Supporting Information

Rabosky and Glor 10.1073/pnas.1007606107

SI Materials and Methods

Phylogenetic Methods. Phylogenetic trees were generated by Mahler et al. (1). Mahler et al. conducted multiple partitioned analyses in MrBayes v3.1.2; the consensus tree from these analyses was then used as the starting tree for the Bayesian relaxed clock method implemented in BEAST 1.4.7. All non-Greater Antillean taxa were then pruned from the trees in the posterior distribution of the BEAST analysis.

Diversification Analyses. The equations for modeling state-dependent diversification dynamics follow from Maddison et al. (2), and we used simple modifications of the original equations. The reader is referred to the original paper for a clear and detailed explanation of the derivation of the two-state BiSSE model (2). In this framework, we track—for each character state—two key probabilities: $D_i(t)$, the probability that a lineage in state i at time t gives rise to a clade like that in the observed data; and $E_i(t)$, the probability that a lineage in state i at time t goes extinct before the present (along with all of its descendants). Here, t is time measured from the present backward in time and Δt is an incremental time step. Let $\lambda_{i,t}$ and $\mu_{i,t}$ be the speciation and extinction rates for a lineage in state i at time t and $q_{ij,t}$ represent the rate of transition from state i to state j at time t. As in Maddison et al. (2) and FitzJohn et al. (3), we assume that only a single event (speciation, extinction, or character change) can happen in the interval Δt. The differential equation governing $D_i(t + \Delta t)$ is

$$\frac{dD_i}{dt} = -\left(\mu_{i,t} + \lambda_{i,t} + \sum_{j=1}^{j\neq i} q_{ij,t}\right)D_i(t) + \sum_{j=1}^{j\neq i} q_{ij,t}D_j(t) + 2\lambda_{i,t}E_i(t)D_i(t),$$

where

$$-\left(\mu_{i,t} + \lambda_{i,t} + \sum_{j=1}^{j\neq i} q_{ij,t}\right)$$

is the rate at which $D_i(t)$ decreases due extinction, speciation, or character change (from state i to all other states) in some infinitesimal amount of time Δt as we move backward down the branch (toward the root), and

$$q_{ij,t}D_j(t)$$

is the rate at which $D_i(t)$ increases due to character change from state i to state j [multiplied by the probability that, having undergone character change, the lineage will go on to evolve into a clade like the observed data, $D_j(t)$]. This term must be summed over all $n - 1$ character states, hence the summation term. Finally, a lineage at some point in time can undergo a speciation event and still give rise to the observed data, provided one of the descendant branches and all of its descendants go extinct before the present (thus, we would never observe a speciation event in a reconstructed tree). This process occurs with rate

$$\lambda_{i,t}E_i(t)D_i(t)$$

and it must be multiplied by 2 (as either the left or the right descendant branches and all their descendants could go extinct, if there is a speciation event on Δt).

The differential equation governing $E_i(t)$ is

$$\frac{dE_i}{dt} = \mu_{i,t} - \left(\mu_{i,t} + \lambda_{i,t} + \sum_{j=1}^{j\neq i} q_{ij,t}\right)E_i(t) + \sum_{j=1}^{j\neq i} q_{ij,t}E_j(t) + \lambda_{i,t}E_i(t)^2$$

and includes the rate at which lineages go extinct on Δt, or $\mu_{i,t}$. The equation includes the rate at which lineages undergo no events on the time interval δt, yet go extinct at some point in the future, or

$$\mu_{i,t} + \lambda_{i,t} + \sum_{j=1}^{j\neq i} q_{ij,t}$$

plus the rate at which lineages in state i switch states and subsequently go extinct, or

$$q_{ij,t}E_j(t),$$

which must be summed over all n character states. Finally, a lineage might undergo speciation on Δt, but for such an event to occur and yet result in a clade that goes extinct before the present, both descendant branches and all their descendants must go extinct. These events occur at rate

$$\lambda_{i,t}E_i(t)^2.$$

A total of $2n$ equations must thus be solved simultaneously to describe the dynamics of diversification and character change through time. The initial conditions, for a tree with complete taxon sampling, are $D_i(0) = 1$ and $E_i(0) = 0$. The probability that a lineage existing in the present will give rise to the observed data —a single lineage—is necessarily 1, unless we are accounting for incomplete taxon sampling. In this case, the initial state is simply the probability of sampling the lineage (3). We solved these systems of equations by numerically integrating backward along each branch using the Fortran LSODA integrator as implemented in the R package deSolve (4). At each interior node in the tree, probabilities $D_i(t)$ for right and left descendant branches were combined as in Maddison et al. (2), and these combined values became the initial conditions for integration backward down the parent branch. At the root node, we combined root-state probabilities using the weighting scheme from FitzJohn et al. (3). Virtually identical results were obtained using alternative weighting methods that assumed equal weights to the individual probabilities $D_1, D_2, \ldots, D_n$.

