Supporting Information

Rabosky and Matute 10.1073/pnas.1305529110

SI Materials and Methods

Modeling the Rate of Reproductive Isolation Evolution. Our general inference framework is essentially a random-effects model, where the random effects themselves are the key quantities of interest. Given a species-specific velocity of reproductive isolation (RI) evolution, ψ, we can define three functional models for the evolution of RI with respect to genetic distance. In addition to a simple linear model, we considered an asymptotic model and a quadratic model. The asymptotic model allowed the rate of RI accumulation to decelerate as a function of genetic distance, and the quadratic model allowed for the possibility of an acceleration in the rate of RI accumulation. For *Drosophila*, the asymptotic prediction model for reproductive isolation (Y) for sympatric pairs of species i and j is given by

$$Y_{ij} = \frac{\left(\beta_{0,S} + \left(\beta_{1,S} + \psi_i + \psi_j\right)X_{ij}\right)}{\left(1 + \left(\beta_{1,S} + \psi_i + \psi_j\right)X_{ij}\right)} + \varepsilon,$$

where $\beta_{0,S}$ and $\beta_{1,S}$ are the corresponding parameters for sympatric species pairs and X is the corresponding genetic distance. Thus, the simplest version of the model for the *Drosophila* datasets, where all species have identical ψ values ($\psi = 0$), contains five parameters: two shape parameters for sympatric and allopatric species pairs, respectively, plus the variance parameter for the error distribution. For birds, the corresponding model contained three parameters, because the data did not distinguish between sympatric and allopatric species pairs. The same general logic underlies our implementation of the linear and quadratic functional models. For the special case with clade-specific ψ values, all species within a particular clade are simply constrained to have identical ψ values. The full model, with a separate ψ parameter for each species, accounts for statistical nonindependence of observations (e.g., a given species may be represented by multiple crosses). However, we generally found that clade-specific models outperformed models with species-specific ψ values as well as models with $\psi = 0$ (with the exception of *Drosophila* premating isolation).

We used a censored model to estimate species-specific and clade-specific values of ψ, such that $0 \leq Y \leq 1.0$. This enables us to account for the contribution of ψ values to total reproductive isolation in fully probabilistic framework. In the censored model, the likelihood of the model parameters given the data are given by

$$L(\theta|D) = \prod_{y \in L} F(y_L = 0) \prod_{y \in U} [1 - F(y_U = 1)] \prod_{y \notin L,U} f(y),$$

where y_L and y_U are the lower and upper observable values (0 and 1), and $f(...)$ and $F(...)$ are the probability density functions and cumulative distribution functions for each observation as specified by parameters β_0, β_1, ψ, and ε.

We implemented the models described above and in the main text in a Bayesian framework. The posterior distribution of the parameters given the data D can be written as

$$f(\beta, \psi, \theta|D) \propto f(D|\beta, \psi, \theta) f(\beta) f(\psi) f(\theta),$$

where β and θ denote shape and error parameters for the model and ψ denotes species or clade-specific factors that affect the rate at which RI evolves. We specified a normal ($\mu = 0$, $\sigma = 1$) prior on the distribution of species- and clade-specific ψ values, to reflect our belief that these values should, on average, be equal to zero. From the perspective of our analyses, the variance of the prior distribution on ψ is largely irrelevant: We are concerned with the relative rates of evolution of RI for individual species or clades. We also placed normal ($\mu = 0$, $\sigma = 1$) priors on all β_0 parameters, and normal ($\mu = 0$, $\sigma = 5$) priors on β_1 parameters. Finally, we assumed a lognormal prior on the error variance with a log-transformed mean of −1 and a log SD of 1.

All models were analyzed by simulating the joint posterior density of model parameters using Markov chain Monte Carlo (MCMC). For our MCMC implementation, all unbounded parameters (β and ψ) were updated using a sliding window proposal mechanism. The sliding window proposal involves the addition of a small uniformly distributed (−U, +U) random variable to the current value of a particular parameter, where U is a tuning parameter that can be arbitrarily changed to facilitate efficient simulation of the posterior. The proposal ratio for the uniform sliding window proposal is 1.0. The error variance θ was updated using a proportional shrinking-expanding mechanism (1), such that a proposal involved choosing a new parameter value as

$$\theta' = \theta V = \theta e^{z(r-0.5)},$$

where z is a tuning parameter and r is sampled from a uniform (0, 1) distribution. The proposal ratio for this update is V. MCMC simulations were performed multiple times using a range of starting values to ensure that posterior distributions for parameters and marginal likelihood estimates converged on similar values. All MCMC chains were run for 20 million generations, updating β, ψ, and ε parameters with relative frequency 50:1:1. Convergence and appropriate burn-in thresholds were estimated from simulation output by computing the effective sample sizes for each parameter using the CODA library for the R programming environment.

For models with clade-specific rates of RI evolution, we numerically maximized the posterior probability of the data using standard numerical methods for optimization. This provided a direct estimate of the parameter set with the maximum a posteriori (MAP) probability. This approach was not feasible for models with species-specific ψ values, owing to the large number of parameters in the model. All results presented in the text for clade-specific models used MAP parameter estimates optimized in this fashion, but credible intervals on parameters (e.g., Fig. 3) were derived from the marginal posterior densities for each parameter simulated using MCMC. For each model with clade-specific ψ values, we estimated MAP parameter values from 2,000 independent optimizations using randomly sampled starting parameters.

***Drosophila* Diversification Rates.** Using the BEAST-derived estimates of crown-clade age, we computed estimators of net speciation rates (2) under relative extinction rates of 0 and 0.95. These estimators make the strong assumption that speciation rates have been constant through time and have been shown to perform poorly in many real datasets when age and species richness are decoupled (3–6). Within the full set of nine *Drosophila* clades, (log-transformed) species richness and clade age are not significantly correlated ($r_p = 0.33$, $P = 0.39$; $r_s = 0.49$, $r_s = 0.18$). However, this lack of relationship is largely attributable to the inclusion of the Hawaiian *Drosophila* radiation, which is a clear outlier in species richness ($n > 1{,}000$). When this young but

exceptionally species-rich radiation is excluded, crown clade age is significantly correlated with log-transformed species richness ($r_p = 0.73$, $P = 0.041$; $r_s = 0.72$, $P = 0.045$). This suggests that crown-based estimators of speciation rates provide some information about the underlying dynamics of speciation in major *Drosophila* clades. We used crown clade estimators of net diversification because our taxon sampling generally included most or all representatives of subsubgroups within each of the nine focal clades.

