Title: Robust estimates of divergence times and selection with a Poisson random field model: a
case study of comparative phylogeographic data

Authors: Amei Amei*1, Brian Tilston Smith§2

Affiliations:
*Department of Mathematical Sciences, University of Nevada, Las Vegas, Las Vegas, Nevada,
89154, USA.
§ Museum of Natural Science, Louisiana State University, Baton Rouge, Louisiana, 70803,
USA.
2 Current Address: Department of Ornithology, American Museum of Natural History, New
York, NY 10024, USA.
Short running title: Divergence and selection in birds

Key words: speciation, bird, multilocus, Isthmus of Panama, Neotropical region

Corresponding author:
Amei Amei
Department of Mathematical Sciences
University of Nevada, Las Vegas
4505 Maryland Parkway, Box 454020
Las Vegas, Nevada 89154
Phone: 702-895-5159
Email: amei.amei@unlv.edu
Abstract

Mutation frequencies can be modeled as a Poisson random field (PRF) to estimate speciation times and the degree of selection on newly arisen mutations. This approach provides a quantitative theory for comparing intraspecific polymorphism with interspecific divergence in the presence of selection and can be used to estimate population genetic parameters. Although the original PRF model has been extended to more general biological settings to make statistical inference about selection and divergence among model organisms, it has not been incorporated into phylogeographic studies that focus on estimating population genetic parameters for non-model organisms. Here, we modified a recently developed time-dependent PRF model to independently estimate genetic parameters from a nuclear and mitochondrial DNA data set of 22 sister pairs of birds that have diverged across a biogeographic barrier. We found that species that inhabit humid habitat had more recent divergence times and larger effective population sizes than those that inhabit drier habitats, and divergence time estimated from the PRF model were similar to estimates from a coalescent species-tree approach. Selection coefficients were higher in sister pairs that inhabited drier habitats than humid habitats, but overall the mitochondrial DNA was under weak selection. Our study indicates that PRF models are useful for estimating various population genetic parameters and serve as a framework for incorporating estimates of selection into comparative phylogeographic studies.
**Introduction**

A range of analytical approaches have been developed to estimate divergence times and effective population sizes from DNA data (Takahata *et al.* 1995; Yang 1997; Yang 2002; Hickerson *et al.* 2003; Rannala and Yang 2003; Hey and Nielsen 2004; Heled and Drummond 2010). These types of models have been applied to an array of empirical systems and used to reconstruct population history (e.g. Jennings and Edwards 2005; Won and Hey 2005, Lee and Edwards 2008, Hurt *et al.* 2009; Hailer *et al.* 2012), but the extension of these analyses to incorporate selection has been limited. A potential means of including selection estimates into studies examining population or phylogeographic history is the implementation of approaches that compare polymorphic and fixed nucleotide patterns within and between species (Hudson *et al.* 1987). Using this method, McDonald and Kreitman (1991) proposed a statistical test for homogeneity of synonymous and nonsynonymous DNA mutations among closely related species and demonstrated that an excess of nonsynonymous differences among fruit fly species was the result of the fixation of beneficial mutations.

The McDonald-Kreitman test was extended and incorporated into a quantitative model used to compare intraspecific polymorphism with interspecific divergence by modeling mutation frequencies as a Poisson random field (PRF; Sawyer and Hartl 1992). The PRF models the number of synonymous and nonsynonymous fixed differences and polymorphic sites as independent Poisson random variables. The expected values of the Poisson counts are derived by approximating discrete time discrete state Markov chains using diffusion processes. Maximum likelihood and Bayesian inference have been used to estimate synonymous and nonsynonymous mutation rates, the selection coefficient of nonsynonymous mutations, and interspecific
divergence times by comparing theoretical expected values with observed counts in PRF models (Sawyer and Hartl 1992; Bustamante et al. 2002).

The first developed PRF model assumed equal fitness for amino acid changes on the same locus, constant population sizes, random mating, no migration between species, genic selection, independence among nucleotide sites and species are at mutation-selection-drift equilibrium after divergence (Sawyer and Hartl 1992). These assumptions are likely violated in natural populations and several studies have attempted to make the PRF model more biologically realistic by relaxing model assumptions. Extensions of the original model include incorporating migration in an island model of subdivision (Wakeley 2008), a random effects model (Sawyer et al. 2003; Sawyer et al. 2007), relaxation of the genic selection assumption (Williamson et al. 2004), and a time-inhomogeneous model (Williamson et al. 2005; Boyko et al. 2008). The assumption of mutation-selection-drift equilibrium has been integrated into a time-dependent PRF model (Amei and Sawyer 2010) and shown to yield robust divergence time and selection estimates (Amei and Sawyer 2012). The greater biological realism of PRF models may make them attractive analyses for comparative phylogeographic studies that focus on estimating population genetic parameters in non-model organism.