Biogeographic Model Selection. A large number of biogeographic scenarios could be considered to model transitions between character states (e.g., dispersal between islands). Under the simplest model, dispersal (transition) rates between all islands might be identical. This model would specify a symmetric, one-rate transition matrix between states. At the most complex end of the spectrum, each pair of character states might have separate asymmetric transition rates ($q_{ij} \neq q_{ji}$), and all rates might vary linearly through time. This model would have a full 24-parameter transition matrix between character states (12 initial rates at the root node and 12 ending rates in the present). Between these scenarios, a very large number of models are possible. We used the state-dependent diversification framework described above and subsequent model-fitting analyses to identify an appropriate background model for subsequent diversification analyses, rather than simply assuming the validity of a one-rate symmetric transition matrix or any other model.

We assumed that all character states (islands) were associated with a particular time-constant rate of speciation, with extinction set to zero. We then used maximum likelihood to fit three biogeographic models to the *Anolis* MCC tree: (*i*) a model with time-invariant, equal rates between all islands (1 parameter, model qsymm1); (*ii*) a model with time-varying dispersal between islands, but with a single rate for all islands at any point in time (2 parameters, model qsymm1_TD); and (*iii*) a model with separate asymmetric transition rates between all island pairs, with rates constant in time (12 parameters, model qasymm12). Many entries in the 12-parameter transition matrix were estimated as zero with a high degree of confidence, owing to the clear lack of obvious dispersal events between many island pairs (e.g., Jamaica and Puerto Rico).

We thus generated a series of secondary "dropped-zero" transition matrices by dropping all of the estimated zero entries in the 12-parameter model. Our analyses of the 12-parameter model recovered two stable solutions that differed in <0.5 log-likelihood units: (*i*) a model with two-way dispersal between Cuba and Hispaniola and one-way dispersal from Cuba to Jamaica and from Hispaniola to Puerto Rico, with all other rates equal to zero (model suffix ch/hc/cj/hp); and (*ii*) an "out of Cuba" model with one-way dispersal from Cuba to all other islands, with all other rates equal to zero (model suffix cuba). We considered a third dropped-zero model by essentially merging these two models, thus generating (*iii*) a model that allowed dispersal from Cuba to all other islands and from Hispaniola to all other islands (model suffix ch:hc).

For all of these secondary dropped-zero models, we considered the following scenarios: (*i*) a one-rate constant, time-invariant rate for all nonzero rate entries of the rate matrix (1 parameter), (*ii*) a *k*-rate model with separate time-invariant rates between all nonzero entries of the rate matrix (*k* parameters), and (*iii*) a model with a single rate that varied through time for all nonzero entries of the matrix (two parameters). All 12 of these models are illustrated in Fig. S3.

Of all models we considered, the greatest improvement in model likelihoods came from simply relaxing the assumption of time-invariant dispersal between islands (Table S1). For example, the log-likelihood of the 12-parameter time-invariant model (qasymm12) was −48.78, yet a simple 2-parameter model with equal rates between all island pairs at any point in time (qsymm1_TD) had a much higher log-likelihood (−42.68). The overall best model was the 2-parameter dropped-zero model allowing only dispersals between Cuba and Hispaniola and from Cuba to Jamaica and from Hispaniola to Puerto Rico. Results using this model as a biogeographic background model are shown in Table 1.

We recognize that our inference about diversification processes is conditional on the underlying biogeographic model used to account for transitions between character states. Ideally, one would consider all possible biogeographic models—of which there are 12, each with 10 diversification models—leading to a minimum of 120 diversification/biogeographic model combinations to be considered. Rather than consider all possible models, we fit the 10 diversification models (Table 1) against two simpler biogeographic models to verify that our results were not simply a function of choosing the qsmm.ch/hc/cj/hp_TD model (Fig. S3). These results, using the poor-fitting one-rate symmetric model (qsymm1) and the much better time-varying model qsymm1_TD provided nearly identical results to those given in Table 1, indicating that our results are not sensitive to the underlying biogeographic model used for inference. These results are summarized in Tables S2 and S3.

Ancestral State Reconstruction and Lineage Accumulation Curves. Ancestral state (island occupancy) probabilities are essentially computed during the implementation of the model described above. At each node in the phylogeny, the probabilities $D_i(t)$ for each descendant branch are combined. By normalizing these state probabilities, we obtain an estimate of the probability that a given node k was in any of the N possible character states. The probability that node k is in state i is given by

$$p_{k,i} = \frac{D_{i,k}}{\sum_{j=1}^{N} D_{j,k}}.$$

We used these node probabilities to estimate lineage accumulation curves for each island (Fig. 2). The approximate number of lineages in state i at time T is then given by

$$\Phi_{T,i} = \sum_{j=1}^{B} p_{j,i} z_j,$$

where $p_{k,i}$ is the probability of state k at node i and z is an indicator variable; and $z_i = 0$ if node i is younger than node x (e.g., $t > T$) and 1 otherwise (B is the total number of nodes).

