

## FAST TRACK

# Calibrating the avian molecular clock

J. T. WEIR\*† and D. SCHLUTER†

*\*Department of Ecology and Evolution, University of Chicago, Chicago, IL 60637, USA, †Department of Zoology, and Center for Biodiversity Research, University of British Columbia, Vancouver, British Columbia, Canada V6T 1Z4*

## Abstract

**Molecular clocks are widely used to date phylogenetic events, yet evidence supporting the rate constancy of molecular clocks through time and across taxonomic lineages is weak. Here, we present 90 candidate avian clock calibrations obtained from fossils and biogeographical events. Cross-validation techniques were used to identify and discard 16 inconsistent calibration points. Molecular evolution occurred in an approximately clock-like manner through time for the remaining 74 calibrations of the mitochondrial gene, cytochrome *b*. A molecular rate of approximately 2.1% ( $\pm 0.1\%$ , 95% confidence interval) was maintained over a 12-million-year interval and across most of 12 taxonomic orders. Minor but significant variance in rates occurred across lineages but was not explained by differences in generation time, body size or latitudinal distribution as previously suggested.**

*Keywords:* birds, cytochrome *b*, mitochondrial DNA, molecular clock, molecular evolution, phylogenetics

*Received 9 November 2007; revision accepted 18 February 2008*

## Introduction

The molecular clock is one of the most widely used tools to investigate the timing of phylogenetic events. In the absence of other methods of dating, molecular clocks have been used not only to date the origin of taxonomic groups but also to determine the impacts of climatic and geological events on diversification (Cooper & Penny 1997; Weir & Schluter 2004; Weir 2006), estimate rates of speciation and extinction (Baldwin & Sanderson 1998; Barraclough & Vogler 2002; Weir & Schluter 2007), determine the timing of dispersal events (Drovetski 2003; Mercer & Roth 2003) and to date the origin of gene families (Robinson-Rechavi *et al.* 2004; Vandepoele *et al.* 2004). However, the premise that molecular evolution occurs at a steady rate through time and across taxonomic lineages is controversial (Garcia-Moreno 2004; Lovette 2004; Ho *et al.* 2005; Penny 2005). Currently, only a small number of calibrations have been obtained (e.g. summarized in Garcia-Moreno 2004; Lovette 2004) and are insufficient to test the general validity of the clock. Here, we generate 78 new calibrations of the avian molecular clock for a single mitochondrial gene, cytochrome

*b*. These calibrations were obtained from dated fossils and the ages of oceanic islands, mountain ranges and land bridges. These were combined with re-evaluation of 12 previously published rates to give a total of 90 clock calibrations spanning 12 taxonomic orders and 12 million years (Myr). We used this data set to test for rate constancy through time and across avian taxonomic orders.

## Materials and methods

### *Calibration sources*

Calibration points were obtained from fossils and biogeographical events. For a justification of each calibration, see Table S1 and Appendix S1, Supplementary material. The oldest calibration point used was 12 Myr. While potential new and pre-existing calibration points older than 12 Myr exist (e.g. Galliformes, Pereira & Baker 2005), these were excluded because cytochrome *b* sequences become saturated, and estimates of sequence divergence are likely to be underestimated for older calibration points, even after correcting for homoplasy using detailed models of sequence evolution. As such, we test for rate constancy only during the past 12 Myr, and suggest that slow-evolving genes be used for older calibration points. Likewise, additional avian calibrations have been obtained using sequence data

Correspondence: Jason T. Weir, Department of Ecology and Evolution, University of Chicago, 1101 E 57th Street, Chicago, IL 60637, USA. Fax: 604-822-2416; E-mail: jtweir@uchicago.edu

for other genes (e.g. Arbogast *et al.* 2006) but cytochrome *b* sequence data were not available for these.

Fossils were used to date the node connecting the taxon represented by the fossil to its sister taxon. Fossils identified to species were used to calibrate the split between the species and its sister group. Likewise, the oldest fossils within a genus or family were used to calibrate the split between that genus or family and its sister group. If the fossil record is incomplete, the oldest known fossil may postdate the origin of a clade, resulting in an overestimation of the molecular rate. Overestimation of molecular rates is more likely to affect young rather than old calibration points. For example, if all fossil-based calibration points postdated the clades they represent by a given amount of time, then calibrated rates for old fossil calibration points would be little affected but recent calibration points would be greatly affected. To minimize such biases, we only used fossil-based calibration points more than 2 Myr old. The 2-Myr cut-off point allowed us to exclude the host of Holocene and mid- to late-Pleistocene fossils, which represent the only fossils known for many groups having poor fossil records. We used the avian fossil record, drawn primarily from North American sources (Becker 1987; Emslie 1998), to obtain potential fossil-based calibration points dating between 2 and 12 Ma. All such calibration points were included provided that the cytochrome *b* sequence data were available for the group. Our sampling of the Old World fossil record was incomplete and additional fossil calibrations may be available for Old World forms.