Phylogeographic studies often use multilocus data sets comprised of mitochondrial (mtDNA) and nuclear DNA to estimate population parameters (Brito and Edwards 2009). The fast mutation rate and maternal inheritance of mtDNA makes it a presumably more informative marker than nuclear DNA for identifying population structure in species (Zink and Barrowclough 2008). However, mtDNA and nuclear DNA can exhibit discordant
phylogeographic patterns due to sex-biased dispersal (Peters et al. 2012), introgression (Bryson et al. 2010; Reid et al. 2012), and/or incomplete lineage sorting (Ballard et al. 2004). Further, the utility of mtDNA as a phylogenetic marker has been questioned because of evidence of selective sweeps causing mtDNA genetic diversity to be incongruent with species abundance estimates (Bazin et al. 2006; Nabholz et al. 2008; Nabholz et al. 2009). Since phylogeographic studies often do not take measure selection the usage of PRF models to estimate divergence times, effective population sizes, and selection coefficients provides a means of incorporating the informativeness of mtDNA and at the same time assessing the strength of selection on mtDNA markers.

Comparative phylogeography allows for the testing of hypotheses that have shaped the diversification history of multiple co-distributed taxa. There are two general analytical approaches for estimating and inferring comparative phylogeographic patterns (Hickerson et al. 2010). A model-based approach that compares the likelihood of alternative diversification scenarios and attempts to find the model that best explains a phylogeographic pattern. The other approach relies on estimating phylogeographic parameters from phylogenetic trees and coalescent modeling, and then assesses the level of congruence in parameters among co-distributed species. Both of these methodological frameworks have been used to evaluate the historical assembly of bird faunas from a range of systems including the Neotropics (Burney and Brumfield 2009; Smith et al. 2013), Afrotropics (Voelker et al. 2013), Australia (Huang et al. 2011, Dolman and Joseph 2012), montane forests in Mexico (Barber and Klicka 2010), the Malay Archipelago (Lim et al. 2011), North American deserts (Zink et al. 2001), and the Caribbean (Sly et al. 2011). An overall goal of both analytical approaches is to evaluate whether
patterns inferred from neutral genetic variation are consistent with the predictions of geological changes or climatic fluctuations. While comparative phylogeographic studies serve as a powerful approach for detecting shared evolutionary histories among species, the overall role of selection in the historical assembly of these bird faunas remains unclear.

In this study, we used a recently developed PRF inference method (Amei and Sawyer 2012) and applied it to a published comparative phylogeographic data set. We modified this PRF model by applying a two-step approach that uses both maximum likelihood and a Bayesian framework to independently estimate genetic parameters in 22 sister pairs of birds that have diverged across the Isthmus of Panama, a biogeographic barrier dominated by lowland rainforest that connects North and South America. Based on the mammalian fossil record, it has been proposed that during periods of climate change this rainforest corridor contracted allowing drier habitat taxa to more readily disperse between continents (Webb and Rancy 1996). We tested this hypothesis by using 22 sister pairs of birds that have been delimited into climatic assemblages of either humid rainforest habitat or drier and more open habitat (Smith et al. 2012). If episodes of contracting and expanding rainforest were responsible for speciation then humid and dry habitat taxa should exhibit temporal differences in divergence times. Similar to the pattern in the mammalian fossil record, we would expect more recent divergence times for humid habitat taxa. Previous work using a multispecies coalescent approach inferred that these humid and dry habitat assemblages of taxa may have diverged across the Isthmus of Panama during alternating episodes of contracting and expanding rainforest (Smith et al. 2012). However, taxa likely experience differing selective pressures and demographic histories in dry and humid habitats, and the magnitude of these differences is unclear. Here, we use a PRF model to estimate
divergence times, effective population sizes, and selection coefficients to test whether humid and
dry habitat assemblages have differing evolutionary histories. Understanding the timing of
speciation between assemblages of taxa that inhabit humid and dry habitats may help explain the
processes that generated high diversity in tropical regions.