Evaluation of Model Adequacy. To assess whether the fitted models could recover patterns of species richness and branch lengths consistent with those observed for *Anolis*, we simulated phylogenetic trees and character state data under maximum-likelihood parameter estimates for GlobalConstant, IslandConstant, GlobalVariable, and IslandVariable models. We then used three summary statistics to test whether predictions under the fitted models match patterns in the observed data. We first simply tabulated the species richness values for each island at the end of each simulation and compared these values to the observed data (Fig. 4 and Table S4).

The remaining statistics we computed assessed patterns of branch length variation in simulated trees. We are most interested in how well branch lengths associated with character state *i* match branch lengths associated with character state *i* in the *Anolis* data. Because we do not know with certainty which internal branch lengths can be assigned to which character state in *Anolis*, we considered only terminal branch lengths. We thus computed the mean terminal branch length for all terminals with character state *i* (Fig. S6), as well as the coefficient of variation in terminal branch lengths (Fig. S7). These results suggest that the best-fit model by likelihood analysis (IslandVariable) consistently fits the observed data for all three summary statistics better than the alternative models with time-invariant diversification or global declines in speciation through time.

Model Comparisons. We used Akaike weights to estimate the conditional probability of each model (5). Given the set of AIC scores corresponding to the M candidate models, we first computed the difference ($\Delta\mathrm{AIC}_i$) between each AIC score and the overall best (lowest) AIC score. These quantities are then used to estimate the probability of model i conditional on the AIC scores observed in the candidate set. This probability is

$$\Delta_i = \frac{\exp\left(-\frac{\Delta\mathrm{AIC}_i}{2}\right)}{\sum_{k=1}^{M} \exp\left(-\frac{\Delta\mathrm{AIC}_k}{2}\right)},$$

where the term in the numerator is known as the Akaike weight of the *i*th model.

1. Mahler DL, Revell LJ, Glor RE, Losos JB (2010) Ecological opportunity and the rate of morphological evolution in the diversification of Greater Antillean anoles. *Evolution* 64:2731–2745.
2. Maddison WP, Midford PE, Otto SP (2007) Estimating a binary character's effect on speciation and extinction. *Syst Biol* 56:701–710.
3. FitzJohn R, Maddison WP, Otto SP (2009) Estimating trait-dependent speciation and extinction rates from incompletely resolved phylogenies. *Syst Biol* 58:595–611.
4. Soetaert K, Petzoldt T, Setzer RW (2009) deSolve: General solvers for initial value problems of ordinary differential equations (ODE), partial differential equations (PDE) and differential algebraic equations (DAE), R package version 1.5Available at http://CRAN.R-project.org/package=deSolve.
5. Burnham KP, Anderson DR (2002) *Model Selection and Multimodel Inference—A Practical Information-Theoretic Approach* (Springer, New York).

