Robust Estimates of Divergence Times and Selection with a Poisson Random Field Model: A Case Study of Comparative Phylogeographic Data

Amei Amei^{*1} and Brian Tilston Smith^{†2}

^{*}Department of Mathematical Sciences, University of Nevada, Las Vegas, Nevada 89154, and

[†]Museum of Natural Science, Louisiana State University, Baton Rouge, Louisiana 70803

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 nonmodel 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 habitats 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 in those in 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.

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, 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 poly-

morphic 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).

Copyright © 2014 by the Genetics Society of America

doi: 10.1534/genetics.113.157776

Manuscript received September 20, 2013; accepted for publication October 11, 2013; published Early Online October 18, 2013.

Supporting information is available online at <http://www.genetics.org/lookup/suppl/doi:10.1534/genetics.113.157776/-/DC1>.

¹Corresponding author: Department of Mathematical Sciences, University of Nevada, Las Vegas, 4505 Maryland Pkwy., Box 454020, Las Vegas, NV 89154.
E-mail: amei.amei@unlv.edu

²Present Address: Department of Ornithology, American Museum of Natural History, New York, NY 10024.

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 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, 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 for analyses of comparative phylogeographic studies that focus on estimating population genetic parameters in nonmodel organisms.

Phylogeographic studies often use multilocus data sets composed of mitochondrial DNA (mtDNA) and nuclear DNA to estimate population parameters (Brito and Edwards 2009). The fast mutation rate and maternal inheritance of mtDNA make 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 and Whitlock 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, 2009). Since phylogeographic studies often do not 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 codistributed taxa. There are two general analytical approaches for estimating and inferring comparative phylogeographic patterns (Hickerson *et al.* 2010). The first one is a model-based approach that compares the likelihoods 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 codistributed species. Both of these methodological frameworks have been used to evaluate the historical assem-

bly of bird faunas from a range of systems, including the Neotropics (Burney and Brumfield 2009; Smith *et al.* 2013), the 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 was composed 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, 13 pairs were assigned to the humid habitat group and 9 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 [Supporting Information, 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), β -fibrinogen intron 5 (β Fib5), myoglobin intron 2, 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 to 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 noncoding 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 noncoding 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 McDonald–Kreitman (M-K) table generated from the single coding gene.

In the time-dependent PRF model, the traditional 2×2 M-K table was generalized to a 2×3 contingency table that included two different types of polymorphisms (Amei and Sawyer 2010, 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 to be 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 $t_{\text{div}}N_e$ generations ago, and the synonymous mutation rate of sites in a noncoding gene in the two lineages is θ_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 θ_s (explicit expressions are in [File S1](#)). We applied a Bayesian approach by treating the two model parameters t_{div} and θ_s as random variables with prior distributions uniform and gamma, respectively. Given the observed counts K_{s_i} , O_{s_i} , and H_{s_i} at each locus i , $1 \leq i \leq M$, the full-likelihood function based on the sampling formulas (Equations S1–S3 in [File S1](#)) and the prior distributions is given by

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

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

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

Numerical solutions for integrations are obtained by using the Crank–Nicolson method for integrals involving the transition density $p(t, x, y)$ (defined in [File S1](#)) and Gauss–Legendre quadrature for all other types (Press *et al.* 2007). All sites in the five noncoding genes are treated as silent sites to calculate observed values of K_{s_i} , O_{s_i} , and H_{s_i} , for $1 \leq i \leq 5$. Setting the hyperparameters α_s and β_s to a small value (~ 0.001) and T to a fixed large value, posterior distributions of the model parameters t_{div} and θ_{s_i} , $1 \leq i \leq 5$, are obtained by MCMC simulations with convergence ensured 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, θ_{s_i} 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 $\tilde{\theta}_{s_i}$. Based on our estimates, the synonymous mutation rates of the five loci are similar; thus we used the mean of the five loci, $\tilde{\theta}_s = 1/5 \sum_{i=1}^5 \tilde{\theta}_{s_i}$, as a final estimate of the synonymous mutation rate per synonymous site per N_e generations. A substitution rate of 1.35×10^{-3} per site per million years (MY), 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 MY and the haploid effective