Three categories of biogeographical events were used to provide calibration points: formation of islands, mountains and land bridges. Because colonization of islands, mountain ranges, and permanent land bridges may occur anytime after their initial formation (e.g. Fleischer *et al.* 1998), dates of biogeographical events might predate the actual date of colonization and thus result in an underestimation of molecular rates. Underestimation of molecular rates is more likely to affect young rather than old calibration points. To minimize this bias, we only included new biogeographical calibration points older than 1 Myr; however, we retained in the data set previously published calibration points younger than 1 Myr.

For volcanic islands, we used the ages of the oldest exposed lavas to estimate when the islands emerged above sea level. Colonization may not be possible during the earliest stages of island formation, adding uncertainty to the calibrated date. For several islands that experienced extended periods of volcanic activity (i.e. Marquesas Islands), we used both the dates of island emergence and the dates when volcanic activity ceased, and then we used the cross-validation method developed below to determine which dates were most consistent with the ages of other calibration points in the data set. In addition, several continental shelf islands were used as calibration points. In the Mediterranean, islands (e.g. Cyprus and Balearic islands) reformed at

5.33 ( $\pm$  0.02) million years ago (Ma) at the end of Messinian salinity crisis (Krijgsman *et al.* 1999), which had left the Mediterranean Sea dry. This date has previously been used as a calibration point for several Mediterranean island endemics (Bohning-Gaese *et al.* 2006).

Calibrations were obtained for avian taxa that are believed to have used land bridges to disperse between continents. The Central American land bridge connecting Central America to Colombia was completed approximately 3.5 Ma (Keigwin 1978; Coates *et al.* 1992, 2004; Coates & Obando 1996; Knowlton & Weigt 1998) and is known to have promoted interchange in mammals (Stehli & Webb 1985) and birds (J. T. Weir, E. Bermingham & D. Schluter, in prep.). We used the phylogenetic reconstructions by Weir *et al.* (submitted) for tanagers (Thraupidae), blackbirds (Icteridae) and woodcreepers (Dendrocolaptidae) as well as phylogenies for *Myadestes* (Miller *et al.* 2007) and *Myioborus* (Perez-Eman 2005) to identify monophyletic sister lineages in which one or more species was endemic to one side of the land bridge (i.e. Central or South America) and the other lineage was endemic to the opposite side. A land bridge connecting the Malay Peninsula with mainland Asia was bisected by a seaway along the Isthmus of Kra approximately 5 Ma (Woodruff 2003) and has previously been used to calibrated molecular rates in woodpeckers (Fuchs *et al.* 2006).

Uplift of highland regions is often extended over long time periods, precluding their use as calibration points. However, we included the uplift of the Talamanca highlands of Costa Rica and western Panama in our study because they experienced rapid uplift approximately 4.5 Ma, when the Cocos Plate was subducted beneath the Caribbean Plate (Abratis & Worner 2001; Grafe *et al.* 2002). In this case, the date when uplift first began is probably only slightly older than the date when colonization by highland species would have been possible. As the Central American land bridge had not yet connected the Talamanca to South America, we used the age of these highlands to date only those endemic highland species with northern ancestors (five such pairs were identified from published phylogenies). Our analysis included informative fossil and biogeographical calibrations regardless of whether their calibrated rates agreed with the traditional 2% clock.

### Phylogenetic analysis

Molecular clock calibrations were obtained exclusively from the mitochondrial cytochrome *b* gene, eliminating possible sources of rate variation among genes. Cytochrome *b* sequences were obtained from GenBank (major sources in Table S1). Sequences of multiple individuals were utilized when possible, provided sequences were of similar length within each calibration.

Multiple mutations at a single site ('saturation') confound genetic distance estimates (Arbogast *et al.* 2002). Some

previously published clock calibrations underestimated the degree of mutation saturation by using overly simplistic models of sequence evolution to generate genetic distances (Arbogast *et al.* 2002). We corrected for saturation using the six-parameter general time reversible (GTR) model (Rodriguez *et al.* 1990). We used the GTR model because it provides the most accurate available model of sequence evolution. The GTR model, or a similar model, is the one used predominantly by researchers estimating phylogenetic trees with molecular data. A gamma correction ( $\Gamma$ ) with four rate categories was applied to correct for among-site rate variation (Yang 1993). PAUP\* version 4.0b10 (Swofford 2002) was used to generate corrected distances. For each calibration, sequences from closely related species (congeneric or con-familiar) were used to generate neighbour-joining trees. These trees were used simply to obtain maximum-likelihood estimates of the GTR- $\Gamma$  model parameters in PAUP. This allowed separate model parameters to be estimated for each calibration point.

Sequence divergence ( $P$ ) measures the point in time when two sequences coalesce. This will predate lineage splitting if ancestral populations possessed DNA sequence polymorphisms (Edwards & Beerli 2000). We corrected for ancestral polymorphism using a standard protocol (Nei & Li 1979; Avise *et al.* 1998) as follows:

$$P_{\text{cor}} = P - \frac{1}{2}(P_A + P_B) \quad (1)$$

where  $P_{\text{cor}}$  is the corrected sequence divergence (GTR- $\Gamma$  distance) and  $P_A$  and  $P_B$  are the average sequence divergences within current populations of each of the daughter lineages. When multiple sequences were not available for one or both daughter lineages, then the correction was made with a single lineage or with the average corrections obtained from other calibration points (0.37%). Molecular rates were determined by dividing  $P_{\text{cor}}$  by the age of the calibration point. The molecular rate for the entire data set of 74 calibrations was estimated by the slope of a regression of  $P_{\text{cor}}$  vs. time with  $y$ -intercept forced through the origin.

