www.sciencemag.org/cgi/content/full/315/5818/1574/DC1

Supporting Online Material for

The Latitudinal Gradient in Recent Speciation and Extinction Rates of Birds and Mammals

Jason T. Weir* and Dolph Schluter

*To whom correspondence should be addressed. E-mail: weir@zoology.ubc.ca

Published 16 March, *Science* **315**, 1574 (2007)
DOI: 10.1126/science.1135590

This PDF file includes:

Other Supporting Online Material for this manuscript includes the following:
(available at www.sciencemag.org/cgi/content/full/315/5818/1574/DC1)

Materials and Methods

Data

New World sister species pairs were chosen from complete or nearly complete published molecular phylogenies for terrestrial bird and mammal taxa and are listed along with their midpoint latitudes and estimated ages in Database S1. In addition, a few sister species pairs were obtained from phylogenies we generated from Genbank sequences. Occasionally, a species may not be monophyletic and may contain a daughter species nestled within it. In such cases, we used the age at which the daughter split from its parent species. Sister species pairs in which one or both species were endemic to oceanic islands were excluded, with the exception of continental shelf islands that were connected to the continent during Pleistocene periods of low sea level. At arctic latitudes, some pairs of sister species are circumpolar in distribution and were included as long as both members reside in the New World. Marine mammal and marine and aquatic bird families were excluded.

In this paper the speciation process is defined as beginning at population splitting and is completed when reproductive isolation evolves. Following Avise et al (14, 15), we used the maximum age of mitochondrial haplotypes within New World species to provide a lower bound estimate of the duration of the speciation process (lag-time to speciation). Avise et al primarily used species that were geographically segregated into monophyletic groups of mitochondrial haplotypes termed phylogroups. Using only species possessing phylogroups however may overestimate the duration of the speciation process because it represents a non-random sample of species. At tropical latitudes a large proportion of

species analyzed in this study possessed phylogroups (see Database S1). However, at high latitudes, many species are very young and have not had time to form phylogroups. At high latitudes the mean ages of phylogroups in our dataset was older than the mean age of sister species and thus do not accurately estimate the duration of the speciation process. To provide better estimates, we measured maximum haplotype divergence within widespread species possessing multiple morphologically differentiated subspecies that may or may not have diagnosed phylogroups but nevertheless do possess haplotype variation. When phylogroups were lacking, we used the maximum sequence divergence (GTR-gamma distance) between haplotypes and a molecular clock to obtain estimates of the dates when extant haplotypes first began to diverge. We present both phylogroup (Fig. 1c) and haplotype (including both species with and without phylogroups; Fig. 1b) datasets across the latitudinal gradient. Estimates of haplotype variation within species were generated primarily from molecular datasets associated with published phylogeographic studies. We sequenced additional sisterspecies and phylogroups (Supplementary Table 1).

Approximate dates of splitting were estimated for sister species and intraspecific haplotype splits from GTR-gamma distances obtained from the mitochondrial cytochrome b DNA sequences. A few very young pairs of sister species were not reciprocally monophyletic for cytochrome b haplotypes because haplotypes have not had enough time to become reciprocally monophyletic following population splitting. In such cases, we used the average divergence between species as a rough approximation of their ages. However, we did not estimate maximum haplotype ages for species that were not monophyletic. Sequences were obtained from Genbank or were sequenced

(Supplementary Table 1) using standard protocols (22) with primers O-H16065 (5'-AGTCTTCAATCTTTGGCTTACAAGAC-3') and O-L14851 (5'-CCTACCTAGGATCATTCGCCCT-3'), which we developed specifically for oscine passerines, or S-L14987 (5'- CCATCAAACATYTCAGCYTGATG -3') for suboscine passerines and non-passerines.

PAUP 4.0b10 (S1) was used to generate GTR-gamma distances. Model parameters were estimated separately by PAUP (S1) for the bird and mammal dataset using maximum likelihood. Though some rate variation occurs, an average molecular clock of 2.0% sequence divergence per million years (1% per lineage) holds true across a large sample of avian orders (12, S2-S4) (calibrated in reference 12 for cytochrome b using GTR-gamma distances). We used this clock for all avian dates. We remeasured the rate of evolution using GTR-gamma distances for 21 published mammalian clocks representing seven of the eight mammalian orders included in this study (Supplementary Table 2). These rates varied between 2% and 4% for all orders tested except Rodentia and Lagomorpha, for which calibrated rates ranged as high as 8%. We used the average calibrated rate for each order to convert GTR-gamma distances into time estimates. Because no calibration was available for the order Erinaceomorpha, we used the rate obtained for the closely related order Soricomorpha to date the single sister species pair of Erinaceomorphid included in this study.