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

population size N_e ($N_e/2$ for diploid population size) was obtained by assuming a generation time of 1 year. Consequently, the species divergence time in units of MY is given by multiplying the model-estimated t_{div} by 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 Equations S1–S3 in File S1 but contain the selection coefficient γ . The magnitude and sign of γ 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 γ 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,

$$\frac{E(O_r) + E(H_r)}{E(K_r)} = \frac{(O_r)_{obs} + (H_r)_{obs}}{(K_r)_{obs}}. \quad (2)$$

Since the mutation parameter θ_r (nonsynonymous mutation rate) cancels in the ratio of the expected counts, Equation 2 becomes an equation of γ and t_{div} . Using the estimated t_{div} from step 1, the numerical solution of the selection coefficient γ is obtained by the Newton–Raphson method and it can be shown that the obtained value is the same as its maximum-likelihood estimate (Sawyer and Hartl 1992). In

addition to the point estimate of γ , we applied a 95% credible interval (CI) of t_{div} , estimated from step 1, to Equation 2 to obtain an interval estimate of γ . Specifically, for each of the 22 species pairs, the lower and upper bounds of the interval estimate of γ 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 (Figure 1). Overall, N_e estimates are smaller in dry habitat pairs ($N_e = 440,000 - 1,670,000$) than in 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 noncoding nuclear loci. Median t_{div} estimates of the posterior distributions ranged from 0.79 to 4.83 MY ago for dry habitat pairs and 0.36 to 2.81 MY ago for humid habitat pairs (Table S1). The side-by-side boxplot in Figure 1 shows that the interquartile ranges (height of the box) of the t_{div} are 3.12 MY ago for the 9 dry habitat bird species and 1.14 MY for the 13 humid habitat species. Although the interquartile ranges of the t_{div} estimates differ by almost

2 MY between the two habitat groups, the median values of the two groups are quite close. Median values were 1.88 MY ago for dry habitat pairs and 1.13 MY ago for humid habitat pairs. The median value for all pairs from the PRF model was 1.23 MY ago (Table S1) in comparison to 1.92 MY 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; Figure 2). Median divergence times estimated from the PRF model between the dry habitat group and the humid habitat 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 γ 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 1×10^{-7} to 8.42×10^{-5} and the other 3 have negative selection coefficients of -5×10^{-7} , -6×10^{-7} , and -8×10^{-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 $< 1 \times 10^{-5}$, 14 pairs' selection coefficients were $< 2 \times 10^{-5}$, and 17 had $s < 3 \times 10^{-5}$. Overall, selection coefficients were relatively larger both in magnitude and in variation for dry habitat birds as opposed to small and consistent selection for humid habitat birds. A comparison of selection coefficients between habitat groups (Figure 1) indicated that the estimated s values in dry habitat bird pairs ranged from -5×10^{-7} to 5.68×10^{-5} with an interquartile range of 1.88×10^{-5} . In contrast, the s values for the 13 humid habitat pairs ranged from -8×10^{-7} to 8.42×10^{-5} with a much smaller interquartile range of 0.88×10^{-5} . Two species pairs had larger s values ($s = 5.68 \times 10^{-5}$, *Campylorhynchus chiapensis* and *C. griseus*, a dry habitat pair; and $s = 8.42 \times 10^{-5}$, *Cyanocopsa cyanoides*, a humid habitat pair) than the other species pairs. To analyze the

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

effect of the biased divergence time estimate on the selection coefficient, we generated an interval estimate for s (Table 1; Figure 3). Except for the abovementioned two species pairs, mean lengths of the interval estimates of the selection coefficient are small, ranging from 1.59 for the dry habitat group to 0.93 for the humid habitat group. There are 6 species pairs whose interval estimates contain zero and the rest are higher than the neutral line (Figure 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 codistributed species. Overall, we found that the PRF model is useful for comparative phylogeographic studies interested in both the

Table 1 Point and interval estimates of selection coefficients for each sister pair

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

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

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 MY ago and continued to 0.8 MY ago. In contrast, speciation of humid habitat birds started much later and went through less variation with an estimated t_{div} of 2.81 MY ago to a latest time of 0.36 MY ago. The older divergence times in dry habitat birds correspond 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_{div} estimates were more recent (Figure 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, 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 to those 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 (Edwards and Beerli 2000; Takahata *et al.* 1995) and provides the size of extant populations. We found that humid habitat pairs had larger median N_e than dry habitat pairs (Figure 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).

