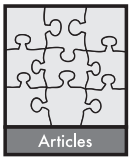

Journal of Avian Biology 47: 84–91, 2016
doi: 10.1111/jav.00766

Subject Editor: Darren Irwin. Editor-in-Chief: Jan-Åke Nilsson. Accepted 17 June 2015

Calibrating the molecular clock beyond cytochrome *b*: assessing the evolutionary rate of COI in birds

Pablo D. Lavinia, Kevin C. R. Kerr, Pablo L. Tubaro, Paul D. N. Hebert and Darío A. Lijtmaer

P. D. Lavinia (pablodlo23@gmail.com), P. L. Tubaro and D. A. Lijtmaer, División Ornitología, Museo Argentino de Ciencias Naturales 'Bernardino Rivadavia', Avenida Ángel Gallardo 470, Ciudad Autónoma de Buenos Aires, Argentina, C1405DJR. – K. C. R. Kerr, Toronto Zoo, Conservation, Education, and Wildlife Division, 361A Old Finch Avenue, Toronto, ON M1B 5K7, Canada. – P. D. N. Hebert, Biodiversity Inst. of Ontario, Univ. of Guelph, Guelph, ON N1G 2W1, Canada.

Estimating the age of species or their component lineages based on sequence data is crucial for many studies in avian evolutionary biology. Although calibrations of the molecular clock in birds have been performed almost exclusively using cytochrome *b* (cyt *b*), they are commonly extrapolated to other mitochondrial genes. The existence of a large, standardized cytochrome *c* oxidase subunit I (COI) library generated as a result of the DNA barcoding initiative provides the opportunity to obtain a calibration for this mitochondrial gene in birds. In this study we compare the evolutionary rate of COI relative to cyt *b* across ten different avian orders. We obtained divergence estimates for both genes from nearly 300 phylogenetically independent pairs of species through the analysis of almost 5000 public sequences. For each pair of species we calculated the difference in divergence between COI and cyt *b*. Our results indicate that COI evolves on average 14% slower than cyt *b*, but also reveal considerable variation both among and within avian orders, precluding the use of this value as a standard adjustment for the COI molecular clock for birds. Our findings suggest that this variation is partially explained by a clear negative relationship between the difference in divergence in these genes and the age of species. Distances for cyt *b* are higher than those for COI for closely related species, but the values become similar as the divergence between the species increases. This appears to be the result of a stronger pattern of negative time-dependency in the rate of cyt *b* than in that of COI, a difference that could be related to lower functional constraints on a small number of sites in cyt *b* that allow it to initially accumulate mutations more rapidly than COI.

Most current knowledge of the tempo and mode of avian diversification relies on the estimation of the age of species or their intraspecific lineages and whether the timing is compatible with a particular evolutionary scenario. Genetic data has now become the standard tool to date events in studies of ecology and evolutionary biology through the application of molecular clocks. The underlying assumption is that nucleotide differences between two DNA sequences accumulate in a regular manner (i.e. a clock-like fashion) over time. At least one calibration point (i.e. an independent time reference such as a biogeographic event or the fossil record) and DNA sequence divergence information are needed to estimate substitution rates (Lovette 2004, Weir and Schluter 2008).

Although calibrations are available for several groups of organisms, the avian mitochondrial molecular clock is one of the most investigated and discussed (García-Moreno 2004, Lovette 2004, Weir and Schluter 2008). Calibrations of the molecular clock in birds have focused almost exclusively on cytochrome *b* (cyt *b*) with detailed studies focusing on a few avian groups (Lovette 2004, Weir and Schluter 2008). However, the evolutionary rate for this gene has commonly been applied to different taxonomic groups and loci (Ho 2007), despite evidence for substantial rate variation among genes and lineages (see below). This practice creates the potential for errors in molecular dating beyond those due to the under- or over-estimation of substitution rates as a consequence of the incomplete nature of the fossil record and the uncertainty concerning the precise dates of geological events thought to be responsible for lineage isolation (García-Moreno 2004, Weir and Schluter 2008, Ho et al. 2011).

Despite the regular use of the 2.1% sequence divergence per million years rate for cyt *b* as the universal mitochondrial clock in birds (Shields and Wilson 1987, Paxinos et al. 2002, Weir and Schluter 2008), there is evidence for considerable rate heterogeneity across taxa and protein-coding loci (Lovette 2004, Pereira and Baker 2006, Patané et al. 2009, Eo and DeWoody 2010, Pacheco et al. 2011). As reviewed by Ho and Duchene (2014), this heterogeneity could be explained a) by differences in the absolute rate of different coding and non-coding regions (i.e. gene effects) mostly because of differing proportions of sites under selection in each locus (Kimura and Ohta 1974, Mouchiroud et al. 1995), b) by disparities in species longevity, generation time, metabolic rate, and body size (i.e. lineage effects; Martin and Palumbi 1993, Mooers and Harvey 1994, Nabholz et al.

2009), or c) by a combination of both effects (Muse and Gaut 1997).

As well, molecular rate estimates appear to be negatively dependent on the depth of the calibration point used for estimation, being higher within species and for recently diverged lineages than for older taxa. The time dependency pattern may arise as a consequence of the difference between short-term mutation rates and long-term substitution rates, which reflect different biological processes (García-Moreno 2004, Ho et al. 2005, 2011). Rates estimated over short evolutionary periods will be mostly based on nucleotide changes that represent all type of mutations (except for the lethal ones) before the deleterious or slightly deleterious changes are removed by natural selection. These rates will therefore resemble the spontaneous mutation rate. On the other hand, rates obtained over long timescales will be mostly derived from the few mutations that were fixed (i.e. the substitutions) after the removal of deleterious mutations, resulting in rates closer to the substitution rate (Ohta 1992, Subramanian et al. 2009, Ho et al. 2011).

