Title

2 Skeletal trait measurements for thousands of bird species

3

Authors

5 Brian C. Weeks[1], Zhizhuo Zhou[2], Charlotte M. Probst[1], Jacob S. Berv[1,3], Bruce O'Brien[1], Brett W.
6 Benz[4], Heather R. Skeen[5,6], Mark Ziebell[1], Louise Bodt[7], David F. Fouhey[8,9]

7

Affiliations

9

10 1. School for Environment and Sustainability, University of Michigan; Michigan, USA
11 2. School of Computer Science, Carnegie Mellon University; Pennsylvania, USA
12 3. Michigan Institute for Data Science, University of Michigan; Michigan, USA
13 4. Department of Ecology and Evolutionary Biology and Museum of Zoology, University of
14 Michigan; Michigan, USA
15 5. Ecology and Evolutionary Biology, University of Connecticut; Connecticut, USA
16 6. Negaunee Integrative Research Center, Field Museum of Natural History; Illinois, USA
17 7. Committee on Evolutionary Biology, University of Chicago; Illinois, USA
18 8. Computer Science, Courant Institute of Mathematical Sciences, New York University; New
19 York, USA
20 9. Electrical & Computer Engineering, Tandon School of Engineering, New York University; New
21 York, USA

22

23 Corresponding author: Brian Weeks (bcweeks@umich.edu)

24

Abstract

26 Large comparative datasets of avian functional traits have been used to address a wide range
27 of questions in ecology and evolution. To date, this work has been constrained by the limited
28 availability of skeletal trait datasets that include extensive inter- and intra-specific sampling.
29 We use computer vision to identify and measure bones from photographs of museum skeletal
30 specimens to assemble an extensive dataset of functionally important skeletal elements in
31 birds. The dataset spans 2,057 species of birds (Aves: Passeriformes) and includes
32 measurements of 12 skeletal elements from 14,419 individuals. In addition to the trait values
33 directly measured from photographs, we leverage the multi-dimensional nature of our dataset
34 and known phylogenetic relationships of the species to impute missing data under an
35 evolutionary model. To facilitate use of the dataset, the taxonomy has been reconciled with
36 an existing comprehensive avian phylogeny and an additional dataset of external functional
37 traits for all birds.

38

Background & Summary

40

41 Understanding large-scale patterns of diversity, and the ecological and evolutionary origins
42 and consequences of these patterns, is of growing interest. These efforts have historically
43 been constrained by the limited availability of comparative quantitative trait datasets at large
44 spatial and taxonomic scales[1]. As large datasets have become available, they have stimulated
45 significant advances in macroecology and macroevolution.

46

47 While large-scale trait datasets are increasingly available across a range of taxa (e.g., vascular
48 plants[2], lizards[3], and freshwater fish[4]), birds are a model system in macroecology and
49 macroevolution, with their well-known distributions[5], extinction risks[5], phylogenetic
50 relationships[6,7], ecological niches[8,9], life history strategies[10], nesting biologies[11], and external
51 morphologies[12]. These diverse datasets have been integrated to answer a wide range of
52 questions spanning ecology and evolution (e.g.[13–17]) and are increasingly being used to

understand human impacts on natural systems (e.g.[18–22]). Although much has been learned from existing large-scale datasets, in animals, the availability of trait data spanning multiple anatomical systems would open new avenues of research and could allow for more mechanistic understanding of morphological patterns. Bird skeletons, which are well-represented in natural history collections, present an underutilized opportunity to develop such a dataset.

The accumulation of comparative skeletal trait data for many traits and across many species and individuals has lagged far behind the generation of data from the measurement of external traits[23]. In birds, aspects of the skeleton provide key insights into bird locomotion[24], the physics of flight[25,26], directional evolution[27], phylogenetic relationships[28], and responses to environmental change[29], and are often used to better understand the morphologies of fossil birds[30]. Further, the utility of skeletal traits expands significantly when they can be studied in conjunction with other types of phenotypic data. For example, although the tendency for appendages to be longer in warmer climates (i.e., Allen's Rule[31]) is a classic pattern in macroecology and has been the focus of intensive research for over a century, the integration of skeletal and plumage trait data revealed a novel morphological trend that generated new insights into the mechanistic basis underlying Allen's Rule[32]. As such, the availability of comparative skeletal data may provide novel insight into macro-scale patterns in avian morphology and improved mechanistic understanding of the drivers of bird morphological variation across space and time.