"ananassae" *clade.* Our sampling includes representatives of main complexes, *ananassae* (*ananassae*) plus *bipectinata* (*bipectinata*), as well as *Drosophila ercepeae*.

"melanogaster" *clade.* We included representatives of the major subgroups: *ficusphila* (*ficusphila*), *takahashii* (*takahashii*), *suzukii* (*lucipennis*), *melanogaster* (*melanogaster* and all other species), *eugracilis* (*eugracilis*), and *elegans* (*elegans*). Our analyses recovered these groups as sister lineages to the core eight species of the *melanogaster* species group. This result is consistent with Flybase taxonomy (www.flybase.org) and systematic studies (7).

"montium" *clade.* Our analyses included representatives of *serrata* (*serrata*), *auraria* (*auraria*), *kikkawai* (*kikkawai*), and *bakoue* (*tsacasi*) groups.

"willistoni" *clade.* *Drosophila nebulosa* and *Drosophila willistoni* are believed to span crown for *D. willistoni* species group (based on the Flybase taxonomy). These are representatives from the *bocainensis* subgroup and *willistoni* subgroup).

"obscura" *clade.* We included one member of each of the three main subgroups that seem to span the crown (*obscura*, *affinis,* and *pseudoobscura*).

"virilis" *clade.* We included representatives from each of the two main subgroups: *montana* (*Drosophila montana*) and *virilis* (*Drosophila virilis*).

Hawaii clade. We included five representatives from the Hawaiian *Drosophila* radiation in our analyses. These representatives span the core Hawaiian radiation in Van der Linde et al. (8). Previous studies have supported the monophyly of the Hawaiian radiation (9).

"repleta" *clade.* We included representatives from the four major subclades within this group: *hydei* (*Drosophila hydei*), *mulleri* (*Drosophila mulleri*), *repleta* (*Drosophila repleta*), and *mercatorum* (*Drosophila mercatorum*).

immigrans. *Drosophila hypocausta* and *Drosophila immigrans* seem to span the crown of this diverse clade (10) and are included in our analysis (along with three other species).

Species richness estimates for each clade were taken from several literature sources (10–12). For phylogenetic generalized least-squares (PGLS) analyses involving *Drosophila*, we pruned the *Drosophila* maximum clade credibility (MCC) tree to include only representatives from each of the nine major clades described above. We fixed the terminal branch lengths for each clade to be equal to one-half of the stem-clade age for each group.

Avian Phylogeny and Speciation Rates. We used two phylogenies as a framework for estimating speciation rates in birds, as well as for PGLS analyses reported in the main text (Table 1). Both are derived from the Jetz et al. (13) time-calibrated phylogeny for all birds. We analyzed the MCC trees for both the Hackett et al. (14) and Ericson (15) backbone trees used by Jetz et al. (13). These trees contained only the 6,670 species for which genetic data were available; hence, we did not include those species whose phylogenetic positions were estimated from taxonomic information alone (13).

To estimate branch-specific rates of speciation in birds, we applied a new Bayesian method [Bayesian Analysis of Macroevolutionary Mixtures (BAMM)] for the analysis of speciation and extinction rates to the time-calibrated avian phylogeny from Jetz et al. (13). The method (16) uses reversible-jump MCMC to move between model subspaces that vary in the number of distinct diversification regimes. The BAMM model assumes that a given phylogeny has been shaped by a mixture of distinct, potentially time-varying processes of diversification. The model proposes that "events," or rate shifts, are added to or removed from the phylogeny according to a compound Poisson process. The occurrence of an event on a particular branch v_e defines an "event subtree" τ_e: All nodes and branches descended from v_e inherit the collection of evolutionary processes Φ_1 (defined by the event at v_e), until the event is terminated. An event terminates at terminal branches, or at the next downstream event. Speciation rates within each event were modeled as

$$\lambda(t) = \lambda_0 e^{kt},$$

where t is the elapsed time from the initial occurrence of the event, λ_0 is the initial speciation rate, and k is the rate at which the speciation rate changes through time. Extinction rates (μ) were assumed to be constant in time within a given event subtree. Each event subtree τ_e is thus associated with three parameters: λ, k, and μ. A full description of the model and likelihood calculations when rates are constant through time within each event subtree is given in Rabosky et al. (16).

We implemented the model in a Bayesian framework, integrating over prior distributions to obtain marginal distributions of evolutionary rates for each branch in the avian tree. Under a compound Poisson process, the number of events on the tree is a Poisson-distributed random variable that occurs with rate Λ. These events can change position, can be deleted from the tree, and occur on any branch with probability proportional to the length of the branch. There is no upper bound on the number of events, because multiple events can occur on each branch. The full model contains the following parameters: (*i*) Λ, the overall event rate Λ; (*ii*) the root rates λ_R, μ_R, and k_R; and (*iii*) λ, μ, and k and location parameters for each event. The BAMM model implements the Metropolis–Hastings–Green algorithm (17) to construct a Markov chain that allowed movement between model spaces of different dimensionality, as occurs when events are added to or deleted from the tree.

When an event was added to the tree, all branches and nodes downstream of the new event "belong" to the event, and they are governed by the evolutionary rates [e.g., $\Phi_1 = (\lambda_1, \mu_1, k_1)$] associated with the event. The addition of a new event on some subtree would result in a new collection of processes or rates (Φ_1). Following the proposal of a new event, new rate parameters were sampled from prior distributions for λ, μ, and k. This leads to a Jacobian matrix with an absolute value of 1 for the bijection between a state with N events and a state with $N + 1$ events. Likewise, deletion of an event collapses rates on all branches governed by a given event to the rates associated with the parent event. In the preceding example, deletion of the single event would involve setting all branches/nodes governed by Φ_1 to the parent process at the root (Φ_R). A tree with 0 events is governed strictly by the evolutionary processes at the root and would only contain three parameters in the current model formulation (λ_R, μ_R, k_R). There is no upper bound on the number of events, because multiple events can occur on each branch. The Jacobian for transdimensional moves (additions and deletions) is 1; all proposal ratios are described previously (16, 18).

In summary, the BAMM model assumes that changes in evolutionary regimes occur across the branches of phylogenetic trees under a compound Poisson process model of rate variation and explicitly allows rates to vary both through time and among lineages. We placed a prior expectation of 0.5 events on the avian tree. With this prior expectation, the 95% credible interval on the number of events was 20–24. We computed the mean of the posterior distribution of speciation rates across each branch in the full avian tree. We used these branch-specific speciation rate

estimates in all subsequent analyses and to compute mean rates of speciation for each family of birds in our analyses.