As a result of the DNA barcoding initiative (Hebert et al. 2003a), a large and well-curated library of cytochrome *c* oxidase subunit I (COI) sequences is available. Given its increased availability, this gene has become a frequent choice, alone or in combination with other loci, in phylogenetic and phylogeographic analyses (Campagna et al. 2010, Weir and Price 2011, Kerr and Dove 2013, Lavinia et al. 2015), as well as for the study of general evolutionary patterns (Kerr et al. 2009, Lijtmaer et al. 2011, Tavares et al. 2011, Milá et al. 2012). However, using this marker for molecular dating is risky because the COI molecular clock has not been carefully calibrated in birds. The present study addresses this gap through a large-scale assessment of the COI evolutionary rate in birds relative to that of cyt *b* in 556 species from ten avian orders. This constitutes, to our knowledge, the most taxonomically comprehensive analysis of the evolutionary rate of COI taking into account phylogenetic information, heterogeneity across avian orders, and time-dependency.

Material and methods

We first created a list of all avian species with publicly available COI data (as of July 2012) in the Barcode of Life Data Systems (BOLD, Ratnasingham and Hebert 2007) for all orders for which we had detailed phylogenetic information (see below), therefore restricting our analyses of this gene to its 'barcode region' (Hebert et al. 2004). Subsequently, we searched for cyt *b* sequences for the same species in GenBank (<www.ncbi.nlm.nih.gov>), and downloaded all records for the species for which both genes were available. We discarded all sequences shorter than 500 bp for COI and 950 bp for cyt *b*. Since the COI region used for DNA barcoding is standardized, COI sequences overlap almost completely, but 950 bp are needed to assure at least 550 bp of overlap for cyt *b*. This resulted in a final dataset of over 2800 COI and 2100 cyt *b* sequences representing 556 species from ten avian orders (Table 1).

Phylogenetic references were obtained from the published literature avoiding, whenever possible, phylogenetic trees based on morphological traits, outdated reconstructions, and phylogenies with few taxa from the avian group under analysis. Using this phylogenetic information, we generated as many phylogenetically independent pairs of species (Felsenstein 1985) as possible within each order to compare the divergence levels in COI and cyt *b* sequences (Supplementary material Appendix 1, Table A1). Uncorrected p-distances were estimated in MEGA 5.0 (Tamura et al. 2011) and then used to calculate an index that indicates the difference in divergence between the genes expressed as a percentage (cyt *b*/COI index) for each pair of species. This index was calculated so that the differences between the loci were always relative to the gene with the lowest divergence using the following formula: [(higher – lower)/(lower)] $\times 100$, where 'higher' and 'lower' correspond to genetic p-distances. This formula assures that, given a certain difference in divergence between the genes, the index value obtained from it will be the same regardless of which of them is the one with the higher divergence in each pair of species. Lastly, when the distance for COI was higher than that for cyt *b*, we added a minus sign so positive index values indicate higher divergence in cyt *b*, while negative values indicate a higher genetic distance for COI. Apart from considering values for individual comparisons, the index was averaged to obtain mean values for each order and for the whole dataset. The global difference in divergence between the genes was calculated after weighting the mean index values for the different orders by the number of pairs of species compared within each of them.

For most species in our dataset, the COI and cyt *b* sequences used did not derive from the same individual. To test for any bias introduced by this procedure, we created a Paired Gene Dataset restricted to those species for which both genes were sequenced from the same specimen. To increase the size of this dataset we selected tissue samples stored at the Museo Argentino de Ciencias Naturales 'Bernardino Rivadavia' for which we had already obtained a COI sequence (as a result of our participation in the project to DNA barcode the birds of southern South America; Kerr et al. 2009) and obtained their cyt *b* sequence. We extracted genomic DNA following a glass fiber-based extraction protocol developed by Ivanova et al. (2006) and then amplified cyt *b* using a shortened version of the primer L14841 and the primer H16065 (Lougheed et al. 2000). Thermocycling conditions for cyt *b* consisted of 3 min at 94°C; 40 cycles of 45 s at 94°C, 30 s at 55°C and 1 min at 72°C; final extension of 10 min at 72°C. Sequencing was performed bidirectionally at Macrogen (Seoul, Korea) with the same primers used for amplification. GenBank accession numbers for these sequences are: KP965490–KP965532. This generated a dataset that included 150 species (354 individuals in total) from eight avian orders (Supplementary material Appendix 1, Table A2).

Since the use of p-distances could potentially underestimate the true divergence between species due to mutational saturation (i.e. multiple hits), especially over long evolutionary times, we performed two different analyses to control for this. In the first place, we also estimated genetic distances for the entire dataset using the Kimura two-parameter model (K2P, Kimura 1980), a simple but commonly applied correction for genetic distances which assigns different weights to transitions and transversions. Because the results were

Table 1. Mean difference in divergence between cyt *b* and COI (cyt *b*/COI index) across ten avian orders obtained from the comparison of uncorrected p-distances in 278 phylogenetically independent pairs of species. Positive index values indicate higher divergence in cyt *b* than in COI. The last two columns indicate the total number of COI and cyt *b* sequences analyzed within each order. Weighted average correspond to the global difference in divergence between the genes for the complete dataset, calculated after weighting the mean index values obtained for the different orders by the number of pair of species compared within each of them.