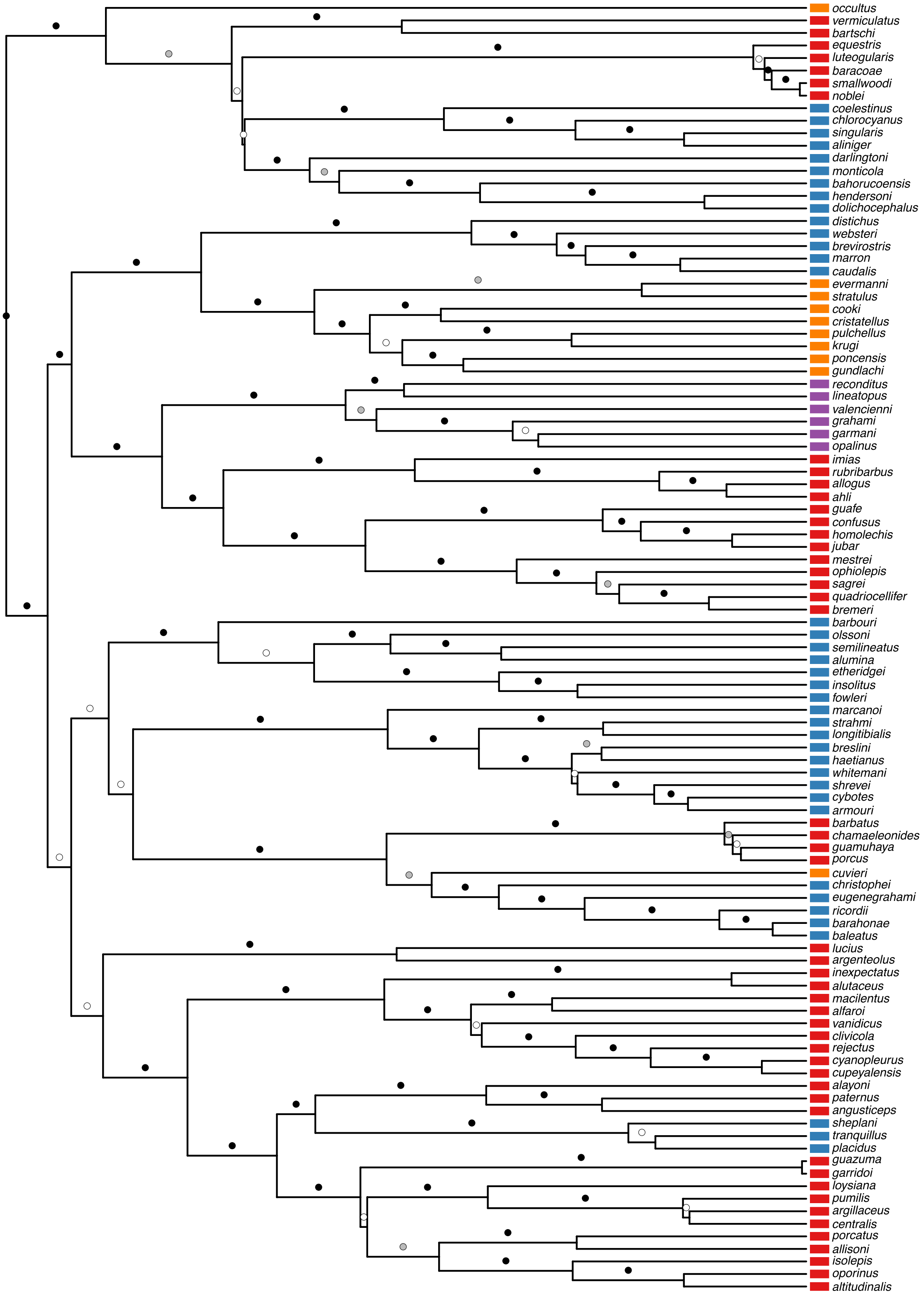

Fig. S1. Maximum clade credibility (MCC) tree resulting from BEAST analyses. Posterior probability (pp) values from the MrBayes analyses used to generate the starting tree for BEAST are indicated by circles above branches (black, $pp > 0.95$; gray, $0.95 > pp < 0.70$; white, $pp < 0.70$). Island occupancy is indicated across the tips of the tree (green, Cuba; blue, Hispaniola; yellow, Puerto Rico; orange, Jamaica).