We also predicted that the maximum speciation rates within clades might be more likely to respond to variation in ψ. If species richness within clades is regulated by ecological or other extrinsic factors, then simple "constant rate" estimators of speciation might not reflect the true tempo of speciation (6, 19). Thus, for birds, we also considered the maximum observed speciation rate within clades, predicting that maximum speciation rates would be more likely to reflect the intrinsic speciation potential of clades and be less likely to show evidence for ecological dampening, as predicted under diversity-dependent models of cladogenesis (20–25). Finally, to reduce the sensitivity of our family-level analyses to errors in estimating maximum speciation rates, we estimated the 0.80 quantile of the distribution of rates within each family. This metric, λ_{q80}, should also reflect the intrinsic potential for speciation of a given clade while reducing the effect of ecological dampening on overall clade-specific rates of speciation.

Power Simulations. We conducted a series of simulations to assess whether our analyses would have had sufficient power to detect a true correlation between ψ and the rate of speciation. We simulated datasets with a known correlation between the rate of speciation and ψ and used these ψ values to generate pseudodatasets that were structurally identical to the actual data.

The first step of our simulation protocol was to generate a distribution of ψ values and speciation rates with a particular correlation. It is straightforward to generate bivariate normal data with a fixed Pearson correlation r_p. However, this is more difficult for nonnormal data. In the present example, speciation rate estimates across clades do not follow a normal distribution. For *Drosophila*, the distribution of λ_0 is better fit by a lognormal distribution than a normal distribution [Akaike information criterion (AIC), lognormal = 42.4; AIC, normal distribution = 46.8]. The lognormal also fits speciation rates better across avian families (for the Hackett λ: AIC, lognormal: −51.8; AIC, normal = −46.7). Hence, our protocol involved drawing bivariate normal data with a fixed correlation structure, then mapping those values onto distributions of ψ and speciation that were consistent with the observed data. Finally, we reestimated the correlation between ψ and speciation, because the correlation coefficient is not invariant under such transformation.

For each dataset, a single simulation consisted of the following procedure. We first sampled random deviates for each clade ($n = 9$ for *Drosophila*; $n = 30$ for birds) from a bivariate normal (mean = 0; SD = 1.0) distribution with fixed correlation r_p, where r_p was sampled from a uniform (0.25, 1.0) distribution. This gives a simulated draw of speciation and ψ for each clade, but the values are on a scale that is dissimilar to that of the true data. We scaled the vector of simulated ψ values to have a variance identical to the observed distribution of ψ values across clades. We then mapped the vector of simulated speciation rates onto a lognormal distribution that was identical in shape and variance to the observed distribution. To do so, we first converted all simulated speciation rates (standard normal deviates, at this point) to quantiles of the standard normal distribution. We then found the corresponding expected values of those quantiles from a lognormal distribution that was parameterized to the observed distribution of speciation rates for each group. We then recomputed the Pearson correlation between the speciation rate and ψ. We recorded this value as the "true" underlying correlation between speciation and ψ.

Each dataset simulated in this fashion was then used to generate a pseudodataset of RI. Using clade-specific (simulated) values of ψ and the MAP estimated error variance, we simulated a new dataset that contained exactly the same number of crosses per clade as the observed data. To be clear, each simulated observation also contained a noise term drawn from an error distribution with variance identical to the one estimated for the real dataset. We then estimated clade-specific ψ values from each simulated dataset. Finally, we computed the correlation between the estimated ψ values and (fixed) speciation rates for the simulated data. We performed these power simulations under a linear model for the two postzygotic datasets and the asymptotic model for the premating dataset, owing to the poor fit of the linear model in the case of *Drosophila* premating isolation. Results are shown in Fig. S6 (left column) and indicate that power to detect a true correlation between ψ and speciation is at least moderate for both postzygotic datasets.

The preceding analyses are designed to address the following: Given a true (known) correlation between ψ and the rate of speciation, what distribution of P values would we expect to observe, given the structure of the data? These simulations also enable us to ask a complementary question: Given the predicted distribution of correlation coefficients for a given simulation parameterization, how plausible are the observed values? Given a generating model with true correlation coefficient $\rho = x$, we are interested in the probability of obtaining our observed correlation y, or Pr $(y \mid \rho = x)$. For example, all of our estimated Pearson and Spearman correlation coefficients for the *Drosophila* postzygotic dataset were negative but nonsignificant. One interpretation of Fig. S6 is that we should have had only low to moderate power to obtain a significant result at the $\alpha = 0.05$ level if the true Pearson correlation between ψ and the rate of speciation was 0.50 or less. Using our simulation results described above, we estimated the one-tailed probability of the observed correlation coefficients under a model with a true positive correlation between ψ and speciation.

Our results indicate that, for postzygotic isolation in both *Drosophila* and birds, the observed correlation coefficients (both Pearson and Spearman) are generally too low relative to the distributions predicted if a true positive correlation existed (Fig. S6, right column). For postzygotic isolation in *Drosophila*, the observed negative correlations between ψ and speciation rate are unlikely under all correlational scenarios examined. The probability of obtaining a Pearson correlation equal to −0.31 (the observed value for both λ_0 and λ_{95}) if the true correlation is 0.40 is less than 0.01. Similar results were obtained for birds, such that the probability of the observed correlations was also low under a model with a true positive correlation. For prezygotic isolation in *Drosophila*, the observed correlation coefficients would have been plausible under a broad range of true correlations, consistent with low overall power for this dataset. In summary, at least for postzygotic isolation, we would have had only low to moderate power to detect a true correlation between speciation and ψ, if this correlation was not strong. However, our observed correlations are generally lower than we would expect if there was a substantial (positive) correlation between ψ and speciation.

The interpretation of these simulations should be treated provisionally for several reasons. First, our simulation protocol tests whether error in and/or the structure of the RI data could have obscured a true relationship between ψ and the speciation rate. As such, it ignores other forms of error that may have influenced the reconstructed correlation between these parameters. Second, we have no information about "reasonable" values for the correlation between speciation rates and ψ. If the rate at which RI evolves is the dominant control on macroevolutionary speciation dynamics, then we would expect this correlation to be strong. However, many other factors might affect this relationship and we can at best provide some insight into the possible range of correlations between ψ and λ that would have been likely to leave a signal in the data. Finally, we lack an appropriate model for generating speciation rates under a particular correlation structure with ψ. In the simulations described above, we sampled clade-specific speciation rates from a distribution

with variance identical to the observed (estimated) rates for clades, but it is possible that other models might be more appropriate and/or lead to higher or lower power than that which we report here.

