Mitochondrial substitution rates estimation for molecular clock analyses in modern birds based on full mitochondrial genomes

Authors: Angel Arcones^{1,2,3*}, Raquel Ponti^{1,2,3}, David R. Vieites^{1*}

1 Department of Biogeography and Global Change. Museo Nacional de Ciencias Naturales –CSIC. C/José Gutiérrez Abascal 2, 28006, Madrid, Spain.

2 Department of Evolutionary Biology, Ecology, and Environmental Sciences, Faculty of Biology; Universitat de Barcelona, Spain.

3 Biodiversity Research Institute (IRBIO), Universitat de Barcelona. Av. Diagonal 643, 08028 Barcelona, Spain.

*corresponding authors: angelarcones@gmail.com, vieites@mncn.csic.es

Abstract

Mitochondrial DNA (mtDNA) is a very popular resource in the study of evolutionary processes in birds, and especially to infer divergence times between lineages. These inferences rely on rates of substitution in the mtDNA genes that, ideally, are specific for the studied taxa. But as such values are often unavailable many studies fixed rate values generalised from other studies, such as the popular “standard molecular clock”. However the validity of these universal rates across all bird lineages and for the different mtDNA has been severely questioned. Thus, we here performed the most comprehensive calibration of the mtDNA molecular clock in birds, with the inclusion of complete mitochondrial genomes for 622 bird species and 25 reliable fossil calibrations. The results show variation in the rates between lineages and especially between markers, contradicting the universality of the standard clock. Moreover, we provide specific rates for every mtDNA marker (except D-loop) in each of the sampled avian orders, which should help improve future estimations of divergence times between bird species or populations.

Introduction

Mitochondrial DNA (mtDNA) is a useful and accessible source of genetic data to explore the recent evolutionary histories of species. Its extended use has been the cornerstone of many phylogeographic studies on bird species (Avice & Walker, 1998; Zink & Barrowclough, 2008), although it has also raised concerns around the reliability of specific methodologies (Ballard & Whitlock, 2004; Ho, 2007; Edwards & Bensch, 2009).

One of the most popular uses of mtDNA in ornithology is the estimation of divergence times. Given the poor fossil record for many avian families and orders (Mayr, 2009), mtDNA mutation rates have been primarily used to estimate divergence times between lineages via molecular clock analyses. However, a large proportion of the studies based their estimations on the use of a “universal” standard molecular clock. This popular rate proposes a fixed value of 2% divergence per million years between any two lineages, which translates into approximately 0.01 substitutions per site per lineage per million years (s/s/l/My). This rate was initially estimated based on data from humans and chimpanzees (Brown *et al.*, 1979) and was supported

by similar rate estimations for other vertebrates (Wilson *et al.*, 1985) and a couple genera of geese (Shields & Wilson, 1987). More recently, a reanalysis extended the applicability of this general rate to all birds (Weir & Schluter, 2008). This standard rate has served as a foundation of many studies involving dating divergence events on birds, such as the study of the intra and interspecific diversification during the Pleistocene (Klicka & Zink, 1997; Avise & Walker 1998; Johnson & Cicero, 2004, Weir & Schluter, 2004).

Despite of its widespread use, the validity of this standard molecular clock for any mtDNA loci in any bird lineage has been heavily criticized (Garcia-Moreno, 2004; Lovette, 2004; Pereira & Baker, 2006; Ho, 2007). Previous analyses of the mtDNA of birds highlight a lack of “standard” rates, finding variation between the different mtDNA genes as well as across different avian lineages (Pereira & Baker, 2006; Pacheco *et al.*, 2011; Lavinia *et al.*, 2016; Nabholz *et al.*, 2016; Nguyen & Ho, 2016). Some of these studies also attempted to calibrate the molecular rates of the different protein-coding mitochondrial genes across the avian phylogeny using complete mitochondrial genomes (Pereira & Baker, 2006; Pacheco *et al.*, 2011; Nabholz *et al.*, 2016). However, these studies show high variation between their results, and more importantly, they often fail to provide rates that are specific for each mtDNA marker in each lineage of the tree. This is of particular importance as proper divergence time estimations rely on using substitution rates as specific as possible in each study case (Garcia-Moreno, 2004; Ho *et al.*, 2007, 2015).