This dataset encompasses a large portion of the diversity within Passeriformes, the most diverse order of Neornithes (modern, living birds). The 2,057 species in the dataset comprise 34% of passerine species and represent 89% of passerine families[6]. The sampling is also spatially expansive and includes specimens from all continents except Antarctica (Fig 1). Multiple individuals were measured per species when possible, resulting in a dataset that includes 14,419 individuals. We targeted twelve skeletal elements for each specimen. Combined with our estimates for missing values our dataset includes 173,028 unique values. The data are presented in three formats: 1) a specimen-level dataset that only includes trait values that were directly measured, 2) a specimen-level dataset with no missing data that includes both the directly measured trait values and imputed trait values, and 3) a complete species-level dataset derived by applying a multivariate evolutionary model. The taxonomy in the datasets has been unified to the Birdlife Version 3 taxonomy to facilitate integration with existing largescale datasets and to simplify conversion to other widely used taxonomies using recently published taxonomic crosswalks[12]. As such, it should be straightforward to integrate our data with data on the phylogenetic history of birds[6], bird range maps[5], IUCN threat statuses[5], and existing comprehensive external trait data[12]. Importantly, the methods used to generate these data are open source[23] and easily applied, enabling future expansion of the dataset.

Methods

Sampling

The large majority of our data come from skeleton specimens held in the University of Michigan Museum of Zoology (UMMZ), one of the largest and most diverse bird skeletal collections in the world, where we effectively photographed and measured the entirety of the UMMZ's passerine skeletal collection (N = 12,421, number of species = 1,881). We supplemented this dataset with specimens at the Field Museum of Natural History (FMNH; N = 1,998, number of species = 438), targeting families that are well represented in both the UMMZ and FMNH collections, with an emphasis on species found in the Neotropics.

104   Ultimately, we photographed 14,419 specimens spanning 2,057 species, from 86 families (Fig.
105   1). Every trait measurement has an associated specimen catalogue number that can be used
106   to link each measurement to the specimen, housed in UMMZ or FMNH.

107   *Photographing*

108   Trait measurements were generated using Skelevision[23], a deep neural network-based
109   approach for identifying and measuring skeletal elements in photographs of bird skeleton
110   specimens. In this approach, museum skeleton specimens are first removed from their
111   containers and spread randomly on a standard background, except for the keel and the skull,
112   which are consistently oriented to display their profile. They are then photographed from a
113   fixed distance before being returned to their boxes. Each specimen is photographed
114   individually, independently, and in its entirety.

115   The photographs were taken with the same imaging equipment that was used in the
116   Skelevision methods paper[23]. All images were collected from ~400mm above the specimen,
117   using a SONY IX183 sensor on a FLIR Blackfly S camera. This generated photographs with a
118   pixel size of 0.07mm.

119   *Trait Measurement*

120   We applied the Skelevision method[23] for segmenting, identifying, and measuring
121   target bones to each photograph. This method integrates a U-Net[33] and Mask R-CNN[34] trained
122   on images annotated by hand[35,36] to identify pixels in the images that are bone, determine
123   which element the pixels belong to, and then segment the elements. The pipeline then takes
124   segmented masks from the images for all elements and measures their longest linear
125   dimension (i.e., the longest linear length of each element)[23]. We use this method to measure
126   12 traits: the lengths of the tibiotarsus, humerus, tarsometatarsus, ulna, radius, keel,
127   carpometacarpus, 2nd digit 1st phalanx, furcula, and femur; the maximum outer diameter of
128   the sclerotic ring, and the length from the back of the skull to the tip of the bill (treating the
129   rhamphotheca as part of the bill when it remains present on the specimen).