1. Yang Z (2006) *Computational Molecular Evolution* (Oxford Univ Press, Oxford).
2. Magallón S, Sanderson MJ (2001) Absolute diversification rates in angiosperm clades. *Evolution* 55(9):1762–1780.
3. Rabosky DL (2009) Ecological limits on clade diversification in higher taxa. *Am Nat* 173 (5):662–674.
4. Rabosky DL (2010) Primary controls on species richness in higher taxa. *Syst Biol* 59(6): 634–645.
5. Rabosky DL, Adams DC (2012) Rates of morphological evolution are correlated with species richness in salamanders. *Evolution* 66(6):1807–1818.
6. Rabosky DL, Slater GJ, Alfaro ME (2012) Clade age and species richness are decoupled across the eukaryotic tree of life. *PLoS Biol* 10(8):e1001381.
7. Lemeunier F, Davis J, Tsacas L, Ashburner M (1986) The melanogaster species group. *The Genetics and Biology of Drosophila*, eds Ashburner M, Carson HL, Thompson J (Academic, London), pp 147–256.
8. van der Linde K, Houle D, Spicer GS, Steppan SJ (2010) A supermatrix-based molecular phylogeny of the family Drosophilidae. *Genet Res* 92(1):25–38.
9. Remsen J, O'Grady PM (2002) Phyogeny of Drosophilinae (Diptera: Drosophilidae), with comments on combined analysis and character support. *Mol. Phy. Evol.* 24: 249–264.
10. Katoh T, Nakaya D, Tamura K, Aotsuka T (2007) Phylogeny of the Drosophila immigrans species group (Diptera: Drosophilidae) based on Adh and Gpdh sequences. *Zoolog Sci* 24(9):913–921.
11. Da Lage JL, et al. (2007) A phylogeny of the Drosophilidae using the Amyrel gene: questioning the Drosophila melanogaster species group boundaries. *J Zoological Syst Evol Res* 45:47–63.
12. Markow TA, O'Grady PM (2006) *Drosophila: A Guide to Species Identification and Use* (Academic, London).
13. Jetz W, Thomas GH, Joy JB, Hartmann K, Mooers AO (2012) The global diversity of birds in space and time. *Nature* 491(7424):444–448.
14. Hackett SJ, et al. (2008) A phylogenomic study of birds reveals their evolutionary history. *Science* 320(5884):1763–1768.
15. Ericson PGP (2012) Evolution of terrestrial birds in three continents: Biogeography and parallel radiations. *J Biogeogr* 39:813–824.
16. Rabosky DL, et al. (2013) Rates of speciation and morphological evolution are correlated across the largest vertebrate radiation. *Nat Commun* 4:1958, 10.1038/ncomms2958.
17. Green PJ (1995) Reversible jump Markov chain Monte Carlo computation and Bayesian model determination. *Biometrika* 82:711–732.
18. Huelsenbeck JP, Larget B, Swofford D (2000) A compound poisson process for relaxing the molecular clock. *Genetics* 154(4):1879–1892.
19. Rabosky DL (2009) Ecological limits and diversification rate: Alternative paradigms to explain the variation in species richness among clades and regions. *Ecol Lett* 12(8): 735–743.
20. Etienne RS, Haegeman B (2012) A conceptual and statistical framework for adaptive radiations with a key role for diversity dependence. *Am Nat* 180(4):E75–E89.
21. Etienne RS, et al. (2012) Diversity-dependence brings molecular phylogenies closer to agreement with the fossil record. *Proc Biol Sci* 279(1732):1300–1309.
22. Phillimore AB, Price TD (2008) Density-dependent cladogenesis in birds. *PLoS Biol* 6(3): e71.
23. Rabosky DL, Glor RE (2010) Equilibrium speciation dynamics in a model adaptive radiation of island lizards. *Proc Natl Acad Sci USA* 107(51):22178–22183.
24. Rabosky DL, Lovette IJ (2008) Density-dependent diversification in North American wood warblers. *Proc Biol Sci* 275(1649):2363–2371.
25. Sepkoski JJ (1978) A kinetic model of Phanerozoic taxonomic diversity I. Analysis of marine orders. *Paleobiology* 4:223–251.