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 s . Two species pairs each from one group are excluded due to their exceptionally large values.

Selection on mtDNA

We inferred that the ND2 gene showed evidence of positive selection in most of the sister pairs (Figure 3). The magnitude of the estimated selection coefficients was $<3 \times 10^{-5}$ for most pairs and only three pairs had negative selection coefficients with absolute values $<8 \times 10^{-7}$. There were two pairs in our data set, 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 bounds 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; Figure 3). But, for the pairs that contained zero in their interval estimate we could not distinguish between selection and neutral evolution (Table 1; Figure 3). The effective population size N_e and the scaled selection coefficient γ are estimated independently and do not bias the γ -value. However, an over- or underestimated 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) or species history (Bensch *et al.* 2006) or is uncorrelated with life history traits (Nabholz *et al.* 2008), but other work suggests 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.* 2009). The variances in the magnitude and type of selection between our study and previous work is unclear, 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 multispecies 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 data sets become larger. The development of next-generation sequencing technologies now makes 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.

Acknowledgments

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

Literature Cited

- Amei, A., and S. Sawyer, 2010 A time-dependent Poisson random field model for polymorphism within and between two related biological species. *Ann. Appl. Probab.* 20: 1663–1696.
- Amei, A., and S. Sawyer, 2012 Statistical inference of selection and divergence from a time-dependent Poisson random field model. *PLoS ONE* 7: e34413.
- Axelsson, E., N. G. C. Smith, H. Sundstrom, S. Berlin, and H. Ellegren, 2004 Male biased mutation rate and divergence in autosomal, Z-linked and W-linked introns of chicken and turkey. *Mol. Biol. Evol.* 21: 1538–1547.
- Ballard, J. W. O., and M. C. Whitlock, 2004 The incomplete natural history of mitochondria. *Mol. Ecol.* 13: 729–744.
- Barber, B. R., and J. Klicka, 2010 Two pulses of diversification across the Isthmus of Tehuantepec in a montane Mexican bird fauna. *Proc. Biol. Sci.* 277: 2675–2681.
- Bazin, E., S. Glémin, and N. Galtier, 2006 Population size does not influence mitochondrial genetic diversity in animals. *Science* 312: 570–572.
- Bensch, S., D. E. Irwin, J. H. Irwin, L. Kvose, and S. Akesson, 2006 Conflicting patterns of mitochondrial and nuclear DNA diversity in *Phylloscopus* warblers. *Mol. Ecol.* 15: 161–171.
- 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: e1000083.
- Brito, P. H., and S. V. Edwards, 2009 Multilocus phylogeography and phylogenetics using sequence-based markers. *Genetica* 135: 439–455.
- Bryson, R. W. Jr., A. N. M. De Oca, J. R. Jaeger, and B. R. Riddle, 2010 Elucidation of cryptic diversity in a widespread Nearctic treefrog reveals episodes of mitochondrial gene capture as frogs diversified across a dynamic landscape. *Evolution* 64: 2315–2330.
- Burgess, R., and Z. Yang, 2008 Estimation of hominoid ancestral population sizes under Bayesian coalescent models incorporating mutation rate variation and sequencing errors. *Mol. Biol. Evol.* 25: 1979–1994.
- Burney, C. W., and R. T. Brumfield, 2009 Ecology predicts levels of genetic differentiation in Neotropical birds. *Am. Nat.* 174: 358–368.