Methods

Data
We used a published multilocus data set of sister pairs of birds that exhibit genetic breaks across
the Isthmus of Panama (Smith et al. 2012). The original data set comprised of 29 sister pairs, but
some of the pairs exhibited considerable genetic structure that was generated after divergence
across the Isthmus of Panama. This genetic structure violates the random mating assumption of
the PRF model, so we used only 22 of the 29 pairs that do not have deep genetic structure in
populations on each side of the Isthmus of Panama. Based on the results of a climatic assemblage
assignment test, thirteen pairs were assigned to the humid habitat group and nine were assigned
to a drier habitat group, but there was greater climatic variance among and within pairs assigned
to the dry habitat group (Smith et al. 2012). Taxon names and corresponding climatic
assemblages are listed in Table S1, and further details on the taxa and data are available in Smith
et al. (2012). For each species pair, we used sequence alignments of five non-protein coding
markers from the nuclear genome: aconitase 1 (ACO1), beta-fibrinogen intron 5 (βFib5),
myoglobin intron 2 (MY), eukaryotic translation elongation factor 2 (EEF2), and ornithine
decarboxylase (ODC) and one protein coding gene from mitochondrial genome: NADH
dehydrogenase subunit 2 (ND2) to generate McDonald-Kreitman tables for all loci in the
program DnaSP (Librado and Rozas 2009). The number of alleles across the six loci ranged from 1-18.

**Model**

We followed a two-step procedure to estimate species divergence time and infer selection on nonsynonymous mutations. First, assuming that selective effects of segregating mutations of non-coding genes are neutral, we used Markov chain Monte Carlo simulations to estimate divergence time for each of the 22 species pairs by applying a time-dependent PRF model to DNA alignments of the five non-coding genes. Second, given an estimated divergence time of a species pair, we inferred the strength and direction of selection by comparing theoretical values with observed counts of the corresponding M-K table generated from the single coding gene.

In the time-dependent PRF model, the traditional 2x2 M-K table was generalized to a 2x3 contingency table which included two different types of polymorphisms (Amei and Sawyer 2010 and 2012). Adopting the conventional notations, we used $K_s$ to represent the number of synonymous sites that are fixed between a pair of species in a sample of aligned DNA sequences from the two species, $O_s$ to represent the number of synonymous sites that are polymorphic in only one species sample, and $H_s$ the number of synonymous sites that are polymorphic in both samples. For each species pair, we assumed that the two species have the same haploid effective population size $N_e$, they diverged $t_{div}N_e$ generations ago, and the synonymous mutation rate of sites in a non-coding gene in the two lineages is $\theta_s$ per gene per $N_e$ generations. Now, suppose that we have $m$ aligned DNA sequences from one species and $n$ from the other species. Under the assumption that nucleotide sites evolve independently, the counts $K_s$, $O_s$, and $H_s$ are distributed
as independent Poisson random variables with means expressed by functions of the genetic parameters \( t_{div} \) and \( \theta_s \) (explicit expressions are in the Supporting Information). We applied a Bayesian approach by treating the two model parameters \( t_{div} \) and \( \theta_s \) as random variables with prior distributions of uniform and gamma respectively. Given the observed counts \( K_{si}, O_{si}, \) and \( H_{si} \) at each locus \( i, 1 \leq i \leq M, \) the full likelihood function based on the sampling formulas (Eqns (S1)-(S3) in the Supporting Information) and the prior distributions is given by

\[
L(\theta_{si}, t_{div}, K_{si}, O_{si}, H_{si}) = \prod_{i=1}^{M} \{ \text{Poi}_1(K_{si} | t_{div}, \theta_{si}, m_i, n_i) \text{Poi}_2(O_{si} | t_{div}, \theta_{si}, m_i, n_i) \text{Poi}_3(H_{si} | t_{div}, \theta_{si}, m_i, n_i) \Gamma(\theta_{si} | \alpha_s, \beta_s) \} U(t_{div} | 0, T)
\]

(1)

where \( M \) is the number of loci \((M = 5 \) in our sample), \( \Gamma(y | \alpha, \beta) \) and \( U(y|0,T) \) are gamma and uniform densities, and

\[
\text{Poi}_j(c_j | t, \theta, m, n) = \frac{e^{-\lambda_j} (\lambda_j)^{c_j}}{c_j!}, \quad j = 1,2,3; \quad \begin{cases} c_1 = K, \lambda_1 = E(K) \\ c_2 = O, \lambda_2 = E(O) \\ c_3 = H, \lambda_3 = E(H) \end{cases}
\]

Numerical solutions for integrations are obtained by using the Crank-Nicolson method for integrals involving the transition density \( p(t,x,y) \) (defined in the Supporting Information) and Gauss-Legendre quadrature for all other types (Press et al. 2007). All sites in the five non-coding genes are treated as silent sites to calculate observed values of \( K_{si}, O_{si}, \) and \( H_{si}, \) for \( 1 \leq i \leq 5. \)