130   Skelevision estimates the probability that Skelevision's classification of each element
131   is correct, given the classification options ('bprob'). Because elements that are classified with
132   a lower certainty (i.e., a low bprob) are at a higher risk of false positives[23], we filtered out all
133   trait estimates with a bprob < 0.95; this has been found to result in a relatively low rate of
134   false negatives without increasing the risk of false positives[23]. For specimens with multiple
135   high-confidence estimates of a trait (e.g., if two femurs were confidently identified and
136   measured from a specimen), we combined these measures by taking the mean. In this way, a
137   single high-quality estimate of each trait was made for each specimen whenever at least one
138   example of an element was confidently identified (*Skelevision-Only Dataset*). For those traits
139   that did not have at least one high confidence trait estimate (i.e., if there was not at least one
140   trait measure with a bprob ≥ 0.95), the trait was marked as missing data (given a value of 'NA'
141   in the *Skelevision-Only Dataset*).

142   *Phylogenetic Data Imputation and Validation*

143

144   To generate a 100% complete dataset, we imputed values for all missing data in the
145   *Skelevision-Only Dataset* using Rphylopars, a maximum likelihood approach for fitting
146   multivariate phylogenetic models and estimating missing values in comparative data[37]. An
147   advantage of this approach is that it can model variation at the level of individual specimens
148   along with variation among species, providing estimates of missing values at the individual

149 level. This approach allowed us to leverage the dataset's large size and dimensionality (12
150 dimensions x 14,419 individuals), along with the expected hierarchical structure due to
151 phylogeny, to estimate missing values. To approximate phylogenetic relationships among the
152 included species, we downloaded 1,000 trees from the posterior distribution of a phylogeny
153 for all birds[6,38] and constructed a consensus tree following Rubolini *et al.*[39]. Using this
154 consensus tree, we used Rphylopars to estimate variance-covariance structures (both within
155 and between species) according to the expectations of a multivariate Brownian Motion model
156 ('mvBM') process. All data were log-transformed before model fitting with Rphylopars, but
157 otherwise, we left all user options set at their defaults (see *Code Availability* for details on
158 accessing the 'Trait_Imputation.R' script).

159 To evaluate the accuracy of trait imputation, we validated our approach by iteratively
160 withholding 10-90% of the non-missing data for each trait to use as test data. We then
161 estimated the mvBM model, excluding the test data in the model fitting, and imputed the
162 withheld test data. For each level of missing data, we repeated the analyses ten times,
163 selecting a different random sample of the dataset to use as test data for each replicate. We
164 then compared the known true values of the test data to their estimated values by calculating
165 the root-mean-squared error (RMSE) and percentage bias (p-bias; Fig. 2).

166 We present the data from these analyses as complete datasets both at the specimen
167 level (*Complete Trait Dataset*) and at the species level using species averages estimated by
168 Rphylopars, which includes both the averages and the estimated standard error associated
169 with each species mean (*Species-level Data Estimates*).

170

Data Records

172

173 For all datasets, we include species binomials following the BirdLifeV3 taxonomy. For the
174 specimen-level data, we also include museum specimen catalogue numbers. The data are
175 presented in three datasets:

176

Skelevision-Only Dataset

178 This dataset includes all the Skelevision data measured with high confidence (bprob ≥ 0.95). It
179 does not include any imputed trait measurements. The data are available for download as a
180 comma-separated values file, "Skelevision_Only_Dataset_v1.csv". DOI to be provided
181 following peer review.

182

Complete Trait Dataset

184 This dataset includes all the Skelevision data measured with a high degree of confidence and
185 imputed trait values for all missing data. The data are available for download as a comma-
186 separated values file, "Complete_Trait_Dataset_v1.csv". DOI to be provided following peer
187 review.

188

Species-level Data Estimates

190 This dataset includes species means from the model fit using Rphylopars. We also provide
191 estimates of the species-level standard error, variance, and 95% confidence intervals around
192 trait means for downstream analyses. The data are available for download as a comma-
193 separated values file, "Skelevision_species_complete_v1". DOI to be provided following peer
194 review.