To overcome this problem, we here performed a comprehensive mitogenomic calibration of the avian molecular clock to obtain substitution rates for each mtDNA gene across the phylogenetic tree of avian species. We hope that these new rates will help future evolutionary and phylogeographic studies on birds to improve their molecular clock dating analyses and better understand the evolutionary processes within the different taxa of interest.

67

68

69 **Materials and methods:**

70

71 Complete mitochondrial genomes

72 We assembled a comprehensive dataset of complete bird mitochondrial genomes
73 available from Genbank by November 2016. To retrieve the genomes, we modified the
74 Mitobank script (Abascal *et al.*, 2007) in BioPerl, providing complete or nearly complete bird
75 genomes for 622 species representing 33 modern out of 38 bird orders, including some extinct
76 species. Of these 622 species sampled 17 belonged to palaeognaths (Ostrich, tinamous, and

allies), 94 to the Galloanserae clade (chickens, quails, geese, fowls and allies) and 511 to the Neoaves (all other avian groups), of which 259 were Passeriformes (Accession Numbers: see Supplementary material). Only five orders were left unrepresented due to the lack of species' mt genome available data: Mesitornithiformes (Mesites), Pterocliiformes (Sandgrouses and allies), Opisthocomiformes (Hoatzin), Leptosomiformes (Cuckoo rollers) and Cariamiformes (Seriemas and allies).

In two cases where the original genomes lack one locus, those were completed with sequence data from other individuals of the same species: *Psittacus erithacus* for ND6 (accession number KM611474) and *Accipiter gularis* for ND1 (EU583261).

After excluding the D-loop and the tRNAs, the result was a combined dataset of 12950 base pairs. Alignments were done gene by gene as global alignments are not possible because of gene rearrangements and complexity of the genomes (Mueller & Boore, 2005). The alignment was performed using the software Genious v.9 (<https://www.geneious.com>). Due to the high variability of some regions of the mitochondrial DNA, and especially in 3rd codon positions, we applied a translation alignment. This implies translating the sequences into aminoacids for the alignment and then returning to the original nucleotide sequence. The genes were then concatenated into a single sequence per species.

We computed a phylogenetic tree using RaxML (Stamatakis, 2014) as a test to detect troublesome sequences that could be introducing errors in the process. This resulted in the exclusion of *Larus vegae* (GenBank accession number: NC_029383) from further analyses, as the sequence does not belong to a gull.

98

99 Fossil time constraints

In order to accurately determine divergence times and evolutionary rates, analyses must include many independent fossil constraints to time-calibrate the phylogenetic tree (Near & Sanderson, 2004; Benton & Donoghue, 2007; Magallón *et al.*, 2013). In our case, these calibrations are based on the fossil record. Following recent recommendations (Ho & Duchêne, 2014; Zheng & Wiens, 2015), we aimed to include the higher number of calibrations as possible across the tree to maximize the accuracy of our results. We included 25 calibration constraints based on their informative value and compatibility with our dataset. Following the recommendations from Parham *et al.* (2012) and Fourment & Holmes (2014), we used a conservative minimum age based on the chronostratigraphic evidence of each fossil, and we did not impose any hard-maximum ages. The calibrated nodes with their corresponding minimum ages are: (1) split *Corvus* – *Urocisa*, 13.6 Mya (Scofield *et al.*, 2017); (2) split Psittaciformes –