Setting the hyper parameters \( \alpha_s \) and \( \beta_s \) to a small value \((\sim 0.001)\) and \( T \) a fixed large value, posterior distributions of the model parameters \( t_{div} \) and \( \theta_{si}, 1 \leq i \leq 5, \) are obtained by MCMC simulations with convergence assured by trace-plots and Gelman-Rubin statistics < 1.001 (Gelman 1996). For each species pair, Markov chain Monte Carlo simulations are implemented 120,000 times, the first 20,000 burn-in iterations are disregarded, and model parameters are estimated from 10,000 samples taken every 10 iterating steps. Based on the derivation of the PRF model, \( \theta_{si} \) is an aggregated synonymous mutation rate at locus \( i, \) per \( N_e \)
generations (Amei and Sawyer 2010). We divided this quantity by the total number of synonymous sites at locus $i$ to get the rate of substitution per synonymous site per $N_e$ generations and denote it by $\bar{\theta}_{si}$. Based on our estimates, the synonymous mutation rates of the five loci are similar, thus we used the mean of the five loci, $\bar{\theta}_s = 1/5 \sum_{i=1}^{5} \bar{\theta}_{si}$, as a final estimate of the synonymous mutation rate per synonymous site per $N_e$ generations. A substitution rate of $1.35 \times 10^{-3}$ per site per million years (Myr), based on a published intron rate for birds (Axelsson et al. 2004; Ellegren 2007), was used to express the $N_e$ generations in a unit of Myr and the haploid effective population size $N_e$ ($N_e/2$ for diploid population size) was obtained by assuming a generation time of one year. Consequently, the species divergence time in unit of Myr is given by multiplying the model estimated $t_{dte}$ with $N_e$.

Next, we compared the ND2 M-K table with theoretical expected values to infer selection. Let $K_r, O_r$, and $H_r$ be the amino acid nonsynonymous counterparts of the previously defined $K_s, O_s$, and $H_s$. Under the same model assumptions, the variables are independent Poisson random variables and their expected values are given explicitly in Amei and Sawyer (2012). The mathematical expressions of these Poisson means are similar to those given in Eqns (S1)-(S3) in the Supporting Information but contain the selection coefficient $\gamma$. The magnitude and sign of $\gamma$ describe the strength and direction of the selection intensity of amino acid nonsynonymous mutations. We assumed that polymorphic or fixed mutations have the same selection coefficients within a gene and estimated $\gamma$ by setting the ratio of the expected number of nonsynonymous polymorphic sites to the expected number of nonsynonymous fixed differences to equal the ratio of the numbers of observed nonsynonymous polymorphic sites and fixed differences. That is,

$$\frac{E(O_r) + E(H_r)}{E(K_r)} = \frac{(O_r)_{obs} + (H_r)_{obs}}{(K_r)_{obs}}$$ (2)
Since the mutation parameter $\theta_r$ (nonsynonymous mutation rate) cancels in the ratio of the expected counts, Eqn (2) becomes an equation of $\gamma$ and $t_{div}$. Using the estimated $t_{div}$ from step 1, the numerical solution of the selection coefficient $\gamma$ is obtained by Newton-Raphson method and it can be shown that the obtained value is same as its maximum likelihood estimate (Sawyer and Hartl 1992). In addition to the point estimate of $\gamma$, we applied a 95% credible interval (CI) of $t_{div}$, estimated from step 1, to Eqn (2) to obtain an interval estimate of $\gamma$. Specifically, for each of the 22 species pairs, the lower and upper bounds of the interval estimate of $\gamma$ were calculated by substituting the upper and lower bounds of the 95% CI of the corresponding $t_{div}$.

Results

Effective population size

Haploid effective population size $N_e$ ($N_e/2$ for diploid population size) for the 22 bird species pairs ranged from 440,000 to 3,640,000 (Table S1). Median (middle line of the box) $N_e$ values are 900,000 for dry habitat pairs and 1,550,000 for humid habitat pairs (Fig. 1). Overall, $N_e$ estimates are smaller in dry habitat pairs ($N_e = 440,000 - 1,670,000$) than humid habitat pairs ($N_e = 620,000 - 2,980,000$). One exception was Troglodytes aedon, which has a much larger $N_e$ than the other dry habitat pairs.