Order	Pairs of species	cyt *b*/COI index (%)	COI sequences	cyt *b* sequences
Anseriformes	23	49.5	340	104
Apodiformes	3	2.2	19	28
Charadriiformes	51	15.1	450	258
Columbiformes	20	6.4	135	69
Falconiformes	4	12.0	57	40
Accipitriformes	18	16.4	122	115
Galliformes	8	37.6	104	237
Piciformes	25	10.7	147	207
Strigiformes	12	1.8	117	104
Passeriformes	114	7.8	1338	976
Total	278		2829	2138
Weighted average		13.9%		

almost identical (both the patterns and the magnitude of the index for each order), we report only those obtained using p-distances. The second, more thorough analysis to compare the results obtained using p-distances and corrected divergences was performed using the Paired Gene Dataset. We first selected the best-fit model of nucleotide substitution for each gene and order using the Bayesian information criterion (BIC) implemented in jModelTest 2.1.1 (Darriba et al. 2012), then re-calculated genetic distances using PAUP* 4.0 (Swofford 2001), and finally repeated all the analyses with the new set of distances (Supplementary material Appendix 1, Table A3).

The association between the divergences estimated from both genes across all pairs of species and its statistical significance were assessed using Pearson product-moment correlations and reduced major axis (RMA) regressions performed in Statistica 8.0 (StatSoft, Tulsa, OK, USA) and RMA 1.21 (Bohonak and van der Linde 2004), respectively. The RMA regression is the appropriate analysis when both variables are assumed to have an associated error term (LaBarbera 1989, Quinn and Keough 2002 and references therein). All tests were two-tailed.

Results

Based on the comparison of uncorrected p-distances for COI and cyt *b* in almost 300 phylogenetically independent pairs of species across ten avian orders, we found that COI evolves, on average, 14% slower than cyt *b* (Table 1). Applying the widely used cyt *b* molecular clock (2.1% sequence divergence per million years; Weir and Schluter 2008) our result implies that COI evolves at an average rate of around 1.8% sequence divergence per million years. Since passerines represent around 40% of our total dataset, we also recalculated this global difference in divergence after excluding members of this speciose order. The result was very similar, with COI evolving on average 18% slower than cyt *b*.

Nevertheless, we found considerable variance in the cyt *b*/COI index among avian orders. Divergence values for COI and cyt *b* were, on average, fairly similar for groups like Strigiformes and Apodiformes, but cyt *b* divergences were considerably higher for Anseriformes and Galliformes, with all other orders showing intermediate index values (Table 1). COI was not found to evolve faster than cyt *b* on average (i.e. negative index values) in any of the orders examined.

Closer inspection revealed that variation within orders was also high and that it was associated with the level of divergence between the species compared (see Table 2 for the case example of Galliformes). To further investigate this pattern we used a reduced major axis regression to compare the divergence in both markers across all pairs of species within our dataset. Even though the association between the distances for both genes is, as expected, highly significant ($r = 0.93$, $p < 0.0001$) and close to the 1:1 relationship (equal rates), the distance for cyt *b* is higher than that for COI for pairs of species with low COI divergence, whereas values tend to equalize as divergence increases (Fig. 1). The same pattern was observed when we analyzed passerine and non-passerine species pairs independently (Supplementary material Appendix 2, Fig. A1), although it was slightly clearer for the non-passerines, most likely due to a higher proportion of non-passerine comparisons in the region of low divergences.

Table 2. Phylogenetically independent pairs of Galliformes species used for the comparison of cyt *b* and COI uncorrected p-distances. The table shows the variation in the cyt *b*/COI index observed within the order and its association with the depth of the divergence between the species that are being compared. The distance for cyt *b* is higher than that for COI (positive index values) for species pairs with low COI divergence, whereas the opposite is true for more distantly related species.

Pair of species	cyt *b*/COI index (%)	COI divergence (%)
Mitu salvini vs *Mitu tuberosum*	188.9	0.9
Perdix perdix vs *Perdix dauurica*	28.6	3.5
Tetrao urogallus vs *Tetrao tetrix*	36.2	4.7
Lagopus lagopus vs *Lagopus muta*	29.2	4.8
Alectoris chukar vs *Alectoris philbyi*	72.2	2.2
Colinus virginianus vs *Oreortyx pictus*	–16.8	12.5
Coturnix japonica vs *Gallus gallus*	–9.7	14.7
Phasianus colchicus vs *Falcipennis canadensis*	–28.1	17.3
Average	37.6	