- Bustamante, C. D., R. Nielsen, S. A. Sawyer, M. D. Purugganan, K. M. Olsen *et al.*, 2002 The cost of inbreeding: fixation of deleterious genes in *Arabidopsis*. *Nature* 416: 531–534.
- Dolman, G., and L. Joseph, 2012 A species assemblage approach to comparative phylogeography of birds in southern Australia. *Ecol. Evol.* 2: 354–369.
- Drummond, A. J., and A. Rambaut, 2007 BEAST: Bayesian evolutionary analysis by sampling trees. *BMC Evol. Biol.* 7: 214.
- Drummond, A. J., S. Y. Ho, M. J. Phillips, and A. Rambaut, 2006 Relaxed phylogenetics and dating with confidence. *PLoS Biol.* 4: e88.
- Edwards, S. V., and P. Beerli, 2000 Gene divergence, population divergence, and the variance in coalescence time in phylogeographic studies. *Evolution* 54: 1839–1854.
- Edwards, S., and S. Bensch, 2009 Looking forwards or looking backwards in avian phylogeography? A comment on Zink and Barrowclough 2008, 2009. *Mol. Ecol.* 18: 2930–2933.
- Ellegren, H., 2007 Molecular evolutionary genomics of birds. *Cytogenet. Genome Res.* 117: 120–130.
- Frankham, R., 1996 Relationships of genetic variation to population size in wildlife. *Conserv. Biol.* 10: 1500–1508.
- Galtier, N., B. Nabholz, S. Glémin, and G. D. D. Hurst, 2009 Mitochondrial DNA as a marker of molecular diversity: a reappraisal. *Mol. Ecol.* 18: 4541–4550.
- Gelman, A., 1996 Inference and monitoring convergence, pp. 131–143 in *Markov Chain Monte Carlo in Practice*, edited by W. Gilks, S. Richardson, and D. Spiegelhalter. Chapman & Hall, London.
- Hailer, F., V. E. Kutschera, B. M. Hallström, D. Klassert, S. R. Fain *et al.*, 2012 Nuclear genomic sequences reveal that polar bears are an old and distinct bear lineage. *Science* 336: 344–347.
- Heled, J., and A. J. Drummond, 2010 Bayesian inference of species trees from multilocus data. *Mol. Biol. Evol.* 27: 570–580.
- Hey, J., and R. Nielsen, 2004 Multilocus methods for estimating population sizes, migration rates and divergence time, with applications to the divergence of *Drosophila pseudoobscura* and *D. persimilis*. *Genetics* 167: 747–760.
- Hickerson, M. J., M. A. Gilchrist, and N. Takebayashi, 2003 Calibrating a molecular clock from phylogeographic data: moments and likelihood estimators. *Evolution* 57: 2216–2225.
- Hickerson, M. J., B. C. Carstens, J. Cavender-Bares, K. A. Crandall, C. H. Graham *et al.*, 2010 Phylogeography's past, present, and future: 10 years after *Avice*, 2000. *Mol. Phylogenet. Evol.* 54: 291–301.
- Huang, W., N. Takebayashi, Y. Qi, and M. J. Hickerson, 2011 MTMLmsBayes: approximate Bayesian comparative phylogeographic inference from multiple taxa and multiple loci with rate heterogeneity. *BMC Bioinformatics* 12: 1.
- Hudson, R. R., M. Kreitman, and M. Aguade, 1987 A test of neutral molecular evolution based on nucleotide data. *Genetics* 116: 153–159.
- Hurt, C., A. Anker, and N. Knowlton, 2009 A multilocus test of simultaneous divergence across the Isthmus of Panama using snapping shrimp in the genus *Alpheus*. *Evolution* 63: 514–530.
- Jennings, W. B., and S. V. Edwards, 2005 Speciation history of Australian grass finches (*Poephila*) inferred from thirty gene trees. *Evolution* 59: 2033–2047.
- Lee, J. Y., and S. V. Edwards, 2008 Divergence across Australia's Carpentarian barrier: statistical phylogeography of the red-backed fairy wren (*Malurus melanocephalus*). *Evolution* 62: 3117–3134.
- Librado, P., and J. Rozas, 2009 DnaSP v5: A software for comprehensive analysis of DNA polymorphism data. *Bioinformatics* 25: 1451–1452.
- Lim, H. C., M. A. Rahman, S. L. Lim, R. G. Moyle, and F. H. Sheldon, 2011 Revisiting Wallace's haunt: coalescent simulations and comparative niche modeling reveal historical mechanisms that promoted avian population divergence in the Malay Archipelago. *Evolution* 65: 321–334.