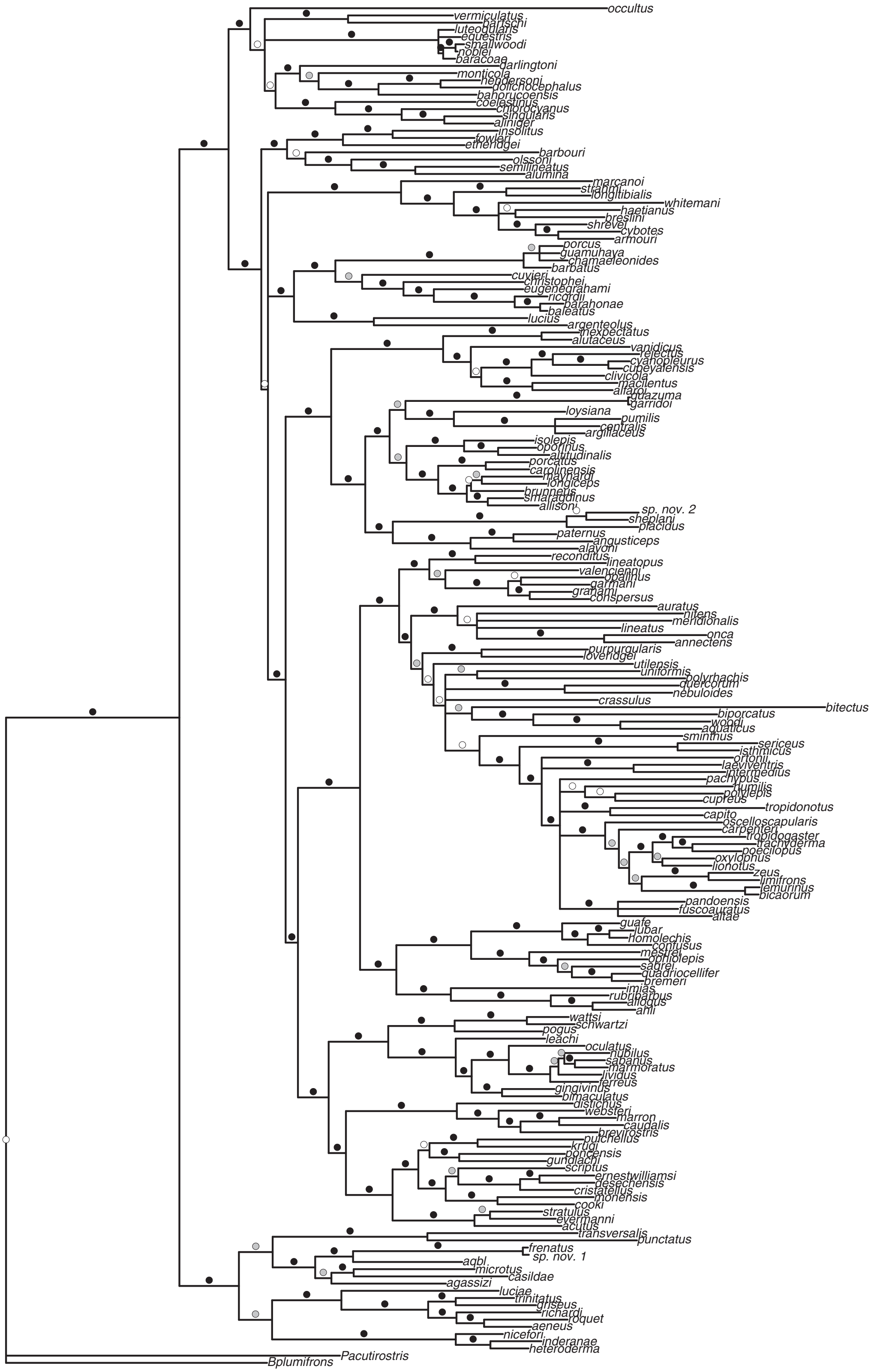

Fig. S2. Consensus tree and branch lengths generated from posterior distribution of MrBayes analysis using the sumt command. This tree includes non-Greater Antillean taxa that were pruned before our analyses. Posterior probability values are indicated by circles above branches (black, pp > 0.95; gray, 0.95 > pp < 0.70; white, pp < 0.70).