Divergence times

We report medians and 95% credible intervals for divergence times estimated from the time-dependent PRF model using the five non-coding nuclear loci. Median $t_{div}$ estimates of the posterior distributions ranged from 0.79 to 4.83 Myr ago for dry habitat pairs and 0.36 to 2.81
Myr ago for humid habitat pairs (Table S1). The side by side boxplot in Fig.1 shows that the interquartile ranges (height of the box) of the $t_{div}$ are 3.12 Myr for the nine dry habitat bird species and 1.14 Myr for the 13 humid species. Although the interquartile ranges of the $t_{div}$ estimates differ by almost 2 million years between the two habitat groups, the median values of the two groups are quite close. Median values were 1.88 Myr ago for dry habitat pairs and 1.13 Myr ago for humid pairs. The median value for all pairs from the PRF model was 1.23 Myr ago (Table S1) in comparison to 1.92 Myr ago from the *BEAST divergence times generated in Smith et al. (2012). For comparison, the *BEAST estimates in Smith et al. (2012) were generated using the following settings: relaxed uncorrelated molecular clocks based on published substitution rates (Axelsson et al. 2004), a Yule process on the species-tree prior, and best-fit models of sequence evolution. The PRF model estimates were typically younger than the *BEAST time estimates, but were not significantly different (two tailed Mann-Whitney test, $u = 0.939$, $P = 0.358$). The 95% credible intervals from the PRF model and *BEAST overlapped for all species except for Xenops minutus (Table S1; Fig.2). Median divergence times estimated from the PRF model between the dry habitat group and the humid group were not statistically significant (two sided Mann-Whitney test, $u=82$, $P=0.125$).

Selection coefficients

For selection coefficients we report values from the protein coding mitochondrial gene ND2. The selection parameter $\gamma$ of nonsynonymous mutation was scaled in terms of the haploid effective population size $N_e$, that is $\gamma = sN_e$, where $s$ is the same selection coefficient per generation. For each of the 22 species pairs, we obtained a point and an interval estimate of the selection coefficient $s$ by applying the ND2 M-K table as well as the estimated values of $t_{div}$ and $N_e$ from
step 1. Selection coefficients varied across the 22 sister pairs (Table 1) and were different between the two climatic assemblages. Among the point estimates of the $s$ of the 22 species pairs, 19 have their $s > 0$ changing from $1 \times 10^{-7}$ to $8.42 \times 10^{-5}$ and the other three have negative selection coefficients of $-5 \times 10^{-7}$, $-6 \times 10^{-7}$, and $-8 \times 10^{-7}$ respectively. For the majority of the species pairs, selection coefficients were positive for nonsynonymous changes, but the strength of selection was weak. There are 11 species pairs whose selection coefficients were less than $1 \times 10^{-5}$, 14 pairs’ selection coefficients were less than $2 \times 10^{-5}$ and 17 had $s < 3 \times 10^{-5}$. Overall, selection coefficients were relatively larger both in magnitude and variation for dry habitat birds as opposed to small and consistent selection for humid birds. A comparison of selection coefficients between habitat groups (Fig. 1) indicated that the estimated $s$ values in dry habitat bird pairs, ranged from $-5 \times 10^{-7}$ to $5.68 \times 10^{-5}$ with an interquartile range of $1.88 \times 10^{-5}$. In contrast, the $s$ values for the 13 humid pairs ranged from $-8 \times 10^{-7}$ to $8.42 \times 10^{-5}$ with a much smaller interquartile range of $0.88 \times 10^{-5}$. Two species pairs had larger $s$ values ($s = 5.68 \times 10^{-5}$, *Campylorhynchus chiapensis* & *C. griseus*, a dry habitat pair; $s = 8.42 \times 10^{-5}$, *Cyanocompsa cyanoides*, a humid habitat pair) than the other species pairs. In order to analyze the effect of biased divergence time estimate on selection coefficient, we generated interval estimate for $s$ (Table 1; Fig. 3). Except for the above mentioned two species pairs, mean lengths of the interval estimates of the selection coefficient are small, ranging from 1.59 for the dry habitat group and 0.93 for the humid group. There are six species pairs whose interval estimates contain zero and the rest are higher than the neutral line (Fig. 3).

**Discussion**
We developed a two-step approach to evaluate divergence times and selection patterns in humid and dry assemblages of birds that are distributed across the Isthmus of Panama. Our study detected temporal differences in divergence times, effective population sizes, and selection coefficients between the taxa that inhabit humid and drier habitats. There was greater variance in all three parameter estimates among the dry habitat pairs that had older divergence times, smaller effective populations, and larger selection coefficients than the humid habitat pairs. The PRF model as a comparative phylogeographic tool provides an analytical framework similar to tree-based methods that yield independent estimates of population genetic parameters that can be used to assess congruence among co-distributed species. Overall, we found that the PRF model is useful for comparative phylogeographic studies interested in studying both the timing of divergence and the role of selection in shaping the evolutionary history of communities.