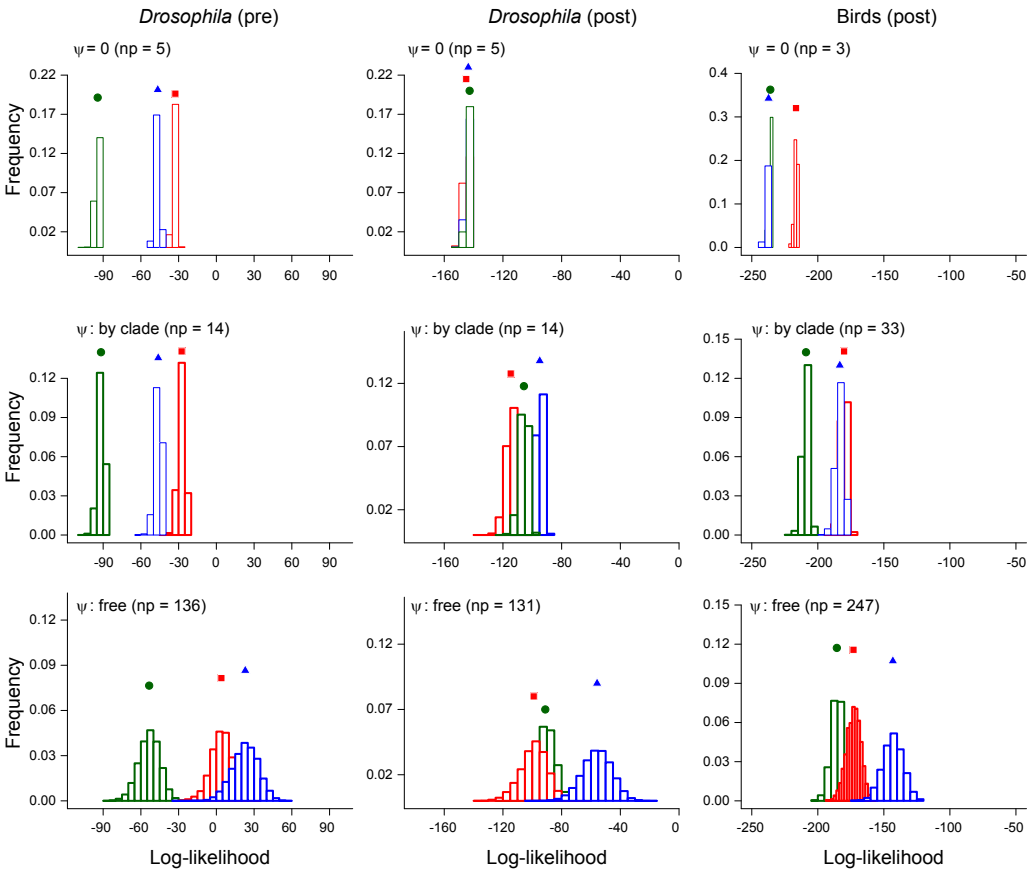

Fig. S1. Posterior distributions of log-likelihoods from Markov chain Monte-Carlo analysis of models for the evolution of reproductive isolation as a function of genetic distance. Each column shows the results of fitting models to a different dataset (*Left*, *Drosophila*, premating RI; *Center*, *Drosophila*, postzygotic RI; *Right*, birds, postzygotic RI). Each panel corresponds to a particular model parameterization with respect to ψ, the species-specific or clade-specific velocity parameter for the rate of evolution of RI (top row, $\psi = 0$; middle row, clade-specific ψ values; bottom row, species-specific ψ values). Histograms give the posterior distribution of log-likelihoods for each of three functional models for the accumulation of RI under each ψ parameterization (green, quadratic or "snowball" model; red, asymptotic model; blue, linear model). Colored symbols denote mean posterior log-likelihoods for each distribution (red squares, asymptotic; green circles, quadratic; blue triangles, linear). The number of parameters in each model is given by np. Table S1 shows comparison of corresponding AIC_M scores for each model.

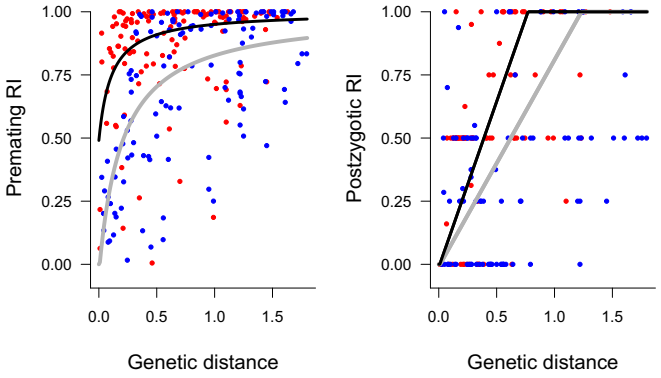

Fig. S2. Accumulation of premating (*Right*) and postzygotic (*Left*) reproductive isolation in *Drosophila* for sympatric (red) and allopatric (blue) species pairs. Lines show the corresponding model-predicted values under the best-fit model for each dataset (premating, asymptotic; postzygotic, linear). Black and gray lines denote predictions for sympatric and allopatric pairs, respectively. All models assumed censoring of RI observations outside of the [0, 1] range.

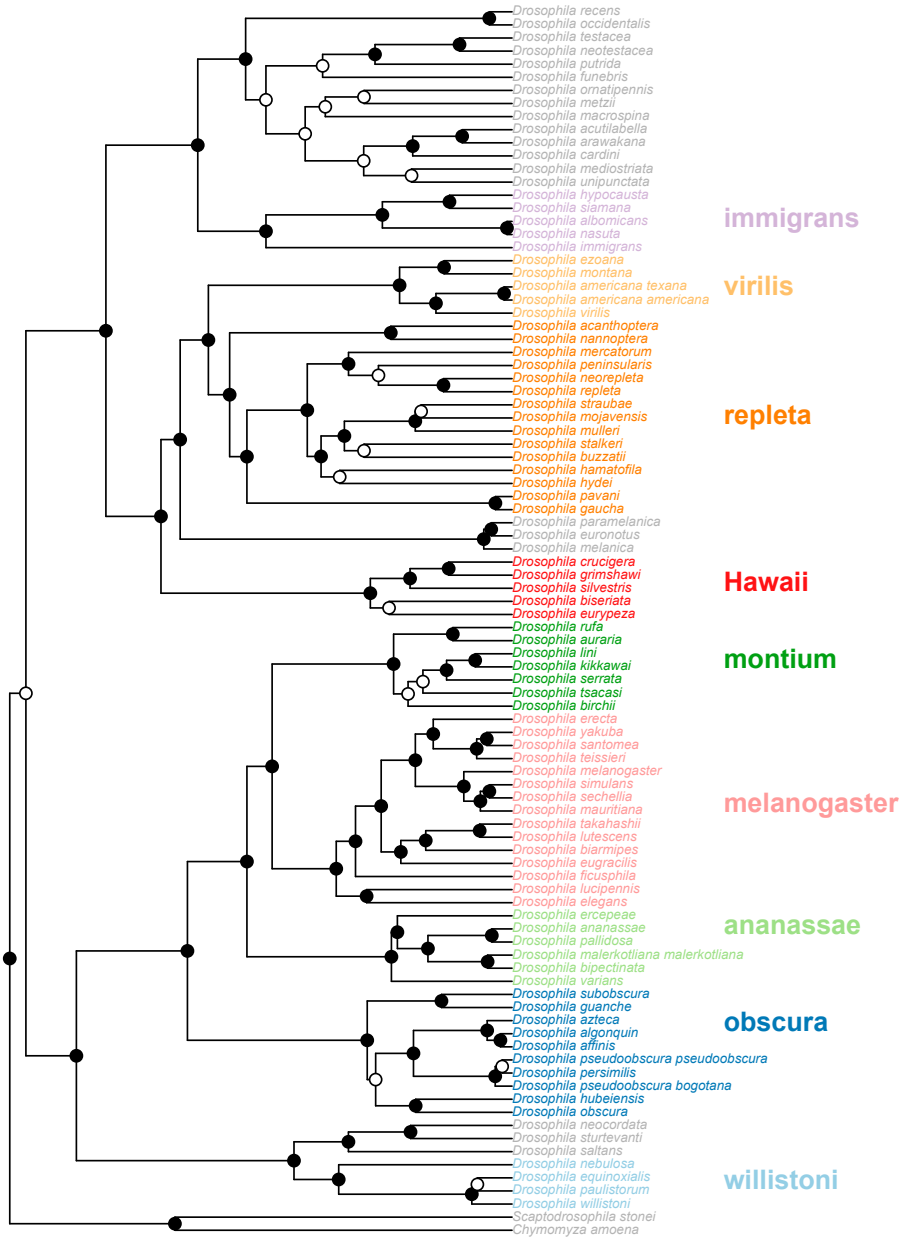

Fig. S3. Maximum clade credibility tree for 94 *Drosophila* species from BEAST analysis, showing the nine major clades for which clade-specific ψ values were estimated. Black circles on interior nodes denote Bayesian posterior probabilities >0.95.