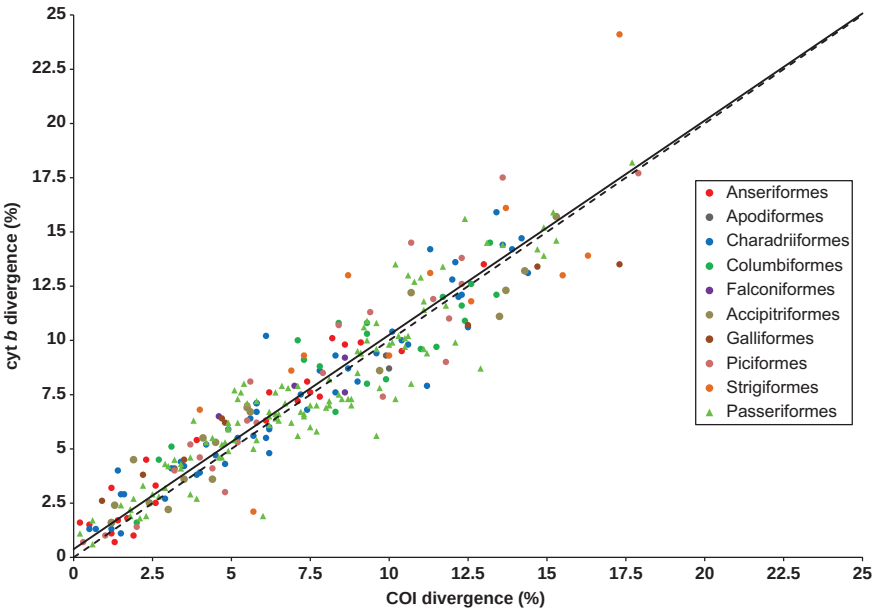

Figure 1. Comparison of uncorrected p-distances for COI and cyt *b* in 278 phylogenetically independent pairs of species. The dashed line shows equal divergence between the two loci. The solid line corresponds to the reduced major axis regression (cyt $b = 0.9879 \times \text{COI} + 0.3774$, $r^2 = 0.87$, where cyt *b* and COI correspond to genetic distances in each gene). Pearson correlation coefficient (r) = 0.93, $p < 0.0001$. Shape and colour of the symbols indicate the order to which each pair of species belongs.

This influence of the level of divergence between pairs of species is even more evident when the cyt *b*/COI index is plotted against COI divergence (Fig. 2). The marked decay of the index with species divergence shows that genetic distances for cyt *b* tend to be higher than those in COI for recently diverged pairs of species. The same tendency was observed when passerine and non-passerine species were analyzed separately (Supplementary material Appendix 2, Fig. A2). Lastly, the index was found to be negatively and statistically significantly correlated with mean COI divergence between the pairs of species of each order ($r = -0.72$, $p < 0.001$; Fig. 3), suggesting that differences in the index among orders can be partially explained by the mean divergence among the pairs of species in them. This result did not change when passerines were excluded from the analysis (Supplementary material Appendix 2, Fig. A3).

Finally, analyses performed with the Paired Gene Dataset, both before and after correction for mutational saturation (Supplementary material Appendix 1, Table A2 and A3), yielded very similar results to those obtained with the entire

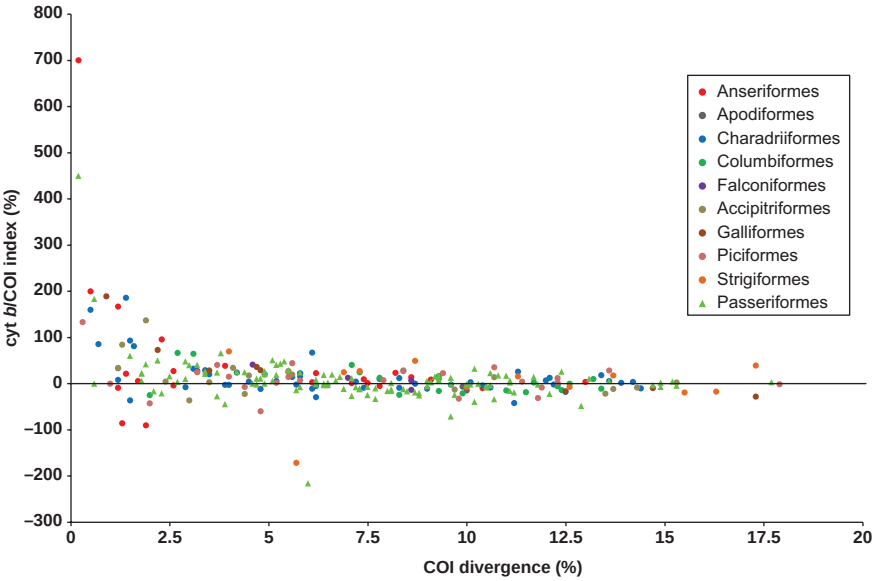

Figure 2. Decay of the difference in divergence between cyt *b* and COI (cyt *b*/COI index) as the uncorrected p-distance for COI between the species being compared increases. Positive index values indicate higher divergence in cyt *b* than in COI, whereas negative values are the reverse. Shape and colour of the symbols indicate the order to which each pair of species belongs.