111 Passeriformes, 53.5 Mya (Ksepka & Clarke, 2015); (3) crown Meleagridinae + Tetraoninae, 16
 112 Mya ptarmigan (Stein *et al.*, 2015); (4) crown Galliformes, 51.6 Mya (Mayr & Weidig, 2004);
 113 (5) Crown Cracidae + Numididae + Phasianidae, 33.9 Mya (van Tuinen & Dyke, 2004); (6)
 114 crown *Gallus* + *Coturnix*, 23.03 Mya (van Tuinen & Dyke, 2004); (7) Crown Anatidae, 66.5
 115 Mya (Ksepka & Clarke, 2015); (8) split Apodidae – Trochilidae, 53 Mya (Harrison, 1984); (9)
 116 crown Trochilidae, 30 Mya (Mayr, 2004); (10) split Charadriiformes – sister taxa, 46.5 Mya
 117 (Mayr, 2000); (11) crown Laromorphae, 23.6 Mya (Smith, 2015); (12) split Jacanidae – other
 118 Scolopacids, 30 Mya (Smith, 2015); (13) split Pan-Alcidae – Stercorariidae, 34.2 (Smith, 2015);
 119 (14) crown Calidrinae, 11.62 (Smith, 2015); (15) split Phoenicopteriformes – Podicipediformes,
 120 27.8 Mya (Mayr & Smith, 2002); (16) split Fregatidae – Suloidea, 51.8 Mya (Smith & Ksepka,
 121 2015); (17) split Phalacrocoracidae – Anhingidae, 24.5 Mya (Smith & Ksepka, 2015); (18)
 122 crown Balaenicipitidae, 30 Mya (Smith, 2013); (19) split Sphenisciformes – Procellariiformes,
 123 60.5 Mya (Ksepka & Clarke, 2015); (20) split *Spheniscus* – *Eudyptula*, 11 Mya (Göhlich,
 124 2007); (21) crown Uppupidae + Phoeniculidae, 46.6 Mya (Mayr, 2006; Ksepka & Clarke,
 125 2015); (22) split Trogoniformes – sister taxa, 49 Mya (Mayr, 2005); (23) split Falconidae –
 126 Polynorinae, 15.9 Mya (Becker, 1987); (24) crown Columbiformes, 20.3 Mya (Gradstein &
 127 Ogg, 2004); (25) split *Casuarius* – *Dromaius*, 23 Mya (Boles, 1992).

128 Rate estimates

129 We performed Bayesian analyses as implemented in BEAST2 v.2.4.7 (Bouckaert *et al.*,
 130 2014) to estimate the substitution rate per locus. The choice of this software was based on its
 131 capability to accommodate different parameters into the analysis (such as substitution models
 132 for each codon position and multiple options of relaxed molecular clock), as well as the
 133 capability of recovering a rate for each gene in all the nodes of the tree. We set independent
 134 partitions for every gene to estimate the molecular clock rate per locus, and for every codon
 135 position to fit the nucleotide substitution models selected. We used PartitionFinder (Lanfear *et*
 136 *al.*, 2012) to estimate the best substitution model for each codon position in each gene.

137 We selected the lognormal relaxed clock implemented in BEAST2. In order to avoid
 138 misleading results due to clock selection, we also ran an analysis under a random local clock
 139 model to compare between results. Due to the fast rate of change in mtDNA, saturation becomes
 140 an issue affecting the topology at deeper nodes. To minimize this problem, we constrained the
 141 tree topology by grouping species' sequences into their corresponding orders, as well as the
 142 phylogenetic relationships between orders based on the results from Prum *et al.* (2015).
 143 Relationships between Galliformes families were also constrained to fit that topology. The
 144 phylogenetic tree was computed using all genes combined, following a birth-death Yule model.

Given the nature of the dataset, with an extensive taxon sampling and basepairs, the MCMC chain needed a large number of generations in order to reach stability. We ran the chain for 1 billion (1×10^9) generations during several months, discarding the first 500 million as burnin and sampling every 5000 states. To confirm that the previous results were not stuck on a local maximum, we ran a second 6×10^8 independent chain to compare with. The random local clock analysis also ran for 5×10^8 generations. Additionally, we ran an analysis using only the information from the priors, to determine whether the posterior probability of the results is driven solely by the priors, or if the data are also introducing information into the final results.

We also ran an analysis based on aminoacids and using the same specifications as in the nucleotide analyses, with substitution model MTREV for each of the partitioned markers as implemented in BEAST2. Aminoacid sequences are more conserved than nucleotide sequences, allowing an assessment of the potential effect of mtDNA saturation of nucleotides in the topology and divergence time estimates.