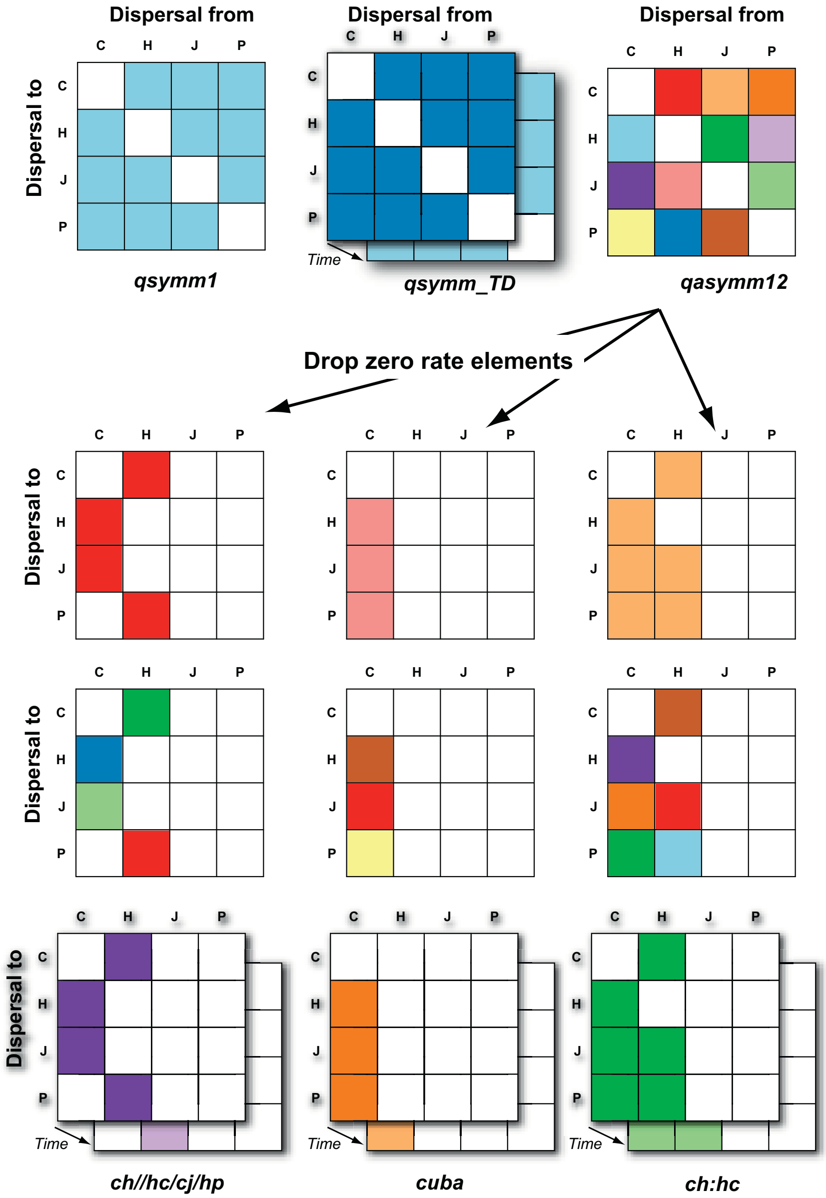

Fig. S3. Biogeographic dispersal models considered for *Anolis*. We considered 12 possible transition matrices representing dispersals between island states. Colored entries denote rate parameters that were free to vary under the model; white entries denote rate parameters that were set equal to zero. Parameters with identical colors are constrained to have the same value (e.g., qsymm1, *Upper Left*, with identical dispersal rates between all islands). The *Lower* three columns of submatrices (ch/hc/cj/hp, cuba, and ch:hc) were formed by eliminating rate parameters from the fitted qasymm12 model if rates under that model were estimated as zero ($<10^{-10}$). Models with time-varying dispersal parameters fitted the data much better than models with time-invariant dispersal (Table S1).

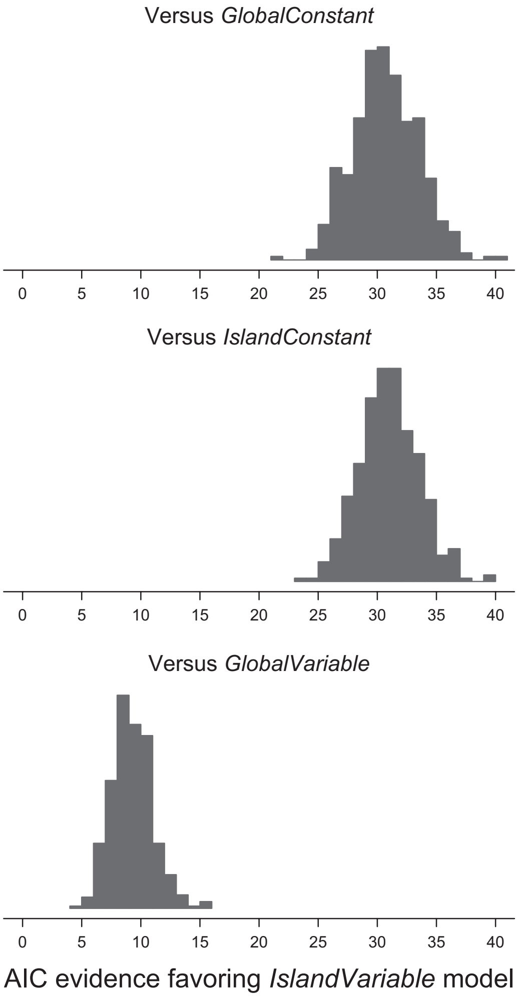

Fig. S4. Robustness of results to phylogenetic uncertainty. Histograms show AIC evidence favoring the IslandVariable model (separate linear change in speciation rates through time on each island) relative to GlobalConstant, IslandConstant, and GlobalVariable models, where all models were fitted to 400 trees sampled randomly from the posterior distributions of trees from the BEAST analysis. The IslandVariable model consistently fits better than GlobalVariable.

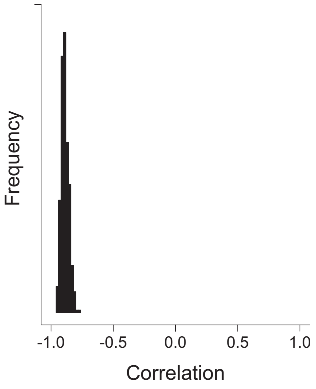

Fig. S5. Pearson correlation between log(island area) and the rate decline parameter ($-\lambda_0/K$) under the best-fit model (IslandVariable) tabulated from 400 trees sampled randomly from the posterior distributions of trees from the BEAST analysis. The rate decline parameter is the slope of the relationship between the speciation rate and time. A negative correlation implies that larger islands have slower changes in speciation rates with respect to time.