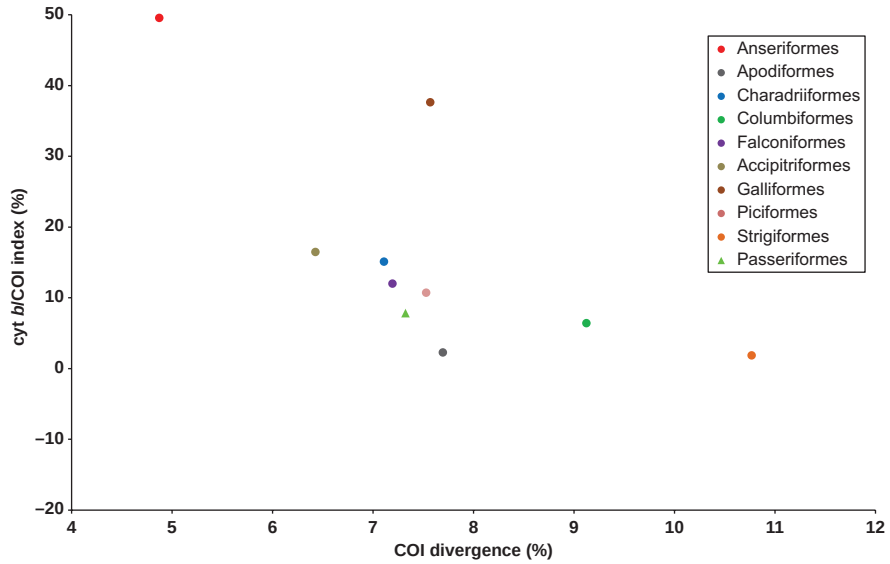

Figure 3. Association between the mean difference in divergence between cyt *b* and COI (cyt *b*/COI index) and the mean COI uncorrected p-distance for ten avian orders. Pearson correlation coefficient (r) = –0.72, $p < 0.001$. The correlation was performed after weighting the mean index values of each order by the number of pair of species compared within each of them. Positive index values indicate higher divergence in cyt *b* than in COI.

dataset (Supplementary material Appendix 2 Fig. A4 to A9), with differences only in the magnitude of the index (Supplementary material Appendix 1, Table A4 and A5). Given the considerable variation within and among orders in the full dataset, differences in the value of the indices (both for each order and for the whole dataset) were expected when analyzing the Paired Gene Dataset. However, we were not interested in comparing the absolute values of the indices for each dataset, but rather the patterns within them, which were highly congruent. This consistency across datasets indicates that no bias was introduced by the comparison of COI and cyt *b* sequences from differing individuals or by the use of uncorrected p-distances to quantify divergence. The latter was expected, though, given the fact that we found COI and cyt *b* to differ only at low divergences and that mutational saturation is a crucial factor over long evolutionary periods rather than over short time frames (Ho et al. 2011, Molak and Ho 2015).

Discussion

Molecular dating is a key, widely accepted component in evolutionary studies, and a tool frequently employed to probe avian diversification history. However, its utility relies on the careful calibration and proper application of molecular substitution rates, which are currently derived mainly from mitochondrial DNA (mostly from cyt *b*) and only to a much lesser extent from nuclear DNA (Lovette 2004, Ellegren 2007, Weir and Schluter 2008). Motivated by the rapidly increasing volume of COI data available for birds, we performed a large-scale comparison of the rates of molecular evolution in COI and cyt *b*. Our findings confirm past evidence for rate heterogeneity across loci, and indicate that differences in evolutionary rates in these two genes are partially explained by the depth of the divergence between the species being compared and a stronger pattern of time dependency in the rate of cyt *b* than in that of COI.

Rate heterogeneity and the age of species

The present results establish that COI evolves on average 14% slower than cyt *b* across ten avian orders. However, the substantial variation in this value within and among orders hinders its application as a standard adjustment for the COI molecular clock for birds. Previous studies have similarly stressed that substitution rates may vary across loci and lineages (Pereira and Baker 2006, Eo and DeWoody 2010, Pacheco et al. 2011), questioning the existence of a universal mitochondrial clock for birds (García-Moreno 2004, Lovette 2004; but see Weir and Schluter 2008). Our results are consistent with rate variation between these loci since at least one of them must show rate heterogeneity among taxa to produce the pattern in the cyt *b*/COI index detected in this study.

Genetic distances for cyt *b* were generally higher than those for COI for recently diverged pairs of species, resulting in a marked decrease of the cyt *b*/COI index as divergence increased. Despite the fact that our study does not compare absolute rate estimates through time but rather assess the effect of time on the relationship between cyt *b* and COI divergences, the decay curve of the index (Fig. 2) does resemble that of the pattern of time-dependency in rate estimates where substitution rates estimated from younger calibration points (less than two million years ago) are faster than those calibrated with older time references (García-Moreno 2004, Ho et al. 2005). Our results suggest that this standard effect (i.e. accelerated short-term rates) might be stronger in cyt *b* than in COI.

Prior studies have provided compelling evidence for the time-dependency of both coding and non-coding mitochondrial loci across animal taxa (García-Moreno 2004, Ho et al.

2005, 2007, Molak and Ho 2015; but see Weir and Schluter 2008), although little is known about the avian COI in particular. Pacheco et al. (2011) reported that apart from being the slowest evolving mitochondrial gene among Neoaves, COI has less rate heterogeneity across lineages, and better fits a clock-like model. Similarly, though with a narrower taxonomic focus (*Ramphastos* toucans – Piciformes), Patané et al. (2009) found that COI was the only mitochondrial protein-coding gene to approximate the 'universal' clock of 2%, while cyt *b* and ND2 (the other loci that were examined) showed markedly higher rates. However, there is still no evidence on whether COI evolutionary rate estimates vary with the age of the calibration point used for estimation, as it happens in the case of cyt *b*. Therefore the pattern reported here could arise either because cyt *b* rate shows a higher degree of time-dependent variation than does COI rate, or because COI rate is fairly constant. Further studies are needed to discriminate between these two possibilities.

We show here that COI and cyt *b* divergences are similar except for closely related species pairs. The question then arises as to why these two genes evolve at different rates over short evolutionary periods.