*Detection of incompatible calibration points.* Variation in molecular rates may reflect genuine rate variation in different taxonomic lineages or may be generated by erroneous or imprecise fossil or biogeographical calibrations. Fossil or biogeographical events may be placed in an incorrect phylogenetic context or they may greatly pre-date (biogeographical events) or postdate (fossils) the nodes they represent. A method that discriminates between sources of rate variation is necessary to eliminate faulty calibration points while retaining genuine rate variation across lineages. The method we used is a simplified version of the cross-validation procedure developed by Near & Sanderson (2004). In their method, a phylogenetic tree is generated with branch lengths proportional to time and that includes each of the

calibration points being tested. The tree is calibrated using a single calibration point and the predicted ages of the remaining calibration points are determined and compared to their fossil or biogeographical age estimates. The calibration process is repeated for all calibration points in the data set. Calibration points that poorly predict the ages of other calibration points are discarded and those that provide accurate predictions are retained. The cross-validation test is a powerful technique for detecting and deleting erroneous calibration points (Near & Sanderson 2004).

To implement this method, we used BEAST version 1.4.6 (Drummond & Rambaut 2006) to generate both a global clock tree and a relaxed-clock tree (with rate variation following a log-normal distribution) that included all the taxa involved in the clock calibrations. Tree topology was determined from published phylogenies (relationships between orders and passerines families were determined from the nuclear DNA phylogenies of Ericson *et al.* 2006 and Barker *et al.* 2004, respectively, and relationships within families were determined from a series of family level phylogenetic studies) and was fixed in the BEAST analysis so that only branch lengths were estimated using cytochrome *b* sequence data and the GTR- $\Gamma$  model of sequence evolution. Analyses were run for 30 million generations for the global clock tree and 16 million generations for the relaxed clock tree. The first ten or five million generations were excluded as a burn-in for the global and relaxed-clock trees, respectively. Following the burn-in, trees were sampled every 5000 generations and were used to construct mean branch lengths.

Using both the global clock and relaxed-clock phylogenies, we then carried out the following cross-validation steps for each individual fossil or biogeographical date  $\chi$ . First, we used the fixed date for  $\chi$  to calculate molecular estimates of divergence dates for all other nodes in the phylogeny that are associated with calibration points. At each of these other nodes  $i$ , the discrepancy between the molecular estimate of the date for  $i$  ( $MA_i$ ) and the fossil or biogeographical date of  $i$  ( $FA_i$ ) was calculated as

$$D_i = \frac{MA_i - FA_i}{\sqrt{0.5(MA_i + FA_i)}} \quad (2)$$

This measure of discrepancy scales the difference between molecular and fossil age estimates to the square root of their average. Near & Sanderson (2004) used the alternative formula  $D_i = MA_i - FA_i$  which is appropriate if calibration points are approximately equal in age, but doesn't take account of the expectation that the absolute difference between  $MA_i$  and  $FA_i$  should increase with older dates.

Second, all the  $D_i$  values corresponding to fossil or biogeographical calibration  $\chi$  were squared and summed to yield

$$SS_\chi = \sum_{i \neq \chi} D_i^2 \quad (3)$$

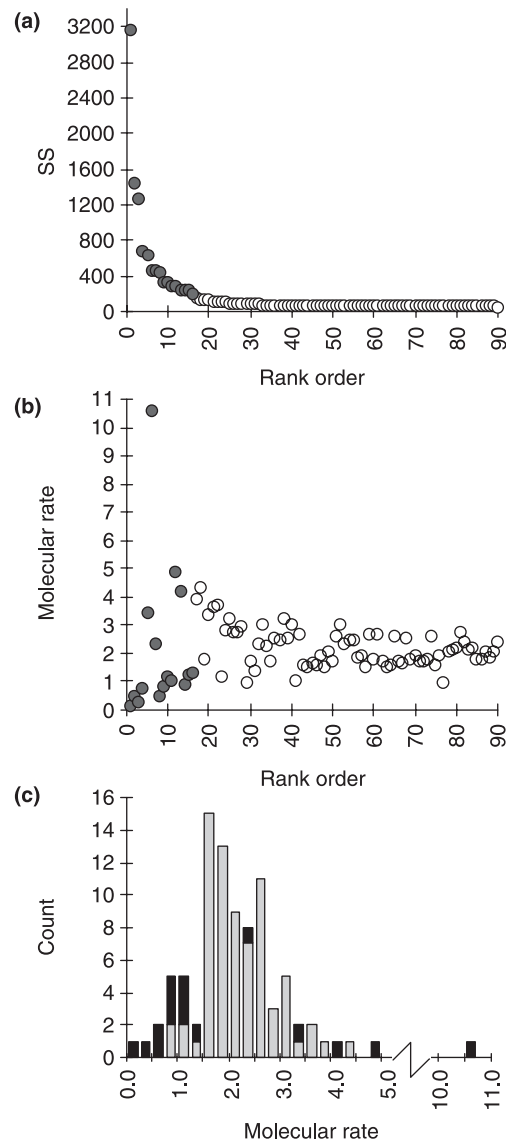
These steps were repeated for every calibration  $\chi$ . Calibrations with low values of  $SS$  are concordant with other molecular and fossil/biogeographical date estimates, whereas calibration nodes with high values of  $SS$  indicate a discrepancy between the given calibration and molecular and other fossil and biogeographical date estimates.