**Estimating divergence time with a PRF model**

Our estimates of species divergence times indicated that dry habitat birds began diverging as early as 4.8 Myr ago and continued to 0.8 Myr ago. In contrast, speciation of humid habitat birds started much later and went through less variation with an estimated $t_{div}$ of 2.81 Myr ago to a latest time of 0.36 Myr ago. The older divergence times in dry habitat birds corresponds to the pattern observed in the mammalian fossil record that drier habitat taxa were the first to disperse across the Isthmus of Panama (Webb and Rancy 1996), but the median divergence times among groups were not significantly different. This finding is consistent with the conclusions of several comparative phylogeographic studies that recovered multiple episodes of divergence across barriers within a single assemblage of birds (Barber and Klicka 2010; Lim et al. 2011; Sly et al. 2011). Also, the assignment of sister pairs into discrete climatic groups may have limited our
ability to detect significant differences among groups. Climatic preferences across the 22 sister pairs represent a gradient from humid to dry adapted species, and binning the pairs into discrete groups may mask more nuanced diversification patterns.

The divergence times estimated using the PRF model detected qualitatively similar speciation patterns to the results based on a species-tree approach (Smith et al. 2012), although the PRF model based $t_{\text{div}}$ estimates were more recent (Fig. 2). The observed differences between the models are likely attributable to *BEAST having more parameters because it incorporates more complex models of sequence evolution and a relaxed molecular clock (Drummond and Rambaut 2007; Heled and Drummond 2010). The incorporation of variable mutation rates (Yang 1997; Yang 2002; Burgess and Yang 2008) or a relaxed molecular clock (Drummond et al. 2006) in our model would more accurately capture the mutational process, but the similar results between *BEAST and the PRF model along with simulation studies (Amei and Sawyer 2012) suggest that our model provides accurate parameter estimates. Despite the current limitations of the PRF model, our approach yields reconstructions of population history comparable to patterns estimated from other models and recovered similar complex patterns that are often observed in empirical systems (e.g. Lim et al. 2011).

Comparative population sizes

The estimation of effective population size in divergence time models accounts for the impact of ancestral population size on divergence time estimation (Takahata et al. 2005; Beerli and Edwards 2000) and provides the size of extant populations. We found that humid habitat pairs had larger median $N_e$ than dry habitat pairs (Fig. 1). This finding is consistent with the
differences in the distributions of humid and dry tropical biomes in Central and South America. Dry biomes have smaller geographic distributions and are highly fragmented in comparison to humid biomes that are more continuously distributed from Mexico to Bolivia (Prado and Gibbs 1993). The larger geographic area that humid biomes encompass presumably allows birds occurring in these habitats to have larger population sizes. Our inferred estimates of $N_e$ are consistent with the observation that genetic diversity is correlated with range size (Frankham 1996).

**Selection on mtDNA**

We inferred that the ND2 gene showed evidence of positive selection in most of the sister pairs (Fig. 3). The magnitude of the estimated selection coefficients were less than $3 \times 10^{-5}$ for most pairs and only three pairs had negative selection coefficients with absolute values smaller than $8 \times 10^{-7}$. There were two pairs in our dataset, each from one climatic group, estimated to have much larger positive selection coefficients than the other pairs. Based on our results it is not directly apparent why these sister pairs would be under greater selection than the other species. One possible explanation is that nonsynonymous mutations are being fixed at an accelerated rate in these taxa because they have much smaller $N_e$ (Smith and Klicka 2013). Additional analyses would be necessary to investigate the role of $N_e$ because the PRF model assumes equal $N_e$ for both lineages. However, in the case of *C. chiapensis* (from the *C. chiapensis/C. griseus* pair), the species likely has a small $N_e$ because it has a small geographic range along the coast of Mexico. The observed differences in selection coefficients between dry and humid habitat birds require further inquiry. For example, species that inhabit drier habitats may be under greater selection
and larger multilocus data sets may detect more apparent evidence of selection observed in mtDNA, such as adaptive selective sweeps (Ballard and Whitlock 2004; Bazin et al. 2006).