195

196

Technical Validation

198

Skelevision Accuracy

200   The accuracy of the processed data (*Skelevision-Only Dataset*) generated from the
201   UMMZ image capture setup and specimens has been quantified previously. Weeks et al (2023)
202   compared 100 handmade measurements of each trait (except the furcula) to Skelevision
203   estimates of the same traits on the same specimen and found a mean RMSE of 0.89mm, with
204   some variation in error across bone types (Table 1).
205   Because there is a risk that Skelevision will perform differently across the different
206   image-capturing contexts (e.g., locations with variation in lighting) and some variation in
207   specimen preparation between UMMZ and FMNH (e.g., differences in the degree to which
208   bones remain articulated), and to collect validation data for the furcula, we conducted a
209   similar validation test with a subset of the FMNH data. For a random sample of 30 specimens
210   from FMNH, a single person measured each trait of interest from the photographs of the
211   specimens using ImageJ software[40]. We then compared these handmade measurements of
212   the trait values to the Skelevision measurements for the same trait on the same skeleton. As
213   with the UMMZ samples, we find Skelevision is accurate, with a mean RMSE of 1.78 mm across
214   all traits. This is higher than the RMSE of the UMMZ specimens but remains comparable to
215   inter-human measurement error. The errors are not uniform among the element types, and
216   while many have similar or lower errors compared to the UMMZ data, a few have elevated
217   error levels, albeit on a similar scale to the expected range of human measurement error
218   (Table 1)[9].
219
220   *Trait Imputation Accuracy*
221
222   The RMSE of the imputed data was uniformly low across replicates and increasing
223   levels of additional missing data (Table 2; Figure 2). This suggests potential uncertainty in
224   phylogenetic relationships has a negligible impact on the trait imputation accuracy. The
225   maximum mean RMSE (~ 1.15mm) was observed for estimates of the length of the second
226   digit; notably this was only when we simulated maximal amounts of additional missing data
227   (90%; or ~85% missing data overall). This maximum mean error is like that observed from
228   human measurement[9]. We also observe very low P-bias (< 1%) in estimated values
229   throughout the range of evaluated levels of missing data, lending further credence to the
230   validity of our approach. Maximum P-bias values ranged from ~ -0.4% (radius) to 0.5% (second
231   digit) and thus were essentially negligible, particularly at the levels of missing data within the
232   dataset (Table 2); most traits had a mean P-bias centered near zero.
233   In general, when anatomical traits have strong phylogenetic signals and multivariate
234   correlations, we expect to be able to estimate missing values with high accuracy and precision
235   under mvBM. Our approach highlights the power of multivariate phylogenetic models to
236   generate complete datasets at the level of individual specimens and should provide a useful
237   framework for future research.
238

Code Availability

240   The code to generate Skelevision trait estimates is publicly available (DOI:
241   10.5281/zenodo.6402893). We have also provided all the code used to impute the trait data,
242   'Trait_Imputation.R', with the data on Dryad (DOI to be provided following peer review). The
243   validation data and the script used to quantify RMSE are also included with the data on Dryad
244   (DOI to be provided following peer review). All trait imputation was done in R[41]. Images of all
245   UMMZ specimens are available through the Deep Blue Data Repository (DOI to be provided
246   following peer review).

247
248

Acknowledgements

Author contributions

BCW: Conceptualization, Methodology, Validation, Formal analysis, Investigation, Resources, Data Curation, Project Administration, Writing – Original Draft

ZZ: Methodology, Formal analysis, Investigation, Writing – Review and Editing

CP: Methodology, Validation, Investigation, Writing – Review and Editing

JB: Methodology, Validation, Formal analysis, Writing – Original Draft

BO: Investigation, Writing – Review and Editing

BB: Project Administration, Writing – Review and Editing

HS: Investigation, Writing – Review and Editing

MZ: Investigation, Writing – Review and Editing

LB: Investigation, Writing – Review and Editing

DFF: Conceptualization, Methodology, Formal Analysis, Resources, Project Administration, Writing – Review and Editing