The cross-validation procedure involves eliminating calibrations which poorly predict other calibration dates in the data set – those with high values of  $SS$ . Near & Sanderson (2004) used an additional step that involved removing the most inconsistent fossil (with the highest  $SS$  value) and calculating the mean squared deviation of  $D_i$  values for all remaining calibrations. This step was repeated sequentially until further removals failed to result in a significant decline in the mean squared deviation of  $D_i$  (for full details see Near & Sanderson 2004). Instead, we favour using the raw  $SS$  values and excluded this additional step (although results were similar when this step was included). Because  $SS$  values are not normally distributed (Fig. 1a), we used a general criterion based on the box plot (Tukey 1977) to determine which  $SS$  values to exclude. First, we calculated the first and third quartiles and interquartile range of  $SS$  values. Then  $SS$  values were classified as mild or extreme outliers if they exceeded the third quartile by more than 1.5 or 3 times the magnitude of the interquartile range, respectively.

The clock cross-validation method was implemented in R (package *CCVAL* 1.0) and is available upon request from the lead author.

**Rate variation across lineages.** We tested for rate variation across lineages in the Bayesian uncorrelated log-normal relaxed clock phylogeny generated above for the cross-validation procedure. For each of the 2000 sampled Bayesian phylogenies, *BEAST* version 1.4.6 (Drummond *et al.* 2007) provides the log of the standard deviation in rates (*uclid.stdev* statistic). When no rate variation across lineages is observed, the *uclid.stdev* statistic is zero and conforms perfectly to a global clock model. Values equal to one indicate that the standard deviation in rate is equal to the mean rate while values greater than one demonstrate excessive rate heterogeneity across lineages.

Additionally, the degree of correlation between branch length values calculated under the uncorrelated log-normal relaxed-clock and global clock models was calculated. Given the wide taxonomic coverage, most deep nodes in these mitochondrial DNA phylogenies are likely to be heavily saturated, even when applying the GTR-gamma model of sequence evolution. However, the nodes used for clock calibrations occur close to the tips of the tree where complete saturation is unlikely and rate variation can be detected. As such, we also compared the degree of correlation in branch lengths between the relaxed- and global clock models using only branches that occur between the time period of the oldest calibrated node (12 Ma) and the present.



**Fig. 1** Cross-validation results for each calibration point. (a)  $SS$  values ranked from highest (worst fit) to lowest (best fit)  $SS$  value, (b) calibrated rates (genetic distance divided by time) plotted by cross-validation rank. The first 16 data points (dark circles) in a and b poorly predicted the data set as a whole and were highlighted as outliers by the cross-validation test. (c) Histogram of molecular rate (genetic distance divided by time) estimates for supported calibrations (grey) and unsupported calibrations (black).

## Results

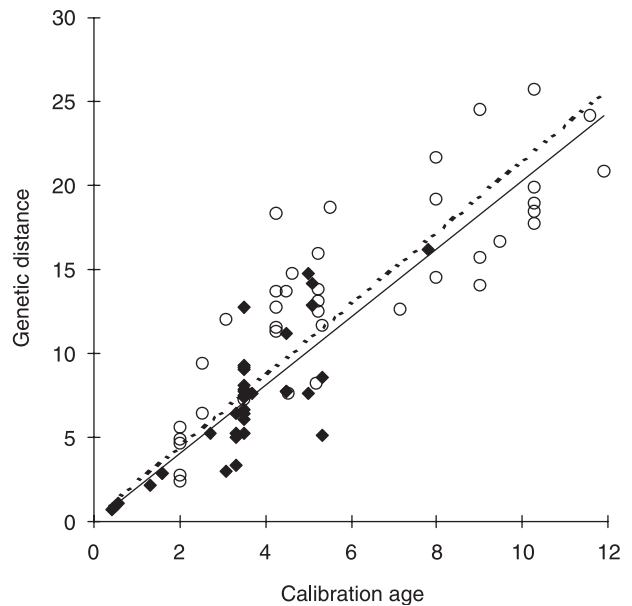
The 90 calibration points are listed in Table S1. These include 78 new molecular clock calibrations and re-analysis of 12 previously published calibrations less than 12 Ma for which we had corresponding cytochrome *b* sequences. Due to intense and extended episodes of volcanism in the Marquesas Islands, age of cessation of volcanism may provide a more realistic calibration point than the age of these islands.

When island age was used to calibrate the five endemic taxa of *Pomarea*, three of five calibrations became extreme outliers and all possessed much higher *SS* values than when dates of cessation of volcanism were used (Table S1). The cessation of volcanism more accurately predicted the ages of other calibrations in our data set under the cross-validation procedure, and so we use these dates in all further analysis.