- McDonald, J. H., and M. Kreitman, 1991 Adaptive protein evolution at the *Adh* locus in *Drosophila*. *Nature* 351: 652–654.
- Nabholz, B., J. F. Mauffrey, E. Bazin, N. Galtier, and S. Glemin, 2008 Determination of mitochondrial genetic diversity in mammals. *Genetics* 178: 351–361.
- Nabholz, B., S. Glemin, and N. Galtier, 2009 The erratic mitochondrial clock: variations of mutation rate, not population size, affect mtDNA diversity across birds and mammals. *BMC Evol. Biol.* 9: 54.
- Peters, J., K. McCracken, C. Pruett, S. Rohwer, S. Drovetski *et al.*, 2012 A parapatric propensity for breeding precludes the completion of speciation in common teal (*Anas crecca*, sensu lato). *Molecular Ecology* 21:4563–4577.
- Prado, D. E., and P. E. Gibbs, 1993 Patterns of species distributions in the dry seasonal forests of South America. *Ann. Mo. Bot. Gard.* 80: 902–927.
- Press, W. H., S. A. Teukolsky, W. T. Vetterling, and B. P. Flannery, 2007 *Numerical Recipes: The Art of Scientific Computing*. Cambridge University Press, Cambridge, UK.
- Rannala, B., and Z. Yang, 2003 Bayes estimation of species divergence times and ancestral population sizes using DNA sequences from multiple loci. *Genetics* 164: 1645–1656.
- Reid, N., J. R. Demboski, and J. Sullivan, 2012 Phylogeny estimation of the radiation of western North American chipmunks (*Tamias*) in the face of introgression using reproductive protein genes. *Syst. Biol.* 61: 44–62.
- Sawyer, S., and D. L. Hartl, 1992 Population genetics of polymorphism and divergence. *Genetics* 132: 1161–1176.
- Sawyer, S. A., R. J. Kulathinal, C. D. Bustamante, and D. L. Hartl, 2003 Bayesian analysis suggests that most amino acid replacements in *Drosophila* are driven by positive selection. *J. Mol. Evol.* 57: S154–S164.
- Sawyer, S. A., J. Parsch, Z. Zhang, and D. L. Hartl, 2007 Prevalence of positive selection among nearly neutral amino acid replacements in *Drosophila*. *Proc. Natl. Acad. Sci. USA* 104: 6504–6510.
- Sly, N. D., A. K. Townsend, C. C. Rimmer, J. M. Townsend, S. C. Latta *et al.*, 2011 Ancient islands and modern invasions: disparate phylogeographic histories among Hispaniola's endemic birds. *Mol. Ecol.* 20: 5012–5024.
- 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.
- Smith, B. T., A. Amei, and J. Klicka, 2012 Evaluating the role of contracting and expanding rainforest in initiating cycles of speciation across the Isthmus of Panama. *Proc. Biol. Sci.* 279: 3520–3526.
- Smith, B. T., M. G. Harvey, B. C. Faircloth, T. C. Glenn, and R. T. Brumfield, 2013 Target capture and massively parallel sequencing of ultraconserved elements (UCEs) for comparative studies at shallow evolutionary time scales. *Syst. Biol.* DOI: 10.1093/sysbio/syt061.
- Stewart, J. B., C. Freyer, J. L. Elson, and N. G. Larsson, 2008 Purifying selection of mtDNA and its implications for understanding evolution and mitochondrial disease. *Nat. Rev. Genet.* 9: 657–662.
- Takahata, N., Y. Satta, and J. Klein, 1995 Divergence time and population size in the lineage leading to modern humans. *Theor. Popul. Biol.* 48: 198–221.
- Voelker, G., B. D. Marks, C. Kahindo, U. A'genonga, F. Bapeamoni *et al.*, 2013 River barriers and cryptic biodiversity in an evolutionary museum. *Ecol. Evol.* 3: 536–545.
- Wakeley, J., 2008 Polymorphism and divergence for island-model species. *Genetics* 163: 411–420.
- Webb, S. D., and A. Rancey, 1996 Late Cenozoic evolution of the Neotropical mammal fauna, pp. 335–358 in *Evolution and Environment in Tropical America*, edited by J. B. C. Jackson, A. F. Budd, and A. G. Coates. University of Chicago Press, Chicago.
- Williamson, S. H., A. F. Alon, and C. D. Bustamante, 2004 Population genetics of polymorphism and divergence for diploid selection models with arbitrary dominance. *Genetics* 168: 463–475.
- Williamson, S. H., R. Hernandez, A. F. Alon, L. Zhu, R. Nielsen *et al.*, 2005 Simultaneous inference of selection and population growth from patterns of variation in the human genome. *Proc. Natl. Acad. Sci. USA* 102: 7882–7887.
- Won, Y. J., and J. Hey, 2005 Divergence population genetics of chimpanzees. *Mol. Biol. Evol.* 22: 297–307.
- Wright, S. D., L. N. Gillman, H. A. Ross, and D. J. Keeling, 2009 Slower tempo of microevolution in island birds: implications for conservation biology. *Evolution* 63: 2275–2287.
- Yang, Z., 1997 On the estimation of ancestral population sizes of modern human. *Genet. Res.* 69: 111–116.
- Yang, Z., 2002 Likelihood and Bayesian estimation of ancestral population sizes in hominoids using data from multiple loci. *Genetics* 162: 1811–1823.
- Zink, R. M., 2005 Natural selection on mitochondrial DNA in *Parus* and its relevance for phylogeographic inference. *Proc. Biol. Sci.* 272: 71–78.
- Zink, R. M., and G. F. Barrowclough, 2008 Mitochondrial DNA under siege in avian phylogeography. *Mol. Ecol.* 17: 2107–2121.
- Zink, R. M., A. E. Kessen, T. V. Line, and R. C. Blackwell-Rago, 2001 Comparative phylogeography of some aridland bird species. *Condor* 103: 1–10.