It should be noted that sequences may begin to diverge before the actual splitting events if ancestral populations possessed sequence polymorphisms at the time of splitting. However, the discrepancy between coalescent times estimated from cytochrome b sequences and splitting times average only 2 to 3 hundred thousand years for temperate

and tropical bird taxa (12, S5), and as such we treat coalescent dates as close approximations to the actual dates of splitting.

Midpoint latitude for each sister species pair was calculated as follows:

$$\text{midpoint latitude} = (|A| + |B|)/2 \qquad \text{Equation 3}$$

where A and B are midpoint latitudes for each member of the sister species pair and | | indicates absolute value. Likewise, midpoint latitudes were obtained for each species in the lag-time dataset. Latitudinal ranges were obtained from digitized range maps for all New World birds and mammals (S6) with a few exceptions as noted in Database S1. Latitudinal limits are well known for most New World birds. Greater uncertainty exists for some mammal groups. Additional information on range limits may alter midpoint latitudes presented in Database 1, but should not affect our results greatly.

Estimation

To estimate speciation and extinction rates, probability distributions of the ages of sister species were generated from simulations of phylogenetic trees under a stochastic birth-death process in which speciation and extinction rates are constant through time (17, S7). At any point in time in a birth death process, the waiting time to the next speciation event (t_λ) follows an exponential distribution with mean equal to the inverse of the speciation rate (λ) multiplied by the total number of lineages extant in the tree. Likewise, the waiting time to the next extinction event (t_μ) follows an exponential distribution with mean equal to the inverse of the extinction rate (μ) multiplied by the total number of lineages extant in the tree.

Phylogenetic trees were simulated in R (code submitted as package PhySim 1.0 (S8)) for 10 time units (t) to represent 10 million years. Starting with a single lineage at t = 0, the next event in the tree can be either a speciation or extinction event. Waiting times to the next speciation and extinction event were drawn randomly from distributions of waiting times as described above. If the waiting time to speciation was shorter than to extinction a speciation event (bifurcation) was added to a randomly chosen lineage in the tree at time = $t + t_{\lambda}$. If the waiting time to extinction was shortest, then an extinction event was added to a randomly chosen lineages at time = $t + t_{\mu}$. This process was repeated until either t = 10 or the entire lineage went extinct leaving no descendents at the present. In order to produce distributions of sister species ages for a given speciation and extinction rate, between 1000 and 15000 such simulations were performed.

Probability distributions were simulated for 18 different values of birth (λ; 0.05, 0.1 0.15, 0.2…0.9) where units are the number of new lineages per lineage per million years. For birth rates less than 0.5, 12 death rates (μ) were simulated (0λ, 0.1λ, 0.2λ, …0.8λ, 0.9λ, 0.95λ, 0.99λ). For birth rates greater than 0.5, the same death rates were used as long as λ-μ < 0.5. This restriction was necessary because phylogenetic trees became very large and computationally expensive when the net diversification rates exceeded 0.5.

The lag-time to speciation was modeled as having an exponential probability distribution with rate inversely proportional to the average lag-time. We used an exponential distribution because distributions of lag-times were highly skewed with most lag-times young in age. Beginning at the root of a simulated tree and moving to the tips, each node was classified as either a “species level” or “intraspecific split”. This was done

by a drawing a lag-time from the corresponding exponential distribution. If the node age was greater than the lag-time it was classified as a species level node and was retained. Otherwise, the node was classified as an "intraspecific split" and all descendants were pruned from the tree. The resulting pruned trees contained only "species level" nodes. Sister species ages were then extracted from these trees to generate probability distributions of sister species ages.

Maximum haplotype divergence within bird and mammal species suggests that the average lag-time ranges between close to 0 (at high latitudes) and 2 million years (near the equator). Thus, for each set of trees simulated under different combinations of birth and death rates, 21 different probability distributions of sister species were extracted, each with a different mean lag-time (0, 0.1, 0.2…1.9, 2.0). In total, 3927 simulated sister species distributions were obtained.

For each set of simulated sister species distributions, the probability density function was obtained using the locfit package in R (S9). The probability of drawing a sister species with age t from a simulated sister species age distribution equals the probability density at time t in the simulated distribution. For a given set of values for b_λ, b_μ, c_λ, and $\underline{c_\mu}$, the appropriate simulated distribution for a sister species with latitude L was determined by solving for λ and μ in Equations 1 and 2. For each value of b_λ, b_μ, c_λ, and c_μ, the likelihood was obtained by multiplying the probabilities of each sister species. The values of b_λ, b_μ, c_λ, and c_μ with the highest likelihood are the maximum likelihood estimates. The likelihood support intervals (equivalent to the 95% confidence interval) includes all parameter combinations within 2 log likelihood units of the maximum likelihood estimate.