Of the 90 calibration points, cross-validation supported the inclusion of 74 (global clock tree) or 75 (relaxed-clock tree) of these (Fig. 1a). Cross-validation results were identical except that *Spheniscus* was an outlier in the global clock tree but not in the relaxed-clock tree. We used results from the global clock tree in all further analysis. Figure 1(b) illustrates the relationship between the molecular rate for a calibration point and its ranking based on *SS* for the global clock analysis. The 16 'outliers' (12 of which are extreme 'outliers') exhibit a wide range of molecular rates (Fig. 1b, c) ranging from 0% to more than 10% sequence divergence per million years (0–5% per lineage per million years). Some of these were similar to the average rate supported by the data set as a whole (Fig. 1c), demonstrating that cross-validation is able to identify faulty calibration points even when their rates appear to agree with the consensus rate. In contrast, calibration points not excluded by the cross-validation test exhibited a reduced range of molecular rates (Fig. 1b, c) as expected under clock-like evolution.

A strong linear relationship exists between sequence divergence and the age of calibration points supported by the cross-validation procedure (Fig. 2). This linear relationship suggests that mutations accumulated steadily through time across the 12-Myr interval. The slope of this relationship yielded a rate of 2.13% ( $\pm 0.065\%$  SEM,  $F = 1084$ , d.f. = 1 and 73,  $P < 2.2 \times 10^{-16}$ ;  $r^2 = 0.94$ ; 95% confidence interval 2.02% to 2.22%). Scatter around this line may be generated by the stochastic nature of the clock, artefacts of the calibration process, or by genuine rate variation across lineages. First, we tested for rate variation by looking at the *uclid.stdev* statistic in our relaxed-clock Bayesian analysis. Although significantly greater than zero, *uclid.stdev* was small in magnitude (mean = 0.258, highest posterior density = 0.218–0.300) suggesting that sequences are evolving in an almost clock-like fashion and that rate variation across lineages, although present, is minor. Furthermore, branch lengths in the relaxed-clock and global clock phylogenies were highly correlated ( $r^2 = 0.992$ ), suggesting little rate variation across lineages. This was also true during the last 12 Myr ( $r^2 = 0.994$ ) when analysing only those branch lengths that date between the oldest calibrated node (12 Ma) and the present.

The type of calibration (fossil or biogeographical) explained some of the variance in sequence divergence through time. On average, sequence divergence for fossils was 2.3% higher than for biogeographical events when date was held constant ( $F = 9.35$ , d.f. = 1, 71,  $P = 0.003$ ; ANCOVA



**Fig. 2** Relationship between molecular divergence and time for 74 calibrations of the avian molecular clock supported by cross-validation. Biogeographical calibrations shown by black diamonds and solid least squares regression lines and fossil calibrations shown by open circles and dashed lines. Regression lines for fossils (slope = 2.16,  $r^2 = 0.62$ ) and biogeographical events (slope = 2.03,  $r^2 = 0.64$ ).

comparing regressions having equal slopes), as expected if fossils provide minimum calibrations and biogeographical events provide maximum calibrations. However, over the time period covered by this analysis, this difference was too small to affect estimated rates of molecular evolution significantly (2.16% and 2.03% per million years for fossil and biogeographical calibrations, respectively;  $F = 0.61$ , d.f. = 1, 72,  $P = 0.44$ ; comparison of linear regressions with *y*-intercepts fitted through the origin).

Some of the variance in sequence divergence with time appears to be partitioned among the 13 avian orders tested ( $F = 2.24$ ,  $P = 0.021$ , d.f. = 13, 60, ANCOVA) although not quite significant after simultaneously correcting for calibration type ( $F = 1.88$ , d.f. = 13, 60,  $P = 0.061$ ). Over the dates covered, this did not lead to significant differences among the orders in the estimated rates of molecular evolution ( $F = 1.42$ , d.f. = 12, 62,  $P = 0.14$ ; comparison of linear regressions with *y*-intercepts fitted through the origin). With the exception of three avian orders (Cuculiformes, Piciformes, Psittaciformes), each of which possessed only one or two calibrations, average rates within orders clustered tightly around 2.1% (Table 1).

Differences in generation time (approximated here by age at first breeding), body mass and latitude (binned as tropical, temperate, with mixed temperate and tropical or tropical migrants excluded) are also thought to contribute to different rates of molecular evolution (Laird *et al.* 1969;

**Table 1** Average molecular rates for 12 avian orders

Order	Rate (standard deviation)	Range	Slowest fossil calibration	Fastest biogeographical calibration	<i>n</i>
Anseriformes	2.29 ( $\pm$ 0.48)	1.75–2.77	1.75	2.77	6
Apodiformes	1.72			1.72	1
Charadriiformes	2.61 ( $\pm$ 0.81)	1.59–4.31	1.59		14
Ciconiiformes	2.10 ( $\pm$ 0.61)	1.57–3.02	1.57		5
Columbiformes	1.96 ( $\pm$ 0.10)	1.89–2.07		2.07	3
Cuculiformes	1.03			1.03	1
Falconiformes (Accipitridae)	1.81		1.81		1
Galliformes	2.38		2.38		1
Passeriformes	2.07 ( $\pm$ 0.20)	0.95–3.74	1.20	3.01	36
Piciformes	3.30 ( $\pm$ 0.47)	2.96–3.63		3.63	2
Procellariiformes	1.89 ( $\pm$ 0.35)	1.51–2.20	1.94	1.51	3
Psittaciformes	3.40		3.4		1
Combined	2.21 ( $\pm$ 0.68)	0.95–4.31	1.20	3.63	74