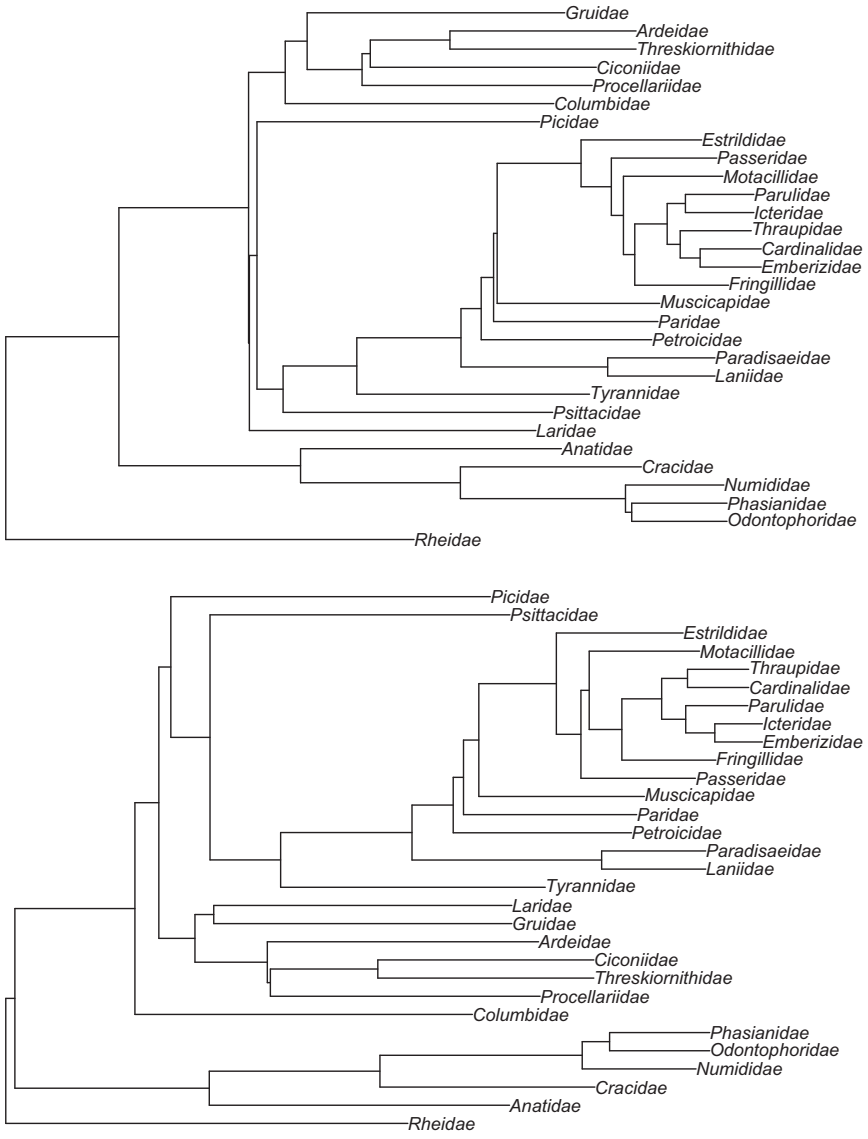

Fig. S4. Pruned MCC trees used for PGLS analyses of avian families. (*Upper*) Hackett backbone tree. (*Lower*) Ericson backbone tree. Length of terminal branches was set equal to 50% of the stem clade age for each family.

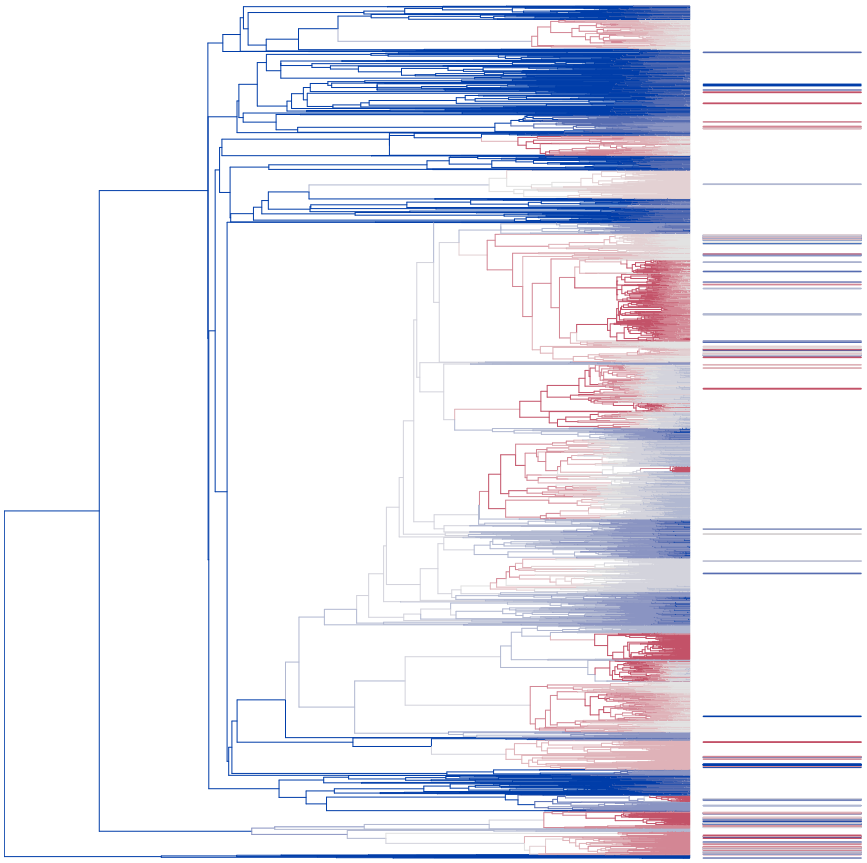

Fig. S5. "Hackett" tree with branch-specific estimates of speciation rates across 6,670 extant bird species from BAMM analysis. Red, gray, and blue denote fast, intermediate, and slow speciation rates, respectively. Horizontal bars to right of tree denote corresponding species-specific estimates of ψ, using the same color scheme. If ψ and speciation rates are correlated, then clades with fast speciation (red) should tend to have an excess of fast (red) ψ values. No such pattern occurs in the data (see also Fig. 4).