Bird and mammal sister species datasets had similar distributions across the latitudinal gradient and a linear regression resulted in almost identical slopes and y-intercepts (Figure 1a). Due to their similarity, these datasets were pooled when estimating diversification rates. However, lag-time datasets (Fig. 1b and c) were not pooled because the relationship between latitude and age of intraspecific splits was less steep in mammals than in birds. As such, separate corrections for the waiting time to species recognition were applied to bird and mammal sister species.

These diversification rate estimates obtained from many independent data points (i.e. sister species) account for the fact that many lineages alive 10 million years ago went extinct leaving no descendents at the present. As a result, we are able to estimate fauna-wide extinction rates despite the fact that we posses information only for extant species.

Phylogenetic Signature in Sister Species Distributions

The shape of sister species age distributions contains the phylogenetic signature of speciation and extinction. In a pure birth model (no extinction), simulated distributions of sister species ages appear exponentially distributed (Fig. S1a). The means of these distributions are inversely proportional to the speciation rate. At low speciation rates, age distributions have relatively large means, and the tails are spread over broad time intervals. In contrast, when speciation rates are high, the mean of the distributions are low and the tails are narrow. As extinction rate increases, sister species age distributions depart more strongly from an exponential distribution (Fig. S1b) with more species dating in the tails and heads and fewer near the mean. This results in every combination of birth and death rates producing a distribution with a unique mean and shape, where the

phylogenetic signals of the speciation and extinction rates are contained in the mean and shape of the distribution respectively. Applying a lag time correction to sister species age distributions simulated under a pure birth model creates a mode in the distribution near the mean lag time (Fig S1c).

Supplemental References

S1. D. L. Swofford, PAUP*4.0b10: phylogenetic analysis using parsimony (*and other methods) (Sinauer, Sunderland, MA, 2002).

S2 R. C. Flischer, C. E. McIntosh, C. L. Tarr, Molecular Ecology **7**, 533 (1998)

S3 J. Klicka, R. M. Zink, Science **277**, 1666 (1997)

S4 J. Garcia-Moreno, J. Avian Biology **35**, 465 (2004).

S5. S. V. Edwards, P. Beerli, Evolution **54**, 1839 (2004).

S6. R. S. Ridgely, et al., Digital Distribution Maps of the Birds of the Western Hemisphere, version 2.1. NatureServe, Arlington, Virginia, USA (2005).

S7. S. Nee, R. M. May, Phil. Trans. R. Soc. Lond. B **344**, 305 (1994)

S8. J. T. Weir, D. Schluter. PhySim: Phylogenetic Tree Simulation Package. R package version 1.0. http://CRAN.R-project.org/ (2007).

S9. Locfit. http://www.locfit.info/ (2005).

S10. C. Lalueza-Fox, et al., BMC Evol. Biol. **5**, 70 (2005).

S11. R. K. Wayne, E. Geffen, D. J. Girman, K. P. Koepfli, L. M. Lau, C. R. Marshall, Syst. Bio. **46**, 622 (1997).

S12. B. Stadelmann, L. K. Lin, T. H. Kunz, M. Ruedi, Mol. Phylogenet. Evol. in press.

S13. C. Steiner, M. Tilak, E. J. P. Douzery, F. M. Catzeflis, Mol. Phylogenet. Evol. **35**, **363** (2005).

S14. N. Yu, C. Zheng, Y. Zhang, W. Li, Mol. Phylogenet. Evol. **16**, 85 (2000).

S15. B. R. Riddle, D. J. Hafner, L. F. Alexander, Mol. Phylogenet. Evol. **17**, 161 (2000).

S16. T. S. DeWalt, P. D. Sudman, M. S. Hafner, S. K. Davis, Mol. Phylogenet. Evol. **2**, 193 (1993).

S17. S. J. Steppan, R. M. Adkins, J. Anderson Syst. Biol. **53**, 533 (2004).

S18. C. J. Conroy, J. A. Cook, J. of Mam. **81**, 344 (2000).

S19. R. G. Harrison, S. M. Bogdanowicz, R. S. Hoffmann, E. Yensen, P. W. Sherman, J. Mam. Evol. **10**, 249 (2003).

S20. S. V. Brant, G. Orti, Mol. Phylogenet. Evol. **22**, 163 (2002).

S21. L. Fumagalli, et al, Mol. Phylogenet. Evol. **11**, 222 (1999).