Results

Molecular rate estimates

BEAST analyses provided a fully resolved tree topology with most nodes recovered as highly supported (>98% of nodes above 0.9 posterior probability). The Maximum Clade Credibility method implemented in TreeAnnotator forcefully prevents polytomies, but no signs of conflict were found in the resolution of any of the nodes of the tree. Results from the sample from prior differed from those from the main analyses, indicating that the recovered posterior probabilities are not just a product of the priors. The random clock analysis recovered a phylogeny where most of the nodes of the tree were younger than 1 Mya. Thus, we considered this method unreliable and discarded it as an alternative. The topology of the amino acid tree (not shown) provided similar results to the obtained from nucleotides, except minor differences at the position of particular tips, affecting less than 5% of the species.

The resulting phylogenetic tree had overall high support, with over 98% of the nodes showing a posterior probability above 0.9. We recovered a root age (Palaeognatha / Neognatha split) of 88.3Mya (95% CI = 90.9 – 85.4 Mya). The split between Neoaves and Galloanserae was recovered at 87.9 Mya (90.7 – 85.3 Mya). The divergence between all extant avian orders occurred before or around the transition between the Cretaceous and the Paleogene, 66Mya, and most of their diversification took place during the Paleogene and Neogene.

178 In general, the fossil-calibrated nodes returned ages several million years older than the
179 prior minimum age of the fossils. The only two exceptions were the calibrations for the
180 Sphenisciformes / Procellariiformes split (minimum age = 60.5Mya) and the *Dromaius* /
181 *Casuarius* split (min. age = 23.03Mya). In both cases, the posterior distribution was constrained
182 by the hard minimum bound of the prior, but still displayed a normal distribution and the mean
183 ages were slightly older than the minimum constraint (not shown).

184 Our approach allowed us to obtain substitution rates for each gene across each node of the
185 tree. We found variation in the rates between genes, and to a lesser extent, between lineages in
186 the tree in each gene. The highly-conserved ribosomal genes 12S and 16S showed the lowest
187 mean substitution rate over the whole tree (0.00107 and 0.00043 s/s/l/My, respectively), while
188 ND2, ATP6, and ATP8 returned the highest overall rate values (0.00227, 0.00215 and 0.00260
189 s/s/l/My, respectively) (Fig. 1).

190 Overall, the rate values in each gene showed minimal variation across the branches of the
191 tree, and there were no significant differences in the overall mitochondrial rates between orders
192 (Fig. 2). Only the order Otidiformes showed values distributed mostly below the overall
193 average, although this order is represented by just a single species in the sampling (*Otis tarda*),
194 and could be affected by a punctual underestimation of the rates due to insufficient taxon
195 coverage. The values of the rates per gene for the avian orders sampled can be found in the
196 Supplementary material.

197

198 Discussion

199

200 Our analyses represent the most comprehensive calibration of the mitochondrial
201 molecular clock performed to date in birds, both in terms of species coverage and fossil
202 calibrations used. We provide the first lineage-specific rates for each of the mtDNA genes for
203 almost all of the extant bird orders (Fig. 2, Supplementary material).

204 The results from previous mitogenomic studies (Pereira & Baker, 2006; Pacheco *et al.*,
205 2011; Nabholz *et al.*, 2016) showed significant variability in the recovered evolutionary rates. In
206 all cases they found rate values lower than the so-called standard molecular clock of 0.01
207 s/s/l/My (Fig 3). The substitution rate values obtained by us fall within the range of the variation
208 observed between previous works (Fig. 3). Moreover, the differences in mean rate values for the
209 mitochondrial genes in our study follow a similar pattern, although with lower value than the
210 rates obtained by Nabholz *et al.* (2016) (Fig. 3, blue line), which until now was the most
211 taxonomic-rich and phylogenetically updated analysis.

The differences in mutation rates between studies can be the consequence of several causes, from the inclusion or elimination of third codon positions that are generally saturated at deeper phylogenetic levels, to the impact of hard maximum bounds in fossil time constraints that force younger ages in the phylogeny (Ho *et al.*, 2005, 2007, 2008), and even the choice of software or other specific parameters could be introducing variation (Drummond & Rambaut, 2015). Given the variety of methodologies that have been used in previous mitogenomic rate estimations for birds (Pereira & Baker, 2006; Pacheco *et al.*, 2011; Nabholz *et al.*, 2016), we are unable to assess whether the differences between our results and those from previous studies are caused by one or several of the mentioned causes, or by any other cause. In any case, we are confident that given the more extensive taxonomic coverage, the more and reliable fossil calibrations included, and the methodological approach (e.g. no maximum-time constraints, very long Bayesian MCMC runs until reaching true stabilization), our results provide a reliable estimation of the molecular clock substitution rates in birds.