Selection coefficients were sensitive to the uncertainty surrounding divergence time. To evaluate the influence divergence time had on the selection coefficient, we estimated $s$ using different point estimates of $t_{div}$, the lower and upper bound of the 95% CI of $t_{div}$. When the time estimate was larger than the median estimate the selection coefficient decreased, and vice versa, the selection coefficient increased when divergence time was smaller than the median time estimate. There were six sister pairs whose selection coefficient interval estimates contained zero and the rest were above the neutral line. Thus, for the majority of sister pairs the direction of selection remained the same when we used the interval estimate (Table 1; Fig. 3). But, for the pairs that contained zero in their interval estimate we could not distinguish between selection and neutral evolution (Table 1; Fig. 3). The effective population size $N_e$ and the scaled selection coefficient $\gamma$ are estimated independently and do not bias the $\gamma$ value. However, an over- or under-estimated $N_e$ would decrease or increase the non-scaled selection coefficient $s$, due to the relation $s = \gamma/N_e$. Estimated selection coefficients were influenced by the uncertainty surrounding other parameters, but the general patterns remained robust.

The strength of selection on mtDNA has important implications for using mtDNA in phylogeographic studies (Edwards and Bensch 2009; Galtier et al. 2009). Previous work has found that selection on mtDNA makes it a poor indicator of population size (Bazin et al. 2006; Nabholz et al. 2009), species history (Bensch et al. 2006), or is uncorrelated with life history traits (Nabholz et al. 2008), but other work suggest that selection is not strong enough to bias
phylogenetic inference (Zink 2005). Our results that mtDNA was under weak positive selection differ from previous work showing purifying selection reduced mtDNA genetic variability (Bazin et al. 2006; Stewart et al. 2008), but evidence of positive selection on mtDNA is not a unique finding. A comparison of mainland-island birds found that the mtDNA from large mainland populations showed evidence of positive selection (Wright et al. 2008). The variances in the magnitude and type of selection among our study and previous work is not clear, but the discrepancies may be attributed to the use of different mtDNA genes (ND2 vs. cytochrome b), locus length, or calculating selection at differing phylogenetic scales. The overall weak selection we recovered suggests that mtDNA might reflect population history more accurately than previous studies have suggested, but our analyses cannot rule out that even low-levels of selection can reduce mtDNA diversity-levels within populations, thus limiting the utility of the marker.

Conclusions

Phylogeographic studies, particularly those on birds, have focused on estimating parameters from neutrally evolving loci to infer patterns of diversification, while neglecting the role of selection (Hickerson et al. 2010). We demonstrated that the PRF model presented here provides an analytical framework for the incorporation of estimates of selection along with standard population genetic parameters into a comparative phylogeographic study. Our approach provided similar divergence time patterns to those estimated using a multi-species coalescent framework, despite the PRF model assuming a strict molecular clock, a single mutation rate across loci and sites, and the same effective population size for ancestral and daughter populations. These limiting factors can be incorporated into more complex PRF models and are poised to become
increasingly useful as genetic datasets become larger. The development of Next-Generation Sequencing technologies now make it feasible to generate thousands of loci for phylogeographic studies (e.g. Smith et al. 2013) and PRF models can efficiently estimate parameters from large multilocus data sets (Amei and Sawyer 2012). The incorporation of selection into phylogeographic studies will yield a more comprehensive view on genetic patterns across species ranges and the processes responsible for shaping their evolution.

Acknowledgements

We thank J. Klicka for his help generating the multilocus dataset. We also thank D. Fletcher, three anonymous reviewers and N. Rosenberg for providing comments and feedback that greatly improved this manuscript.

Literature Cited


Boyko, A. R., S. H. Williamson, A. R. Indap, J. D. Degenhardt, R. D. Hernandez et al., 2008 Assessing the evolutionary impact of amino acid mutations in the human genome. PLoS Genet 4:
el000083.


Smith, B. T., and J. Klicka, 2013 Examining the role of effective population size on mitochondrial and multilocus divergence time discordance in a songbird. PLoS ONE 8: e55161.


Figure Legends

**Figure 1** Boxplots of PRF model parameter estimates. Haploid effective population size $N_e$ (millions of individuals), species divergence time $t_{div}$ (million years ago), and selection coefficient $s$ ($10^5$) for dry habitat pairs (left) and humid habitat pairs (right). Medians are indicated by thick horizontal bars. Boxes include middle 50% of the distributions.

**Figure 2** Comparison of PRF model and *BEAST estimates of divergence times across the Isthmus of Panama in the 22 bird sister pairs. Error bars represent 95% credible intervals for PRF model and*BEAST estimates.