Supplemental Table 1

Species	**Location**	**Museum**	**Tissue No.**	**Accession No.**
Campylorhamphus pusillus	PANAMA, Bocas del Toro, Continental Divide	STRI	JTW094	EF202815
Corvus caurinus	CANADA, British Columbia, Vancouver	CO	none	EF210778
Dendrocolaptes sanctithomae	PANAMA, Bocas del Toro, Cerro Chalite	STRI	JTW251	EF212895
Dendrocolaptes sanctithomae	PANAMA, Bocas del Toro, Cerro Chalite	STRI	IJL079	EF212896
Heterospingus rubrifrons	PANAMA, Bocas del Toro, Cerro Chalite	STRI	JTW278	EF202820
Manacus vitellinus	PANAMA, Panama Province, Achiote Road	STRI	PA-MVI-PC16	EF202819
Pseudocolaptes lawrencii	COSTA RICA, San Jose	LMN	B19934	EF202814
Tityra semifasciata	PANAMA, Bocas del Toro, Cerro Chalite	STRI	JTW298	EF212894
Trogon viridis	PANAMA, Panama Province, confluence of Rio Charges and Rio Chagrecito	STRI	PA-TVI2034	EF202818

Supplemental Table 2

ORDER (Family)	*Calibration Taxon*	Date of Split	GTR-gamma distance	Rate (%)	Calibration Reference
ARTIODACTYLA					
Bovidae	split *Myotragus balearicus* and *Ovis*	5.4	0.1757	3.28	*S10*
CARNIVORA					
Canidae	earliest split within *Vulpes*	9.5	0.1921	2.02	*S11*
	split *Canis* and *Lycaon*	6.7	0.1716	2.56	*S11*
	split *Canis latrans / C. lupus* and *C. simensis*	3.5	0.0823	2.35	*S11*
CHIROPTERA					
Vespertilionidae	split *Myotis nattereri* and *M. schaubi*	6	0.2209	3.68	*S12*
	split *Myotis daubentonii* from *M. bechsteinii*	5	0.1929	3.86	*S12*
DIDELPHIMORPHIA					
Didelphidae	split *Micoureus* and *Marmosa murina / M. lepida*	14.1	0.2884	2.05	*S13*
	split *Didelphis* and *Philander*	5.9	0.1758	2.98	*S13*
LAGOMORPHA					
Ochotonidae	first split within extant species of *Ochotona*	5.5	0.2769	5.04	*S14*
PRIMATE					
Hominidae	split *Homo* and *Pan*	5.4	0.1490	2.76	
RONDENTIA					
Geomyidae	split *Perognathus / Chaetodipus* and *Dipodomys / Microdipodops*	16.5	0.5800	3.52	*S15*
	split *Geomys* and *Cratogeomys / Pappogeomys*	6	0.3447	5.75	*S16*
	split *Pappogeomys* and *Cratogeomys*	4	0.2546	6.36	*S16*
	split Thomomyini and Geomyini	5.6	0.4399	7.93	*S16*
Muridae	split *Batomys* and all other Murine genera	12	0.3492	2.91	*S17*
	split *Microtus californicus* and *M. mexicanus*	2.1	0.1532	7.30	*S18*
Sciuridae	split *Marmota* and sister clade of *Spermophilus / Cynomys*	7.7	0.2038	2.65	*S19*
	split *Cynomys* and sister clade of *Spermophilus*	2.7	0.1504	5.57	*S19*
	First split within *Spermophilus*, *Marmota* and *Cynomys*	16.5	0.2680	1.62	*S19*
SORICOMORPHA					
Soricidae	split *Cryptotis* and *Blarina*	9	0.2354	2.62	*S20*
	split Crocidurinae and Soricinae	20	0.5125	2.56	*S21*

Supplemental Table 1

Cytochrome b sequences submitted to Genbank. Museums: Field Museum of Natural History (FM), Cowan Vertebrate Museum (CO), Smithsonian Tropical Research Institute (Naos tissue collection; STRI), Louisiana Museum of Natural History (LMN).

Supplemental Table 2

Molecular clock calibrations for Mammalian orders.

Supplemental Figure 1

Simulated probability distributions of sister species ages under different rates of speciation and extinction. A) Speciation rate equals 0.4 (red) and 0.1 (blue) lineages / lineage / million years, no extinction and no correction for lag time to speciation. B) the effect of adding extinction. Speciation rate equals 0.4 and death rate equals 0 (red) and 0.38 (blue). C) speciation and extinction rates (rates same as in b) with a mean lag time of 2.0 Ma.

Figure S1