Differences between COI and cyt *b*

Negative or purifying selection (i.e. the removal from populations of functionally deleterious or slightly deleterious mutations; Kimura and Ohta 1974) may be insufficient to explain the time-dependency of molecular rate estimates (Woodhams 2006, Peterson and Masel 2009, Ho et al. 2011), but differing intensities of this effect could explain the differences observed between COI and cyt *b* for closely related species. Kerr (2011) showed that COI is the least variable gene in the mitochondrial genome and that this is most likely a consequence of strong purifying selection and functional constraint rather than of selective sweeps (i.e. positive selection). More precisely, it has been shown that non-synonymous substitution rates are much lower for COI than for other protein-coding mitochondrial loci, whereas synonymous substitution rates are similar (Eo and DeWoody 2010, Kerr 2011).

Cytochrome *b* might initially accumulate mutations more rapidly than COI if it possesses some hypervariable sites where substitutions are only weakly exposed to purifying selection while COI lacks such sites. Mutational hotspots have been reported in protein-coding mitochondrial DNA from mammals (Galtier et al. 2006), and their position seemed to change with time and across lineages. Conversely, Ho et al. (2005) and Denver et al. (2000) showed that substitutions are uniformly distributed among cyt *b* sequences of primate taxa and across the mitochondrial genome of *Caenorhabditis elegans* respectively.

Examination of some closely related species pairs (with less than 2.5% COI divergence) in our dataset revealed no clear signs of persistent mutational hotspots in the cyt *b* sequence, but suggested that some hypervariable sites may exist. We found 252 base substitutions across 192 different sites, 86% of them involving changes in third-codon positions. This means that 47 substitutions (96% at third-codon positions) occurred in the same site in more than one species pair. Following Ho et al. (2005), we also divided the cyt *b* sequence into five sections and tested whether substitution counting fit a Poisson distribution (i.e. they were randomly distributed across the length of the sequence). Although there were slightly more substitutions in the last two regions (686–914 bp and 915–1143 bp) of the gene, we found no significant differences across regions (Chi-square test, $\chi^2 = 3.21$, $p = 0.36$) and therefore could not reject the null hypothesis of random distribution. However, a deeper and more thorough analysis (e.g. Galtier et al. 2006), taking into account the co-occurrence of intraspecific polymorphisms between closely related species, is necessary to rigorously test the existence of mutational hotspots in the cyt *b* sequence that could explain its higher substitution rate over short evolutionary intervals.

Implications of our results and the use of COI in evolutionary studies

As the 5' region of COI has been adopted as the marker of choice for DNA barcoding in the animal kingdom (Hebert et al. 2003a, b), the number of publicly available COI sequences has increased dramatically, prompting its use in phylogenetic and phylogeographic studies. Since our results show that cyt *b* and COI behave differently only for recently diverged species and most species pairs do not fall into the 'danger zone' of low divergences, both COI or cyt *b* could be used to estimate species ages with similar results in most cases. However, when a sole calibration is used to assess the age of different nodes in phylogenetic analyses (therefore covering a wide range of species divergences), COI could be more appropriate than cyt *b* because its evolutionary rate appears relatively more uniform, making the application of molecular clocks (either strict or relaxed) more legitimate. On the other hand, COI is particularly constrained between recently diverged taxa (Aliabadian et al. 2009, Lijtmaer et al. 2011, this study), so it is probably not the best choice to gain resolution of lineages in studies below the species level, especially in the assessment of recent or ongoing speciation. Instead, loci with higher rates should be used, although with knowledge that the application of 'universal' evolutionary rates (such as the widely used 2.1% per million years for cyt *b*; Weir and Schluter 2008) should be avoided whenever possible (see Cabanne et al. 2008 for one of the few examples were molecular rates were corrected for time-dependency).

Further directions

Several approaches have been developed to minimize the effects of rate heterogeneity and to deal with the impact of rate uncertainty (Ho and Duchene 2014) while we wait for new calibrations to become available for unexplored avian groups and alternative loci. Although it is more likely that the debate on molecular dating and the factors that could affect molecular rate estimates will continue, we believe that it is worth revisiting past calibrations (which are mostly based on cyt *b* sequence data) in order to estimate rates of molecular evolution for other genes, like COI. Assessing whether the patterns reported here are sustained when comparing absolute (i.e. not relative) evolutionary rates, together with a deeper investigation into the existence of mutational hotspots across the avian mitochondrial genome, would

advance our understanding of evolutionary rate variation among loci and taxonomic groups.

Acknowledgements – We thank all the researchers that through their effort have generated all public sequences used in this study. In particular we are grateful to Mark Stoeckle for coordinating the barcoding projects aimed at obtaining worldwide avian COI sequences. We also thank the collectors and staff of the División Ornitología of the Museo Argentino de Ciencias Naturales 'Bernardino Rivadavia' for their participation in several collection expeditions that made available the tissue samples used here. We also thank Darren Irwin for his valuable comments and suggestions that improved our manuscript. We thank the Administración de Parques Nacionales and Secretaría de Ambiente y Desarrollo Sustentable de la Nación from Argentina for granting the permits. This study was funded by the Univ. de Buenos Aires, the Consejo Nacional de Investigaciones Científicas y Técnicas, and the Agencia Nacional de Promoción Científica y Tecnológica from Argentina, and the International Development Research Centre of Canada.

References

Aliabadian, M., Kaboli, M., Nijman, V. and Vences, M. 2009. Molecular identification of birds: performance of distance-based DNA barcoding in three genes to delimit parapatric species. – PLoS One 4: e4119.

Bohonak, A. J. and van der Linde, K. 2004. RMA: software for reduced major axis regression, Java version. – <www.kimvdlinde.com/professional/rma.html>.