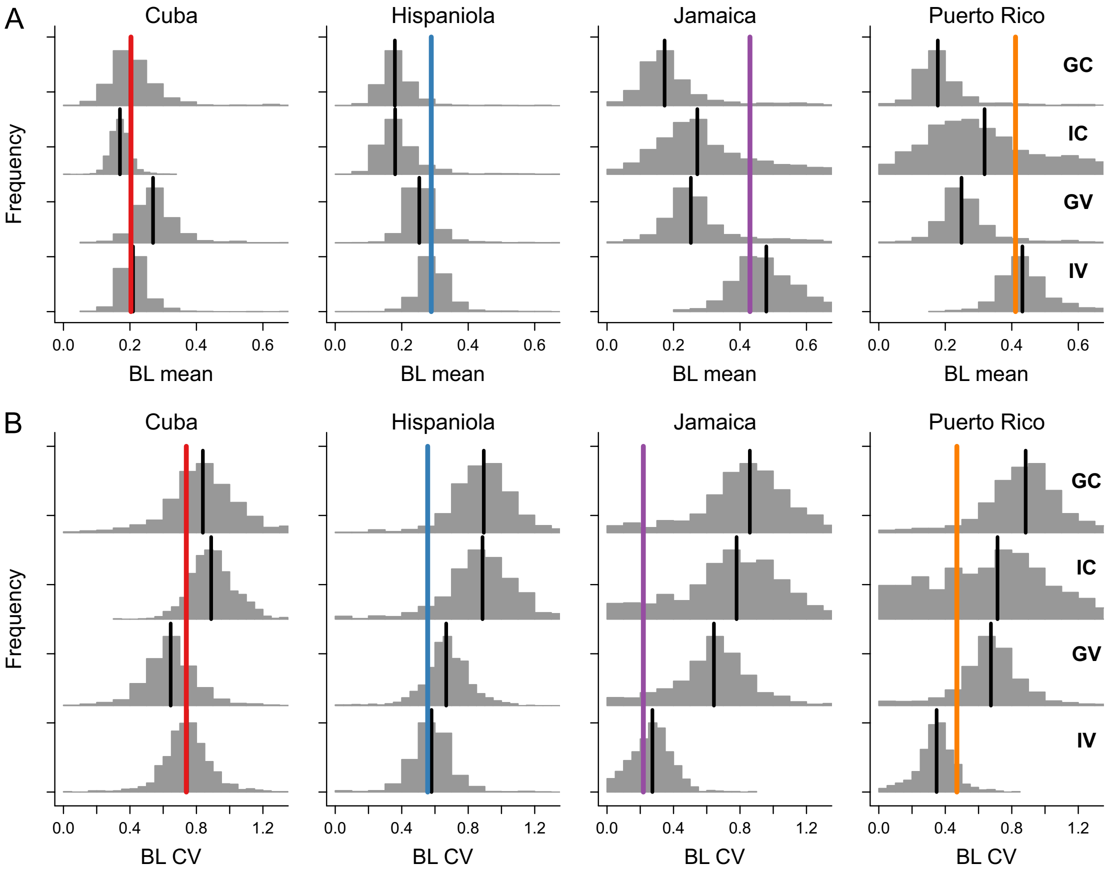

Fig. S6. Distribution of mean terminal branch lengths (*A*) and the coefficient of variation in terminal branch lengths (*B*) for trees simulated under four fitted diversification models, partitioned by terminal character state (Cuba, Hispaniola, Jamaica, or Puerto Rico). Models are (*Top* to *Bottom*) as follows: GC, GlobalConstant; IC, IslandConstant; GV, GlobalVariable; IV, IslandVariable. Colored lines denote mean values for the *Anolis* MCC tree; black lines and histograms represent mean values and distributions tabulated from 2,000 datasets simulated under maximum-likelihood parameter estimates for each model. The IslandVariable model provides a much closer match to the observed data than the three alternative models.

Table S1. Biogeographic models evaluated (see Fig. S2 for graphical description)

Model	Description	np	LogL	AIC	ΔAIC
qasymm12	Asymmetric transition rates between all island pairs.	16	−48.78	129.56	35.88
qsymm1	Single transition rate between all islands.	5	−54.9	119.8	26.12
qsymm.ch:hc	"Out of Cuba/out of Hispaniola" model. Single rate for all transitions from Cuba to all other islands and from Hispaniola to other islands (q_{12}, q_{13}, q_{14}, q_{21}, q_{23}, q_{24}); 0 for all others.	5	−53.9	117.8	24.12
qasymm.ch:hc6	Out of Cuba/out of Hispaniola model. Different rates for q_{12}, q_{13}, q_{14}, q_{21}, q_{23}, q_{24}; 0 for all others.	10	−48.78	117.56	23.88
qsymm.cuba	"Out of Cuba" model. Single rate for q_{12}, q_{13}, q_{14}; 0 for all others.	5	−52.36	114.72	21.04
qasymm.ch/hc/cj/hp.4	Different rates for q_{12}, q_{21}, q_{13}, q_{24} (between Cuba and Hispaniola and from Cuba to Jamaica and from Hispaniola to Puerto Rico); 0 for all others.	8	−48.8	113.6	19.92
qasymm.cuba3	Out of Cuba, with different rates for q_{12}, q_{13}, q_{14}; 0 for all others.	7	−49.67	113.34	19.66
qsymm.ch/hc/cj/hp	Single rate for q_{12}, q_{21}, q_{13}, q_{24}; all other rates 0.	5	−51.4	112.8	19.12
qsymm.ch:hc_TD	Single time-varying transition parameter for all transitions from Cuba and Hispaniola to all other islands.	6	−43.17	98.34	4.66
qsymm1_TD	Single time-varying transition parameter between all islands.	6	−42.68	97.36	3.68
qsymm.cuba_TD	Single time-dependent out of Cuba transition parameter (q_{12}, q_{13}, $q_{14} > 0$); all others = 0.	6	−42.67	97.34	3.66
qsymm.ch/hc/cj/hp_TD	Shared time-varying parameter for q_{12}, q_{21}, q_{13}, q_{24}.	6	−40.84	93.68	0