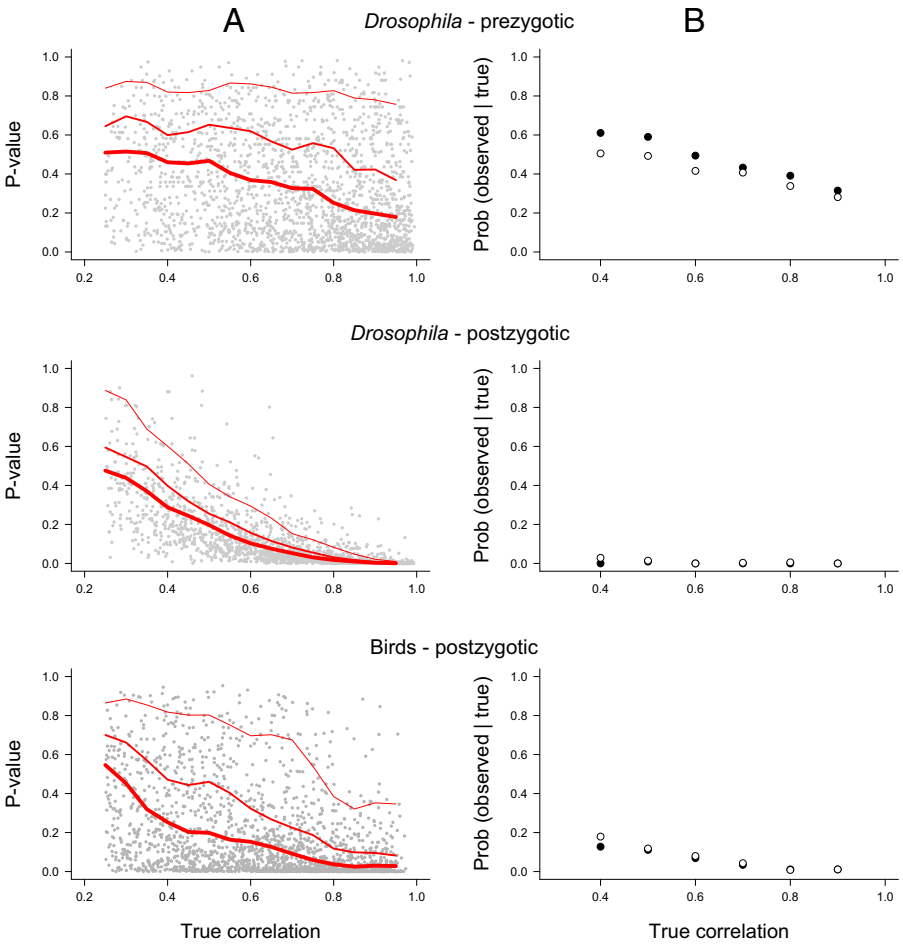

Fig. S6. Power to detect a true correlation between speciation and the rate of evolution of reproductive isolation for each dataset. The *x*-axis shows the true (Pearson) correlation between the speciation rate and clade-specific values of ψ in simulated datasets with the same error structure. (*A*, left column) The *y*-axis shows estimated minimum *P* value (Pearson or Spearman) for the correlation after numerically finding MAP estimates for clade-specific ψ values for each simulated dataset. (*B*, right column) One-tailed probability of the observed correlation coefficients between ψ and speciation for all three datasets, assuming a true positive relationship. Black and white circles denote results for Pearson and Spearman correlation coefficients, respectively. For postzygotic isolation in *Drosophila* and birds, the observed correlation coefficients (e.g., Table 1) are generally inconsistent with the possibility that ψ and speciation are substantially correlated. Power to detect an effect is low for the *Drosophila* premating dataset, where no significant difference in ψ values among clades was observed. However, power is considerably greater for both postzygotic datasets. The estimated power curve (*A*) suggests that for moderate-to-high correlations between speciation and ψ, statistical analyses should have trended toward significance. Our empirical analyses show no such trend and the estimated correlation coefficients are not consistent with a true positive relationship (*B*). For *Drosophila* postzygotic isolation, the minimum observed (two-tailed) *P* value for the Pearson correlation between ψ and any metric of speciation rate was 0.41, and the corresponding minimum Spearman *P* value was 0.39; moreover, the estimated correlation coefficient was actually negative for all speciation metrics and models (Table 1 and Table S3), despite the fact that we should have had greatest power to detect a true effect for this dataset. For birds, the minimum observed Pearson *P* value across 18 statistical tests was 0.5 (mean *P* value 0.82), and the minimum Spearman correlation was 0.23 (mean *P* value across 18 tests 0.55).

Table S1. Relative fit of nine models for the accumulation of RI through time for three datasets

Group	Model	Type	NP*	AIC_M	ΔAIC_M
Birds	Asymptotic	Clade	33	374	0
Birds	Linear	Clade	33	386	12
Birds	Linear	Full	247	404.9	30.9
Birds	Asymptotic	Full	247	405.4	31.4
Birds	Quadratic	Full	247	408.3	34.3
Birds	Quadratic	Clade	33	431	57
Birds	Asymptotic	$\psi = 0$	3	435.9	61.9
Birds	Quadratic	$\psi = 0$	3	475.1	101.1
Birds	Linear	$\psi = 0$	3	477.8	103.8
Drosophila (post)	Linear	Clade	14	203.9	0
Drosophila (post)	Quadratic	Clade	14	228.9	25
Drosophila (post)	Asymptotic	Clade	14	253.3	49.4
Drosophila (post)	Quadratic	Full	131	269.4	65.5
Drosophila (post)	Quadratic	$\psi = 0$	5	291.1	87.2
Drosophila (post)	Linear	$\psi = 0$	5	292.7	88.8
Drosophila (post)	Asymptotic	$\psi = 0$	5	295.1	91.2
Drosophila (post)	Linear	Full	131	310.3	106.4
Drosophila (post)	Asymptotic	Full	131	367.3	163.4
Drosophila (pre)	Asymptotic	Clade	14	69.4	0
Drosophila (pre)	Asymptotic	$\psi = 0$	5	70.4	1
Drosophila (pre)	Linear	$\psi = 0$	5	98.1	28.7
Drosophila (pre)	Linear	Clade	14	105.7	36.3
Drosophila (pre)	Asymptotic	Full	136	135.3	65.9
Drosophila (pre)	Linear	Full	136	171.1	101.7
Drosophila (pre)	Quadratic	$\psi = 0$	5	193.7	124.3
Drosophila (pre)	Quadratic	Clade	14	196.6	127.2
Drosophila (pre)	Quadratic	Full	136	250.1	180.7

The three datasets are birds, postzygotic; *Drosophila*, postzgotic; and *Drosophila*, premating. Each functional model for the relationship between RI and genetic distance (asymptotic, linear, quadratic) was parameterized with species-specific ψ values, clade-specific ψ values, and $\psi = 0$, for a total of nine fitted models per dataset. NP* is the maximum number of parameters in each model. Because models were fitted in a hierarchical modeling framework with prior distributions on parameters, the effective number of parameters in the model need not correspond to this value. AIC_M values can be interpreted as approximate AIC scores. ΔAIC_M values indicate the difference in ΔAIC_M scores between each model and the overall best-fit model for each dataset. Fig. S1 shows posterior log-likelihood distributions.