Cabanne, G. S., d'Horta, F. M., Sari, E. H. R., Santos, F. R. and Miyaki, C. Y. 2008. Nuclear and mitochondrial phylogeography of the Atlantic Forest endemic *Xiphorhynchus fuscus* (Aves: Dendrocolaptidae): biogeography and systematic implications. – Mol. Phylogenet. Evol. 49: 760–773.

Campagna, L., Lijtmaer, D. A., Kerr, K. C. R., Barreira, A. S., Hebert, P. D. N., Lougheed, S. C. and Tubaro, P. L. 2010. DNA barcodes provide new evidence of a recent radiation in the genus *Sporophila* (Aves: Passeriformes). – Mol. Ecol. Resour. 10: 449–458.

Darriba, D., Taboada, G. L., Doallo, R. and Posada, D. 2012. jModelTest 2: more models, new heuristics and parallel computing. – Nat. Methods 9: 772.

Denver, D. R., Morris, K., Lynch, M., Vassilieva, L. L. and Thomas, W. K. 2000. High direct estimate of the mutation rate in the mitochondrial genome of *Caenorhabditis elegans*. – Science 29: 2342–2344.

Ellegren, H. 2007. Molecular evolutionary genomics of birds. – Cytogenet. Genome Res. 117: 120–130.

Eo, S. H. and DeWoody, J. A. 2010. Evolutionary rates of mitochondrial genomes correspond to diversification rates and to contemporary species richness in birds and reptiles. – Proc. R. Soc. B 277: 3587–3592.

Felsenstein, J. 1985. Phylogenies and the comparative method. – Am. Nat. 125: 1–15.

Galtier, N., Enard, D., Radondy, Y., Bazin, E. and Belkhir, K. 2006. Mutation hotspots in mammalian mitochondrial DNA. – Genome Res. 16: 215–222.

García-Moreno, J. 2004. Is there a universal mtDNA clock for birds? – J. Avian Biol. 35: 465–468.

Hebert, P. D. N., Cywinska, A., Ball, S. L. and DeWaard, J. R. 2003a. Biological identifications through DNA barcodes. – Proc. R. Soc. B 270: 313–321.

Hebert, P. D. N., Ratnasingham, S. and DeWaard, J. R. 2003b. Barcoding animal life: cytochrome *c* oxidase subunit 1 divergences among closely related species. – Proc. R. Soc. B 270 (Suppl. 1): S96–S99.

Hebert, P. D. N., Stoeckle, M. Y., Zemlak, T. S. and Francis, C. M. 2004. Identification of birds through DNA barcodes. – PLoS Biol. 2: 1657–1663.

Ho, S. Y. W. 2007. Calibrating molecular estimates of substitution rates and divergence times in birds. – J. Avian Biol. 38: 409–414.

Ho, S. Y. W. and Duchene, S. 2014. Molecular-clock methods for estimating evolutionary rates and timescales. – Mol. Ecol. 23: 5947–5965.

Ho, S. Y. W., Phillips, M. J., Cooper, A. and Drummond, A. J. 2005. Time dependency of molecular rate estimates and systematic overestimation of recent divergent times. – Mol. Biol. Evol. 22: 1561–1568.

Ho, S. Y. W., Shapiro, B., Phillips, M., Cooper, A. and Drummond, A. J. 2007. Evidence for time dependency of molecular rate estimates. – Syst. Biol. 56: 515–522.

Ho, S. Y. W., Lanfear, R., Brohman, L., Phillips, M. J., Soubrier, J., Rodrigo, A. G. and Cooper, A. 2011. Time-dependent rates of molecular evolution. – Mol. Ecol. 20: 3087–3101.

Ivanova, N. V., Dewaard, J. R. and Hebert, P. D. N. 2006. An inexpensive, automation-friendly protocol for recovering high-quality DNA. – Mol. Ecol. Notes 6: 998–1002.

Kerr, K. C. R. 2011. Searching for evidence of selection in avian DNA barcodes. – Mol. Ecol. Resour. 11: 1045–1055.

Kerr, K. C. R. and Dove, C. J. 2013. Delimiting shades of gray: phylogeography of the northern fulmar, *Fulmarus glacialis*. – Ecol. Evol. 3: 1915–1930.

Kerr, K. C. R., Lijtmaer, D. A., Barreira, A. S., Hebert P. D. N. and Tubaro, P. L. 2009. Probing evolutionary patterns in neotropical birds through DNA barcodes. – PLoS One 4: e4379.

Kimura, M. 1980. A simple method for estimating evolutionary rate of base substitutions through comparative studies of nucleotide sequences. – J. Mol. Evol. 15: 111–120.

Kimura, M. and Ohta, T. 1974. On some principles governing molecular evolution. – Proc. Natl Acad. Sci. USA 71: 2848–2852.

LaBarbera, M. 1989. Analyzing body size as a factor in ecology and evolution. – Annu. Rev. Ecol. Syst. 20: 97–117.

Lavinia, P. D., Escalante, P., García, N. C., Barreira, A. S., Trujillo-Arias, N., Tubaro, P. L., Naoki, K., Miyaki, C. Y., Santos, F. R. and Lijtmaer, D. A. 2015. Continental-scale analysis reveals deep diversification within the polytipic red-crowned ant tanager (*Habia rubica*, Cardinalidae). – Mol. Phylogenet. Evol. 89: 182–193.