Martin & Palumbi 1993; Wright *et al.* 2006), but none of these factors significantly explained the variance in sequence divergence with time while simultaneously correcting for calibration type (generation time,  $F = 1.23$ , d.f. = 1, 68,  $P = 0.27$ ; mass,  $F = 0.026$ , d.f. = 1, 60,  $P = 0.87$ ; latitude,  $F = 2.59$ , d.f. = 1, 35,  $P = 0.12$ ; ANCOVA).

## Discussion

A mitochondrial DNA clock of approximately 2% has been widely used for birds, mammals, and other vertebrate groups (Brown *et al.* 1979; Klicka & Zink 1997; Garcia-Moreno 2004; Lovette 2004). This traditional rate was based on relatively few calibrations and disparate methods of estimating genetic distances [i.e. distances based on third codon positions, all codon positions or restriction fragment length polymorphism (RFLP) data; see review in Lovette 2004]. Nevertheless, using consistent methods of calibration (all codon positions for a single gene) and an unbiased method for choosing calibration points, our analysis supports an average molecular rate of 2.1% with little variation in mean rate across most orders (Table 1). These results suggest that the 2% rate is highly conserved in birds.

Recently, several authors have suggested that molecular rates are time dependent with the fastest rates obtained from calibration points near the recent (Garcia-Moreno 2004; Ho *et al.* 2005; Penny 2005). Two exceptionally fast calibrations dating to the past 2 Myr (10.6% *Nectarinia* and 4.9% *Branta*) were highlighted as outliers by cross-validation and most other calibrations dating to this time period agree with a rate close to 2%. As errors in assigning biogeographical or fossil dates to nodes and correcting for ancestral polymorphism may have a much greater effect on rate estimates for young than for old calibration points, we suspect apparent faster rates exhibited by some calibration points near the recent are an artefact.

After excluding the 16 calibration points which poorly predicted other calibrations in the data set, molecular calibrations ranged between 1.0% and 4.3% (Table 1; Fig. 1c). This fairly broad range suggests that substantial variation in rates may occur. A considerable proportion of the observed variation can be explained by the stochastic nature of sequence evolution. For example, 37 of the 74 calibrations fall within the expected range of variation for their given sequence lengths under a molecular clock of 2.1% (assuming a Poisson process of molecular divergence in which variance in the number of substitutions after a given amount of time is equal to the mean number of substitutions). The remaining rates differed significantly from a rate of 2.1% suggesting that rate variation does occur.

The imprecise nature of the calibration process likely inflates estimates of rate variation, even after exclusion of the most imprecise calibrations using cross-validation. Fossil-based calibrations have a tendency to overestimate molecular rates because poorly preserved fossil records result in the ages of calibration points being underestimated (Fig. 2). The fastest supported rates (4.3%) came from fossil calibrations. Biogeographical based calibrations have a tendency to underestimate molecular rates (Fig. 2) because colonization may occur anytime after the formation of islands, mountain ranges or land bridges. The slowest supported rates (1.0%) came from biogeographical calibrations. A more accurate estimate of the range of rate variation is derived from the slowest rates calibrated for fossils (1.2%) and the fastest rates calibrated for biogeographical events (3.6%; Table 1).

Dated phylogenies are best produced using multiple ingroup calibration points with a relaxed-clock model as implemented in a growing number of phylogenetic packages. The best-supported calibrations presented in Table S1 (e.g. those associated with small SS values) can be used as fixed age constraints. When less accurate calibrations

(those associated with larger *SS* values) are used, we suggest treating fossil-based calibrations as minimum age constraints and biogeographical-based calibrations as maximum age constraints. Given the absence of suitable internal calibration points for many groups of birds, the consensus estimate of molecular rates (2.1%) presented here can be used, provided stochastic error is allowed for. The 2% rule has previously been based on incredibly little data, and most avian studies have relied on the robust calibrations for Hawaiian honeycreepers (Fleischer *et al.* 1998) and geese (Shields & Wilson 1987). Our results demonstrate that average rates are fairly constant through time, similar across taxonomic orders (Table 1), and in general support the 2% rule.

### Acknowledgements

This manuscript benefitted from discussion with Trevor Price and from three anonymous reviewers. J.T.W. acknowledges a postdoctoral fellowship from the Natural Sciences and Engineering Research Council of Canada (NSERC) and D.S. acknowledges support from the NSERC and from the Canada Foundation for Innovation.