Our results, along with those from previous analyses, contradict the validity of the standard mitochondrial molecular clock (Pereira & Baker, 2006; Pacheco *et al.*, 2011; Nabholz *et al.*, 2016) and support the criticism against its generalized use (Garcia-Moreno, 2004; Lovette, 2004; Ho, 2007, Lavinia *et al.*, 2016). These new substitution rates per locus provided in this study should improve the accuracy and reliability of future molecular clock analyses, making available rates for more clades that can be used instead of standard rates. We expect that the implementation of reliable well-calibrated rates will significantly benefit future research on the recent evolutionary history of bird groups without fossils to time-calibrate their genealogical or phylogenetic relationships.

References

- Abascal, F., Posada, D., & Zardoya, R. (2007). MtArt: A new model of amino acid replacement for Arthropoda. *Molecular Biology and Evolution*, 24(1), 1–5.
- Awise, J., & Walker, D. (1998). Pleistocene phylogeographic effects on avian populations and the speciation process. *Proceedings of the Royal Society B: Biological Sciences*, 265, 457–463.
- Becker, J. J. (1987b). Revision of "*Falco*" *ramenta* Wetmore and the Neogene evolution of the Falconidae. *The Auk*, 104(2), 270–276.
- Benton, M. J., & Donoghue, P. C. J. (2007). Paleontological evidence to date the tree of life. *Molecular Biology and Evolution*, 24(1), 26–53.
- Boles, W. E. (1992). Revision of *Dromaius gidju* (Patterson & Rich, 1987) from Riversleigh, northwestern Queensland, Australia, with a reassessment of its generic position. *Natural History Museum of LA County Science Series*, 36, 195–208.
- Bouckaert, R., Heled, J., Kühnert, D., Vaughan, T., Wu, C. H., Xie, D., ... & Drummond, A. J. (2014). BEAST 2: a software platform for Bayesian evolutionary analysis. *PLoS computational biology*, 10(4), e1003537.
- Brown, W. M., George, M., & Wilson, A. C. (1979). Rapid evolution of animal mitochondrial DNA.

250 *Proceedings of the National Academy of Sciences of the United States of America*, 76(4), 1967–
251 1971.

252 Drummond, A. J., & Rambaut, A. (2015). *Bayesian evolutionary analysis with BEAST 2*. Cambridge
253 University Press.

254 Fourment, M., & Holmes, E. (2014). Novel non-parametric models to estimate evolutionary rates and
255 divergence times from heterochronous sequence data. *BMC Evolutionary Biology*, 14(1), 163.

256 Garcia-Moreno, J. (2004). Is there a universal mtDNA clock for birds? *Journal of Avian Biology*, 35(6),
257 465–468.

258 Gradstein, F., & Ogg, J. (2004). Geologic time scale 2004—why, how, and where next!. *Lethaia*, 37(2),
259 175–181.

260 Harrison, C. J. O. (1984). A revision of the fossil swifts (Vertebrata, Aves, Suborder, Apodi), with
261 descriptions of three new genera and two new species. *Mededelingen van de Werkgroep voor*
262 *Tertiaire en Kwartaire Geologie*, 21(4), 157–177.

263 Ho, S. Y.W., (2007). Calibrating molecular estimates of substitution rates and divergence times in birds.
264 *Journal of Avian Biology*, 38(4), 409–414.

265 Ho, S. Y.W., & Duchêne, S. (2014). Molecular-clock methods for estimating evolutionary rates and
266 timescales. *Molecular Ecology*, 23(24), 5947–5965.

267 Ho, S. Y., Duchêne, S., Molak, M., & Shapiro, B. (2015). Time-dependent estimates of molecular
268 evolutionary rates: evidence and causes. *Molecular ecology*, 24(24), 6007–6012.

269 Ho, S. Y. W, Phillips, M. J., Cooper, A., & Drummond, A. J. (2005). Time dependency of molecular rate
270 estimates and systematic overestimation of recent divergence times. *Molecular Biology and*
271 *Evolution*, 22(7), 1561–1568.