Communicating editor: N. A. Rosenberg

GENETICS

Supporting Information

<http://www.genetics.org/lookup/suppl/doi:10.1534/genetics.113.157776/-/DC1>

Robust Estimates of Divergence Times and Selection with a Poisson Random Field Model: A Case Study of Comparative Phylogeographic Data

Amei Amei and Brian Tilston Smith

The explicit expressions of expected values of the three Poisson counts K_s , O_s , and H_s are

$$E(K_s) = \theta_s \left[\int_0^1 (I(x, m)K(x, n) + I(x, n)K(x, m))(1 - x)m(dx) \right. \\ \left. + 2 \left(t - \int_0^t \int_0^1 \left(\lim_{x \rightarrow 0} \frac{p(u, x, y)}{x} \right) ym(dy)du \right) + L(m) + L(n) \right] \quad (S1)$$

$$E(O_s) = \theta_s \int_0^1 (2 - x^m - (1 - x)^m - x^n - (1 - x)^n - 2J(x, m)J(x, n))(1 - x)m(dx) \quad (S2)$$

$$E(H_s) = \theta_s \int_0^1 J(x, m)J(x, n)(1 - x)m(dx) \quad (S3)$$

where

$$I(x, m) = 1 - x - \int_0^1 p(t, x, y)(1 - y - (1 - y)^m)m(dy)$$

$$J(x, m) = \int_0^1 p(t, x, y)(1 - y^m - (1 - y)^m)m(dy), \quad J(x, 1) = 0$$

$$K(x, m) = x + \int_0^1 p(t, x, y)(y^m - y)m(dy)$$

$$L(m) = \frac{1}{m} - \int_0^1 \int_0^1 p(t, x, y) y^m (1 - x)m(dy)m(dx)$$

and $m(dx) = \frac{dx}{x(1-x)}$, $t = t_{div}$. The transition density $p(t, x, y)$ is smooth, symmetric, and such

that the integral $u(t, x) = \int_0^1 p(t, x, y)f(y)m(dy)$ is the solution of the diffusion equation

$$\begin{cases} \frac{\partial}{\partial t} u(t, x) = x(1 - x) \frac{\partial^2}{\partial^2 x} u(t, x), & t > 0, \quad 0 < x < 1 \\ u(t, 0) = u(t, 1) = 0, & u(0, x) = f(x) \end{cases} \quad (S4)$$

for any continuous function $f(x)$ on $0 \leq x \leq 1$ (see mathematical derivations in Amei and Sawyer (2010)).