Lijtmaer, D. A., Kerr, K. C. R., Barreira, A. S., Hebert, P. D. and Tubaro, P. L. 2011. DNA barcode libraries provide insight into continental patterns of avian diversification. – PLoS One 6: e20744.

Lougheed, S. C., Freeland, J. R., Handford, P. and Boag, P. T. 2000. A molecular phylogeny of warbling-finches (*Poospiza*): paraphyly in a Neotropical emberizid genus. – Mol. Phylogenet. Evol. 3: 367–378.

Lovette, I. J. 2004. Mitochondrial dating and mixed support for the "2% rule" in birds. – Auk 121: 1–6.

Martin, A. P. and Palumbi, S. R. 1993. Body size, metabolic rate, generation time, and the molecular clock. – Proc. Natl Acad. Sci. USA 90: 4087–4091.

Milá, B., Tavares, E. S., Saldaña, A. M., Karubian, J., Smith, T. B. and Baker, A. J. 2012. A trans-Amazonian screening of mtDNA reveals deep intraspecific divergence in forest birds and suggests a vast underestimation of species diversity. – PLoS One 7: e40541.

Molak, M. and Ho, S. Y. W. 2015. Prolonged decay of molecular rate estimates for metazoan mitochondrial DNA. – PeerJ 3: e821.

Mooers, A. O. and Harvey, P. H. 1994. Metabolic rate, generation time, and the rate of molecular evolution in birds. – Mol. Phylogenet. Evol. 3: 344–350.

Mouchiroud, D., Gautier, C. and Bernardi, G. 1995. Frequencies of synonymous substitutions in mammals are gene-specific and

correlated with frequencies of nonsynonymous substitutions. – J. Mol. Evol. 40: 107–113.
Muse, S. V. and Gaut, B. S. 1997. Comparing patterns of nucleotide substitution rates among chloroplast loci using the relative ratio test. – Genetics 146: 393–399.
Nabholz, B., Glémin, S. and Galtier, N. 2009. The erratic mitochondrial clock: variations of mutation rate, not population size, affect mtDNA diversity across birds and mammals. – BMC Evol. Biol. 9: 54.
Ohta, T. 1992. The nearly neutral theory of molecular evolution. – Annu. Rev. Ecol. Syst. 23: 263–286.
Pacheco, M. A., Battistuzzi, F. U., Lentino, M., Aguilar, R. F., Kumar, S. and Escalante, A. A. 2011. Evolution of modern birds revealed by mitogenomics: timing the radiation and origin of major orders. – Mol. Biol. Evol. 28: 1927–1942.
Patané, J. S. L., Weckstein, J. D., Aleixo, A. and Bates, J. M. 2009. Evolutionary history of *Ramphastos* toucans: molecular phylogenetics, temporal diversification, and biogeography. – Mol. Phylogenet. Evol. 53: 923–934.
Paxinos, E. E., James, H. F., Olson, S. L., Sorenson, M. D., Jackson, J. and Fleischer, R. C. 2002. mtDNA from fossils reveals a radiation of Hawaiian geese recently derived from the Canada goose (*Branta canadensis*). – Proc. Natl Acad. Sci. USA 99: 1399–1404.
Pereira, S. L. and Baker, A. J. 2006. A mitogenomic timescale for birds detects variable phylogenetic rates of molecular evolution and refutes the standard molecular clock. – Mol. Biol. Evol. 23: 1731–1740.
Peterson, G. I. and Masel, J. 2009. Quantitative prediction of molecular clock and K_a/K_s at short timescales. – Mol. Biol. Evol. 26: 2595–2603.

Supplementary material (Appendix JAV-00766 at <www.avianbiology.org/appendix/jav-00766>). Appendix 1–2.

Quinn, G. P. and Keough, M. J. 2002. Experimental design and data analysis for biologists. – Cambridge Univ. Press.
Ratnasingham, S. and Hebert, P. D. N. 2007. BOLD: the Barcode of Life Data System (<www.barcodinglife.org>). – Mol. Ecol. Notes 7: 355–364.
Shields, G. F. and Wilson, A. C. 1987. Calibration of mitochondrial DNA evolution in geese. – J. Mol. Evol. 24: 212–217.
Subramanian, S., Denver, D. R., Millar, C. D., Heupink, T., Aschrafi, A., Emslie, S. D., Baroni, C. and Lambert, D. M. 2009. High mitogenomic evolutionary rates and time dependency. – Trends Genet. 25: 482–486.
Swofford, D. L. 2001. PAUP*. Phylogenetic analysis using parsimony (*and other methods). Version 4. – Sinauer Associates.
Tamura, K., Peterson, D., Peterson, N., Stecher, G., Nei, M. and Kumar, S. 2011. MEGA5: molecular evolutionary genetics analysis using maximum likelihood, evolutionary distance, and maximum parsimony methods. – Mol. Biol. Evol. 28: 2731–2739.
Tavares, E. S., Gonçalves, P., Miyaki, C. Y. and Baker, A. J. 2011. DNA barcode detects high genetic structure within Neotropical bird species. – PLoS One 6: e28543.
Weir, J. T. and Schluter, D. 2008. Calibrating the avian molecular clock. – Mol. Ecol. 17: 2321–2328.
Weir, J. T. and Price, M. 2011. Andean uplift promotes lowland speciation through vicariance and dispersal in *Dendrocincla* woodcreepers. – Mol. Ecol. 20: 4550–4563.
Woodhams, M. 2006. Can deleterious mutations explain the time dependency of molecular rate estimates? – Mol. Biol. Evol. 23: 2271–2273.