272 Ho, S. Y.W., Saarma, U., Barnett, R., Haile, J., & Shapiro, B. (2008). The effect of inappropriate
273 calibration: Three case studies in molecular ecology. *PLoS ONE*, 3(2).

274 Ho, S. Y. W, Shapiro, B., Phillips, M. J., Cooper, A., & Drummond, A. J. (2007). Evidence for Time
275 Dependency of Molecular Rate Estimates. *Systematic Biology*, 56(3), 515–522.

276 Johnson, N. K., & Cicero, C. (2004). New mitochondrial DNA data affirm the importance of Pleistocene
277 speciation in North American birds. *Evolution*, 58(5), 1122–1130.

278 Klicka, J., & Zink, R. M. (1997). The importance of recent ice ages in speciation: a failed paradigm.
279 *Science*, 277(5332), 1666–1669.

280 Ksepka, D., & Clarke, J. (2015). Phylogenetically vetted and stratigraphically constrained fossil
281 calibrations within Aves. *Palaeontologia Electronica*, 18(1), 1–25.

282 Lanfear, R., Calcott, B., Ho, S. Y., & Guindon, S. (2012). PartitionFinder: combined selection of
283 partitioning schemes and substitution models for phylogenetic analyses. *Molecular biology and*
284 *evolution*, 29(6), 1695–1701.

285 Lavinia, P. D., Kerr, K. C., Tubaro, P. L., Hebert, P. D., & Lijtmaer, D. A. (2016). Calibrating the
286 molecular clock beyond cytochrome b: assessing the evolutionary rate of COI in birds. *Journal of*
287 *Avian Biology*, 47(1), 84–91.

288 Lovette, I. J. (2004). Mitochondrial dating and mixed support for the “2% rule” in birds. *The Auk*, 121(1),
289 1–6.

290 Magallón, S., Hilu, K. W., & Quandt, D. (2013). Land plant evolutionary timeline: Gene effects are
291 secondary to fossil constraints in relaxed clock estimation of age and substitution rates. *American*
292 *Journal of Botany*, 100(3), 556–573.

293 Mayr, G. (2000). Charadriiform birds from the Early Oligocene of Céreste (France) and the Middle
294 Eocene of Messel (Hessen, Germany). *Geobios*, 33, 625-636

295 Mayr, G. (2004). Old World fossil record of modern-type hummingbirds. *Science*, 304(5672), 861-864.

296 Mayr, G. (2005). New trogons from the early Tertiary of Germany. *Ibis*, 147, 512-518.

297 Mayr, G. (2009). *Paleogene fossil birds*. Springer Science & Business Media.

298 Mayr, G., & Weidig, I. (2004). The Early Eocene bird *Gallinuloides wyomingensis*—a stem group
299 representative of Galliformes. *Acta Palaeontologica Polonica*, 49(2), 211–217.

300 Mayr G, Smith R. (2002) Avian remains from the lowermost Oligocene of Hoogbutsel (Belgium).
301 *Bulletin de l'Institut Royal des Sciences Naturelles de Belgique Sciences de la Terre*, 72, 139–150.

302 Mueller, R. L., & Boore, J. L. (2005). Molecular mechanisms of extensive mitochondrial gene
303 rearrangement in Plethodontid salamanders. *Molecular Biology and Evolution*, 22(10), 2104–2112.

304 Nabholz, B., Lanfear, R., & Fuchs, J. (2016). Body mass-corrected molecular rate for bird mitochondrial
305 DNA. *Molecular Ecology*, 4438–4449.

306 Near, T. J., & Sanderson, M. J. (2004). Assessing the quality of molecular divergence time estimates by
307 fossil calibrations and fossil-based model selection. *Philosophical Transactions of the Royal Society*
308 *B: Biological Sciences*, 359(1450), 1477–1483.

309 Nguyen, J. M., & Ho, S. Y. (2016). Mitochondrial rate variation among lineages of passerine birds.
310 *Journal of Avian Biology*, 47(5), 690-696.

311 Pacheco, M. A., Battistuzzi, F. U., Lentino, M., Aguilar, R. F., Kumar, S., & Escalante, A. A. (2011).
312 Evolution of modern birds revealed by mitogenomics: Timing the radiation and origin of major
313 orders. *Molecular Biology and Evolution*, 28(6), 1927–1942.