### References

- Abratis M, Worner G (2001) Ridge collision, slab-window formation, and the flux of Pacific asthenosphere into the Caribbean realm. *Geology*, **29**, 127–130.
- Arbogast BS, Edwards SV, Wakeley J, Beerli P, Slowinski JB (2002) Estimating divergence times from molecular data on phylogenetic and population genetic timescales. *Annual Review of Ecology and Systematics*, **33**, 707–740.
- Arbogast BS, Drovetski SV, Curry RL *et al.* (2006) The origin and diversification of Galapagos mockingbirds. *Evolution*, **60**, 370–382.
- Avise JC, Walker D, Johns GC (1998) Speciation durations and Pleistocene effects on vertebrate phylogeography. *Proceedings of the Royal Society B: Biological Sciences*, **265**, 1707–1712.
- Baldwin BG, Sanderson MJ (1998) Age and rate of diversification of the Hawaiian silversword alliance (Compositae). *Proceedings of the National Academy of Sciences, USA*, **95**, 9402–9406.
- Barker FK, Cibois A, Schikler P, Feinstein J, Cracraft J (2004) Phylogeny and diversification of the largest avian radiation. *Proceedings of the National Academy of Sciences, USA*, **101**, 11040–11045.
- Barraclough TG, Vogler AP (2002) Recent diversification rates in North American tiger beetles estimated from a dated mtDNA phylogenetic tree. *Molecular Biology and Evolution*, **19**, 1706–1716.
- Becker J (1987) *Neogene Avian Localities of North America*. Smithsonian Institution Press, Washington, D.C.
- Bohning-Gaese K, Caprano T, van Ewijk K, Veith M (2006) Range size: disentangling current traits and phylogenetic and biogeographic factors. *American Naturalist*, **167**, 555–567.
- Brown WM, George M, Wilson AC (1979) Rapid evolution of animal mitochondrial DNA. *Proceedings of the National Academy of Sciences, USA*, **76**, 1967–1971.
- Coates AG, Obando JA (1996) The geologic evolution of the Central American isthmus. In: *Evolution and Environment in Tropical America* (eds Jackson JBC, Budd AF, Coates AG), pp. 21–56. University of Chicago Press, Chicago, Illinois.
- Coates AG, Jackson JBC, Collins LS *et al.* (1992) Closure of the Isthmus of Panama: the near-shore marine record of Costa Rica and western Panama. *Geological Society of America Bulletin*, **104**, 814–828.
- Coates AG, Collins LS, Aubry MP, Berggren WA (2004) The geology of the Darien, Panama, and the late Miocene–Pliocene collision of the Panama arc with northwestern South America. *Geological Society of America Bulletin*, **116**, 1327–1344.
- Cooper A, Penny D (1997) Mass survival of birds across the Cretaceous–Tertiary boundary: molecular evidence. *Science*, **275**, 1109–1113.
- Drovetski SV (2003) Plio-Pleistocene climatic oscillations, Holarctic biogeography and speciation in an avian subfamily. *Journal of Biogeography*, **30**, 1173–1181.
- Drummond AJ, Rambaut A (2006) *BEAST v1.4.6*. Available from URL: <http://beast.bio.ed.ac.uk/>.
- Drummond AJ, Ho SYW, Rawlence N, Rambaut A (2007) *A Rough Guide to BEAST 1.4*. University of Auckland, New Zealand.
- Edwards SV, Beerli P (2000) Perspective: gene divergence, population divergence, and the variance in coalescence time in phylogeographic studies. *Evolution*, **54**, 1839–1854.
- Emslie SD (1998) *Avian Community, Climate, and Sea-Level Changes in the Plio-Pleistocene of the Florida Peninsula*. American Ornithologists' Union, Washington, D.C.
- Ericson PG, Anderson CL, Britton T *et al.* (2006) Diversification of Neoaves: integration of molecular sequence data and fossils. *Biology Letters*, **2**, 543–547.
- Fleischer RC, McIntosh CE, Tarr CL (1998) Evolution on a volcanic conveyor belt: using phylogeographic reconstructions and K-Ar-based ages of the Hawaiian Islands to estimate molecular evolutionary rates. *Molecular Ecology*, **7**, 533–545.
- Fuchs J, Ohlson JI, Ericson PGP, Pasquet E (2006) Molecular phylogeny and biogeographic history of the piculets (Piciformes: Picumninae). *Journal of Avian Biology*, **37**, 487–496.
- García-Moreno J (2004) Is there a universal mtDNA clock for birds? *Journal of Avian Biology*, **35**, 465–468.
- Grafe K, Frisch W, Villa IM, Meschede M (2002) Geodynamic evolution of southern Costa Rica related to low-angle subduction of the Cocos Ridge: constraints from thermochronology. *Tectonophysics*, **348**, 187–204.
- Ho SYW, Phillips MJ, Cooper A, Drummond AJ (2005) Time dependency of molecular rate estimates and systematic over-estimation of recent divergence times. *Molecular Biology and Evolution*, **22**, 1561–1568.
- Keigwin LD (1978) Pliocene closing of Isthmus of Panama, based on biostratigraphic evidence from nearby Pacific Ocean and Caribbean Sea cores. *Geology*, **6**, 630–634.
- Klicka J, Zink RM (1997) The importance of recent ice ages in speciation: a failed paradigm. *Science*, **277**, 1666–1669.
- Knowlton N, Weigt LA (1998) New dates and new rates for divergence across the Isthmus of Panama. *Proceedings of the Royal Society B: Biological Sciences*, **265**, 2257–2263.
- Krijgsman W, Hilgen FJ, Raffi I, Sierro FJ, Wilson DS (1999) Chronology, causes and progression of the Messinian salinity crisis. *Nature*, **400**, 652–655.
- Laird CD, McConaughy BL, McCarthy BJ (1969) Rate of fixation of nucleotide substitutions in evolution. *Nature*, **224**, 149–154.
- Lovette IJ (2004) Mitochondrial dating and mixed-support for the '2% rule' in birds. *Auk*, **121**, 1–6.
- Martin AP, Palumbi SR (1993) Body size, metabolic rate, generation time, and the molecular clock. *Proceedings of the National Academy of Sciences, USA*, **90**, 4087–4091.