Competing interests

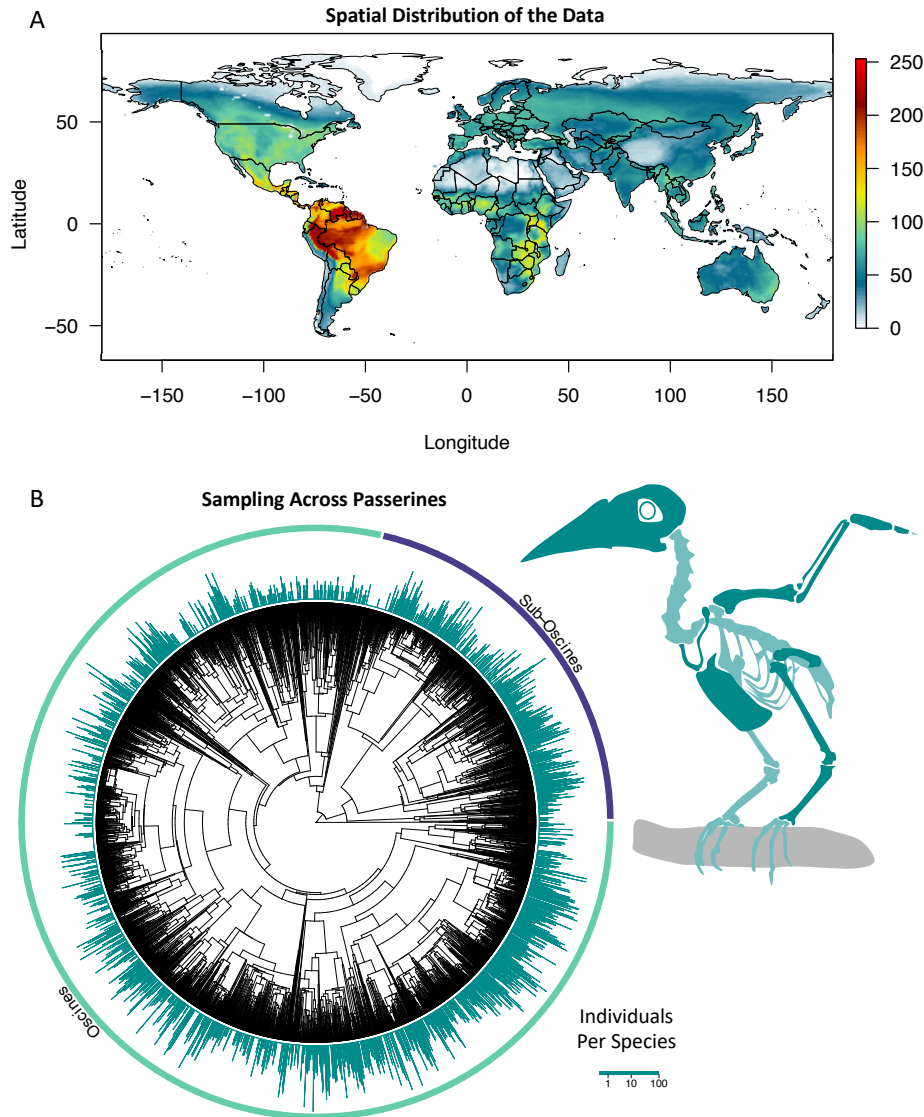The authors declare no competing interests.