Table S2. ANOVA analyses of clade-specific differences in ψ, using species-specific ψ estimates from each functional model for the accumulation of RI

Group	Model	df	*F*	*P*	Group variance fraction
Drosophila (pre)	Asymptotic	8, 122	1.87	0.07	0.11
Drosophila (pre)	Linear	8, 122	0.61	0.77	0.04
Drosophila (pre)	Quadratic	8, 122	1.76	0.09	0.1
Drosophila (post)	Asymptotic	8, 117	4.48	<0.001	0.23
Drosophila (post)	Linear	8, 117	7.43	<0.001	0.34
Drosophila (post)	Quadratic	8, 117	7.37	<0.001	0.33
Birds (post)	Asymptotic	29, 214	4.03	<0.001	0.35
Birds (post)	Linear	29, 214	2.67	<0.001	0.27
Birds (post)	Quadratic	29, 214	2.48	<0.001	0.25

Group variance fraction is the percentage of total variance in species-specific ψ values explained by clade (9 *Drosophila* clades or 30 avian family-level clades). Species-specific ψ values were computed as the mean of the marginal posterior density for each species simulated using MCMC. Pre, premating; post, postzygotic.

Table S3. Correlations between estimates of clade-specific rates of reproductive isolation (prezygotic or postzygotic) and metrics of macroevolutionary dynamics in birds and flies

Dataset	Model	Metric	r_p	p (r_s)	r_s	p (r_p)	Slope	t-statistic (PGLS)	PGLS P value	N
Drosophila (post)	Asymptotic	λ_0	−0.3	0.43	−0.33	0.39	−0.82	−1.29	0.24	9
Drosophila (post)	Linear	λ_0	−0.31	0.42	−0.08	0.84	−1.22	−1.38	0.21	9
Drosophila (post)	Quadratic	λ_0	−0.21	0.58	−0.12	0.78	−0.77	−1.09	0.31	9
Drosophila (post)	Asymptotic	λ_{95}	−0.27	0.48	−0.32	0.41	−0.42	−0.87	0.42	9
Drosophila (post)	Linear	λ_{95}	−0.31	0.41	−0.1	0.81	−0.71	−1.05	0.33	9
Drosophila (post)	Quadratic	λ_{95}	−0.2	0.61	−0.05	0.91	−0.39	−0.72	0.49	9
Drosophila (pre)	Asymptotic	λ_0	0.3	0.43	0.23	0.55	1.03	0.8	0.45	9
Drosophila (pre)	Linear	λ_0	0.36	0.34	0.17	0.68	25.64	1.38	0.21	9
Drosophila (pre)	Quadratic	λ_0	0.45	0.22	0.57	0.12	2.42	0.58	0.58	9
Drosophila (pre)	Asymptotic	λ_{95}	0.24	0.53	0.58	0.11	0.86	0.94	0.38	9
Drosophila (pre)	Linear	λ_{95}	0.28	0.46	0.47	0.21	19.59	1.46	0.19	9
Drosophila (pre)	Quadratic	λ_{95}	0.4	0.29	0.37	0.34	1.77	0.58	0.58	9
Birds (H)	Asymptotic	λ	0.04	0.83	0.05	0.79	−0.01	−0.43	0.67	30
Birds (H)	Linear	λ	−0.01	0.95	0	0.98	−0.01	−0.25	0.81	30
Birds (H)	Quadratic	λ	0	0.99	0.07	0.71	0	−0.21	0.84	30
Birds (E)	Asymptotic	λ	0.06	0.75	0.13	0.49	0.01	0.31	0.76	30
Birds (E)	Linear	λ	0.06	0.74	0.14	0.45	0.02	0.56	0.58	30
Birds (E)	Quadratic	λ	0.12	0.54	0.22	0.24	0.02	0.68	0.5	30
Birds (H)	Asymptotic	λ_{MAX}	−0.01	0.96	0.12	0.54	−0.04	−0.39	0.7	30
Birds (H)	Linear	λ_{MAX}	−0.03	0.88	0.05	0.77	−0.03	−0.25	0.81	30
Birds (H)	Quadratic	λ_{MAX}	0.04	0.84	0.13	0.49	−0.04	−0.39	0.7	30
Birds (E)	Asymptotic	λ_{MAX}	0.03	0.87	0.13	0.48	0.01	0.09	0.93	30
Birds (E)	Linear	λ_{MAX}	0.1	0.61	0.14	0.45	0.05	0.61	0.55	30
Birds (E)	Quadratic	λ_{MAX}	0.13	0.5	0.22	0.24	0.02	0.29	0.77	30
Birds (H)	Asymptotic	λ_{q80}	−0.01	0.98	0.08	0.66	−0.01	−0.34	0.73	30
Birds (H)	Linear	λ_{q80}	−0.03	0.87	0.04	0.82	−0.01	−0.3	0.77	30
Birds (H)	Quadratic	λ_{q80}	0.02	0.91	0.12	0.51	−0.01	−0.14	0.89	30
Birds (E)	Asymptotic	λ_{q80}	0.01	0.97	0.12	0.52	0.01	0.22	0.83	30
Birds (E)	Linear	λ_{q80}	0.03	0.87	0.13	0.48	0.03	0.46	0.65	30
Birds (E)	Quadratic	λ_{q80}	0.09	0.62	0.23	0.23	0.03	0.54	0.59	30

λ denotes speciation rate; λ_0 and λ_{95} denote estimates of speciation for *Drosophila* clades assuming relative extinction rates of 0 and 0.95, respectively. For birds, λ is the mean estimated rate for each family; λ_{MAX} is the maximum estimated rate; λ_{q80} is the estimated 0.80 quantile of the distribution of rates for each family. This latter value was chosen to reduce sensitivity to outlying estimates for the maximum rate. Analyses are conducted at the family level for birds (n = 30) and level of major clade for *Drosophila* (n = 9). r_p and r_s denote Pearson and Spearman correlations, respectively; p (r) is the corresponding two-tailed P value. Slope is the estimated slope from PGLS regression using estimated phylogeny shown in Fig. S3. N is the number of clades considered in each analysis. Birds (H) and birds (E) refer to analyses of the avian dataset conducted using Hackett (H) and Ericson (E) backbone phylogenies, respectively.