All biogeographic scenarios were evaluated against a background diversification model with island-specific differences in speciation, but no rate variation through time (IslandConstant model). Models with linear time-dependent change in the transition rate ("TD") fitted the data much better than corresponding models without time-dependent transition rates. States 1, 2, 3, and 4 correspond to Cuba, Hispaniola, Jamaica, and Puerto Rico, respectively.

Table S2. Model-fitting results under alternative biogeographic background models (Table S1: qsymm_TD and qsymm)

Diversification model	NP ($\mu = 0$)	LogL ($\mu = 0$)	ΔAIC ($\mu = 0$)	NP ($\mu \geq 0$)	LogL ($\mu \geq 0$)	ΔAIC ($\mu \geq 0$)	Δ_i
Biogeographic model: Symmetric dispersal between all islands, but rate varies through time (qsymm_TD)							
IslandConstant	6	−42.68	29.88	10	−42.67	37.86	0
GlobalConstant	3	−46.57	31.66	4	−46.56	33.64	0
GlobalVariable	4	−34.87	10.26	6	−34.63	13.78	0.006
IslandVariableFull	10	−25.97	4.46	18	−22.1	12.72	0.091
IslandVariable	7	−26.74	0	12	−24.3	5.12	0.903
Biogeographic model: Symmetric dispersal between all islands, but rate is constant through time (qsymm)							
IslandConstant	5	−54.91	30.04	9	−54.91	38.04	0
GlobalConstant	2	−58.22	30.66	3	−58.19	32.6	0
GlobalVariable	3	−46.58	9.38	5	−45.5	11.22	0.004
IslandVariableFull	9	−38.11	4.44	17	−27.96	0.14	0.460
IslandVariable	6	−38.89	0	11	−35.4	3.02	0.536

All diversification models are identical to those in Table 1. The biogeographic model assumes equal transition rates between all pairs of islands at any point in time, but with (*i*) a linear change in the rate of transitions through time (*Upper* row, *Center*, in Fig. S3), or (*ii*) identical time-invariant rates through time (*Upper* row, *Left*, in Fig. S3). Results are shown for models without ($\mu = 0$) and with ($\mu \geq 0$) extinction. np, number of parameters; Δ_i, conditional probability of each model given the candidate set of models. Models without extinction ($\mu = 0$) fit the data better than the corresponding model with extinction in all cases but one (IslandVariableFull under qsymm). Models specifying island-specific changes in λ and/or μ account for $P = 0.994$ of the total probability of the data taken across all models, and the GlobalVariable model fits much more poorly than the IslandVariable model.

Table S3. Summary statistics for island species richness in datasets simulated under maximum-likelihood parameters of GlobalConstant, IslandConstant, GlobalVariable, and IslandVariable models (Table 1)

Model	Island	Observed	Mean	Median	q_{25}	q_{75}
GlobalConstant	Cuba	61	6.93	4	0	10
GlobalConstant	Hispaniola	40	9.56	6	1	14
GlobalConstant	Jamaica	6	4.3	0	0	5
GlobalConstant	Puerto Rico	10	6.8	3	0	10
IslandConstant	Cuba	61	12.42	7	0	18
IslandConstant	Hispaniola	40	9.23	6	1	14
IslandConstant	Jamaica	6	2.3735	0	0	4
IslandConstant	Puerto Rico	10	2.53	1	0	4
GlobalVariable	Cuba	61	23.6	15	4	33.25
GlobalVariable	Hispaniola	40	32.17	23	7	47
GlobalVariable	Jamaica	6	13.508	4	0	17
GlobalVariable	Puerto Rico	10	24.284	13	2	35
IslandVariable	Cuba	61	47.8	30	9	69
IslandVariable	Hispaniola	40	35.95	26	10	52
IslandVariable	Jamaica	6	6.5	2	0	8
IslandVariable	Puerto Rico	10	15	9	2	21.5

Observed denotes observed species richness on each island. q_{25} and q_{75} denote 0.25 and 0.75 percentiles of the distribution of richness values under each simulation model.