Figures

Figure 1. Dataset coverage. (**A**) The dataset includes species spanning all continents except Antarctica. The ranges of all species included in the dataset are plotted, with colour indicating the number of species included in our dataset at each point in space. (**B**) The species in the dataset span Passeriformes, the most diverse taxonomic order within modern birds. Each bar associated with a tip on the phylogeny represents a species that is included in the dataset, with the height of the bar indicating the number of individuals of that species that were measured and included in the dataset. The bird skeleton highlights the bones that were measured in darker green.
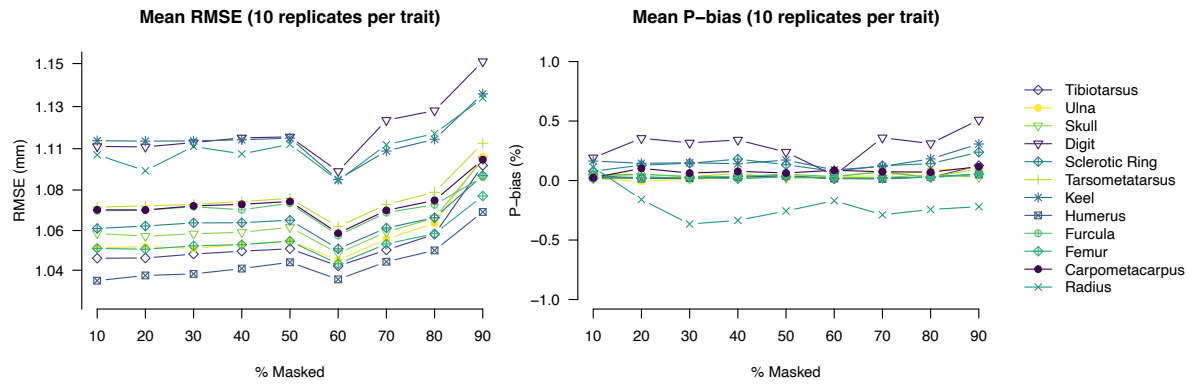
Figure 2. Variation in trait imputation accuracy across different levels of missing data. We masked an increasing percentage of the data generated by Skelevision (10-90% in intervals of 10%) and then imputed the missing values. We estimate the RMSE and P-Bias across ten randomized sets of data for each percentage quantile and present the mean per trait. RMSE and P-Bias are generally low and stable across the range of missing data quantiles.

8

295
296 **Tables**
297
298

Skeletal Element	UMMZ-Human RMSE[†]	UMMZ-Human P-bias[†]	FMNH-Human RMSE	FMNH-Human P-bias
Skull	1.20 mm	-1.8%	0.68 mm	-0.7%
Sclerotic Ring	0.34 mm	2.4%	0.25 mm	1.0%
Humerus	0.47 mm	-0.2%	0.53 mm	2.3%
Radius	0.51 mm	-0.5%	1.39 mm	-0.7%
Ulna	0.96 mm	-1.6%	0.86 mm	-0.6%
Carpometacarpus	1.11 mm	0.5%	1.64 mm	1.3%
2nd Digit 1st Phalanx	0.49 mm	3.3%	3.14 mm	20.4%
Furcula	NA	NA	3.31 mm	4.0%
Keel	2.23 mm	-6.1%	4.63 mm	8.5%
Femur	0.46 mm	0.0%	0.39 mm	0.6%
Tibiotarsus	1.17 mm	-0.7%	0.77 mm	2.1%
Tarsometatarsus	0.84 mm	0.6%	3.74 mm	9.6%

[†]Estimates for the UMMZ-human RMSE and P-bias are taken from [23]

299
300 **Table 1.** Estimated error of Skelevision-generated data. Data for 100 specimens photographed
301 at UMMZ were compared to measurements made by a single measurer, who measured the
302 same bone by hand using digital callipers. The root mean square error (RMSE) and percent
303 bias (P-bias) of the Skelevision estimates compared to the handmade measurements, taken
304 from[23], are presented. For 30 FMNH specimens, a single measurer used ImageJ software[40] to
305 measure each of the traits in the photographs; as with the UMMZ specimens, we present the
306 RMSE and P-bias for the Skelevision estimates compared to the handmade estimates from the
307 same specimen.
308
309
310

Skeletal Element	Specimens with Direct Measurements	Percent Missing Data	Predicted RMSE	Predicted P-bias
Skull	13,197	8%	1.06 mm	0.04%
Sclerotic Ring	8,703	40%	1.07 mm	0.18%
Humerus	12,306	15%	1.04 mm	0.02%
Radius	7,274	50%	1.11 mm	-0.26%
Ulna	8,574	41%	1.06 mm	0.04%
Carpometacarpus	7,681	47%	1.08 mm	0.06%
2nd Digit 1st Phalanx	4,635	68%	1.12 mm	0.36%
Furcula	5,473	62%	1.07 mm	0.07%
Keel	10,842	25%	1.11 mm	0.15%
Femur	12,584	13%	1.05 mm	0.02%
Tibiotarsus	12,579	13%	1.05 mm	0.02%
Tarsometatarsus	8,921	38%	1.08 mm	0.06%

311 **Table 2. Estimated error of imputed data.** We round up the percent missing data to the
312 nearest ten and include the predicted RMSE and P-bias for that trait at that level of missing
313 data (Fig. 2).