**Figure 3** Comparison of selection coefficients ($s$) between dry and humid habitat groups with the species pairs sorted by the values of the estimates (medians). Error bars represent interval estimates of the $s$. Two species pairs each from one group are excluded due to their exceptional large values.
Table 1 Point and interval estimates of selection coefficients for each sister pair.

<table>
<thead>
<tr>
<th>Sister Pair</th>
<th>Taxa</th>
<th>$s$</th>
<th>Interval estimate of $s$</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td><em>Campylorhynchus chiapensis</em> &amp; <em>C. griseus</em></td>
<td>5.7</td>
<td>(2.33-11.52)</td>
</tr>
<tr>
<td>2</td>
<td><em>Cardinalis sinuatus</em> &amp; <em>C. phoeniceus</em></td>
<td>0.8</td>
<td>(0.4-1.49)</td>
</tr>
<tr>
<td>3</td>
<td><em>Forpus cyanopygius</em> &amp; <em>F. passerinu</em></td>
<td>2.2</td>
<td>(1.03-3.74)</td>
</tr>
<tr>
<td>4</td>
<td><em>Granatellus venustus</em> &amp; <em>G. pelzelni</em></td>
<td>0.4</td>
<td>(0.23-0.76)</td>
</tr>
<tr>
<td>5</td>
<td><em>Icterus graceannae</em> &amp; <em>I. pectoralis</em></td>
<td>2.2</td>
<td>(0.99-4.09)</td>
</tr>
<tr>
<td>6</td>
<td><em>Pheugopedius sclateri</em> &amp; <em>P. felix</em></td>
<td>2.1</td>
<td>(1.24-3.65)</td>
</tr>
<tr>
<td>7</td>
<td><em>Quiscalus mexicanus</em> &amp; <em>Q. lugubris</em></td>
<td>0.1</td>
<td>(-0.27-1.2)</td>
</tr>
<tr>
<td>8</td>
<td><em>Rhodothraupis celaeno</em> &amp; <em>Periporphyrus erythromelas</em></td>
<td>1.3</td>
<td>(0.69-1.97)</td>
</tr>
<tr>
<td>9</td>
<td><em>Troglodytes aedon</em></td>
<td>-0</td>
<td>(-0.08-0.02)</td>
</tr>
<tr>
<td>10</td>
<td><em>Baryphthengus martii</em></td>
<td>0.2</td>
<td>(0.11-0.38)</td>
</tr>
<tr>
<td>11</td>
<td><em>Cantorichilus nigricapillus</em></td>
<td>0.5</td>
<td>(0.12-1.1)</td>
</tr>
<tr>
<td>12</td>
<td><em>Caryothraustes canadensis</em> &amp; <em>C. poliogaster</em></td>
<td>-0</td>
<td>(-0.16-0.09)</td>
</tr>
<tr>
<td>13</td>
<td><em>Cyanocompsa cyanoides</em></td>
<td>8.4</td>
<td>(3.75-20.63)</td>
</tr>
<tr>
<td>14</td>
<td><em>Hylophilus ochraceiceps</em></td>
<td>-0</td>
<td>(-0.22-0.39)</td>
</tr>
<tr>
<td>15</td>
<td><em>Microbates cinereiventris</em></td>
<td>0.4</td>
<td>(0.16-0.79)</td>
</tr>
<tr>
<td>16</td>
<td><em>Notharchus tectus</em></td>
<td>0.5</td>
<td>(0.17-0.82)</td>
</tr>
<tr>
<td>17</td>
<td><em>Pionus senilis</em> &amp; <em>P. chalcopterus</em></td>
<td>0.8</td>
<td>(-0.08-3.17)</td>
</tr>
<tr>
<td>18</td>
<td><em>Saltator atripennis</em> &amp; <em>S. atriceps</em></td>
<td>1</td>
<td>(0.54-1.6)</td>
</tr>
<tr>
<td>19</td>
<td><em>Saltator maximus</em></td>
<td>1</td>
<td>(0.57-1.71)</td>
</tr>
<tr>
<td>20</td>
<td><em>Trogon rufus</em></td>
<td>0</td>
<td>(-0.05-0.12)</td>
</tr>
<tr>
<td>21</td>
<td><em>Xenops minutus</em></td>
<td>1.2</td>
<td>(0.77-1.72)</td>
</tr>
<tr>
<td>22</td>
<td><em>Xiphorhynchus susurrans</em></td>
<td>0.6</td>
<td>(0.16-1.27)</td>
</tr>
</tbody>
</table>

Sister pairs 1-9 belong to the dry habitat group and 10-22 belong to the humid habitat group.