IMa2 (left) and *BEAST (right) divergence time estimates using 50 loci chosen at random (dataset 1). Divergence times between nodes in each figure are proportional to values calculated in IMa2 [τ1=0.062 (0.049 - 0.084), τ2=0.853 (0.676 - 1.044)] and *BEAST [τ1=0.0016 (0.0007 – 0.0026), τ2=0.0034 (0.0031 – 0.0044)]. Species names are abbreviated as in Fig. 3.

Fig. 6

65x35mm (300 x 300 DPI)