314

315

References

1. Tobias, J. A. A bird in the hand: Global-scale morphological trait datasets open new frontiers of ecology, evolution and ecosystem science. *Ecology Letters* vol. 25 573–580 Preprint at https://doi.org/10.1111/ele.13960 (2022).

2. Kattge, J. *et al.* TRY plant trait database – enhanced coverage and open access. *Glob Chang Biol* (2020) doi:10.1111/gcb.14904.

3. Meiri, S. Traits of lizards of the world: Variation around a successful evolutionary design. *Global Ecology and Biogeography* **27**, 1168–1172 (2018).

4. Brosse, S. *et al.* FISHMORPH: A global database on morphological traits of freshwater fishes. *Global Ecology and Biogeography* **30**, 2330–2336 (2021).

5. BirdLife International. IUCN Red List for birds. http://www.birdlife.org (2015).

6. Jetz, W., Thomas, G. H., Hartmann, K. & Mooers, A. O. The global diversity of birds in space and time. *Nature* **491**, 444–448 (2012).

7. Prum, R. O. *et al.* A comprehensive phylogeny of birds (Aves) using targeted next-generation DNA sequencing. *Nature* **526**, 569–573 (2015).

8. Wilman, H. *et al.* EltonTraits 1.0: Species-level foraging attributes of the world's birds and mammals. *Ecology* **95**, 2027–2027 (2014).

9. Pigot, A. *et al.* Macroevolutionary convergence connects morphological form to ecological function in birds. *Nat Ecol Evol* **4**, 230–239 (2020).

10. Tobias, J. A. & Pigot, A. L. Integrating behaviour and ecology into global biodiversity conservation strategies. *Philosophical Transactions of the Royal Society B: Biological Sciences* **374**, 20190012 (2019).

11. Sheard, C. *et al.* Nest traits for the world's birds. *Global Ecology and Biogeography* **33**, 206–214 (2023).

12. Tobias, J. A. *et al.* AVONET: morphological, ecological and geographical data for all birds. *Ecol Lett* **25**, 581–597 (2022).

13. Sheard, C. *et al.* Ecological drivers of global gradients in avian dispersal inferred from wing morphology. *Nat Commun* **11**, 1–9 (2020).

14. Hughes, E. C. *et al.* Global biogeographic patterns of avian morphological diversity. *Ecology Letters* vol. 25 598–610 Preprint at https://doi.org/10.1111/ele.13905 (2022).

15. Freeman, B. G., Weeks, T., Schluter, D. & Tobias, J. A. The latitudinal gradient in rates of evolution for bird beaks, a species interaction trait. *Ecology Letters* vol. 25 635–646 Preprint at https://doi.org/10.1111/ele.13726 (2022).

16. Crouch, N. M. A. & Tobias, J. A. The causes and ecological context of rapid morphological evolution in birds. *Ecology Letters* vol. 25 611–623 Preprint at https://doi.org/10.1111/ele.13962 (2022).

17. Drury, J. P. *et al.* Limited ecological opportunity influences the tempo of morphological evolution in birds. *Current Biology* (2024) doi:10.1016/j.cub.2023.12.055.

18. Weeks, B. C., Naeem, S., Lasky, J. R. & Tobias, J. A. Diversity and extinction risk are inversely related at a global scale. *Ecol Lett* **25**, 697–707 (2022).

19. Stewart, P. S. *et al.* Global impacts of climate change on avian functional diversity. *Ecology Letters* vol. 25 673–685 Preprint at https://doi.org/10.1111/ele.13830 (2022).

20. Ali, J. R., Blonder, B. W., Pigot, A. L. & Tobias, J. A. Bird extinctions threaten to cause disproportionate reductions of functional diversity and uniqueness. *Funct Ecol* **37**, 162–175 (2023).

21. Martínez-Núñez, C., Martínez-Prentice, R. & García-Navas, V. Land-use diversity predicts regional bird taxonomic and functional richness worldwide. *Nat Commun* **14**, (2023).