314 Parham, J. F., Donoghue, P. C. J., Bell, C. J., Calway, T. D., Head, J. J., Holroyd, P. A., ... Benton, M. J.
315 (2012). Best practices for justifying fossil calibrations. *Systematic Biology*, 61(2), 346–359.

316 Pereira, S. L., & Baker, A. J. (2006). A mitogenomic timescale for birds detects variable phylogenetic
317 rates of molecular evolution and refutes the standard molecular clock. *Molecular Biology and*
318 *Evolution*, 23(9), 1731–1740.

319 Prum, R. O., Berv, J. S., Dornburg, A., Field, D. J., Townsend, J. P., Lemmon, E. M., & Lemmon, A. R.
320 (2015). A comprehensive phylogeny of birds (Aves) using targeted next-generation DNA
321 sequencing. *Nature*, 526, 569–573.

322 Scofield, R. P., Mitchell, K. J., Wood, J. R., De Pietri, V. L., Jarvie, S., Llamas, B., & Cooper, A. (2017).
323 The origin and phylogenetic relationships of the New Zealand ravens. *Molecular phylogenetics*
324 *and evolution*, 106, 136-143

325 Shields, G. F., & Wilson, A. C. (1987). Calibration of Mitochondrial DNA Evolution in Geese. *Journal of*
326 *Molecular Evolution*, 24, 212–217.

327 Smith, N.D. & Ksepka, D.T. (2015) Five well-supported fossil calibrations within the “Waterbird”
328 assemblage (Tetrapoda, Aves). *Palaeontologia Electronica*, 18, 1–21.

329 Stein, R. W., Brown, J. W., & Mooers, A. Ø. (2015). A molecular genetic time scale demonstrates
330 Cretaceous origins and multiple diversification rate shifts within the order Galliformes (Aves).
331 *Molecular phylogenetics and evolution*, 92, 155-164.

332 Stamatakis, A. (2014). RAxML Version 8: A tool for Phylogenetic Analysis and Post-Analysis of Large
333 Phylogenies. *Bioinformatics*, 30 (9), 1312-1313

334 van Tuinen, M. and G. J. Dyke (2004). Calibration of galliform molecular clocks using multiple fossils
335 and genetic partitions. *Molecular Phylogenetics and Evolution* 30(1): 74-86.

- 336 Weir, J. T., & Schluter, D. (2004). Ice sheets promote speciation in boreal birds. *Proceedings. Biological*
337 *Sciences / The Royal Society*, 271(1551), 1881–1887.
- 338 Weir, J. T., & Schluter, D. (2008). Calibrating the avian molecular clock. *Molecular Ecology*, 17(10),
339 2321–2328.
- 340 Wilson, A. C., Cann, R. L., Carrii, S. M., George, M., Gyllenstein, U. L. F. B., Kathleen, M., ... Sage, R.
341 D. (1985). Mitochondrial DNA and two perspectives on evolutionary genetics. *Biological Journal*
342 *of the Linnean Society*, 26, 375–400.
- 343 Zheng, Y., & Wiens, J. J. (2015). Do missing data influence the accuracy of divergence-time estimation
344 with BEAST? *Molecular Phylogenetics and Evolution*, 85, 41–49.
- 345 Zink, R. M., & Barrowclough, G. F. (2008). Mitochondrial DNA under siege in avian phylogeography.
346 *Molecular Ecology*, 17(9), 2107–2121.

347

348

349

350 Figures

351

352

353

354 Figure 1: boxplot of the substitution rate values (in substitutions / site / lineage / million years) from every
355 node of the mitogenomic tree, for each of the 15 analyzed mitochondrial genes .

356

357

358 Figure 2: boxplot of the average substitution rate values (in substitutions / site / lineage / million years) of
359 all 15 mitochondrial genes for each of the orders of birds.

360

361

362 Figure 3: Comparison of the mean substitution rate estimates for each mitochondrial gene, between our
363 results (black line) and previous mitogenomic calibrations of the molecular clock, with the number of
364 species and fossil time constraints included in each analysis.

365