Table S1 PRF model and *BEAST parameter estimates for 22 sister pairs that exhibit a genetic break across the Isthmus of Panama.

Sister Pair	Taxa	Habitat Group	t _{div}	95%CI	t _{div} *BEAST	95% CI *BEAST	N _e
1	<i>Campylorhynchus chiapensis</i> & <i>C. griseus</i>	Dry	0.84	(0.43-1.92)	1.11	(0.50-1.94)	0.44
2	<i>Cardinalis sinuatus</i> & <i>C. phoeniceus</i>	Dry	2.85	(1.73-4.79)	4.92	(3.43-6.47)	0.9
3	<i>Forpus cyanopygius</i> & <i>F. passerinus</i>	Dry	3.82	(2.27-7.21)	4.09	(2.72-5.53)	0.67
4	<i>Granatellus venustus</i> & <i>G. pelzelni</i>	Dry	4.29	(2.75-6.53)	6.13	(3.99-8.28)	1.33
5	<i>Icterus graceannae</i> & <i>I. pectoralis</i>	Dry	1.88	(1.01-3.71)	2.19	(1.15-3.19)	0.8
6	<i>Pheugopedius sclateri</i> & <i>P. felix</i>	Dry	1.02	(0.63-1.72)	1.15	(0.29-2.15)	1.67
7	<i>Quiscalus mexicanus</i> & <i>Q. lugubris</i>	Dry	0.79	(0.36-1.81)	1.05	(0.57-1.55)	0.78
8	<i>Rhodothraupis celaeno</i> & <i>Periporphyrus erythromelas</i>	Dry	4.83	(3.15-8.26)	5.2	(3.07-7.72)	1.25
9	<i>Troglodytes aedon</i>	Dry	1.67	(1.15-2.38)	2.14	(1.47-2.83)	3.64
10	<i>Baryphthengus martii</i>	Humid	2.16	(1.51-3.11)	3.24	(1.92-4.60)	2.98
11	<i>Cantorchilus nigricapillus</i>	Humid	0.83	(0.45-1.41)	0.88	(0.31-1.50)	1.89
12	<i>Caryothraustes canadensis</i> & <i>C. poliogaster</i>	Humid	1.26	(0.79-2.04)	1.3	(0.45-2.22)	1.8
13	<i>Cyanocompsa cyanoides</i>	Humid	0.41	(0.19-0.87)	0.76	(0.24-1.37)	0.87
14	<i>Hylophilus ochraceiceps</i>	Humid	1.07	(0.58-2.08)	2.23	(1.11-3.31)	1.23
15	<i>Microbates cinereiventris</i>	Humid	1.2	(0.75-1.92)	1.69	(0.84-2.64)	1.81
16	<i>Notharchus tectus</i>	Humid	2.45	(1.53-4.19)	3.09	(1.62-4.79)	1.29
17	<i>Pionus senilis</i> & <i>P. chalcopterus</i>	Humid	0.36	(0.15-0.75)	0.31	(0.06-0.63)	0.62
18	<i>Saltator atripennis</i> & <i>S. atriceps</i>	Humid	1.76	(1.11-2.93)	3.08	(1.75-4.48)	1.33
19	<i>Saltator maximus</i>	Humid	0.81	(0.48-1.31)	0.56	(0.26-0.91)	1.55
20	<i>Trogon rufus</i>	Humid	2.81	(1.96-4.16)	3.93	(2.46-5.53)	2.27
21	<i>Xenops minutus</i>	Humid	1.13	(0.70-1.73)	0.42	(0.15-0.74)	2.5
22	<i>Xiphorhynchus susurrans</i>	Humid	0.89	(0.50-1.55)	1.05	(0.51-1.61)	1.17

Sister pairs referred to taxa that exhibit a genetic break across the Isthmus of Panama. Habitat group corresponds to preferred habitat of each sister pair. Median divergence time estimated using PRF model (t_{div}) and 95% credible interval are reported in units of millions of years. Mean divergence times and 95% credible interval using *BEAST from Smith *et al.* (2012) are shown. PRF model estimates of effective population sizes (N_e) in units of a million are reported.