22. Howard, C. *et al.* Local colonisations and extinctions of European birds are poorly explained by changes in climate suitability. *Nat Commun* **14**, (2023).

367 23. Weeks, B. C. *et al.* A deep neural network for high throughput measurement of
368   functional traits on museum skeletal specimens. *Methods Ecol Evol* **14**, 347–359 (2023).
369 24. Lowi-Merri, T. M. *et al.* Reconstructing locomotor ecology of extinct avialans: a case
370   study of Ichthyornis comparing sternum morphology and skeletal proportions.
371   *Proceedings of the Royal Society B: Biological Sciences* **290**, (2023).
372 25. Harvey, C., Baliga, V. B., Wong, J. C. M., Altshuler, D. L. & Inman, D. J. Birds can transition
373   between stable and unstable states via wing morphing. *Nature* **603**, 648–653 (2022).
374 26. Sullivan, T. N., Meyers, M. A. & Arzt, E. Scaling of bird wings and feathers for efficient
375   flight. *Sci Adv* **5**, eaat4269 (2019).
376 27. Wright, N. A., Steadman, D. W. & Witt, C. C. Predictable evolution toward flightlessness
377   in volant island birds. **113**, 4765–4770 (2016).
378 28. Field, D. J., Benito, J., Chen, A., Jagt, J. W. M. & Ksepka, D. T. Late Cretaceous
379   neornithine from Europe illuminates the origins of crown birds. *Nature* **579**, 397–401
380   (2020).
381 29. Weeks, B. C. *et al.* Shared morphological consequences of global warming in North
382   American migratory birds. *Ecol Lett* **23**, 316–325 (2020).
383 30. Field, D. J., Lynner, C., Brown, C. & Darroch, S. A. F. Skeletal correlates for body mass
384   estimation in modern and fossil flying birds. *PLoS One* **8**, (2013).
385 31. Allen, J. A. The influence of physical conditions in the genesis of species. *Radical Review*
386   **1**, 108–140 (1887).
387 32. Weeks, B. C. *et al.* Skeletal morphology of bird wings is determined by
388   thermoregulatory demand for heat dissipation in warmer climates. *bioRxiv* 1–11 (2023)
389   doi:https://doi.org/10.1101/2023.02.06.527306.
390 33. Ronneberger, O., Fischer, P. & Brox, T. U-net: Convolutional networks for biomedical
391   image segmentation. in *International conference on medical image computing and
392   computer-assisted intervention* 234–241 (Springer, Cham, 2015). doi:10.1007/978-3-
393   319-24574-4_28.
394 34. He, K., Gkioxari, G., Dollár, P. & Girshick, R. Mask R-CNN. *IEEE Trans Pattern Anal Mach
395   Intell* (2020) doi:10.1109/TPAMI.2018.2844175.
396 35. Kar, A. *et al.* Toronto Annotation Suite. Preprint at (2021).
397 36. Dutta, A., Gupta, A. & Zisserman, A. VGG image annotator (VIA). Preprint at
398   http://robotics.ox.ac.uk/~vgg/software/via/ (2016).
399 37. Goolsby, E. W., Bruggeman, J. & Ané, C. Rphylopars: fast multivariate phylogenetic
400   comparative methods for missing data and within-species variation. *Methods Ecol Evol*
401   **8**, 22–27 (2017).
402 38. Hackett, S. J. *et al.* A phylogenomic study of birds reveals their evolutionary history.
403   *Science (1979)* **320**, 1763–1768 (2008).
404 39. Rubolini, D., Liker, A., Garamszegi, L. Z., Møller, A. P. & Saino, N. Using the birdtree.org
405   website to obtain robust phylogenies for avian comparative studies: a primer. *Curr Zool*
406   **61**, 959–965 (2015).
407 40. Schneider, C. A., Rasband, W. S. & Eliceiri, K. W. NIH Image to ImageJ: 25 years of image
408   analysis. *Nature Methods* vol. 9 671–675 Preprint at
409   https://doi.org/10.1038/nmeth.2089 (2012).
410 41. R Core Team. R: A language and environment for statistical computing. Preprint at
411   http://www.r-project.org/ (2018).
412
413