- Mercer JM, Roth VL (2003) The effects of Cenozoic global change on squirrel phylogeny. *Science*, **299**, 1568–1572.
- Miller MJ, Bermingham E, Ricklefs RE (2007) Historical biogeography of the New World solitaires (*Myadestes* spp.). *Auk*, **124**, 868–885.
- Near TJ, Sanderson MJ (2004) Assessing the quality of molecular divergence time estimates by fossil calibrations and fossil-based model selection. *Philosophical Transactions of the Royal Society of London. Series B, Biological Sciences*, **359**, 1477–1483.
- Nei M, Li WH (1979) Mathematical model for studying genetic variation in terms of restriction endonucleases. *Proceedings of the National Academy of Sciences, USA*, **76**, 5269–5273.
- Penny D (2005) Evolutionary biology — relativity for molecular clocks. *Nature*, **436**, 183–184.
- Pereira SL, Baker AJ (2005) A molecular timescale for galliform birds accounting for uncertainty in time estimates and heterogeneity of rates of DNA substitutions across lineages and sites. *Molecular Phylogenetics and Evolution*, **38**, 499–509.
- Perez-Eman JL (2005) Molecular phylogenetics and biogeography of the Neotropical redstarts (*Myioborus*; Aves, Parulinae). *Molecular Phylogenetics and Evolution*, **37**, 511–528.
- Robinson-Rechavi M, Boussae B, Laudet V (2004) Phylogenetic dating and characterization of gene duplications in vertebrates: the cartilaginous fish reference. *Molecular Biology and Evolution*, **21**, 580–586.
- Rodriguez F, Oliver JL, Marin A, Medina JR (1990) The general stochastic model of nucleotide substitution. *Journal of Theoretical Biology*, **142**, 485–501.
- Shields GF, Wilson AC (1987) Calibration of mitochondrial DNA evolution in geese. *Journal of Molecular Evolution*, **24**, 212–217.
- Stehli FG, Webb SD (1985) *The Great American Interchange*. Plenum Press, New York.
- Swofford DL (2002) *PAUP\*4 0b10: Phylogenetic Analysis Using Parsimony (\*and Other Methods)*. Sinauer & Associates, Sunderland, Massachusetts.
- Tukey JW (1977) *Exploratory Data Analysis*. Addison-Wesley, Reading, Massachusetts.
- Vandepoole K, De Vos W, Taylor JS, Meyer A, van de Peer Y (2004) Major events in the genome evolution of vertebrates: paranome age and size differ considerably between ray-finned fishes and land vertebrates. *Proceedings of the National Academy of Sciences, USA*, **101**, 1638–1643.
- Weir JT (2006) Different timing and patterns of species accumulation in lowland and highland Neotropical birds. *Evolution*, **61**, 842–845.
- Weir JT, Schluter D (2004) Ice sheets promote speciation in boreal birds. *Proceedings of the Royal Society of London B Biological Science*, **271**, 1881–1887.
- Weir JT, Schluter D (2007) The latitudinal gradient in recent speciation and extinction rates in birds and mammals. *Science*, **315**, 1928–1933.
- Woodruff DS (2003) Neogene marine transgressions, palaeogeography and biogeographic transitions on the Thai-Malay Peninsula. *Journal of Biogeography*, **30**, 551–567.
- Wright S, Keeling J, Gillman L (2006) The road from Santa Rosalia: a faster tempo of evolution in tropical climates. *Proceedings of the National Academy of Sciences, USA*, **103**, 7718–7722.
- Yang Z (1993) Maximum likelihood estimation of phylogeny from DNA sequences when substitution rates differ over sites. *Molecular Biology and Evolution*, **10**, 1396–1401.

---

Jason T. Weir is interested in how speciation and extinction rates change across latitudinal gradients and contribute to the latitudinal diversity gradient. Dolph Schluter is interested in the roles of ecology and natural selection in driving diversifications.

---

### Supplementary material

The following supplementary material is available for this article:

**Table S1** Excel file of calibration data set contains genetic distances, calibration dates and paleontological or biogeographical references for each calibration point

**Appendix S1** Justification of clock calibrations. The complete data set of clock calibrations and associated biogeographical, paleontological and phylogenetic references are presented in Table S1. Additional details of calibration points are presented here

This material is available as part of the online article from:

<http://www.blackwell-synergy.com/doi/abs/10.1111/j.1365-294X.2008.03742.x>

(This link will take you to the article abstract).

Please note: Blackwell Publishing are not responsible for the content or functionality of any supplementary materials supplied by the authors. Any queries (other than missing material) should be directed to the corresponding author for the article.